Generation and analysis of a gene regulatory network for *Chlamydomonas reinhardtii*

Dissertation an der Fakultät für Biologie
der Ludwig-Maximilians-Universität München
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

Mayra Anna Sauer

München, 2023

Generation and analysis of a gene regulatory network for *Chlamydomonas reinhardtii*

Dissertation an der Fakultät für Biologie
der Ludwig-Maximilians-Universität München
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

Mayra Anna Sauer

München, 2023

Diese Dissertation wurde angefertigt
unter der Leitung von Prof. Dr. Pascal Falter-Braun
am Institut für Netzwerkbiologie
des Helmholtz Zentrums München.

Erstgutachter: Prof. Dr. Pascal Falter-Braun

Zweitgutachter: Prof. Dr. Jörg Nickelsen

Dissertation eingereicht am: 12.06.2023

Tag der mündlichen Prüfung: 07.02.2024

Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt ist.

Die vorliegende Dissertation wurde weder ganz, noch teilweise bei einer anderen Prüfungskommission vorgelegt. Ich habe noch zu keinem früheren Zeitpunkt versucht, eine Dissertation einzureichen oder an einer Doktorprüfung teilzunehmen.

München, den 12.06.2023

Mayra Anna Sauer

♦ Dedication ♦

To Paul and Philipp.

Abstract

Stopping climate change requires renewable energy sources. Microalgae are a promising biotechnological resource for biofuels. Under stress conditions they produce large quantities of storage lipids as triacylglycerides (TAGs), which can be further converted into biodiesel. To optimize the biotechnological production of valuable compounds such as TAGs in microalgae, a thorough system-level understanding of the underlying regulatory mechanisms is required. Current knowledge is based on mutant screens, co-expression analyses, and predictions but a systems-level perspective is missing. The present work is dedicated to expanding the existing knowledge of fundamental regulatory mechanisms in microalgae and enabling advanced biotechnological use of microalgae.

All regulatory processes in a cell are mediated and controlled by interactions between macromolecules, such as protein-protein interactions (PPI), protein-DNA interactions (PDI), or interactions between proteins and RNA or small molecules. These interactions are organized in complex networks, and the systematic mapping of these interactomes and regulomes is essential for understanding complex biological relationships in an organism. In order to enable large-scale mapping of PPIs and PDIs for the model organism of microalgae, *Chlamydomonas reinhardtii* (*C. reinhardtii*), a comprehensive set of protein-encoding open-reading frames (ORFs) was generated. This ORFeome allows for an unbiased and high-throughput approach to performing proteomic studies. The successfully generated ORFeome collection consists of 8,627 Gateway® Entry clones. It enables systematic investigations of fundamental biological and regulatory processes in *C. reinhardtii*.

PDIs are mediated by transcription factors (TFs) that bind to specific genomic DNA sequences in promoter regions, enhancers, and other cis-regulatory elements, thereby regulating the expression of target genes. These interactions are described in gene regulatory networks (GRN). Using the Yeast-1-hybrid (Y1H) method, the first experimental-based large-scale GRN of *C. reinhardtii* was constructed, consisting of 1,451 interactions between 142 TFs binding to 200 promoter regions. To determine the quality of the Y1H network, a second orthogonal assay, called PampDAP-seq, was performed. This modified DAP-seq method identified 320 unique PDIs for six TFs. The comparison of DAP-seq interactions with found Y1H interactions showed an overlap of 23.8 %, which highly

supports the Y1H results.

The generated network map was hierarchically structured according to the outdegree of the TFs. 15 TFs represent the highest level in this network and are responsible for 66.7 % of all interactions. In addition, multiple TFs, including previously known stress-related TFs, were found to form an intermediate level of regulation that combines various signals from other TFs to mediate a subsequent cellular response. Moreover, a distributed regulation structure for metabolic pathway promoters was observed. 84.5 % of metabolic promoters are targeted by more than two TFs. This indicates that each metabolic pathway is regulated by a large number of TFs, a distinct characteristic of higher plants like *Arabidopsis thaliana* (*A. thaliana*). In addition, four promising key regulators of the *C. reinhardtii* lipid metabolism were identified.

The successfully generated ORFeome collection provides a resource for the in-depth exploration of fundamental biological processes. Furthermore, the first experimental-based GRN for any microalgae provides valuable insights into the global structure of transcriptional regulation and can pave the way for advanced biotechnological engineering in microalgae.

Zusammenfassung

Die globale Klimakrise erfordert die Energiegewinnung aus erneuerbaren Rohstoffen. Mikroalgen stellen eine solche vielversprechende Energieressource dar, die zur Biodieselherstellung genutzt werden kann. Unter Stressbedingungen produzieren Mikroalgen große Mengen an Lipiden, so genannte Triacylglyceride (TAGs), die in Biodiesel umwandelbar sind. Die Optimierung der biotechnologischen Produktion von TAGs hängt entscheidend von einem fundierten, globalen Verständnis der zugrunde liegenden Regulationsmechanismen ab. Dieses Wissen ist für Mikroalgen noch lückenhaft und basiert bisher vorwiegend auf Mutanten Screenings und Genexpressionsanalysen.

Die vorliegende Arbeit widmet sich dem Ziel, das vorhandene Wissen zu ergänzen und auszubauen, um grundlegende regulatorische Mechanismen in Mikroalgen zu verstehen und eine zukunftsweisende, biotechnologische Nutzung von Mikroalgen zu ermöglichen.

Zelluläre Stoffwechselreaktionen und regulatorische Prozesse werden durch Wechselwirkungen von Makromolekülen gesteuert, dazu zählen Protein-Protein Interaktionen (PPI), Protein-DNA Interaktionen (PDI) und Interaktionen zwischen Proteinen und RNA oder niedermolekularen Verbindungen. Diese Interaktionen werden in komplexen Netzwerken, so genannten Interaktomen und Regulomen dargestellt. Interaktome und Regulome sind für ein systematisches Verständnis und die Analyse komplexer biologischer Zusammenhänge in einem Organismus unerlässlich.

Diese Netzwerke können nun generiert werden, da im Zuge dieser Arbeit eine umfangreiche proteinkodierende Open-Reading-Frame (ORF) Kollektion (ORFeome) für den Modellorganismus der Mikroalgen, *Chlamydomonas reinhardtii* (*C. reinhardtii*), aufgebaut wurde. Die ORFeome-Sammlung beinhaltet 8627 Gateway® Entry-Klone und bildet die Grundlage für proteomische Studien von PPIs und PDIs zur Aufklärung grundlegender biologischer und regulatorischer Prozesse in *C. reinhardtii*.

PDIs definieren die Bindung von Transkriptionsfaktoren (TFs) an spezifischen genomischen DNA-Sequenzen in Promotorregionen, Enhancern und anderen cis-regulatorischen Elementen. Diese Interaktionen regulieren die Expression von Zielgenen und werden in einem genregulatorischen Netzwerk (GRN) abgebildet.

Mit Hilfe der Yeast-1-Hybrid (Y1H)-Methode wurde das erste experimentell bestimmte GRN von *C. reinhardtii* aufgebaut, welches aus 1451 Interaktionen zwischen 142 TFs und 200 Promotorregionen besteht. Um die Qualität des Y1H-Netzwerks zu bestätigen, wurde ein zweiter Assay zur PDI-Bestimmung, PampDAP-seq genannt, durchgeführt. Diese modifizierte DAP-seq-Methode identifizierte 320 einzigartige PDIs für sechs TFs. Der Vergleich der Interaktionsdaten zeigte eine Übereinstimmung von 23,8 % und bestätigte damit die Aussagekraft der Y1H Ergebnisse.

Das Netzwerk wurde entsprechend der Anzahl der ausgehenden Verbindungen der TFs hierarchisch strukturiert. 15 TFs repräsentieren die höchste Ebene und bilden 66,7 % der gefundenen Interaktionen. Ein Großteil der TFs, unter anderem bereits bekannte Stress-assoziierte TFs, formen eine mittlere regulatorische Ebene, die Signale bündeln und gezielt weiterleiten. Weiterhin konnte eine dezentralisierte Regulation für die Promotoren der Stoffwechselwege beobachtet werden. 84,5 % der Stoffwechselpromotoren werden von mehreren TFs gebunden. Dies zeigt, dass jeder Stoffwechselweg von einer großen Anzahl an TFs reguliert wird, was ein charakteristisches Merkmal für höherentwickelte Pflanzen wie *Arabidopsis thaliana* (*A. thaliana*) ist. Außerdem wurden vier erfolgversprechende Schlüsselregulatoren des Lipidstoffwechsels von *C. reinhardtii* identifiziert.

Das erste ORFeome für *C. reinhardtii* stellt eine bedeutende Ressource für die zukünftige Erforschung grundlegender biologischer Prozesse dar. Das erste experimentell bestimmte genregulatorische Netzwerk für Mikroalgen liefert wertvolle Einblicke in die globale Struktur der transkriptionellen Regulation in *C. reinhardtii* und ebnet damit den Weg für eine fortschrittliche biotechnologische Verwendung von Mikroalgen.

List of Publications, Presentations and Posters

List of publications

Sauer M, Pandiarajan R, Lin C.-W, Weller B, Rothballer S, Nigish S, Strobel A, Einhaus A, Wobbe L, Falter C, Dementyeva P, Kruse O, Salehi-Ashtiani K, Falter-Braun P (in preparation). ChlamyRegNet: The first gene regulatory network for *C. reinhardtii**

Sauer M, Weller B, Lin C.-W, Strobel A, Rothballer S, Nigish S, Kang YJ, Falter C, Wobbe L, Kruse O, Falter-Braun P (in preparation). The *Chlamydomonas reinhardtii* ORFeome: a resource for high-throughput proteomics

Weller B, Lin CW, Pogoutse O, **Sauer M**, Marin-de la Rosa N, Strobel A, Young V, Knapp J, Rayhan A, Falter C, Kim DK, Roth F, Falter-Braun P (in revision). **A resource of human coronavirus protein coding sequences in a flexible, multi-purpose Gateway Entry clone collection.** *G3: Genes|Genomes|Genetics*.

Kim DK, Weller B, Lin CW, Sheykhkarimli D, Knapp JJ, Dugied G, Zanzoni A, Pons C, Tofaute MJ, Maseko SB, Spirohn K, Laval F, Lambourne L, Kishore N, Rayhan A, **Sauer M**, Young V, Halder H,Marín-de la Rosa N, Pogoutse O, Strobel A, Schwehn P, Li R, Rothballer ST, Altmann M, Cassonnet P, Coté AG, Vergara LE, Hazelwood I, Liu BB, Nguyen M, Pandiarajan R, Dohai B, Rodriguez Coloma PA, Poirson J, Giuliana P, Willems L, Taipale M, Jacob Y, Hao T, Hill DE, Brun C, Twizere JC, Krappmann D, Heinig M, Falter C, Aloy P, Demeret C, Vidal M, Calderwood MA, Roth FP, Falter-Braun P (2023). **A proteome-scale map of the SARS-CoV-2-human contactome.**
Nature Biotechnology 41 (1): 140–49. doi: 10.1038/s41587-022-01475-z

Altmann M, Altmann S, Rodriguez P, Weller B, Elorduy Vergara L, Palme J, Marín-de la Rosa N, **Sauer M**, Wenig M, Villaécija-Aguilar J A, Sales J, Lin C-W, Pandiarajan R, Young V, Strobel A, Gross L, Carbonnel S, Kugler K G, Garcia-Molina A, Bassel G W, Falter C, Mayer K F X, Gutjahr C, Vlot A C, Grill E, Falter-Braun P (2020). **Extensive signal integration by the phytohormone protein network.**
Nature 583 (7815): 271–76. doi: 10.1038/s41586-020-2460-0

List of presentations and posters

09/2021	Oral presentation at the 19th International Conference on the Cell and Molecular Biology of Chlamydomonas
09/2019	Poster presentation at the Yeast Munich Meeting 2019
02/2019	Poster presentation at the Genomics and Systems Biology IX Conference
09/2018	Poster and oral presentation at the e:Bio Innovation Competition Systems Biology Seminar
06/2018	Poster presentation at the 18th International Conference on the Cell and Molecular Biology of Chlamydomonas
04/2018	Poster presentation at the EMBO Workshop: Integrating Systems Biology: From Networks to Mechanism to Models
06/2016	Oral presentation at the 17th International Conference on the Cell and Molecular Biology of Chlamydomonas

Contents

List of Figures

List of Tables

Nomenclature

3-AT	3-amino-1,2,4 triazole
3-PG	3-phosphoglycerate
A. thaliana	*Arabidopsis thaliana*
C. elegans	*Caenorhabditis elegans*
C. reinhardtii	*Chlamydomonas reinhardtii*
E. coli	*Escherichia coli*
H. sapiens	*Homo sapiens*
O. sativa	*Oryza sativa*
S. cerevisiae	*Saccaromyces cerevisiae*
Z. mays	*Zea mays*
aa	amino acids
ACCase	acetyl-CoA carboxylase
ACP	acyl carrier protein
AD	activation domain
AMP	adenosine monophosphate
AP-MS	affinity-purification mass spectrometry
AP3	APETALA3 transcription factor
ASQD	2′-O-acyl-sulfoquinovosyldiacylglycerol
ATP	adenosine triphosphate
BC	bar-coded

BFG-Y2H	Barcode Fusion Genetics-Yeast Two-Hybrid
bHLH TF	basic helix-loop-helix domain TF
bp	base pairs
bZIP TF	Basic Leucine Zipper domain TF
CAM	chloramphenicol
Carb	carbenicillin
CCM	carbon concentrating mechanism
cDNA	complementary DNA
CDS	coding sequences
CEN	centromere sequence/region
CGL86	Cre12.g534450, FHA-TF, orthologue for nuclear inhibitor of PP1
ChIP-seq	chromatin-immunoprecipitation sequencing
CHX	cycloheximide
COVID-19	coronavirus disease 2019
CREs	*cis*-regulatory elements
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRP	cAMP receptor protein
CTAB	cetyltrimethylammonium bromide
DAG	diacylglycerol
DAP-seq	DNA-affinity purification sequencing
DB	DNA binding domain
DCW	dry cell weight
DGAT	diacylglycerol acyltransferase
DGDG	digalactosyldiacylglycerol
DGTS	diacylglycerol-*N,N,N*-trimethylhomoserine
DMAP1	DNA methyltransferase 1-associated protein 1

DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DYRK	Dual-Specificity Tyrosine-Phosphorylation-Regulated Kinase
DYRKP	plant-specific DYRK
ER	endoplasmatic reticulum
eY1H	enhanced Y1H
F2,6BP	fructose-2,6-biphosphate
FA	fatty acid
FAD	fatty acid desaturase
FAS	fatty acid synthase
FAT	fatty-ACP thioesterases
FBP4	fructose-2,6-bisphosphate 2-phosphatase / phosphofructokinase-2
FFA	free fatty acid
FHA TF	forkhead-associated domain TF
G3P	glycerol-3-phosphate
Gal4-AD	yeast Gal4 transcription factor activation domain
Gal4-DB	yeast Gal4 DNA-binding domain
GAP	glyceraldehyde 3-phosphate
GNAT	Gcn5-related N-acetyltransferases
GOI	gene of interest
GPAT	glycerol-3-phosphate acyltransferase
GRN	gene regulatory network
GWAS	genome-wide association studies
HMG TF	high mobility group transcription factor
HMG_6	Cre16.g672300, a HMG-TF
HT	high-throughput

HYDA1	[FeFe]-hydrogenase isoform 1
IPTG	Isopropyl-ß-D-thiogalactopyranosid
IRE1	Inositol-Requiring Enzyme 1
kb	kilobase (= 1000 bp)
LACS	long-chain acyl-CoA synthetase
LB	lysogeny broth
LHCII	light-harvesting complex II
LPA	lysophosphatidic acids
LPAT	lysophosphatidic acid acyl-transferase
LRL1	Lipid remodeling regulator 1
LSD1	Lesion Simulation Disease 1
MGDG	monogalactosyldiacylglycerol
mQ	milli-Q water, ultrapure water
mRNA	messenger RNA
MS_1	Cre10.g450500, TF with unknown DNA-binding domain, wrongly assigned as MYB TF
Mtoe	millions of tonnes of oil equivalent
MUFAs	monounsaturated fatty acids
MYB TF	myeloblastosis TF
MYB_9	Cre03.g144907, MYB TF
N	nitrogen
NAD/H	nicotinamide adenine dinucleotide
NADP/H	nicotinamide adenine dinucleotide phosphate
NGS	next-generation sequencing
NIPP1	nuclear inhibitor of protein phosphatase-1
NRR1	Nitrogen Response Regulator 1

OK1	different name for PWD, phosphoglucan, water dikinase
ORF	open reading frame
ORFeom	collection of protein-encoding ORFs in a plasmid backbone
OTC1	ornithine carbamoyltransferase 1
P	phosphate
PA	phosphatidic acid
PampDAP-seq	promoter PCR amplified DAP-seq
PAP	phosphatidic acid phosphatase
PBM	protein binding microarrays
PCR	polymerase chain reaction
PDAT	phospholipid:diacylglycerol acyltransferase
PDI	protein-DNA interaction
PFK-1	phosphofructokinase-1
PGD1	plastid galactoglycerolipid degradation 1
PI	PISTILLA transcription factor
POI	protein of interest
PP1	protein phosphatase 1
PPI	protein-protein interaction
PPP	pentose phosphate pathway
PRS	positive reference set
PsaH	photosystem I reaction center subunit VI
PSI	photosystem I
PSII	photosystem II
PSR1	Phosphorus starvation response 1
PtdEtn	phosphatidylethanolamine
PtdGro	phosphatidylglycerol

PtdIns	phosphatidylinositol
PUFAs	polyunsaturated fatty acids
PWD	phosphoglucan, water dikinase
RCD1	Radical-induced cell death 1
RNA	ribonucleic acid
RRS	random reference set
RuBisCo	ribulose-1,5-bisphosphate carboxylase-oxygenase
S	sulfur
SAC1	Sulfur Acclimation 1
SANT domain	switching-defective protein 3 (Swi3), adaptor 2 (Ada2), nuclear receptor co-repressor (N-CoR), transcription factor (TF)IIIB domain
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
SDS	sodium dodecyl sulfate
SFAs	saturated fatty acids
sn	stereospecifically numbered
SNRK2.2	SNF1-Related Protein Kinase 2.2
SQD1	UDP-sulfoquinovose synthase 1
SQD2	diacylglycerol sulfoquinovosyl transferase 2
SQDG	sulfoquinovosyldiacylglycerol
Swc4, EAF2	SWR1-complex 4
SWR1-C	SWR1 chromatin remodeling complex
TAG	triacylglycerol
TAR1	triacylglycerol accumulation regulator 1
TB	terrific broth
TE buffer	Tris/EDTA-buffer
TF	transcription factor

TFBS	transcription factor binding site
TR	transcriptional regulator
TRAF	tumor necrosis factor receptor-associated factor
TRN	transcriptional regulated network
TSS	transcription start site
UAS	Upstream Activation Sequence
Y1H	yeast one-hybrid
Y2H	yeast two-hybrid
Yak1	Yet another Kinase 1

1 Introduction

1.1 Stopping climate change requires renewable energy resources

In February 2022, the Intergovernmental Panel on Climate Change (IPCC) stated that climate change threatens human life on Earth and the health of the planet, and therefore 'urgent action' is required [1]. The leading causes of global warming are the rise of emissions and accumulation of green-house gases, such as carbon dioxide (CO_2) caused by human activity. For this reason, new ways must be found to reduce these emissions. The greenhouse gases are generated by industrial processes such as the combustion of fossil fuels, gas, coal, and oil for energy production [1, 2, 3]. In 2018, the demand for energy in 2018 was reported at 14,550 Millions of tonnes of oil equivalent (Mtoe), and it is expected to rise in the coming decades [4]. Given that the energy generated by burning fossil fuels is the primary driver of climate change, the world must shift to the use of renewable energy sources [5, 6].

To date, renewable energy sources include hydro-, solar-, wind-, geothermal-, wave-, tidal power, and modern biofuels [5, 7]. Biofuels have been more favored of late since they are produced from biomass in a short period, while fossil fuels are produced over millions of years. Compared with other forms of renewable energy, biofuels can be easily stored and used directly with existing diesel engines and distribution infrastructure [8]. The most common biofuels are bioethanol and biodiesel. Bioethanol is produced by fermenting biomass, specifically from carbohydrates of starch-based feedstocks such as corn and sugarcane or from non-food cellulosic biomass such as grass or trees. Biodiesel, on the other hand, is made by catalytic transesterification from glycerolipids with methanol or ethanol, most commonly using triacylglycerides (TAGs) derived from vegetable oils or animal fats [9, 10, 11]. Biodiesel can be classified into 1^{st} and 2^{nd} generation depending on the feedstock used for their production. The 1^{st} generation biodiesel can be produced from vegetable oil extracted from seeds or pulp of various oil plants such as oil palm, canola, sunflower, peanut, soybean, and coconut palm [11]. They all must be grown in agriculture, require fertile soil, and are food crops. This raises crucial questions about sustainability. The increased demand for these crops for biofuel production negatively affects their prices on the food market [12, 13]. At

the same time, it also has a negative environmental impact through additional deforestation, eutrophication from intensive agricultural practices, and increased energy demand [13]. For example, the cultivation and processing of canola and soy plants require 50 % of the energy that is obtained by biodiesel production from them [14]. To reduce the problems associated with 1st generation biodiesel production, fats from non-edible resources, like oil from non-edible plants, waste cooking oil, and animal fats, has been used for biodiesel production. They are the feedstocks for 2nd generation biodiesel [15]. Even though these feedstocks do not compete with agriculture, they have limitations. Oil and fat waste are inexpensive, but the need for a centralized collection system and the impurities they contain limit their industrial application [16]. In addition, animal fats are a mixture of lipids, proteins, water and minerals that require additional purification before subsequent use in biodiesel production [17]. Therefore, alternatives to these 1st and 2nd generation biodiesel feedstocks are sought [14]. Algae, or more precisely microalgae, circumvent all these limitations and are thus a promising alternative [18, 19, 13, 20, 21].

1.2 Benefits of algae

Algae are a large and diverse group of around 300,000 photosynthetic species [14, 22]. Algae are eukaryotic organisms, characterized by living in aquatic biotopes, using sunlight for photosynthesis, and being primary autotrophs. They can also grow under mixotroph and heterotroph conditions using external carbon sources. Algae vary considerably in size and habitat. Unicellular microalgae can be easily cultivated in bioreactors, while large multicellular algae species require more space to grow and are unsuitable for growing in tanks or railways [23, 24, 25].

Microalgae have immense potential as a resource for biofuel production, as they can synthesize substantial amounts of biomass and oil, more than comparable amounts of conventional crops [26]. Some microalgae species produce large amounts of TAGs, neutral storage lipids, which can be easily converted into biodiesel [13, 27]. They can reach lipid amounts of 20-60 % of the algae's dry cell weight (DCW) [19]. In contrast, the usual lipid amount in conventional oil plants accounts for only 0.5-25 % of their DCW [28].

In addition, the life cycle of algae is faster compared to conventional oil plants. Algae have a rapid doubling time. They can perform 1-3 doublings per day and produce up to 10-fold more biomass per area than comparable terrestrial plants [13]. This biomass can, in turn, be used to produce bioethanol [19, 18, 13, 20, 21]. Microalgae can be cultivated in saline or brackish water in ponds and tanks, such as raceways or photobioreactors [29]. This provides another enormous advantage of algae: they can be cultivated on barren lands, such as deserts, unsuitable for conventional agriculture [30]. It reduces competition for agricultural land and freshwater with conventional food agriculture and production. Algae utilize growth nutrients such as nitrogen and phosphorus from

various wastewater sources like agricultural wastewater or municipal sewage, thereby providing an additional benefit of wastewater bioremediation. Consequently, microalgae represent a more promising solution for biofuels production as 1st and 2nd generation biofuel sources [31].

1.3 Biodiesel quality relies on fatty acid composition

Biodiesel consists of a variety of monoalkyl esters of fatty acids (FA) released from TAGs by a transesterification reaction [13, 27]. Oils from different sources vary in their FA composition and, thus, in their biodiesel properties, as these are influenced by the different FA esters. The structure of different FA esters in terms of chain length and degree of saturation affects low-temperature operability, e.g., viscosity, oxidative stability, and combustion efficiency [32, 33]. The lipid composition of these FAs is determined by the proportion of saturated FAs (SFA) and unsaturated FAs. Unsaturated FAs can be further divided into monounsaturated FAs (MUFA) and polyunsaturated FAs (PUFA). Biodiesel quality depends on the ratio of SFAs and unsaturated FAs and the carbon chain length of the respective FAs [34]. SFAs are resistant to degradation and, thus, increase the longevity of biodiesel, e.g., palmitic (C16:0) acid. MUFAs and PUFAs, such as oleic (C18:1), palmitoleic (C16:1) acid, linoleic (C18:2), and linolenic (C18:3) acids, improved low-temperature flow proprieties such as viscosity by lowering the mixture's overall melting point and improve the ignition properties of the biodiesel [32, 33, 35]. On the other hand, PUFAs are prone to auto-oxidation. However, the auto-oxidation rate depends on the number and position of double bonds, which means that not all unsaturated FAs are susceptible to oxidation processes. Linoleic (C18:2) and linolenic (C18:3) acids, e.g., are highly sensitive to oxidation as bis-allylic FAs [32, 36]. The applicability of the produced biodiesel depends on the climate condition under which the biodiesel fuel will be used. In summary, biodiesel quality strongly depends on the composition of the FAs it contains. A proportion of 50 % MUFAs and a balanced ratio of SFAs and PUFAs are required to obtain high-quality biodiesel.

C. reinhardtii, as well as several other algae, e.g., *Chlorella*, and *Dunaliella* species exhibit a high proportion of PUFAs ($>$ 38 %, see Table A.1 in the appendix) and have a balanced ratio between SFA and MUFAs. However, the lipid profile of most algae does not fit to an optimal profile for high-quality biodiesel. Thus, they need to be blended to achieve the proper lipid profile that satisfies cold and hot flow requirements [12, 36, 34]. Alternatively, the yield and FA composition of the microalgae oils need to be improved or modified for direct use in biodiesel production [18, 37]. Successful industrial biodiesel production from microalgae depends on the lipid composition of an algae strain. Therefore, the first essential decision for successful biodiesel production from algae is the selection of a specific microalgae strain. Many global screening programs have searched for suitable algal species in various locations. The pioneering study 'Aquatic Species Program' from 1998 of the US Department of Energy was the first program that searched for suit-

able new algae species for biodiesel production [38]. Although there is a wide variety of algae, intensive research has focused on a small number of fast-growing microalgal species that can accumulate significant amounts of lipids and other components under certain conditions. In particular, the model organism *C. reinhardtii* has been extensively studied.

C. reinhardtii is used as a model organism for lipid metabolism [39, 40] but is also used to investigate fundamental biological processes such as photosynthesis, the CO_2-concentration mechanism, light perception, chloroplast biogenesis, cell-cell recognition, cell cycle control and structure, function, and biogenesis of cilia [41, 42, 43]. It is often referred to as 'green yeast' due to its many beneficial properties for molecular genetic studies [44]. Its fast growth rate, colony formation on agar plates, a haploid genome, and high-frequency transformation of nuclear, chloroplast, and mitochondrial genomes make it suitable for cell studies [42, 45, 46, 47, 48, 49]. Additional genetic and molecular tools have been developed, such as gene silencing with artificial microRNA [50] or CRISPR-mediated targeted gene disruption [51], making *C. reinhardtii* an excellent model organism. Further, a sequenced nuclear genome [52], resources for gene annotations [53, 54], as well as a genome-wide collection of insertion mutants is publicly available from the Chlamydomonas Resource Center [55, 56]. This mutant collection consists of over 60,000 mutant strains, covering 83 % of nuclear protein-coding genes [55]. *C. reinhardtii* is the microalgae with most genetic tools and the best studied metabolism. However, studies on various *Dunaliella*, *Chlorella*, and *Nannochloropsis* species also contributed significantly to a better understanding of lipid metabolism and criteria for microalgae cultivation [14, 57, 58, 59, 60].

In addition to deciding which algae strain to use, downstream processes such as harvesting, drying, lipid extraction, and biodiesel conversions determine the harvesting and labor costs of the entire process. Compared to small-scale laboratory applications, the productivity of large-scale cultures is often affected by contamination, evaporation, flooding, and lack of control over temperature and light distribution in open ponds. Limitations in constant light intensities or oxygen build-ups can impact the productivity of large-scale cultures in closed photobioreactors [61, 62, 12].

1.4 Lipid metabolism in *C. reinhardtii*

Optimizing technical requirements alone is not sufficient for effective biodiesel production from microalgae. It is crucial to consider several biological factors as well, e.g., the growth rate and lipid productivity of a specific microalgae strain. Lipids in *C. reinhardtii* are composed of neutral lipids, such as TAGs and DAGs, and polar lipids, such as membrane lipids. Lipid productivity, i.e., TAG synthesis and accumulation (see section 1.4.1), is closely connected to various metabolic pathways and products, including membrane lipids (see section 1.4.2), and strongly depends on cultivation parameters,

such as light intensity, temperature, cultivation time, availability of carbon dioxide and nutrients, e.g., carbon, nitrogen, phosphorus, and sulfur [63] (see section 1.4.3).

1.4.1 TAG synthesis in *C. reinhardtii*

TAGs are a glycerol backbone esterified with three FAs [39]. They are mainly stored in extra plastidic lipid droplets in the cytosol, some close to the chloroplast. TAGs can be produced by the acyl-CoA-dependent pathway, the Kennedy pathway, and the acyl-CoA-independent pathway. The acyl-CoA-dependent pathway occurs in the endoplasmatic reticulum (ER) and with another activated acyl donor molecule also in the chloroplast [63, 64]. The Kennedy pathway in the ER uses acyl-CoA varieties as substrates to acylate the glycerol backbone [65]. Acyl-chains covalently bound to an acyl-carrier protein (ACP) are directly used for TAG synthesis in the chloroplast.

TAG synthesis starts with glycerol-3-phosphate (G3P) and is shown in Figure 1.1. Glycerol-3-phosphate dehydrogenase (GPDH) catalyzes the reversible reduction of dihydroxyacetone phosphate (DHAP) into G3P. G3P is converted into lysophosphatidic acids (LPA) by incorporating one FA species by the glycerol-3-phosphate acyltransferase enzyme (GPAT). LPA is then again acylated with a FA by lysophosphatidic acid acyltransferase (LPAT), resulting in phosphatidic acid (PA). The phosphatidic acid phosphatase (PAP) catalyzes the dephosphorylation of PA to DAG. The last step of the TAG synthesis, also believed to be the rate-limiting step, is the acylation of DAG to TAG via a diacylglycerol acyltransferase (DGAT) [63, 66, 67].

There are two ways to generate acyl chains for TAG synthesis. One way is the *de novo* FA synthesis. It starts in the chloroplast by converting acetyl-CoA to malonyl-CoA by acetyl-CoA carboxylase (ACCase). This is the rate-limiting step in FA synthesis. Afterward, malonyl-CoA is transferred to an ACP by malonyl-CoA transacylase. Malonyl-ACP is used with acetyl-CoA by the fatty acid synthase (FAS), a multi-enzymatic complex, to form new acyl chains through elongation. The generated acyl-ACP molecules are directly used for further lipid synthesis steps in the chloroplast or are hydrolyzed by different fatty-ACP thioesterases (FAT) into three free FAs (FFA) and a glycerol backbone [63, 64, 68]. These FFAs are exported from the chloroplast and processed further in the extra-chloroplastic compartment. The enzyme long-chain acyl-CoA synthetase (LACS) links these FFAs with coenzyme A (CoA) under ATP consumption, building the acyl-CoA substrates used for TAG synthesis in the ER [69]. *De novo* FA synthesis of acyl chains is essential for TAG synthesis, as TAG accumulation decreased by almost 80 % when the FA synthesis inhibitor cerulenin was used to block the synthesis [70].

The acyl-CoA-independent pathway, the second TAG synthesis route, uses acyl-lipids from membrane lipids to acylate DAG to TAG in the last synthesis step.

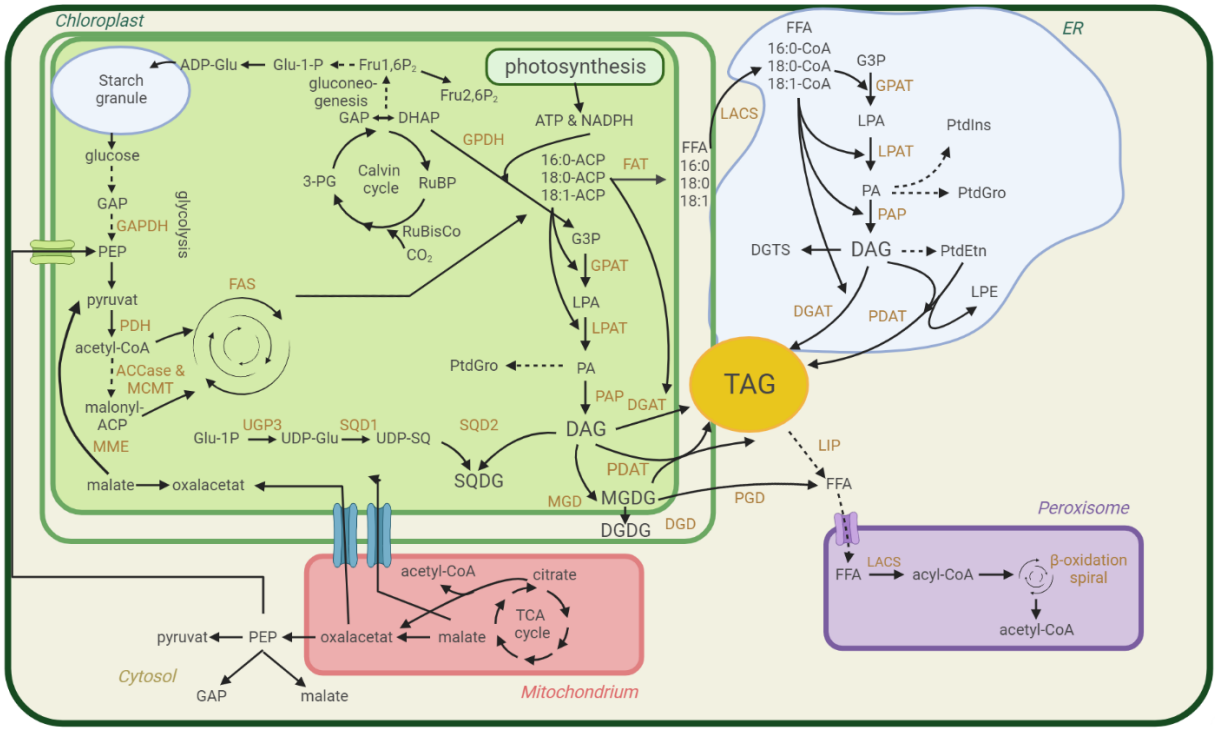


Figure 1.1: Metabolic overview of *C. reinhardtii*

Overview of the central metabolic pathways related to lipid and starch synthesis in *C. reinhardtii*. The metabolic substrates are represented in black, the enzymes in orange. Dashed lines represent multiple reactions. Intermediate reactions were omitted for clarity. 3-PG: 3-phosphoglycerate; ACCase: acetyl-CoA carboxylase complex (composed of carboxyltransferase, carboxylase and Acetyl-CoA biotin carboxyl carrier protein); ACP: acyl carrier protein; ADP-Glu: ADP-Glucose; ATP: adenosine triphosphate; CoA: coenzyme A; CO_2: carbon dioxide; DAG: diacylglycerol; DGAT: diacylglycerol acyltransferase; DGD: digalactosyldiacylglycerol synthase; DGTS: diacylglycerol-N,N,N-trimethylhomoserine; DHAP: dihydroxyacetone phosphate; ER: endoplasmatisches reticulum; FAS: fatty acid synthase; FAT: fatty acyl-ACP thioesterase; FFA: free fatty acid; Fru-1,6-P_2: fructose-1,6-bisphosphate; Fru-2,6-P_2: fructose-2,6-bisphosphate; G3P: glycerol 3-phosphate; GAP: glyceraldehyde 3-phosphate; GAPDH: glyceraldehyde 3-phosphate dehydrogenase; GPDH: glycerol-3-phosphate dehydrogenase; GPAT: glycerol 3-phosphate acyltransferase; Glu-1P: glucose 1-phosphate; LACS: long-chain acyl-CoA synthetase; LIP: TAG lipase; LPA: lysophosphatidic acid; LPAT: lysophosphatidic acid acyltransferase; LPE: lysophosphatidylethanolamine; MCMT: malonyl-CoA:acyl carrier protein malonyltransferase; MGD: monogalactosyldiacylglycerol synthase; NADPH: nicotinamide adenine dinucleotide phosphate; PA: phosphatidic acid; PAP: phosphatidic acid phosphatase; PDAT: phospholipid:diacylglycerol acyltransferase; PDH: pyruvate dehydrogenase; PEP: phosphoenolpyruvate; PGD: galactoglycerolipid lipase ; PtdEtn: phosphatidylethanolamine; PtdGro: phosphatidylglycerol; PtdIns: phosphatidylinositol; RuBP: ribulose 1,5-bisphosphate; RuBisCo: ribulose-1,5-bisphosphate carboxylase-oxygenase; SQDG: sulphoquinovosyl diacylglycerol; SQD1: UDP-sulfoquinovose synthase; SQD2: Sulfolipid synthase; UDP-Glu: uridine diphosphate glucose; UDP-SQ: UDP-sulfoquinovose; UGP1: UDP-glucose pyrophosphorylase; TAG: triacylglycerol. The pathways are simplified to highlight the major routes in central carbon and lipid metabolism leading to triacylglyveride and starch storage. Figure was adapted from Goncalves et al. (2016) and created with BioRender.com. References for the described pathways are: [71, 72, 64, 73, 74, 63, 75, 76, 77, 40, 66].

Thus, in addition to FAs that are newly synthesized and subsequently used for TAG synthesis, a large proportion of pre-existing FAs are converted from membrane lipids to be incorporated into TAGs. This final step is catalyzed by phospholipid:diacylglycerol acyltransferase (PDAT). It was shown that the PDAT enzyme in *C. reinhardtii* has a broad substrate specificity and can use a variety of phospholipids as well as galactolipids for the transacylation reaction [76]. Depending on the membrane lipid used, this synthesis pathway occurs in the ER or in the chloroplast [63, 64].

1.4.2 Membrane lipid synthesis in *C. reinhardtii* and their remodeling routes

Membrane lipids in *C. reinhardtii* can be divided into two major lipid categories: phospholipids and glycolipids. Membrane lipids are defined by different hydrophilic head groups attached to the *sn*-3 position of the DAG backbone. These headgroups define the physicochemical properties of the different membrane lipids [78]. In *C. reinhardtii*, phosphatidylglycerol (PtdGro), phosphatidylethanolamine (PtdEtn), and only one species of phosphatidylinositol (PtdIns) represent the phospholipids, whereas monogalactosyldiacylglycerol (MGDG), digalactosyldiacylglycerol (DGDG) and sulfoquinovosyldiacylglycerol (SQDG) are the glycolipids. In contrast to other algae or terrestrial plants like *A. thaliana*, *C. reinhardtii* lacks phosphatidylcholine (PtdCho), which usually represents the major membrane lipid of plastid membranes [79, 40, 80]. The most abundant membrane lipid in *C. reinhardtii* is MGDG [40, 81]. In addition, *C. reinhardtii* possess a betaine lipid, diacylglycerol-*N,N,N*-trimethylhomoserine (DGTS), which has similarities in structure and biophysical properties to PtdCho and which might fulfill the function of the missing PtdCho [40, 70, 63].

In addition to the different head groups, all phospholipids and glycolipids are esterified with two FAs on the glycerol backbone. Since *C. reinhardtii* and *A. thaliana* derive from the same ancestor, they share a similar FA profile [52]. Their FA profile consists of FAs with 16 and 18 carbons, which are saturated, monounsaturated, and polyunsaturated. The main species in *C. reinhardtii* are C16:0, C18:1, C18:3, and C16:4 (see Table A.2 in the appendix). They account for about 64 % of all FAs in *C. reinhardtii* [63, 82, 81].

The FAs vary in length and degree of saturation. They influence the membrane's physical properties such as fluidity, permeability, bilayer thickness, charge and intrinsic curvature [78]. The conversion of saturated into unsaturated FAs is catalyzed by fatty acid desaturases (FADs), localized in the chloroplast and ER. FADs introduce double bonds into saturated FAs and thereby generate unsaturated FAs [63]. Membrane lipids in *C. reinhardtii* contain enriched PUFAs, like C18:3, C16:4 and C18:2. Especially C16:4, a unique characteristic of *C. reinhardtii*, is almost exclusively present in MGDGs [83, 40]. In contrast, TAGs in *C. reinhardtii* usually contain SFAs and MUFAs, such as C16:0, C16:1, and C18:1 [82, 83]. As previously described, *C. reinhardtii* contains almost equal

amounts of SFAs (24 %) and MUFAs (23 %) and around 50 % of PUFAs. A proportion of over 50 % of MUFAs and a balanced ratio between SFAs and PUFAs are required to produce high-quality biodiesel. Since MUFAs and SFAs are the main parts of TAGs, TAG synthesis in *C. reinhardtii* needs to be favored over membrane lipid synthesis to shift the lipid composition toward biodiesel production. However, not completely, since PUFAs with long-chain FAs are necessary to maintain the low-temperature flow proprieties as viscosity by lowering the biodiesel's overall melting point [32, 33].

As mentioned before, in addition to *de novo* FA synthesis, alternative acyl chain flux routes are described that provide acyl chains for TAG synthesis in *C. reinhardtii*. One route uses preexisting acyl chains of membrane lipids that are extracted from membrane lipids and then incorporated into TAGs. Another route uses newly synthesized acyl chains that are first esterified in membrane lipids as intermediates before being remodeled and re-extracted to be incorporated into TAGs [75, 84]. All FA flux routes were validated in several studies using labeling experiments with ^{13}C isotopes. It was shown that preexisting FAs of membrane lipids contribute much more to TAG synthesis than previously assumed. More precisely, up to 35-45 % of incorporated acyl chains from membrane lipids are remodeled and integrated into TAGs [85]. The remodeling and extraction of FAs from membrane lipids are accompanied by a reduction of membranes in the cell. A report of a corresponding reduction in chloroplast membranes during nutrient stress confirmed this [77]. Interestingly, this mechanism is reversible, resulting in cleavage of TAGs again into DAG and FFAs by hydrolase activity [76]. This recycling phenomenon was also predicted through transcriptomic analysis in *Coccomyxa subellipsoidea* [86]. Quantification of this phenomenon in *C. reinhardtii* confirmed that acyl chains from TAGs contribute up to 63 % to the resynthesis of membrane lipid within the first 12h after nutrient resupply [85, 87]. Thus, lipid remodeling pathways enable the exchange of FAs and the restructuring of FA composition in the cell [85]. Still, it remains to be determined how these flux routes can be guided toward each other.

1.4.3 How cultivation conditions affect microalgae's lipid metabolism and improve TAG production

TAG production and accumulation can be increased by several factors, in particular cultivation conditions, e.g., high temperatures and bright light affect microalgae's growth rate and productivity. Too high temperatures impair the growth of microalgae, as they damage cells and lead to cell death [38, 15, 88]. High light intensities, on the other hand, can accelerate lipid synthesis. For instance, *C. reinhardtii* accumulates high amounts of TAG and neutral lipids when exposed to saturated or high light conditions, i.e. > 200 µmol m^{-2} s^{-1} at constant CO_2 supply [89]. Nonetheless, overfull light exposure can also have the opposite effect. Studies with the microalgae species *Chlorella sp.* and *Monoraphidium dy.* have shown that photosynthesis, especially photosystem II (PSII) activity, declines at higher light intensities [90]. Excessive light intensities can lead to ox-

idative stress by forming reactive oxygen species (ROS) that damage the photosynthetic machinery. This mechanism, known as photoinhibition, can impair the metabolism of microalgae [90, 91, 92].

Since lipid synthesis requires large amounts of NAD(P)H along with acetyl-CoA and ATP to reduce acyl chains, an increased TAG accumulation under high light conditions is thought to prevent over-reduction of the photosynthetic electron transport chain. Thus, elevated TAG accumulation prevents photooxidative damage by using photosynthetic reductants and fixating carbon [19, 85]. This indicates that the TAG biosynthesis plays an active role in the light stress response of microalgae. Thus, TAGs have an additional purpose besides their natural function as carbon and energy storage.

Another important factor is the cultivation time that influences TAG production. Usually, high cell density is achieved by longer growth times. With a higher cell density, a higher biomass is achieved, which is often accompanied by a high lipid yield. Nevertheless, microalgae proliferate in the log phase and form metabolites in the stationary state. Hence, long growth times, e.g., over 20 days, can lead to higher lipid accumulation than active growing cultures in the log stage [93, 94, 88]. However, none of these parameters affects microalgal TAG accumulation and metabolism to the same extent as changes in carbon, nitrogen, phosphorus, or sulfur nutrients in the media. Overproduction of lipids allows microalgae to buffer extreme changes in environmental conditions [19, 95, 96, 97], which explains why lipid accumulation changes during stress.

Nutrient starvation or limitation such as nitrogen (N), sulfur (S), or phosphorus (P) deprivations are the main lipid inducer in green microalgae. Nutrient limitation is the most intensively studied and successful strategy to increase TAG accumulation plus it can be easily controlled, manipulated, and monitored [19, 82, 70, 74, 98, 99].

Besides carbon, N is *C. reinhardtii*'s most important nutrient for growth and division [100]. *C. reinhardtii* can utilize various N sources, such as inorganic substances like nitrate, nitrite, and ammonium, and organic substances like urea, urate, and amino acids. N starvation leads to increased TAG accumulation but decreased protein production. The enhanced TAG production is a survival strategy to store reduced carbon when insufficient N nutrients exist for growth and cell division [19, 82, 101, 102, 103, 104, 105]. The immediate transcriptomic response to N starvation is a strong gene expression upregulation of transporters for N, ammonium, and complex N-containing organic compounds and activation of degradation pathways to release N from existing proteins. A recycle mode is activated by which abundant proteins like Rubisco and ribosomes are degraded to release bound N [104, 106, 107]. Also, photosynthesis-related proteins, such as PSII, cytochrome b_6f, photosystem I (PSI), ATP synthases and molecular chlorophyll are degraded causing the cellular phenotype of chlorosis [108]. This reduces the amount of total protein by ~50 % and total RNA by ~60 % per cell. Furthermore, N-deprivation leads to a reduction in membrane lipids and a remodeling of the released FAs into TAGs,

as described before, thereby increasing TAG concentration in the cell [85]. However, this remodeling process is rescued after N resupply. Nevertheless, the rapid degradation of chlorophyll and photosynthesis-related proteins, as well as the reduction of membranes restricts the overall growth, limits expansions in lipid yield and, thus, biotechnological productivity.

In contrast, P and S starvation can lead to an enhanced TAG accumulation without disrupting photosynthesis integrity [109]. These stress responses are also characterized by a gene expression change and increased mobilization and transport mechanisms for P and S, respectively. Furthermore, a general metabolic slowdown, cell division stagnation, and the reduction of photosynthesis activity are also common stress responses to P and S limitations [110].

An additional stress reaction to N depletion is an increased starch formation in *C. reinhardtii*. Starch is composed of amylopectin and amylose, two differently structured glucose-unit polysaccharide molecules made up of glucose-1-phosphate, generated via gluconeogenesis from triose phosphates [111]. *C. reinhardtii* produces starch under phototrophic conditions by photosynthetic carbon fixation and also under heterotrophic and mixotrophic conditions by using carbon-reduced compounds. Starch is produced faster than TAGs but re-consumed and converted into glucose after a few hours of N starvation [112]. TAGs are an energy reservoir that is 2.25 times more energetic than starch [112]. TAG and starch synthesis compete for carbon C3 precursors since their metabolism starts from the identical pool of glyceraldehyde 3-phosphate (GAP) or dihydroxyacetone phosphate (DHAP) and 3-phosphoglycerate (3-PG). The metabolic connections between lipid and starch metabolism are illustrated in Figure 1.1. The competition between starch and lipids leads to the partitioning of the carbon precursors and to the still-unknown regulatory question of which synthesis pathway - starch or TAG - is preferred. TAG production in *C. reinhardtii* is increased by 3.5 - 8-fold compared to wild type when starch biosynthesis is impaired, e.g., in the *C. reinhardtii* starchless mutant *sta6* [64, 113, 104]. However, complete inhibition of starch synthesis, e.g., by a disruption mutation of the starch synthesis enzymes, does not lead to an increased TAG accumulation, as expected, but has harmful effects on the growth of *C. reinhardtii* [114, 95]. This hinders the biotechnological application of this mutant, as total TAG production is limited due to growth stagnation.

In summary, the entire lipid metabolism of *C. reinhardtii* composed of *de novo* FA synthesis, FA activation and desaturation, the synthesis and turnover of membrane lipids, FA β-oxidation, and the FA activation and transport influence TAG biosynthesis and the lipid composition in *C. reinhardtii*. In addition it is highly interwoven and dynamic [68]. Various environmental influences, e.g., N starvation and readdition, trigger specific cellular responses leading to different lipid compositions and an adaptation of primary metabolism in the cell [86]. Because metabolic processes are complex, they must be highly controlled to maintain homeostasis in the cell [115].

Especially, a systematic understanding of the regulation of the different TAG and lipid synthesis routes and the remodeling mechanisms are required to be able to influence them. It would be highly beneficial for finding targets to optimize metabolic processes in microalgae, especially for producing substantial amounts of TAGs under normal growth conditions without stress [82].

1.5 Different layers of regulation control all cellular processes

Regulatory mechanisms are essential for an organism to handle the effects of perturbation, maintain its homeostasis and modulate its constitutive dynamics [116]. Cellular regulation happens on different regulation layers and are mediated by the synergy of proteins with other proteins, DNA, RNA and small molecules. Most metabolic processes are regulated through various mechanisms, including short-term and long-term regulation. Short-term regulation involves rapid, reversible changes in enzyme activity or substrate availability that can alter metabolic flux within minutes or seconds. Allosteric regulation and protein phosphorylation are crucial examples of short-term regulation [117, 118]. One prominent example for allosteric regulation can be found in the starch synthesis pathway. The major rate-controlling step in starch synthesis is ADP-glucose synthesis by ADP-glucose pyrophosphorylase. The allosteric binding of 3-PG activates the enzyme through a conformational change, whereas the binding of orthophosphate inhibits it [119, 120]. Since 3-PG is a precursor substrate for the formation of glucose-1-phosphate, which is converted to ADP-glucose by ADP-glucose pyrophosphorylase, this is a classical feed-forward regulation loop.

The enzyme phosphofructokinase-1 (PFK-1) represents another well-known example of an allosterically controlled regulator. It is sensitive to the adenosine triphosphate (ATP) level of the cell and can be regulated by numeric allosteric effectors, including ATP, adenosine monophosphate (AMP), fructose-6-phosphate (Frc-6-P), fructose-2,6-bisphosphate (frc-2,6-P_2) and citrate. It is one of the most important regulatory enzymes of glycolysis. It catalyzes the transfer of ATP to fructose-6-phosphate to build fructose-1,6-bisphosphate and can, thus, determine the rate of glycolysis in response to the cell's energy requirements. Interestingly, fructose-2,6-bisphosphate, a product of phosphofructokinase-2 (PFK-2), also known as *FBP4* in *C. reinhardtii*, is the most potent activator of PFK-1. PFK-2 which, in turn, is regulated by phosphorylation. Phosphorylation represents another important short-term regulatory mechanism. In this process, a phosphate group is added to a hydroxyl group residue of the enzyme, which can also have either an activating or inhibiting effect [117]. The dephosphorylation of PFK-2 (*FBP4*) activates its kinase activity, resulting in enhanced synthesis of frc-2,6-P_2 and subsequently increased activity of PFK-1 with consecutive accelerated glycolysis rate.

Since PFK-2 (*FBP4*) is a bifunctional enzyme, its phosphorylation leads to the opposite reaction. In its phosphorylated form, PFK-2 expresses phosphatase activity and causes a breakdown of Frc-2,6-P_2 to Frc-6-P, which leads to an inhibition of glycolysis and stimulates gluconeogenesis [121].

Long-term regulation, conversely, involves changes in gene expression and protein synthesis, leading to alterations in the abundance and activity of enzymes and proteins. Transcriptional regulation occurs over a longer time scale, typically hours or days, and is essential for adaptation to changes in environmental conditions or developmental stages. Transcriptional regulation acts directly at the DNA level. It determines when and which gene is expressed, whereas other types of regulation often adjust or reduce gene expression rather than determine it [122].

Transcriptional regulation is guided by multiple components, first, by transcription factors (TF). TFs are proteins that can bind to *cis*-regulatory elements (CREs), which are *cis*-acting non-coding DNA regions, and induce or repress thereby the transcription of target genes [122]. The gene expression of other TFs is also regulated by TFs. Furthermore, multiple TFs often regulate the fine-tuning of a defined cellular response. Second, by transcriptional regulators (TR) that do not carry characterized DNA binding domains but influence transcription by other mechanisms, such as chromatin remodeling [123]. In a dynamic process, they change the chromatin architecture by introducing chromatin modifications, e.g., histone acetylation and methylations. The opening of chromatin enables the transcriptional machinery to access specific genomic regions and start the gene expression of a desired gene and vice versa [124].

This can result in a specific temporal gene expression of the desired genes. Specific gene expression is necessary to control development and physiology in a cell and allows the cell to react to environmental changes. Gene expression of *gap* and *pair-rule* in *Drosophila* embryos is a famous example for transcriptional regulation. Spatiotemporal gene expression is crucial for defining the segmentation pattern, and further development of the *Drosophila* embryo [122, 125]. The MADS-box transcription factors in *A. thaliana* represent another prominent example of tissue-specific gene expression. The transcription factors (TF), APETALA3 (AP3) and PISTILLA (PI) are required for specific petal and stamen development in the *A. thaliana* flower [126, 127, 128]. When mutated, petals are transformed into sepals and stamens into carpels. Tissue-specific expression of the MADS-box genes is critical for the proper flower organ development in *A. thaliana*, but their gene expression is tightly regulated by a complex network of other transcriptional regulators to ensure accurate flower development [128].

The lipid metabolism in *C. reinhardtii* has been intensively studied. Nevertheless, it is very challenging to comprehensively describe the interplay between the various metabolic pathways and their respective regulation [71]. As indicated, several external influences are known that lead to increased TAG accumulation, but how the transcrip-

tional regulation of these metabolic processes occurs remains to be unraveled. Nevertheless, over the past decade, some distinct regulatory factors affecting lipid metabolism in *C. reinhardtii* were identified using transcriptomic analysis and phenotypic characterization of *C. reinhardtii* mutants [74, 101, 108, 129, 130, 131, 132, 109].

1.6 Known regulators of lipid metabolism in *C. reinhardtii*

The squamosa promoter-binding protein Nitrogen Response Regulator 1 (NRR1) is one TF that was identified as a positive regulator of lipid accumulation under N depletion by transcriptomic analysis. The insertion mutant *nrr-1-1* accumulates only 50 % of TAGs compared to its parental strain CC-4425 under N deprivation conditions. NRR1 seems to be a specific regulator during N deficiency, as other nutrient deficiencies do not affect the abundance of its transcript, which is increased under N deficiency [74]. In addition, three ER stress-sensing proteins that rapidly accumulate TAGs under ER stress conditions were identified, the Inositol-Requiring Enzyme 1 (IRE1) and the two basic leucine zipper (bZIP) TFs bZIP1 and bZIP2 [133, 134, 135].

Another TF that regulates carbon storage metabolism under P and N deficiency by influencing gene expression of specific starch and lipid synthesis genes is Phosphorus Starvation Response 1 (PSR1), a MYB-like (myeloblastosis) TF. PSR1 was initially introduced as a TF that regulates P uptake and extracellular mobilization [136, 137]. Later, transcriptomic analyses revealed that it is rather a general regulator of lipid metabolism under N starvation [138]. In particular, the *psr1* mutants showed a disturbed gene expression of a high number of starch and lipid metabolism genes and subsequent inhibition of starch and lipid synthesis. Overexpression of *PSR1* resulted in an increased starch and TAG concentration, most likely triggered by transcriptional control of specific starch and lipid metabolic-related genes [138, 139].

Another MYB TF, the Lipid Remodeling Regulator 1 (LRL1), was identified as a regulator for TAG synthesis and membrane remodeling under P starvation. *lrl1* is upregulated under P deprivation, and the mutant strain showed apparent repression of genes typically associated with lipid remodeling, like diacylglycerol sulfoquinovosyl transferase 2 (SQD2) and diacylglycerol acyltransferases (DGATs) (see Figure 1.1). Consequently, TAG and SQDG accumulation was suppressed in P-depleted mutants [109]. MYB1, also a MYB TF, was characterized in 2022 as a key regulator of lipid metabolism under stress conditions, especially N depletion. The *myb1* mutants accumulate 35-40 % fewer TAGs and 65-70 % fewer total FAs than their parental strain during N starvation. Indeed, genes that drive lipid transport from the chloroplast to the ER were strongly expressed in the wild-type strain but not in the *myb1-1* mutant under N depletion [140]. The transcriptomic analysis revealed further that MYB1 is upregulated upon TAG-inducing conditions and coexpressed with many other lipid-metabolism associated genes. The analysis of the MYB1 orthologue in microalgae *Chromochloris zofingiensis* strengthens the

hypothesis that it regulates these genes since the MYB1 orthologue binds to multiple lipid metabolism-related promoters in a yeast-one-hybrid (Y1H) assay [141]. It appears that MYB TFs play an essential role in *C. reinhardtii*, as three TFs of the known key regulators belong to this TF family.

In addition to TFs, two members of the Dual-Specificity Tyrosine-Phosphorylation-Regulated Kinase (DYRK) family also regulate lipid metabolism under N depletion. First, Triacylglycerol Accumulation Regulator 1 (TAR1), a member of Yet another Kinase 1 (Yak1) subfamily and the plant-specific DYRK (DYRKP). Knock-out mutant *tar1-1* accumulates only 10 % TAGs compared to its parental strain under mixotrophic N-limited conditions [142] but accumulates more TAGs than the wild-type during photoautotrophic N starvation [143]. In contrast, the *drykp* mutant exhibits a different phenotype than the *tar1-1* mutant. It produces more TAGs after N deprivation, regardless of the presence of an external carbon source [144].

Since S deficiency triggers a different stress response in *C. reinhardtii*, it is only natural that other global regulators could be identified in *C. reinhardtii* that mediate the sulfur-depletion related response: Sulfur Acclimation 1 (SAC 1) and SNF1-Related Protein Kinase 2.2 (SNRK2.2). SAC 1 is a high affinity sulfate transporter at the plasma membrane that can sense a shortage of S and then transduce this signal to upregulate the transcript level of a particular set of genes [145, 98, 146]. The SNRK2.2 protein (also known as SAC3) belongs to the serine/threonine kinase family and regulates the physiological response to environmental S-levels [98]. During S deficiency, SAC1 and SNRK2.2 promote the degradation of SQDG. SGDG, a specific thylakoid membrane lipid, stabilizes the photosynthetic apparatus. Its degradation leads to the release of S and supports the resynthesis of other proteins in the early phase of S deficiency [147, 148]. In addition, SAC1 and SNRK2.2 appear to induce the expression of several TAG synthesis related genes that direct metabolic flux toward TAG synthesis. However, the exact mechanism of how they mediate this response has yet to be discovered [98].

Despite the tremendous progress and although many regulators have been already identified, overall systematic knowledge about the regulatory correlations, metabolic mechanisms, and regulatory pathways of stress-induced lipid accumulation remains obscure. In addition, the regulatory responses to various stresses use different regulation patterns and involve various regulators. Different gene expression profiles, multiple regulators, and varying phenotypes support this conclusion.

Therefore, a systematic approach is the logical consequence to unravel the complexity of transcriptional regulation in *C. reinhardtii*. In particular, only a systematic understanding of the coordination and regulatory relationships of lipid metabolism in *C. reinhardtii* will help to specifically influence metabolic processes and, thus, the production of desired products. Furthermore, it allows the identification of new regulatory targets and regulators.

1.7 Network maps describe biological relationships

A systematic analysis provides information about correlations and connection between interacting partners, such as proteins with DNA or proteins with each other. The mapping and analysis of protein-protein interactions (PPIs) and protein-DNA interactions (PDIs) on a small- or large-scale has contributed significantly to unraveling fundamental questions in biology [149, 150, 151, 152, 153]. Performed in a high-throughput manner, PPI and PDI screens investigate the interactions of a large number of proteins to each other or, respectively, to specific DNA molecules.

1.7.1 Interactomes are very advantageous for understanding complex biological relationships

The yeast-two-hybrid (Y2H) system identifies binary protein-protein interactions and was the first genetic assay for PPIs [154]. Today, it is one of the most widely used cost-effective methods for detecting PPIs [149]. The interaction partners of a Y2H assay are divided into bait and prey, but both are proteins. One protein (prey) is fused to the N-terminal activation domain of the yeast Gal4 transcription factor (Gal4-AD). The second protein (bait) is fused to the DNA binding domain, here the Gal4 DNA-binding domain (Gal4-DB), the C-terminal part of the yeast Gal4 transcription factor [155]. Expression of both fusion proteins in yeast plus their interaction result in the reconstitution of a functional Gal4 TF (see Figure 1.2). This leads to an activation of one or several reporter genes, like *HIS3* [154].

A Y2H screen variant, called Barcode Fusion Genetics (BFG)-Y2H, differs from state-of-the-art Y2H methods and uses deep sequencing to improve the identification of PPIs in pools [156]. BFG-Y2H enables screening a whole matrix of protein pairs in a single multiplexed yeast pool [156]. Every ORF destination vector contains two unique barcodes. Cre recombination fuses these DNA barcodes of the interacting diploid yeast strains, generating chimeric protein-pair barcodes that can be identified using next-generation sequencing (NGS) [156]. The efficiency of protein matrix screening is highly increased by using BFG-Y2H.

The resulting network maps are very advantageous for understanding complex biological relationships that are usually not visible at first glance. PPIs from a large-scale study are mapped in an interactome. Interactomes have helped answer fundamental questions in plant biology. For example, an interactome study of virulence effector proteins to their host plant proteins identified common hubs, which are highly connected nodes in the network that are targeted collectively by these evolutionary distant pathogen [158]. The independently evolved pathogens interacted with the same plant target proteins, which was only visible through the systematic approach and interactome analysis. The identified plant protein targets included well-connected cellular hubs. The

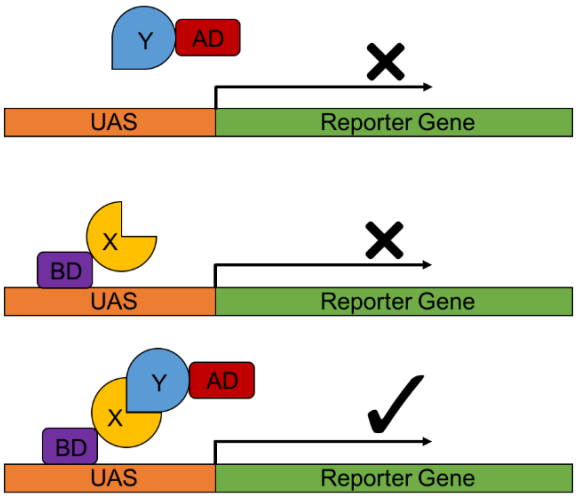


Figure 1.2: Schematic diagram of a Y2H assay

Y (blue) represents the prey protein, a fusion protein with the Gal4-AD domain, indicated as red AD. X (yellow) stands for the bait protein, that is fused to the Gal4-BD domain (purple). Prey and bait alone can not activate the gene expression of the reporter gene. Only through the interaction of prey and bait protein does the activation domain get close enough to the DNA to initiate gene expression of the reporter gene. Adapted from Lifeasible® [157].

genetic validation supports the concept that effectors likely converge on interconnected host machinery to suppress effective host defense and increase disease susceptibility of *A. thaliana* [158, 149].

Another *A. thaliana* interactome study has revealed higher interconnectivity between plant signaling pathways than previously reported and expected. Plant hormones regulate stress responses to environmental cues and have been examined by genetic screens, hypothesis-driven approaches, and interactome studies of selected pathways [159]. Interestingly, much fundamental information about crosslinks between the different pathways was missing. A systematic PPI map of *A. thaliana* phytohormone signaling network discovered more than 2,000 binary interactions in a highly interconnected network. Pathway communities and hundreds of unknown contact points between the different signaling pathways were identified, representing a crosstalk network, indicating that a large majority of signaling proteins function pleiotropically in several pathways [150].

The first human interactome network map highlights the beneficial potential of PPI networks for understanding biological concepts. Analysis of the PPI network map showed that proteins involved in the same disease were more interconnected and formed subnetworks [160, 161]. These valuable insights into the complex protein relationships in a cell provide clues for therapeutic targets and help to develop new hypotheses.

The contactome of 26 viral SARS-CoV-2 proteins against 17,472 human proteins showed that SARS-CoV-2 proteins interact enriched with human proteins, involved in immune signaling, inflammation, protein ubiquitination, and membrane trafficking [162]. Furthermore, it was found that the SARS-CoV-2 target network communities were enriched for human proteins genetically associated with comorbidities of severe illness and long COVID. The contactome links viral-human protein interactions to genome-wide association studies (GWAS) that identified genetic risk loci for severe coronavirus disease 2019 (COVID-19) and proposes potential therapeutic targets [163].

1.7.2 Gene regulatory networks describe transcriptional regulation

PDI networks, also known as gene regulatory networks (GRN), are equally crucial for studying fundamental biological processes. PDI networks reflect regulatory relationships and how TFs control biological processes in a cell [164]. They provide information about the regulatory and hierarchical composition of TFs, like cross links and connections between multiple TFs and DNA regions that can only be detected in a network [165, 166, 167].

Since TFs and TRs can act as activators, repressors, or both [122], different TFs can interact with the same DNA region up and downstream of a target gene, resulting in a broad and comprehensive PDI network. PDI network maps have been used to investigate regulatory relationships in *C. elegans* [168, 169], *A. thaliana* [170, 171, 166, 172, 173, 174, 167, 175, 176] and humans [177]. Plant defense mechanism, plant cell regeneration after wounding, and primary and secondary metabolism regulation of *A. thaliana* were subjects of large PDI studies. They revealed novel regulators and correlations in the metabolic regulation of *A. thaliana*.

GRN can be generated experimentally or computationally. Transcriptomic data is used to construct gene correlation networks, i.e., computationally generated co-expression networks. This computational approach groups genes with highly correlated gene expression levels into clusters. Then, the network is used to identify similar biological functions in the gene clusters [178]. The relationships between TFs and corresponding genes are predicted based on their similar gene expression and then clustered. In this way, computational GRNs model physical interactions between genes and their regulators based on related gene expression [122, 179]. Thanks to NGS technology and several *C. reinhardtii* genome annotations, a large amount of transcriptomic data for *C. reinhardtii* is accessible. The web-based tool 'ChlamyNET' [180] allows users to search for one or multiple *C. reinhardtii* genes of interest (GOI) and provides information about co-expressed neighboring genes under specific expression conditions. Unfortunately, the network does not contain all *C. reinhardtii* genes and is limited due to the number of expression profiles [180]. Interestingly, only a few gene regulatory co-expression networks for *C. reinhardtii* have been published so far [181, 182, 183, 184, 180, 185, 186].

Many investigate the regulatory adaptations that *C. reinhardtii* undergoes in response to different light and CO_2 conditions and the associated CO_2-concentration mechanism [186, 183, 184, 185].

Another co-expression study used a combined omics approach that includes transcriptome, proteome, and metabolome data to predict a regulatory network that controls lipid metabolism in *C. reinhardtii* cells during N deprivation [182]. The gene expression patterns of several TFs and TRs in response to a shift from N-replete to N-deficient media were analyzed and divided into two phases related to TAG synthesis: 'before TAG synthesis' and 'after TAG synthesis'. Additional TFs and TRs were correlated with genes involved in metabolic processes different from TAG synthesis, such as N assimilation and photosynthesis, to link these TFs to a specific biological function. Some TFs were identified as hubs and subsequently categorized into specific and permanent hubs according to their temporal gene expression behavior [182]. Unfortunately, only a few TFs could be identified as regulatory hubs for the TAG synthesis. Moreover, even though this prediction of regulatory hubs is very valuable, there is neither an experimental grounded global-scale GRN nor an experimental validation of this prediction.

Experimental methods to generate a GRN and map physical PDIs can be divided into two approaches: TF-centered ('protein-to-DNA') and gene-centered ('DNA-to-protein') methods (see Figure 1.3). In TF-centered methods, all genomic DNA fragments bound by a given TF are identified [122]. In this way, genome-wide TF binding sites (TFBS) can be detected for a specific TF.

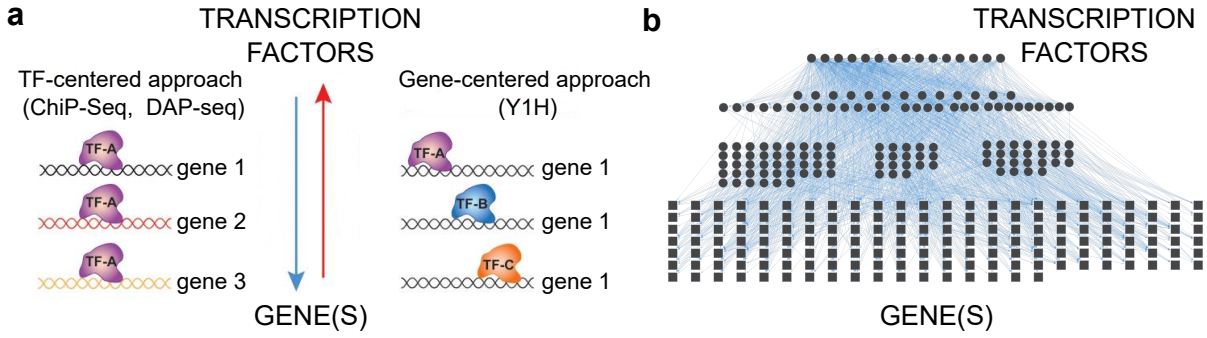


Figure 1.3: Methods to identify protein-DNA interactions

a Adapted from Arda et al. (2010) [122]. Two conceptually different experimental methodologies to identify protein-DNA interactions. TF-centered approaches begin with a TF of interest and identify DNA fragments, that are bound by this TF. Examples are, e.g., ChIP-seq and DAP-seq. Gene-centered approaches investigate a defined DNA fragment collection and identify different TFs that interact with these DNA fragments.

b Schematic representation of a gene regulatory network. Different hierarchical levels are indicated. The top layer represents the key regulators, TFs and TRs (circles), whereas the middle layers represent TFs that are targeted by the top layer TFs and also interact with downstream promoters and thus regulate them. The last level represents the DNA promoters of genes (squares).

Chromatin-immunoprecipitation followed by high-throughput sequencing (ChIP-seq), DNA adenine methyltransferase identification (DamID), and *in vitro* DNA-binding experiments like DNA-affinity purification sequencing (DAP-seq) and protein binding microarrays (PBM) are well-known TF-centered methods [122, 115, 187, 188, 189, 190].

Gene-centered methods, on the other hand, start from one or more specific DNA fragments and identify multiple TFs that can interact with this DNA fragment [122]. The yeast-1-hybrid (Y1H) method is the classical gene-centered approach that can also be performed in a high-throughput manner to detect PDIs and generate large GRNs. It reveals TF binding to specific regulatory DNA regions and is based on two components: 'DNA baits' and 'protein preys'. DNA baits are specific DNA fragments that are considered regulatory. Large DNA fragments such as gene promoters of 500 - 2,000 base pairs (bp) or small DNA fragments can be used as DNA baits [191]. They are placed upstream of one or two Y1H reporter genes: *HIS3* and *LacZ*. Each DNA bait:reporter construct is randomly integrated into the yeast genome, resulting in a reporter bait yeast strain (see chapter 4). A plasmid encoding the TF of interest fused to the activation domain of the yeast transcription factor Gal4 (Gal4-AD) is transformed into a prey yeast strain. Mating of the two complementary yeast strains is followed by the selection for diploid yeast cells. When the TF binds to the DNA fragment, the Gal4-AD moiety activates the expression of the reporter gene. In the case of the *HIS3* reporter and interaction between TF prey and DNA bait, the yeast can grow auxotrophic on selection media containing 3-amino-triazol (3-AT) (see Figure 1.4) [192]. The *LacZ* reporter reveals a positive interaction by conversion of X-gal into 5-bromo-4-chloro-3-hydroxyindole. Two 5-bromo-4-chloro-3-hydroxyindole molecules react and oxidize to 5,5'-dibromo-4,4'-dichloro-indigo, which stains the growing yeast dark blue [193].

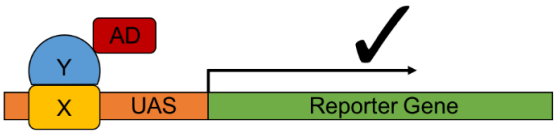


Figure 1.4: Schematic diagram of a Y1H assay

X (yellow) describes a DNA bait, a promoter sequence, here in front of an Upstream Activation Sequence (UAS) and a reporter gene. Y (blue) indicates the TF prey, that is fused to the Gal4-AD domain (red). A positive interaction occurs when prey and bait interact with each other. It leads to the gene expression activation of the reporter.
Adapted from Lifeasible® [157].

The described methods have advantages and limitations but are powerful tools for generating GRNs. They complement each other and advance our understanding of regulatory relationships in biological systems. Despite previous successes and systems-level studies for *C. reinhardtii*, there is an urgent need for a deeper understanding of metabolic and regulatory processes to fully exploit its biotechnological potential.

1.7.3 Impact of a *C. reinhardtii* ORFeome

Y1H and Y2H assays rely on the expression of a protein of interest (POI) in a controlled experimental environment. To express a protein, its DNA coding sequence, a so-called open reading frame (ORF), needs to be available and possible to clone into various reporter constructs, usually as covalent fusion proteins with well-characterized protein domains, called 'tags'. For Y1H and Y2H, the Gal4-AD domain represents such a 'tag'. These 'tags' provide important experimental anchors or functional properties that enable the detection and analysis of the desired ORF [194]. In particular, for high-throughput studies, a collection of ORFs, an ORFeome, is required. It allows functional genomic applications, such as Y2H analysis, and affinity-purification mass spectrometry (AP-MS) analysis [195]. They are perfect for conducting experiments to determine the functional roles of specific proteins, such as TF regulatory studies, enzymatic activity experiments and systematic *in vivo* protein localization assays. The generation of an ORFeome collection bridges 'whole genome sequencing' efforts to downstream 'omics' applications' [196].

Many ORFeome collections used the Gateway® cloning system. It supports flexible recombinant transfer of ORFs from cloning vectors into Gateway® destination vectors and thereby high-throughput cloning and expression for any GOI in multiple experimental setups [197]. Large sets of cloned ORFs were established for many model organisms [194], however, such a comprehensive ORFeome collection has yet to be generated for *C. reinhardtii*, as well as a TF collection. A *C. reinhardtii* ORFeome collection would be highly beneficial for the *C. reinhardtii* research community as it enables high-throughput functional experiments and improves the understanding of *C. reinhardtii* biology. It enables the investigation of the dynamic regulatory relationships of TFs and regulators that control *C. reinhardtii*'s metabolism. It will improve our understanding of the regulatory and metabolic processes *C. reinhardtii* undergoes under certain stress conditions. Furthermore, it can pave the way for synthetic biology approaches to alter metabolic regulation such that, for example, lipid accumulation can be highly increased without nutrient deprivation.

1.8 Objective of this thesis

Microalgae represent a promising resource for biofuel production, but the biotechnological use of microalgae is not yet efficient enough. Neutral lipids, e.g., TAGs, produced by microalgae can be easily converted into biodiesel. The FA composition of these lipids directly influences the efficiency of biofuel conversion and its quality. However, most oleaginous algae studied do not have the optimal lipid composition for direct conversion into high-quality biodiesel but have the potential to produce significantly more lipids than conventional oil plants.

Several attempts to increase lipid accumulation, particularly TAG accumulation, were addressed without leading to a final solution for cost-effective and high-efficient production. According to current knowledge, the increased TAG accumulation can only be triggered by external stress factors, e.g., nutrient deficiency. Nutrient starvation, however, limits the biotechnological use of microalgae, as it usually leads to growth arrest and further stress reactions of the alga cell. Other approaches must be found to control and manipulate the metabolism of algae in order to exploit the full biotechnological potential of microalgae.

C. reinhardtii has been intensively studied as a model organism for microalgae [39, 40]. Therefore, there is already extensive knowledge about metabolic processes [41, 42, 43]. However, a thorough system-level understanding of the underlying regulatory mechanisms is missing. Regulation can occur at the DNA level, such as in transcriptional regulation, or at the protein level, such as in protein modifications mediated by protein interactions with other proteins. Most biological processes in a cell are controlled and structured by the synergy of these interactions.

To improve the knowledge about cellular regulation, a systematic approach is needed to describe PDIs and PPIs, which can be mapped in large networks to understand these complex biological relationships. A global ORFeome of *C. reinhardtii* represents a resource that enables these proteomic studies. Accordingly, one goal of the thesis aims to create the first genome-scale ORFeome for *C. reinhardtii*.

The next step is to create a TF collection from the ORFeome that will enable a large-scale PDI screen, mapping transcriptional regulation relationships in *C. reinhardtii*, another goal of this work. However, to perform a PDI screen, a selection of DNA promoter sequences representing the second element of the Y1H screen is required. To gain a global understanding of the transcriptional regulatory processes in *C. reinhardtii*, all promoters of *C. reinhardtii* TFs will be included in the selection of DNA sequences. In addition, more profound insights into the transcriptional regulation of metabolic processes are sought. Accordingly, additional promoters of metabolic genes will be added to the selection of DNA sequences. The promoters of metabolic genes will be selected based

on RNA-seq data and literature search. The RNA-seq data will be filtered for growth conditions with altered nutrients, such as N and S depletion conditions, as these are known to alter overall metabolism, e.g., leading to a reduction in photosynthesis and an increased TAG accumulation in *C. reinhardtii*. The network could provide valuable clues for the regulation of metabolic changes in *C. reinhardtii*.

2 Results

2.1 Generation of a *C. reinhardtii* ORFeome collection

For many model organisms, e.g., *E. coli* [198], *S. cerevisiae* [153], *C. elegans* [199], *A. thaliana* [200], *O. sativa* [201], and *H. sapiens* [202, 160, 151], ORFeome collections are available [196], which enable the expression of proteins in high-throughput functional genomics assays, such as large-scale Y2H screens. The model organism *C. reinhardtii* lacks such a resource, which prevents systematic functional studies for this important model. Towards the goal of my thesis to generate a systematic regulatory network map, it was necessary to generate an ORFeome collection.

2.1.1 Overall cloning strategy

Due to the diversity of coding sequences studied in large-scale experiments, the necessary high-throughput cloning requires a restriction-enzyme independent cloning system. The Gateway cloning system is based on the site-specific recombination by which bacteriophage Lambda integrates an ORF into a plasmid [203]. This system enables the directional cloning of any ORF into a resource vector, a Gateway Entry plasmid, from which it can be moved into expression vectors for any experimental system, using a simple and generic enzymatic reaction [196, 204].

A high-throughput cloning pipeline consisting of a 2-stage polymerase chain reaction (PCR) approach followed by a Gateway® cloning procedure was established (see Figure 2.1). In the 1st PCR, the ORF is amplified from a complex complementary DNA (cDNA) library with gene-specific primers. In the 2nd PCR, Gateway® cloning sites are added to the amplicon using general primers containing Gateway® recombination sites. After a quality control by gel electrophoresis, the ORF is cloned into an Entry® vector using a Gateway® recombination reaction followed by a final quality control using colony PCR and sequence verification.

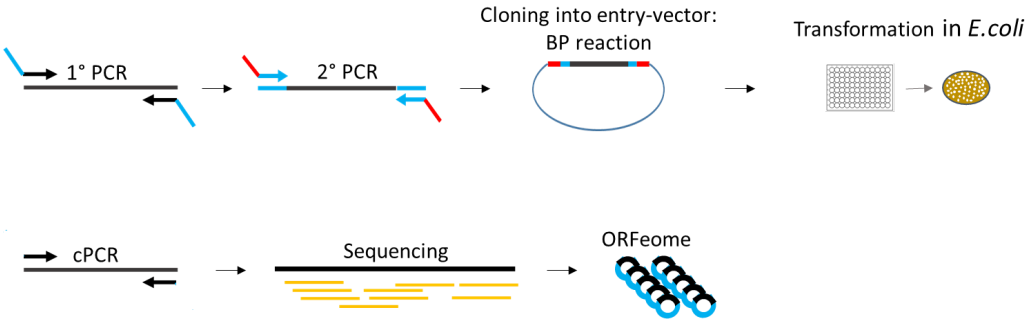


Figure 2.1: Cloning pipeline

Cloning pipeline: ORF capturing in the 1st PCR. Gateway® recombination sites are added in the 2nd PCR. The ORF is cloned using a BP® cloning reaction into a Gateway® donor vector, followed by a bacterial transformation in DH5α *E.coli* cells. A colony PCR confirms the cloning of the ORF into the Entry vector. After DNA isolation, the cloned ORF is sequenced and processed according to further application.

2.1.2 Optimization of high-throughput *C. reinhardtii* ORFeome cloning

Since many ORFs are to be cloned in parallel and the cloning conditions can not be adjusted for each individual ORF, the PCR condition must work reliably for the majority of ORFs. 30 ORFs that span a size range from 700 bp to 3.0 kb were selected for the initial optimization of the PCR conditions. PCR amplification for *C. reinhardtii* ORFs is challenging due to the high GC content of 65 % in the *C. reinhardtii* genome [52]. Therefore, it was first addressed how the choice of DNA polymerase affects the success rate, specificity, and yield of *C. reinhardtii* ORF amplifications.

Six different polymerases were tested, namely PCRBIO Hifi polymerase, TaKaRa Taq DNA polymerase, AmpliTag Gold®(TM) 360 polymerase, Q5®, High-Fidelity DNA polymerase, Phusion® High Fidelity DNA polymerase and KOD Hot Start polymerase, as these polymerases were reported to work with a broad range of GC-rich targets. Selected results of seven different polymerases are displayed in Figure 2.2 a. Although optimized for high GC content, two polymerases, i.e., PCRBIO Hifi polymerase and TaKaRa Taq, could not amplify any of the ORFs tested. AmpliTag Gold®TM 360 polymerase and Q5® High-Fidelity DNA polymerase amplified fragments up to 1.0 kb, which did not correspond to the correct length of the tested ORFs. The Phusion® High Fidelity DNA polymerase could correctly amplify some of the amplicons but did not achieve the performance of the KOD Hot Start Polymerase, which showed the best and cleanest overall PCR success rate. However, even though the KOD Hot Start DNA polymerase performed best, additional parameters had to be modified to improve the PCR yield. Therefore, the addition of PCR additives was tested.

The addition of betaine has been reported to improve the PCR amplification and the yield of GC-rich DNA sequences by reducing the formation of secondary structures caused by GC-rich regions [205]. DMSO has been reported to have the same effect on PCR amplification of GC-rich amplicons. In many protocols, using both reagents is recommended to increase the efficiency of PCR [206, 207, 208, 209]. The concentration of Mg^{2+} also affects the polymerase performance and, thereby the PCR amplification. Mg^{2+} ions stabilize primer-template complexes and promote the catalytic activity of the DNA polymerase. Thermostable DNA polymerases require Mg^{2+} ions as a cofactor during the reaction process, which improves DNA amplification and thus allows better amplification of longer amplicons. However, an excessive amount of Mg^{2+} ions can reduce accuracy and specificity of DNA polymerase and interfere with the complete denaturation of DNA strands during amplification. Furthermore, too high Mg^{2+} levels can affect the primer hybridization, causing the primers to anneal to incorrect sites on the DNA template and consequently resulting in unwanted PCR products [210, 211, 212]. Hence, optimizing the Mg^{2+} concentration is necessary to achieve enhanced and correct amplification of longer amplicons. As shown in Figure 2.2 b, the addition of betaine (1.2 M) and DMSO (5 %) resulted in a significant increase in individual PCR amplicons and yield. However, the overall amplification background also became higher, but this decreased significantly in the second PCR. Therefore, the addition of betaine and DMSO was considered better than the amplification without these additions.

An increased yield was achieved with $MgSO_4$ concentration of 1.25 mM for ORFs smaller than 1.5 kb and 1.5 mM for ORFs larger than 1.5 kb (see Figure 2.2 c). ORFs larger than 4.0 kb were often amplified with artifacts, even with KOD Hot Start Polymerase. Therefore, KOD Xtreme Hot Start DNA Polymerase was tested (see Figure 2.2 d), which is designed for PCR applications of long stranded, challenging and highly GC-rich templates. KOD Xtreme Hot Start polymerase reduced unspecific amplification artifacts in the 1^{st} PCR and improved the amplification of long ORFs and, thus, also the subsequent cloning reaction. 79.2 % of the first test ORFs were successfully cloned. Sequencing confirmed the identity of the desired ORFs. Therefore, a high-throughput approach was started using these optimized PCR conditions.

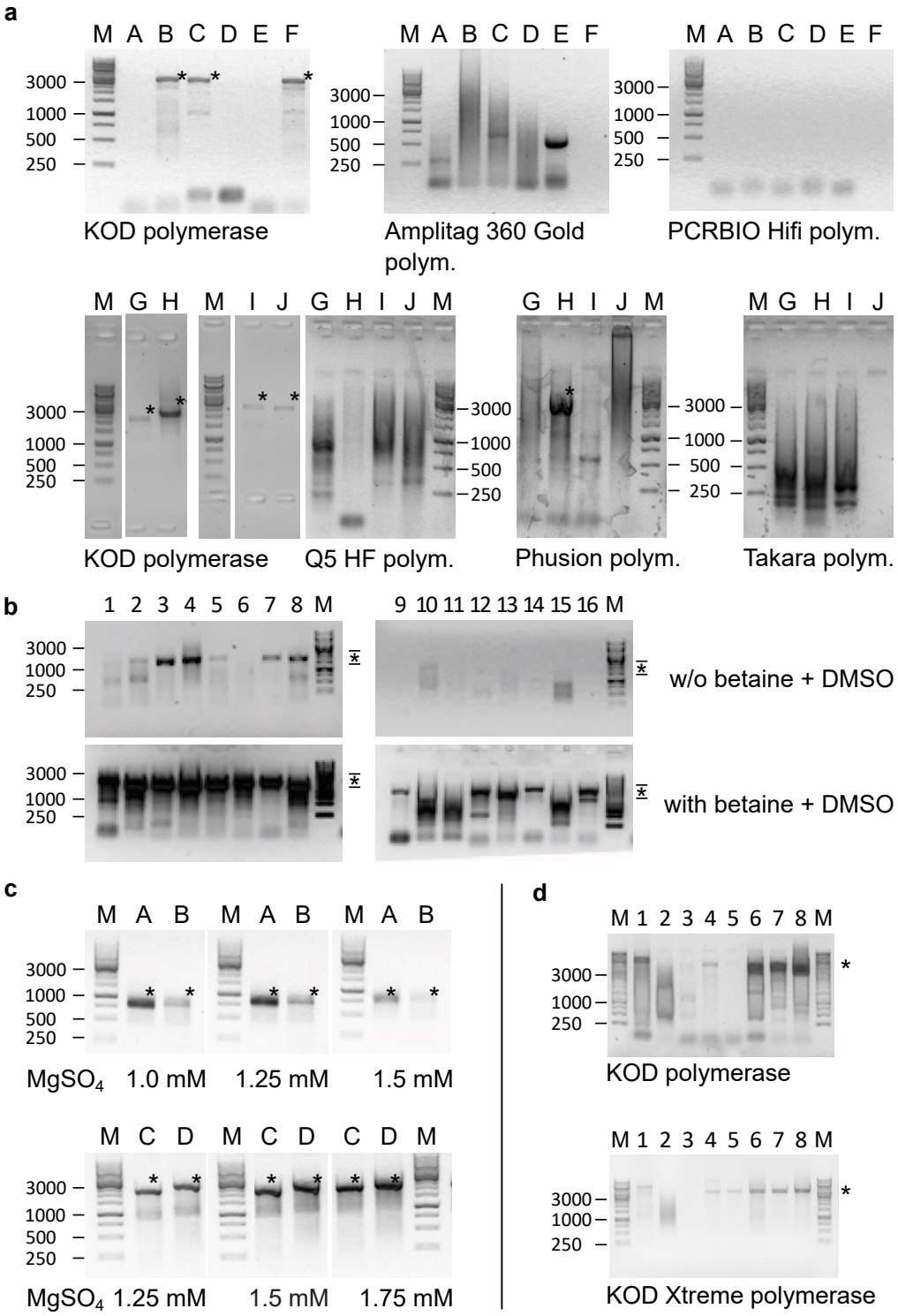


Figure 2.2: PCR optimization

Optimization of PCR additives. Correct amplicon size or size range is marked with an asterisk. Detailed ORFs (1-16, A-H) for a-d and success rates for the different polymerases of a are listed in the appendix in Table A.3 and A.4. **a** Comparison of different polymerases. The KOD Hot Start polymerase (named KOD polymerase) performed best. Selected data shown. M: 1 kb ladder, polym.: polymerase; Q5 HF polym.: Q5 High-Fidelity DNA polymerase. **b** Differences of PCR amplification with and without the addition of 1.2 M betaine and 5 % DMSO with KOD polymerase. ORF size range from 1,600 - 2,300 bp. **c** Testing of different $MgSO_4$ concentrations for ORFs < 1.5 kb and > 1.5 kb. **d** Amplification of ORFs larger than 3.0 kb could be increased using KOD Xtreme polymerase.

2.1.3 Over 50 % cloning success for *C. reinhardtii* ORFs

4,903 *C. reinhardtii* ORFs were selected for the first high-throughput cloning batch. This test set was designed to prove the reliability of the PCR conditions and cloning pipeline. The 4,903 ORFs include genes like enzymes, transporters, metabolic polypeptides, kinases, transcription factors, regulatory proteins, and cytoskeletal and motor proteins. They were selected based on their gene model annotation in version v5.5 and v4.3 of the databases KEGG, COG, Pfam, GO, EC, and PANTHER [52, 54, 213]. The lengths of the 4,903 ORFs differ widely. Therefore, they were divided into four categories according to their size to ensure a parallel procession of multiple samples (see Table 4.4 in chapter 4).

The success rate of the 1st PCR of the 4,903 ORFs was 59.2 % for amplicons with the expected size, in total 2,905 ORFs. An additional 14.1 % of PCR amplicons, in total 693 ORFs, were larger or smaller than expected. Out of 4,903 starting ORFs, 4,158 ORFs were further processed in the cloning pipeline to save resources. The 2nd PCR resulted in 53.6 % of amplicons with the correct size, in total 2,228 ORFs. An additional 4.7 % of amplicons could be amplified with a different size than expected. The PCR success rates varied among the four size categories, with lower success rates for larger ORF sizes (see Figure A.1 in the appendix). This was expected as the Gateway ® cloning success rate decreases with increasing ORF length [202].

The cloned ORFs may carry mutations that arose during PCR amplification. In addition, after recombination, the clones may also contain an empty Gateway® donor vector in which the toxic ccdB gene that normally prevents growth of the empty vector clone is no longer functional due to a mutation or the insertion of a PCR artifact [196, 203, 204]. Therefore, the 4,062 ORFs were sent for NGS to verify the full-length ORF sequence. The full-length sequence was verified for 2,309 of 4,158 ORFs. This corresponds to an overall success rate of 55.5 %, a considerably good result according to the literature. Thus, the high-throughput cloning attempt was successful. Nevertheless, it was observed that the cloning BP reaction in the first cloning set did not always work reliably despite good amplification in the 2nd PCR. Therefore, a PCR purification was tested to determine whether components of the PCR mix affected the cloning reaction. At the same time, sequencing of the miscloned plasmids revealed that short primer dimers were cloned into the plasmid. The PCR purification removed interfering PCR artifacts, such as primer dimers from the PCR mixture, which could improve cloning efficiency. Hence, a purification step after the 2nd PCR was introduced to the cloning pipeline for further experiments.

Based on this first successful high-throughput cloning, the cloning of the entire ORF collection was planned. Primers had to be designed for the remaining ORFs to clone a complete ORFeome. Successful primer design requires an accurate gene model. Therefore, Dr. Yang Jae Kang, a colleague, used an existing dataset of paired-end, stranded

RNA-seq data [138] to validate gene models of all 17,741 protein-coding genes of *C. reinhardtii* using intrinsic features and comparisons to the much better annotated genome of *A. thaliana*. Distinct differences appeared by comparison of the intergenic regions and the coding sequences (CDS) length. The average length of the intergenic region in *C. reinhardtii* is much shorter than in *A. thaliana*. A comparison of CDS length distribution showed that *C. reinhardtii* contains more genes with a long CDS than *A. thaliana*.

Additionally, the algorithm suggested some isoforms or new transcript variants. Therefore, in close collaboration with Dr. Yang Jae Kang, an evaluation scheme for existing gene model annotations was developed based on four criteria extracted from RNA-seq data. These criteria were, first, whether poly-A tail information was available for a transcript. Second, if RNA-seq reads cover the gene model well. Third, if the gene model was conserved among the species, and fourth if the gene transcript could be reconstructed *de novo* from the RNA-seq data (see Figure 2.3 a). Transcripts with a high score in this model showed a higher validation rate by PCR than transcripts with less RNA-seq evidence (see Figure 2.3 b). Furthermore, 54 transcripts and splice isoforms originally differently designated were identified and validated by comparing the sequencing data of the first cloning set with the newly developed algorithm. This result demonstrates that not all genes in *C. reinhardtii* are well annotated and that some isoforms seem to be cloned easier than the annotated gene version.

After reviewing the genome annotation, a primer set for additional 10,631 ORFs was designed. This represents the second cloning set. Primers with defined properties (see section 4.2.4) could not be designed for 2,584 genes. Hence, these genes were excluded from the selection. The extended set of 10,631 ORFs was processed similarly to the first set except for two steps. First, the quality control for the 1st PCR was reduced due to the similar amplification results in the 1st and 2nd PCR of the first cloning set. Not all amplicons of the 1st PCR were evaluated by gel analysis, but only individual strips. The second PCR was analyzed thoroughly and completely by gel analysis. This reduction saved a considerable amount of time during the process. Second, the already mentioned PCR purification was applied for all ORFs in this second set. In total, 49.3 % of 10,667 ORFs (5,256) could be amplified with the expected size in the 2nd PCR. Additionally, 14.3 % of amplicons smaller or larger than expected were amplified (1,525) (see Figure 2.4 a). Similar to the first ORF set, cloning success highly depended on the ORF length.

Due to the high sequencing cost, not all samples could be Sanger sequenced. In addition, the company seqWell, which NGS sequenced the ORFs of the first cloning set, discontinued its sequencing service. Therefore, not all Entry® clones of the second set were full-length sequenced. Nevertheless, a Sanger sequencing quality control was performed for a subset of processed ORFs. The isolated plasmid DNA of eight ORFs of 24 out of 117 randomly selected 96-well plates were Sanger sequenced. For 22 out of 24 investigated microtiter plates, the sequencing results were analogous to the 2nd PCR amplification results and confirmed successful cloned ORFs. This corresponds to the

observation from the first 4,158 ORFs, where the sequencing results also confirmed the 2nd PCR results. In addition, 2,710 ORFs could be verified by full-length NGS sequencing.

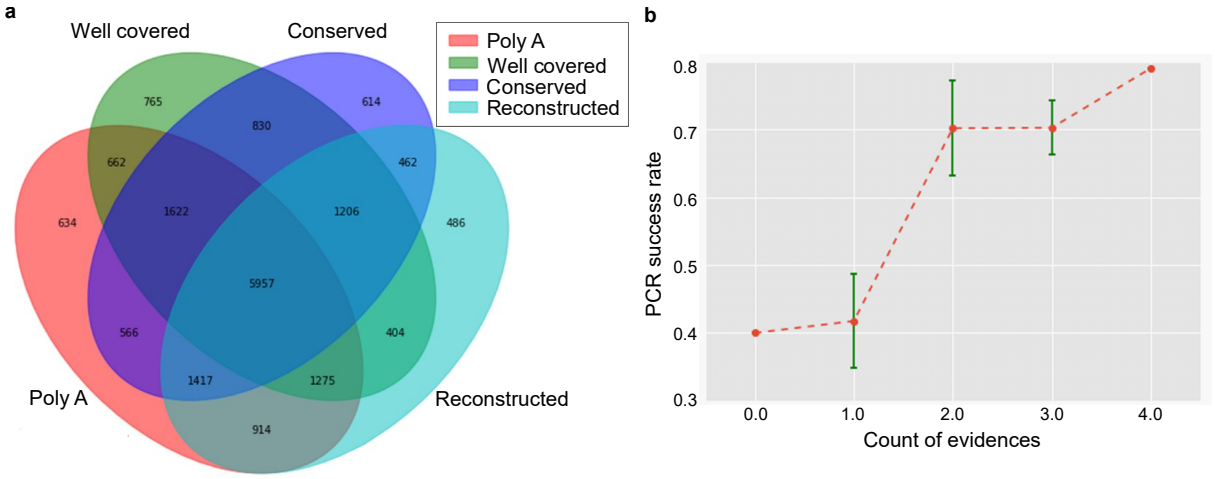


Figure 2.3: Gene model analysis

a Venn diagram of the number of transcripts supported by different evidence extracted from RNA-seq data. **b** PCR success rate of an ORF according to evidence number from RNA-seq data. Graph designed by Dr. Yang Jae Kang.

2.1.4 Higher cloning success rate for more abundant genes

The ORFs were PCR amplified using extracted mRNA converted into cDNA. This leads to the hypothesis that highly expressed genes are cloned with a higher success rate than lower expressed genes because the mRNA level of a transcript is crucial for the cDNA concentration of this transcript in the PCR mix. Hence, the general gene expression of 3,000 randomly picked *C. reinhardtii* genes was compared to 3,000 cloned genes. Figure 2.4 b shows that the cloned genes are significantly higher expressed than the randomly selected genes. That suits our expectations and reflects the previously reported success rates of other ORFeome projects [201].

The gene model analysis revealed a different CDS length distribution concerning larger ORFs in *C. reinhardtii* than in *A. thaliana*. Based on this result, the question arose whether the cloned ORFs correspond to the general CDS length distribution of *C. reinhardtii* genes or whether some gene lengths are overrepresented in the ORFeome. Therefore, all cloned ORFs were compared in length with all annotated *C. reinhardtii* genes. Most of the cloned genes have CDS lengths between 150 to 2.5 kb, but 679 ORFs larger than 4.0 kb were also successfully cloned. Figure 2.4 c shows that the cloned ORFs correspond to the overall CDS length distribution of *C. reinhardtii* genes and fits the assumption that the larger the ORF, the lower the cloning efficiency.

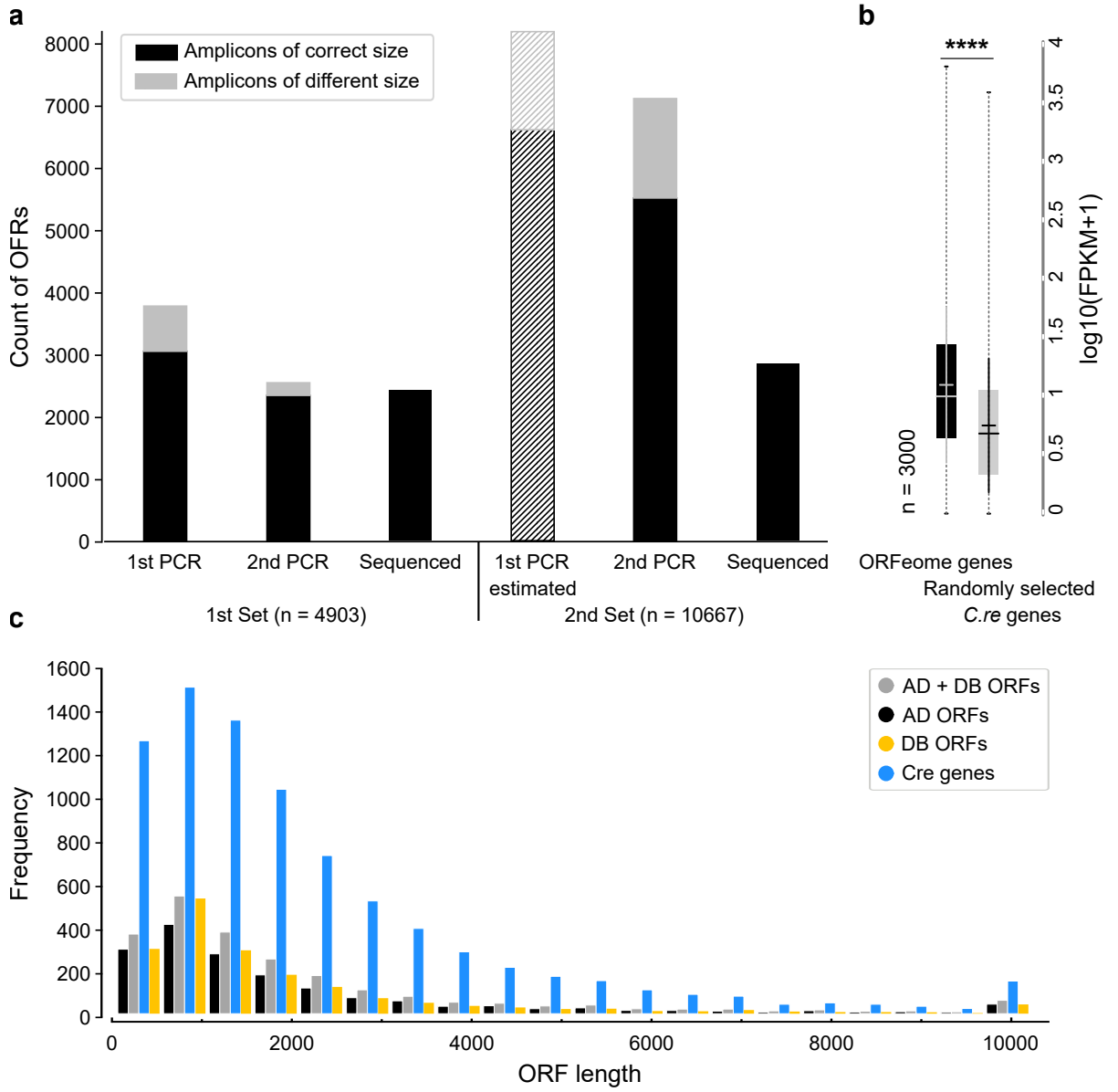


Figure 2.4: ORFeome cloning

a The amount of cloned ORFs of the 1st and 2nd ORF sets are shown. They are divided into 1st and 2nd PCR and the sequenced validated number of ORFs. Black shows all amplicons of expected size, grey marks amplicons with a different size. The bar graph with dashed lines indicates an estimation of 1st PCR amplicons. **c** FPKM values of 3,000 cloned genes are compared with the gene expression of 3,000 randomly picked *C. reinhardtii* genes. The gene expression of cloned ORFs is significantly higher than the FPKM values of randomly picked genes. **c** The CDS length distribution of all *C. reinhardtii* genes (blue) together with the cloned ORFs in total (grey) plus the ORFs cloned into the AD (black) and DB (yellow) destination vectors are displayed.

2.1.5 Sequenced ORF library in BFG-Y2H vectors

The Gateway Entry® clones of the first and second cloning sets were consolidated to facilitate their processing in further experiments. One such application is a binary Y2H assay. The BFG-Y2H method uses deep sequencing to improve the identification of PPIs in pools [156]. It enables screening of a whole matrix of protein pairs in a single multiplexed yeast pool, thereby increasing the screen's efficiency. To perform the Y2H assay, the examined ORFs must be present in at least one of two fusion reporter vectors (AD or DB destination vector). Therefore, 8,627 ORFs that scored positively in the 2nd PCR were selected to be cloned into BFG-Y2H destination vectors. DNA extracted from all Entry® clones was divided into eight DNA pools, each containing the DNA of approximately 1.100 ORFs. Each DNA pool was used for one *en masse* cloning reaction, resulting in eight *en masse* reactions with AD (activation domain) destination plasmids and eight *en masse* cloning reactions with DB (DNA binding domain) destination plasmids.

After the bacterial transformation of the cloning reaction, a twofold oversampling was performed for AD and DB destination vectors to capture, in the best case, each ORF with two different DNA barcode sets. Single colonies were picked and arrayed into 96-well microtiter plates. A colony-PCR was performed with barcoded labeled primers to identify the ORFs. Next-generation sequencing (NGS) identified 4,735 out of 8,627 ORFs (54.9 %) in either AD or DB destination vectors. 3,590 ORFs were identified as AD-fusion destination vectors and 3,648 ORFs as DB-fusion destination vectors. The number of identified ORFs cloned into the BFG-Y2H destination vectors after *en masse* LR was far lower than expected. Here, a twofold oversampling was performed and the *en masse* LR was not repeated. Yachie et al. demonstrated that a fourfold oversampling with multiple rounds of cherry-picking leads to a higher cloning success rate [156]. Thus, more ORFs could be cloned by repeating the colony picking and the *en masse* LR for missing ORFs. Nevertheless, these vectors are ready to be used in an upcoming Y2H screen that was outside the scope of this work.

The *C. reinhardtii* ORFeome cloning resulted in a consolidated Gateway® Entry clone collection of 8,627 ORFs plus 3,590 ORFs cloned into BFG-Y2H AD-fusion destination vectors and 3,648 ORFs cloned into the BFG-Y2H DB-fusion destination vectors. The generated ORFeome collection is the first broad collection of *C. reinhardtii* ORFs and represents a highly beneficial resource for the *C. reinhardtii* community.

2.2 Gene regulatory network

This comprehensive ORFeome collection of *C. reinhardtii* builds the foundation for my further work. A large-scale PDI network reflects the global structure of transcriptional connectivity, the regulatory and hierarchical composition of TFs and how they control biological processes in a cell [164, 165, 166, 167]. For such a comprehensive high-

throughput PDI approach, a TF and promoter bait collection are required to screen a high number of TFs against a large set of promoters in parallel. The TF collection could be assembled as most of the *C. reinhardtii* TFs were part of the ORFeome cloning attempt. Together with a selection of DNA promoters the first experimental-based gene regulatory network of *C. reinhardtii* was generated.

2.2.1 Consolidation of *C. reinhardtii* transcription factors and regulators

To assemble a *C. reinhardtii* TF collection, all available TF and TR information from the two plant TF databases 'PlantTFDB' and 'PlnTFDB' [214, 215, 216, 217, 123, 218] were extracted and combined. In addition, *C. reinhardtii* orthologues of *A. thaliana* TFs (2,296) were identified, extracted, and included in the selection. In total, 475 *C. reinhardtii* TFs and TRs were defined. They are grouped into 61 families, of which 19 are singletons. An overview of all defined TFs and the assigned names are listed in Table A.5 in the appendix. Since many TFs are not yet fully annotated, I assigned names to the TFs present in the PDI screen that correspond to their respective DNA-binding domain or TF family. Figure 2.5 displays all TFs according to their assigned families. Their CDS length varies between 200 bp to 20 kb.

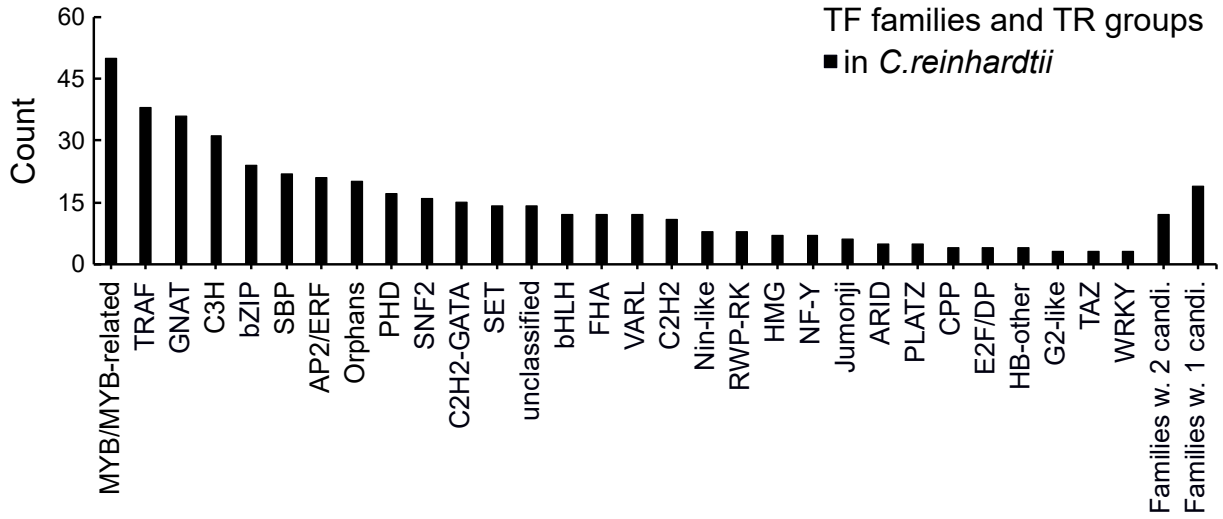


Figure 2.5: Summary of *C. reinhardtii* TFs and TRs

All defined *C. reinhardtii* TFs and TRs split by their TF family or TR group allocation (61 in total). 12 TF families with two independent candidates and 19 TF families with only one candidate are shown in abbreviated form due to space limitations. An extended TF list can be found in the appendix.

TF cloning was challenging, with 195 of 475 TFs having a CDS length larger than 3.0 kb and many GC-rich repeats. 161 *C. reinhardtii* TFs were cloned and could be extracted from the ORFeome collection. To expand the scope of the TF collection, the missing

genes were commercially synthesized. Only TFs with a CDS length of up to 3.2 kb could be selected for gene synthesis, as this was the maximum possible synthesis length that Twist Bioscience could synthesize at that time. Longer TFs are missing in the final collection of synthesized TFs. GC-rich repeats and repetitive sequences in the TFs sequences hampered gene synthesis of the remaining TFs. However, after codon optimization for yeast, 96 additional TFs were synthesized. A total of 241 unique TFs, representing 50.8 % of all putative *C. reinhardtii* TFs and TRs, were transformed into yeast. Since TFs can also bind non-specifically to vector DNA, an auto-activation screen was performed to exclude these TFs before the actual screening (see section 4.2.19). 220 TFs remained for the Y1H screening after 21 auto-activating TFs were excluded. They represent 46.3 % of all characterized and putative *C. reinhardtii* TFs.

2.2.2 Promoter selection and cloning

The DNA bait collection was chosen to address two biological questions:
First, which TFs regulate other TFs, and how is this regulatory TF network structured?
Second, can this Y1H screen identify general regulators of lipid metabolism, particularly of stress-adaptive metabolism in nutrient deficiency in *C. reinhardtii*?

To answer the first question, all promoters of the consolidated TFs and TRs, in total 475, were included in the promoter selection. In addition, around 400 candidate genes were preselected to address whether general regulators of lipid and primary metabolism can be found. Two criteria defined these candidate genes. First, the genes were associated in the literature with lipid metabolism, lipid remodeling, and metabolism remodeling during nitrogen and sulfur deprivation. Second, the gene expression profiles of these genes were analyzed using the RNA-seq data set of Ngan et al. [138]. Ngan et al. performed a deep RNA-seq analysis throughout the course of N- and S depletion up to 48 h after starvation. Under these conditions, lipid accumulation is induced and pronounced. They performed a chromatin signature and transcriptomic analysis to decipher the regulation of lipid biosynthesis in *C. reinhardtii* [138]. All preselected candidate genes that showed a log2 fold change alteration in the time series of sixteen time points during N- or S-deprivation in the study of Ngan et al. were included in the promoter selection. In total, 139 candidate genes were chosen for cloning.

The 475 promoter regions of the TFs and TRs and the 139 promoter regions for candidate genes of the stress-involved deprivation metabolism, add up to a final number of 614 promoters. These promoters represent the Y1H baits and were used for bait cloning (see Figure 2.7 a). Since previous Y1H studies reported that different promoter lengths function in the Y1H screen (see Table 3.1), the question of the suitable promoter length for *C. reinhardtii* arose. Yu et al. showed that the positional distribution of transcription factor binding sites (TFBS) of *A. thaliana* lies in the regions from -1.000 bp to +200 bp to the transcription start side (TSS) [219]. Furthermore, Scranton et al. designed synthetic promoters capable of driving robust gene expression in *C.reinhardtii* with a promoter

region of -1.000 bp and +50 bp [220]. Accordingly, in this study, promoter regions were defined as DNA sequences that were +1.000 bp upstream and +100 bp downstream of the annotated TSS of the selected genes.

472 promoters (76 %) could be successfully amplified in the first and second PCR. Promoter PCR amplicons were Gateway® cloned into the pDONR P4-P1R vector (Entry vector) and afterwards upstream of the *HIS3* reporter gene into the pMW#2 vector, representing the destination and reporter vector for Y1H screenings. The correct sequence was confirmed by Sanger sequencing. The cloning efficiency of this reaction was much lower than the previously described cloning into the pDONR223 vector. Most likely due to the different attP sites present in the pDONR P4-P1R vector. Up to four individual colonies were picked for each promoter construct to increase the low cloning efficiency. Cloning efficiency could be further improved by picking more colonies for each promoter construct.

340 promoter constructs were successfully cloned into the pMW#2 vector and verified by sequencing. The vectors were linearized by restriction enzymes, followed by a randomized integration into the yeast genome to generate chromatinized DNA-bait strains. Since the promoter fragments are 1.100 bp long and randomly integrated into the yeast genome, they may also contain binding elements to which yeast TFs could bind and trigger transcription of the reporter. Therefore, all DNA baits were tested for auto-activation. 37 DNA promoters showed auto-activation during the self-activation test. These promoter bait strains were excluded from the upcoming experiments. 303 of 340 DNA promoter bait strains remained for the primary screening. 220 DNA bait strains contained TF promoters, and 83 DNA bait strains contained promoters of structural genes.

2.2.3 Primary Y1H pool screening identified 503 interactions

The high number of 303 promoters and the TF collection of 220 TFs resulted in an extensive screening matrix. State of the art Y1H studies uses the enhanced Y1H (eY1H) method for their Y1H studies. eY1H works with a high-density format, where a robotic platform arrays 1.536 yeast spots on one plate [193]. This eY1H approach was not possible for this study. Hence, a different high-throughput Y1H approach was developed.

A major difference between the developed approach and the published eY1H method is the cultivation of the yeast strains in liquid media during the procedure. This enables the usage of a Tecan Fluent liquid handling robot and an easy and reliable transfer of TF prey and DNA promoter bait strains. However, this resulted in the necessity to adapt the experimental Y1H pipeline. As indicated in Figure 2.7 b, prey and bait yeast strains were mated in yeast extract peptone dextrose (YEPD) and afterward double selected for diploid yeast cells in SC - His - Trp (SC-H-W) minimal media. This double selection is different from eY1H but was necessary due to the liquid media cultivation setup. After

one selection round, some haploid yeast cells still survived and could falsify the result. After a second selection round, all haploid cells were eliminated, and only diploid yeast cells were present in the medium.

To efficiently screen all promoters against all TFs and to save resources, the Y1H screen was divided into two parts: the primary screening and the verification screening. The primary screening is a pre-screening followed by a 4-fold one-by-one verification screening.

It has been previously reported that TFs from the same TF family that are structurally related often bind to similar DNA-binding motifs [221]. Therefore, all 220 TFs were subdivided into 16 TF mini pools according to their TF family affiliation. 13-20 TF yeast strains of a particular TF family were mixed in one TF mini pool. Then, all 16 TF mini pools were tested against all 303 DNA promoter bait strains. This distribution in different mini pools should minimize competitive effects between TFs from different TF families. Furthermore, it increases the possibility of finding multiple individual interactions of different TFs of one family after only one primary interaction of the TF mini pool.

For practical reasons, the 220 TF DNA promoter strains were screened first, followed by the 83 structural gene promoter strains. Figure 2.7 c shows one example of a yeast readout plate of the primary screening. Eight DNA promoter baits (A-H) are mated against TF mini pools 1-8 and an AD-GFP control yeast strain (column 9 and 10). The AD-GFP control yeast strain expresses an AD-GFP fusion protein instead of an AD-TF fusion protein. This negative control defines the mating background for each DNA promoter bait strain on the readout plate. A primary interaction was counted when yeast colony growth was detected for one TF mini pool with a DNA promoter bait and the associated negative controls showed no growth.

682 protein-DNA interactions could be detected in the primary screening of 16 TF pools against 220 TF promoter bait strains. The interactions, shown in Table 2.1, are allocated to the 16 TF pools. Three TF pools (pool numbers 1, 10, and 14) showed very strong yeast colony growth with a majority of DNA promoter bait strains. For TF pool 1, 150; for TF mini pool 10, 184; for TF mini pool 14, 145 positive primary interactions were detected.

Table 2.1: Interactions found in the primary screening of 220 TF promoter baits

TF mini pool	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
PDIs	150 (6)	5	8	15	27	97	2	33
TF mini pool	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
PDIs	2	184 (16)	3	6	1	145 (15)	3	1

Table 2.2: Interactions found in the primary screening of 83 structural gene promoter baits

TF mini pool	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
INT	8	6	35	11	23	25	7	35
TF mini pool	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
INT	12	11	12	54	7	6	17	31

The increased number of primary interactions compared to the other 13 TF pools suggests that these three TF mini pools contain one or two auto-activating TFs. It suggests that these TFs bind unspecifically to the bait vector backbone and were not detected in the previously performed auto-activation screen. Therefore, the auto-activator screen was repeated and confirmed the suspicion. 7 TFs were identified as auto-activators and eliminated from the screen. New TF mini pools were prepared and used for further primary screens. 240 primary interactions remained for the TF promoters after removing the auto-activators and screening the baits again against the three new TF pools 1, 10, and 14. 300 primary pool interactions could be detected in the primary screening of 83 structural gene promoter bait strains, shown in Table 2.2. In total, 540 primary interactions were identified in the primary screening.

2.2.4 Verification screen confirmed 1.451 interactions

Each primary interaction was tested in a 4-fold one-by-one verification screen in which all individual TFs included in one TF mini pool were screened one-by-one against the specific DNA promoter bait in four replicates. Figure 2.7 d shows one yeast verification screen readout plate. Strict criteria for the readout were used to minimize the number of false positives. An interaction was only counted if it occurred in at least three repeats, and the mating control plate (see Figure A.2 in the appendix) indicated a positive mating, plus, the negative and positive controls met the expectations (see section 4.2.20.1).

In total, 1.451 interactions between 142 TFs and 200 DNA promoter baits were identified in the one-by-one verification Y1H screen. Interactions were revealed for 64.5 % of the studied TFs with 66.0 % of the investigated promoters. The identification of PDIs for 66.0 % of promoters is below expectations from reported studies. One reason for this could be the immense size of promoter samples, which made a systematic retesting of individual DNA promoter strains in the primary screening impossible. However, finding interactions for 64.5 % of TFs is higher than expected.

Y1H can only detect interactions of monomeric or homodimeric TFs [122]. Furthermore, several TFs need post-translational modifications to bind DNA, and thus limiting the Y1H assay [168, 192]. To address the question of whether a particular TF family does not interact at all in the Y1H screen, a comparison of TFs transformed in yeast and

TFs interacted in the screen was performed. Figure 2.6 shows the comparison of TFs transformed in yeast and TFs that interacted in the Y1H screen. It is obvious that not all TFs interacted, but that members of most TF families were found as interactors in the screen and that the dropout of TFs from different families is homogeneous, indicating that the Y1H system does not have an inherent preference for or against certain types of TFs. Finding PDIs for 142 TFs represent DNA-binding evidence for approximately 30 % of all annotated *C. reinhardtii* TFs, an excellent coverage [168, 222].

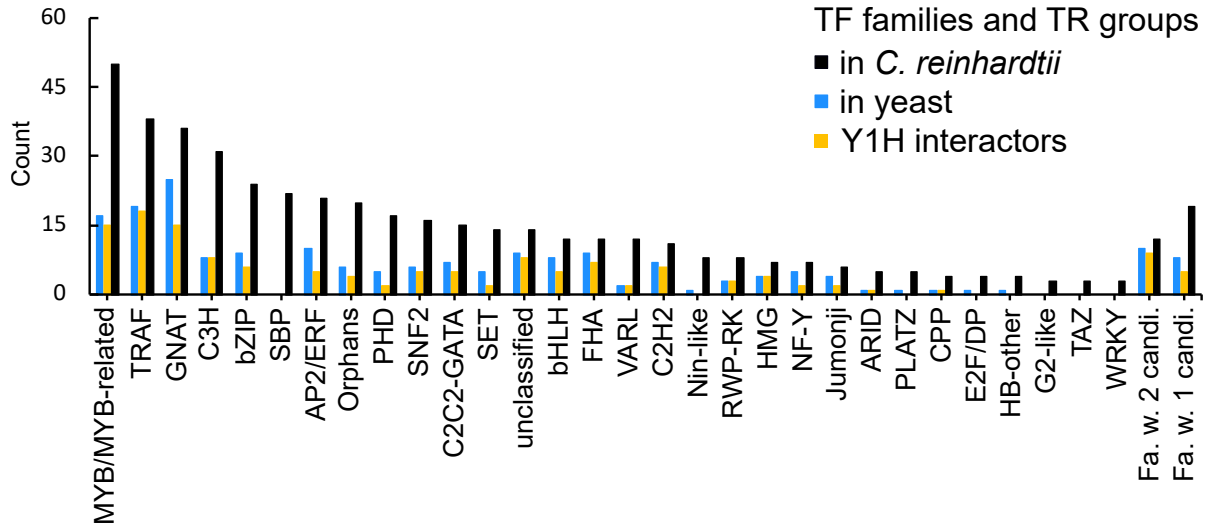


Figure 2.6: *C. reinhardtii* TFs and TRs transformed into yeast and Y1H interactors
Black bars are the TF and TRs defined in *C. reinhardtii*. Blue bars represent the TFs and TRs cloned and screened in the Y1H screen. Yellow bars indicate the number of TFs and TRs that interacted with promoters in the PDI screen. Surprisingly, almost all TF families and TR groups interacted in the Y1H screen.

A saturation assay was performed to assess the sampling sensitivity of the primary screening, since no single assay can detect every interaction [223]. It provides information on how many interactions the assay can detect. Therefore, a subset of 48 baits was randomly chosen and tested against all 16 TF mini pools. This experiment was repeated in four independent repeats. Afterward, the found interactions were compared and a saturation model was calculated (see Figure 2.7 e). The model revealed that about 50 % of all interactions could be found after four repeats. The primary screening performed with all DNA baits identified 34 % of interactions. Many possible interactions could be uncovered by repeating the primary screening. The Y1H PDIs were mapped into a network, the first experimental-based large-scale PDI network for *C. reinhardtii*.

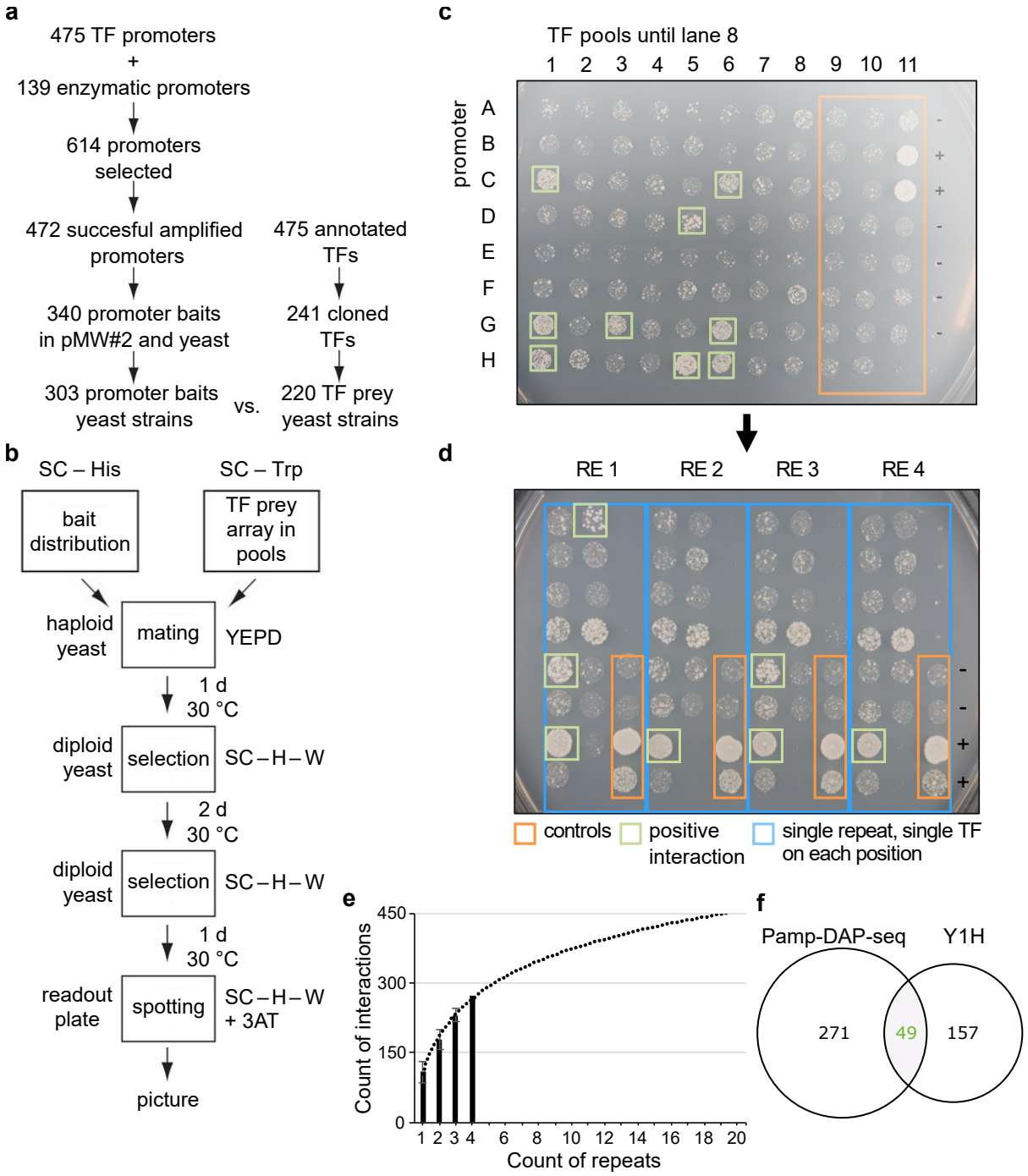


Figure 2.7: Y1H assay setup

a Description of Y1H prey and baits. **b** Experimental Y1H pipeline based on liquid cultures. The yeast cultures are grown in 96-well microtiter plates. The prey and bait strains are mated in YEPD and selected twice in SC-His-Trp media. If a PDI occurs the diploid yeast can survive on SC-His-Trp +3AT readout plates. Adapted from [224]. **c** A cut-out of one yeast readout plate of the primary screening. The plate contains 8 DNA promoter baits mated against the TF mini pools 1-8. Positive interactions are marked with a green square. The controls are framed in an orange rectangle. Positive controls in column 11 are supposed to grow, negative controls not. Columns 9 and 10 contain the AD-GFP control yeast strain that defines the background of the mating. **d** The picture shows a cut-out of a verification yeast readout plate of the TF mini pool 5 with DNA promoter of Cre14.g621050. An interaction was counted when it was present in three repeats. The blue rectangle highlights the four different repeats. **e** Sample sensitivity model: black bars represent the verified interactions of the four saturation assay repeats. The black dashed line shows the screen saturation model based on four repeats. **f** Venn diagram shows the PDI overlap of all found interactions of six *C. reinhardtii* TFs in Y1H and PampDAP-seq.

2.2.5 PampDAP-seq validated PDIs for six *C. reinhardtii* TFs

A second orthogonal test is required to determine the quality of the network. Therefore, a modified ampDAP-seq experiment was performed as biophysical validation to evaluate the quality of the detected Y1H interactions. DAP-seq is an *in vitro* TF-DNA binding assay that detects TF binding sites in genomic DNA with *in vivo* expressed TFs [190]. The modified DAP-seq version ampDAP-seq uses a PCR-amplified genomic DNA library as a template for the TF-binding. Here, ampDAP-seq was further modified. The DNA promoter sequences cloned for the Y1H screen were PCR amplified and used for the DNA library preparation. A detailed description of the DNA library preparation is displayed in section 4.2.24.1. Hereafter, the method is referred to as PampDAP-seq, promoter PCR amplified DAP-seq. It is characterized by the fact that PCR amplified promoter sequences were sonicated and subsequently tagmented to minimize the size of the fragments. Afterward, all fragments bound by a given affinity-purified TF were NGS-sequenced and analyzed. Enriched DNA fragments can be identified by mapping the NGS reads to the reference sequences of the used promoters. The pile-up of DNA fragments enriched on a specific DNA position is called a peak. Identified peaks represent a PDI of a particular TF with a promoter sequence. The computational pipeline for peak detection and analysis was developed by my colleague Dr. Chung-Wen Lin, and is described in section 4.2.24.3.

Initially, 25 TFs were selected, depending on their outdegree within the regulatory network, to be tested by PampDAP-seq. 19 of the 25 TFs have been successfully cloned into the expression vector pGEX6 (further described in section 4.1.3). Ten TFs could be expressed, purified, and processed in the PampDAP-seq experiment. Peaks in two biological replicates were identified for six out of ten TFs tested. Two additional TF proteins showed peaks only in one replicate and were therefore excluded from further analysis. In summary, 320 unique PDIs were detected for six TFs in PampDAP-seq.

For yeast-driven experiments, the detection profiles of two assays are different in most cases [223]. Consequently, the profile of interactions detected by two different assays will only partially overlap. By comparing the PDIs of Y1H and PampDAP-seq, an overlap of 23.8 % was identified, supporting the Y1H-binding results (see Figure 2.7 f). The PDI overlap and the PDI interactions are listed in Table A.9 in the appendix.

Given the lack of previously reported Y1H studies compared with DAP-seq, a comparison of previously reported ChIP-seq and eY1H data was performed to relate the results to other regulatory networks. The PDI overlap of this study is of the same magnitude as other previous studies and even a bit higher. A fisher test was performed to show the significance of the overlap between the Y1H and PampDAP-seq data. The fisher test resulted in a non-significant p-value of 0.11 for the overlap, which could be explained by a subsequent power analysis showing that the experiment was underpowered due to the insufficient sample size in PampDAP-seq. Even if the overlap is not significant, the

validation rate is similar to previously PDIs reported in the literature and underscores the high quality of the detected PDIs.

2.2.6 Analysis of the network structure

This *C. reinhardtii* GRN displays interactions between 303 interaction partners (see Figure 2.8) and dramatically expands the number of known DNA binding evidence for *C. reinhardtii* TFs. The network is directed, i.e., connections between nodes are directional. In contrast, Y2H networks are undirected networks. Connected nodes in a Y2H network are proteins interacting with each other. In contrast, connected nodes in a Y1H network represent TFs, the active part of the network that binds to other nodes, which represent DNA sequences, the targeted element in the network. This results in a different topology. The outgoing connections of a TF to DNA promoters are described as the outdegree of this TF. Indegree is the opposite, the number of interactions directed to a node. The edge count represents the sum of the out- and indegree of every node. Here, the GRN comprises a special characteristic; it contains promoters of TFs, which are also present as active proteins in the screen. This experimental feature will help to reveal the transcriptional structure of these TFs and how they regulate other TFs. 89 TFs are only present as promoters and not as active TFs. Hence, they have an outdegree of zero because they can not target other promoters. The network in Figure 2.8 is structured according to the TF outdegree. Figure A.5 in the appendix shows the network with all annotated and assigned names.

The lowest layer of the network is composed of the promoters since their outdegree is zero. TFs that interact with one to nine promoters are shown in the layer above the promoters with an outdegree of zero and represent the third layer in the network. 69 TFs interact with one or two promoters (see Figure 2.9). Another 45 TFs interact with three to nine promoters. These 114 TFs account for 19.9 % of interactions and could represent specialized TFs that regulate the gene expression of only a few target genes or performed poorly in the Y1H screen. The second layer represents TFs with an outdegree between 10 - 24. They account for 13.4 % of interactions. These TFs are already well-connected and seem to represent a middle layer in the network. 15 TF comprise the highest level of the network. They form 66.7 % of all interactions in the network. TFs with the highest outdegree are the following:

Six TFs belong to the zinc finger family, but subdivide into four C3H zinc finger TFs (Cre06.g250950, Cre10.g460400, Cre04.g231124, Cre06.g278136) and two C2H2 zinc finger TFs (NFXL1 (Cre17.g696300), Cre12.g538801). Further, three basic leucine zipper (bZIP) domain TFs, BLZ1 (Cre13.g590350), BLZ5 (Cre07.g321550), and Cre07.g344668; two MYB/MYB-like TFs, called MYB_9, (Cre03.g144907) and LRL1 (lipid remodeling regulator 1, Cre03.g197100); one unannotated TF, called MS_1, (Cre10.g450500); one forkhead-associated (FHA) domain TF CGL86 (Cre12Fang.g534450)); one basic helix-loop-helix (bHLH) domain TF, called bHLH_6 (Cre16.g667150), and one tumor necrosis

factor receptor-associated factor (TRAF) TF, called TRAF_19 (Cre04.g219600) are part of this hierarchical layer.

Based on previous studies, the top outdegree scorers, known as hubs, can be considered as important 'regulators', whereas nodes with a high indegree represent promising 'modulator' candidates [225, 226]. Hubs usually play essential roles in the networks and fulfill critical functions in the biological context of the examined organism [160]. The studied promoters showed interactions with 1-40 TFs, of which 84.5 % are targeted by more than two TFs.

Interestingly, 31 % of promoters are targeted by six or seven different TFs. 82.3 % of these promoters belong to TFs (51 of 62), and only 11 promoters are from structural genes. This strengthens the assumption that *C. reinhardtii* has a middle layer of TFs that are very well connected with other TFs but do not represent large hubs. It could reflect the convergence of several regulation strands merged in the targeted promoter. It could be a fine-tuning effect, concentrating the incoming signal from different TFs to respond precisely to a specific event.

The known stress-associated TFs Lesion Simulation Disease Resistance 1 (LSD1) [227] and Radical-induced cell death 1 (RCD1) [228] are part of these TF promoters, which are targeted by more than six other TFs. Two other stress-associated TFs are the heat-shock TF HSF2 and a cold-shock domain containing protein named CSD_2. Seven of nine promoters of Gcn5-related N-acetyltransferases (GNAT), which represent histone acetyltransferases, are also part of this middle-layer. These examples support the hypothesis that different signal strands converge in one TF.

For previous networks, it has been shown that most biological networks are scale-free, not randomly distributed, and follow a power-law distribution [229]. Therefore, the edge count and the out-and-indegree distribution were tested for their distribution behavior. All three are not normally distributed (tested with the Kolmogorov-Smirnov test). They follow a power law distribution, but the network does not have the power to show significance due to the relatively low node number.

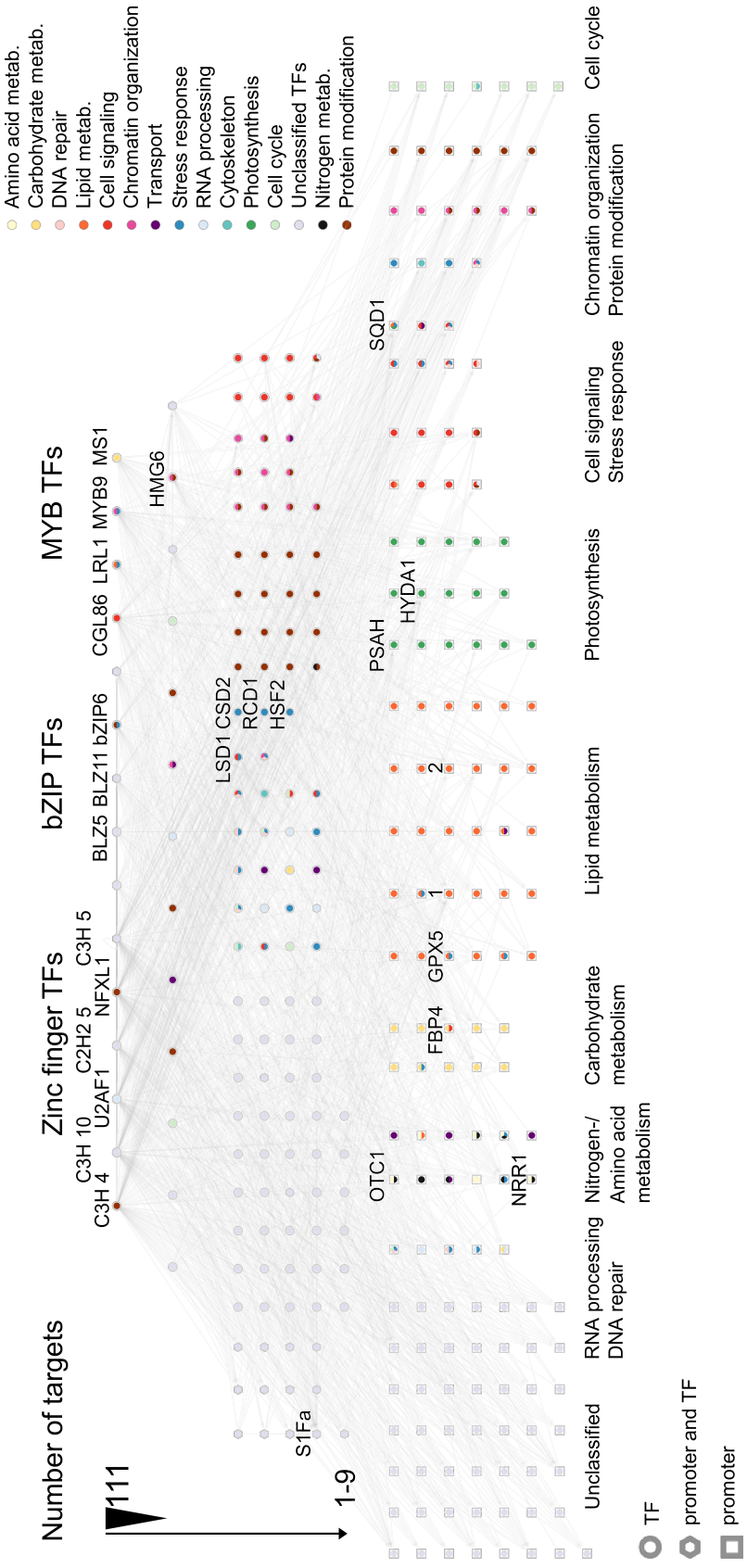


Figure 2.8: *C. reinhardtii*'s gene regulatory network

C. reinhardtii GRN represents all found Y1H PDIs. The network is structured according to the TF outdegree. TFs are shown as circles, promoters as squares, and interaction partners that function as promoter and TF are shown as hexagons. The assigned colors are designated to the different GO groups. Network was analyzed and displayed by using Cytoscape [230]. 1: Cre07.g334150, 2: Cre10.g425100.

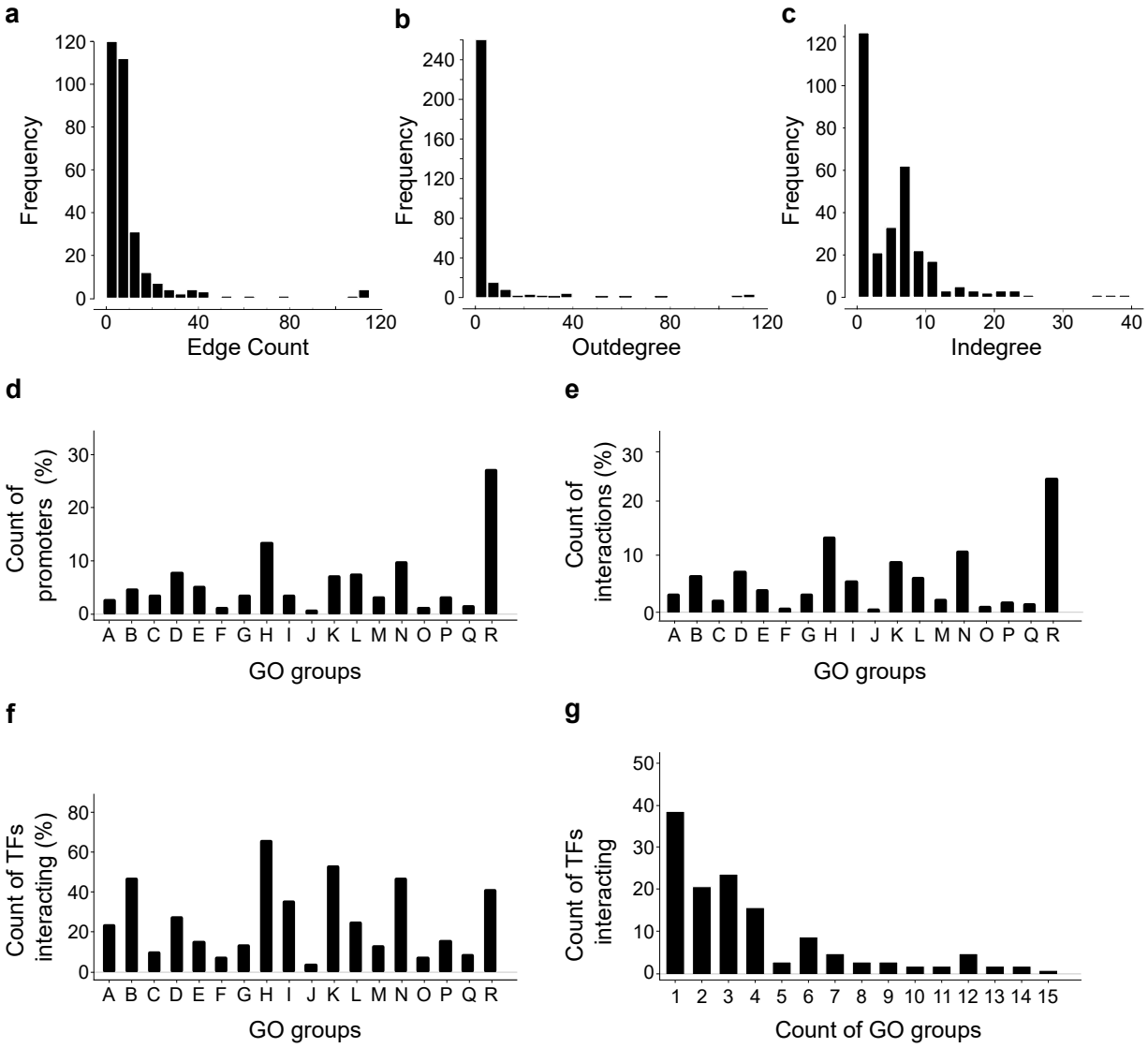


Figure 2.9: Network attributes

a Edge count distribution **b** Outdegree distribution **c** Indegree distribution **d** Count of promoters for each GO group **e** Count of PDIs for each GO group **f** Count of TFs targeting promoters of a certain GO group **g** Count of TFs corresponding to their interaction profile with multiple GO groups.
Classified GO group terms: A: Amino acid metabolism; B: Carbohydrate metabolism; C: Cell cycle; D: Cell signaling; E: Chromatin organization; F: Cytoskeleton; G: DNA repair; H: Lipid metabolism; I: Nitrogen metabolism; J: Oxidative phosphorylation; K: Photosynthesis; L: Protein modification; M: RNA processing; N:Stress response; O: Translation; P: Transport; Q: Development; R: Unclassified

2.2.7 GO group analysis of network partners

To analyze the functional relationships within the regulatory network, it is crucial to know the biological function or the process in which the interaction partner is involved. Therefore, all 303 interaction partners were manually assigned to Gene Ontology (GO)

Table 2.3: GO terms

• amino acid metabolism	• DNA repair	• protein modification
• carbohydrate metabolism	• development	• RNA processing
• cell cycle	• lipid metabolism	• stress response
• cell signaling	• nitrogen metabolism	• translation
• chromatin organization	• oxidative phosphorylation	• transport
• cytoskeleton	• photosynthesis	• unclassified

Biological Process (BP) terms [231, 232]. Further description of the GO dedication is stated in section 4.2.25. The GO terms of the interaction partners were grouped into 18 functional GO groups as multiple GO terms describe similar biological functions or categories (see Table 2.3).

Some genes in the screen are assigned to two or three GO terms, e.g., GPX5, which is assigned to 'lipid metabolism' and 'stress response'. GPX5 is essential in oxidative stress homeostasis during lipid biosynthesis and belongs to both GO terms. However, 113 TFs in the screen could not be grouped in any GO group because they have no known functional or metabolism-associated annotations. They were labeled as 'unclassified'. The network attributes corresponding to the different GO groups were analyzed to determine whether some TFs bind only promoters of single biological functions or multiple metabolic areas (see Figure 2.9). Therefore, all TF-promoter interactions were also assigned to their respective GO term. Plot d of Figure 2.9 shows the number of promoters belonging to the different GO terms. The number of gene promoters classified as 'lipid metabolism' and 'unclassified' is higher than the number of promoters that belong to other GO groups. One explanation is that the promoter selection primarily included promoters of all *C. reinhardtii* TFs, also carrying uncharacterized TFs and promoters of genes involved in overall lipid metabolism. Plot e in Figure 2.9 visualizes the total number of interactions found in the Y1H screen. A general positive correlation exists between the number of interactions and the number of promoters screened for each functional group except for promoters of genes assigned to the GO term protein modification. Figure 2.9 f shows the number of TFs that target promoters of a specific GO term. There is a trend that many TFs are targeting promoters of the lipid, carbohydrate, and nitrogen metabolism, photosynthesis, and stress-response related genes in this screen (see Figure 2.9 f).

72.5 % of TFs in the network bind to promoters of two or more GO groups, meaning that only 27.5 % of TFs bind to promoters of genes within one GO group. In addition, it was observed that *C. reinhardtii* TFs are strongly interconnected, and multiple TFs can bind to the same promoter. This indicates that *C. reinhardtii*'s primary metabolism is

mediated by interconnected transcriptional regulation. It suggests that TFs regulate the gene expression of multiple genes in different metabolic processes to coordinate these interdependent processes. It allows more precise regulation of the promoter gene and an adaptation to different environmental circumstances.

2.2.8 Identifications of regulator candidates in *C. reinhardtii*

To identify key TFs regulating *C. reinhardtii*'s primary metabolism, particular lipid metabolism, 83 promoters of enzyme-coding genes involved in lipid metabolism, photosynthesis, amino acid-, carbohydrate-, and nitrogen metabolism were screened against all 220 cloned TFs. As indicated, nodes with high out- or indegree represent promising regulator candidates in a network [225, 226]. Eleven promoters were targeted by more than 19 TFs, ranking them the most targeted promoters in the Y1H screen (95th percentile). Several of these highly targeted promoters belong to well-studied genes (see Table 2.4).

Table 2.4: Highly targeted promoters in the Y1H

Promoter	Indegree	**Gene ID**	**Description**
NRR1	23	Cre16.g673250	Known key regulator (TF) involved in TAG accumulation during N-starvation stress response [74].
FBP4	25	Cre16.g652150	Encodes for fructose-2,6-bisphosphatase/ 6-phosphofructo-2-kinase. This bifunctional enzyme regulates the fructose-2,6-bisphosphate concentration in the cell by catalyzing the formation and degradation of the allosteric regulator, Fru-2,6,-P_2. By influencing Fru-2,6,-P_2 concentration, carbohydrate metabolism is guided towards glycolysis or gluconeogenesis [233].
SQD1	21	Cre16.g656400	Encodes for DP-sulfoquinovose synthase, the major enzyme of SQDG syntheses. Mutations in this gene have a high impact of photosynthesis and sulfolipid synthesis. The mutant contains no SQDGs [234].
HYDA1	37	Cre03.g199800	[FeFe]-hydrogenase isoform 1 that is catalyzing H_2 production from fermentation or photosynthetic pathways [235].

PsaH	23	Cre07.g330250	Chloroplastic photosystem I reaction center subunit H that binds light-harvesting complex II (LHCII) and stabilizes photosystem I subunit D [236].
GPX5	22	Cre10.g458450	Thioredoxin-dependent glutathione peroxidase that catalyzes the reduction of hydrogen peroxide and organic hydroperoxides during photooxidative stress. Reactive oxygen species (ROS) can be formed during high light conditions and can cause severe damage to the cell [237, 238].
OTC1	19	Cre12.g489700	Encodes for ornithine carbamoyltransferase 1 that catalyzes the conversion of ornithine to citrulline, the first step of arginine biosynthesis in *C. reinhardtii*. It seems to play an important role in N assimilation during N deficiency, as it was one of eight genes whose gene expression was strongly impaired in the *tar1-1* mutant [239, 142].
Cre07.g334150	40	Cre07.g334150	not well characterized
S1Fa-like_2	35	Cre09.g386753	not well characterized
MYB_21	21	Cre08.g364050	not well characterized, MYB TF with SANT-domain
Cre10.g425100	20	Cre10.g42510 0	not well characterized

Seven of these 11 genes are well characterized in *C. reinhardtii* and play important roles in its metabolism. To be targeted by multiple TFs could give *C. reinhardtii* the possibility to adjust the activity of these enzymes and TFs very precisely, according to different environmental conditions and cellular needs. Four genes of the identified promoters are not well characterized but based on the other identified highly targeted promoters, they may also have important roles in *C. reinhardtii*.

One is an MYB-related TF (MYB_21, Cre08.g364050) of unknown function that contains a SANT-domain, which enables chromatin remodeling [240]. Two other TFs are not yet annotated. They belong to the genes Cre10.g425100 and Cre07.g334150. Cre10.g425100 has a predicted acyl transferase/lysophospholipase domain, a histone deacetylation domain, and a leucine-rich repeat. It could play a role in chromatin remodeling because deacetylation of chromatin and histones lead to a condensed chromatin structure.

This chromatin structure usually leads to gene repression [241]. Cre07.g334150 is predicted to encode for a TRAF TF but is also annotated with a triglyceride lipase activity to degrade TAGs. The fourth uncharacterized TF is the DNA binding protein S1FA (Cre09.g386753) [242].

S1FA was present as a promoter and active DNA-binding protein in the screen. S1FA is targeted by 35 other TFs and targets three other promoters. These three promoters belong to genes that have functions in carbohydrate and lipid metabolism. One promoter is additionally involved in the light stress response. The first targeted promoter of S1FA is the promoter of *ALD6* (Cre12.g501050). ALD6 is an aldehyde dehydrogenase that catalyzes the oxidation of acetaldehyde and coenzyme A together with NAD^+ to acetyl-CoA and NADH + H^+ and *vice versa*. This reaction supplies NADH + H^+ and acetyl-CoA, both important substrates for lipid and carbohydrate metabolism. The second promoter belongs to PDH2 (Cre03.g194200), an E2 subunit of the pyruvate dehydrogenase complex. It catalyzes the pyruvate and coenzyme A reaction to acetyl-CoA and CO_2. It is a critical reaction in carbohydrate metabolism and one starting point for the *de novo* FA synthesis. The third promoter belongs to GPX5 (Cre10.g458450). The GPX5 gene codes for a dual-targeted protein that can be localized in the chloroplast and the cytoplasm. GPX5 was shown to protect the cell from oxidative stress (ROS). The gpx5 mutant showed inhibited lipid biosynthesis with high ROS levels. In addition, down-regulation of photosynthesis was observed in the gpx5 mutant to minimize ROS accumulation. Therefore, GPX5 plays an important role in the homoeostasis of oxidative stress during lipid biosynthesis [237, 238]. All three genes, ALD6, PDH2, and GPX5, are part of different metabolic pathways and belong to different GO groups. The binding of S1FA to these three gene promoters belonging to different processes indicates that S1FA is involved in their regulation and strengthens the hypothesis that a TF does not target only promoters of one GO group. Furthermore, the fact that S1FA is targeted by 35 other TFs makes it an essential hub in the middle regulatory layer of the network and illustrates as an example that these four proteins may be involved in essential processes.

It was observed that many TFs target gene promoters belonging to different biological processes. Therefore, the question arose whether there are also TFs that target promoters of different GO terms more than expected or interact enriched with promoters of one GO term. To address this question a hypergeometric test was performed for all interacting TFs. Significant enrichment or depletion (p-value < 0.05) in terms of targeting promoters of a particular GO group was detected for 34 TF out of 142 TFs. For additional 25 TFs, the p-value for enrichment or depletion was highly significant (p-value < 0.01). Enrichment or depletion was observed for the GO groups amino acid metabolism, carbohydrate metabolism, lipid metabolism, nitrogen metabolism, photosynthesis, and unclassified promoters. The matrix for enrichment and depletion of single TFs is shown in Figure 2.10).

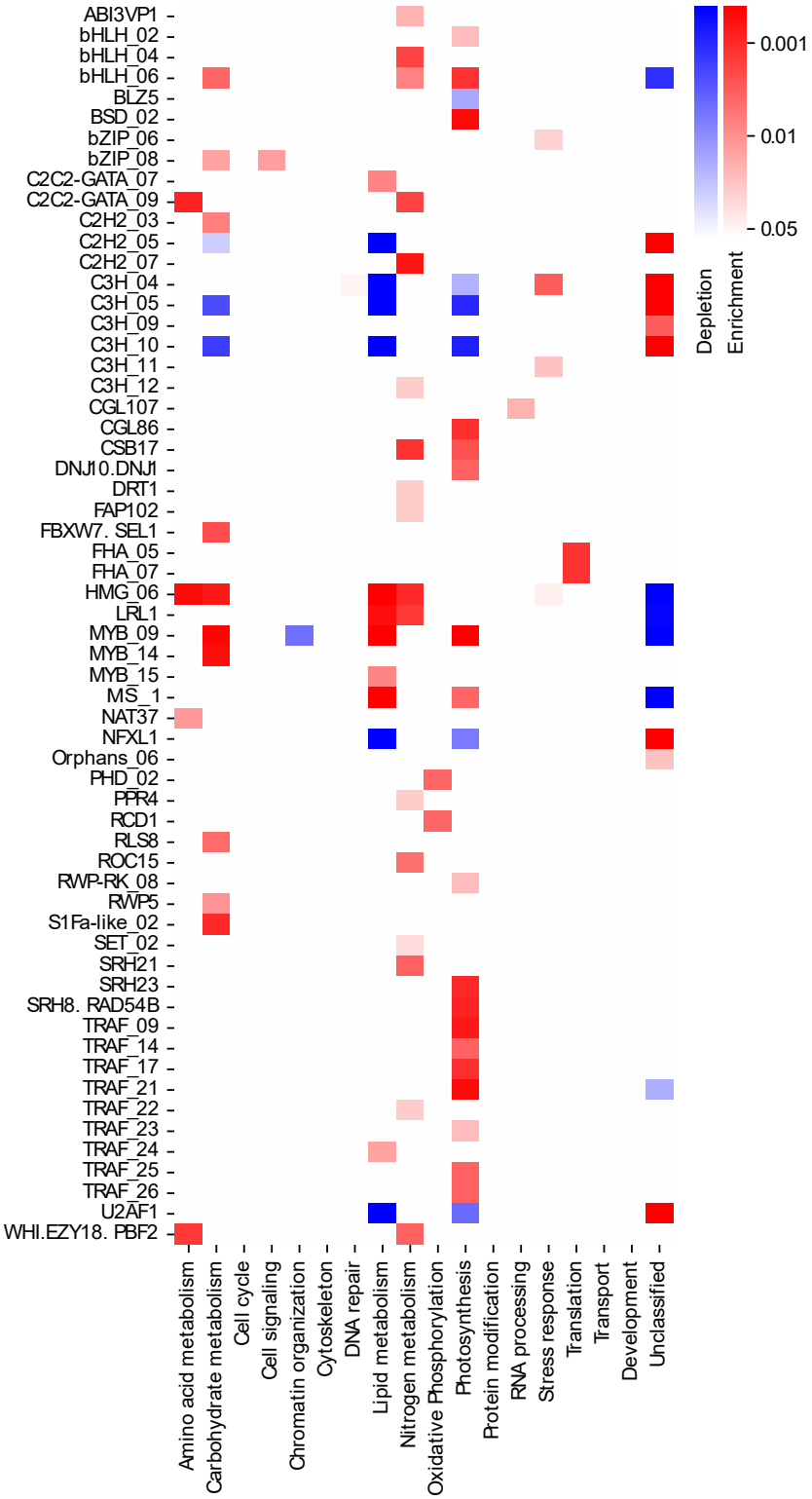


Figure 2.10: Enrichment analysis of TF interactions

Enrichment analysis of single TFs and their interaction profile with promoters of different GO groups. A significant depletion for interactions with promoters of a specific GO group is colored in blue. A significant enrichment is colored in red (p-value < 0.05). 59 TFs were observed to have a higher or lower interaction number, as expected, after the hypergeometric test.

Some TFs preferentially interact with promoters of genes from specific metabolic processes. Therefore, the question arose whether entire TF families are more involved in regulating specific metabolic processes than others.

2.2.9 TF families are involved in specific metabolic processes

The GO enrichment analysis confirmed that two TF families are more involved in targeting of promoters of metabolic processes than others. In addition, one TF family appears to be more involved in the general regulation of other TFs. The C3H-TF family, represented by 8 TFs in the Y1H screen and responsible for 602 interactions, targets significantly fewer promoters of the primary metabolism but have many more interations with promoters of other TFs than it would be expected by chance (see Figure 2.10). In contrast, MYB and MYB-related TFs show significant enrichment in their interactions for promoters of lipid and carbohydrate metabolism and photosynthesis. In addition, they display a significant depletion visible for interactions with promoters of unclassified TFs, protein modification, RNA processing and stress response. This indicates that the MYB and MYB-related TFs, tested in the Y1H screen, are more involved in the specific regulation of metabolic pathways than in the overall transcriptional regulation.

The TRAF TR family represents another interesting TF family. So far, not much is known about TRAFs in *C. reinhardtii*, even if they represent one of the largest TR families in *C. reinhardtii*. 19 of the 38 members were present in this Y1H screen, and caused 90 PDIs. Two TRAF TRs (Cre04.g219600 and Cre02.g145602) interacted with a high number of promoters (25 and 13, respectively). The remaining 11 TRAFs have an average interaction number of three. As indicated in Figure 2.10 and Figure A.4 in the appendix, the interactions are significantly enriched for photosynthesis promoters. This suggests that *C. reinhardtii* TRAF TRs may play a role in the photosynthesis regulation.

2.2.10 Promising TF regulators for *C. reinhardtii* lipid metabolism

The enrichment analysis revealed some TFs that interacted significantly more with promoters of specialized metabolic processes than others. A key regulator for a specific process is expected to interact with multiple promoters of this metabolic process. Section 2.2.8 describes promising regulator candidate genes depending on their indegree. Here, the outdegree is used to find promising regulator candidates. In the search for promising new key regulators of lipid metabolism, TFs characterized by a high outdegree were analyzed to determine whether they exhibit enrichment behavior for promoters of lipid metabolism.

MYB_9, MS_1, CGL86, HMG_6 and the known key regulator LRL1 all have a high outdegree of 24 or higher (see Figure 2.8), indicating a role as TF hubs in the Y1H

screen. LRL1 (Cre03.g197100) represents a TF regulator for lipid remodeling under certain stress conditions in *C. reinhardtii* [109]. It interacted with 26 promoters in the screen and is part of the highest network layer (see Figure (2.8). Together with the four other TFs, it targeted many promoters of lipid metabolism and other primary metabolism processes. All interaction partners of LRL1, MYB_9, MS_1, HMG_6 and CGL86 are shown in a subnetwork in Figure A.3 in the appendix. MYB_9, MS_1, and HMG_6 are, to date, uncharacterized TFs with unknown functions. They all belong to different TF families.

MYB_9 (Cre03.g144907) belongs to the MYB/MYB-like TF family like LRL1. MS_1 (Cre10.g450500) is an unknown TF. CGL86 (Cre12.g534450) contains a forkhead-associated domain (FHA) and belongs to the FHA TF family and HMG_6 (Cre16.g672300) is part of the high mobility group (HMG) TF family. MYB_9 interacted with 53 promoters, MS_1 with 37 promoters, HMG_6 with 24 promoters, and CGL86 with 60 promoters. The enrichment analysis for these proteins showed that LRL1, MYB_9, MS_1 and HMG_6 interacted significantly enriched ($p < 0.01$) with promoters of lipid metabolism. In addition, LRL1, MYB_9, MS_1, and HMG_6 showed significant depletion of interaction with promoters of uncharacterized TFs. More precisely, 13 %, 16 %, and 4 % of their interactions were with other TF promoters, respectively (MYB_9, MS_1, HMG_6). All other interactions were with promoters of functional genes involved in metabolism. CGL86 interacted with 13 of 33 lipid metabolism promoters but also interacted with 47 promoters of genes classified in other GO groups. Therefore, the hypergeometric analysis did not reveal an enrichment for lipid metabolism promoters, even if CGL86 targets almost 40 % of them. CGL86 showed a significant enrichment only for promoters involved in photosynthesis. Interestingly, 52 % of CGL86 targets were promoters of other TFs. Significant enrichment ($p < 0.05$) of all five hubs for other GO groups, in addition to lipid metabolism, are shown in Table 2.5.

Table 2.5: Enrichment analysis of interaction behavior of TF hub candidates

GO term	MYB_9	MS_1	HMG_6	LRL 1	CGL86
lipid metabolism	x	x	x	x	
carbohydrate metabolism	x		x		
photosynthesis	x	x			x
nitrogen metabolism			x	x	
aminoacid metabolism			x		

Since MYB_9, MS_1, and HMG_6 and CGL86 interacted with many promoters for lipid metabolism and are therefore potential regulators or lipid metabolism, the question arose whether they have a similar interaction profile to LRL1, a known important regulator of lipid metabolism. In addition, a comparison with PSR1, another known regulator of lipid metabolism, would have been interesting. However, PSR1 could not be cloned and was therefore not part of the Y1H screen. By comparing the interaction pattern of MYB_9, MS_1, HMG_6, and CGL86 with LRL1, it became obvious that these four

TFs share a high number of similar interactions with it (shown in Figure 2.11). MYB_9 and MS_1 interact with 96 % of the promoters targeted by LRL1, plus with additional promoters. CGL86 and HMG_6 share 12 interactions (46 %) with LRL1.

Comparison of the associated Jaccard index also showed that MYB_9 and MS_1 have high similarity with LRL1 (see Table A.6 in the appendix). However, MYB_9, MS_1, and CGL86 have significantly more interactions than LRL1, influencing the Jaccard index calculation.

Table 2.6: Overlap of interaction

***C. reinhardtii* gene ID**	**TF**	**PDIs**	**MYB_9**	**LRL1**	**MS_1**	**CGL86**	**HMG_6**
Cre03.g144907	MYB_9	53	-	25	35	26	15
Cre03.g197100	LRL1	26	25	-	25	12	12
Cre10.g450500	MS_1	37	35	25	-	18	13
Cre12.g534450	CGL86	60	26	12	18	-	12
Cre16.g672300	HMG_6	24	15	12	13	12	-

It is known that the stress adaptation of lipid metabolism in *C. reinhardtii* affects many other metabolic processes such as carbohydrate and nitrogen metabolism, as well as photosynthesis. Assuming a distributed regulation model for the metabolic regulation, as it was indicated in the GRN, a key metabolic regulator is also expected to interact with multiple promoters of different metabolic processes to influence them in addition to its interactions with promoters of lipid metabolism. Since this is the case for MYB_9, MS_1 and HMG_6, they represent new key regulator candidates of lipid metabolism and related processes.

MYB_9 and MS_1 interacted with more targets than HMG_6 and showed a higher interaction overlap with LRL1 targets. Therefore, the individual target promoters of MYB_9 and MS_1 were listed and analyzed for biological relationships (see Table 2.7). The target genes of MYB_9 are illustrated in their biological context in Figure 2.11. For a broader overview, a detailed list of all targets of LRL1, MYB_9, MS_1, HMG_6, and CGL86 are shown in the appendix in Table A.7.

Table 2.7: Description of single promoters targeted by MYB_9 and MS_1

TF	**gene promoters**	**biological process**
MYB_9, MS_1	KAR1	*de novo* FA biosynthesis
MYB_9	BCC2	*de novo* FA biosynthesis
MYB_9, MS_1	FAD7	FA desaturation
MYB_9, MS_1	ADH11	very long chain FA biosynthesis

MYB_9, MS_1	Cre14.g615050	very long chain FA biosynthesis
MYB_9	Cre08.g363500	very long chain FA biosynthesis
MYB_9	MECR, NRBF1	very long chain FA biosynthesis
MYB_9, MS_1	Cre08.g377300	TAG biosynthesis
MYB_9, MS_1	Cre08.g365950	TAG biosynthesis
MYB_9, MS_1	Cre07.g334150	TAG degradation
MYB_9, MS_1	Cre01.g038900	TAG degradation
MYB_9, MS_1	TGL18	TAG degradation
MYB_9	PHC45	TAG degradation
MYB_9, MS_1	ACOX2	FA activation, export and β-oxidation
MYB_9, MS_1	Cre03.g182050	FA activation, export and β-oxidation
MYB_9, MS_1	PLAP2	LD structural proteins
MYB_9, MS_1	PLAP7	LD structural proteins
MYB_9, MS_1	LBP1	sulfolipid synthesis
MYB_9	SQD1	sulfolipid synthesis
MYB_9, MS_1	Cre10.g425100	phosphatidylcholine acyl editing
MS_1	GGPS1	terpenoid biosynthesis
MYB_9, MS_1	GAP1	carbohydrate metabolism
MYB_9, MS_1	FBP4	carbohydrate metabolism
MYB_9	PDH2	carbohydrate metabolism
MYB_9	DLA2	carbohydrate metabolism
MYB_9	TAL1	pentose phosphate pathway
MYB_9	RPE2	pentose phosphate pathway
MYB_9, MS_1	ACLA1	TCA cycle
MYB_9, MS_1	LHCBM9	photosynthesis
MYB_9, MS_1	PSBP1	photosynthesis
MYB_9, MS_1	PETM	photosynthesis
MYB_9, MS_1	PSBP6	photosynthesis
MYB_9	CGL30	photosynthesis
MYB_9	LHCBM5	photosynthesis
MYB_9	PSBX	photosynthesis
MYB_9	FNR1	photosynthesis
MYB_9, MS_1	LHCB4	photosynthesis
MYB_9, MS_1	PsaH	photosynthesis
MYB_9, MS_1	CAH3	photosynthesis + CCM
MYB_9	NII1	nitrogen metabolism
MYB_9, MS_1	OTC1	nitrogen metabolism
MYB_9	NRR1	nitrogen metabolism
MYB_9, MS_1	NAR1.2	nitrogen metabolism
MS_1	GLN1	nitrogen + amino acid metabolism
MYB_9	HAS1	amino acid + ACP metabolism
MYB_9	HYDA1	stress response
MYB_9, MS_1	CYP51G1	stress response

MYB_9, MS_1	GPX5	stress response
MYB_9, MS_1	Cre09.g386753	TF of unknown function
MYB_9, MS_1	Cre04.g216200	TF of unknown function
MYB_9, MS_1	Cre08.g364050	TF of unknown function
MYB_9	Cre04.g216200	TF of unknown function
MYB_9, MS_1	APK1	cell signaling
MYB_9, MS_1	Cre07.g351550	cell signaling
MYB_9, MS_1	SHR12	DNA repair
MYB_9, MS_1	NAT5	chromatin remodeling

MYB_9 and MS_1 bound to many gene promoters of different metabolic processes in the cell. They interacted with promoters of key enzymes involved in carbohydrate metabolism, nitrogen metabolism, and photosynthesis (see Figure 2.11. Further, they targeted promoters involved in various parts of the lipid metabolism as *de novo* FA biosynthesis, FA activation, export, desaturation, and degeneration; TAG biosynthesis; TAG degeneration; phospholipid synthesis and structural protein synthesis for lipid droplet formation.

Their targets reflect a diversity and interconnectivity of the different metabolic processes. MYB_9 and MS_1 appear to regulate many important processes in the cell, which only strengthens the hypothesis that they are potential key regulators for altered cell metabolism during stress. CGL86 seems to be more involved in the upper-level regulation of other TFs.

As promising regulatory candidates, MYB_9, MS_1, HMG_6, CGL86, and LRL1 were part of the selection of TFs for the previously described PampDAP-seq experiment performed to validate interactions from the Y1H screen. The cloning of HMG_6 into the expression vector pGEX6 failed. MYB_9 also dropped out of the collection because the bacterial transformation into Rosetta™ (DE3) competent cells failed. It appeared that MYB_9 had a lethal effect on the cells as the DNA concentration and the transformation procedure corresponded to those of the other TFs that were successfully transformed into Rosetta™ (DE3) competent cells. Cloning of MYB_9 was repeated several times but never worked. This reinforces the assumption that MYB_9 harmed the cells. Protein expression of MS_1 as GST-fusion protein failed or the protein expression was so low that it could not be detected. Probably, protein expression failed due to the large protein size of around 100 kDa of MS_1.

CGL86 and LRL1 could be expressed, purified, and examined in the PampDAP-seq experiment. For CGL86, the most considerable overlap of interactions of all TFs was observed. 26 of 60 Y1H interactions were also detected in PampDAP-seq. For LRL1, 7 out of 26 Y1H interactions could be confirmed by PampDAP-seq. This corresponds to an overlap of interactions of almost 44 % and 30 %, respectively.

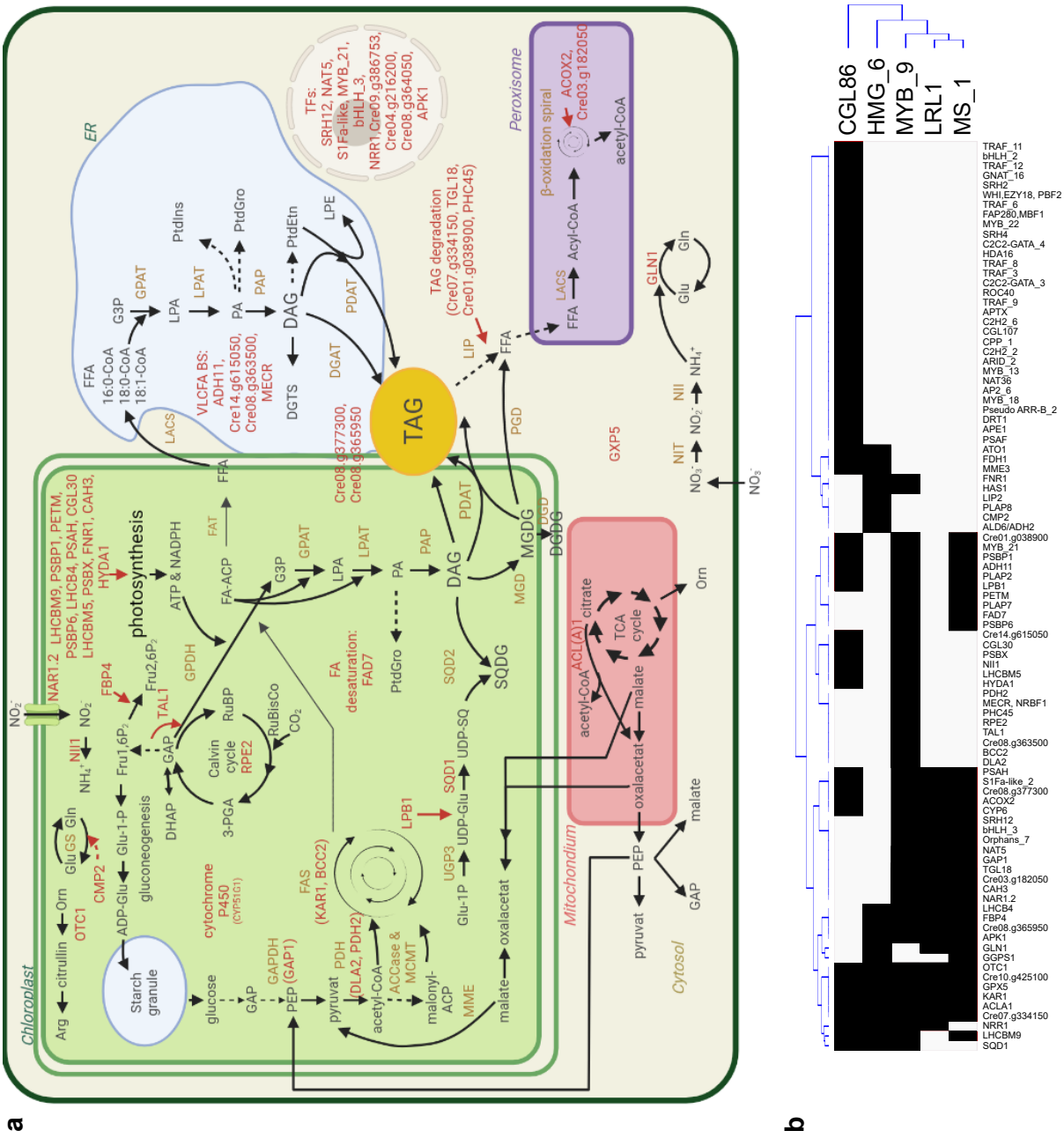


Figure 2.11: Target promoters of key candidates

a Promoter target genes of MYB_9 are illustrated in their metabolic context. Target genes are shown in red, enzymes in brown, metabolic substrates in black. 'Photosynthesis' includes PSI, b6f, PSII and ATP-synthesis. The pathways are simplified to emphasize the reactions that involve target genes of TF MYB_9. **b** Interaction pattern of MYB_9, MS_1, HMG_6, LRL1, and CGL86. Black represents an interaction. All five TFs share multiple interactions.

2.2.11 Search for insertion mutants

Subsequent analyses would include *in vivo* experiments to demonstrate the biological function of the proposed key regulators candidates. Therefore, knock-out mutants for MYB_9, MS_1, CGL86, and HMG_6 were sought, since a vast mutant library is available for *C. reinhardtii*.

No mutants were available for MYB_9, CGL86, and HMG_6. Only one mutant was available for MS_1, whereas the insertion mutation is close to the 3'-end of the ORF. The protein's function is most likely not affected by this mutation. The absence of mutants for MYB_9, CGL86, and HMG_6, may support the hypothesis that they are essential in regulating primary metabolic processes in *C. reinhardtii*. Propably, an insertional mutation in these genes has a too strong effect on the regulation of metabolism in the cell and is, therefore, lethal.

2.2.12 Overexpression of key candidates in *C. reinhardtii*

Another way to test whether these TFs exhibit a biological effect in *C. reinhardtii* is to overexpress them *in vivo*. In collaboration with the research group of Prof. Olaf Kruse of the University Bielefeld, the overexpression mutants for LRL1, CGL86, MYB_9, and MS_1 were planned to be generated. For exogenous protein expression in *C. reinhardtii*, all sequences had to be codon-optimized. Furthermore, intron insertions into the expression sequence are required for functional protein expression in *C. reinhardtii* [243, 244, 245].

All sequences, except for MS_1, were gene synthesized by GenScript and cloned into a YFP-fusion encoding vector construct for transient transformation by Alexander Einhaus (Prof. Kruse's laboratory). Alexander Einhaus performed the transformation of the constructs several times but could only find cells with minimal expression of the regulatory proteins. Most transformed cells showed a 1-3x fold protein expression compared to the parental strain. In a first experiment for lipid quantification with Nile red staining no significant change in lipid concentration could be detected. Overexpression was no longer detectable after the transformed cell populations were exposed to nutrient stress. However, transcriptional regulation can occur even at low regulator concentrations since an mRNA can be translated multiple times. For *E. coli*, the average translation rate is around 40 translations per transcript [246]. Further, it was reported that the ratio of mRNA and protein is exponential to each other [247].

Therefore, the overexpressed cell lines were sent to our collaborators in Prof. Salehi-Ashtiani's laboratory at NYU Abu Dhabi for further analysis. The fatty acid composition of the overexpression cell lines and the overall FA content will be investigated. The results of these experiments were not yet available at the time this work was completed.

3 Discussion

The global climate crisis requires the use of renewable energy instead of conventionally used fossil fuels. Microalgae represent a promising resource for biofuel production, as neutral lipids, e.g., TAGs can be directly converted into biodiesel. However, most oleaginous algae studied do not have the optimal lipid composition for direct conversion to high quality biodiesel, and even if they could produce significantly more TAGs than conventional oil plants, increased lipid accumulation can so far only be triggered by external stress factors such as nutrient deficiencies. This, in turn, has adverse effects on the biotechnological production of TAGs in microalgae, leading to growth arrest and stress reactions in the alga cell. To date, no cost-effective and high-efficiency production has been established. Nevertheless, with a deeper global understanding of microalgae metabolism, genetic modification, and targeted manipulation, higher productivity of biotechnological lipids may be possible [95]. To systematically study *C. reinhardtii*, an ORFeome is immensely useful because it allows large-scale functional genomic studies.

This work aimed to generate the first ORFeome for the model organism *C. reinhardtii*, an essential resource for fundamental studies of microalgal biology. It will help to improve our understanding of algal metabolism and biology and allow conducting proteomic studies (section 3.1). Despite a broad knowledge of overall metabolism in microalgae, there is a high demand for a more detailed understanding of the regulation and diverse interdependencies of metabolism, especially for lipid metabolism, without stressing the cell. To date, knowledge of general transcriptional regulation and specific master regulators of essential cellular processes in *C. reinhardtii* is limited.

The second goal of this thesis was to generate a large-scale GRN that serves as a model to predict regulatory relationships and master regulators. This first experimental-based GRN for *C. reinhardtii* provides a comprehensive overview of the interrelationships in transcriptional regulation (section 3.2). In addition, the network analysis revealed exciting biological insights and suggested four promising regulator candidates for the regulation of lipid metabolism (section 3.3).

3.1 The *C. reinhardtii* ORFeome represents a powerful tool

The first global ORFeome for *C. reinhardtii* is essential for accomplishing future high-throughput functional studies such as Y2H or AP-MS. 8,627 of 15,157 attempted ORF could be successfully cloned, corresponding to a success rate of 57 %. The first cloning attempts of the *C. elegans* ORFeome resulted in 55 % of predicted protein-encoding ORFs [199]. Further cloning attempts were required to complete the ORFeome. The Rice ORFeome contained less than 10 % of all predicted *O. sativa* genes, more precisely 2,300 of 30,000-50,000 predicted protein-coding genes of *O. sativa* were cloned [201]. The *A. thaliana* ORFeome started with about 8,000 cloned ORFs (29 %). It was expanded several times and now comprises approximately 13,000 ORFs, which account for about 47 % of all protein-coding genes of *A. thaliana* [158, 248, 150]. The *S. cerevisiae* ORFeome published in 2005 included 5,854 ORFs [249], representing 93 % of all ORFs known up to that time. Further genome analysis increased the total ORF number to 6,769 ORFs, thus, the ORFeome collection corresponds to 87 % of all known ORFs. The *S. cerevisiae* ORFeome construction benefited greatly from previous cloning attempts and library constructions that already validated many ORFs and the genome annotation [250, 251, 252, 253].

Compared to other cloning attempts of previous ORFeoms, the success rate of the *C. reinhardtii* ORFeome was on a par or better, with the exception of the *S. cerevisiae* ORFeome. Interestingly, the cloning success for *C. reinhardtii* is more comparable to multicellular organisms than to the unicellular organism *S. cerevisiae*, even though *C. reinhardtii* is a unicellular organism. One possible reason could be the high GC content of the *C. reinhardtii* genome. Many polymerases had difficulties amplifying the *C. reinhardtii* ORFs, despite being reported to work with GC-rich templates and difficult amplicons. Further, the PCR conditions could not be adjusted for each ORF, which may also have resulted in a loss of a few amplicons. Nevertheless, KOD Hot Start polymerase performed best and enabled high-throughput cloning.

Another limitation for ORFeome cloning is that the cloning efficiency decreased with increasing ORF length. ORFs up to 1.0 kb in length were amplified with a 70 % success rate, whereas only approximately 30 % of ORFs larger than 4.0 kb could be amplified (see Figure A.1 in the appendix). This phenomenon has been observed in previous ORF cloning projects, e.g., the human ORFeome project [202], and was, therefore, to be expected.

Furthermore ORF cloning is dependent on the genome annotation. Here, the genome assembly and annotation v5.5 was used, which contains gaps, unplaced scaffolds, and misannotated genes. However, the quality and level of annotation play a crucial role in generating an ORFeome. Continued efforts to improve the ORFeome of *C. elegans* could increase the size of the first ORFeome by approximately 20 % in the second cloning at-

tempt. More accurate gene predictions have been shown to improve cloning success immensely [199].

In 2022, a new genome annotation for *C. reinhardtii*, called v6, was released, which proved that the gene annotation v5.5 contained incorrect annotations. The preprint of Craig et al. compares and annotates the *C. reinhardtii* genes. The total number of nuclear genes dropped from 17,741 to 16,795 genes [254]. More interesting, however, is the increased number of alternative transcripts, which rose from 1,789 to 14,874 alternative transcripts [254]. There is a possibility that the discrepancy in the length of the ORF detected in 14 % of the 1st PCR amplicons is due to this misannotation. In fact, the ORFeome provides systematic and experimental annotation evidence for more than 8,000 ORFs. It would be interesting to re-align the sequencing data of the cloned ORFs with the new genome annotation. Some genes may have been cloned but not assigned to the correct genes because of incorrect reference annotation.

For most predicted genes in *C. reinhardtii*, the function is uncharacterized. Only 4,408 *C. reinhardtii* genes have assigned gene names, and their functional role is known. An additional 1,119 genes have a description and functional prediction. Of 8,627 ORFeome Entry clones, only 2,450 have an assigned gene name within the new genome annotation v6.1. This implies that 6,177 ORFs have been cloned, but their function is unknown or only predicted. Even if the function of thousands of genes or precisely proteins are already known, information at the structural, functional, and dynamic levels is needed to develop new models for cellular systems and to understand the underlying principles of cellular systems [196].

The ORFeome can be used to map PPI and regulatory networks and can help to provide this information. 4,893 *C. reinhardtii* ORFs have already been successfully transferred into the Y2H target vectors and can be screened in a BFG-Y2H screen. Hence, the ORFeome represents a powerful resource for further fundamental functional studies.

3.2 Revealing key elements of the *C. reinhardtii* gene regulatory network

A key question of this thesis was to explore the regulatory relationships within the *C. reinhardtii* metabolism using systematic approaches. Once the ORFeome was available, I used this potent resource to map transcriptional regulation in *C. reinhardtii*. Although there are some RNA-seq-based co-expression networks for *C. reinhardtii*, it lacks an experimental GRN. The Y1H method is designed for large-scale mapping of PDIs between TFs and promoters of target genes and was used here to generate the first experimental GRN for *C. reinhardtii* global map of these regulatory interactions in *C. reinhardtii*.

3.2.1 Y1H identified PDIs for 30 % of annotated TFs

The experimental setup for this Y1H screen differs from commonly used eY1H setups. All high-throughput Y1H assays provide a tremendous increase in performance compared to small scale Y1H assays. eY1H, in particular, was used to identify and analyze *C. elegans*, *A. thaliana* and human regulatory PDIs [177, 222, 193, 255, 168, 167, 256, 257]. Here, a different growth strategy for the yeast cultures was established as they were mated and grown in liquid media. A TECAN Fluent liquid handling robot was used in place of an HDA RotoR robot (Singer Instruments), a commonly used array pinning robot for eY1H [193, 256]. The cultivation in liquid media resulted in an incomplete selection of diploid cells and a background of unmated haploid cells that remained in the solution and falsified the result. Therefore, the Y1H screening pipeline was extended by a selection step to ensure that only diploid yeast cells survive. After implementing the adapted Y1H method in the laboratory, the screen revealed 1,451 interactions between 142 TFs and 200 DNA promoter baits. It utilized a newly compiled TF collection representing 46 % of all annotated *C. reinhardtii* TF. The resulting network describes interactions between 64.5 % of TFs with 66 % of promoters.

The detection of interactions found for 66 % of promoters was at the lower end of those reported in previous studies, where PDIs were identified for 65 - 96 % of investigated promoters [257, 177, 168, 258, 169]. However, in most previous Y1H studies, fewer DNA promoters were analyzed than were examined here. Deplancke et al., e.g., screened 116 promoters of *C. elegans* against 167 TFs, and Gaudinier et al. screened 98 promoters against 345 TFs [168, 175]. Only Fuxman Bass et al. screened more than 300 promoters in a single Y1H study [259]. The fact that interactions were not found for every promoter may be due to the extensive promoter matrix of 303 promoter strains, which did not allow retesting of individual promoter strains for technical reasons.

Furthermore, there might be biological reasons as the TF promoters in this study covered only 46 % of all *C. reinhardtii* TFs and they were also tested against only a subset of TFs (also 46 %). Brady et al. screened a subset of tissue-specific *A. thaliana* TFs against promoters of genes that encode these TFs and identified even fewer PDIs for the TF promoters screened, namely, 25 % [173].

Finding interactions for 64.5 % of investigated TFs is higher than reported in other studies, in which the proportion of interacting TFs was below 45 % (see Table 3.1 and [224, 222, 171, 259, 174, 175]). Furthermore, no inherent preference for a particular TF family could be seen in the Y1H screen (see Figure 2.6). This phenomenon was also observed before, e.g., for *C. elegans* [122, 168, 258], but it is still surprising since the Y1H assay has limitations.

Y1H is conducted in yeast, which can affect the results due to its genetic background. Furthermore, Y1H can only detect interactions of monomeric or homodimeric TFs.

Table 3.1: Previously reported Y1H studies

Publication	Organism	TFs	Promoters	found PDIs	Interactors	Promoter size
Deplancke et al. 2006, Cell [168]	*C. elegans*	167	116	283	117 TFs vs. 72 baits	2,500 bp
Reece-Hoyes et al. 2011, Nature Methods [224]	*C. elegans*	865	50	769	160 TFs vs. 48 baits	300 - 2,000 bp
Gaudinier et al. 2011, Nature [222]	*A. thaliana*, root	653	13	158	85 TFs vs. 13 baits	2,000 bp
Taylor-Teeples et al. 2014, Nature [171]	*A. thaliana*	467	50	617	from 242 genes (not specified)	450 - 3,000 bp
Fuxman Bass et al. 2015, Cell [259]	*H. sapiens*, disease related	1,086	360	2,230	283 TFs vs. 246 baits	400 - 2,400 bp
Yang et al. 2017, Mol. Plants [174]	*Z. mays*	1,901	54	1,100	568 TFs vs. 54 baits	1,000 bp
Ikeuchi et al. 2018, PCP [167]	*A. thaliana*	559	48	1,162	252 TFs vs. 48 baits	2,000 bp
Gaudinier et al. 2018, Nature [175]	*A. thaliana*	345	98	1,660	not specified	2,000 bp
Tang et al. 2021, Mol. Syst. Biol. [176]	*A. thaliana*	2,036	224	27,485	220 baits vs. 1,930 TFs	2,000 bp

Since many TFs bind DNA as heterodimeric dimers, these interactions will be missed by the Y1H screen. TFs that need post-translational modifications absent in yeast to bind to DNA, or are misfolded in yeast when fused to the AD-domain, can also be missed in the screen [260, 256, 122]. However, PDIs and, thus, DNA-binding evidence for around 30 % of all annotated *C. reinhardtii* TFs were detected. This suggests excellent coverage since other primary gene-centered PDIs networks identified interactions for 10-20 % of annotated TFs in their respective organisms [168, 222, 259].

The total PDI number of 1,451 is comparable to the previously mentioned Y1H studies in Table 3.1. Only two PDI studies [259, 176] identified more interactions than in this screening. Tang et al. found significantly more PDIs on an entirely different scale. Nevertheless, comparing the PDI to active TF ratio, this Y1H screen here performed very well. 1,451 PDIs identified for 142 TFs represent an average ratio of 10.2 PDIs/TF. The network of Fuxman Bass et al. [259] has an average ratio of 7.88 PDIs/TF and Tang et al. a ratio of 14.2 PDI/TF [176].

However, since a Y1H network is a directed network, these ratios are not necessarily biologically relevant and do not reflect connectivity or density of the network per se, since a TF can also interact with, e.g., 509 promoters and only one promoter [176]. Moreover, the various PDI studies were performed with different biological questions, and as already seen in Brady et al. [173], the choice of TFs and promoters to be studied has a significant influence on the outcome. Thus, these total interaction numbers are not easily comparable in a biological manner. In addition, varying amounts of TFs and promoters were investigated using different technical pipelines and parameters.

To compare Y1H results of different experiments more consistently, it would be of great benefit to establish an empirical framework of quality parameters for a network, as developed by Venkatesan et al. for mapping binary interactomes [261]. However, this would have been beyond the scope of this work and therefore a saturation assay and a second orthogonal assay were performed to firstly determine the sample sensitivity of the primary screening and secondly to be able to assess the quality of the detected Y1H interactions. The performed saturation assay confirmed the assumption that only a subset of interactions was detected. The saturation assay indicates that the primary screening detected around 34 % of possible interactions. Nevertheless, it demonstrated a high technical quality of the Y1H primary screening due to the high reproducibility between independent repeats.

A second independent method was used to validate the Y1H data set and to estimate the data quality. Indeed, once it has been shown that a data set consists of reliable interaction data, the interactions can be analyzed for their biological and structural significance [223]. The second orthogonal interaction assay performed was a modified ampDAP-seq assay called PampDAP-seq. DAP-seq is a cost-effective method that detects TF binding sites in a high-throughput manner, similar to Y1H. The PampDAP-seq method takes

advantage of the already cloned and amplified DNA promoters, which serve as a DNA library for the binding assay. Protein stability or particular requirements for protein activity, like protein partners, cofactors, or post-translational modification affects the TF performance in this assay. It is likely the reason why PDIs could not be identified for all ten proteins but for six of ten TFs in two biological repeats. A total of 320 unique PDIs were detected in the PampDAP-seq assay.

To put these results in context, they were compared with previous studies. Due to the lack of previously reported comparisons between Y1H and DAP-seq experiments, a comparison of ChIP-Seq and eY1H was performed to relate the results to other regulatory networks. As no single assay can detect every interaction, the interaction profile detected by two different assays will only partially overlap [223]. Fuxman-Bass et al., Reece-Hoyes et al., and Santoso et al. describe comprehensive GRNs for *C. elegans* and *H. sapiens* TFs. Their studies revealed a significant PDI overlap of 11.7 % [259], 17.5 % [262], 20.0 % [255], and 20.3 % [257] between eY1H and ChIP-seq. This study's PDI overlap between Y1H and PampDAP-seq is 23.8 %, which is on a par with previous reports and thus indicates a high quality of the generated Y1H network map.

3.3 Biological insights

After evaluating the quality of the network, a central question of this thesis was whether new biological insights could be gained through the network.

3.3.1 The Network displays a distributed regulatory system for metabolic promoters

The *C. reinhardtii* regulatory network reflects the structure of a scale-free network, consisting of a few highly connected hubs and many nodes with low degrees [263]. The network topology of GRN of *E. coli* and *S. cerevisiae* show a pyramid-shaped hierarchical structure. They consist of four layers: top, middle, bottom, and target. The top, middle, and bottom layer consists of TFs, and the lowest layer is composed of target genes [264, 265]. Prokaryotic promoters, e.g., promoters of *E. coli*, were believed to be regulated by a single specific regulatory protein, either a repressor or an activator. This specific regulation was initially identified in the lac operon regulation by the LacI repressor [266]. Accordingly, it has been believed that a single TF regulates one specific gene or a small number of transcription units [267, 268]. However, more advanced techniques have enabled screening of entire genomes for regulatory DNA sequences bound by TFs. Most *E. coli* promoters were found to be regulated by multiple TFs [269]. Consequently, the term regulon was introduced for a regulatory unit of a set of multi-target regulators that control a large number of genes [270]. However, also single global regulators were identified for specific metabolic processes that regulate an entire pathway or process, e.g., the TF cAMP receptor protein (CRP) that regulates genes for transport

and utilization of carbon sources in *E. coli* [270, 271]. Additional studies on *E. coli* and *S. cerevisiae* led to a model in which different transcription units control specific cellular processes or metabolic pathways [272, 273, 274, 275, 276, 176].

A critical question was how plant metabolism is regulated to maximize performance under various biotic and abiotic environmental stresses. The aliphatic glucosinolate metabolism plays an essential role in the defense against microorganisms and herbivores and was used to study this question [277]. Several key TFs were identified to regulate the aliphatic GSL pathways in *A. thaliana*. Mutants of the TF genes *MYB28*, *MYB29*, and *MYB76* resulted in impaired accumulation of GLS metabolites. However, it seemed that these key regulators do not regulate the entire metabolic pathway, as their mutations only led to a gene expression reduction of some, but not all, GLS synthesis genes [277, 278, 279]. Therefore, they did not correspond to the classical definition of a master transcriptional regulator and suggested the necessary involvement of other TFs. A large-scale Y1H study by Li et al. identified numerous additional TFs that bind the aliphatic GLS enzyme promoters and could also regulate GLS accumulation [170, 280].

The transcriptional regulatory network in *A. thaliana* is much more interconnected and complex than previously thought. Several large-scale PDI screens confirmed this. Taylor-Teeples et al. identified hundreds of novel regulators of the developmental regulation of secondary cell wall differentiation [171]. Another large-scale Y1H study on plant regeneration in *A. thaliana* also revealed a multilayered transcriptional network that is highly interwoven [167]. Nevertheless, they could also identify critical regulators of certain underlying subnetworks that regulate specific processes. These key regulators play essential roles in transmitting signals in their subnetwork [171, 167].

The GRN for nitrogen-associated metabolism in *A.thaliana* consisted of highly interconnected interactions and demonstrated that central TFs regulate multiple processes. They demonstrated that TFs regulating nitrogen metabolic pathways also regulate growth-promoting processes related to nitrogen metabolism [175]. The high interconnectivity was also observed in the GRN for *A. thaliana* primary and specialized metabolism. It showed that 90 % of TFs bound to promoters of genes in two or more metabolic pathways instead of promoters of one particular pathway. Consequently, a model for *A. thaliana* was proposed in which each metabolic pathway is influenced and regulated by a large number of TFs [176]. Such a distributed TF regulation system provides the ability to coordinate several pathways together as a unit and to fine-tune a precise response to either environmental or developmental events. Further, it increases the flexibility of a particular pathway.

The *C. reinhardtii* network displays a clear connectivity hierarchy. Only 15 TFs, described as the first layer in section 2.2.6, constitute for around 66.7 % of interactions in the network. However, the majority of TFs (80.3 %) bound only few promoters and form 19.9 % of all interactions. Moreover, after analyzing the interaction behavior of all TFs, it

becomes clear that 72.5 % of TFs in the network bind to promoters of two or more functional GO groups (Figure 2.9 d). This indicates that these TFs regulate genes of different metabolic processes and pathways. Prominent examples are the identified key regulators described in section 2.2.10. They all target promoters of very different metabolic processes, such as the different pathways of lipid and carbohydrate metabolism, and promoters of photosynthesis. This supports the assumption of a distributed regulation system in *C. reinhardtii* as it exists in *A. thaliana*.

It appears that the *C. reinhardtii* regulatory system additionally contains key regulators that regulate multiple metabolic pathways together as a module to allow a precise, fast, and flexible adaptation to external influences and stresses. A large proportion of promoters are targeted by six and seven different TFs. Remarkably, 82 % of these promoters belong to TF promoters (51 of 61), and only 11 are from enzymatic genes. This indicates that *C. reinhardtii* has a middle regulatory layer. These are well connected with other TFs but do not represent large hubs. Several TFs, known to concentrate external signals and trigger a response, are part of these middle TF layers. RCD1 is known to be involved in the intersection of mitochondrial and chloroplastic retrograde signaling cascades. It integrates ROS-dependent signals from the chloroplast and mitochondria for establishing transcriptional control over metabolic processes in both organelles [228, 281]. LSD1 is a negative regulator of programmed cell death but also plays an important role in the response and acclimation of plants to a wide range of abiotic stresses such as drought, UV-C radiation, and cold and light stress [282, 283, 284, 285, 227]. If LSD1 in *A. thaliana* is mutated, the cell can not stop cell death once activated [283]. LSD1 concentrates different external stimuli and mediates, through crosstalk with other TFs, global acclimatory and defense mechanisms in *A. thaliana* cells [284]. Heat shock TFs sense temperature changes and activate the transcription of heat shock proteins, which mediate a heat response, a survival mechanism of the cell. All of these TFs are known essential mediators that detect external influences and induce intracellular stress responses. These examples of known TFs support the hypothesis of an intermediate regulation level that converges multiple signals of several regulation strands and TFs and mediates a subsequent cellular response.

3.3.2 TRAF TF family showed enriched interaction behavior

Here, the network suggests a distributed regulation system together with key regulators. Further, several PDI studies have shown that TFs from different TF families regulate metabolic pathways [171, 167, 175, 222, 176]. The enrichment analysis also revealed multiple TFs of different TF families with an enrichment or depletion for promoters of a certain GO group. Whether entire TF families preferentially interact with promoters of a particular GO group and have thereby a unique role in *C. reinhardtii* was investigated by analyzing the interaction behavior of these TF families.

The MYB TF family showed enrichment interaction behavior. Nevertheless, it was not limited to one metabolic process. The enrichment of interaction could be detected for promoters of lipid-, carbohydrate-, and photosynthesis metabolism (see Figure A.4 in the appendix). This means the MYB TFs are not exclusively involved in regulating one particular metabolic process. Four of 15 MYB TFs tested in the Y1H screen, interacted with more than 15 promoters. These results support the assumption of a distributed regulatory system and reinforce the suggestion that MYB TFs play a primary role in the regulatory coordination of several metabolic processes.

Only one TF family could be allocated to a specific metabolic process. The TRAF family represents this exception. They represent one of the largest TR families in *C. reinhardtii*, with up to 38 members. This is surprising because they represent a minor TR group in other organisms. *C. elegans* has one member [286], *D. melanogaster* two [287], and mammals have six TRAF TRs [288]. Not much is known about TRAF TFs, but they have been described as molecular adapters or E3 ubiquitin ligases that facilitate the signal transduction from cell surface receptors to their downstream effectors in plants [289]. 19 of the 38 members were studied in the Y1H screen, and 90 PDIs detected. Two TRAF TRs showed many interactions (25 and 13) in the Y1H screen with various promoters. The other 16 studied TRAF TRs interacted with an average of three promoters and were responsible for 52 interactions. The enrichment analysis of their interactions revealed that TRAF TFs interacted significantly more with photosynthesis promoters in the Y1H screen than expected. Transcriptional regulators usually do not interact directly with DNA but with other TFs or proteins, thereby regulating transcriptional activity [123]. The *C. reinhardtii* TRAF TFs, which showed interactions in the screen, contain additional to their E3 ubiquitin ligase domain also a DNA-binding domain. Thus they can interact with DNA, and furthermore they could also act as chromatin remodeling proteins due to their ubiquitinylation function. Thus, their biological function, possibly as regulators of photosynthesis, and why they are so expanded in *C. reinhardtii* [182] remain to be determined. The fact that not only individual TFs but also entire plant TF families are involved in the regulation of specific processes has been described before and supports that the TRAF TF family in *C. reinhardtii* may have a specialized role.

The WRKY-TF gene family, e.g., plays an important role in regulating transcriptional reprogramming in the context of plant stress responses [290, 291]. The WRKY TF is one of the largest TF families in plants. They are involved in the response to abiotic stress and defense regulation but also contribute significantly to plant growth and development [292]. In addition, the AP2/ERF (ethylene-responsive factor), NAC (NAM, ATAF1/2, CUC1/2), and MYB TF families are also reported to play essential roles in controlling plant-specific regulation [293, 294, 295, 296, 297].

3.3.3 Identification of key regulator candidates of the lipid metabolism

The GO enrichment analysis showed that three TFs, specifically MYB_9, MS_1 and HMG_6 target significantly enriched promoters of lipid and primary metabolism in *C. reinhardtii*. Could these TFs represent novel key regulators of the lipid metabolism in *C. reinhardtii*?

The interaction pattern of MYB_9 and MS_1 contained, among other promoters with 96 % concordance, the target promoters of LRL1, a known TF regulator of lipid remodeling under certain stress conditions in *C. reinhardtii*. HMG_6 and CGL86 shared 46 % of LRL1 targets. MYB_9, MS_1, and HMG_6 interact with multiple promoters of lipid-associated genes and promoters of photosynthesis and central carbon metabolism or nitrogen-associated metabolism. Half of the promoter targets of CGL86 belong to enzymatic genes. The other half are TF promoters.

I propose that all four TFs (MYB_9, MS_1, HMG_6, CGL86) play essential roles in regulating lipid metabolism, particularly in stress-induced TAG accumulation and metabolic adaptation to stress. The metabolic adaptation of *C. reinhardtii* to stress, especially nutrient stress, is highly complex. N, S, or P deficiency result in increased TAG accumulation and a significant lipid metabolism remodeling. This remodeling involves the *de novo* biosynthesis of lipids, the degradation of existing membrane lipids, and the conversion of existing FAs. Other metabolic processes, such as photosynthesis, starch metabolism and the associated carbohydrate metabolism, are also involved in this altering process.

The precursors of lipid biosynthesis originate from carbohydrate metabolism. Thus, there is a direct competition between starch metabolism and lipid biosynthesis, necessitating close corregulation of the carbohydrate metabolism. Furthermore, protein and lipid synthesis are also in competition with each other. All regulator candidates target promoters of several important key enzymes of different metabolic pathways. The detailed description of the different targets is intended to illustrate the complexity of the regulation of these TFs.

MYB_9 interacts with the promoters of *PDH2* and *DLA2*. PDH2 and DLA2 are pyruvate decarboxylation enzymes that catalyze the conversion of pyruvate to acetyl-CoA. By regulating the gene expression of pyruvate decarboxylation enzymes, their protein concentrations are changed, which directly influences the acetyl-CoA production in the cell. Another target of MYB_9 and all other mentioned potential key regulators (LRL1, MS_1, CGL86, HMG_6) is the promoter of the enzyme ACL(A)1. ACL(A)1 is a subunit of the ATP citrate lyase in the TCA cycle. ATP citrate lyase catalyzes the reverse enzymatic step of the citrate synthase. It converts citrate into acetyl-CoA and oxaloacetate under the consumption of ATP. Assuming that ATP citrate lyase activity is increased, this leads to a raised acetyl-CoA concentration in the cell, which could subsequently be used in FA biosynthesis. Deng et al. showed that the *C. reinhardtii* citrate synthase

mutant, which can not synthesize citrate of acetyl-CoA and oxaloacetate anymore, significantly accumulated increased TAG levels (169 %). This solidifies the assumption that there is substrate competition in green algae [298] and that the different biosynthesis routes of acetyl-CoA play important roles in the multifactor-dependent biosynthesis of lipids. Since acetyl-CoA, which is needed for FA de novo synthesis, is also used as a substrate to form oxaloacetate or enters directly as acetyl-CoA the TCA cycle, thus also provides the substrates and energy for aminoacid synthesis.

Another enzyme that directly influences the acetyl-CoA concentration and is also a target of MYB_9, MS_1, LRL1, and CGL86 is the promoter of *ACOX2*. ACOX2 is an acyl-CoA oxidase/dehydrogenase that catalyzes the first step of FA β-oxidation, the breakdown process of FAs into acetyl-CoA units in the peroxisome. An impaired FA β-oxidation significantly impacts FA turnover in *C. reinhardtii*. The mutant strain accumulated 20 % more neutral lipids than the wild-type strain under N deprivation [299]. The regulation of the acetyl-CoA level by controlling these promoters could lead to an increased acetyl-CoA concentration and a subsequent enhanced FA biosynthesis.

Two enzymes directly involved in the *de novo* FA synthesis are KAR1 and BCC2. BCC2 is part of the ACCase complex that catalyzes the ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA in a multistep reaction. It represents the first step in FA synthesis and is the rate-limiting step for the pathway [300]. Experiments where the activity of ACCase was inhibited in fungi resulted in FA depletion and subsequent rapid cell death [301]. KAR1 represents one subunit of the multi-subunit fatty acid synthetase complex (FAS complex). In several steps, the FAS complex forms from acetyl-CoA and malonyl-CoA, a fatty acid-ACP molecule, a precursor for DAG- and TAG-biosynthesis. Previous overexpression experiments of cyanobacterial KAR in the chloroplast of red algae *Cyanidioschyzon merolae* could increase TAG accumulation while maintaining the cellular growth of the algae [302]. MYB_9 interacts with the promoters of both genes, MS_1, LRL1, and HMG_6, only with the promoter of *KAR1*. The regulation of both enzymes could directly affect FA synthesis and, in turn, TAG biosynthesis.

Another exciting target of MYB_9 and MS_1 is the promoter of *fad7*. FAD7 is the only reported plastid desaturase in *C. reinhardtii*, while multiple isoforms are known for other plants [68]. FAD7 is responsible for the production of PUFAs [303]. Since PUFAs account for more than 50 % of total FAs in *C. reinhardtii*, the FA desaturase enzyme FAD7 is crucial. The *fad7* mutant strain exhibits a reduction of more than 65 % of PUFAs compared to the wild-type *C. reinhardtii* strain but no difference for SFAs or MUFAs was observed. The gene restoration by nuclear complementation resulted in a wild-type fatty acid profile [303]. Since *C. reinhardtii* carries a too high level of PUFAs for high-quality biodiesel, controlling the expression of the *fad7* gene, e.g., by MYB_9 or MS_1, is a potential solution to enhance the quality of biodiesel derived from *C. reinhardtii*, as it can help to reduce the level of PUFAS.

The promoters of two other known regulatory genes are also targets of almost all key regulators mentioned : NRR1 and FBP4. MYB_9, LRL1, HMG_6, and CGL86 target the promoter of *NRR1*, and FBP4 is targeted by all except for CGL86. NRR1, a known TF regulator in N deprivation, was identified based on increased expression levels during N starvation and subsequent mutant analysis, which showed a 50 % reduction in TAG accumulation [74]. However, the exact process of how NRR1 influences and regulates lipid metabolism is not known.

FBP4 is a bifunctional kinase/phosphatase enzyme that regulates the metabolism of carbohydrates toward glycolysis or gluconeogenesis [233] (see Table 2.4). Thus, it strongly influences the overall carbohydrate metabolism of the cell. Moreover, together with *TAL1*, it represents one of the central genes upregulated in the *sta6* mutant during N-starvation [101]. Interestingly, *TAL1* is targeted only by MYB9, in contrast to *FBP4*, which is targeted by MYB_9, MS_1, HMG_6, and LRL1. *TAL1* encodes a transaldolase in the pentose phosphate pathway (PPP), which, together with a transketolase, provides a link between the glycolytic and PPP. The PPP produces reduction equivalents in the oxidative phase, which are required for reductive FA biosynthesis. It also yields glyceraldehyde-3-phosphate (G3P), the starting point for DAG or TAG biosynthesis.

All proposed regulator candidates also target the previously mentioned promoters of two lipid degradation enzymes, Cre07.g334150 and Cre10.g425100. As indicated in section 2.2.8, both proteins possess a lipase domain but also DNA-binding domains. GPX5, which is involved in stress response in *C. reinhardtii*, is also a target of all four regulator candidates plus LRL1. MYB_9 and MS_1 also interact with promoters of some unknown genes, e.g., with two diacylglycerol o-acyltransferases (Cre08.g377300 and Cre08.g365950). These genes are predicted to be involved in TAG biosynthesis. At least one gene, Cre08.g365950, displays a substantial increase in gene expression during N and S deprivation in the time course of Ngan et al. [138].

These examples highlight that MYB_9 and MS_1 may play important regulatory roles in lipid and carbohydrate metabolism in *C. reinhardtii*. However, besides the lipid and carbohydrate metabolism, MYB_9, MS_1, and HMG_6 interact with multiple promoters involved in the photosynthesis light reaction of PS I and II, e.g., LHCBM9, PSBP1, PETM, PSBP6, LHCB4, PSAH, PSBX, LHCBM5, FNR1. Photosynthesis is usually downregulated to avoid photooxidative stress during substrate shortage caused by nutrient deficiency. Since MYB_9, MS_1, and HMG_6 appear to be involved in the regulation of photosynthesis-related genes, they may represent the regulatory nodes that link all these processes.

This biological analysis of the targets shows that the proposed regulator candidates are important for the regulation of *C. reinhardtii*. Since all TFs, with the exception of LRL1, are not well characterized, I searched for the orthologs and analyzed the protein domain structure to find further clues to their function.

MYB_9 contains a SANT/MYB domain, and additionally, a domain that is a homolog of yeast SWR1-complex 4 (Swc4/Eaf2), the human DNA methyltransferase 1-associated protein 1 (DMAP1), and the SWC4 protein in *A. thaliana*. SWC4 is one subunit of the SWR1 chromatin remodeling complex (SWR1-C). It is a DNA-binding protein that recruits SWR1-C to the target DNA by interacting with the other subunits of SWR1-C and bridges SWR1-C with the nucleosome [304]. The SWR1-C in *A. thaliana* catalyzes the exchange of H2A-H2B histone dimers with H2A.Z-H2B dimers. This replacement is known to affect the nucleosome stability and chromatin structure. Although it was described that the deposition of H2A.Z in *A. thaliana* genes plays a repressive role in transcription, Sura et al. showed that the enrichment of H2A.Z close to the TSS activates gene expression of multiple genes during stress responses [305]. Thus, SWR1-C can positively and negatively regulate gene expression in *A. thaliana* in an active process. Various studies have shown that it is involved in plant development, DNA damage repair, stress responses, perception of ambient temperature, and flowering time regulation [305, 306, 307, 304, 308, 309]. As MYB_9 is the orthologue of the *A. thaliana* protein SWC4 and interacts with promoters of essential enzymes involved in the remodeled metabolism during N-deprivation stress conditions, I assume it plays an important role as a metabolic regulator. It is likely activating or repressing the gene expression of its targets by regulating the chromatin architecture around the target genes by interacting with other chromatin-remodeling proteins.

Another potential key TF seems to be MS_1, a to date uncharacterized protein. MS_1 was mistakenly annotated as MYB/MYB-related TF in the plant database 'PlnTFDB' [123]. It was, thereby, part of the TF selection for the Y1H screen. A closer analysis of the protein sequence revealed no known DNA-binding domain. Nevertheless, it interacted with 37 promoters in the Y1H screen, demonstrating that it can bind to DNA. It has an annotated carbohydrate-binding-like fold and consists of 761 aa, of which only 105 aa at the N-terminus belong to the carbohydrate-binding domain. It is a homolog of the *A. thaliana* protein OK1, also known as phosphoglucan, water dikinase (PWD). PWD represents one of two important kinases in the starch degradation pathway that phosphorylates starch molecules at the C3 position of the glucose units. Together with α-glucan water dikinase (GWD), the second kinase, it phosphorylates the starch granule surface and stimulates starch degradation via hydrolytic enzymes. This homology emphasizes a role in starch degradation. Since it interacts with many promoters of genes related to lipid and carbohydrate metabolism, it supports the hypothesis that it is involved in regulating the starch and lipid metabolism in *C. reinhardtii*. Lipid and starch metabolism are closely connected and in constant competition for carbon precursors. During nutrient stress, there is a switch from starch to increased lipid synthesis, for which the regulation is still unknown. Since MS_1 can bind to starch, stimulate the degradation mechanism and bind multiple important enzymatic promoters, it could be one regulator that controls this connection.

The third TF that shares 50 % the interaction pattern of LRL1 and represents a potential new regulator of lipid metabolism is HMG_6. It belongs to the high mobility group TR family. For *C. reinhardtii*, very little is known about this group of transcriptional regulators. They are regulators that regulate chromatin dynamics in eukaryotes. They are not thought to regulate by intrinsic transcriptional activity but by altering chromatin structure to allow TFs to bind to DNA [310]. Interestingly, in this study, HMG_6 interacted directly with 24 promoters, all enzymes' promoters. The enrichment analysis of the interaction profile revealed significant enrichment for promoters of the lipid and carbohydrate metabolism and for nitrogen and amino acid metabolism ($p < 0.01$). Other HMG proteins in the screen interacted only with one to three promoters. Their interacting behavior matches the reported abilities of HMG regulators. HMG_6 behaved differently than the other HMG proteins in the screen and different than expected. This indicates that HMG_6 has specific properties and may be a regulator for the interconnection of the lipid, carbohydrate and nitrogen metabolism.

The FHA TF CGL86 also belongs to the highly connected TFs, which share 50 % interactions with LRL1. CGL86 is an ortholog to the nuclear inhibitor of protein phosphatase-1 (NIPP1), which is expressed in both plants and animals [311, 312, 313]. It targets 60 promoters in the Y1H screen, of which 30 belong to promoters of other TFs, 28 to promoters of enzymes, and two belong to the previously described promoters, Cre07.g334150 and Cre10.g425100, that are annotated to have regulatory abilities as well as enzymatic activity. The FHA TF CGL86 was proposed by Lopez Garcia de Lomana et al. to be one transcriptional regulator during N starvation according to their gene expression profile described in their transcriptional regulatory network (TRN) [87]. They categorized it into the mid-stage response, representing the time period of 24 to 60 minutes after applying nutrient stress to the cells, named the 'metabolic state transition'. In this cluster, they categorized many differential regulated transcripts that allocate to proteins that scavenge N from amino acids, nucleotides, and polyamines. Their described mid-stage serves as an adaptation process to mitigate starvation using internal N resources for cellular homeostasis. CGL86 interacts with many promoters of known important TFs, like ROC40 (Cre06.g275350) and NRR1. ROC40 is a TF involved in circadian rhythm regulation and is also known to be involved in the N-deprivation response of *C. reinhardtii* [44, 72]. The *C.reinhardtii roc40* mutant accumulates less neutral lipids than the compared wild-type strain. As CGL86 was suggested before as a regulator during N starvation, this study's findings emphasize the hypothesis that it is a key regulator in stress response to N deprivation. Interestingly, there was no significant enrichment detectable for lipid or carbohydrate metabolism promoters, only a significant enrichment for promoters of photosynthesis.

All of this evidence suggests that the four TFs presented are essential in regulating *C. reinhardtii*. MYB_9, MS_1, and HMG_6 seem to have a significant role in regulating the adaptation of lipid metabolism and the associated metabolic pathways. CGL86 instead seems to act as a key regulator at a higher level of transcriptional regulation,

e.g., by controlling gene expression of other TFs, which then regulate downstream enzymatic genes. However, without *in vivo* experiments, there is no evidence to what extent or in which conditions the candidates regulate lipid metabolism or other related pathways. Therefore, validation experiments are currently being conducted to elucidate their biological function. This first experimental-based GRN for *C. reinhardtii* provides a valuable resource for selecting candidates for further studies and gives an overview of the transcriptional connectivity and regulation in *C. reinhardtii*.

3.4 Concluding remarks

A deeper understanding of the metabolic processes in microalgae is required to exploit their full biotechnological potential, e.g., as a renewable energy resource. Since *C. reinhardtii* is the model organism for microalgae, it was the logical choice for this study.

Metabolic processes are mediated and controlled by PPIs or PDIs that regulate the transcriptional expression of target genes. This work has generated multiple resources allowing a systems-level analysis of the *C. reinhardtii* metabolism. The first genome-scale ORFeome for *C. reinhardtii*, generated as one goal of this work, will enable proteomic studies and help to unravel fundamental biological questions. It represents a powerful resource for further high-throughput functional assays.

Another resource compiled from the ORFeome was a *C. reinhardtii* TF collection consisting of 241 *C. reinhardtii* TFs. A TF collection allows the investigation of regulatory relationships in *C. reinhardtii*. Investigating how transcriptional regulation is structured and organized in *C. reinhardtii* and how *C. reinhardtii*'s TFs are interconnected was the second goal of this work. I also intended to find new key regulators controlling primary metabolism, particularly lipid metabolism in *C. reinhardtii*, during its stress response to nutrient deficiencies. For this purpose, I generated the first experimental-based gene regulatory network for *C. reinhardtii*.

The network exhibits a clear hierarchy, as expected for a unicellular organism [272, 273, 314, 315, 274]. Furthermore, the network is highly interconnected and features a middle regulatory layer of TFs that receives input from other TFs and mediates a subsequent cellular response. The regulation of the metabolic pathways differs from the master-regulatory and pathway-specific regulation system described for unicellular organisms. Over 72 % of TFs, including members of almost all TF and TR families, interact with promoters of various metabolic processes and 84.5 % of all metabolic promoters are targeted by more than two TFs. Only the TRAF TR family interacted predominantly and significantly enriched with promoters of photosynthesis (see Figure A.4 in the appendix).

The results propose a distributed regulatory system hypothesis for the unicellular alga *C. reinhardtii*, which is similarly described for multicellular organisms such as *A. thaliana* [176, 170, 280]. It allows a more precise, fast, and flexible adaptation to external influences. Lipid metabolism in *C. reinhardtii* is complex and linked to many other metabolic pathways. From previously reported studies, it is known that the stress response of *C. reinhardtii* to nutrient deficiency is accompanied by considerable metabolism remodeling. Processes such as photosynthesis, starch, and carbohydrate metabolism are significantly altered. Key regulators that mediate this remodeling as a stress response must co-regulate all these processes to ensure proper functioning and cell survival. Based on my network, I identified multiple potent key regulator candidates of the primary

metabolism: MYB_9, MS_1, and HMG_6. These regulators interact with promoters of different metabolic pathways. All target genes of these TFs exhibit changes in their gene expression upon nutrient starvation. Several of the found and described interaction partners of MYB_9, MS_1, and HMG_6 represent essential enzymes in their respective metabolic pathways. Consequently, these three TFs represent promising candidates that are thought to mediate the different cellular responses to stress in lipid metabolism and related processes. Further experiments are needed to prove their regulatory role in the altered stress-induced lipid metabolism. CGL86, another identified and previously predicted key regulator during N starvation, seems to regulate the primary metabolism from an upper transcriptional level by interacting with many other TFs hierarchically below it.

The GRN provides an excellent starting point for further analysis. To overcome certain limitations of the Y1H method, the PDI data could be merged with data from other methods, such as transcriptomic co-expression data or further physical PDI determination methods, to gain more insights from the network or to refine it.

However, the network substantially improves our understanding of transcriptional regulation in *C. reinhardtii*, although it remains to be completed as it reflects about 34 % of the possible interactions. It provides in-depth insights into the global structure and cross talk of transcriptional regulation in microalgae and provides regulatory context information that raises ideas and candidates for further research. Furthermore, the GRN can pave the way for future *in vivo* biotechnological experiments to manipulate the primary metabolism in *C. reinhardtii*.

4 Material and Methods

This chapter gives an overview of all used materials and describes the used methods and protocols.

4.1 Material

4.1.1 *C.reinhardtii* strain

The RNA and genomic DNA extraction samples were derived from the *C. reinhardtii* strain cc-1883 cw15 NIT+ mt-. The Chlamydomonas reinhardtii culture was grown on a shaker in 20 ml TAP media (recipe according to ḧttps://www.chlamycollection.org/, [316]) at 22°C and constant light conditions (300-350 µMol $m^{-2\ -1}$) and under N- and S-deprivation conditions. The *C. reinhardtii* RNA and genomic DNA were kindly provided by the research group of Prof. Dr. Olaf Kruse at the University of Bielefeld. Dr. Lutz Wobbe extracted the RNA. Dr. Polina Dementyeva extracted the genomic *C. reinhardtii* DNA.

4.1.2 Bacterial strains

DH5α *E.coli* cells were used for bacterial transformations of cloned constructs, described in section 4.2.12. DB3.1 *E.coli* cells propagated empty Gateway® plasmids containing the cell suicide gene *ccdB*. Rosetta^{TM} (DE3) competent cells were used for bacterial transformation of TF-expression constructs. The genotype is the following: F$^-$ *ompT hsd*S_B(r_B^- m_B^-) *gal dcm* (DE3) pRARE (Cam^R).

4.1.3 Plasmids

The following Gateway® plasmids were used to clone the *C. reinhardtii* ORFs for the ORFeome collection: the Gateway® pDONR plasmid pDONR223. The Gateway® plasmids pDONR P4-P1R, pDEST AD 2µ (adapted from pDEST22 (Invitrogen)) and pMW#2 [168] were used to clone the promoters and TFs for the Y1H screen. The pGEX-6P-1 plasmid was modified by introducing Gateway® cloning sites by Prof. Pascal Falter-Braun. This modified plasmid, hereafter referred to as pGEX6, was used to clone several TF

from entry clones into this expression plasmid for protein expression in E.coli. Further details are explained in section 4.2.21.

4.1.4 Yeast strains

For the Y1H assay, two *S.cerevisiae* strains, Yα1867 and Y1H-aS2, were used. All Gal4-AD-TF constructs were transformed into Yα1867, referred to prey yeast strain (Matα SUC2, gal2, mal, mel, flo1, flo8-1, hap1, ho, bio1, bio6, ura3-52, ade2-101, trp1-901, his3-Δ200). The DNA baits were genome integrated into Y1H-aS2, the DNA bait yeast strains. Y1H-aS2 originated from the *S.cerevisiae* strain YM4271 (MATa, ura3-52, his3-Δ200, ade2-101, ade5, lys2-801, leu2-3, 112, trp1-901, tyr1-501, gal4Δ, gal80Δ, ade5::hisG) [192] and was modified by change of the *his3-Δ200* locus with *his3-Δ1* (190 bp deletion) [193].

4.1.5 Antibiotics

The antibiotics listed in Table 4.1 were used to select correct clones on agar-agar media plates.

Table 4.1: Antibiotics

Antibiotics	Supplier Company	Cat.number
Ampicillin sodium salt	Roth	K029
Carbenicillin disodium	Roth	6344
Gentamycin sulfate	Duchefa	G0124
Kanamycin sulfate	Carl Roth	T832
Spectinomycin HCl pentahydrate	Duchefa	SO188
Chloramphenicol	Carl Roth	3886.2

4.1.6 Chemicals

The chemicals listed in Table 4.2 were used to prepare media. A detailed description of the amounts and concentrations of the different chemicals used is explained in the corresponding sections.

Table 4.2: Chemicals, reagents, kits for media and buffer preparation and common applications

Name	Supplier Company	Cat.number
Yeast extract	Carl Roth	2363.5
Tryptone/Peptone	Carl Roth	8952.5
Sodium chloride (NaCl)	Carl Roth	9265.3

Potassium chloride (KCl)	Carl Roth	6781.1
Magnesium sulfate ($MgSO_4$)	Carl Roth	682.1
Magnesium chloride ($MgCl_2$)	Carl Roth	KK36.3
Ethanol	HMGU Lagermaterial	5000006
Agarose	VWR-Peqlab	732-2789
Agar Agar	Applichem	A0949, 5000
Glycerol	Carl Roth	3783.2
Dipotassium hydrogenphosphate (K_2HPO_4)	Carl Roth	P749.2
Potassium dihydrogenphosphate (KH_2PO_4)	Carl Roth	3904.1
Glucose	Carl Roth	X997.5
Sodium hydroxide (NaOH)	Carl Roth	6771.1
Isopropyl-ß-D-thiogalactopyranosid (IPTG)	PanReac AppliChem	A1008,0005
Coomassie Blue R-250	Santa Cruz	sc-24972
SERVA Tris-Glycine/SDS sample buffer (2x)	SERVA ELECTROPORESIS GmbH	42527.01
Glycine	Sigma-Aldrich Chemie GmbH	G7126-500G
Tris base	Carl Roth	4855.2
Sodium dodecylsulfate (SDS) solution (10 %)	Applichem	A0676
30 % NF-Acylamide/Bis	Carl Roth	A124.1
TEMED	Carl Roth	2367.1
10 % APS	Carl Roth	9592.1
Methanol	Carl Roth	4627.6
Tween-20	Merck Millipore	655205-250ML
SUPERSIGNAL® WEST FEMTO	VWR	PIER34095
Midori Green Advance	Biozym Scientific GmbH	617004
DEOXYNUCLEOTIDE SET, 100 mM	Sigma-Aldrich Chemie GmbH	DNTP100-1KT
Nucleospin Plasmid No Lid	Macherey - Nagel	5001251
Gateway® LR Clonase II Mix	Life Technologies GmbH	11791100
Gateway® BP Clonase II Enzyme Mix	Life Technologies GmbH	11789100
GeneRuler 1kb DNA Ladder	Life Technologies GmbH	SM0312
KOD HOT START DNA POLYMERASE	Merck Chemicals GmbH	71086-4
Dream Taq Green DNA Polymerase	Thermo Fisher Scientific GmbH	5001171
Quant-it™ PicoGreen™	Thermo Fisher Scientific GmbH	P7589

4.1.7 Yeast media

4.1.7.1 YEPD

The YEPD media contains 2 % tryptone/peptone and 1 % yeast extract. The pH is adjusted to 5.9 with acidic acid or sodium hydroxide. After autoclaving the media, 50 ml 40 % glucose and 15 ml 65 mM adenine are added. For solid YEPD plates, 20 g of agar-agar is added before autoclaving the medium. The YEPD plates must dry at RT for one week before yeast can be spotted onto them. This prevents the spotted cultures from running into each other.

4.1.7.2 SC-media

SC media is prepared by mixing 2.6 g amino acid mix (contains of 6.0 g of each of the following aminoacids, all are purchased by Sigma-Aldrich Chemie GmbH: alanine (A7627), arginine (A5006), aspartic acid (A9256), asparagine (A8381), cysteine (C7352), glutamic acid (G5889), glycine (G7126), isoleucine (I2752), lysine (L5626), methionine (M9625), phenylalanine (P2126), proline (P0380), serine (S4500), threonine (T8625), tyrosine (T3754), and valine (V0500)). 5.0 g ammonium sulfate (Carl Roth, 9212.1) and 6.7 g yeast nitrobase mix (Sigma-Aldrich Chemie GmbH, Y0626) is also added. Everything is solved in 950 ml milli-Q (mQ) H_2O. The pH is adjusted to 5.9 with sodium hydroxide and hydrochloric acid. Afterward the media is autoclaved. For solid plates 20 g agar agar is added before the autoclaving process. The media can be stored up to six month at RT and solid plates were stored at 4°C for up to three month.
If the selective amino acid uracil (U0750) is needed in the media, 6.0 g of uracil is added to the amino acid mix.
Stock solutions are prepared for the selective amino acids tryptophan (T0254), histidine (H8125), leucine (I2752) and adenine. They are dissolved at 65 °C for 30 minutes and then sterile filtered using a 0.2 µM filter.

Table 4.3: Stock solutions for yeast media

Solution	concentration
tryptophan solution	40 mM
histidin solution	100 mM
leucine solution	100 mM
adenine solution	65 mM

Before using the media, 50 ml 40 % glucose, 15 ml 65 mM adenine stock solution and 8 ml of the respective selective amino acid stock solution were added.

For SC-Trp: add 8 ml Leu and 8 ml His. For SC-His: add 8 ml Trp and 8 ml Leu. For SC-Leu: add 8 ml Trp and 8 ml His. For SC-Trp-His: add only 8 ml Leu. For SC-Leu-Trp: add only 8 ml His.
For plates that contained 3-amino-1,2,4 triazole (3-AT), 4.2 g 3-AT (Sigma, A8056) was solved in 50 ml destilled H_2O to obtain a 1 M stock solution. The solution was sterile filtrated using a 0.2 µM or 0.45 µM filter.
For plates that contained cycloheximide (CHX) (Sigma, C4859), 100 mg CHX were solved in 10 ml 100 % ethanol and stored at the -20 °C.

4.1.8 Bacterial media

4.1.8.1 LB

LB media is prepared by mixing 5 g yeast extract (Carl Roth, 2363.5), 10 g tryptone/peptone (Carl Roth, 8952.5), and 10 g NaCl (Carl Roth, 9265.3). Everything is solved in 950 ml mQ H_2O. The pH is adjusted to 7.0 with sodium hydroxide and hydrochloric acid. Afterward, the media is autoclaved. For solid plates, 10 g agar agar is added to 1 l media before autoclaving. The media can be stored for up to six months at RT, and solid plates were stored at 4°C for up to three months.

4.1.8.2 TB

Terrific broth (TB) media is prepared by mixing 24 g yeast extract (Carl Roth, 2363.5), 12 g tryptone/peptone (Carl Roth, 8952.5), and 40 ml glycerol (Carl Roth, 3783.2). Everything is dissolved in 900 ml mQ H_2O. Usually, the pH is around 7.0 and does not need further adjustment. Afterward, the media is autoclaved. The media can be stored for up to six months at RT. Immediately before usage, 100 ml phosphate buffer is added.
The phosphate buffer (0.017 M) for TB media is prepared by mixing 2.31 g KH_2PO_4 (Carl Roth, 3904.1) and 12.54 g K_2HPO_4 (Carl Roth, P749.2) in 1,000 ml mQ H_2O followed by autoclaving.

4.2 Methods

4.2.1 ORFeome selection and cloning strategy

To clone ORFs for the *C. reinhardtii* ORFeome collection, a subset of 4,903 ORFs was selected. The annotated gene models need to be reliable by two criteria. First, the gene model annotation of the ORF was consistent between the *C. reinhardtii* gene annotation Chlre v4.3 and v5.5. Second, the ORF size did not show a 2-fold difference in *A.thaliana* orthologous. The first *C. reinhardtii* subset of 4,903 ORFs represented functional genes like enzymes, transporters, metabolic polypeptides, kinases, transcription factors, regulatory proteins, and cytoskeletal and motor proteins. These genes were selected with corresponding gene model annotations from KEGG, KOG, pfam, GO, EC, and PANTHER of genome annotation version Chlre v5.5 and v4.3. To extend the first set of *C. reinhardtii* ORFs, all remaining of the 17,741 protein-coding *C. reinhardtii* ORFs were reviewed by my colleague Dr. Yang Jae Kang, and the accuracy of their gene model was examined by using stranded RNA-seq data from Ngan et al. [138]. The Gateway® Cloning strategy (Invitrogen) was used for cloning. It is described in more detail in section 4.2.10.

4.2.2 Definition and consolidation of *C. reinhardtii* TF and TRs

To select transcription factors (TFs) and transcriptional regulators (TR) of *C. reinhardtii* all annotated TF for *C. reinhardtii* from the plant TF database 'PlantTFDB' [214, 215, 216, 217] (PlantTFDB 3.0 and 4.0) were extracted. This refers to 230 TFs (206 loci) divided into 29 families. Their analysis was based on the *C. reinhardtii* v5.5 gene annotation [52]. A second plant TF database 'PlnTFDB' [123, 218] provided 346 TF protein sequences for *C. reinhardtii* TFs and transcriptional regulators, arranged into 35 TF families and 19 groups of transcriptional regulators. A protein was defined as a transcriptional regulator if one or more domains were present, which does not allow their classification into a defined family. This also occurred if a combination of two different TF family motifs were present. The *C. reinhardtii* v4.0 gene annotation was used to annotate the 346 protein models. To determine the overlap and differences between the two databases, the selected protein IDs and sequences of *C. reinhardtii* v4.0 were blasted against the v5.5 annotation [53] to identify the corresponding v5.5 gene IDs. Moreover, all 2296 *A.thaliana* TFs (tair 10, 1,717 loci) were compared (blasted) to the *C. reinhardtii* v5.5 gene annotation to detect orthologous. Further, all corresponding orthologous groups for the *A.thaliana* TFs (2,296) were extracted from Plaza Diacots 4.0 (last modified 2017/11/14) [317] together with the corresponding orthologous group for *C. reinhardtii* and analyzed. By combining all information and deleting the duplicates, a list of 475 TFs remained. All TFs are listed in Table (A.5).

161 TFs could be cloned and extracted from the ORFeome collection. 96 additional ORFs were gene synthesized from the company Twist Bioscience. To that date, Twist

Bioscience was the only company that offered a gene synthesis service for genes with a length of up to 3.2 kb. ORFs with a CDS length smaller than 1.8 kb were synthesized as fragments that needed to be cloned into a vector. ORFs larger than 1.8 kb were delivered as 'clonal genes' already incorporated into a vector, ready to be cloned into the Gateway® system. Many TFs contained GC rich repeats and repetitive sequences, interfering with the gene synthesis. Therefore all coding sequences were codon optimized for yeast and could thus be synthesized.

4.2.3 Promoter selection

The promoters for the Y1H screen were selected according to two categories. First, all promoter regions for the characterized 475 TF and regulators were selected. Second, genes associated in the literature with lipid metabolism, lipid remodeling and metabolism remodeling during nitrogen and sulfur deprivation were chosen. The gene expression profiles of this subset of around 400 genes were analyzed using the RNA-seq data set of Ngan et al. [138]. All candidate genes that showed a log 2-fold change alteration in the time series of the 16 time points were included in the promoter selection. Finally, promoter regions of 139 genes were selected for the cloning pipeline. In the study of Yu CP et al., [219], they examined the positional distribution of TFBS (transcription factor binding sites) in *A. thaliana* and found that 86 % of the TFBS are in the regions from -1,000 bp to +200 bp with respect to the transcription start side. Scranton MA et al. [220] designed synthetic promoters capable of driving robust gene expression in *C. reinhardtii* with a promoter region of -1,000 bp and +50 bp. This study defined DNA sequences -1,000 bp upstream and +100 bp downstream of the annotated transcription start site as promoter region. A total of 614 promoters were chosen for the protein-DNA interaction screen. Primers were designed according to section 4.2.4 by Dr. Chung-Wen Lin and subsequently ordered.

4.2.4 Primer design

To amplify a *C. reinhardtii* ORF or promoter of interest, the following criteria were used to design primers for the PCR. The primer length was set to 15 - 30 bp, and the annealing temperature range to 54 - 63°C. The software Primer3 [318, 319, 320] was adapted by Dr. Yang Jae Kang and Dr. Chung-Wen Lin to design the primers for all ORFs and promoters. The adapted script of Dr. Yang Jae Kang and Dr. Chung-Wen Lin can be found at github: https://github.com/k821209/seq2primer.

Since an accurate gene model is required for successful primer design, my colleague Dr. Yang Jae Kang used paired-end stranded RNA-seq data from Ngan et al. [138] to review the gene models of all 17,741 protein-coding *C. reinhardtii* ORFs. He developed an algorithm using reference-based assembly, RNA-seq read coverage, and Poly-A tail information to score the predicted *C. reinhardtii* CDS and reconstructs the ORF sequences.

A detailed list of all used primers can be found in the appendix in Table A.2. Primers for amplification of *C. reinhardtii* ORFs and *C. reinhardtii* promoters were ordered as mixed primer plates from the company Eurofins Genomics. The concentration of the mixed plates was set to 200 µM from Eurofins Genomics. The specific primers of the *C. reinhardtii* ORFs were sorted by annealing temperature for the first PCR and according to the ORF sizes, so that all ORFs in one plate have similar annealing temperatures and would also require similar elongation time in the PCR. Since the promoter amplicon length was fixed to 1,100 bp, the primer for the amplification of *C. reinhardtii* promoters were sorted only by the annealing temperature.

The used primer for *C. reinhardtii* ORFs were extended by a linker sequence between the start - and end of the ORF and the recombination sites. This extends the sequences by four amino acids. The use of a linker sequence can increase the stability of fusion protein, the flexibility of certain domains caused by more distance between the two fused proteins and therefore can improve the biological activity of the fusion protein and the ability to interact with other proteins more easier [321]. The primers were designed that the stop codon at the end of the *C. reinhardtii* ORF is eliminated during the PCR amplification. This allows to clone the ORFs into destination vectors with an N-terminal and a C-terminal fusion domain. For amplifying *C. reinhardtii* ORFs like TFs, the following linker sequence were added to all specific primers:

Linker adapters forward (ORFs): 5´-GGA-GGC-GGA-GGT-3´

Linker adapters reverse (ORFs): 5´-GGT-TCC-ACC-ACC-ACC-3´

The PCR primers for the first and second PCR are the following:

1st PCR primer forward (ORFs):
5´-CA-GGA-GGC-GGA-GGT-ATG- ORF-specific-sequence-3´
1st PCR primer reverse (ORFs):
5´-GGT-TCC-ACC-ACC-ACC-ORF-specific-sequence-3´

2nd PCR attB1 forward primer:
5´-GGGG-ACA-AGT-TTG-TAC-AAA-AAA-GCA-GGC-TCA-GGAGGCGGAGGT-ATG-3´
2nd PCR attB2 reverse primer:
5´-GGGG-AC-CAC-TTT-GTA-CAA-GAA-AGC-TGG-GTC-GGTTCCACCACCACC-3´

To amplifying *C. reinhardtii* promoters, no additional linker sequence was placed between the promoter sequences and the recombination site. The PCR primers were added to the specific primers:

1st PCR primer forward (promoter):
5´-GTA-TAG-AAA-AGT-TGC-C-3´

1st PCR primer reverse (promoter):
5´-TTT-TTG-TAC-AAA-CTT-GT-3´

2nd PCR attB4 forward primer:
5-GGGG-ACA-ACT-TTG-TAT-AGA-AAA-GTT-GCC-3´
2nd PCR attB1R reverse primer:
5´-GGG-GAC-TGC-TTT-TTT-GTA-CAA-ACT-TGT-3´

4.2.5 RNA extraction

The *C. reinhardtii* RNA and genomic DNA was kindly provided by the the research group of Prof. Dr. Olaf Kruse at the university of Bielefeld. The RNA extraction was done by Dr. Lutz Wobbe. Dr. Polina Dementyeva was extracting the genomic *C. reinhardtii* DNA.

4.2.6 Genomic DNA isolation

As mentioned above, all described steps were performed by Dr. Polina Dementyeva at the university of Bielefeld *C. reinhardtii* cells were grown to a logarithmic phase of OD_{750} 0.6-0.8. The cells were harvested by centrifugation at 3,000 g for 5 minutes and resuspended in 1 ml water afterward. After another centrifugation step at 16,000 g for 5 minutes, the supernatant was discarded. Cells were frozen at -80°C or lysed directly with 350 µl CTAB-buffer (100 mM Tris-HCl, 20 mM EDTA, 1.4 M NaCl, 2 % CTAB, pH 8.0), 50-100 µl proteinase K and 25-50 µl 20 % SDS in a water bath for 2 h at 55°C. Afterward the sample was cooled down on ice, and 50-100 µl of 5 M potassium acetate was added, followed by incubation on ice for 30 minutes. The centrifuged sample (13,000 g, 15 minutes) was mixed with an equal volume of phenol/chloroform/isoamyl alcohol. The aqueous phase was mixed with 1 ml -20°C cold ethanol to precipitate the genomic DNA. The DNA pellet (centrifugation: 16,000 g, 5 minutes) was washed with 500 µl 70 % ethanol, followed by another centrifugation step at 10,000 g for 5 minutes. After the removal of ethanol, the DNA pellet was resuspended in 50 µl TE buffer (25 mM Tris-HCl, 5 mM EDTA, pH 7.5) plus 1 µl RNase (1 µg/µl).

4.2.7 cDNA synthesis for ORF amplification

500 ng total mRNA (in 5 µl) was used in a 25 µl approach for the cDNA synthesis. It was mixed with 0.1 µl 50 µM random primers (Promega, C1181), 0.5 µl (50 µM) oligo$(dT)_{16}$ (Life Technologies, N8080128), and 1.5 µl 10 mM dNTPs (dNTP mix, consists of dATP, dGTP, dCTP, and dTTP) (Roth, K039.1). The following steps were performed in a thermocycler. The reaction mixture is heated for 5 minutes at 70°C followed by 10 minutes at 21°C. A second reaction mixture is prepared by mixing 5 µl 5x reaction buffer (250 mM Tris-HCl (pH 8.3), 375 mM KCl, 15 mM $MgCl^2$), 2.5 µl 100 mM DTT, 5.0 µl 5 M Betaine, 1.25 µl RNasin Ribonuclease Inhibitor (40U/µl) (Promega, N2511), 1.25 µl SuperScript III (SSIII) (200 U/µl) (Thermo Fisher, 18080-x) and 1.25 µl water. Both reaction mixtures were mingled together and the following thermocycler steps were applied to them: 10 minutes at 21°C, 120 minutes at 42°C. After this time additional 250 U of SSIII were added to the reaction mix and the sample was treated for 30 minutes at 55°C

followed by a heating step of 15 minutes at 70°C. The cDNA was stored at -20°C before being used to amplify *C. reinhardtii* ORFs.

4.2.8 Amplification of desired ORF by use of nested PCR

For the amplification of *C. reinhardtii* ORFs via PCR, 10 different DNA polymerases were tested. The best high throughput test result was achieved with the KOD Hot Start DNA polymerase (Merck, 71086). 0.5 U of KOD Hot Start DNA polymerase (1U/µl in 50 mM Tris-HCl, 1 mM DTT, 0.1 mM EDTA, 50 % glycerol, 0.001 % Nonidet P-40, 0.001 % Tween 20, pH 8.0) and 500 ng cDNA as template were used per reaction. 10 mM dNTPs, 25 mM $MgSO_4$ and 10x KOD reaction buffer were added according to manufacturing recommendations. Due to the high C+G content in the coding sequence of *C. reinhardtii* additional substances needed to be added to the recommended protocol for the KOD Hot Start DNA polymerase. DMSO to a final concentration of 5 % together with 3 M betaine (final concentration of 1.2 M) were added to the reaction mix. The specific primers for the PCR were diluted to a concentration of 10 µM and 2 µl were added to the first PCR [final concentration: 0.66 µM].

The amplicon of the first PCR was used as template (2µl of the reaction mix) for the second PCR. The Gateway® recombination cloning sites were added onto the first PCR product in the second PCR. A primer end concentration of 0.2 mM was used. All other additives described for the first PCR were also added in the second PCR. For amplification fragments larger than 2,000 bp, the $MgSO_4$ concentration was adapted to 1.5-1.75 mM. The used thermocycler program is based on the KOD manual and differs according to the annealing temperatures of the primer of the first PCR. The annealing temperature for the first PCR primers varies from 54°C - 63°C. Table A.2 lists all primer sequences for the cloned ORFs and promoters.

PCRs were performed in 96-well microtiter plates, and if the PCR success rate of that 96 ORFs was lower than 40 % or if it was apparent that the PCR failed, the whole PCR was repeated. The vector DNA was isolated afterward as described in section 4.2.12.

For the second PCR, the primer annealing Tm is 54°C. The elongation time depends on the size of the amplicon. All *C. reinhardtii* ORFs were devided in five sets. As described in 4.4, the elongation time depends on the ORF size. All ORFs of one category, e.g., P1, were amplified with the same elongation time, e.g., 30 seconds. All PCR amplicons of the PCR were analyzed by gel electrophoresis. 3 µl of 6x loading dye were added to 5 µl of the finished reaction and loaded onto agarose gel containing Midori Green (Biozym, 617004) for analysis.

Table 4.4: Elongation time according to ORF size

	P1	**P2**	**P4**	**P6**	**P9**
ORF size	$<$ 1.0kb	1.0-2.5 kb	2.5-4kb	4.0-6.0 kb	$>$ 6.0kb
annealing time	20 sec	30 sec	30 sec	30 sec	30 sec
elongation time	0.5 min	1.5 min	3 min	6 min	10 min

4.2.9 Amplicon purification with magnetic beads

This PCR purification was done for the extended ORF set (10,667 ORFs) and the promoter cloning. PCR amplicons of the 2nd PCR were purified before they were cloned into a destination vector. This decreased undesired cloning products and increased cloning efficiency. The purification procedure was performed manually or with the help of the Tecan liquid handling robot called Fluent Automation Workstation. The size selection purification of the PCR product was done with magnetic beads (Steinbrenner, MDKT00010075) and accordingly to the Steinbrenner manual for DNA purification. Therefore, a specific ratio of PCR-magnetic beads was used to exclude DNA particles and primer from the PCR mix and to exchange the reaction mix with Tris-HCl, pH 8.0. In Table 4.5, three different PCR amplicon-magnetic bead ratios are further described.

Table 4.5: Ratio between PCR fragment size and amount of magnetic beads

Ratio	PCR-fragment size	magnetic beads
1 : 1	0 - 300 bp	25 µl
1 : 0.8	301 - 1000 bp	20 µl
1 : 0.6	from 1000 bp on	15 µl

The larger the PCR amplicon, the fewer magnetic beads are needed for the size exclusion purification. The PCR mix and the appropriate magnetic beads ratio are mixed. The magnetic beads bind the DNA amplicon during incubation of 5 minutes at room temperature. After another incubation step on the magnetic separation plate of 5 minutes, the supernatant was removed, and the beads were washed twice with 100 µl freshly prepared 80 % ethanol. The ethanol was removed by pipetting and completely evaporated by drying at the sterile bench. To eluate the DNA amplicons of the magnetic beads, the beads were resuspended in 5 mM Tris-HCl, pH 8.5 (25-40 µl according to starting material of the PCR and handling procedure. Purification manually: 25-30 µl, purification on the Tecan liquid handling robot: 40 µl). The elution mix is placed back into a magnetic separation plate to separate the eluted DNA amplicons from the magnetic beads. The bead-free eluate is transferred after an incubation of 5 minutes to a fresh reaction tube or reaction plate.

4.2.10 Gateway® cloning

For cloning, the established and commercially available Gateway® Cloning System from Thermo Fisher Scientific was used. In the first cloning step, 3 µl PCR product was mixed with 1 µl pDONR plasmid (50 ng/µl), 0.5 µl TE buffer (10 mM Tris pH 8.0, 1 mM EDTA) and 0.5 µl BP II clonase mix (Thermo Fisher Scientific, 11789) to perform the BP reaction. Table 4.6 describes which pDONR plasmid was used for cloning of *C. reinhardtii* promoter and *C. reinhardtii* ORFs. The BP reaction was performed in a 96-well microtiter plate by incubation overnight at 25°C in a thermocycler.
2-2.5 µl of the BP reaction were transformed into high chemical competent *E.coli* cells as described in section 4.2.12.
For the LR reaction 2.5 µl of plasmid DNA containing the entry clone (35-100 ng/µl) was mixed with 1 µl of pDEST (150 ng/µl), 0.5 TE buffer (10 mM Tris pH 8.0, 1 mM EDTA) and 0.5 µl LR II

Table 4.6: Plasmids used in the Gateway® Cloning System

BP		
	cloning	**pDONR plasmid**
	promoter	pDONR P4-P1R
	ORF/TFs	pDONR223
	ORF	pDONR223
	protein expression	pDONR 223/pDONR Twist
LR		
	cloning	**pDEST plasmid**
	promoter	pMW#2
	ORF/TFs	pDEST AD 2µ
	ORF	pDEST AD/pDEST DB
	TF	pGex6

clonase mix (Thermo Fisher Scientific, 11791), followed by a 4-16 hours incubation at 25°C in a thermocycler. 2-2.5 µl of LR reaction were transformed into high chemical competent *E.coli* cells as described in section 4.2.12. For all cloning reactions, different controls were performed to detect contamination and destination vectors without functional selection markers. As a negative control the BP reaction master mix was tested without a PCR amplicon. Instead of 3 µl PCR mix, 3 µl of TE buffer was added to the reaction mix. This control contains the empty pDONR or pDEST vector, which should not show bacterial growth after the transformation because the selection gene ccdB should kill the cells.

4.2.11 Gateway® cloning into BFG-Y2H plasmids

All entry clones of successfully amplified 2_{nd} PCR amplicons were consolidated and freshly spotted on agar plates, in total 8627 ORFs. Afterward, the spotted bacteria spots of 12 x 96 ORFs were pooled into one bacterial culture pool, and the DNA was subsequently isolated. By this procedure, eight DNA pools were created, containing roughly 1,152 ORFs. Afterward, an en-masse LR Gateway® reaction was performed. As destination vectors, two different vector pools were used:

- AD-destination BC plasmid pool
- DB-destination BC plasmid pool

All vector plasmids were previously generated and reported by Yachie et al. [156]. Each of the destination vectors contains two different barcodes (BC) and parts of an ARS along with part of a centromere sequence (CEN). These sequences enable the plasmid to replicate itself independently in the yeast cells and are considered as low copy vectors. All BFG-Y2H plasmid are CEN-vectors. A different type of yeast vectors are yeast episomal plasmids. They contain a fragment from the 2-micron circle that allows a plasmid copy number of more than 50 per cell, thereby considered a high copy number. These vectors are hereafter referred to as 2µ-vectors.

This study used them for the TF transformation for the Y1H screen.

After the en-masse LR reaction with the AD - and DB-destination BC plasmid pools, the LR reaction was transformed into bacteria and plated on LB-Carb agar plates. After one day of growth, for each DNA pool, around 2500 single colonies (26 x 96-well plates) were picked with the QPix^{TM} 420 Colony Picker from Molecular Devices. In total, 416 plates (218 AD-containing plasmids, 216 DB-containing plasmids) per DNA pool. Glycerol stock plates were produced for all plates and frozen at -80°. The following BFG-Y2H screen still needs to be performed.

4.2.12 Bacterial transformation and plasmid extraction

The cloned constructs were transformed into competent *E. coli* cells to get the plasmid constructs amplified from the bacterial cells. The transformation protocol was adapted from a single reaction tube to a 96-well microtiter plate. The frozen competent cells were thawed on ice. 20 µl of competent cells (>1x10^8 cfu/µg) were used per well. If the transformation efficiency of the competent cells was less than 1x10^8 cfu/µg 30 µl of competent cells were used. For the transformation of BP or LR reaction, 2.5 µl of the reaction mixture was added to each well using an 8-channel pipette or the Integra liquid pipette VIAFLO 96. The microtiter plate was sealed and incubated for 30 minutes on ice, followed by a heat shock of 30 seconds (max. 60 seconds) at 42°C in a thermocycler or water bath, plus a 2 minutes incubation on ice. To recover the cells, 100 µl of SOB media (0.5 % yeast extract, 2 % tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$, 10 mM $MgCl_2$) were added to each well, followed by incubation at 37°C for 45-60 minutes without shaking. Afterward, the cell solution was transferred into a prepared deep-well microtiter plate, filled with 1.7 ml TB media (2.4 % yeast extract, 1.2 % tryptone, 4 % glycerol, 10 % phosphate buffer (0.017 M KH_2PO_4, 0.017 M K_2HPO_4)) or 1.7 ml LB media (0.5 % yeast extract, 1 % tryptone, 1 % NaCl) with the respective antibiotics and incubated for 16 - 18 hours on a plate shaker at 37°C.

4.2.13 Determination of the presence of insert DNA in plasmid constructs via colony PCR

To verify that a particular cloned plasmid construct was incorporated into *E.coli* cells, a colony PCR was performed. Therefore, one colony of transformed *E.coli* was resuspended in 15 µl mq H_2O or directly with the PCR mix. If the colony was resuspended in water, 3 µl of this mix was used for the subsequent colony PCR. 5 µl of the resuspended sample was spotted on the respective selection plate to maintain the clone for further experiments. The 30 µl PCR reaction containing 3 µl sample, 4 µl H_2O, 3 µl dNTPs (2.5 mM each), 3 µl DreamTaq buffer (20 mM $MgCl_2$), 1.5 µl forward primer (10 µM), 1.5 µl reverse primer (10 µM), 1.5 µl DMSO, 12 µl betaine (3M), 0.5 µl DreamTaq were subjected to the following thermal cycles: 10 minutes at 98°C, followed by 35 cycles of step 1: 95°C for 30 seconds, step 2: 56°C for 30 seconds, step 3: 72°C elongation time depending to the ORF size, see 4.4) and finally 5 minutes at 72°C. 3 µl of 6x loading dye were added to 5 µl of the finished reaction and loaded onto agarose gel containing Midori Green for analysis.

4.2.14 Glycerol stocks and plasmid extraction

For long-term storage the bacterial cultures were frozen in glycerol stocks. After the growth of 16 - 18 hours 75 µl/well of the bacterial culture is transferred into a skirted 96-well u-bottom microtiter plate and gently mixed with 75 µl/well 40 % glycerol. The glycerol stocks are kept at -80 °C. The remaining over-night culture was pelleted by centrifugation at 1000xg for 10 minutes. The supernatant was aspirated with a 8-channel pipet or a liquid handling robot. The remaining cell pellet was used directly for plasmid isolation or frozen at -20°C for later usage. For plasmid isolation commercially available plasmid purification kits were used from Macherey-Nagel for single tube plasmid purification and kits from Qiagen (962241) or 7Bioscience (EM21-BR, custom-made product for us) were used with the Qiagen liquid handling BioRobot according to manufacturer´s recommendations.

4.2.15 ORF sequencing

The first set of functional ORFs was verified by next-generation sequencing. To verify the full-length ORF sequence and not the vector backbone, a PCR onto the isolated plasmid DNA was performed to amplify the ORF sequence. 4062 ORFs were sent for sequencing to the company seqWell. To that date, seqWell provided a service to serve full-length next-generation sequencing (NGS) for plasmids. With NGS, a vast number of short sequences can be sequenced in one reaction. For this, the input DNA is randomly cleaved into short fragments, followed by 5´and 3´adapter ligation. For Illumina sequencing, the cleaved fragments are between 100 - 150 bp long. Afterward, these fragments are loaded into a flow cell on surface-bound oligos complementary to the library adapters. The fragments are then amplified to clonal clusters and sequenced afterward. To provide the full-length information of an ORF with more than 150 bp, seqWell performed a tagmentation reaction during the library preparation of all samples. During the tagmentation reaction, DNA samples get fragmented, and two unique barcodes with Illumina NGS linkers are added to these DNA fragments. The barcodes are unique for each well position and each 96-well microtiter plate. As the cost for a single sample is high, three ORF PCR samples were pooled in one sequencing sample. 318 ORF PCR samples were pooled 4:1 instead of 3:1. By this pooling, the sequencing sample size could be decreased from 4062 to 1328 samples. Due to the pooling of three ORF PCRs in one well, these three amplicons carried the identical two unique barcodes for the well and plate position. As the three ORFs had different coding sequences and the tagmentation occurs randomly, the short sequencing fragments could be aligned afterward to the *C. reinhardtii* reference genome. The data analysis and the alignment of paired-end reads of the 4062 ORFs were performed by Dr. Yang-Jae Kang. A minimum of 10 paired-end reads was determined as the threshold for a sequenced ORF. 96 ORFs were only Sanger sequenced.

4.2.16 Yeast transformation

The Y1H assay is based on DNA promoter bait constructs and TF-AD constructs. The TF-AD constructs are transformed into the Y1H-S2a strain. The promoter bait constructs were integrated into the genome of the Yα1867. Both strains are further described in 4.1.4. The transformation and integration were performed according to previously reported protocols [322, 192, 191]. The linearization of the plasmid construct is necessary to integrate the DNA bait constructs.

Otherwise they will not be integrated into the yeast genome. The plasmid constructs were linearized by the restriction enzymes AflII, XhoI, and SbfI. The transformation of the TF-AD constructs and the integration of the DNA promoter bait-reporter constructs into opposite mating types of yeast enables yeast mating and reduces the workload compared to co-transformation protocols.

The yeast strains were streaked out on separate 145 mM YEPD plates (see 4.1.7) and incubated at 30°C for 48-72 hours to obtain isolated colonies. One day before the transformation, a preculture was inoculated. Therefore 20-50 ml of YEPD media (2 % bacto peptone, 1 % yeast extract, 2 % glucose) was inoculated with 5-10 single colonies of the respective yeast strain and incubated for 14-18 hours at 30°C shaking (150-180 rpm). 200 ml YEPD media per 96-well microtiter plate that should be transformed is inoculated with the preculture to an OD_{600} of 0.1. When the OD_{600} of the culture reached 0.4-0.6, the culture was harvested by centrifugation (5 minutes, 800 g, RT), the supernatant was discarded, and the cell pellet was gently resuspended in 40 ml H_2O. After another centrifugation step (5 minutes, 800g, RT), the cell pellet was resuspended in 10 ml TE/LiAc buffer (10 mM Tris, pH 8.0, 0.5 mM EDTA, 100 mM LiAc), followed by a third centrifugation step. The supernatant was discarded again, and the pellet was gently resuspended in 2 ml TE/LiAc buffer (see above). 200 mg of single-stranded carrier DNA (Sigma-Adrich Chemie GmbH, D1626) (was boiled before for 5 minutes and cooled down on ice) together with 10 ml TE/LiAc/PEG solution (10 mM Tris, pH 8.0, 0.5 mM EDTA, 100 mM LiAc, 80 % of 44 % PEG stock solution) were added to cell solution. 120 µl of this cell suspension was then aliquoted into each well a 96-well microtiter plate using an 8-channel pipette. This solution now carried the competent yeast cells. 0.5-1 µg of plasmid DNA (10-15 µl) per well were added to yeast cell suspension with an 8-channel pipette or with the Integra liquid pipette VIAFLO 96. The cell-DNA mix was then incubated for 30 minutes at 30°C without shaking, followed by a 15 minutes heat shock at 42°C. The plate was centrifuged for 5 minutes at 800g, and the supernatant was discarded. The cells were washed in 100 µl sterile mQ H_2O, followed by another centrifugation for 5 minutes and 800 g. If the yeast was transformed with a TF-AD plasmid construct, the transformed cells were resuspended in 15 µl sterile mQ H_2O and then spotted on a respective selection plate (SC-Trp (see 4.1.7)). If the yeast was transformed with a DNA promoter bait construct, the transformed cells were resuspended in 50-100 µl of sterile mQ H_2O and then streaked out on a selection plate (SC-His (see 4.1.7)) to obtain single colonies after the integration transformation. The plates were incubated for 48-72 hours at 30°C. The TF-AD yeast transformants were inoculated afterward into a u-bottom cikroplate containing 170 µl SC-Trp media and incubated on a vibration plate shaker for 72 hours at 30°C, following the preparation of glycerol stocks. Therefore, 75 µl of grown cells were gently mixed with 40 % glycerol, sealed, and stored at -80°C. Usually, two glycerol copies were prepared, one for common use and one copy as archival stock. For the integrated DNA baits, 8-12 single colonies were picked and inoculated in SC-His media to prepare glycerol stocks and perform a self-activation test (described below 4.2.18).

4.2.17 Yeast lysis and subsequent colony PCR of lysed yeast cells

To examine if a transformation of a cloned plasmid was transformed or integrated into yeast, a yeast lysis with subsequent colony PCR of the lysed yeast cells need to be done. Therefore

two approached were performed depending on the availability of reagents. One approach was the yeast lysis with Zymolyase (amsbio, 120491-1). The other approach was the yeast lysis with sodium hydroxide. For the lysis with Zymolase, 2.5 mg of Zymolase (2500 U/g) were solved in 1 ml 0.1 M phosphate buffer (solve 5 ml 10M stock solutions KH_2PO_4 and 5 ml K_2HPO_4 in 500 ml H_2O, adjust pH to 7.4). This solution can be kept for maximal two weeks. One colony was resuspend in 15 µl of zymolase solution and subjected the following thermocycler program: 45-60 minutes at 37°C, followed by 15-20 minutes at 95°C. 85 µl mq H_2O was added, followed by a centrifugation step at 800g for 10 minutes. 3 µl of lysed sample was used for the subsequent colony PCR.

The second approach was performed when the zymolase enzym has lost their activity and was not functional. Therefore a small yeast cell pellet (grown in respective liquid media) were washed twice with 100 µl H_2O to remove media components. In case of yeast colonies grown on a solid plate, one colony was used for the lysis, resuspended in 25 µl H_2O. Afterwards the cell pellet was solved in 25 µl mq H_2O. 25 µl of 40 mM sodium hydroxide solution was added, followed by a 30 minutes treatment at 95°C in a thermocycler, followed by a centrifugation at 1000g for 5 minutes. 3 µl of cell lysate was used for the subsequent colony PCR.

The colony PCR was performed as follows. The 30 µl PCR reaction containing 3µl sample, 3.2 µl H_2O, 3 µl dNTPs (2.5 mM each), 3 µl 10x KOD Hot Start buffer, 1.8 µl $MgSO_4$ (25 mM), 1 µl forward primer (10 µM), 1 µl reverse primer (10 µM), 1.5 µl DMSO, 12 µl betain (3M), 0.5 µl KOD Hot Start polymerase were subjected to the following thermal cycles: 10 minutes at 98°C, followed by 35 cylces of step 1: 95°C for 20-30 seconds (according to Table 4.4), step 2: respective annealing temperatur depending on the used primers for 30 seconds, step 3: 70°C elongation time depending on the ORF size, see Table 4.4) and finally 5 minutes at 72°C. 3 µl of 6x loading dye were added to 5 µl of the finished reaction and loaded onto agarose gel containing Midori Green for analysis.

4.2.18 Self-activation test of yeast bait strains

After the yeast transformation of the DNA promoter bait constructs, the strains must be tested for their self-activation. Every incorporated plasmid construct has a minor background activity because of the minimal promoter in the plasmid reporter construct. This activity needs to be determined and the yeast clone with the least self-activation is chosen for the actual Y1H screen afterward.

For each yeast DNA promoter strain, eight single colonies (if available) were picked and grown in 150 µl SC-His media for 24 hours. Afterward, the yeast cultures were diluted 1:4 (120µl media + 40 µl grown culture), and 5 µl of the dilution was spotted on SC-His plates containing 10, 20, and 40 mM 3-AT. The dilution was done on cell cultures with an OD_{600} of 0.65-0.9, meaning that cells with an OD_{600} of around 0.2 were spotted onto the selection plates were spotted. After two days of inoculation at 30°C, pictures of the plates were taken and analyzed. The yeast clone with the least growth on the plate with the lowest 3-AT concentration was picked and grown again in 150 µl SC-His media to prepare glycerol stocks (described in section 4.2.16).

4.2.19 Auto-activation test for TF prey strain

Since TFs can also bind nonspecifically to vector DNA, an auto-activation screen is performed with all TF prey strains. Therefore, all TF prey strains are mated with a yeast strain that contains an integrated pMW#2 vector without any additional promoter sequence. This pMW#2 vector is called 'an empty pMW#2 vector'. This assay can detect nonspecific binding to the pMW#2 vector. If the mated yeast can grow after selection, the studied TF binds unspecifically to the vector backbone and is therefore excluded for the following Y1H screen.

4.2.20 Y1H pipeline

For the Y1H assay, a DNA promoter is cloned upstream of the His3 reporter. This reporter construct is integrated into the yeast genome of Yα1867. The corresponding interaction partners, the TFs, are transformed into the Y1H-S2a yeast strain. After picking the DNA promoter bait strain clone with the least self-activation (one clone per promoter), this bait strain clone will be tested in a primary screening against a mini-pool of various TFs. The 16 TF mini-pools contain between 13 and maximal 20 TFs (detailed description of the TFs, see Table A.5). The positive interactions found in the primary screening are verified in a one-by-one analysis Y1H screen, where each bait strain clone is mated against one TF strain in four repeats.

4.2.20.1 Y1H primary screening

The Tecan liquid handling robot distributed all used DNA promoter bait strains. One bait strain was inoculated into position 1-10 of one row of a 96-well microtiter plate (filled with 170 µl SC-His media), indicated in Figure 4.1 (red circles). These DNA promoter bait strain plates were grown on a vibration plate shaker for 72 hours at 30°C. 75 µl of grown culture was mixed with 75 µl 40 % glycerol, sealed and stored at -80°C. Two glycerol copies were prepared, one for common use and one copy as archival stock.

The TFs were picked according to the distribution pattern, illustrated into Figure 4.3, into a 96-well u-bottom plate, filled with 170 µl SC-Trp media and grown on a vibration shaker for 24 hours at 30°C. Afterwards they were spotted on solid SC-Trp plates and grown for 72 hours at 30°C. The various TFs of one pool were scraped from one plate and resuspended in 20 ml of SC-Trp media. TF mini-pool 1-8 were filled into the TF_pool_all_1 plate, TF mini-pool 9-16 were filled into the TF_pool_all_2 plate (clarified in Figure 4.1). Two glycerol stocks were made from each TF mini-pool mix, sealed and stored at -80°C.

The Y1H primary screening is performed by mating of one DNA promoter bait strain plate with a TF_pool_all plate. Therefore 5 µl of both plates (DNA promoter bait plate and TF_pool_all plate) were inoculated with the Tecan liquid handling robot in a 96-well u-bottom microtiter plate, filled with 150 µl of YEPD, followed by 24 hours shaking on a vibration shaker at 30°C. To select for diploid yeast cells, 5 µl of the mated yeast culture were transferred into a new 96-well u-bottom microtiter plate, filled with 150 µl SC-Trp -His media, followed by 48 hours incubation on a vibration shaker at 30°C. On the third day after the mating, 5 µl of the diploid yeast cells were transferred again into a new 96-well u-bottom microtiter plate, filled with SC -Trp -His media, to obtain a second generation of diploid cells. The plates were incubated on a vibration

shaker for 24 hours at 30°C. The next day, the diploid selected yeast cells were diluted 1:5 with fresh SC - Trp - His media to an OD_{600} of 0.2 and 5 µl were spotted on SC -Trp -His + 3-AT plates (containing 1 mM, 2.5 mM, 5.0 mM, 10 mM 3-AT). Two days after the spotting pictures were taken from all plates. The read-out plates were scored manually. Afterwards a python script, written by Patrick Schwehn, was trained to recognize and score the positive interactions. The score was categorized into 5 categories, from 1, almost no yeast growth to 5, strong yeast growth. After the automated analysis the results were double-checked and the positive scored interactions were selected for the Y1H verification.

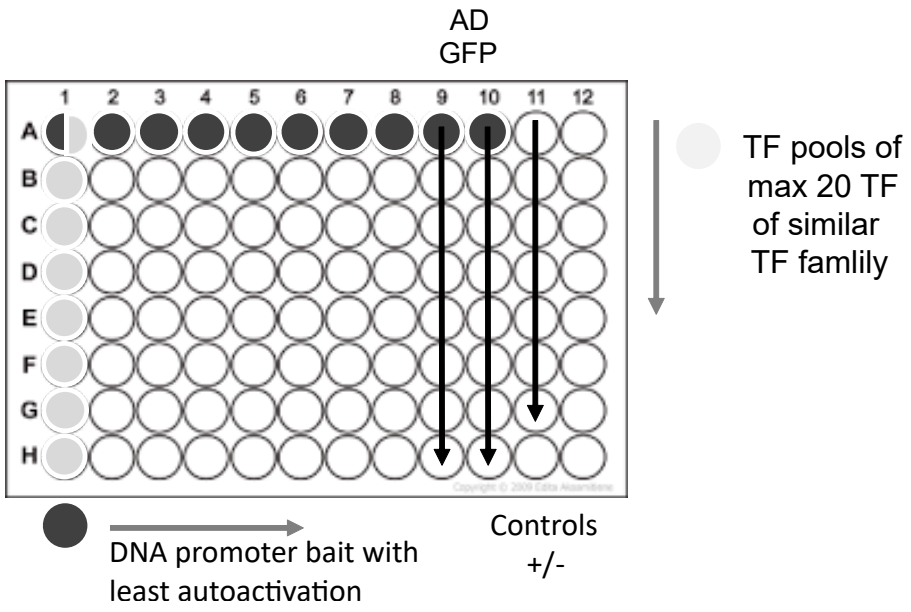


Figure 4.1: 96-well microtiter plate layout for primary screening

One DNA promoter bait strain was distributed into one row of a 96-well u-bottom shape microtiter plate (grey circles), leading to eight different DNA promoter bait strains filled into one microtiter plate. TF mini-pools were distributed into columns, 8 mini-pools were distributed over one microtiter plate (black circles). Column 9 and 10 were filled with Y1H-S2a strain transformed with pDEST AD 2µ GFP and an empty pDEST AD 2µ vector as negative mating control. Column 11 was filled with several negative and positive controls.

4.2.20.2 Saturation assay

For the saturation assay, four independent primary screens with 12 DNA promoter bait plates (contain each 8 DNA baits) against all TF mini pools were performed.

4.2.20.3 Y1H verification

The TFs of one mini-pool were distributed over three columns of a 96-well microtiter plate for one repeat. All four repeats of the same bait strain clone with the respective TFs were performed

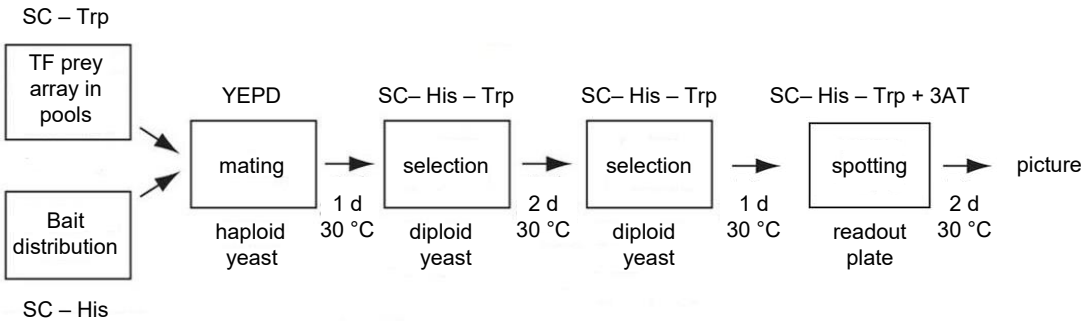


Figure 4.2: Experimental setup of Y1H pipeline

The Y1H workflow is based on liquid cultures. They are grown in 96-well u-bottom shape microtiter plates and transfered as described in the text. The prey and bait strains are mated in YEPD and afterwards double selected in SC-His-Trp selection media. For the subsequent analysis the diploid yeast is spotted on SC-His-Trp +3AT readout plates. Adapted from Reece-Hoyes et al. [193].

on one 96-well microtiter plate. Figure 4.3 illustrate the distribution pattern of the four repeats of single TFs on one 96-well microtiter plate. For the verification of primary positive interactions, each DNA promoter bait strain was inoculated separately into 8 ml of SC-His media and grown shaking for 48-72 hours at 30°C and for the mating distributed over a 96-well u-bottom microtiter plate to be able to use the Tecan liquid handling robot for the mating. As mating partners, the TF_pool_single_plates 1-16 were inoculated in 170 µl SC-Trp media and grown on a vibration shaker for 48-72 hours. 5 µl of DNA promoter bait strain culture and 5 µl of the TF_pool_single_plate culture were inoculated into 150 µl YEPD media, filled in a 96-well u-bottom plate with the Tecan liquid handling robot. After 12-24 hours incubation on a vibration shaker at 30°C, 5 µl of the mated yeast cells were transferred into the selection media SC -Trp - His (150 µl in a 96-well u-bottom plates) to select for diploid cells. 48 hours later, 5 µl of diploid cells were transferred again into fresh 150 µl SC -Trp - His media, followed by 24 hours incubation on a vibration shaker at 30°C. 5 µl of the mated yeast cells were 1:5 diluted and spotted on dried YEPD and SC - Trp - His + 1 mM 3-AT plates for the readout. Two days after, pictures of all plates were taken and analyzed. Cytoscape plugins were used to analyze and display the network [230].

Position E3, E6, E9 and E12 were filled with a Y1H-S2a strain transformed with pDEST AD 2µ GFP. Position F3, F6, F9 and F12 were filled with a Y1H-S2a strain transformed with an empty pDEST AD 2µ vector. pDEST AD 2µ GFP and pDEST AD 2µ vector are the negative controls for the DNA promoter bait strains, to identify DNA promoter bait auto-activators. Position G3, H3, G6, H6, G9, H9, G12 and H12 were filled with already mated positive control strain.

4.2.21 Protein expression

A subset of TFs were chosen and transformed into the pGEX6 expression vector, listed in Table 4.6. The plasmid vector contains a N-terminal GST-tag, that is fused to the TFs. They were transformed into Rosetta^{TM} (DE3) cells. Rosetta host strains are BL21 derivates that enhance

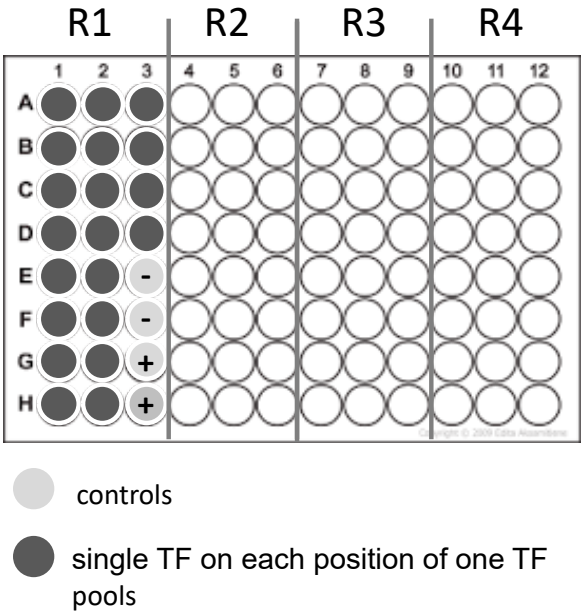


Figure 4.3: Distribution of TFs on 96-well microtiter plate for verification
Each TF pool contain up to 20 TFs of one TF family. For some TF pools two TF families were merged into one pool. For the 4-fold verification each single TF was distributed four times onto a 96-well plate to achieve a four repeat design on one single spotting plate. Each three columns represent one repeat. The last four samples of each third column contain two negative controls and two positive controls.

the expression of eukaryotic proteins in E.coli. The correct clones were sequenced verified and showed a resistance again Carb and CAM. The transformed Rosetta cells were grown over night at 37°C on a orbit shaker and freshly inoculated the next day in LB-Carb/CAM media. With an OD of 0.4-0.6, the cells were induced with a final concentration of 1 mM IPTG. After 4 hours of growth the cells were harvested, spun down and the cell pellet was frozen at -20°C.

4.2.22 Protein lysis and purification

The cell pellet undergoes three freeze and thaw cycles on ice and with liquid nitrogen. Then the cells are resuspended in 10 ml lysis buffer (140 mM NaCl, 10 mM Na_2HPO_4, 2.7 mM KCl, 2.0 mM KH_2PO_4, 1 mM EDTA, 1 mM DTT, 1 mg/ml Lysozym, 0.1 % Triton X-100, 1x Protease Inhibitor cocktail, 1x DNase) and incubated for 30 min at 4°C. For 50 ml cells, 10 ml lysis buffer was used, if less cells were used, the amount of buffer was adjusted. The lysate was spun at 4°C for 10 min at 10000g. The lysate supernatant was transferred to a new tube. MagneGST# Glutathione Particles, 4 ml (Promega, V8611), hereafter referred as MagneGST beads, were used for the purification of the GST-tagged TF-proteins.

100 µl magnetic beads per sample were preequilibrated according to the manufacturer protocol and subsequently added to the centrifugated lysate. Incubate the beads with the lysate at 4°C for 60-120 min on a shaker (120 rpm, the beads should be in constant motion). Using a magnetic rack, the beads are separated from the supernatant, which can be collected as 'flow through' for

further gel analysis. The beads are washed with 1000 µl MagneGST Bind/Wash buffer (140 mM NaCl, 4.2 mM Na_2HPO_4, 10 mM KCl, 2.0 mM KH_2PO_4, 1.0 mM DTT) and incubated and mixed for 5 min at 4°C or RT. The washing step is repeated without incubation two times. The beads remain in the sample tube, while the washing buffer is discarded. For each step the samples are put in the magnetic rack, that the beads are bound by the magnet. For protein elution, the beads are incubated in 50-100 µl elution buffer (50 mM Glutathione, 50 mM Tris-HCl, pH 8.1) for 30-60 min at 4°C on a shaker (1000 rpm). After putting the samples back on the magnetic rack, the eluat can be transferred into a new tube. The eluat contain the purified proteins.

4.2.22.1 Buffer exchange

For protein quantification with the Pierce$_{TM}$ BCA Assay kit (Thermo Fisher, 23227) the buffer of the eluted proteins need to be change to get rid of the glutathione. Otherwise the glutathione would interfere with the BCA Assay. Therefore a buffer exchange by ultrafiltration with Amicon Ultra 0.5 ml 30 K columns was performed. The buffer exchange was performed with 50 mM Tris-HCl because this is part of the elution buffer. The ultrafiltration filters are prerinsed with 500 µl 50 mM Tris-HCl, pH 7.4 and centrifugated at 14000 g for 4 min. Then the eluted proteins (50-100 µl) are loaded onto the ultrafilters and diluted with 350-400 µl 50 mM Tris-HCl, pH 7.4. The ultrafilters are centrifugated with 14000 g for 3-5 min at 4°C. Afterwards the fluid level needs to be checked. It should not run under 50 µl. 450 µl 50 mM Tris-HCl was added again to the ultrafilters and the centrifugation step was repeated. (14000 g, 3-5 min). The fluid level should be now between 50-100 µl. The ultrafilters were then transferred to a new tube and reverse spun at 1000 g, 2 min at 4°C.

4.2.23 Protein detection on SDS-PAGE and Western Blot

To validate if TF proteins were expressed 100-500 µl of cell culture or 10 µl of cell lysate were loaded and separated on 10 % SDS polyacrylamide gels (120V, 60-120 min). Before the samples were resuspended and boiled in 2x Lämmli buffer (2x Tris_Glycerine Buffer + 1 M DTT (end concentration 1x buffer, 10 mM DTT)) for 10 min at 90°C. The SDS-polyacrylamide gels were either stained in Coomassie solution and destained afterwards. Or the proteins were transferred to a nitrocellulose membrane via wet-blotting (100V, constant, 60-90 min). The electrophoresis chambers and the mini Trans-Blot^R Cell of Bio-Rad was used for the separation and transfer. The membranes were blocked in TBS-T 0.1 % + 5 % skim milk for 90 min or over night and subsequently incubated with the primary antibody, diluted in blocking solution (TBS-T 0.1 % + 1 % skim milk) at 4°C over night or for 120-180 min. After washing the membranes three times in TBS-T 0.1 % for 5-15 min, they were incubated with the secondary antibody, diluted in blocking solution (TBS-T 0.1 % + 1 % milk) for 60-180 min and washed again three times. For signal detection the membranes were incubated in Super Signal West Femto substrate analyzed using a Luminescent Image Analyzer. The membranes were stained with Ponceau rouge to control equal loading and transfer.

4.2.24 PampDAP-seq

DAP-seq is an *in vitro* TF-DNA binding assay that detects TF binding sites in genomic DNA with *in vitro* expressed TFs [190]. In a modified version of DAP-seq, called ampDAP-seq, a PCR-amplified genomic DNA library is used as template for the TF-binding. For this study a slightly modified ampDAP-seq assay was used as a biophysical validation to verify PDIs from the *C. reinhardtii* gene regulatory network. The DNA promoter sequences that were cloned for the Y1H screen were PCR amplified and used for the DNA library preparation (detailed description of the DNA library preparation is displayed in section 4.2.24.1). Hereafter referred to as PampDAP-seq, promoter PCR amplified DAP-seq. After PCR amplification the promoter sequences were sheared, fragmented using Nextera tagmentation enzyme and short Illumina DNA adapters were attached to the promoter fragments. This process is called tagmentation reaction.

Tagmentation is a process by which unfragmented DNA is cleaved by Tn5 transposase and tagged with specific labels for further analysis. In a separate reaction, the affinity-purified TFs were prepared by *in vitro* expression and bound to glutathione-coupled magnetic beads. The promoter DNA library was added to the affinity-bound TFs and the unbound DNA was washed away. The TF-bound DNA fraction is eluted, amplified with Illumina Nextera index primers and NGS-sequenced afterwards ([190].

4.2.24.1 DNA promoter library preparation

The PCR-amplified promoter sequences (from Entry clones) were pooled and purified with magnetic beads, according to manufacturer protocol (MagSi-NGS^{PREP} PLUS, Steinbrenner Laborsysteme GmbH). The concentration of this promoter pool library, hereafter referred as DNA library, was quantified using the Quant-it^{TM} PicoGreen^{TM} ds DNA Assay-Kit according to manufacturers protocol (Quant-iT^{TM} PicoGreen, invitrogen). 13 µg pooled DNA library in 130 µl elution buffer (5 mM Tris-HCl, pH 8.5) were tranferred to a Covarids microTUB and sonicated using a Covaris E220 Focused-ultrasonicator system with a program of: 140 W, 'Cycles/Burst' 200, 'Duty Factor' 10 % and a time of 55 sec.The targeted fragment size was 250-500 bp. A fragment size of 200-600 bp was achieved. For further tagmentation, the DNA library was diluted to 6.5 ng/µl treated with Illumina Tagment DNA TD1 Enzyme and buffer kit (Illumina, 20034197).

Table 4.7: Tagmentation reaction

Component	Amount per reaction (µl)	Final amount
Tagmentation TD1 enzyme	0.5	-
Tagmentation buffer	2.0	1x
DNA pool library	1.0	6.5 ng
MQ H_2O	6.5	-

Table 4.8: Tagmentation PCR

Component	Amount per reaction (µl)	Final amount
Phusion Taq polymerase	0.2	0.4 U
Phusion polymerase buffer	4.0	1x
10 mM dNTPs	2.0	1.0 mM
Primer Nextera fw (10 µM)	1.0	0.5 µM
Primer Nextera rv (10 µM)	1.0	0.5 µM
MQ H_2O	1.8	-

Table 4.9: Thermocycler program tagmentation

Cycle number	Denature	Anneal	Extend
1	72°C for 3 min	-	-
2	95°C for 30 sec	-	-
3-13	95°C for 30 sec	55°C for 30 sec	72°C for 2 min
14	-	-	72°C for 10 min

4.2.24.2 PampDAP-seq, protein and DNA binding

25 µl MagneGST beads were washed four times with 1x PBS buffer (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 2 mM KH_2PO_4) to equilibrate the beads. To reduce the unspecifically binding affinity of the beads, the PBS contains 100 ng salmon sperm DNA. 3 µg of purified TF-GST proteins are solved in 400 µl 1x PBS and added to the washed MagneGST beads. The samples were rotated on a rotate shaker for 60 min at 4°C, that the beads are in constant motion and can bind the proteins in the solution. The samples were spun for 5 sec at 3,000 g to collect all liquid on the bottom of the tube. For separation of beads and solution, the samples were placed into a magnetic rack. The beads were washed five times with 400 µl of 1x PBS + NP40 (0.005 %). The supernatant was discarded. Afterwards the beads were washed three times with 400 µl 1x PBS. After the last washing step the beads are resuspended in 40 µl 1x PBS and transferred into a 8-well PCR strip or a 96-well microtiter plate.

75-100 ng of prepared DNA promoter library was diluted in 40 µl 1x PBS and added to the washed beads. The bound proteins and the DNA promoter library were incubated at RT for 60-120 min, shaking on a rotator shaker. Then the beads were washed four or eight times with 180 µl 1x PBS + NP40 (0.005 %) and subsequently transferred to a new tube or plate. Then the beads were washed again two or four times with 180 µl 1x PBS. After the final washing step the beads were transferred to a new tube or plate. The supernatant was discarded and the beads were resuspended in 25 µl elution buffer (5 mM Tris-HCl, pH 8.5). They were then heated at 98°C for 10 min to denaturate the proteins. The eluted DNA was then PCR amplified using Nextera i5 and i7 primers with unique barcodes to differentiate the different samples according to 4.10. The PCR samples were pooled into one sample and gel purified. PCR products between 150 - 400 bp were selected. The sample DNA amount was estimated using the Quant-it™ PicoGreen™ ds DNA Assay-Kit and send for NGS sequencing.

Table 4.10: Thermocycler program DAP PCR

Cycle number	Denature	Anneal	Extend
1	98°C for 2 min	-	-
2	98°C for 30 sec	-	-
3-23	98°C for 15 sec	55°C for 30 sec	72°C for 2 min
24	-	-	72°C for 10 min

4.2.24.3 PampDAP-seq analysis - computation pipeline

NGS data was demultiplexed by the NGS core facility of the Helmholtz Zentrum München. Afterwards the reads were trimmed according their read quality using trimmomatic-0.39 [323]. Afterwards they were aligned onto the reference promoter sequences using the bwa 0.7.17 tool [324]. All mapped reads were indexed and and sorted using the tool Samtools 1.15 [325, 326]. The subsequent peak calling was done with macs2 v2.2.7.1 [327].

For *denovo* binding motif identification and discovery the MEME tool of 'The MEME Suite' was used [328, 329]. For comparison with already known binding motifs of *A.thaliana* the TomTom tool of 'The MEME Suite' was used [330]. The computational pipeline to analyze the DAP-seq data was coded and performed by Dr. Chung-Wen Lin.

4.2.25 GO term analysis

All avaiable GO terms of the 303 interaction partners in the Y1H screen (142 TFs and 220 DNA promoter baits) were extracted from g:Profiler [232]. The GO terms are based on the genome annotation v5.5 [52]. The GO terms for biological processes were used to condense the 303 interaction partners in higher-level categories (hereafter called: GO groups). One example is that 'RNA processing', 'mRNA processing', 'ncRNA processing', 'regulation of mRNA processing' and 'tRNA processing' are summarised by the term 'RNA processing'. For 113 genes out of 303 no GO term could be assigned. These genes were grouped as 'unclassified'. 47 genes were grouped into two categories and 12 had an third identifier. The majority of the 'unclassified' genes are TFs and TRs with unknown function.

4.2.26 Overexpression *C. reinhardtii* cell lines

In collaboration with Alexander Einhaus of the research group of Prof. Olaf Kruse of the University Bielefeld, the overexpression mutants for LRL1, CGL86, MYB_9, and MS_1 should be generated. To achieve a better expression, all four sequences were codon-optimized and the first intron of ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit 2 (rbcS2i1) was inserted sequential multiple times. The sequence optimization was performed with the help of the online tool 'Intronserter' [244]. This method has be shown to dramatically improve the protein expression [245]. After optimization of the gene sequences, they were ordered at the gene synthesis company 'GenScript'. MYB_9, CGL86 and LRL1 could be gene synthesized by 'GenScript', MS_1 not. After the sequences were synthesized, Alexander Einhaus from the Prof. Kruse laboratory generated the vector constructs for transient transformation in *C. reinhardtii*.

Table 4.11: Candidates in GO groups

GO group	Gene count
Amino acid metabolism	6
Biosynthesis of secondary metabolites	5
Carbohydrate metabolism	12
Cell cycle	12
Cell signaling	18
Chromatin organization	20
Cytoskeleton	4
DNA repair	11
Lipid metabolism	33
Nitrogen metabolism	6
Oxidative Phosphorylation	1
Photosynthesis	17
Protein modification	20
RNA processing	8
Stress response	22
Translation	5
Transport	10
Development	1
Unclassified	151

All protein were tagged with YFP N - and C-terminal for an easy visual transformation selection with the microscope. Alexander Einhaus performed the transformation several times but could only find cells with very little expression of the regulatory proteins. He extracted always three different cells for each construct and measured the fluorescence of liquid culture of these cell populations.

The highest gene expression he detected was 3-5x higher than the parental strain (UVM4 strain) he used for the transformation. Most of the transformed cells exhibit a protein expression of 1-3x fold compared to the parental strain. Other overexpression experiments of Alexander Einhaus resulted in a 1,000x higher expression of the desired protein. After applying nutrient stress to the transformed cell populations, no more overexpression was detectable. Previous experiments of the Prof. Kruse lab have shown this reduction of expression proteins during nutrient stress but because in this experiment the overexpression rate was so low, the expression drops almost to one or below during starvation.

Acknowledgement

First and foremost, I would like to express my sincere gratitude to my advisor Professor Dr. Pascal Falter-Braun, for his continuous support of my Ph.D. study during my whole time at the Institute of Network Biology and at the TUM Weihenstephan. His guidance helped me develop and grow as a scientist, giving me much freedom and confidence in my research. I also want to thank Prof. Dr. Nickelsen, Prof. Dr. Werth, and Dr. Mokranjac for supporting this thesis by being active graduate committee members.
My sincere thanks go to all my colleagues at the INET of the Helmholtz Zentrum München. The friendly and kind atmosphere gave me the basis for my work and research. I enjoyed being part of this fantastic team I could always rely on. No matter if I needed some help in the lab or someone listening when I was frustrated due to a failed experiment.

A special thank you goes to my colleagues and friends:

To Ram, with whom I really enjoyed it to work in the lab and loved to create delightful and innovative ideas and visions;
to Benny, who was always there to calm me down when I was overwhelmed by science;
to Patrick, who showed me a lot of Excel tricks and was always there to help me with programming issues;
to Theresa for so much support, many lovely hours and coffee breaks together;
to Chung-Wen for his tremendous support in data analysis;
and last but not least, to Alex, Sarah, and Simin for their tireless help in the laboratory.

I also want to thank all my 'Tübinger' and 'Münchner' friends for their support, the great distractions, and our fantastic trips.

Lastly, I would like to thank my whole family for all their support and encouragement during this thesis, especially my husband, who supported me unremittingly and couldn't be a better father for my kids.

Munich, 12.06.2023

Mayra Sauer

Bibliography

[1] H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.). IPCC 2022: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Technical report, Cambridge University Press, 2022.

[2] Assessing the Global Climate in 2021. https://www.ncei.noaa.gov/news/global-climate-202112, January 2022. Accessed: 2022-8-23.

[3] Greenhouse gas emissions. https://www.umweltbundesamt.de/en/data/environmental-indicators/indicator-greenhouse-gas-emissions, February 2017. Accessed: 2022-8-23.

[4] Jeff Desjardins. The World's Projected Energy Mix, 2018-2040, Source IEA, World Energy Outlook 2019. Modified by Visual Capitalist. https://www.visualcapitalist.com/the-worlds-projected-energy-mix-2018-2040/. Accessed: 2022-8-24.

[5] Hannah Ritchie, Max Roser, and Pablo Rosado. Energy. *Our World in Data*, 2020.

[6] Iea (2021). Global Energy Review 2021. https://www.iea.org/reports/global-energy-review-2021, 2021. Accessed: 2022-8-25.

[7] United Nations. What is renewable energy? | United Nations. https://www.un.org/en/climatechange/what-is-renewable-energy. Accessed: 2022-8-24.

[8] César Morales Bayetero, Carlos Mafla Yépez, Ignacio Benavides Cevallos, and Erik Hernández Rueda. Effect of the use of additives in biodiesel blends on the performance and opacity of a diesel engine. *Materials Today: Proceedings*, 49:93–99, January 2022.

[9] Wikipedia contributors. Biofuel. https://en.wikipedia.org/w/index.php?title=Biofuel&oldid=1103923184, August 2022. Accessed: NA-NA-NA.

[10] Biodiesel. https://afdc.energy.gov/fuels/biodiesel.html. Accessed: 2022-8-24.

[11] Biodiesel. https://www.eubia.org/cms/wiki-biomass/biofuels/biodiesel/. Accessed: 2022-8-25.

[12] Melinda J Griffiths and Susan T L Harrison. Lipid productivity as a key characteristic for choosing algal species for biodiesel production. *J. Appl. Phycol.*, 21(5):493–507, October 2009.

[13] Yusuf Chisti. Biodiesel from microalgae beats bioethanol. *Trends Biotechnol.*, 26(3):126–131, March 2008.

[14] Stuart A Scott, Matthew P Davey, John S Dennis, Irmtraud Horst, Christopher J Howe, David J Lea-Smith, and Alison G Smith. Biodiesel from algae: challenges and prospects. *Curr. Opin. Biotechnol.*, 21(3):277–286, June 2010.

[15] Jee Young Kim, Jong-Min Jung, Sungyup Jung, Young-Kwon Park, Yiu Fai Tsang, Kun-Yi Andrew Lin, Yoon-E Choi, and Eilhann E Kwon. Biodiesel from microalgae: Recent progress and key challenges. *Prog. Energy Combust. Sci.*, 93:101020, November 2022.

[16] Spyros Foteinis, Efthalia Chatzisymeon, Alexandros Litinas, and Theocharis Tsoutsos. Used-cooking-oil biodiesel: Life cycle assessment and comparison with first- and third-generation biofuel. *Renewable Energy*, 153:588–600, June 2020.

[17] Mbalo Ndiaye, Abdellah Arhaliass, Jack Legrand, Guillaume Roelens, and Anthony Kerihuel. Reuse of waste animal fat in biodiesel: Biorefining heavily-degraded contaminant-rich waste animal fat and formulation as diesel fuel additive. *Renewable Energy*, 145:1073–1079, January 2020.

[18] D Ryan Georgianna and Stephen P Mayfield. Exploiting diversity and synthetic biology for the production of algal biofuels. *Nature*, 488(7411):329–335, August 2012.

[19] Qiang Hu, Milton Sommerfeld, Eric Jarvis, Maria Ghirardi, Matthew Posewitz, Michael Seibert, and Al Darzins. Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *Plant J.*, 54(4):621–639, May 2008.

[20] Michael Hannon, Javier Gimpel, Miller Tran, Beth Rasala, and Stephen Mayfield. Biofuels from algae: challenges and potential. *Biofuels*, 1(5):763–784, September 2010.

[21] Nitin Verma and Vivek Kumar. Microbial conversion of waste biomass into bioethanol: current challenges and future prospects. *Biomass Conversion and Biorefinery*, August 2021.

[22] Michael D Guiry. HOW MANY SPECIES OF ALGAE ARE THERE? *J. Phycol.*, 48(5):1057–1063, October 2012.

[23] Claude Courties, André Vaquer, Marc Troussellier, Jacques Lautier, and Hervé Claustre. Smallest eukarotic organism. *Nature*, 370(6487):255–255, July 1994.

[24] Armin Hallmann. Algal transgenics and biotechnology. *Transgenic Plant J*, 1(1):81–98, 2007.

[25] Robert Steneck, M H Graham, B J Bourque, D Corbett, J M Erlandson, J A Estes, and M J Tegner. Kelp Forest Ecosystems: Biodiversity, Stability, Resilience and Future. *Cambridge University Press*, 29(4):436, 2002.

[26] A E Walsby. Microalgae: Biotechnology and Microbiology. By E. W. Becker. Cambridge: Cambridge University Press (1994), pp. 230, £40.00, US$69.95. ISBN 0-521-35020-4. *Exp. Agric.*, 31(1):112–112, January 1995.

[27] H Fukuda, A Kondo, and H Noda. Biodiesel fuel production by transesterification of oils. *J. Biosci. Bioeng.*, 92(5):405–416, 2001.

[28] Randall J Weselake. Chapter 15 - Engineering Oil Accumulation in Vegetative Tissue. In Thomas A McKeon, Douglas G Hayes, David F Hildebrand, and Randall J Weselake, editors, *Industrial Oil Crops*, pages 413–434. AOCS Press, January 2016.

[29] Peter J le B. Williams and Lieve M L Laurens. Microalgae as biodiesel & biomass feedstocks: Review & analysis of the biochemistry, energetics & economics. *Energy Environ. Sci.*, 3(5):554–590, May 2010.

[30] Qingtao Zhang, Jiong Ma, Guoyu Qiu, Li Li, Shu Geng, E Hasi, Cheng Li, Guangyi Wang, and Xiaoyan Li. Potential energy production from algae on marginal land in China. *Bioresour. Technol.*, 109:252–260, April 2012.

[31] Anoop Singh, Poonam Singh Nigam, and Jerry D Murphy. Renewable fuels from algae: an answer to debatable land based fuels. *Bioresour. Technol.*, 102(1):10–16, January 2011.

[32] Gerhard Knothe. Dependence of biodiesel fuel properties on the structure of fatty acid alkyl esters. *Fuel Process. Technol.*, 86(10):1059–1070, June 2005.

[33] Digambar Singh, Dilip Sharma, S L Soni, Sumit Sharma, and Deepika Kumari. Chemical compositions, properties, and standards for different generation biodiesels: A review. *Fuel*, 253:60–71, October 2019.

[34] Ahmad Farhad Talebi, Seyed Kaveh Mohtashami, Meisam Tabatabaei, Masoud Tohidfar, Abdolreza Bagheri, Mehrshad Zeinalabedini, Hossein Hadavand Mirzaei, Mehrdad Mirzajanzadeh, Saeid Malekzadeh Shafaroudi, and Shiva Bakhtiari. Fatty acids profiling: A selective criterion for screening microalgae strains for biodiesel production. *Algal Research*, 2(3):258–267, July 2013.

[35] L C Meher, D Vidya Sagar, and S N Naik. Technical aspects of biodiesel production by transesterification—a review. *Renewable Sustainable Energy Rev.*, 10(3):248–268, June 2006.

[36] María Jesús Ramos, Carmen María Fernández, Abraham Casas, Lourdes Rodríguez, and Angel Pérez. Influence of fatty acid composition of raw materials on biodiesel properties. *Bioresour. Technol.*, 100(1):261–268, January 2009.

[37] Yujin Cao, Wei Liu, Xin Xu, Haibo Zhang, Jiming Wang, and Mo Xian. Production of free monounsaturated fatty acids by metabolically engineered Escherichia coli. *Biotechnol. Biofuels*, 7:59, April 2014.

[38] J Sheehan, T Dunahay, J Benemann, and P Roessler. Look Back at the U.S. Department of Energy's Aquatic Species Program: Biodiesel from Algae; Close-Out Report. Technical report, U.S. Department of Energy, 1998.

[39] Sabeeha S Merchant, Janette Kropat, Bensheng Liu, Johnathan Shaw, and Jaruswan Warakanont. TAG, you're it! Chlamydomonas as a reference organism for understanding algal triacylglycerol accumulation. *Curr. Opin. Biotechnol.*, 23(3):352–363, June 2012.

[40] Bensheng Liu and Christoph Benning. Lipid metabolism in microalgae distinguishes itself. *Curr. Opin. Biotechnol.*, 24(2):300–309, April 2013.

[41] Severin Sasso, Herwig Stibor, Maria Mittag, and Arthur R Grossman. From molecular manipulation of domesticated Chlamydomonas reinhardtii to survival in nature. *Elife*, 7, November 2018.

[42] Elizabeth H Harris. CHLAMYDOMONAS AS A MODEL ORGANISM. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 52:363–406, June 2001.

[43] Elizabeth H Harris. *The Chlamydomonas Sourcebook: Introduction to Chlamydomonas and Its Laboratory Use: Volume 1*. Academic Press, March 2009.

[44] Michael Brunner and Martha Merrow. The green yeast uses its plant-like clock to regulate its animal-like tail. *Genes Dev.*, 22(7):825–831, April 2008.

[45] K L Kindle. High-frequency nuclear transformation of Chlamydomonas reinhardtii. *Proc. Natl. Acad. Sci. U. S. A.*, 87(3):1228–1232, February 1990.

[46] J E Boynton, N W Gillham, E H Harris, J P Hosler, A M Johnson, A R Jones, B L Randolph-Anderson, D Robertson, T M Klein, and K B Shark. Chloroplast transformation in Chlamydomonas with high velocity microprojectiles. *Science*, 240(4858):1534–1538, June 1988.

[47] U W Goodenough. Green yeast. *Cell*, 70(4):533–538, August 1992.

[48] Claire Remacle, Pierre Cardol, Nadine Coosemans, Mauricette Gaisne, and Nathalie Bonnefoy. High-efficiency biolistic transformation of Chlamydomonas mitochondria can be used to insert mutations in complex I genes. *Proc. Natl. Acad. Sci. U. S. A.*, 103(12):4771–4776, March 2006.

[49] Takuya Matsuo and Masahiro Ishiura. Chlamydomonas reinhardtii as a new model system for studying the molecular basis of the circadian clock. *FEBS Lett.*, 585(10):1495–1502, May 2011.

[50] Attila Molnar, Andrew Bassett, Eva Thuenemann, Frank Schwach, Shantanu Karkare, Stephan Ossowski, Detlef Weigel, and David Baulcombe. Highly specific gene silencing by artificial microRNAs in the unicellular alga Chlamydomonas reinhardtii. *Plant J.*, 58(1):165–174, April 2009.

[51] Aron Ferenczi, Douglas Euan Pyott, Andromachi Xipnitou, and Attila Molnar. Efficient targeted DNA editing and replacement in Chlamydomonas reinhardtii using Cpf1 ribonucleoproteins and single-stranded DNA. *Proc. Natl. Acad. Sci. U. S. A.*, 114(51):13567–13572, December 2017.

[52] Sabeeha S Merchant, Simon E Prochnik, Olivier Vallon, Elizabeth H Harris, Steven J Karpowicz, George B Witman, Astrid Terry, Asaf Salamov, Lillian K Fritz-Laylin, Laurence Maréchal-Drouard, Wallace F Marshall, Liang-Hu Qu, David R Nelson, Anton A Sanderfoot, Martin H Spalding, Vladimir V Kapitonov, Qinghu Ren, Patrick Ferris, Erika Lindquist, Harris Shapiro, Susan M Lucas, Jane Grimwood, Jeremy Schmutz, Pierre Cardol, Heriberto Cerutti, Guillaume Chanfreau, Chun-Long Chen, Valérie Cognat, Martin T Croft, Rachel Dent, Susan Dutcher, Emilio Fernández, Hideya Fukuzawa, David González-Ballester, Diego González-Halphen, Armin Hallmann, Marc Hanikenne, Michael Hippler, William Inwood, Kamel Jabbari, Ming Kalanon, Richard Kuras, Paul A Lefebvre, Stéphane D Lemaire, Alexey V Lobanov, Martin Lohr, Andrea Manuell, Iris Meier, Laurens Mets, Maria Mittag, Telsa Mittelmeier, James V Moroney, Jeffrey Moseley, Carolyn Napoli, Aurora M Nedelcu, Krishna Niyogi, Sergey V Novoselov, Ian T Paulsen, Greg Pazour, Saul Purton, Jean-Philippe Ral, Diego Mauricio Riaño-Pachón, Wayne Riekhof, Linda Rymarquis, Michael Schroda, David Stern, James Umen, Robert Willows, Nedra Wilson, Sara Lana Zimmer, Jens Allmer, Janneke Balk, Katerina Bisova, Chong-Jian Chen, Marek Elias, Karla Gendler, Charles Hauser, Mary Rose Lamb, Heidi Ledford, Joanne C Long, Jun Minagawa, M Dudley Page, Junmin Pan, Wirulda Pootakham, Sanja Roje, Annkatrin Rose, Eric Stahlberg, Aimee M Terauchi, Pinfen Yang, Steven Ball, Chris Bowler, Carol L Dieckmann, Vadim N Gladyshev, Pamela Green, Richard Jorgensen, Stephen Mayfield, Bernd Mueller-Roeber, Sathish Rajamani, Richard T Sayre, Peter Brokstein, Inna Dubchak, David Goodstein, Leila Hornick, Y Wayne Huang, Jinal Jhaveri, Yigong Luo, Diego Martínez, Wing Chi Abby Ngau, Bobby Otillar, Alexander Poliakov, Aaron Porter, Lukasz Szajkowski, Gregory Werner, Kemin Zhou, Igor V Grigoriev, Daniel S Rokhsar, and Arthur R Grossman. The Chlamydomonas genome reveals the evolution of key animal and plant functions. *Science*, 318(5848):245–250, October 2007.

[53] David M Goodstein, Shengqiang Shu, Russell Howson, Rochak Neupane, Richard D Hayes, Joni Fazo, Therese Mitros, William Dirks, Uffe Hellsten, Nicholas Putnam, and Daniel S Rokhsar. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.*, 40(Database issue):D1178–86, January 2012.

[54] Ian K Blaby, Crysten E Blaby-Haas, Nicolas Tourasse, Erik F Y Hom, David Lopez, Munevver Aksoy, Arthur Grossman, James Umen, Susan Dutcher, Mary Porter, Stephen King, George B Witman, Mario Stanke, Elizabeth H Harris, David Goodstein, Jane Grimwood, Jeremy Schmutz, Olivier Vallon, Sabeeha S Merchant, and Simon Prochnik. The Chlamydomonas genome project: a decade on. *Trends Plant Sci.*, 19(10):672–680, October 2014.

[55] Xiaobo Li, Weronika Patena, Friedrich Fauser, Robert E Jinkerson, Shai Saroussi, Moritz T Meyer, Nina Ivanova, Jacob M Robertson, Rebecca Yue, Ru Zhang, Josep Vilarrasa-Blasi, Tyler M Wittkopp, Silvia Ramundo, Sean R Blum, Audrey Goh, Matthew Laudon, Tharan Srikumar, Paul A Lefebvre, Arthur R Grossman, and Martin C Jonikas. A genome-wide algal mutant library and functional screen identifies genes required for eukaryotic photosynthesis. *Nat. Genet.*, 51(4):627–635, April 2019.

[56] Chlamydomonas Resource Center Website. `https://www.chlamycollection.org/`, May 2015. Accessed: 2022-4-12.

[57] Cristina González-Fernández, Bruno Sialve, Nicolas Bernet, and Jean-Philippe Steyer. Impact of microalgae characteristics on their conversion to biofuel. Part II: Focus on biomethane production. *Biofuels Bioprod. Biorefin.*, 6(2):205–218, March 2012.

[58] Gonzalo M Figueroa-Torres, Wan M Asyraf Wan Mahmood, Jon K Pittman, and Constantinos Theodoropoulos. Microalgal biomass as a biorefinery platform for biobutanol and biodiesel production. *Biochem. Eng. J.*, 153:107396, January 2020.

[59] Ines Barkia, Nazamid Saari, and Schonna R Manning. Microalgae for High-Value Products Towards Human Health and Nutrition. *Mar. Drugs*, 17(5), May 2019.

[60] Yevhen Maltsev and Kateryna Maltseva. Fatty acids of microalgae: diversity and applications. *Rev. Environ. Sci. Technol.*, 20(2):515–547, June 2021.

[61] O Pulz. Photobioreactors: production systems for phototrophic microorganisms. *Appl. Microbiol. Biotechnol.*, 57(3):287–293, October 2001.

[62] Yuan-Kun Lee. Microalgal mass culture systems and methods: Their limitation and potential. *J. Appl. Phycol.*, 13(4):307–315, August 2001.

[63] Y Li-Beisson, F Beisson, and W Riekhof. Metabolism of acyl-lipids in Chlamydomonas reinhardtii. *Plant J.*, 82(3):504–522, 2015.

[64] E C Goncalves, A C Wilkie, M Kirst, and B Rathinasabapathi. Metabolic regulation of triacylglycerol accumulation in the green algae: identification of potential targets for engineering to improve oil yield. *Plant Biotechnol. J.*, 14(8):1649–1660, 2016.

[65] Alexandro Cagliari, Rogerio Margis, Felipe dos Santos Maraschin, Andreia Carina Turchetto-Zolet, Guilherme Loss, and Marcia Margis-Pinheiro. Biosynthesis of Triacylglycerols (TAGs) in Plants and algae. *International Journal of Plant Biology*, 2(1):e10, December 2011.

[66] Jin Liu, Danxiang Han, Kangsup Yoon, Qiang Hu, and Yantao Li. Characterization of type 2 diacylglycerol acyltransferases in Chlamydomonas reinhardtii reveals their distinct substrate specificities and functions in triacylglycerol biosynthesis. *Plant J.*, 86(1):3–19, April 2016.

[67] Yang Xu, Kristian Mark P Caldo, Dipasmita Pal-Nath, Jocelyn Ozga, M Joanne Lemieux, Randall J Weselake, and Guanqun Chen. Properties and Biotechnological Applications of Acyl-CoA:diacylglycerol Acyltransferase and Phospholipid:diacylglycerol Acyltransferase from Terrestrial Plants and Microalgae. *Lipids*, 53(7):663–688, July 2018.

[68] Yonghua Li-Beisson, Jay J Thelen, Eric Fedosejevs, and John L Harwood. The lipid biochemistry of eukaryotic algae. *Prog. Lipid Res.*, 74:31–68, April 2019.

[69] Fan Bai, Lihua Yu, Jianan Shi, Yonghua Li-Beisson, and Jin Liu. Long-chain acyl-CoA synthetases activate fatty acids for lipid synthesis, remodeling and energy production in Chlamydomonas. *New Phytol.*, 233(2):823–837, January 2022.

[70] Jilian Fan, Carl Andre, and Changcheng Xu. A chloroplast pathway for the de novo biosynthesis of triacylglycerol in Chlamydomonas reinhardtii. *FEBS Lett.*, 585(12):1985–1991, June 2011.

[71] X Johnson and J Alric. Central carbon metabolism and electron transport in Chlamydomonas reinhardtii: metabolic constraints for carbon partitioning between oil and starch. *Eukaryot. Cell*, 12(6):776–793, 2013.

[72] Elton C Goncalves, Jin Koh, Ning Zhu, Mi-Jeong Yoo, Sixue Chen, Takuya Matsuo, Jodie V Johnson, and Bala Rathinasabapathi. Nitrogen starvation-induced accumulation of triacylglycerol in the green algae: evidence for a role for ROC40, a transcription factor involved in circadian rhythm. *Plant J.*, 85(6):743–757, March 2016.

[73] Fabio Facchinelli and Andreas P M Weber. The metabolite transporters of the plastid envelope: an update. *Front. Plant Sci.*, 2:50, September 2011.

[74] Nanette R Boyle, Mark Dudley Page, Bensheng Liu, Ian K Blaby, David Casero, Janette Kropat, Shawn J Cokus, Anne Hong-Hermesdorf, Johnathan Shaw, Steven J Karpowicz, Sean D Gallaher, Shannon Johnson, Christoph Benning, Matteo Pellegrini, Arthur Grossman, and Sabeeha S Merchant. Three acyltransferases and nitrogen-responsive regulator are implicated in nitrogen starvation-induced triacylglycerol accumulation in Chlamydomonas. *J. Biol. Chem.*, 287(19):15811–15825, May 2012.

[75] Xiaobo Li, Eric R Moellering, Bensheng Liu, Cassandra Johnny, Marie Fedewa, Barbara B Sears, Min-Hao Kuo, and Christoph Benning. A galactoglycerolipid lipase is required for triacylglycerol accumulation and survival following nitrogen deprivation in Chlamydomonas reinhardtii - PubMed. *Plant Cell*, 24(11):4670–4686, November 2012.

[76] Kangsup Yoon, Danxiang Han, Yantao Li, Milton Sommerfeld, and Qiang Hu. Phospholipid:diacylglycerol acyltransferase is a multifunctional enzyme involved in membrane lipid turnover and degradation while synthesizing triacylglycerol in the unicellular green microalga Chlamydomonas reinhardtii. *Plant Cell*, 24(9):3708–3724, September 2012.

[77] Jeong-Jin Park, Hongxia Wang, Mahmoud Gargouri, Rahul R Deshpande, Jeremy N Skepper, F Omar Holguin, Matthew T Juergens, Yair Shachar-Hill, Leslie M Hicks, and David R Gang. The response of Chlamydomonas reinhardtii to nitrogen deprivation: a systems biology analysis. *Plant J.*, 81(4):611–624, February 2015.

[78] Linhui Yu, Chao Zhou, Jilian Fan, John Shanklin, and Changcheng Xu. Mechanisms and functions of membrane lipid remodeling in plants. *Plant J.*, 107(1):37–53, July 2021.

[79] W Eichenberger, J-C Schaffner, and A Boschetti. Characterization of proteins and lipids of photosystem I and II particles fromChlamydomonas reinhardi. *FEBS Lett.*, 84(1):144–148, December 1977.

[80] Hong-Wei Xue, Xu Chen, and Yu Mei. Function and regulation of phospholipid signalling in plants. *Biochem. J*, 421(2):145–156, June 2009.

[81] Jennifer Popko. Lipid Composition of Chlamydomonas reinhardtii. In Markus R Wenk, editor, *Encyclopedia of Lipidomics*, pages 1–6. Springer Netherlands, Dordrecht, 2017.

[82] Magali Siaut, Stéphan Cuiné, Caroline Cagnon, Boris Fessler, Mai Nguyen, Patrick Carrier, Audrey Beyly, Fred Beisson, Christian Triantaphylidès, Yonghua Li-Beisson, and Gilles Peltier. Oil accumulation in the model green alga Chlamydomonas reinhardtii: characterization, variability between common laboratory strains and relationship with starch reserves. *BMC Biotechnol.*, 11:7, January 2011.

[83] Pei-Li Shen, Hai-Tao Wang, Yan-Fei Pan, Ying-Ying Meng, Pei-Chun Wu, and Song Xue. Identification of Characteristic Fatty Acids to Quantify Triacylglycerols in Microalgae. *Front. Plant Sci.*, 7:162, February 2016.

[84] Miao Yang, Fantao Kong, Xi Xie, Peichun Wu, Yadong Chu, Xupeng Cao, and Song Xue. Galactolipid DGDG and Betaine Lipid DGTS Direct De Novo Synthesized Linolenate into Triacylglycerol in a Stress-Induced Starchless Mutant of Chlamydomonas reinhardtii. *Plant Cell Physiol.*, 61(4):851–862, April 2020.

[85] Danielle Yvonne Young and Yair Shachar-Hill. Large fluxes of fatty acids from membranes to triacylglycerol and back during N-deprivation and recovery in Chlamydomonas. *Plant Physiol.*, 185(3):796–814, April 2021.

[86] James W Allen, Concetta C DiRusso, and Paul N Black. Triacylglycerol synthesis during nitrogen stress involves the prokaryotic lipid synthesis pathway and acyl chain remodeling in the microalgae Coccomyxa subellipsoidea. *Algal Research*, 10:110–120, July 2015.

[87] A Lopez Garcia de Lomana, S Schauble, J Valenzuela, S Imam, W Carter, D D Bilgin, C B Yohn, S Turkarslan, D J Reiss, M V Orellana, N D Price, and N S Baliga. Transcriptional program for nitrogen starvation-induced lipid accumulation in Chlamydomonas reinhardtii. *Biotechnol. Biofuels*, 8:207, 2015.

[88] Xin Li, Hong-Ying Hu, and Yu-Ping Zhang. Growth and lipid accumulation properties of a freshwater microalga Scenedesmus sp. under different cultivation temperature. *Bioresour. Technol.*, 102(3):3098–3102, February 2011.

[89] Hugh Douglas Goold, Stéphan Cuiné, Bertrand Légeret, Yuanxue Liang, Sabine Brugière, Pascaline Auroy, Hélène Javot, Marianne Tardif, Brian Jones, Fred Beisson, Gilles Peltier, and Yonghua Li-Beisson. Saturating Light Induces Sustained Accumulation of Oil in Plastidal Lipid Droplets in Chlamydomonas reinhardtii. *Plant Physiol.*, 171(4):2406–2417, August 2016.

[90] Qiaoning He, Haijian Yang, Lei Wu, and Chunxiang Hu. Effect of light intensity on physiological changes, carbon allocation and neutral lipid accumulation in oleaginous microalgae. *Bioresour. Technol.*, 191:219–228, September 2015.

[91] Qiang Liao, Hai-Xing Chang, Qian Fu, Yun Huang, Ao Xia, Xun Zhu, and Nianbing Zhong. Physiological-phased kinetic characteristics of microalgae Chlorella vulgaris growth and lipid synthesis considering synergistic effects of light, carbon and nutrients. *Bioresour. Technol.*, 250:583–590, February 2018.

[92] Giuseppe Torzillo and Avigad Vonshak. Environmental stress physiology with reference to mass cultures. In *Handbook of Microalgal Culture*, pages 90–113. John Wiley & Sons, Ltd, Oxford, UK, April 2013.

[93] Gabriela F Ferreira, Luisa F Rios Pinto, Patricia O Carvalho, Mirela B Coelho, Marcos N Eberlin, Rubens Maciel Filho, and Fregolente V Leonardo. Biomass and lipid characterization of microalgae genera Botryococcus, Chlorella, and Desmodesmus aiming high-value fatty acid production. *BIOMASS CONVERSION AND BIOREFINERY*, 2019.

[94] Ling Xia, Rong Huang, Yinta Li, and Shaoxian Song. The effect of growth phase on the surface properties of three oleaginous microalgae (Botryococcus sp. FACGB-762, Chlorella sp. XJ-445 and Desmodesmus bijugatus XJ-231). *PLoS One*, 12(10):e0186434, October 2017.

[95] L D Zhu, Z H Li, and E Hiltunen. Strategies for Lipid Production Improvement in Microalgae as a Biodiesel Feedstock. *Biomed Res. Int.*, 2016:8792548, 2016.

[96] Dawei Yang, Donghui Song, Tobias Kind, Yan Ma, Jens Hoefkens, and Oliver Fiehn. Lipidomic Analysis of Chlamydomonas reinhardtii under Nitrogen and Sulfur Deprivation. *PLoS One*, 10(9):e0137948, September 2015.

[97] Zhi-Yan Du and Christoph Benning. Triacylglycerol Accumulation in Photosynthetic Cells in Plants and Algae. *Subcell. Biochem.*, 86:179–205, 2016.

[98] Atsushi Sato, Rie Matsumura, Naomi Hoshino, Mikio Tsuzuki, and Norihiro Sato. Responsibility of regulatory gene expression and repressed protein synthesis for triacylglycerol accumulation on sulfur-starvation in Chlamydomonas reinhardtii. *Front. Plant Sci.*, 5:444, September 2014.

[99] Emilio Fernandez and Aurora Galvan. Nitrate assimilation in Chlamydomonas. *Eukaryot. Cell*, 7(4):555–559, April 2008.

[100] Nanette R Boyle and John A Morgan. Flux balance analysis of primary metabolism in Chlamydomonas reinhardtii. *BMC Syst. Biol.*, 3:4, January 2009.

[101] I K Blaby, A G Glaesener, T Mettler, S T Fitz-Gibbon, S D Gallaher, B Liu, N R Boyle, J Kropat, M Stitt, S Johnson, C Benning, M Pellegrini, D Casero, and S S Merchant. Systems-level analysis of nitrogen starvation-induced modifications of carbon metabolism in a Chlamydomonas reinhardtii starchless mutant. *Plant Cell*, 25(11):4305–4323, 2013.

[102] V H Work, R Radakovits, R E Jinkerson, J E Meuser, L G Elliott, D J Vinyard, L M Laurens, G C Dismukes, and M C Posewitz. Increased lipid accumulation in the Chlamydomonas reinhardtii sta7-10 starchless isoamylase mutant and increased carbohydrate synthesis in complemented strains. *Eukaryot. Cell*, 9(8):1251–1261, 2010.

[103] Irina A Guschina and John L Harwood. Lipids and lipid metabolism in eukaryotic algae. *Prog. Lipid Res.*, 45(2):160–186, March 2006.

[104] Zi Teng Wang, Nico Ullrich, Sunjoo Joo, Sabine Waffenschmidt, and Ursula Goodenough. Algal lipid bodies: stress induction, purification, and biochemical characterization in wild-type and starchless Chlamydomonas reinhardtii. *Eukaryot. Cell*, 8(12):1856–1868, December 2009.

[105] Liliana Rodolfi, Graziella Chini Zittelli, Niccolò Bassi, Giulia Padovani, Natascia Biondi, Gimena Bonini, and Mario R Tredici. Microalgae for oil: strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. *Biotechnol. Bioeng.*, 102(1):100–112, January 2009.

[106] N C Martin and U W Goodenough. Gametic differentiation in Chlamydomonas reinhardtii. I. Production of gametes and their fine structure. *J. Cell Biol.*, 67(3):587–605, December 1975.

[107] Sabeeha S Merchant and John D Helmann. Elemental economy: microbial strategies for optimizing growth in the face of nutrient limitation. *Adv. Microb. Physiol.*, 60:91–210, 2012.

[108] Stefan Schmollinger, Timo Mühlhaus, Nanette R Boyle, Ian K Blaby, David Casero, Tabea Mettler, Jeffrey L Moseley, Janette Kropat, Frederik Sommer, Daniela Strenkert, Dorothea Hemme, Matteo Pellegrini, Arthur R Grossman, Mark Stitt, Michael Schroda, and Sabeeha S Merchant. Nitrogen-Sparing Mechanisms in Chlamydomonas Affect the Transcriptome, the Proteome, and Photosynthetic Metabolism. *Plant Cell*, 26(4):1410–1435, April 2014.

[109] Nur A Hidayati, Yui Yamada-Oshima, Masako Iwai, Takashi Yamano, Masataka Kajikawa, Nozomu Sakurai, Kunihiro Suda, Kanami Sesoko, Koichi Hori, Takeshi Obayashi, Mie Shimojima, Hideya Fukuzawa, and Hiroyuki Ohta. Lipid remodeling regulator 1 (LRL1) is differently involved in the phosphorus-depletion response from PSR1 in Chlamydomonas reinhardtii. *Plant J.*, 100(3):610–626, November 2019.

[110] Vered Irihimovitch and Shlomit Yehudai-Resheff. Phosphate and sulfur limitation responses in the chloroplast of Chlamydomonas reinhardtii. *FEMS Microbiol. Lett.*, 283(1):1–8, June 2008.

[111] K M Koo, S Jung, B S Lee, J B Kim, Y D Jo, H I Choi, S Y Kang, G H Chung, W J Jeong, and J W Ahn. The Mechanism of Starch Over-Accumulation in Chlamydomonas reinhardtii High-Starch Mutants Identified by Comparative Transcriptome Analysis. *Front. Microbiol.*, 8:858, 2017.

[112] Singh Himanshu, R Shukla Manish, J Rao Basuthkar, and V R Chary Kandala. Regulation of starch, lipids and amino acids upon nitrogen sensing in Chlamydomonas reinhardtii. *Algal Research*, 18:33–44, September 2016.

[113] Yantao Li, Danxiang Han, Guangrong Hu, Milton Sommerfeld, and Qiang Hu. Inhibition of starch synthesis results in overproduction of lipids in Chlamydomonas reinhardtii. *Biotechnol. Bioeng.*, 107(2):258–268, October 2010.

[114] Yantao Li, Danxiang Han, Guongrong Hu, David Dauvillee, Milton Sommerfeld, Steven Ball, and Qiang Hu. Chlamydomonas starchless mutant defective in ADP-glucose pyrophosphorylase hyper-accumulates triacylglycerol. *Metab. Eng.*, 12(4):387–391, July 2010.

[115] Lisa Van den Broeck, Max Gordon, Dirk Inzé, Cranos Williams, and Rosangela Sozzani. Gene Regulatory Network Inference: Connecting Plant Biology and Mathematical Modeling. *Front. Genet.*, 11:457, May 2020.

[116] Leonardo Bich, Matteo Mossio, Kepa Ruiz-Mirazo, and Alvaro Moreno. Biological regulation: controlling the system from within. *Biol. Philos.*, 31(2):237–265, March 2016.

[117] Federica Cariti, Marie Chazaux, Linnka Lefebvre-Legendre, Paolo Longoni, Bart Ghysels, Xenie Johnson, and Michel Goldschmidt-Clermont. Regulation of Light Harvesting in Chlamydomonas reinhardtii Two Protein Phosphatases Are Involved in State Transitions. *Plant Physiol.*, 183(4):1749–1764, August 2020.

[118] George A Khoury, Richard C Baliban, and Christodoulos A Floudas. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep.*, 1, September 2011.

[119] Nathalie Van den Koornhuyse, Nathalie Libessart, Brigitte Delrue, Christophe Zabawinski, André Decq, Alberto Iglesias, Anne Carton, Jack Preiss, and Steven Ball. Control of Starch Composition and Structure through Substrate Supply in the Monocellular Alga Chlamydomonas reinhardtii*. *J. Biol. Chem.*, 271(27):16281–16287, July 1996.

[120] C Zabawinski, N Van Den Koornhuyse, C D'Hulst, R Schlichting, C Giersch, B Delrue, J M Lacroix, J Preiss, and S Ball. Starchless mutants of Chlamydomonas reinhardtii lack the small subunit of a heterotetrameric ADP-glucose pyrophosphorylase. *J. Bacteriol.*, 183(3):1069–1077, February 2001.

[121] M S Patel and R A Harris. Metabolic Regulation. In Ralph A Bradshaw and Philip D Stahl, editors, *Encyclopedia of Cell Biology*, pages 288–297. Academic Press, Waltham, January 2016.

[122] H Efsun Arda and Albertha J M Walhout. Gene-centered regulatory networks. *Brief. Funct. Genomics*, 9(1):4–12, January 2010.

[123] Diego Mauricio Riaño-Pachón, Slobodan Ruzicic, Ingo Dreyer, and Bernd Mueller-Roeber. PlnTFDB: an integrative plant transcription factor database. *BMC Bioinformatics*, 8:42, February 2007.

[124] Oliver J Rando and Howard Y Chang. Genome-wide views of chromatin structure. *Annu. Rev. Biochem.*, 78:245–271, 2009.

[125] D St Johnston and C Nüsslein-Volhard. The origin of pattern and polarity in the Drosophila embryo. *Cell*, 68(2):201–219, January 1992.

[126] K Goto and E M Meyerowitz. Function and regulation of the Arabidopsis floral homeotic gene PISTILLATA. *Genes Dev.*, 8(13):1548–1560, July 1994.

[127] T Jack, L L Brockman, and E M Meyerowitz. The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens. *Cell*, 68(4):683–697, February 1992.

[128] Rebecca S Lamb and Vivian F Irish. Functional divergence within the *APETALA3/PISTILLATA* floral homeotic gene lineages. *Proc. Natl. Acad. Sci. U. S. A.*, 100(11):6558–6563, May 2003.

[129] Nishikant Wase, Paul N Black, Bruce A Stanley, and Concetta C DiRusso. Integrated quantitative analysis of nitrogen stress response in Chlamydomonas reinhardtii using metabolite and protein profiling. *J. Proteome Res.*, 13(3):1373–1396, March 2014.

[130] Ursula Goodenough, Ian Blaby, David Casero, Sean D Gallaher, Carrie Goodson, Shannon Johnson, Jae-Hyeok Lee, Sabeeha S Merchant, Matteo Pellegrini, Robyn Roth, Jannette Rusch, Manmilan Singh, James G Umen, Taylor L Weiss, and Tuya Wulan. The path to triacylglyceride obesity in the sta6 strain of Chlamydomonas reinhardtii. *Eukaryot. Cell*, 13(5):591–613, May 2014.

[131] Mahmoud Gargouri, Philip D Bates, Jeong-Jin Park, Helmut Kirchhoff, and David R Gang. Functional photosystem I maintains proper energy balance during nitrogen depletion in Chlamydomonas reinhardtii, promoting triacylglycerol accumulation. *Biotechnol. Biofuels*, 10:89, April 2017.

[132] Chia-Hong Tsai, Sahra Uygun, Rebecca Roston, Shin-Han Shiu, and Christoph Benning. Recovery from N Deprivation Is a Transcriptionally and Functionally Distinct State in Chlamydomonas. *Plant Physiol.*, 176(3):2007–2023, March 2018.

[133] Yasuyo Yamaoka, Bae Young Choi, Hanul Kim, Seungjun Shin, Yeongho Kim, Sunghoon Jang, Won-Yong Song, Chung H Cho, Hwan Su Yoon, Kenji Kohno, and Youngsook Lee. Identification and functional study of the endoplasmic reticulum stress sensor IRE1 in Chlamydomonas reinhardtii. *Plant J.*, 94(1):91–104, April 2018.

[134] Yasuyo Yamaoka, Seungjun Shin, Bae Young Choi, Hanul Kim, Sunghoon Jang, Masataka Kajikawa, Takashi Yamano, Fantao Kong, Bertrand Légeret, Hideya Fukuzawa, Yonghua Li-Beisson, and Youngsook Lee. The bZIP1 Transcription Factor Regulates Lipid Remodeling and Contributes to ER Stress Management in Chlamydomonas reinhardtii. *Plant Cell*, 31(5):1127–1140, March 2019.

[135] Fan Bai, Yu Zhang, and Jin Liu. A bZIP transcription factor is involved in regulating lipid and pigment metabolisms in the green alga Chlamydomonas reinhardtii. *Algal Research*, 59:102450, November 2021.

[136] D D Wykoff, A R Grossman, D P Weeks, H Usuda, and K Shimogawara. Psr1, a nuclear localized protein that regulates phosphorus metabolism in Chlamydomonas. *Proc. Natl. Acad. Sci. U. S. A.*, 96(26):15336–15341, December 1999.

[137] K Shimogawara, D D Wykoff, H Usuda, and A R Grossman. Chlamydomonas reinhardtii mutants abnormal in their responses to phosphorus deprivation. *Plant Physiol.*, 120(3):685–694, July 1999.

[138] C Y Ngan, C H Wong, C Choi, Y Yoshinaga, K Louie, J Jia, C Chen, B Bowen, H Cheng, L Leonelli, R Kuo, R Baran, J G Garcia-Cerdan, A Pratap, M Wang, J Lim, H Tice, C Daum, J Xu, T Northen, A Visel, J Bristow, K K Niyogi, and C L Wei. Lineage-specific chromatin signatures reveal a regulator of lipid metabolism in microalgae. *Nat Plants*, 1:15107, 2015.

[139] Amit K Bajhaiya, Andrew P Dean, Leo A H Zeef, Rachel E Webster, and Jon K Pittman. PSR1 Is a Global Transcriptional Regulator of Phosphorus Deficiency Responses and Carbon Storage Metabolism in Chlamydomonas reinhardtii. *Plant Physiol.*, 170(3):1216–1234, March 2016.

[140] Bae Young Choi, Donghwan Shim, Fantao Kong, Pascaline Auroy, Yuree Lee, Yonghua Li-Beisson, Youngsook Lee, and Yasuyo Yamaoka. The Chlamydomonas transcription factor MYB1 mediates lipid accumulation under nitrogen depletion. *New Phytol.*, 235(2):595–610, July 2022.

[141] Meicheng Shi, Lihua Yu, Jianan Shi, and Jin Liu. A conserved MYB transcription factor is involved in regulating lipid metabolic pathways for oil biosynthesis in green algae. *New Phytol.*, March 2022.

[142] Masataka Kajikawa, Yuri Sawaragi, Haruka Shinkawa, Takashi Yamano, Akira Ando, Misako Kato, Masafumi Hirono, Naoki Sato, and Hideya Fukuzawa. Algal dual-specificity tyrosine phosphorylation-regulated kinase, triacylglycerol accumulation regulator1, regulates accumulation of triacylglycerol in nitrogen or sulfur deficiency. *Plant Physiol.*, 168(2):752–764, June 2015.

[143] Haruka Shinkawa, Masataka Kajikawa, Yuko Nomura, Mayu Ogura, Yuri Sawaragi, Takashi Yamano, Hirofumi Nakagami, Naoyuki Sugiyama, Yasushi Ishihama, Yu Kanesaki, Hirofumi Yoshikawa, and Hideya Fukuzawa. Algal Protein Kinase, Triacylglycerol Accumulation Regulator 1, Modulates Cell Viability and Gametogenesis in Carbon/Nitrogen-Imbalanced Conditions. *Plant Cell Physiol.*, 60(4):916–930, April 2019.

[144] Miriam Schulz-Raffelt, Vincent Chochois, Pascaline Auroy, Stéphan Cuiné, Emmanuelle Billon, David Dauvillée, Yonghua Li-Beisson, and Gilles Peltier. Hyper-accumulation of starch and oil in a Chlamydomonas mutant affected in a plant-specific DYRK kinase. *Biotechnol. Biofuels*, 9:55, March 2016.

[145] J P Davies, F H Yildiz, and A Grossman. Sac1, a putative regulator that is critical for survival of Chlamydomonas reinhardtii during sulfur deprivation. *EMBO J.*, 15(9):2150–2159, May 1996.

[146] Jeffrey L Moseley, David Gonzalez-Ballester, Wirulda Pootakham, Shaun Bailey, and Arthur R Grossman. Genetic interactions between regulators of Chlamydomonas phosphorus and sulfur deprivation responses. *Genetics*, 181(3):889–905, March 2009.

[147] Koichi Sugimoto, Norihiro Sato, and Mikio Tsuzuki. Utilization of a chloroplast membrane sulfolipid as a major internal sulfur source for protein synthesis in the early phase of sulfur starvation in Chlamydomonas reinhardtii. *FEBS Lett.*, 581(23):4519–4522, September 2007.

[148] Koichi Sugimoto, Mikio Tsuzuki, and Norihiro Sato. Regulation of synthesis and degradation of a sulfolipid under sulfur-starved conditions and its physiological significance in Chlamydomonas reinhardtii. *New Phytol.*, 185(3):676–686, February 2010.

[149] Pascal Braun. Interactome mapping for analysis of complex phenotypes: insights from benchmarking binary interaction assays. *Proteomics*, 12(10):1499–1518, May 2012.

[150] Melina Altmann, Stefan Altmann, Patricia A Rodriguez, Benjamin Weller, Lena Elorduy Vergara, Julius Palme, Nora Marín-de la Rosa, Mayra Sauer, Marion Wenig, José Antonio Villaécija-Aguilar, Jennifer Sales, Chung-Wen Lin, Ramakrishnan Pandiarajan, Veronika Young, Alexandra Strobel, Lisa Gross, Samy Carbonnel, Karl G Kugler, Antoni Garcia-Molina, George W Bassel, Claudia Falter, Klaus F X Mayer, Caroline Gutjahr, A Corina Vlot, Erwin Grill, and Pascal Falter-Braun. Extensive signal integration by the phytohormone protein network. *Nature*, 583(7815):271–276, July 2020.

[151] Katja Luck, Dae-Kyum Kim, Luke Lambourne, Kerstin Spirohn, Bridget E Begg, Wenting Bian, Ruth Brignall, Tiziana Cafarelli, Francisco J Campos-Laborie, Benoit Charloteaux, Dongsic Choi, Atina G Coté, Meaghan Daley, Steven Deimling, Alice Desbuleux, Amélie Dricot, Marinella Gebbia, Madeleine F Hardy, Nishka Kishore, Jennifer J Knapp, István A Kovács, Irma Lemmens, Miles W Mee, Joseph C Mellor, Carl Pollis, Carles Pons, Aaron D Richardson, Sadie Schlabach, Bridget Teeking, Anupama Yadav, Mariana Babor, Dawit Balcha, Omer Basha, Christian Bowman-Colin, Suet-Feung Chin, Soon Gang Choi, Claudia Colabella, Georges Coppin, Cassandra D'Amata, David De Ridder, Steffi De Rouck, Miquel Duran-Frigola, Hanane Ennajdaoui, Florian Goebels, Liana Goehring, Anjali Gopal, Ghazal Haddad, Elodie Hatchi, Mohamed Helmy, Yves Jacob, Yoseph Kassa, Serena Landini, Roujia Li, Natascha van Lieshout, Andrew MacWilliams, Dylan Markey, Joseph N Paulson, Sudharshan Rangarajan, John Rasla, Ashyad Rayhan, Thomas Rolland, Adriana San-Miguel, Yun Shen, Dayag Sheykhkarimli, Gloria M Sheynkman, Eyal Simonovsky, Murat Taşan, Alexander Tejeda, Vincent Tropepe, Jean-Claude Twizere, Yang Wang, Robert J Weatheritt, Jochen Weile, Yu Xia, Xinping Yang, Esti Yeger-Lotem, Quan Zhong, Patrick Aloy, Gary D Bader, Javier De Las Rivas, Suzanne Gaudet, Tong Hao, Janusz Rak, Jan Tavernier, David E Hill, Marc Vidal, Frederick P Roth, and Michael A Calderwood. A reference map of the human binary protein interactome. *Nature*, 580(7803):402–408, April 2020.

[152] Thomas Rolland, Murat Taşan, Benoit Charloteaux, Samuel J Pevzner, Quan Zhong, Nidhi Sahni, Song Yi, Irma Lemmens, Celia Fontanillo, Roberto Mosca, Atanas Kamburov, Susan D Ghiassian, Xinping Yang, Lila Ghamsari, Dawit Balcha, Bridget E Begg, Pascal Braun, Marc Brehme, Martin P Broly, Anne-Ruxandra Carvunis, Dan Convery-Zupan, Roser Corominas, Jasmin Coulombe-Huntington, Elizabeth Dann, Matija Dreze, Amélie Dricot, Changyu Fan, Eric Franzosa, Fana Gebreab, Bryan J Gutierrez, Madeleine F Hardy, Mike Jin, Shuli Kang, Ruth Kiros, Guan Ning Lin, Katja Luck, Andrew MacWilliams, Jörg Menche, Ryan R Murray, Alexandre Palagi, Matthew M Poulin, Xavier Rambout, John Rasla, Patrick Reichert, Viviana Romero, Elien Ruyssinck, Julie M Sahalie, Annemarie Scholz, Akash A Shah, Amitabh Sharma, Yun Shen, Kerstin Spirohn, Stanley Tam, Alexander O Tejeda, Shelly A Trigg, Jean-Claude Twizere, Kerwin Vega, Jennifer Walsh, Michael E Cusick, Yu Xia, Albert-László Barabási, Lilia M Iakoucheva, Patrick Aloy, Javier De Las Rivas, Jan Tavernier, Michael A Calderwood, David E Hill, Tong Hao, Frederick P Roth, and Marc Vidal. A proteome-scale map of the human interactome network. *Cell*, 159(5):1212–1226, November 2014.

[153] Haiyuan Yu, Pascal Braun, Muhammed A Yildirim, Irma Lemmens, Kavitha Venkatesan, Julie Sahalie, Tomoko Hirozane-Kishikawa, Fana Gebreab, Na Li, Nicolas Simonis, Tong

Hao, Jean-François Rual, Amélie Dricot, Alexei Vazquez, Ryan R Murray, Christophe Simon, Leah Tardivo, Stanley Tam, Nenad Svrzikapa, Changyu Fan, Anne-Sophie de Smet, Adriana Motyl, Michael E Hudson, Juyong Park, Xiaofeng Xin, Michael E Cusick, Troy Moore, Charlie Boone, Michael Snyder, Frederick P Roth, Albert-László Barabási, Jan Tavernier, David E Hill, and Marc Vidal. High-quality binary protein interaction map of the yeast interactome network. *Science*, 322(5898):104–110, October 2008.

[154] S Fields and O Song. A novel genetic system to detect protein-protein interactions. *Nature*, 340(6230):245–246, July 1989.

[155] S J Fashena, I Serebriiskii, and E A Golemis. The continued evolution of two-hybrid screening approaches in yeast: how to outwit different preys with different baits. *Gene*, 250(1-2):1–14, May 2000.

[156] Nozomu Yachie, Evangelia Petsalaki, Joseph C Mellor, Jochen Weile, Yves Jacob, Marta Verby, Sedide B Ozturk, Siyang Li, Atina G Cote, Roberto Mosca, Jennifer J Knapp, Minjeong Ko, Analyn Yu, Marinella Gebbia, Nidhi Sahni, Song Yi, Tanya Tyagi, Dayag Sheykhkarimli, Jonathan F Roth, Cassandra Wong, Louai Musa, Jamie Snider, Yi-Chun Liu, Haiyuan Yu, Pascal Braun, Igor Stagljar, Tong Hao, Michael A Calderwood, Laurence Pelletier, Patrick Aloy, David E Hill, Marc Vidal, and Frederick P Roth. Pooled-matrix protein interaction screens using Barcode Fusion Genetics. *Mol. Syst. Biol.*, 12(4):863, April 2016.

[157] Lifeasible. The Yeast Two-Hybrid Assay in Plants. `https://www.lifeasible.com/custom-solutions/plant/analytical-services/gene-function-analysis/the-yeast-two-hybrid-assay-in-plants/`. Accessed: 2022-10-5.

[158] M Shahid Mukhtar, Anne-Ruxandra Carvunis, Matija Dreze, Petra Epple, Jens Steinbrenner, Jonathan Moore, Murat Tasan, Mary Galli, Tong Hao, Marc T Nishimura, Samuel J Pevzner, Susan E Donovan, Lila Ghamsari, Balaji Santhanam, Viviana Romero, Matthew M Poulin, Fana Gebreab, Bryan J Gutierrez, Stanley Tam, Dario Monachello, Mike Boxem, Christopher J Harbort, Nathan McDonald, Lantian Gai, Huaming Chen, Yijian He, European Union Effectoromics Consortium, Jean Vandenhaute, Frederick P Roth, David E Hill, Joseph R Ecker, Marc Vidal, Jim Beynon, Pascal Braun, and Jeffery L Dangl. Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science*, 333(6042):596–601, July 2011.

[159] Shelley Lumba, Shigeo Toh, Louis-François Handfield, Michael Swan, Raymond Liu, Ji-Young Youn, Sean R Cutler, Rajagopal Subramaniam, Nicholas Provart, Alan Moses, Darrell Desveaux, and Peter McCourt. A mesoscale abscisic acid hormone interactome reveals a dynamic signaling landscape in Arabidopsis. *Dev. Cell*, 29(3):360–372, May 2014.

[160] Jean-François Rual, Kavitha Venkatesan, Tong Hao, Tomoko Hirozane-Kishikawa, Amélie Dricot, Ning Li, Gabriel F Berriz, Francis D Gibbons, Matija Dreze, Nono Ayivi-Guedehoussou, Niels Klitgord, Christophe Simon, Mike Boxem, Stuart Milstein, Jennifer Rosenberg, Debra S Goldberg, Lan V Zhang, Sharyl L Wong, Giovanni Franklin, Siming Li, Joanna S Albala, Janghoo Lim, Carlene Fraughton, Estelle Llamosas, Sebiha Cevik,

Camille Bex, Philippe Lamesch, Robert S Sikorski, Jean Vandenhaute, Huda Y Zoghbi, Alex Smolyar, Stephanie Bosak, Reynaldo Sequerra, Lynn Doucette-Stamm, Michael E Cusick, David E Hill, Frederick P Roth, and Marc Vidal. Towards a proteome-scale map of the human protein-protein interaction network. *Nature*, 437(7062):1173–1178, October 2005.

[161] Ulrich Stelzl, Uwe Worm, Maciej Lalowski, Christian Haenig, Felix H Brembeck, Heike Goehler, Martin Stroedicke, Martina Zenkner, Anke Schoenherr, Susanne Koeppen, Jan Timm, Sascha Mintzlaff, Claudia Abraham, Nicole Bock, Silvia Kietzmann, Astrid Goedde, Engin Toksöz, Anja Droege, Sylvia Krobitsch, Bernhard Korn, Walter Birchmeier, Hans Lehrach, and Erich E Wanker. A human protein-protein interaction network: a resource for annotating the proteome. *Cell*, 122(6):957–968, September 2005.

[162] Dae-Kyum Kim, Benjamin Weller, Chung-Wen Lin, Dayag Sheykhkarimli, Jennifer J Knapp, Nishka Kishore, Mayra Sauer, Ashyad Rayhan, Veronika Young, Nora Marinde la Rosa, Oxana Pogoutse, Kerstin Spirohn, Alexandra Strobel, Florent Laval, Patrick Schwehn, Roujia Li, Simin Rothballer, Melina Altmann, Patricia Cassonnet, Guillaume Dugied, Atina G Cote, Lena Elorduy Vergara, Isaiah Hazelwood, Bingruo B Liu, Maria Nguyen, Ramakrishnan Pandiarajan, Patricia A Rodriguez Coloma, Luc Willems, Jean-Claude Twizere, Caroline Demeret, Yves Jacob, Tong Hao, Dave E Hill, Claudia Falter, Marc Vidal, Michael A Calderwood, Frederick P Roth, and Pascal Falter-Braun. A map of binary SARS-CoV-2 protein interactions implicates host immune regulation and ubiquitination. March 2021.

[163] Dae-Kyum Kim, Benjamin Weller, Chung-Wen Lin, Dayag Sheykhkarimli, Jennifer J Knapp, Guillaume Dugied, Andreas Zanzoni, Carles Pons, Marie J Tofaute, Sibusiso B Maseko, Kerstin Spirohn, Florent Laval, Luke Lambourne, Nishka Kishore, Ashyad Rayhan, Mayra Sauer, Veronika Young, Hridi Halder, Nora Marín-de la Rosa, Oxana Pogoutse, Alexandra Strobel, Patrick Schwehn, Roujia Li, Simin T Rothballer, Melina Altmann, Patricia Cassonnet, Atina G Coté, Lena Elorduy Vergara, Isaiah Hazelwood, Betty B Liu, Maria Nguyen, Ramakrishnan Pandiarajan, Bushra Dohai, Patricia A Rodriguez Coloma, Juline Poirson, Paolo Giuliana, Luc Willems, Mikko Taipale, Yves Jacob, Tong Hao, David E Hill, Christine Brun, Jean-Claude Twizere, Daniel Krappmann, Matthias Heinig, Claudia Falter, Patrick Aloy, Caroline Demeret, Marc Vidal, Michael A Calderwood, Frederick P Roth, and Pascal Falter-Braun. A proteome-scale map of the SARS-CoV-2-human contactome. *Nat. Biotechnol.*, 41(1):140–149, January 2023.

[164] Joseph Swift and Gloria M Coruzzi. A matter of time - How transient transcription factor interactions create dynamic gene regulatory networks. *Biochim. Biophys. Acta Gene Regul. Mech.*, 1860(1):75–83, January 2017.

[165] Caihuan Tian, Xiaoni Zhang, Jun He, Haopeng Yu, Ying Wang, Bihai Shi, Yingying Han, Guoxun Wang, Xiaoming Feng, Cui Zhang, Jin Wang, Jiyan Qi, Rong Yu, and Yuling Jiao. An organ boundary-enriched gene regulatory network uncovers regulatory hierarchies underlying axillary meristem initiation. *Mol. Syst. Biol.*, 10:755, October 2014.

[166] Wolfgang Busch, Andrej Miotk, Federico D Ariel, Zhong Zhao, Joachim Forner, Gabor Daum, Takuya Suzaki, Christoph Schuster, Sebastian J Schultheiss, Andrea Leibfried, Silke

Haubeiss, Nati Ha, Raquel L Chan, and Jan U Lohmann. Transcriptional control of a plant stem cell niche. *Dev. Cell*, 18(5):849–861, May 2010.

[167] Momoko Ikeuchi, Michitaro Shibata, Bart Rymen, Akira Iwase, Anne-Maarit Bågman, Lewis Watt, Duncan Coleman, David S Favero, Tatsuya Takahashi, Sebastian E Ahnert, Siobhan M Brady, and Keiko Sugimoto. A Gene Regulatory Network for Cellular Reprogramming in Plant Regeneration. *Plant Cell Physiol.*, 59(4):765–777, April 2018.

[168] Bart Deplancke, Arnab Mukhopadhyay, Wanyuan Ao, Ahmed M Elewa, Christian A Grove, Natalia J Martinez, Reynaldo Sequerra, Lynn Doucette-Stamm, John S Reece-Hoyes, Ian A Hope, Heidi A Tissenbaum, Susan E Mango, and Albertha J M Walhout. A gene-centered C. elegans protein-DNA interaction network. *Cell*, 125(6):1193–1205, June 2006.

[169] H Efsun Arda, Stefan Taubert, Lesley T MacNeil, Colin C Conine, Ben Tsuda, Marc Van Gilst, Reynaldo Sequerra, Lynn Doucette-Stamm, Keith R Yamamoto, and Albertha J M Walhout. Functional modularity of nuclear hormone receptors in a Caenorhabditis elegans metabolic gene regulatory network. *Mol. Syst. Biol.*, 6:367, May 2010.

[170] Baohua Li, Allison Gaudinier, Michelle Tang, Mallorie Taylor-Teeples, Ngoc T Nham, Cyrus Ghaffari, Darik Scott Benson, Margaret Steinmann, Jennifer A Gray, Siobhan M Brady, and Daniel J Kliebenstein. Promoter-based integration in plant defense regulation. *Plant Physiol.*, 166(4):1803–1820, December 2014.

[171] M Taylor-Teeples, L Lin, M de Lucas, G Turco, T W Toal, A Gaudinier, N F Young, G M Trabucco, M T Veling, R Lamothe, P P Handakumbura, G Xiong, C Wang, J Corwin, A Tsoukalas, L Zhang, D Ware, M Pauly, D J Kliebenstein, K Dehesh, I Tagkopoulos, G Breton, J L Pruneda-Paz, S E Ahnert, S A Kay, S P Hazen, and S M Brady. An Arabidopsis gene regulatory network for secondary cell wall synthesis. *Nature*, 517(7536):571–575, January 2015.

[172] Nick W Albert, Kevin M Davies, David H Lewis, Huaibi Zhang, Mirco Montefiori, Cyril Brendolise, Murray R Boase, Hanh Ngo, Paula E Jameson, and Kathy E Schwinn. A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. *Plant Cell*, 26(3):962–980, March 2014.

[173] Siobhan M Brady, Lifang Zhang, Molly Megraw, Natalia J Martinez, Eric Jiang, Charles S Yi, Weilin Liu, Anna Zeng, Mallorie Taylor-Teeples, Dahae Kim, Sebastian Ahnert, Uwe Ohler, Doreen Ware, Albertha J M Walhout, and Philip N Benfey. A stele-enriched gene regulatory network in the Arabidopsis root. *Mol. Syst. Biol.*, 7:459, January 2011.

[174] Fan Yang, Wei Li, Nan Jiang, Haidong Yu, Kengo Morohashi, Wilberforce Zachary Ouma, Daniel E Morales-Mantilla, Fabio Andres Gomez-Cano, Eric Mukundi, Luis Daniel Prada-Salcedo, Roberto Alers Velazquez, Jasmin Valentin, Maria Katherine Mejía-Guerra, John Gray, Andrea I Doseff, and Erich Grotewold. A Maize Gene Regulatory Network for Phenolic Metabolism. *Mol. Plant*, 10(3):498–515, March 2017.

[175] Allison Gaudinier, Joel Rodriguez-Medina, Lifang Zhang, Andrew Olson, Christophe Liseron-Monfils, Anne-Maarit Bågman, Jessica Foret, Shane Abbitt, Michelle Tang, Baohua Li, Daniel E Runcie, Daniel J Kliebenstein, Bo Shen, Mary J Frank, Doreen Ware, and Siobhan M Brady. Transcriptional regulation of nitrogen-associated metabolism and growth. *Nature*, 563(7730):259–264, November 2018.

[176] Michelle Tang, Baohua Li, Xue Zhou, Tayah Bolt, Jia Jie Li, Neiman Cruz, Allison Gaudinier, Richard Ngo, Caitlin Clark-Wiest, Daniel J Kliebenstein, and Siobhan M Brady. A genome-scale TF-DNA interaction network of transcriptional regulation of Arabidopsis primary and specialized metabolism. *Mol. Syst. Biol.*, 17(11):e10625, November 2021.

[177] John S Reece-Hoyes, A Rasim Barutcu, Rachel Patton McCord, Jun Seop Jeong, Lizhi Jiang, Andrew MacWilliams, Xinping Yang, Kourosh Salehi-Ashtiani, David E Hill, Seth Blackshaw, Heng Zhu, Job Dekker, and Albertha J M Walhout. Yeast one-hybrid assays for gene-centered human gene regulatory network mapping. *Nat. Methods*, 8(12):1050–1052, October 2011.

[178] Steve Horvath. *Weighted Network Analysis: Application in Genomics and Systems Biology. ISBN*. Springer New York Dordrecht Heidelberg London, 2011.

[179] Shubhada R Kulkarni, Dries Vaneechoutte, Jan Van de Velde, and Klaas Vandepoele. TF2Network: predicting transcription factor regulators and gene regulatory networks in Arabidopsis using publicly available binding site information. *Nucleic Acids Res.*, 46(6):e31, April 2018.

[180] F J Romero-Campero, I Perez-Hurtado, E Lucas-Reina, J M Romero, and F Valverde. ChlamyNET: a Chlamydomonas gene co-expression network reveals global properties of the transcriptome and the early setup of key co-expression patterns in the green lineage. *BMC Genomics*, 17:227, 2016.

[181] Sunjoo Joo, Yoshiki Nishimura, Evan Cronmiller, Ran Ha Hong, Thamali Kariyawasam, Ming Hsiu Wang, Nai Chun Shao, Saif-El-Din El Akkad, Takamasa Suzuki, Tetsuya Higashiyama, Eonseon Jin, and Jae-Hyeok Lee. Gene Regulatory Networks for the Haploid-to-Diploid Transition of Chlamydomonas reinhardtii. *Plant Physiol.*, 175(1):314–332, September 2017.

[182] Mahmoud Gargouri, Jeong-Jin Park, F Omar Holguin, Min-Jeong Kim, Hongxia Wang, Rahul R Deshpande, Yair Shachar-Hill, Leslie M Hicks, and David R Gang. Identification of regulatory network hubs that control lipid metabolism in Chlamydomonas reinhardtii. *J. Exp. Bot.*, 66(15):4551–4566, August 2015.

[183] Wei Fang, Yaqing Si, Stephen Douglass, David Casero, Sabeeha S Merchant, Matteo Pellegrini, Istvan Ladunga, Peng Liu, and Martin H Spalding. Transcriptome-wide changes in Chlamydomonas reinhardtii gene expression regulated by carbon dioxide and the CO2-concentrating mechanism regulator CIA5/CCM1. *Plant Cell*, 24(5):1876–1893, May 2012.

[184] Andrew J Brueggeman, Dayananda S Gangadharaiah, Matyas F Cserhati, David Casero, Donald P Weeks, and Istvan Ladunga. Activation of the carbon concentrating mechanism by CO2 deprivation coincides with massive transcriptional restructuring in Chlamydomonas reinhardtii. *Plant Cell*, 24(5):1860–1875, May 2012.

[185] Marius Arend, Yizhong Yuan, Nooshin Omranian, Zoran Nikoloski, and Dimitris Petroutsos. Widening the landscape of transcriptional regulation of algal photoprotection. February 2022.

[186] Flavia Vischi Winck, Samuel Arvidsson, Diego Mauricio Riaño-Pachón, Sabrina Hempel, Aneta Koseska, Zoran Nikoloski, David Alejandro Urbina Gomez, Jens Rupprecht, and Bernd Mueller-Roeber. Genome-wide identification of regulatory elements and reconstruction of gene regulatory networks of the green alga Chlamydomonas reinhardtii under carbon deprivation. *PLoS One*, 8(11):e79909, November 2013.

[187] Matthew T Weirauch, Ally Yang, Mihai Albu, Atina G Cote, Alejandro Montenegro-Montero, Philipp Drewe, Hamed S Najafabadi, Samuel A Lambert, Ishminder Mann, Kate Cook, Hong Zheng, Alejandra Goity, Harm van Bakel, Jean-Claude Lozano, Mary Galli, Mathew G Lewsey, Eryong Huang, Tuhin Mukherjee, Xiaoting Chen, John S Reece-Hoyes, Sridhar Govindarajan, Gad Shaulsky, Albertha J M Walhout, François-Yves Bouget, Gunnar Ratsch, Luis F Larrondo, Joseph R Ecker, and Timothy R Hughes. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell*, 158(6):1431–1443, September 2014.

[188] Michael F Berger and Martha L Bulyk. Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors. *Nat. Protoc.*, 4(3):393–411, 2009.

[189] Anna Bartlett, Ronan C O'Malley, Shao-Shan Carol Huang, Mary Galli, Joseph R Nery, Andrea Gallavotti, and Joseph R Ecker. Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.*, 12(8):1659–1672, August 2017.

[190] Ronan C O'Malley, Shao-Shan Carol Huang, Liang Song, Mathew G Lewsey, Anna Bartlett, Joseph R Nery, Mary Galli, Andrea Gallavotti, and Joseph R Ecker. Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell*, 165(5):1280–1292, May 2016.

[191] B Deplancke, D Dupuy, M Vidal, and A J Walhout. A gateway-compatible yeast one-hybrid system. *Genome Res.*, 14(10B):2093–2101, 2004.

[192] B Deplancke, V Vermeirssen, H E Arda, N J Martinez, and A J Walhout. Gateway-compatible yeast one-hybrid screens. *CSH Protoc.*, 2006(5), 2006.

[193] John S Reece-Hoyes, Alos Diallo, Bryan Lajoie, Amanda Kent, Shaleen Shrestha, Sreenath Kadreppa, Colin Pesyna, Job Dekker, Chad L Myers, and Albertha J M Walhout. Enhanced yeast one-hybrid assays for high-throughput gene-centered regulatory network mapping. *Nat. Methods*, 8(12):1059–1064, October 2011.

[194] Eric Phizicky, Philippe I H Bastiaens, Heng Zhu, Michael Snyder, and Stanley Fields. Protein analysis on a proteomic scale. *Nature*, 422(6928):208–215, March 2003.

[195] Alicia L Richards, Manon Eckhardt, and Nevan J Krogan. Mass spectrometry-based protein-protein interaction networks for the study of human diseases. *Mol. Syst. Biol.*, 17(1):e8792, January 2021.

[196] Jean-François Rual, David E Hill, and Marc Vidal. ORFeome projects: gateway between genomics and omics. *Curr. Opin. Chem. Biol.*, 8(1):20–25, February 2004.

[197] D L Cheo, S A Titus, D R Byrd, J L Hartley, G F Temple, and M A Brasch. Concerted assembly and cloning of multiple DNA segments using in vitro site-specific recombination: functional analysis of multi-segment expression clones. *Genome Res.*, 14(10B):2111–2120, 2004.

[198] Seesandra V Rajagopala, Natsuko Yamamoto, Adrienne E Zweifel, Tomoko Nakamichi, Hsi-Kuang Huang, Jorge David Mendez-Rios, Jonathan Franca-Koh, Meher Preethi Boorgula, Kazutoshi Fujita, Ken-Ichirou Suzuki, James C Hu, Barry L Wanner, Hirotada Mori, and Peter Uetz. The Escherichia coli K-12 ORFeome: a resource for comparative molecular microbiology. *BMC Genomics*, 11:470, August 2010.

[199] Philippe Lamesch, Stuart Milstein, Tong Hao, Jennifer Rosenberg, Ning Li, Reynaldo Sequerra, Stephanie Bosak, Lynn Doucette-Stamm, Jean Vandenhaute, David E Hill, and Marc Vidal. C. elegans ORFeome version 3.1: increasing the coverage of ORFeome resources with improved gene predictions. *Genome Res.*, 14(10B):2064–2069, October 2004.

[200] Arabidopsis Interactome Mapping Consortium. Evidence for network evolution in an Arabidopsis interactome map. *Science*, 333(6042):601–607, July 2011.

[201] Shayne D Wierbowski, Tommy V Vo, Pascal Falter-Braun, Timothy O Jobe, Lars H Kruse, Xiaomu Wei, Jin Liang, Michael J Meyer, Nurten Akturk, Christen A Rivera-Erick, Nicolas A Cordero, Mauricio I Paramo, Elnur E Shayhidin, Marta Bertolotti, Nathaniel D Tippens, Kazi Akther, Rita Sharma, Yuichi Katayose, Kourosh Salehi-Ashtiani, Tong Hao, Pamela C Ronald, Joseph R Ecker, Peter A Schweitzer, Shoshi Kikuchi, Hiroshi Mizuno, David E Hill, Marc Vidal, Gaurav D Moghe, Susan R McCouch, and Haiyuan Yu. A massively parallel barcoded sequencing pipeline enables generation of the first ORFeome and interactome map for rice. *Proc. Natl. Acad. Sci. U. S. A.*, 117(21):11836–11842, May 2020.

[202] Philippe Lamesch, Ning Li, Stuart Milstein, Changyu Fan, Tong Hao, Gabor Szabo, Zhenjun Hu, Kavitha Venkatesan, Graeme Bethel, Paul Martin, Jane Rogers, Stephanie Lawlor, Stuart McLaren, Amélie Dricot, Heather Borick, Michael E Cusick, Jean Vandenhaute, Ian Dunham, David E Hill, and Marc Vidal. hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes. *Genomics*, 89(3):307–315, March 2007.

[203] J L Hartley, G F Temple, and M A Brasch. DNA cloning using in vitro site-specific recombination. *Genome Res.*, 10(11):1788–1795, November 2000.

[204] A J Walhout, G F Temple, M A Brasch, J L Hartley, M A Lorson, S van den Heuvel, and M Vidal. GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol.*, 328:575–592, 2000.

[205] W Henke, K Herdel, K Jung, D Schnorr, and S A Loening. Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acids Res.*, 25(19):3957–3958, October 1997.

[206] P R Winship. An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. *Nucleic Acids Res.*, 17(3):1266, February 1989.

[207] Michael A Jensen, Marilyn Fukushima, and Ronald W Davis. DMSO and betaine greatly improve amplification of GC-rich constructs in de novo synthesis. *PLoS One*, 5(6):e11024, June 2010.

[208] Jonghoon Kang, Myung Soog Lee, and David G Gorenstein. The enhancement of PCR amplification of a random sequence DNA library by DMSO and betaine: application to in vitro combinatorial selection of aptamers. *J. Biochem. Biophys. Methods*, 64(2):147–151, August 2005.

[209] Bei Lv, Yunjia Dai, Ju Liu, Qiang Zhuge, and Dawei Li. The Effect of Dimethyl Sulfoxide on Supercoiled DNA Relaxation Catalyzed by Type I Topoisomerases. *Biomed Res. Int.*, 2015:320490, November 2015.

[210] Todd C Lorenz. Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies. *J. Vis. Exp.*, (63):e3998, May 2012.

[211] Agnieszka Kuffel, Alexander Gray, and Niamh Nic Daeid. Impact of metal ions on PCR inhibition and RT-PCR efficiency. *Int. J. Legal Med.*, 135(1):63–72, January 2021.

[212] Kenneth H Roux. Optimization and troubleshooting in PCR. *Cold Spring Harb. Protoc.*, 2009(4):db.ip66, April 2009.

[213] Henrik Nordberg, Michael Cantor, Serge Dusheyko, Susan Hua, Alexander Poliakov, Igor Shabalov, Tatyana Smirnova, Igor V Grigoriev, and Inna Dubchak. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res.*, 42(Database issue):D26–31, January 2014.

[214] Jinpu Jin, Feng Tian, De-Chang Yang, Yu-Qi Meng, Lei Kong, Jingchu Luo, and Ge Gao. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.*, 45(D1):D1040–D1045, January 2017.

[215] Jinpu Jin, He Zhang, Lei Kong, Ge Gao, and Jingchu Luo. PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Res.*, 42(Database issue):D1182–7, January 2014.

[216] Feng Tian, De-Chang Yang, Yu-Qi Meng, Jinpu Jin, and Ge Gao. PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Res.*, 48(D1):D1104–D1113, January 2020.

[217] Jinpu Jin, Kun He, Xing Tang, Zhe Li, Le Lv, Yi Zhao, Jingchu Luo, and Ge Gao. An Arabidopsis Transcriptional Regulatory Map Reveals Distinct Functional and Evolutionary Features of Novel Transcription Factors. *Mol. Biol. Evol.*, 32(7):1767–1773, July 2015.

[218] D M Riano-Pachon, L G Correa, R Trejos-Espinosa, and B Mueller-Roeber. Green transcription factors: a chlamydomonas overview. *Genetics*, 179(1):31–39, 2008.

[219] Chun-Ping Yu, Jinn-Jy Lin, and Wen-Hsiung Li. Positional distribution of transcription factor binding sites in Arabidopsis thaliana. *Sci. Rep.*, 6:25164, April 2016.

[220] Melissa A Scranton, Joseph T Ostrand, D Ryan Georgianna, Shane M Lofgren, Daphne Li, Rosalie C Ellis, David N Carruthers, Andreas Dräger, David L Masica, and Stephen P Mayfield. Synthetic promoters capable of driving robust nuclear gene expression in the green alga Chlamydomonas reinhardtii. *Algal Research*, 15:135–142, April 2016.

[221] Shaun Mahony, Philip E Auron, and Panayiotis V Benos. DNA familial binding profiles made easy: comparison of various motif alignment and clustering strategies. *PLoS Comput. Biol.*, 3(3):e61, March 2007.

[222] Allison Gaudinier, Lifang Zhang, John S Reece-Hoyes, Mallorie Taylor-Teeples, Li Pu, Zhijie Liu, Ghislain Breton, Jose L Pruneda-Paz, Dahae Kim, Steve A Kay, Albertha J M Walhout, Doreen Ware, and Siobhan M Brady. Enhanced Y1H assays for Arabidopsis. *Nat. Methods*, 8(12):1053–1055, October 2011.

[223] Pascal Braun, Murat Tasan, Matija Dreze, Miriam Barrios-Rodiles, Irma Lemmens, Haiyuan Yu, Julie M Sahalie, Ryan R Murray, Luba Roncari, Anne-Sophie de Smet, Kavitha Venkatesan, Jean-François Rual, Jean Vandenhaute, Michael E Cusick, Tony Pawson, David E Hill, Jan Tavernier, Jeffrey L Wrana, Frederick P Roth, and Marc Vidal. An experimentally derived confidence score for binary protein-protein interactions. *Nat. Methods*, 6(1):91–97, January 2009.

[224] John S Reece-Hoyes, Alos Diallo, Bryan Lajoie, Amanda Kent, Shaleen Shrestha, Sreenath Kadreppa, Colin Pesyna, Job Dekker, Chad L Myers, and Albertha J M Walhout. Enhanced yeast one-hybrid assays for high-throughput gene-centered regulatory network mapping. *Nat. Methods*, 8:1059, 2011.

[225] Maryam Nazarieh, Andreas Wiese, Thorsten Will, Mohamed Hamed, and Volkhard Helms. Identification of key player genes in gene regulatory networks. *BMC Syst. Biol.*, 10(1):1–12, September 2016.

[226] Konstantina Dimitrakopoulou, Charalampos Tsimpouris, George Papadopoulos, Claudia Pommerenke, Esther Wilk, Kyriakos N Sgarbas, Klaus Schughart, and Anastasios Bezerianos. Dynamic gene network reconstruction from gene expression data in mice after influenza A (H1N1) infection. *J. Clin. Bioinforma.*, 1:27, October 2011.

[227] Maciej Jerzy Bernacki, Weronika Czarnocka, Magdalena Szechyńska-Hebda, Ron Mittler, and Stanisław Karpiński. Biotechnological Potential of LSD1, EDS1, and PAD4 in the Improvement of Crops and Industrial Plants. *Plants*, 8(8), August 2019.

[228] Alexey Shapiguzov, Julia P Vainonen, Kerri Hunter, Helena Tossavainen, Arjun Tiwari, Sari Järvi, Maarit Hellman, Fayezeh Aarabi, Saleh Alseekh, Brecht Wybouw, Katrien Van Der Kelen, Lauri Nikkanen, Julia Krasensky-Wrzaczek, Nina Sipari, Markku Keinänen, Esa Tyystjärvi, Eevi Rintamäki, Bert De Rybel, Jarkko Salojärvi, Frank Van Breusegem, Alisdair R Fernie, Mikael Brosché, Perttu Permi, Eva-Mari Aro, Michael Wrzaczek, and Jaakko Kangasjärvi. Arabidopsis RCD1 coordinates chloroplast and mitochondrial functions through interaction with ANAC transcription factors. *Elife*, 8, February 2019.

[229] Albert-László Barabási. Scale-free networks: a decade and beyond. *Science*, 325(5939):412–413, July 2009.

[230] P Shannon, A Markiel, O Ozier, N S Baliga, J T Wang, D Ramage, N Amin, B Schwikowski, and T Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, 13(11):2498–2504, 2003.

[231] M Ashburner, C A Ball, J A Blake, D Botstein, H Butler, J M Cherry, A P Davis, K Dolinski, S S Dwight, J T Eppig, M A Harris, D P Hill, L Issel-Tarver, A Kasarskis, S Lewis, J C Matese, J E Richardson, M Ringwald, G M Rubin, and G Sherlock. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, 25(1):25–29, May 2000.

[232] Uku Raudvere, Liis Kolberg, Ivan Kuzmin, Tambet Arak, Priit Adler, Hedi Peterson, and Jaak Vilo. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.*, 47(W1):W191–W198, July 2019.

[233] I J Kurland and S J Pilkis. Covalent control of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase: insights into autoregulation of a bifunctional enzyme. *Protein Sci.*, 4(6):1023–1037, June 1995.

[234] Wayne R Riekhof, Michael E Ruckle, Todd A Lydic, Barbara B Sears, and Christoph Benning. The sulfolipids 2'-O-acyl-sulfoquinovosyldiacylglycerol and sulfoquinovosyldiacylglycerol are absent from a Chlamydomonas reinhardtii mutant deleted in SQD1. *Plant Physiol.*, 133(2):864–874, October 2003.

[235] Jonathan E Meuser, Sarah D'Adamo, Robert E Jinkerson, Florence Mus, Wenqiang Yang, Maria L Ghirardi, Michael Seibert, Arthur R Grossman, and Matthew C Posewitz. Genetic disruption of both Chlamydomonas reinhardtii [FeFe]-hydrogenases: Insight into the role of HYDA2 in H2 production. *Biochem. Biophys. Res. Commun.*, 417(2):704–709, January 2012.

[236] H V Scheller, P E Jensen, A Haldrup, C Lunde, and J Knoetzel. Role of subunits in eukaryotic Photosystem I. *Biochim. Biophys. Acta*, 1507(1-3):41–60, October 2001.

[237] Beat B Fischer, Régine Dayer, Yvonne Schwarzenbach, Stéphane D Lemaire, Renata Behra, Anja Liedtke, and Rik I L Eggen. Function and regulation of the glutathione peroxidase homologous gene GPXH/GPX5 in Chlamydomonas reinhardtii. *Plant Mol. Biol.*, 71(6):569–583, December 2009.

[238] Xiaocui Ma, Haiyan Wei, Yaodong Zhang, You Duan, Wanting Zhang, Yingyin Cheng, Xiao-Qin Xia, and Mijuan Shi. Glutathione peroxidase 5 deficiency induces lipid metabolism regulated by reactive oxygen species in Chlamydomonas reinhardtii. *Microb. Pathog.*, 147:104358, October 2020.

[239] Olivier Vallon and Martin H Spalding. Chapter 4 - Amino Acid Metabolism. In Elizabeth H Harris, David B Stern, and George B Witman, editors, *The Chlamydomonas Sourcebook (Second Edition)*, pages 115–158. Academic Press, London, January 2009.

[240] Laurie A Boyer, Robert R Latek, and Craig L Peterson. The SANT domain: a unique histone-tail-binding module? *Nat. Rev. Mol. Cell Biol.*, 5(2):158–163, February 2004.

[241] Verandra Kumar, Jitendra K Thakur, and Manoj Prasad. Histone acetylation dynamics regulating plant development and stress responses. *Cell. Mol. Life Sci.*, 78(10):4467–4486, May 2021.

[242] D X Zhou, C Bisanz-Seyer, and R Mache. Molecular cloning of a small DNA binding protein with specificity for a tissue-specific negative element within the rps1 promoter. *Nucleic Acids Res.*, 23(7):1165–1169, April 1995.

[243] Yuan Lin, Fanli Meng, Chao Fang, Bo Zhu, and Jiming Jiang. Rapid validation of transcriptional enhancers using agrobacterium-mediated transient assay. *Plant Methods*, 15:21, March 2019.

[244] Daniel Jaeger, Thomas Baier, and Kyle J Lauersen. Intronserter, an advanced online tool for design of intron containing transgenes. *Algal Research*, 42:101588, September 2019.

[245] Thomas Baier, Julian Wichmann, Olaf Kruse, and Kyle J Lauersen. Intron-containing algal transgenes mediate efficient recombinant gene expression in the green microalga Chlamydomonas reinhardtii. *Nucleic Acids Res.*, 46(13):6909–6919, July 2018.

[246] Ido Golding, Johan Paulsson, Scott M Zawilski, and Edward C Cox. Real-time kinetics of gene activity in individual bacteria. *Cell*, 123(6):1025–1036, December 2005.

[247] Julia Mergner, Martin Frejno, Markus List, Michael Papacek, Xia Chen, Ajeet Chaudhary, Patroklos Samaras, Sandra Richter, Hiromasa Shikata, Maxim Messerer, Daniel Lang, Stefan Altmann, Philipp Cyprys, Daniel P Zolg, Toby Mathieson, Marcus Bantscheff, Rashmi R Hazarika, Tobias Schmidt, Corinna Dawid, Andreas Dunkel, Thomas Hofmann, Stefanie Sprunck, Pascal Falter-Braun, Frank Johannes, Klaus F X Mayer, Gerd Jürgens, Mathias Wilhelm, Jan Baumbach, Erwin Grill, Kay Schneitz, Claus Schwechheimer, and Bernhard Kuster. Mass-spectrometry-based draft of the Arabidopsis proteome. *Nature*, 579(7799):409–414, March 2020.

[248] Ralf Weßling, Petra Epple, Stefan Altmann, Yijian He, Li Yang, Stefan R Henz, Nathan McDonald, Kristin Wiley, Kai Christian Bader, Christine Gläßer, M Shahid Mukhtar, Sabine Haigis, Lila Ghamsari, Amber E Stephens, Joseph R Ecker, Marc Vidal, Jonathan D G Jones, Klaus F X Mayer, Emiel Ver Loren van Themaat, Detlef Weigel, Paul Schulze-Lefert, Jeffery L Dangl, Ralph Panstruga, and Pascal Braun. Convergent targeting of a common host protein-network by pathogen effectors from three kingdoms of life. *Cell Host Microbe*, 16(3):364–375, September 2014.

[249] Daniel M Gelperin, Michael A White, Martha L Wilkinson, Yoshiko Kon, Li A Kung, Kevin J Wise, Nelson Lopez-Hoyo, Lixia Jiang, Stacy Piccirillo, Haiyuan Yu, Mark Gerstein, Mark E Dumont, Eric M Phizicky, Michael Snyder, and Elizabeth J Grayhack. Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Dev.*, 19(23):2816–2826, December 2005.

[250] Rama Balakrishnan, Karen R Christie, Maria C Costanzo, Kara Dolinski, Selina S Dwight, Stacia R Engel, Dianna G Fisk, Jodi E Hirschman, Eurie L Hong, Robert Nash, Rose Oughtred, Marek Skrzypek, Chandra L Theesfeld, Gail Binkley, Qing Dong, Christopher Lane, Anand Sethuraman, Shuai Weng, David Botstein, and J Michael Cherry. Fungal

BLAST and Model Organism BLASTP Best Hits: new comparison resources at the Saccharomyces Genome Database (SGD). *Nucleic Acids Res.*, 33(Database issue):D374–7, January 2005.

[251] J Michael Cherry, Eurie L Hong, Craig Amundsen, Rama Balakrishnan, Gail Binkley, Esther T Chan, Karen R Christie, Maria C Costanzo, Selina S Dwight, Stacia R Engel, Dianna G Fisk, Jodi E Hirschman, Benjamin C Hitz, Kalpana Karra, Cynthia J Krieger, Stuart R Miyasato, Rob S Nash, Julie Park, Marek S Skrzypek, Matt Simison, Shuai Weng, and Edith D Wong. Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res.*, 40(Database issue):D700–5, January 2012.

[252] J M Cherry, C Adler, C Ball, S A Chervitz, S S Dwight, E T Hester, Y Jia, G Juvik, T Roe, M Schroeder, S Weng, and D Botstein. SGD: Saccharomyces Genome Database. *Nucleic Acids Res.*, 26(1):73–79, January 1998.

[253] Stacia R Engel, Fred S Dietrich, Dianna G Fisk, Gail Binkley, Rama Balakrishnan, Maria C Costanzo, Selina S Dwight, Benjamin C Hitz, Kalpana Karra, Robert S Nash, Shuai Weng, Edith D Wong, Paul Lloyd, Marek S Skrzypek, Stuart R Miyasato, Matt Simison, and J Michael Cherry. The reference genome sequence of Saccharomyces cerevisiae: then and now. *G3*, 4(3):389–398, March 2014.

[254] Rory J Craig, Sean D Gallaher, Shengqiang Shu, Patrice Salomé, Jerry W Jenkins, Crysten E Blaby-Haas, Samuel O Purvine, Samuel O'Donnell, Kerrie Barry, Jane Grimwood, Daniela Strenkert, Janette Kropat, Chris Daum, Yuko Yoshinaga, David M Goodstein, Olivier Vallon, Jeremy Schmutz, and Sabeeha S Merchant. The Chlamydomonas Genome Project, version 6: reference assemblies for mating type plus and minus strains reveal extensive structural mutation in the laboratory. June 2022.

[255] Juan I Fuxman Bass, Carles Pons, Lucie Kozlowski, John S Reece-Hoyes, Shaleen Shrestha, Amy D Holdorf, Akihiro Mori, Chad L Myers, and Albertha Jm Walhout. A gene-centered C. elegans protein-DNA interaction network provides a framework for functional predictions. *Mol. Syst. Biol.*, 12(10):884, October 2016.

[256] Shaleen Shrestha, Xing Liu, Clarissa Stephanie Santoso, and Juan Ignacio Fuxman Bass. Enhanced Yeast One-hybrid Screens To Identify Transcription Factor Binding To Human DNA Sequences. *J. Vis. Exp.*, (144), February 2019.

[257] Clarissa S Santoso, Zhaorong Li, Sneha Lal, Samson Yuan, Kok Ann Gan, Luis M Agosto, Xing Liu, Sebastian Carrasco Pro, Jared A Sewell, Andrew Henderson, Maninjay K Atianand, and Juan I Fuxman Bass. Comprehensive mapping of the human cytokine gene regulatory network. *Nucleic Acids Res.*, 48(21):12055–12073, December 2020.

[258] Vanessa Vermeirssen, M Inmaculada Barrasa, César A Hidalgo, Jenny Aurielle B Babon, Reynaldo Sequerra, Lynn Doucette-Stamm, Albert-László Barabási, and Albertha J M Walhout. Transcription factor modularity in a gene-centered C. elegans core neuronal protein-DNA interaction network. *Genome Res.*, 17(7):1061–1071, July 2007.

[259] Juan I Fuxman Bass, Nidhi Sahni, Shaleen Shrestha, Aurian Garcia-Gonzalez, Akihiro Mori, Numana Bhat, Song Yi, David E Hill, Marc Vidal, and Albertha J M Walhout. Human gene-centered transcription factor networks for enhancers and disease variants. *Cell*, 161(3):661–673, April 2015.

[260] Juan I Fuxman Bass, John S Reece-Hoyes, and Albertha J M Walhout. Gene-Centered Yeast One-Hybrid Assays. *Cold Spring Harb. Protoc.*, 2016(12):db.top077669, December 2016.

[261] Kavitha Venkatesan, Jean-François Rual, Alexei Vazquez, Ulrich Stelzl, Irma Lemmens, Tomoko Hirozane-Kishikawa, Tong Hao, Martina Zenkner, Xiaofeng Xin, Kwang-Il Goh, Muhammed A Yildirim, Nicolas Simonis, Kathrin Heinzmann, Fana Gebreab, Julie M Sahalie, Sebiha Cevik, Christophe Simon, Anne-Sophie de Smet, Elizabeth Dann, Alex Smolyar, Arunachalam Vinayagam, Haiyuan Yu, David Szeto, Heather Borick, Amélie Dricot, Niels Klitgord, Ryan R Murray, Chenwei Lin, Maciej Lalowski, Jan Timm, Kirstin Rau, Charles Boone, Pascal Braun, Michael E Cusick, Frederick P Roth, David E Hill, Jan Tavernier, Erich E Wanker, Albert-László Barabási, and Marc Vidal. An empirical framework for binary interactome mapping. *Nat. Methods*, 6(1):83–90, January 2009.

[262] John S Reece-Hoyes, Carles Pons, Alos Diallo, Akihiro Mori, Shaleen Shrestha, Sreenath Kadreppa, Justin Nelson, Stephanie Diprima, Amelie Dricot, Bryan R Lajoie, Philippe Souza Moraes Ribeiro, Matthew T Weirauch, David E Hill, Timothy R Hughes, Chad L Myers, and Albertha J M Walhout. Extensive rewiring and complex evolutionary dynamics in a C. elegans multiparameter transcription factor network. *Mol. Cell*, 51(1):116–127, July 2013.

[263] Wilberforce Zachary Ouma, Katja Pogacar, and Erich Grotewold. Topological and statistical analyses of gene regulatory networks reveal unifying yet quantitatively different emergent properties. *PLoS Comput. Biol.*, 14(4):e1006098, April 2018.

[264] Dong-Qing Sun, Liu Tian, and Bin-Guang Ma. Spatial organization of the transcriptional regulatory network of Saccharomyces cerevisiae. *FEBS Lett.*, 593(8):876–884, April 2019.

[265] Liu Tian, Tong Liu, Kang-Jian Hua, Xiao-Pan Hu, and Bin-Guang Ma. The Spatial Organization of Bacterial Transcriptional Regulatory Networks. *Microorganisms*, 10(12), November 2022.

[266] W S Reznikoff. The lactose operon-controlling elements: a complex paradigm. *Mol. Microbiol.*, 6(17):2419–2422, September 1992.

[267] Gong-Hong Wei, De-Pei Liu, and Chih-Chuan Liang. Charting gene regulatory networks: strategies, challenges and perspectives. *Biochem. J*, 381(Pt 1):1–12, July 2004.

[268] Agustino Martínez-Antonio and Julio Collado-Vides. Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr. Opin. Microbiol.*, 6(5):482–489, October 2003.

[269] Akira Ishihama. Prokaryotic genome regulation: multifactor promoters, multitarget regulators and hierarchic networks. *FEMS Microbiol. Rev.*, 34(5):628–645, September 2010.

[270] Akira Ishihama. Prokaryotic genome regulation: a revolutionary paradigm. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.*, 88(9):485–508, 2012.

[271] J G Harman. Allosteric regulation of the cAMP receptor protein. *Biochim. Biophys. Acta*, 1547(1):1–17, May 2001.

[272] Jan Ihmels, Gilgi Friedlander, Sven Bergmann, Ofer Sarig, Yaniv Ziv, and Naama Barkai. Revealing modular organization in the yeast transcriptional network. *Nat. Genet.*, 31(4):370–377, August 2002.

[273] Jan Ihmels, Ronen Levy, and Naama Barkai. Principles of transcriptional control in the metabolic network of Saccharomyces cerevisiae. *Nat. Biotechnol.*, 22(1):86–92, January 2004.

[274] Martin Lempp, Niklas Farke, Michelle Kuntz, Sven Andreas Freibert, Roland Lill, and Hannes Link. Systematic identification of metabolites controlling gene expression in E. coli. *Nat. Commun.*, 10(1):4463, October 2019.

[275] Xu Fang, Cong Wei, Cai Zhao-Ling, and Ouyang Fan. Effects of organic carbon sources on cell growth and eicosapentaenoic acid content of Nannochloropsis sp. *J. Appl. Phycol.*, 16(6):499–503, December 2004.

[276] Tong Ihn Lee, Nicola J Rinaldi, François Robert, Duncan T Odom, Ziv Bar-Joseph, Georg K Gerber, Nancy M Hannett, Christopher T Harbison, Craig M Thompson, Itamar Simon, Julia Zeitlinger, Ezra G Jennings, Heather L Murray, D Benjamin Gordon, Bing Ren, John J Wyrick, Jean-Bosco Tagne, Thomas L Volkert, Ernest Fraenkel, David K Gifford, and Richard A Young. Transcriptional regulatory networks in Saccharomyces cerevisiae. *Science*, 298(5594):799–804, October 2002.

[277] Tamara Gigolashvili, Ruslan Yatusevich, Bettina Berger, Caroline Müller, and Ulf-Ingo Flügge. The R2R3-MYB transcription factor HAG1/MYB28 is a regulator of methionine-derived glucosinolate biosynthesis in Arabidopsis thaliana. *Plant J.*, 51(2):247–261, July 2007.

[278] Ida Elken Sønderby, Bjarne Gram Hansen, Nanna Bjarnholt, Carla Ticconi, Barbara Ann Halkier, and Daniel J Kliebenstein. A systems biology approach identifies a R2R3 MYB gene subfamily with distinct and overlapping functions in regulation of aliphatic glucosinolates. *PLoS One*, 2(12):e1322, December 2007.

[279] Ida Elken Sønderby, Meike Burow, Heather C Rowe, Daniel J Kliebenstein, and Barbara Ann Halkier. A complex interplay of three R2R3 MYB transcription factors determines the profile of aliphatic glucosinolates in Arabidopsis. *Plant Physiol.*, 153(1):348–363, May 2010.

[280] Baohua Li, Michelle Tang, Ayla Nelson, Hart Caligagan, Xue Zhou, Caitlin Clark-Wiest, Richard Ngo, Siobhan M Brady, and Daniel J Kliebenstein. Network-Guided Discovery of Extensive Epistasis between Transcription Factors Involved in Aliphatic Glucosinolate Biosynthesis. *Plant Cell*, 30(1):178–195, January 2018.

[281] Waeil Al Youssef, Regina Feil, Maureen Saint-Sorny, Xenie Johnson, John E Lunn, Bernhard Grimm, and Pawel Brzezowski. Singlet oxygen-induced signalling depends on the metabolic status of the Chlamydomonas cell. August 2022.

[282] T Jabs, R A Dietrich, and J L Dangl. Initiation of runaway cell death in an Arabidopsis mutant by extracellular superoxide. *Science*, 273(5283):1853–1856, September 1996.

[283] R A Dietrich, T P Delaney, S J Uknes, E R Ward, J A Ryals, and J L Dangl. Arabidopsis mutants simulating disease resistance response. *Cell*, 77(4):565–577, May 1994.

[284] Per Mühlenbock, Magdalena Szechynska-Hebda, Marian Plaszczyca, Marcela Baudo, Alfonso Mateo, Philip M Mullineaux, Jane E Parker, Barbara Karpinska, and Stanislaw Karpinski. Chloroplast signaling and LESION SIMULATING DISEASE1 regulate crosstalk between light acclimation and immunity in Arabidopsis. *Plant Cell*, 20(9):2339–2356, September 2008.

[285] Weronika Wituszyńska, Magdalena Szechyńska-Hebda, Mirosław Sobczak, Anna Rusaczonek, Anna Kozłowska-Makulska, Damian Witoń, and Stanisław Karpiński. Lesion simulating disease 1 and enhanced disease susceptibility 1 differentially regulate UV-C-induced photooxidative stress signalling and programmed cell death in Arabidopsis thaliana. *Plant Cell Environ.*, 38(2):315–330, February 2015.

[286] H Wajant, F Mühlenbeck, and P Scheurich. Identification of a TRAF (TNF receptor-associated factor) gene in Caenorhabditis elegans. *J. Mol. Evol.*, 47(6):656–662, December 1998.

[287] Guang-Ho Cha, Kyoung Sang Cho, Jun Hee Lee, Myungjin Kim, Euysoo Kim, Jeehye Park, Sung Bae Lee, and Jongkyeong Chung. Discrete functions of TRAF1 and TRAF2 in Drosophila melanogaster mediated by c-Jun N-terminal kinase and NF-kappaB-dependent signaling pathways. *Mol. Cell. Biol.*, 23(22):7982–7991, November 2003.

[288] Ping Xie. TRAF molecules in cell signaling and in human diseases. *J. Mol. Signal.*, 8(1):7, June 2013.

[289] Hua Qi, Fan-Nv Xia, Shi Xiao, and Juan Li. TRAF proteins as key regulators of plant development and stress responses. *J. Integr. Plant Biol.*, 64(2):431–448, February 2022.

[290] Muneer Ahmed Khoso, Amjad Hussain, Faujiah Nurhasanah Ritonga, Qurban Ali, Muhammed Malook Channa, Rana M Alshegaihi, Qinglin Meng, Musrat Ali, Wajid Zaman, Rahim Dad Brohi, Fen Liu, and Hakim Manghwar. WRKY transcription factors (TFs): Molecular switches to regulate drought, temperature, and salinity stresses in plants. *Front. Plant Sci.*, 13:1039329, November 2022.

[291] Tao Yao, Jin Zhang, Meng Xie, Guoliang Yuan, Timothy J Tschaplinski, Wellington Muchero, and Jin-Gui Chen. Transcriptional Regulation of Drought Response in Arabidopsis and Woody Plants. *Front. Plant Sci.*, 11:572137, 2020.

[292] Chang-Quan Zhang, Yong Xu, Yan Lu, Heng-Xiu Yu, Ming-Hong Gu, and Qiao-Quan Liu. The WRKY transcription factor OsWRKY78 regulates stem elongation and seed development in rice. *Planta*, 234(3):541–554, September 2011.

[293] Jingjing Jiang, Shenghui Ma, Nenghui Ye, Ming Jiang, Jiashu Cao, and Jianhua Zhang. WRKY transcription factors in plant responses to stresses. *J. Integr. Plant Biol.*, 59(2):86–101, February 2017.

[294] Supriya Ambawat, Poonam Sharma, Neelam R Yadav, and Ram C Yadav. MYB transcription factor genes as regulators for plant responses: an overview. *Physiol. Mol. Biol. Plants*, 19(3):307–321, July 2013.

[295] Zhouli Xie, Trevor M Nolan, Hao Jiang, and Yanhai Yin. AP2/ERF Transcription Factor Regulatory Networks in Hormone and Abiotic Stress Responses in Arabidopsis. *Front. Plant Sci.*, 10:228, February 2019.

[296] Xi Yuan, Hui Wang, Jiating Cai, Dayong Li, and Fengming Song. NAC transcription factors in plant immunity. *Phytopathology Research*, 1(1):1–13, January 2019.

[297] Hongbo Shao, Hongyan Wang, and Xiaoli Tang. NAC transcription factors in plant multiple abiotic stress responses: progress and prospects. *Front. Plant Sci.*, 6:902, October 2015.

[298] Xiaodong Deng, Jiajia Cai, and Xiaowen Fei. Effect of the expression and knockdown of citrate synthase gene on carbon flux during triacylglycerol biosynthesis by green algae Chlamydomonas reinhardtii. *BMC Biochem.*, 14:38, December 2013.

[299] Fantao Kong, Yuanxue Liang, Bertrand Légeret, Audrey Beyly-Adriano, Stéphanie Blangy, Richard P Haslam, Johnathan A Napier, Fred Beisson, Gilles Peltier, and Yonghua Li-Beisson. Chlamydomonas carries out fatty acid β-oxidation in ancestral peroxisomes using a bona fide acyl-CoA oxidase. *Plant J.*, 90(2):358–371, April 2017.

[300] Namrata Misra, Prasanna Kumar Panda, Mahesh Chandra Patra, Sukanta Kumar Pradhan, and Barada Kanta Mishra. Insights into molecular assembly of ACCase heteromeric complex in Chlorella variabilis–a homology modelling, docking and molecular dynamic simulation study. *Appl. Biochem. Biotechnol.*, 170(6):1437–1457, July 2013.

[301] David J Triggle John B. Taylor. Acetyl-CoA Carboxylase. `https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/acetyl-coa-carboxylase`, 2007. Accessed: 2022-11-4.

[302] Nobuko Sumiya, Yasuko Kawase, Jumpei Hayakawa, Mami Matsuda, Mami Nakamura, Atsuko Era, Kan Tanaka, Akihiko Kondo, Tomohisa Hasunuma, Sousuke Imamura, and Shin-Ya Miyagishima. Expression of Cyanobacterial Acyl-ACP Reductase Elevates the Triacylglycerol Level in the Red Alga Cyanidioschyzon merolae. *Plant Cell Physiol.*, 56(10):1962–1980, October 2015.

[303] Hoa Mai Nguyen, Stéphan Cuiné, Audrey Beyly-Adriano, Bertrand Légeret, Emmanuelle Billon, Pascaline Auroy, Fred Beisson, Gilles Peltier, and Yonghua Li-Beisson. The green microalga Chlamydomonas reinhardtii has a single ω-3 fatty acid desaturase that localizes to the chloroplast and impacts both plastidic and extraplastidic membrane lipids. *Plant Physiol.*, 163(2):914–928, October 2013.

[304] Ángeles Gómez-Zambrano, Pedro Crevillén, José M Franco-Zorrilla, Juan A López, Jordi Moreno-Romero, Pawel Roszak, Juan Santos-González, Silvia Jurado, Jesús Vázquez, Claudia Köhler, Roberto Solano, Manuel Piñeiro, and José A Jarillo. Arabidopsis SWC4 Binds DNA and Recruits the SWR1 Complex to Modulate Histone H2A.Z Deposition at Key Regulatory Genes. *Mol. Plant*, 11(6):815–832, June 2018.

[305] Weronika Sura, Michał Kabza, Wojciech M Karlowski, Tomasz Bieluszewski, Marta Kus-Slowinska, Łukasz Pawełoszek, Jan Sadowski, and Piotr A Ziolkowski. Dual Role of the Histone Variant H2A.Z in Transcriptional Regulation of Stress-Response Genes. *Plant Cell*, 29(4):791–807, April 2017.

[306] Mohammad Aslam, Beenish Fakher, Bello Hassan Jakada, Shijiang Cao, and Yuan Qin. SWR1 Chromatin Remodeling Complex: A Key Transcriptional Regulator in Plants. *Cells*, 8(12), December 2019.

[307] Kyuha Choi, Chulmin Park, Jungeun Lee, Mijin Oh, Bosl Noh, and Ilha Lee. Arabidopsis homologs of components of the SWR1 complex regulate flowering and plant development. *Development*, 134(10):1931–1941, May 2007.

[308] Wen-Feng Nie, Mingguang Lei, Mingxuan Zhang, Kai Tang, Huan Huang, Cuijun Zhang, Daisuke Miki, Pan Liu, Yu Yang, Xingang Wang, Heng Zhang, Zhaobo Lang, Na Liu, Xuechen Xu, Ramesh Yelagandula, Huiming Zhang, Zhidan Wang, Xiaoqiang Chai, Andrea Andreucci, Jing-Quan Yu, Frederic Berger, Rosa Lozano-Duran, and Jian-Kang Zhu. Histone acetylation recruits the SWR1 complex to regulate active DNA demethylation in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.*, 116(33):16641–16650, August 2019.

[309] Souha Berriri, Sreeramaiah N Gangappa, and S Vinod Kumar. SWR1 Chromatin-Remodeling Complex Subunits and H2A.Z Have Non-overlapping Functions in Immunity and Gene Regulation in Arabidopsis. *Mol. Plant*, 9(7):1051–1065, July 2016.

[310] Rwitie Mallik, Anindya Kundu, and Shubho Chaudhuri. High mobility group proteins: the multifaceted regulators of chromatin dynamics. *Nucleus*, 61(3):213–226, December 2018.

[311] Nobuhiro Tanuma, Sei-Eun Kim, Monique Beullens, Yao Tsubaki, Shinya Mitsuhashi, Miyuki Nomura, Takeshi Kawamura, Kyoichi Isono, Haruhiko Koseki, Masami Sato, Mathieu Bollen, Kunimi Kikuchi, and Hiroshi Shima. Nuclear inhibitor of protein phosphatase-1 (NIPP1) directs protein phosphatase-1 (PP1) to dephosphorylate the U2 small nuclear ribonucleoprotein particle (snRNP) component, spliceosome-associated protein 155 (Sap155). *J. Biol. Chem.*, 283(51):35805–35814, December 2008.

[312] Aleyde Van Eynde, Mieke Nuytten, Mieke Dewerchin, Luc Schoonjans, Stefaan Keppens, Monique Beullens, Lieve Moons, Peter Carmeliet, Willy Stalmans, and Mathieu Bollen. The nuclear scaffold protein NIPP1 is essential for early embryonic development and cell proliferation. *Mol. Cell. Biol.*, 24(13):5863–5874, July 2004.

[313] Hugo Ceulemans, Willy Stalmans, and Mathieu Bollen. Regulator-driven functional diversification of protein phosphatase-1 in eukaryotic evolution. *Bioessays*, 24(4):371–381, April 2002.

[314] Christian L Barrett, Christopher D Herring, Jennifer L Reed, and Bernhard O Palsson. The global transcriptional regulatory network for metabolism in Escherichia coli exhibits few dominant functional states. *Proc. Natl. Acad. Sci. U. S. A.*, 102(52):19103–19108, December 2005.

[315] Xin Fang, Anand Sastry, Nathan Mih, Donghyuk Kim, Justin Tan, James T Yurkovich, Colton J Lloyd, Ye Gao, Laurence Yang, and Bernhard O Palsson. Global transcriptional regulatory network for *Escherichia coli* robustly connects gene expression to transcription factor activities. *Proc. Natl. Acad. Sci. U. S. A.*, 114(38):10286–10291, September 2017.

[316] D S Gorman and R P Levine. Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of Chlamydomonas reinhardi. *Proc. Natl. Acad. Sci. U. S. A.*, 54(6):1665–1669, December 1965.

[317] Michiel Van Bel, Tim Diels, Emmelien Vancaester, Lukasz Kreft, Alexander Botzki, Yves Van de Peer, Frederik Coppens, and Klaas Vandepoele. PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Res.*, 46(D1):D1190–D1196, January 2018.

[318] Triinu Kõressaar, Maarja Lepamets, Lauris Kaplinski, Kairi Raime, Reidar Andreson, and Maido Remm. Primer3_masker: integrating masking of template sequence with primer design software. *Bioinformatics*, 34(11):1937–1938, June 2018.

[319] Triinu Koressaar and Maido Remm. Enhancements and modifications of primer design program Primer3. *Bioinformatics*, 23(10):1289–1291, May 2007.

[320] Andreas Untergasser, Ioana Cutcutache, Triinu Koressaar, Jian Ye, Brant C Faircloth, Maido Remm, and Steven G Rozen. Primer3–new capabilities and interfaces. *Nucleic Acids Res.*, 40(15):e115, August 2012.

[321] X Chen, J L Zaro, and W C Shen. Fusion protein linkers: property, design and functionality. *Adv. Drug Deliv. Rev.*, 65(10):1357–1369, 2013.

[322] Melina Altmann, Stefan Altmann, Claudia Falter, and Pascal Falter-Braun. High-Quality Yeast-2-Hybrid Interaction Network Mapping. *Curr Protoc Plant Biol*, 3(3):e20067, September 2018.

[323] Anthony M Bolger, Marc Lohse, and Bjoern Usadel. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114–2120, August 2014.

[324] Heng Li and Richard Durbin. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14):1754–1760, July 2009.

[325] Heng Li, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16):2078–2079, August 2009.

[326] Petr Danecek, James K Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O Pollard, Andrew Whitwham, Thomas Keane, Shane A McCarthy, Robert M Davies, and Heng Li. Twelve years of SAMtools and BCFtools. *Gigascience*, 10(2), February 2021.

[327] Yong Zhang, Tao Liu, Clifford A Meyer, Jérôme Eeckhoute, David S Johnson, Bradley E Bernstein, Chad Nusbaum, Richard M Myers, Myles Brown, Wei Li, and X Shirley Liu. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, 9(9):R137, September 2008.

[328] T L Bailey and C Elkan. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, 2:28–36, 1994.

[329] Timothy L Bailey, James Johnson, Charles E Grant, and William S Noble. The MEME Suite. *Nucleic Acids Res.*, 43(W1):W39–49, July 2015.

[330] Shobhit Gupta, John A Stamatoyannopoulos, Timothy L Bailey, and William Stafford Noble. Quantifying similarity between motifs. *Genome Biol.*, 8(2):R24, 2007.

[331] Xiao-Nian Ma, Tian-Peng Chen, Bo Yang, Jin Liu, and Feng Chen. Lipid Production from Nannochloropsis. *Mar. Drugs*, 14(4), March 2016.

A Appendix

A.1 Additional figures

Figure A.1 shows the different PCR amplification success rates according to the ORF size of the amplified ORFs.
In Figure A.3 all targeted promoters and TFs of LRL1, MYB_9, MS_1, CGL86 and HMG_6 are described. The interaction partners are colored according to their assigned GO terms.
Figure A.4 shows the hypergeometric test analysis of the interaction pattern of TRAF TFs and MYB TFs. These were the only TF families that showed an enriched or depleted interaction behavior for certain promoters.
The first experimental-based GRN of *C. reinhardtii* with all annotated and assigned names is shown in Figure A.5.

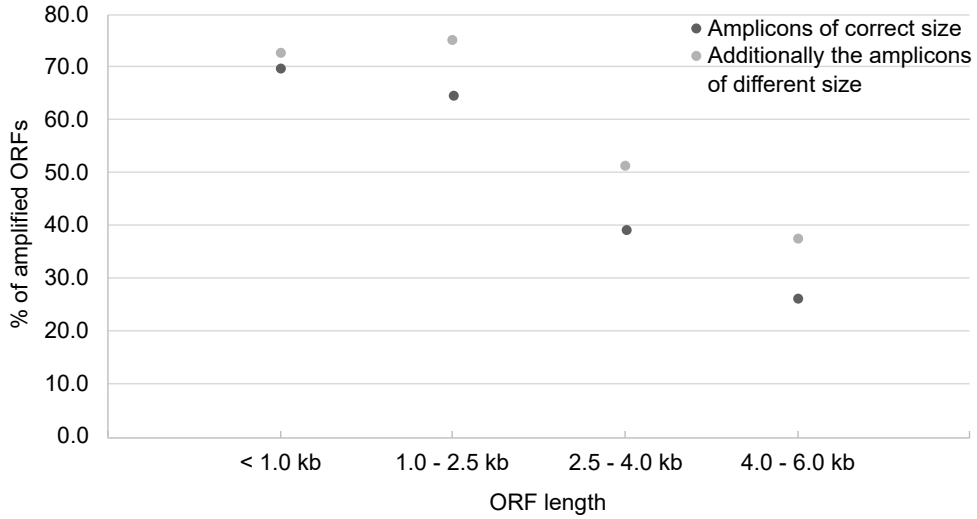


Figure A.1: Different OFR amplification rates according to ORF size
The ORF amplification is highly depended on the ORF length. Larger ORFs are amplified less efficiently in a PCR due to the high GC content of the *C. reinhardtii* genome. The data points reflect the amplification results of the 1st PCR of 4.903 ORFs.

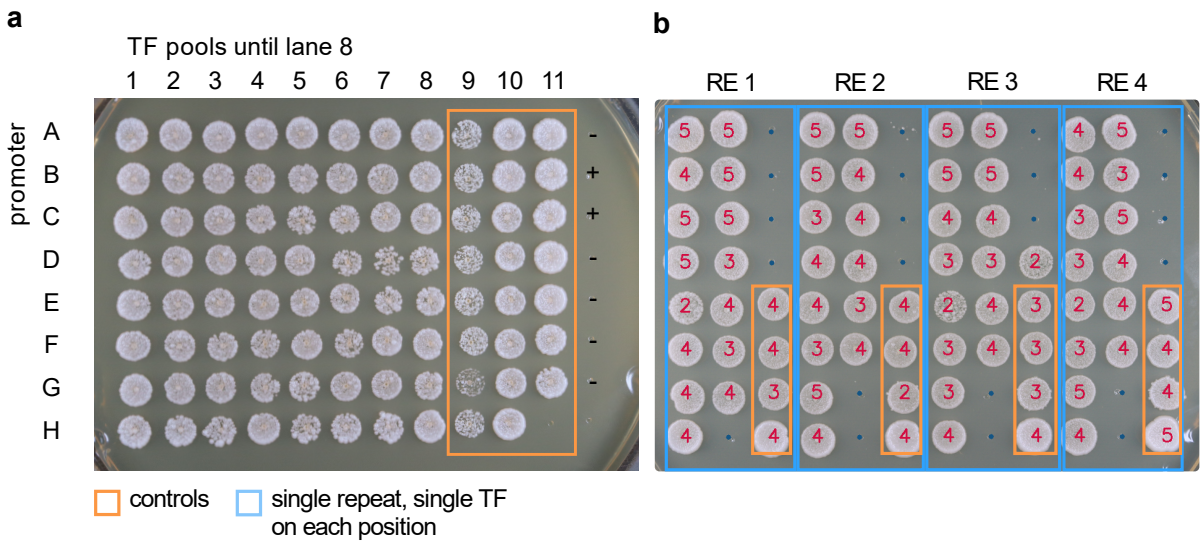


Figure A.2: Yeast mating plate

Yeast mating plate (YEPD) of primary screening (**a**) and verification screening (**b**). Red numbers indicate the automated scoring system, which was used for analysis. Since the scoring system was trained first, all images were also analyzed manually. YEDP plate of primary screening: MS_Bait26 vs TF pool 1, YEPD plate of verification screening: TF mini pool 5 vs. Cre14.g621050 (Bait 24G).

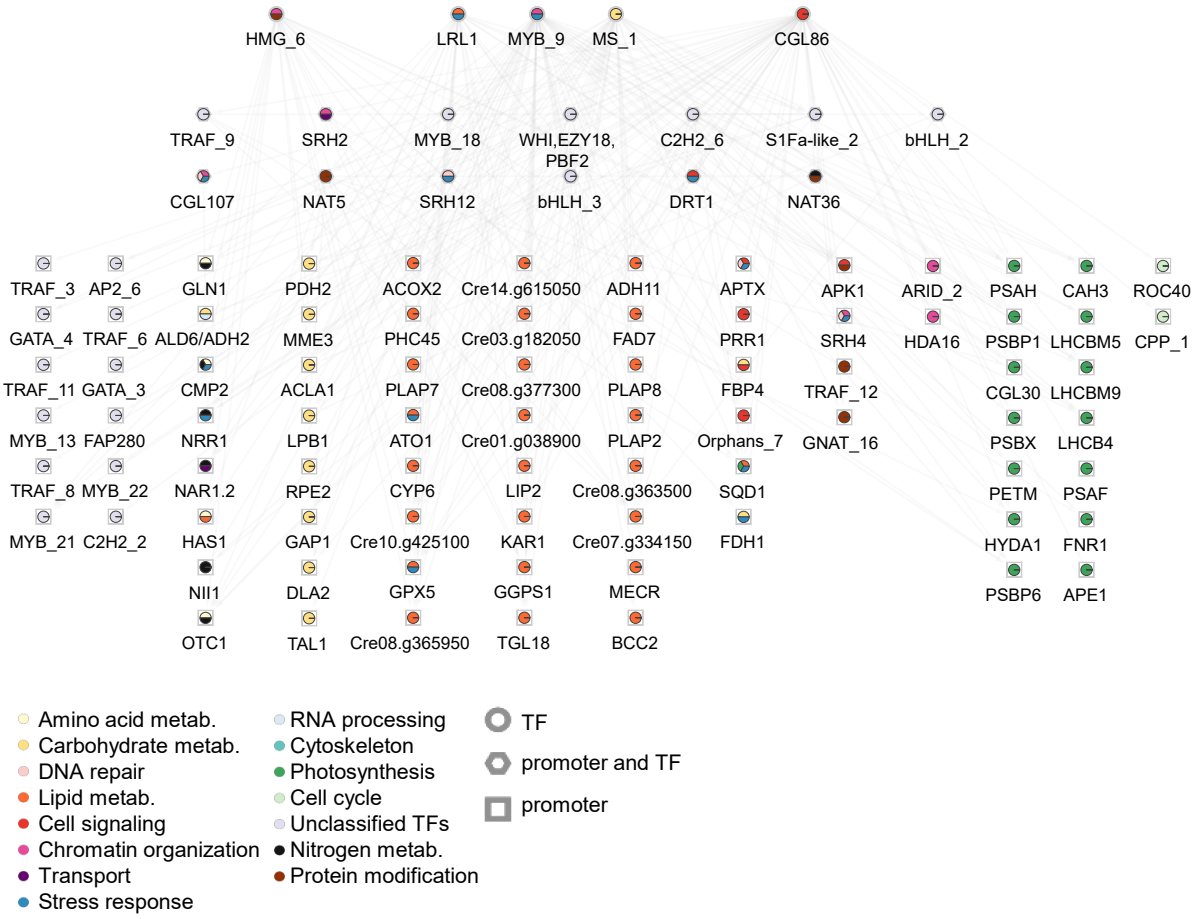


Figure A.3: Interaction partners of LRL1, MYB_9, MS_1, CGL86 and HMG_6

The network shows the targeted promoters of LRL1, MYB_9, MS_1, CGL86 and HMG_6. This subnetwork is structured according to the outdegree of the TFs. TFs are represented as circles, promoters as squares, and interaction partner, present as active TFs and promoters are shown as hexagons. All interaction partners are colored according to their GO terms.

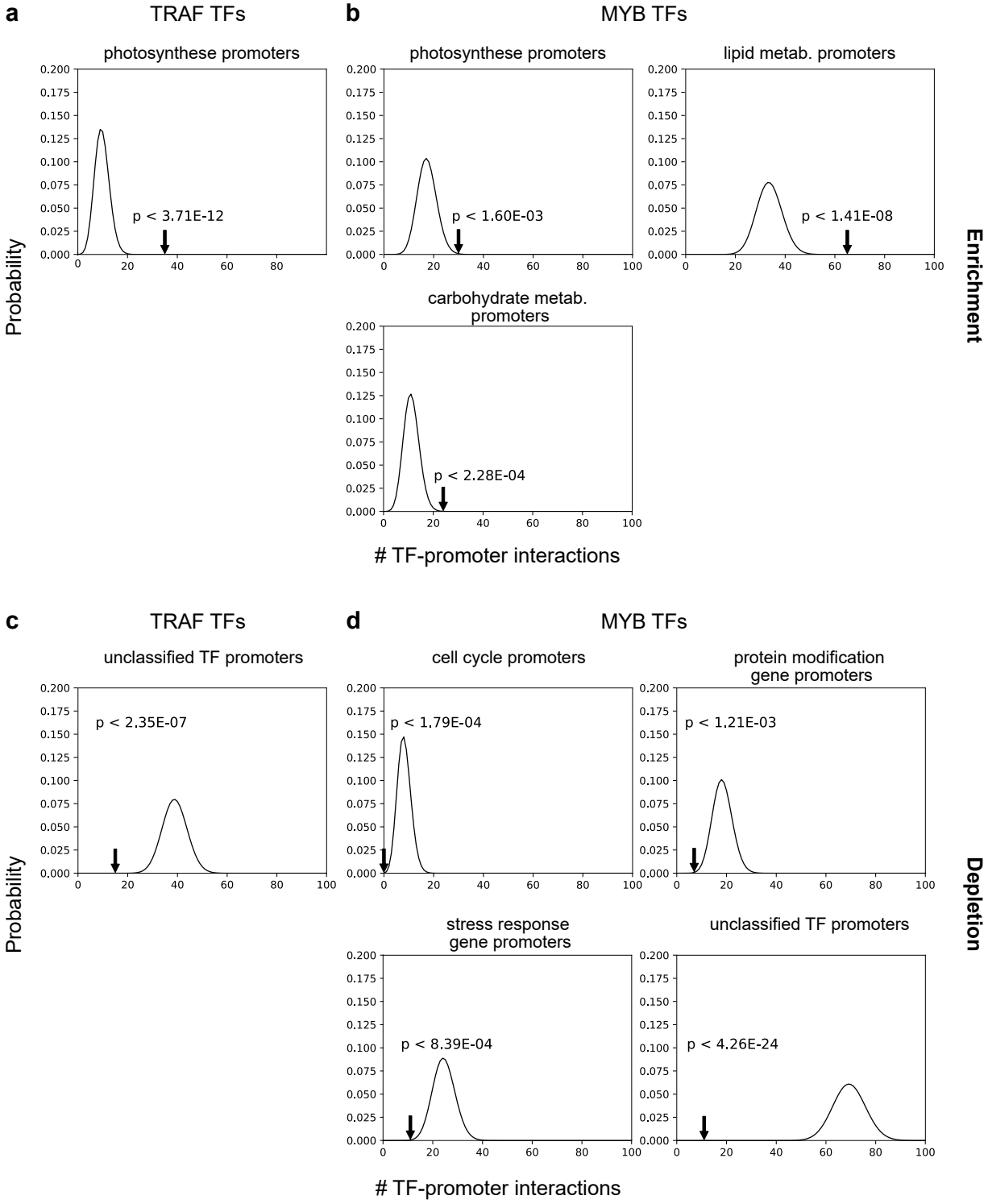


Figure A.4: Enrichment and depletion analysis for TF families

Hypergeometric test analysis of interaction behavior of TRAF and MYB TFs. All plots show the interval from 0 - 100 for count of PDIs. The arrow indicates the number of interactions between the TRAF TFs or MYB TFs with the respective promoters of a specific GO term. The p-value describes the probability of finding this number of PDIs by chance. Plot **a** describes the enrichment interaction behavior for TRAF TF with photosynthesis promoters. Plot **b** shows the enrichment for MYB TFs with promoters of photosynthesis, lipid metabolism and carbohydrate metabolism. **c** Plot shows the depletion of interaction of TRAF TFs with promoters of unclassified TFs, whereas plot **d** describes the depletion of the interaction behavior of MYB TFs with promoters of cell cycle, protein modification, stress response and unclassified TFs.

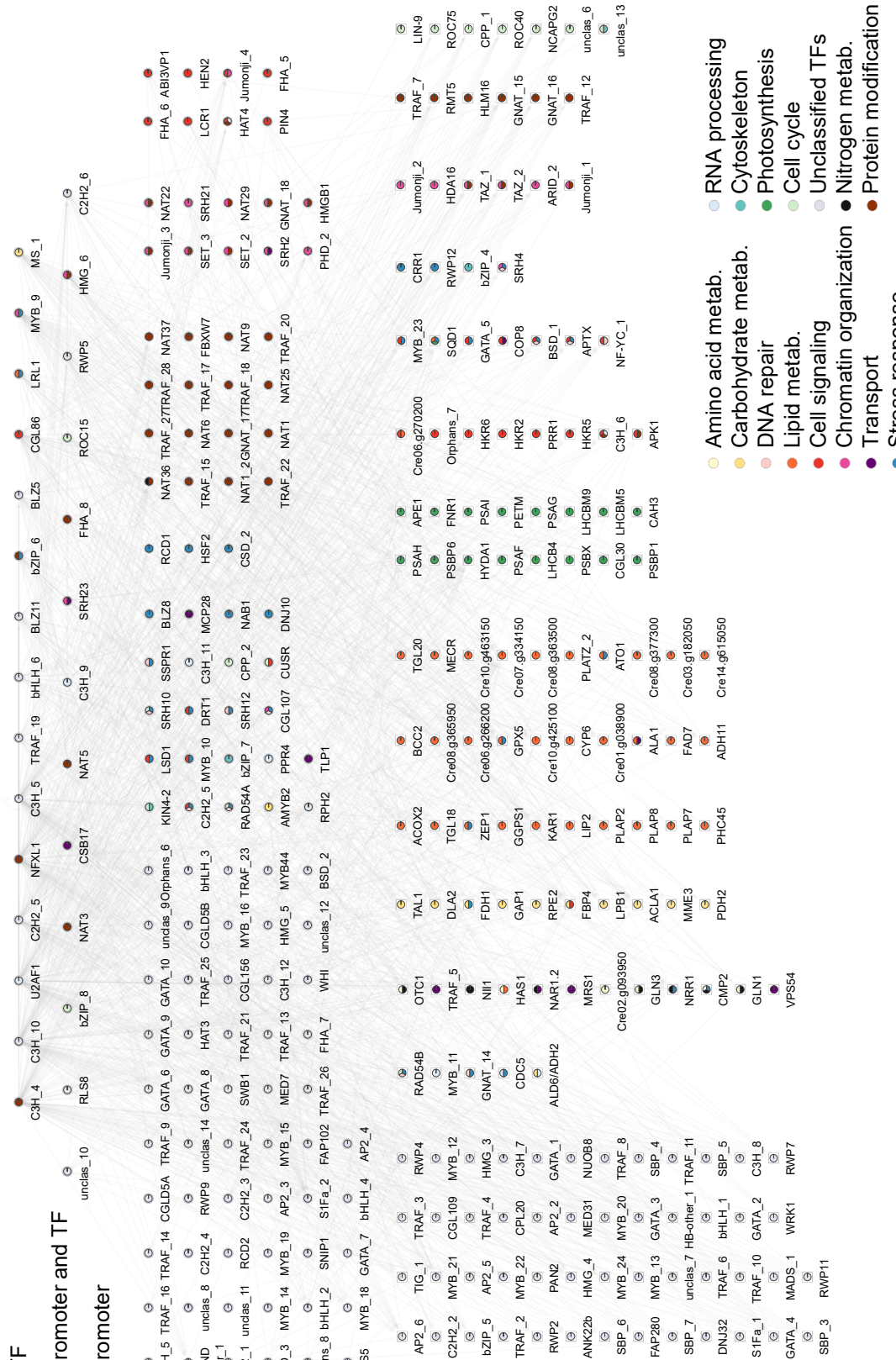


Figure A.5: *C. reinhardtii's* GRN

First experimental-based GRN of *C. reinhardtii*. The network displays all gene names. A name was assigned to the gene by me when there was no gene name in v5.5. All nodes and edges are colored according to their GO terms. The network is structures according to the outdegree of the interaction partners.

A.2 Lists

Table A.1 shows the lipid composition, consisting of SFAs, MUFAs and PUFAs, of various microalgae species, including *C. reinhardtii*, *Chlorella* species, *Dunaliella* species and *Nannochloropsis* species. In addition, the lipid composition of palm, sunflower, rapeseed and soybean are described.
Table A.2 describes the fatty acid composition in *C. reinhardtii cw15*. The table was taken from Popko et al. [81]. The fatty acid composition was determined via GC-FID.

Table A.3 provides information on the gene ID and ORF length of the ORFs used for PCR optimization described in Figure 2.2.

Table A.1: Fatty acid profile of different oil plant organisms

Organism	**SFA**	**MUFA**	**PUFA**	Ref.
Chlamydomonas reinhardtii	24.0	23.0	53.0	[81, 63, 82]
*Chlorella salina**	29.3	18.5	40.7	[34]
Chlorella emersonii	24.6	17.0	38.4	[34]
Chlorella vulgaris	24.0	24.8	45.9	[34]
Dunaliella saline	22.8	22.9	34.5	[34]
Dunaliella sp.	13.5	24.7	48.2	[34]
Nannochloropsis sp.	30.0	42.6	23.7	[331]
Palm	44.7	46.4	8.9	[36]
Sunflower	11.1	25.6	63.3	[36]
Rapeseed	6.5	65.3	28.3	[36]
Soybean	15.3	25.6	59.1	[36]

Table A.2: Fatty acid composition of *C. reinhardtii*

Fatty acid	**Mol %**
C16:0	22
C16:1(7Z)	4.5
C16:1(9Z)	0.5
C16:1(3E)	2
C16:2(7Z,10Z)	2
C16:3(4Z,7Z,10Z)	1
C16:3(7Z,10Z,13Z)	2
C16:4(4Z,7Z, 10Z,13Z)	13
C18:0	2
C18:1(9Z)	13
C18:1(11Z)	3
C18:2(9Z,12Z)	8

Table A.3: ORFs for PCR optimization

Identifier	Gene ID	ORF length	Identifier	Gene ID	ORF length
a: A	Cre03.g155550	1110	a: F	Cre06.g285300	2985
a: B	Cre07.g341850	2901	a: G	Cre01.g035700	2538
a: C	Cre03.g146227	2937	a: H	Cre07.g315350	3018
a: D	Cre02.g146600	1461	a: I	Cre07.g319150	2781
a: E	Cre03.g144867	1653	a: J	Cre07.g355800	2787

Identifier upper panel	Gene ID	ORF length	Identifier lower panel	Gene ID	ORF length
b: 1	Cre03.g144867	1653	b: 9	Cre08.g385050	2223
b: 2	Cre04.g217915	1659	b: 10	Cre06.g265100	2226
b: 3	Cre07.g339750	1659	b: 11	Cre07.g351800	2232
b: 4	Cre17.g713350	1662	b: 12	Cre07.g323850	2235
b: 5	Cre05.g232600	1665	b: 13	Cre12.g484150	2247
b: 6	Cre17.g732050	1665	b: 14	Cre09.g391541	2250
b: 7	Cre03.g188300	1668	b: 15	Cre02.g142426	2253
b: 8	Cre10.g443250	1668	b: 16	Cre13.g582476	2253

Identifier	Gene ID	ORF length	Identifier	Gene ID	ORF length
c: A	Cre13.g584850	729	c: C	Cre12.g514450	1974
c: B	Cre16.g685000	729	c: D	Cre11.g479600	2391

Identifier	Gene ID	ORF length	Identifier	Gene ID	ORF length
d: 1	Cre07.g325739	5886	d: 9	Cre07.g329900	4083
d: 2	Cre12.g557500	5967	d: 10	Cre12.g525300	4116
d: 3	Cre24.g755347	4005	d: 11	Cre10.g456300	4158
d: 4	Cre17.g724950	4020	d: 12	Cre10.g431950	4200
d: 5	Cre07.g329882	4065	d: 13	Cre07.g325754	4203
d: 6	Cre06.g268800	4068	d: 14	Cre03.g154900	4263
d: 7	Cre16.g687800	4071	d: 15	Cre08.g364862	4278
d: 8	Cre01.g009300	4074	d: 16	Cre06.g278138	4299

Table A.4: Success rates of different tested polymerases in Figure 2.2

tested polymerase	correct amplicons
KOD Hot Start polymerase	7/10
PCRBIO Hifi polymerase	0/10
TaKaRa Taq DNA polymerase	0/10
AmpliTag Gold$^{(TM)}$ ®TM 360 polymerase	0/10 (different lengths amplicons)
Q5®R High-Fidelity DNA polymerase	0/10 (different lengths amplicons)
Phusion®R High Fidelity DNA polymerase	3/10

Table A.5 contains all consolidated TFs and TRs of *C. reinhardtii* and their assigned TF or TR families. It also contains the names I assigned them in the Y1H screen, if there were part of the screen.

Table A.5: List of TFs and TRs

No	C.re gene ID	TF family/ TR group	named as	No	C.re gene ID	TF family/ TR group	named as
1	Cre13.g562400	ABI3/VP1	ABI3VP1	2	Cre03.g173350	ANKYRIN	ANK22b
3	Cre05.g245700	ANKYRIN		4	Cre03.g144424	AP2/ERF	AP2_5
5	Cre03.g209953	AP2/ERF	AP2_3	6	Cre16.g667900	AP2/ERF	AP2_6
7	Cre13.g602750	AP2/ERF	AP2_1	8	Cre07.g353500	AP2/ERF	CGLD5A
9	Cre08.g364400	AP2/ERF	AP2_4	10	Cre16.g649433	AP2/ERF	AP2_2
11	Cre16.g674890	AP2/ERF	CGLD5B	12	Cre01.g009575	AP2/ERF	-
13	Cre01.g009650	AP2/ERF	-	14	Cre02.g100100	AP2/ERF	-
15	Cre08.g383000	AP2/ERF	-	16	Cre08.g385350	AP2/ERF	-
17	Cre10.g441300	AP2/ERF	-	18	Cre17.g729600	AP2/ERF	-
19	Cre01.g023100	AP2/ERF	-	20	Cre06.g275500	AP2/ERF	-
21	Cre06.g278181	AP2/ERF	-	22	Cre14.g620500	AP2/ERF	-
23	Cre16.g661650	AP2/ERF	-	24	Cre19.g750947	AP2/ERF	
25	Cre10.g429880	ARID	CGL156	26	Cre13.g571050	ARID	ARID_2
27	Cre06.g305850	ARID	-	28	Cre12.g508050	ARID	-
29	Cre12.g514250	ARID	-	30	Cre05.g243000	ARR-B	-
31	Cre01.g011150	bHLH	bHLH_2	32	Cre04.g216200	bHLH	bHLH_3
33	Cre04.g224600	bHLH	bHLH_4	34	Cre07.g332250	bHLH	bHLH_1
35	Cre07.g353555	bHLH	bHLH_5	36	Cre16.g667150	bHLH	bHLH_6
37	Cre02.g109683	bHLH	-	38	Cre04.g216204	bHLH	-
39	Cre05.g241636	bHLH	-	40	Cre06.g259100	bHLH	-
41	Cre07.g349152	bHLH	-	42	Cre14.g620850	bHLH	-
43	Cre06.g284000	BSD	BSD_2	44	Cre10.g466550	BSD	BSD_1
45	Cre07.g321550	bZIP	BLZ5	46	Cre07.g344668	bZIP	bZIP_6
47	Cre10.g438850	bZIP	bZIP_8	48	Cre12.g501600	bZIP	BLZ8
49	Cre13.g590350	bZIP	BLZ11	50	Cre16.g680232	bZIP	bZIP_4
51	Cre16.g680588	bZIP	bZIP_7	52	Cre17.g746547	bZIP	bZIP_5
53	Cre01.g051174	bZIP	-	54	Cre05.g238250	bZIP	-
55	Cre06.g310500	bZIP	-	56	Cre07.g318050	bZIP	-
57	Cre09.g413050	bZIP	-	58	Cre10.g454850	bZIP	-
59	Cre11.g479450	bZIP	-	60	Cre12.g489000	bZIP	-
61	Cre12.g510200	bZIP	-	62	Cre12.g556950	bZIP	-
63	Cre12.g557300	bZIP	-	64	Cre13.g568350	bZIP	-
65	Cre14.g631750	bZIP	-	66	Cre16.g675700	bZIP	-
67	Cre16.g692250	bZIP	-	68	Cre17.g746947	bZIP	-

69	Cre03.g152150	C2H2	C2H2_3	70	Cre03.g185950	C2H2	C2H2_7
71	Cre05.g240600	C2H2	C2H2_4	72	Cre05.g240650	C2H2	DNJ32
73	Cre06.g305200	C2H2	C2H2_6	74	Cre12.g538801	C2H2	C2H2_5
75	Cre13.g567250	C2H2	C2H2_2	76	Cre17.g696300	C2H2	NFXL1
77	Cre03.g201250	C2H2	-	78	Cre09.g389750	C2H2	
79	Cre11.g481115	C2H2	-	80	Cre01.g025050	C2H2-GATA	C2C2-GATA_6
81	Cre03.g160600	C2H2-GATA	C2C2-GATA_4	82	Cre06.g266850	C2H2-GATA	C2C2-GATA_3
83	Cre06.g266950	C2H2-GATA	C2C2-GATA_5	84	Cre07.g319701	C2H2-GATA	C2C2-GATA_7
85	Cre08.g358534	C2H2-GATA	C2C2-GATA_8	86	Cre08.g378450	C2H2-GATA	C2C2-GATA_9
87	Cre08.g378800	C2H2-GATA	C2C2-GATA_10	88	Cre11.g467581	C2H2-GATA	C2C2-GATA_2
89	Cre12.g538550	C2H2-GATA	C2C2-GATA_1	90	Cre05.g242650	C2H2-GATA	
91	Cre03.g160700	C2H2-GATA		92	Cre05.g242600	C2H2-GATA	
93	Cre08.g358532	C2H2-GATA		94	Cre10.g435450	C2H2-GATA	
95	Cre03.g178550	C3H	FBXW7, SEL1	96	Cre04.g229948	C3H	C3H_11
97	Cre04.g231124	C3H	C3H_5	98	Cre06.g250950	C3H	U2AF1
99	Cre06.g278136	C3H	C3H_4	100	Cre06.g293750	C3H	C3H_9
101	Cre07.g330050	C3H	C3H_8	102	Cre07.g355700	C3H	C3H_7
103	Cre08.g371052	C3H	C3H_6	104	Cre10.g460400	C3H	C3H_10
105	Cre11.g476250	C3H	C3H_12	106	Cre01.g014050	C3H	
107	Cre01.g035150	C3H		108	Cre02.g075650	C3H	
109	Cre02.g097850	C3H		110	Cre02.g101950	C3H	
111	Cre03.g145407	C3H		112	Cre03.g146267	C3H	
113	Cre03.g154900	C3H		114	Cre03.g170601	C3H	
115	Cre03.g172800	C3H		116	Cre04.g229374	C3H	
117	Cre06.g254650	C3H		118	Cre06.g254700	C3H	
119	Cre06.g292650	C3H		120	Cre07.g326150	C3H	
121	Cre11.g476200	C3H		122	Cre12.g523000	C3H	
123	Cre13.g566200	C3H		124	Cre13.g606750	C3H	
125	Cre16.g689535	C3H		126	Cre05.g241634	CCAAT	
127	Cre12.g548050	Coactivator p15		128	Cre06.g278159	CO-like	
129	Cre16.g675100	CO-like		130	Cre08.g361400	CPP	CPP_1
131	Cre12.g550250	CPP	CPP_2	132	Cre11.g481800	CPP	
133	Cre17.g745697	CPP		134	Cre02.g078831	CSD	CSD_2
135	Cre06.g268600	CSD	NAB1	136	Cre05.g247000	DDT	
137	Cre13.g570150	DIRP	LIN-9	138	Cre12.g521150	Dof	
139	Cre01.g052300	E2F/DP		140	Cre07.g323000	E2F/DP	
141	Cre13.g572950	E2F/DP		142	Cre13.g573000	E2F/DP	
143	Cre02.g082550	FHA	ZEP1	144	Cre03.g152450	FHA	FHA_6
145	Cre09.g390430	FHA	SNIP1	146	Cre09.g402350	FHA	FHA_7
147	Cre12.g516050	FHA	FHA_8	148	Cre12.g548250	FHA	FHA_5
149	Cre16.g671900	FHA	PIN4	150	Cre12.g534450	FHA	CGL86
151	Cre02.g119800	FHA		152	Cre03.g189850	FHA	
153	Cre04.g212050	FHA		154	Cre07.g316600	FHA	
155	Cre02.g083750	G2-like	ROC75	156	Cre12.g495100	G2-like	
157	Cre17.g696750	G2-like		158	Cre01.g048800	GNAT	GNAT_16
159	Cre02.g084550	GNAT	GNAT_13	160	Cre03.g184151	GNAT	NAT9
161	Cre06.g256200	GNAT	HAT3	162	Cre06.g269100	GNAT	GNAT_17
163	Cre06.g278108	GNAT	GNAT_18	164	Cre07.g351500	GNAT	GNAT_14
165	Cre07.g354100	GNAT	HAT4	166	Cre08.g364450	GNAT	NAT1
167	Cre09.g392300	GNAT	NAT22	168	Cre10.g464750	GNAT	NAT25
169	Cre13.g572150	GNAT	NAT29	170	Cre13.g583200	GNAT	NAT5
171	Cre13.g586450	GNAT	GNAT_15	172	Cre14.g623800	GNAT	Cre14.g623800
173	Cre16.g662650	GNAT	NAT36	174	Cre17.g705950	GNAT	NAT37
175	Cre17.g729750	GNAT	NAT3	176	Cre01.g002250	GNAT	
177	Cre01.g009550	GNAT		178	Cre02.g101850	GNAT	

179	Cre02.g144800	GNAT		180	Cre07.g331450	GNAT	
181	Cre07.g351850	GNAT		182	Cre07.g353700	GNAT	
183	Cre10.g431450	GNAT		184	Cre11.g475100	GNAT	
185	Cre12.g560200	GNAT		186	Cre13.g581150	GNAT	
187	Cre13.g586950	GNAT		188	Cre14.g614750	GNAT	
189	Cre14.g630376	GNAT		190	Cre14.g630400	GNAT	
191	Cre16.g657150	GNAT		192	Cre16.g662350	GNAT	
193	Cre17.g745947	GNAT		194	Cre03.g200431	HB-other	HB-other_1
195	Cre02.g109650	HB-other		196	Cre03.g174500	HB-other	
197	Cre08.g375400	HB-other		198	Cre01.g029450	HMG	SSPR1
199	Cre06.g261450	HMG	HMGB1	200	Cre11.g480950	HMG	HMG_5
201	Cre12.g558950	HMG	HMG_3	202	Cre17.g702650	HMG	HMG_4
203	Cre16.g672300	HMG	HMG_6	204	Cre11.g481050	HMG	
205	Cre07.g354500	HSF	HSF2	206	Cre09.g387150	HSF	
207	Cre01.g062902	HTH	MED7	208	Cre10.g441000	IWS1	
209	Cre02.g078700	Jumonji	Jumonji_3	210	Cre10.g446850	Jumonji	Jumonji_1
211	Cre12.g485950	Jumonji	Jumonji_4	212	Cre16.g654550	Jumonji	Jumonji_2
213	Cre09.g400441	Jumonji		214	Cre17.g709550	Jumonji	
215	Cre12.g517350	LSD	LSD1	216	Cre11.g467577	MADS	MADS_1
217	Cre06.g253250	MADS		218	Cre02.g108450	MBF1	FAP280,MBF1
219	Cre03.g146027	MED7		220	Cre12.g560750	mTERF	
221	Cre09.g410450	MYB	ROC15	222	Cre02.g108350	MYB	MYB_14
223	Cre03.g144747	MYB	MYB_12	224	Cre03.g149201	MYB	MYB_15
225	Cre03.g197100	MYB	LRL1	226	Cre03.g197350	MYB	CDC5L, CEF1
227	Cre07.g345350	MYB	MYB44	228	Cre10.g450500	MYB	MYB_17
229	Cre14.g621050	MYB	MYB_13	230	Cre16.g677382	MYB	MYB_16
231	Cre17.g714229	MYB	MYB_11	232	Cre01.g034350	MYB	
233	Cre02.g103450	MYB		234	Cre08.g373346	MYB	
235	Cre09.g399067	MYB		236	Cre12.g522400	MYB	
237	Cre14.g632176	MYB		238	Cre03.g168050	MYB	
239	Cre10.g425251	MYB		240	Cre03.g144907	MYB_related	MYB_9
241	Cre02.g146629	MYB_related	MYB_22	242	Cre06.g275350	MYB_related	ROC40
243	Cre06.g289900	MYB_related	HEN2	244	Cre08.g364050	MYB_related	MYB_21
245	Cre09.g399552	MYB_related	LCR1	246	Cre09.g411600	MYB_related	MYB_23
247	Cre10.g430750	MYB_related	MYB_10	248	Cre14.g621172	MYB_related	MYB_24
249	Cre16.g672400	MYB_related	MYB_18	250	Cre16.g686250	MYB_related	MYB_19
251	Cre16.g695600	MYB_related	DNJ10,DNJ1	252	Cre01.g003376	MYB_related	
253	Cre02.g110550	MYB_related		254	Cre02.g111200	MYB_related	
255	Cre02.g113450	MYB_related		256	Cre03.g176651	MYB_related	
257	Cre03.g198800	MYB_related		258	Cre06.g249050	MYB_related	
259	Cre06.g264400	MYB_related		260	Cre09.g391353	MYB_related	
261	Cre09.g397475	MYB_related		262	Cre09.g411950	MYB_related	
263	Cre10.g421021	MYB_related		264	Cre12.g514400	MYB_related	
265	Cre14.g609350	MYB_related		266	Cre14.g633789	MYB_related	
267	Cre15.g639350	MYB_related		268	Cre03.g188200	MYB_related	RPH2
269	Cre12.g507500	MYB_related	MYB_20	270	Cre02.g089150	MYB_related	
271	Cre06.g282251	MYB_related	MYND finger_1	272	Cre03.g163200	NF-X1	
273	Cre07.g341800	NF-YB	CGL107	274	Cre17.g739450	NF-YB	DRT1
275	Cre02.g079200	NF-YB		276	Cre04.g226400	NF-YC	NF-YC_1
277	Cre03.g193900	NF-YC		278	Cre12.g556400	NF-YC	
279	Cre16.g680050	NF-YC		280	Cre01.g000050	Nin-like	
281	Cre03.g153050	Nin-like		282	Cre03.g177700	Nin-like	
283	Cre04.g218050	Nin-like		284	Cre06.g291500	Nin-like	
285	Cre09.g387850	Nin-like		286	Cre10.g453500	Nin-like	
287	Cre14.g612100	Nin-like		288	Cre03.g180800	Orphans	HDA16
289	Cre07.g329900	Orphans	COP8	290	Cre07.g351550	Orphans	Orphans_7

291	Cre10.g446450	Orphans	CUSR	292	Cre13.g571200	Orphans	HKR5
293	Cre12.g508150	Orphans	SRH21	294	Cre01.g012200	Orphans	Orphans_8
295	Cre02.g110150	Orphans	Orphans_6	296	Cre01.g038050	Orphans	
297	Cre02.g074150	Orphans		298	Cre06.g278200	Orphans	
299	Cre07.g329650	Orphans		300	Cre11.g467678	Orphans	
301	Cre11.g469400	Orphans		302	Cre12.g534100	Orphans	
303	Cre12.g534151	Orphans		304	Cre13.g572450	Orphans	
305	Cre13.g583972	Orphans		306	Cre13.g602700	Orphans	
307	Cre15.g635100	Orphans		308	Cre06.g262800	PHD	NCAPG2, LUZP5
309	Cre16.g668200	PHD	PHD_2	310	Cre16.g675246	PHD	PHD_3
311	Cre05.g235400	PHD		312	Cre03.g148250	PHD	
313	Cre08.g377200	PHD		314	Cre09.g388600	PHD	
315	Cre09.g398141	PHD		316	Cre10.g446600	PHD	
317	Cre10.g455600	PHD		318	Cre12.g492200	PHD	
319	Cre12.g500900	PHD		320	Cre12.g515050	PHD	
321	Cre14.g612350	PHD		322	Cre16.g679050	PHD	
323	Cre16.g679151	PHD		324	Cre16.g694208	PHD	
325	Cre12.g502251	PLATZ	PLATZ_2	326	Cre16.g650400	PLATZ	CPL20
327	Cre12.g502650	PLATZ		328	Cre12.g502678	PLATZ	
329	Cre12.g502700	PLATZ		330	Cre16.g676421	Pseudo ARR-B	Pseudo ARR-B_2
331	Cre02.g094150	Pseudo ARR-B	HKR2	332	Cre06.g255450	RB	
333	Cre12.g540350	Rcd1-like	RCD2	334	Cre12.g540400	Rcd1-like	RCD1
335	Cre01.g004600	RWP-RK	RWP12	336	Cre03.g149350	RWP-RK	RWP4
337	Cre03.g149400	RWP-RK	RWP11	338	Cre03.g184600	RWP-RK	RWP7
339	Cre06.g275550	RWP-RK	RWP2	340	Cre06.g285600	RWP-RK	RWP5
341	Cre06.g297150	RWP-RK	RWP9	342	Cre06.g275600	RWP-RK	RWP-RK_8
343	Cre03.g193400	S1Fa-like	S1Fa-like_1	344	Cre09.g386753	S1Fa-like	S1Fa-like_2
345	Cre01.g043550	SBP	SBP_3	346	Cre02.g104700	SBP	SBP_4
347	Cre05.g233551	SBP	SBP_5	348	Cre09.g390023	SBP	CRR1
349	Cre16.g673250	SBP	NRR1	350	Cre16.g683953	SBP	SBP_7
351	Cre16.g692500	SBP	SBP_6	352	Cre01.g033900	SBP	
353	Cre01.g070932	SBP		354	Cre02.g085150	SBP	
355	Cre02.g112750	SBP		356	Cre04.g212700	SBP	
357	Cre06.g278229	SBP		358	Cre06.g300600	SBP	
359	Cre07.g325738	SBP		360	Cre07.g335150	SBP	
361	Cre07.g345050	SBP		362	Cre09.g399289	SBP	
363	Cre10.g421550	SBP		364	Cre10.g433350	SBP	
365	Cre16.g686100	SBP		366	Cre17.g698233	SBP	
367	Cre06.g283200	SET	SET_2	368	Cre07.g352600	SET	SET_3
369	Cre11.g475950	SET	HLM16	370	Cre12.g503800	SET	RMT5
371	Cre02.g089200	SET		372	Cre03.g164650	SET	
373	Cre03.g197700	SET		374	Cre07.g321200	SET	
375	Cre09.g392542	SET		376	Cre12.g541777	SET	
377	Cre17.g702351	SET		378	Cre17.g702451	SET	
379	Cre17.g732901	SET		380	Cre17.g742700	SET	
381	Cre03.g194950	Sigma70-like		382	Cre02.g087600	SNF2	SRH2
383	Cre03.g158550	SNF2	SRH4	384	Cre03.g201850	SNF2	SRH10
385	Cre06.g270850	SNF2	SRH12	386	Cre10.g430950	SNF2	RAD54B
387	Cre14.g614400	SNF2	SRH23	388	Cre03.g189000	SNF2	SRH8, RAD54B
389	Cre03.g162701	SNF2		390	Cre03.g179300	SNF2	
391	Cre03.g183350	SNF2		392	Cre03.g210625	SNF2	
393	Cre06.g287950	SNF2		394	Cre09.g390000	SNF2	
395	Cre10.g455750	SNF2		396	Cre12.g537671	SNF2	

397	Cre16.g647602	SNF2		398	Cre12.g520850	SOH1	MED31, SOH1
399	Cre02.g118250	SWI/SNF-BAF60b	SWB1	400	Cre10.g448700	SWI/SNF-BAF60b	
401	Cre03.g189650	TAZ	TAZ_2	402	Cre03.g212641	TAZ	TAZ_1
403	Cre03.g212977	TAZ		404	Cre10.g436800	TIG	TIG_1
405	Cre07.g340450	TIG		406	Cre01.g001600	TRAF	TRAF_13
407	Cre01.g022950	TRAF	TRAF_11	408	Cre01.g048501	TRAF	TRAF_12
409	Cre02.g101150	TRAF	TRAF_6	410	Cre02.g101300	TRAF	TRAF_14
411	Cre02.g145550	TRAF	TRAF_5	412	Cre02.g145602	TRAF	CSB17
413	Cre03.g174150	TRAF	TRAF_7	414	Cre03.g179000	TRAF	TRAF_2
415	Cre03.g179650	TRAF	TRAF_15	416	Cre04.g211700	TRAF	TRAF_16
417	Cre04.g213650	TRAF	TRAF_17	418	Cre04.g215450	TRAF	TRAF_8
419	Cre04.g218400	TRAF	TRAF_18	420	Cre04.g219600	TRAF	TRAF_19
421	Cre04.g221900	TRAF	TRAF_20	422	Cre04.g221950	TRAF	TRAF_3
423	Cre06.g286700	TRAF	TRAF_9	424	Cre09.g401219	TRAF	TRAF_21
425	Cre11.g479800	TRAF	TRAF_22	426	Cre13.g580450	TRAF	TRAF_10
427	Cre13.g583150	TRAF	TRAF_23	428	Cre14.g610550	TRAF	TRAF_24
429	Cre14.g610582	TRAF	TRAF_25	430	Cre14.g610631	TRAF	TRAF_26
431	Cre16.g653250	TRAF	TRAF_4	432	Cre16.g653258	TRAF	TRAF_27
433	Cre17.g700600	TRAF	TRAF_28	434	Cre02.g101550	TRAF	
435	Cre03.g174200	TRAF		436	Cre04.g212100	TRAF	
437	Cre04.g219500	TRAF		438	Cre04.g220250	TRAF	
439	Cre04.g231516	TRAF		440	Cre06.g278184	TRAF	
441	Cre11.g478850	TRAF		442	Cre12.g532900	TRAF	
443	Cre17.g700550	TRAF		444	Cre12.g520650	TUB	TLP1
445	Cre16.g667450	TUB		446	Cre02.g087551	unclassified	VPS54
447	Cre04.g221200	unclassified	CGL109	448	Cre05.g236300	unclassified	unclassified_11
449	Cre06.g257700	unclassified	unclassified_6	450	Cre06.g281200	unclassified	unclassified_13
451	Cre12.g511400	unclassified	PPR4	452	Cre17.g736850	unclassified	unclassified_8
453	Cre17.g738632	unclassified	unclassified_9	454	Cre03.g182700	unclassified	
455	Cre07.g325722	unclassified		456	Cre10.g417650	unclassified	
457	Cre12.g508644	unclassified		458	Cre16.g683050	unclassified	
459	Cre12.g510150	unclassified		460	Cre09.g393284	VARL	RLS8
461	Cre17.g738600	VARL	RLS5	462	Cre01.g010750	VARL	
463	Cre06.g302150	VARL		464	Cre06.g304250	VARL	
465	Cre09.g393654	VARL		466	Cre10.g418600	VARL	
467	Cre14.g617151	VARL		468	Cre14.g617200	VARL	
469	Cre15.g643700	VARL		470	Cre15.g643702	VARL	
471	Cre17.g709800	VARL		472	Cre02.g091550	Whirly	WHI,EZY18, PBF2
473	Cre04.g228400	WRKY	WRK1	474	Cre10.g464100	WRKY	
475	Cre12.g517451	WRKY					

Table A.6 represents the jaccard similarity index of the interaction overlap of MYB_9, LRL1, MS_1, CGL86, and HMG_6.

Table A.6: Jaccard index of interaction similarity between LRL1, MYB_9, MS_1, CGL86 and HMG_6

C. reinhardtii gene ID	TF	MYB_9	LRL1	MS_1	CGL86	HMG_6
Cre03.g144907	MYB_9	1.0	0.46	0.64	0.30	0.24
Cre03.g197100	LRL1	0.46	1.0	0.66	0.16	0.32
Cre10.g450500	MS_1	0.64	0.66	1.0	0.23	0.27
Cre12.g534450	CGL86	0.30	0.16	0.23	1.0	0.16
Cre16.g672300	HMG_6	0.24	0.32	0.27	0.17	1.0

Table A.7 displays the information which enzymatic promoters are targeted by LRL1 and the four potential key regulators MYB_9, MS_1, HMG_6, and CGL86. A one means that this particular TF interacts with the promoter in the corresponding row. A zero indicates no interaction.

Table A.7: List of target promoters of MYB_9, LRL1, MS_1, HMG_6, CGL86

Gene name	Gene ID	GO terms	LRL1	MYB_9	MS_1	HMG_6	CGL86
ACL(A)1	Cre05.g241850	Carbohydrate metabolism	1	1	1	1	1
ACOX2	Cre11.g467350	Lipid metabolism	1	1	1	0	1
ADH11	Cre11.g476550	Lipid metabolism	0	1	1	0	1
ALD6/ADH2	Cre12.g501050	Carbohydrate metabolism, RNA processing	0	0	0	1	0
AP2_6	Cre16.g667900	Unclassified	0	0	0	0	1
APE1	Cre16.g665250	Photosynthesis	0	0	0	0	1
APK1	Cre16.g677500	Cell signaling, Protein modification	1	1	1	1	0
APTX	Cre06.g297904	Cell signalin, DNA repair, Stress response	0	0	0	0	1
ARID_2	Cre13.g571050	Chromatin organization	0	0	0	0	1
ATO1	Cre17.g723650	Lipid metabolism, Stress response	0	0	0	1	1
BCC2	Cre01.g037850	Lipid metabolism	0	1	0	0	0
bHLH_2	Cre01.g011150	Unclassified	0	0	0	0	1
bHLH_3	Cre04.g216200	Unclassified	1	1	1	0	0
C2C2-GATA_3	Cre06.g266850	Unclassified	0	0	0	0	1
C2C2-GATA_4	Cre03.g160600	Unclassified	0	0	0	0	1
C2H2_2	Cre13.g567250	Unclassified	0	0	0	0	1
C2H2_6	Cre06.g305200	Unclassified	0	0	0	0	1

CAH3	Cre09.g415700	Photosynthesis	1	1	1	0	0
CGL107	Cre07.g341800	Chromatin organization, DNA repair, Stress response	0	0	0	0	1
CGL30	Cre03.g198950	Photosynthesis	0	1	0	0	1
CMP2	Cre06.g308500	Amino acid metabolism, Nitrogen metabolism, Stress response	0	0	0	1	0
CPP_1	Cre08.g361400	Cell cycle	0	0	0	0	1
Cre01.g038900	Cre01.g038900	Lipid metabolism	0	1	1	0	1
Cre03.g182050	Cre03.g182050	Lipid metabolism	1	1	1	0	0
Cre07.g334150	Cre07.g334150	Lipid metabolism	1	1	1	1	1
Cre08.g363500	Cre08.g363500	Lipid metabolism	0	1	0	0	0
Cre08.g365950	Cre08.g365950	Lipid metabolism	1	1	1	1	0
Cre08.g377300	Cre08.g377300	Lipid metabolism	1	1	1	0	1
Cre10.g425100	Cre10.g425100	Lipid metabolism	1	1	1	1	1
Cre14.g615050	Cre14.g615050	Lipid metabolism	0	1	0	0	1
CYP6	Cre02.g092350	Lipid metabolism	1	1	1	0	1
DLA2	Cre03.g158900	Carbohydrate metabolism	0	1	0	0	0
DRT1	Cre17.g739450	Cell signaling, Stress response	0	0	0	0	1
FAD7	Cre01.g038600	Lipid metabolism	0	1	1	0	0
FAP280,MBF1	Cre02.g108450	Unclassified	0	0	0	0	1
FBP4	Cre16.g652150	Carbohydrate metabolism, Cell signaling	1	1	1	1	0
FDH1	Cre12.g543400	Carbohydrate metabolism, Stress response	0	0	0	1	1
FNR1	Cre11.g476750	Photosynthesis	0	1	0	1	0
GAP1	Cre12.g485150	Carbohydrate metabolism	1	1	1	0	0
GGPS1	Cre12.g484200	Lipid metabolism	0	0	1	1	0

GLN1	Cre02.g113200	Amino acid metabolism, Nitrogen metabolism	1	0	1	1	0
GNAT_16	Cre01.g048800	Protein modification	0	0	0	0	1
GPX5	Cre10.g458450	Lipid metabolism, Stress response	1	1	1	1	1
HAS1	Cre01.g018500	Amino acid metabolism, Lipid metabolism	0	1	0	1	0
HDA16	Cre03.g180800	Chromatin organization	0	0	0	0	1
HYDA1	Cre03.g199800	Photosynthesis	0	1	0	0	1
KAR1	Cre03.g172000	Lipid metabolism	1	1	1	1	1
LHCB4	Cre17.g720250	Photosynthesis	1	1	1	1	0
LHCBM5	Cre03.g156900	Photosynthesis	0	1	0	0	1
LHCBM9	Cre06.g284200	Photosynthesis	0	1	1	1	1
LIP2	Cre12.g498750	Lipid metabolism	0	0	0	1	0
LPB1	Cre12.g554250	Carbohydrate metabolism	0	1	1	0	1
MECR, NRBF1	Cre01.g035350	Lipid metabolism	0	1	0	0	0
MME3	Cre14.g629700	Carbohydrate metabolism	0	0	0	1	1
MYB_13	Cre14.g621050	Unclassified	0	0	0	0	1
MYB_18	Cre16.g672400	Unclassified	0	0	0	0	1
MYB_21	Cre08.g364050	Unclassified	0	1	1	0	1
MYB_22	Cre02.g146629	Unclassified	0	0	0	0	1
NAR1.2	Cre06.g309000	Nitrogen metabolism, Transport	1	1	1	0	0
NAT36	Cre16.g662650	Nitrogen metabolism, Protein modification	0	0	0	0	1
NAT5	Cre13.g583200	Protein modification	1	1	1	0	0
NII1	Cre09.g410750	Nitrogen metabolism	0	1	0	0	1
NRR1	Cre16.g673250	Nitrogen metabolism, Stress response	1	1	0	1	1
Orphans_7	Cre07.g351550	Cell signaling	1	1	1	0	0

OTC1	Cre12.g489700	Amino acid metabolism, Nitrogen metabolism	1	1	1	1	1
PDH2	Cre03.g194200	Carbohydrate metabolism	0	1	0	0	0
PETM	Cre12.g546150	Photosynthesis	0	1	1	0	0
PHC45	Cre11.g483250	Lipid metabolism	0	1	0	0	0
PLAP2	Cre03.g188650	Lipid metabolism	0	1	1	0	1
PLAP7	Cre14.g611450	Lipid metabolism	0	1	1	0	0
PLAP8	Cre14.g611700	Lipid metabolism	0	0	0	1	0
PSAF	Cre09.g412100	Photosynthesis	0	0	0	0	1
PSAH	Cre07.g330250	Photosynthesis	1	1	1	0	1
PSBP1	Cre12.g550850	Photosynthesis	0	1	1	0	1
PSBP6	Cre07.g328200	Photosynthesis	0	1	1	0	0
PSBX	Cre02.g082750	Photosynthesis	0	1	0	0	1
Pseudo ARR-B_2	Cre16.g676421	Cell signaling	0	0	0	0	1
ROC40	Cre06.g275350	Cell cycle	0	0	0	0	1
RPE2	Cre02.g116450	Carbohydrate metabolism	0	1	0	0	0
S1Fa-like_2	Cre09.g386753	Unclassified	1	1	1	0	1
SQD1	Cre16.g656400	Lipid metabolism Photosynthesis Stress response	0	1	0	1	1
SRH12	Cre06.g270850	DNA repair Stress response	1	1	1	0	0
SRH2	Cre02.g087600	Chromatin organization Transport	0	0	0	0	1
SRH4	Cre03.g158550	Chromatin organization DNA repair Stress response	0	0	0	0	1
TAL1	Cre01.g032650	Carbohydrate metabolism	0	1	0	0	0
TGL18	Cre12.g530950	Lipid metabolism	1	1	1	0	0
TRAF_11	Cre01.g022950	Unclassified	0	0	0	0	1
TRAF_12	Cre01.g048501	Protein modification	0	0	0	0	1
TRAF_3	Cre04.g221950	Unclassified	0	0	0	0	1
TRAF_6	Cre02.g101150	Unclassified	0	0	0	0	1
TRAF_8	Cre04.g215450	Unclassified	0	0	0	0	1
TRAF_9	Cre06.g286700	Unclassified	0	0	0	0	1
WHI,EZY18, PBF2	Cre02.g091550	Unclassified	0	0	0	0	1

All found and verified interactions in the Y1H Screen are shown in table A.8.

Table A.8: List of found PDIs in Y1H

TF	promoter	TF	promoter	TF	promoter
Cre01.g001600	Cre02.g078831	Cre01.g001600	Cre07.g334150	Cre01.g001600	Cre03.g199800
Cre01.g011150	Cre07.g330250	Cre01.g011150	Cre07.g334150	Cre01.g011150	Cre12.g560950
Cre01.g011150	Cre05.g241850	Cre01.g012200	Cre09.g386753	Cre01.g012200	Cre10.g425100
Cre01.g012200	Cre16.g652150	Cre01.g025050	Cre07.g334150	Cre01.g029450	Cre03.g180800
Cre01.g029450	Cre14.g621172	Cre01.g029450	Cre01.g038600	Cre01.g062902	Cre08.g377300
Cre01.g062902	Cre12.g501050	Cre02.g078700	Cre03.g180800	Cre02.g078700	Cre14.g621172
Cre02.g078700	Cre03.g198950	Cre02.g078700	Cre12.g501050	Cre02.g078700	Cre09.g410750
Cre02.g078700	Cre01.g038600	Cre02.g078831	Cre03.g188650	Cre02.g084550	Cre12.g550850
Cre02.g087600	Cre07.g328200	Cre02.g087600	Cre07.g334150	Cre02.g091550	Cre02.g113200
Cre02.g091550	Cre14.g611700	Cre02.g091550	Cre12.g484200	Cre02.g091550	Cre16.g677500
Cre02.g091550	Cre17.g723650	Cre02.g091550	Cre06.g308500	Cre02.g095043	Cre03.g160600
Cre02.g095043	Cre08.g364050	Cre02.g095043	Cre09.g386753	Cre02.g095043	Cre16.g662650
Cre02.g095043	Cre16.g667900	Cre02.g095043	Cre12.g489700	Cre02.g095043	Cre17.g720250
Cre02.g095043	Cre10.g458450	Cre02.g095043	Cre03.g158900	Cre02.g095043	Cre03.g199800
Cre02.g095043	Cre12.g543400	Cre02.g101300	Cre07.g328200	Cre02.g101300	Cre16.g665250
Cre02.g101300	Cre10.g463150	Cre02.g108350	Cre16.g652150	Cre02.g108350	Cre05.g241850
Cre02.g110150	Cre08.g364050	Cre02.g110150	Cre09.g386753	Cre02.g110150	Cre16.g667900
Cre02.g118250	Cre07.g328200	Cre02.g118250	Cre07.g334150	Cre02.g145602	Cre02.g078831
Cre02.g145602	Cre13.g567250	Cre02.g145602	Cre12.g489700	Cre02.g145602	Cre07.g330250
Cre02.g145602	Cre10.g458450	Cre02.g145602	Cre07.g328200	Cre02.g145602	Cre16.g665250
Cre02.g145602	Cre12.g530600	Cre02.g145602	Cre07.g334150	Cre02.g145602	Cre09.g410750
Cre02.g145602	Cre12.g508550	Cre02.g145602	Cre03.g158900	Cre02.g145602	Cre03.g199800
Cre03.g144907	Cre04.g216200	Cre03.g144907	Cre06.g270850	Cre03.g144907	Cre07.g351550
Cre03.g144907	Cre08.g364050	Cre03.g144907	Cre09.g386753	Cre03.g144907	Cre13.g583200
Cre03.g144907	Cre11.g476750	Cre03.g144907	Cre01.g038900	Cre03.g144907	Cre10.g425100
Cre03.g144907	Cre12.g489700	Cre03.g144907	Cre07.g330250	Cre03.g144907	Cre14.g611450
Cre03.g144907	Cre03.g198950	Cre03.g144907	Cre16.g652150	Cre03.g144907	Cre12.g485150
Cre03.g144907	Cre17.g720250	Cre03.g144907	Cre10.g458450	Cre03.g144907	Cre01.g035350
Cre03.g144907	Cre12.g530950	Cre03.g144907	Cre14.g615050	Cre03.g144907	Cre03.g194200
Cre03.g144907	Cre03.g182050	Cre03.g144907	Cre08.g377300	Cre03.g144907	Cre02.g092350
Cre03.g144907	Cre06.g309000	Cre03.g144907	Cre03.g172000	Cre03.g144907	Cre12.g546150
Cre03.g144907	Cre07.g328200	Cre03.g144907	Cre11.g483250	Cre03.g144907	Cre12.g550850
Cre03.g144907	Cre02.g082750	Cre03.g144907	Cre02.g116450	Cre03.g144907	Cre16.g656400
Cre03.g144907	Cre01.g032650	Cre03.g144907	Cre16.g677500	Cre03.g144907	Cre06.g284200
Cre03.g144907	Cre01.g018500	Cre03.g144907	Cre07.g334150	Cre03.g144907	Cre11.g467350
Cre03.g144907	Cre11.g476550	Cre03.g144907	Cre16.g673250	Cre03.g144907	Cre08.g363500
Cre03.g144907	Cre09.g410750	Cre03.g144907	Cre03.g158900	Cre03.g144907	Cre03.g199800
Cre03.g144907	Cre03.g156900	Cre03.g144907	Cre12.g554250	Cre03.g144907	Cre03.g188650
Cre03.g144907	Cre08.g365950	Cre03.g144907	Cre01.g037850	Cre03.g144907	Cre09.g415700
Cre03.g144907	Cre01.g038600	Cre03.g144907	Cre05.g241850	Cre03.g149201	Cre01.g038900
Cre03.g149201	Cre07.g334150	Cre03.g152150	Cre07.g334150	Cre03.g152150	Cre03.g158900
Cre03.g152150	Cre03.g199800	Cre03.g152150	Cre12.g560950	Cre03.g152150	Cre12.g543400
Cre03.g152450	Cre16.g656400	Cre03.g178550	Cre07.g351550	Cre03.g178550	Cre13.g586450
Cre03.g178550	Cre14.g629700	Cre03.g178550	Cre12.g543400	Cre03.g179650	Cre16.g656400
Cre03.g184151	Cre12.g550850	Cre03.g185950	Cre12.g489700	Cre03.g185950	Cre09.g410750
Cre03.g185950	Cre03.g158900	Cre03.g188200	Cre01.g037850	Cre03.g189000	Cre17.g720250
Cre03.g189000	Cre03.g199800	Cre03.g197100	Cre04.g216200	Cre03.g197100	Cre06.g270850
Cre03.g197100	Cre07.g351550	Cre03.g197100	Cre09.g386753	Cre03.g197100	Cre13.g583200
Cre03.g197100	Cre02.g113200	Cre03.g197100	Cre10.g425100	Cre03.g197100	Cre12.g489700
Cre03.g197100	Cre07.g330250	Cre03.g197100	Cre16.g652150	Cre03.g197100	Cre12.g485150

Cre03.g197100	Cre17.g720250	Cre03.g197100	Cre10.g458450	Cre03.g197100	Cre12.g530950
Cre03.g197100	Cre03.g182050	Cre03.g197100	Cre08.g377300	Cre03.g197100	Cre02.g092350
Cre03.g197100	Cre06.g309000	Cre03.g197100	Cre03.g172000	Cre03.g197100	Cre16.g677500
Cre03.g197100	Cre07.g334150	Cre03.g197100	Cre11.g467350	Cre03.g197100	Cre16.g673250
Cre03.g197100	Cre08.g365950	Cre03.g197100	Cre09.g415700	Cre03.g197100	Cre05.g241850
Cre03.g201850	Cre03.g199800	Cre03.g202000	Cre03.g199800	Cre03.g209953	Cre01.g022950
Cre04.g211700	Cre16.g665250	Cre04.g211700	Cre07.g334150	Cre04.g213650	Cre02.g078831
Cre04.g213650	Cre07.g330250	Cre04.g213650	Cre16.g665250	Cre04.g213650	Cre07.g334150
Cre04.g213650	Cre03.g158900	Cre04.g213650	Cre03.g199800	Cre04.g216200	Cre10.g425100
Cre04.g216200	Cre16.g673250	Cre04.g218400	Cre09.g386753	Cre04.g218400	Cre03.g199800
Cre04.g219600	Cre01.g022950	Cre04.g219600	Cre02.g087600	Cre04.g219600	Cre02.g145550
Cre04.g219600	Cre03.g144424	Cre04.g219600	Cre03.g179000	Cre04.g219600	Cre03.g200431
Cre04.g219600	Cre06.g266850	Cre04.g219600	Cre06.g297904	Cre04.g219600	Cre07.g351500
Cre04.g219600	Cre08.g364050	Cre04.g219600	Cre09.g386753	Cre04.g219600	Cre12.g503800
Cre04.g219600	Cre13.g570150	Cre04.g219600	Cre13.g586450	Cre04.g219600	Cre16.g662650
Cre04.g219600	Cre16.g667900	Cre04.g219600	Cre07.g330250	Cre04.g219600	Cre16.g656400
Cre04.g219600	Cre07.g334150	Cre04.g219600	Cre11.g467350	Cre04.g219600	Cre16.g673250
Cre04.g219600	Cre09.g410750	Cre04.g219600	Cre03.g199800	Cre04.g219600	Cre03.g188650
Cre04.g219600	Cre12.g560950	Cre04.g221900	Cre07.g328200	Cre04.g221900	Cre07.g334150
Cre04.g224600	Cre07.g330250	Cre04.g224600	Cre06.g270200	Cre04.g224600	Cre16.g677500
Cre04.g224600	Cre12.g530600	Cre04.g224600	Cre16.g673250	Cre04.g229948	Cre09.g390023
Cre04.g229948	Cre16.g653250	Cre04.g229948	Cre12.g543400	Cre04.g231124	Cre01.g004600
Cre04.g231124	Cre01.g011150	Cre04.g231124	Cre01.g022950	Cre04.g231124	Cre01.g025050
Cre04.g231124	Cre01.g043550	Cre04.g231124	Cre01.g048501	Cre04.g231124	Cre01.g048800
Cre04.g231124	Cre02.g082550	Cre04.g231124	Cre02.g083750	Cre04.g231124	Cre02.g093950
Cre04.g231124	Cre02.g101150	Cre04.g231124	Cre02.g101500	Cre04.g231124	Cre02.g104700
Cre04.g231124	Cre02.g108450	Cre04.g231124	Cre02.g145550	Cre04.g231124	Cre03.g144747
Cre04.g231124	Cre03.g149350	Cre04.g231124	Cre03.g149400	Cre04.g231124	Cre03.g158550
Cre04.g231124	Cre03.g160600	Cre04.g231124	Cre03.g173350	Cre04.g231124	Cre03.g184600
Cre04.g231124	Cre03.g188200	Cre04.g231124	Cre03.g189650	Cre04.g231124	Cre03.g193400
Cre04.g231124	Cre03.g200431	Cre04.g231124	Cre03.g209953	Cre04.g231124	Cre03.g212641
Cre04.g231124	Cre04.g221200	Cre04.g231124	Cre04.g221950	Cre04.g231124	Cre04.g226400
Cre04.g231124	Cre04.g231124	Cre04.g231124	Cre05.g233551	Cre04.g231124	Cre06.g257700
Cre04.g231124	Cre06.g262800	Cre04.g231124	Cre06.g266850	Cre04.g231124	Cre06.g266950
Cre04.g231124	Cre06.g270850	Cre04.g231124	Cre06.g275350	Cre04.g231124	Cre06.g275550
Cre04.g231124	Cre06.g282251	Cre04.g231124	Cre06.g285600	Cre04.g231124	Cre06.g297904
Cre04.g231124	Cre06.g305200	Cre04.g231124	Cre07.g319701	Cre04.g231124	Cre07.g321550
Cre04.g231124	Cre07.g330050	Cre04.g231124	Cre07.g332250	Cre04.g231124	Cre07.g341800
Cre04.g231124	Cre07.g351500	Cre04.g231124	Cre07.g351550	Cre04.g231124	Cre07.g354100
Cre04.g231124	Cre07.g354500	Cre04.g231124	Cre07.g355700	Cre04.g231124	Cre08.g361400
Cre04.g231124	Cre08.g364050	Cre04.g231124	Cre08.g371052	Cre04.g231124	Cre09.g386753
Cre04.g231124	Cre09.g390023	Cre04.g231124	Cre09.g411600	Cre04.g231124	Cre10.g429880
Cre04.g231124	Cre10.g436800	Cre04.g231124	Cre10.g466550	Cre04.g231124	Cre11.g467577
Cre04.g231124	Cre11.g479800	Cre04.g231124	Cre12.g485950	Cre04.g231124	Cre12.g502251
Cre04.g231124	Cre12.g503800	Cre04.g231124	Cre12.g507500	Cre04.g231124	Cre12.g517350
Cre04.g231124	Cre12.g520850	Cre04.g231124	Cre12.g538550	Cre04.g231124	Cre12.g540400
Cre04.g231124	Cre13.g571050	Cre04.g231124	Cre13.g571200	Cre04.g231124	Cre13.g580450
Cre04.g231124	Cre13.g586450	Cre04.g231124	Cre13.g590350	Cre04.g231124	Cre14.g611000
Cre04.g231124	Cre14.g621172	Cre04.g231124	Cre14.g623800	Cre04.g231124	Cre16.g649433
Cre04.g231124	Cre16.g650400	Cre04.g231124	Cre16.g653250	Cre04.g231124	Cre16.g662650
Cre04.g231124	Cre16.g667900	Cre04.g231124	Cre16.g668200	Cre04.g231124	Cre16.g672400
Cre04.g231124	Cre16.g678900	Cre04.g231124	Cre16.g680232	Cre04.g231124	Cre16.g683953
Cre04.g231124	Cre16.g686250	Cre04.g231124	Cre16.g692500	Cre04.g231124	Cre17.g702650
Cre04.g231124	Cre17.g705950	Cre04.g231124	Cre17.g714229	Cre04.g231124	Cre17.g746547
Cre04.g231124	Cre12.g489700	Cre04.g231124	Cre16.g652150	Cre04.g231124	Cre12.g485150
Cre04.g231124	Cre03.g172000	Cre04.g231124	Cre16.g656400	Cre04.g231124	Cre16.g665250

Cre04.g231124	Cre06.g284200	Cre04.g231124	Cre12.g498750	Cre04.g231124	Cre01.g018500
Cre04.g231124	Cre11.g467350	Cre04.g231124	Cre16.g673250	Cre04.g231124	Cre03.g156900
Cre05.g236300	Cre12.g543400	Cre05.g240600	Cre03.g158900	Cre06.g250950	Cre01.g004600
Cre06.g250950	Cre01.g011150	Cre06.g250950	Cre01.g022950	Cre06.g250950	Cre01.g025050
Cre06.g250950	Cre01.g043550	Cre06.g250950	Cre01.g048501	Cre06.g250950	Cre01.g048800
Cre06.g250950	Cre02.g082550	Cre06.g250950	Cre02.g083750	Cre06.g250950	Cre02.g093950
Cre06.g250950	Cre02.g101150	Cre06.g250950	Cre02.g101500	Cre06.g250950	Cre02.g104700
Cre06.g250950	Cre02.g108450	Cre06.g250950	Cre02.g145550	Cre06.g250950	Cre03.g144747
Cre06.g250950	Cre03.g149350	Cre06.g250950	Cre03.g149400	Cre06.g250950	Cre03.g158550
Cre06.g250950	Cre03.g160600	Cre06.g250950	Cre03.g173350	Cre06.g250950	Cre03.g184600
Cre06.g250950	Cre03.g188200	Cre06.g250950	Cre03.g189650	Cre06.g250950	Cre03.g193400
Cre06.g250950	Cre03.g200431	Cre06.g250950	Cre03.g209953	Cre06.g250950	Cre03.g212641
Cre06.g250950	Cre04.g221200	Cre06.g250950	Cre04.g221950	Cre06.g250950	Cre04.g226400
Cre06.g250950	Cre04.g231124	Cre06.g250950	Cre05.g233551	Cre06.g250950	Cre06.g257700
Cre06.g250950	Cre06.g262800	Cre06.g250950	Cre06.g266850	Cre06.g250950	Cre06.g266950
Cre06.g250950	Cre06.g270850	Cre06.g250950	Cre06.g275350	Cre06.g250950	Cre06.g275550
Cre06.g250950	Cre06.g282251	Cre06.g250950	Cre06.g285600	Cre06.g250950	Cre06.g297904
Cre06.g250950	Cre06.g305200	Cre06.g250950	Cre07.g319701	Cre06.g250950	Cre07.g321550
Cre06.g250950	Cre07.g330050	Cre06.g250950	Cre07.g332250	Cre06.g250950	Cre07.g341800
Cre06.g250950	Cre07.g351500	Cre06.g250950	Cre07.g351550	Cre06.g250950	Cre07.g354100
Cre06.g250950	Cre07.g354500	Cre06.g250950	Cre07.g355700	Cre06.g250950	Cre08.g361400
Cre06.g250950	Cre08.g364050	Cre06.g250950	Cre08.g371052	Cre06.g250950	Cre09.g386753
Cre06.g250950	Cre09.g411600	Cre06.g250950	Cre10.g429880	Cre06.g250950	Cre10.g436800
Cre06.g250950	Cre10.g466550	Cre06.g250950	Cre11.g467577	Cre06.g250950	Cre11.g479800
Cre06.g250950	Cre12.g485950	Cre06.g250950	Cre12.g502251	Cre06.g250950	Cre12.g503800
Cre06.g250950	Cre12.g507500	Cre06.g250950	Cre12.g517350	Cre06.g250950	Cre12.g520850
Cre06.g250950	Cre12.g538550	Cre06.g250950	Cre12.g540400	Cre06.g250950	Cre13.g571050
Cre06.g250950	Cre13.g571200	Cre06.g250950	Cre13.g580450	Cre06.g250950	Cre13.g586450
Cre06.g250950	Cre13.g590350	Cre06.g250950	Cre14.g611000	Cre06.g250950	Cre14.g621172
Cre06.g250950	Cre14.g623800	Cre06.g250950	Cre16.g649433	Cre06.g250950	Cre16.g650400
Cre06.g250950	Cre16.g662650	Cre06.g250950	Cre16.g667900	Cre06.g250950	Cre16.g668200
Cre06.g250950	Cre16.g672400	Cre06.g250950	Cre16.g678900	Cre06.g250950	Cre16.g680232
Cre06.g250950	Cre16.g683953	Cre06.g250950	Cre16.g686250	Cre06.g250950	Cre16.g692500
Cre06.g250950	Cre17.g702650	Cre06.g250950	Cre17.g705950	Cre06.g250950	Cre17.g714229
Cre06.g250950	Cre17.g746547	Cre06.g250950	Cre12.g489700	Cre06.g250950	Cre16.g652150
Cre06.g250950	Cre12.g485150	Cre06.g250950	Cre10.g458450	Cre06.g250950	Cre03.g172000
Cre06.g250950	Cre16.g656400	Cre06.g250950	Cre16.g665250	Cre06.g250950	Cre12.g530600
Cre06.g250950	Cre06.g284200	Cre06.g250950	Cre12.g498750	Cre06.g250950	Cre01.g018500
Cre06.g250950	Cre11.g467350	Cre06.g250950	Cre16.g673250	Cre06.g250950	Cre03.g158900
Cre06.g250950	Cre03.g199800	Cre06.g250950	Cre03.g156900	Cre06.g250950	Cre12.g543400
Cre06.g256200	Cre07.g334150	Cre06.g261450	Cre03.g180800	Cre06.g261450	Cre08.g377300
Cre06.g268600	Cre03.g188650	Cre06.g269100	Cre13.g586450	Cre06.g270350	Cre09.g386753
Cre06.g270350	Cre03.g199800	Cre06.g270350	Cre12.g543400	Cre06.g270850	Cre07.g328200
Cre06.g270850	Cre03.g158900	Cre06.g275600	Cre09.g386753	Cre06.g275600	Cre17.g720250
Cre06.g275600	Cre03.g199800	Cre06.g275600	Cre12.g543400	Cre06.g278108	Cre03.g199800
Cre06.g278136	Cre01.g004600	Cre06.g278136	Cre01.g022950	Cre06.g278136	Cre01.g025050
Cre06.g278136	Cre01.g043550	Cre06.g278136	Cre01.g048501	Cre06.g278136	Cre01.g048800
Cre06.g278136	Cre02.g082550	Cre06.g278136	Cre02.g093950	Cre06.g278136	Cre02.g101150
Cre06.g278136	Cre02.g101500	Cre06.g278136	Cre02.g104700	Cre06.g278136	Cre02.g108450
Cre06.g278136	Cre02.g145550	Cre06.g278136	Cre03.g144747	Cre06.g278136	Cre03.g149350
Cre06.g278136	Cre03.g149400	Cre06.g278136	Cre03.g158550	Cre06.g278136	Cre03.g173350
Cre06.g278136	Cre03.g184600	Cre06.g278136	Cre03.g188200	Cre06.g278136	Cre03.g189650
Cre06.g278136	Cre03.g193400	Cre06.g278136	Cre03.g200431	Cre06.g278136	Cre03.g209953
Cre06.g278136	Cre03.g212641	Cre06.g278136	Cre04.g221200	Cre06.g278136	Cre04.g221950
Cre06.g278136	Cre04.g226400	Cre06.g278136	Cre04.g231124	Cre06.g278136	Cre05.g233551
Cre06.g278136	Cre06.g262800	Cre06.g278136	Cre06.g266850	Cre06.g278136	Cre06.g266950

Cre06.g278136	Cre06.g270850	Cre06.g278136	Cre06.g275350	Cre06.g278136	Cre06.g275550
Cre06.g278136	Cre06.g282251	Cre06.g278136	Cre06.g285600	Cre06.g278136	Cre06.g297904
Cre06.g278136	Cre06.g305200	Cre06.g278136	Cre07.g319701	Cre06.g278136	Cre07.g321550
Cre06.g278136	Cre07.g330050	Cre06.g278136	Cre07.g332250	Cre06.g278136	Cre07.g341800
Cre06.g278136	Cre07.g351500	Cre06.g278136	Cre07.g351550	Cre06.g278136	Cre07.g354100
Cre06.g278136	Cre07.g354500	Cre06.g278136	Cre07.g355700	Cre06.g278136	Cre08.g361400
Cre06.g278136	Cre08.g364050	Cre06.g278136	Cre08.g371052	Cre06.g278136	Cre09.g386753
Cre06.g278136	Cre09.g390023	Cre06.g278136	Cre09.g411600	Cre06.g278136	Cre10.g429880
Cre06.g278136	Cre10.g436800	Cre06.g278136	Cre10.g466550	Cre06.g278136	Cre11.g467577
Cre06.g278136	Cre11.g479800	Cre06.g278136	Cre12.g485950	Cre06.g278136	Cre12.g503800
Cre06.g278136	Cre12.g507500	Cre06.g278136	Cre12.g517350	Cre06.g278136	Cre12.g520850
Cre06.g278136	Cre12.g538550	Cre06.g278136	Cre12.g540400	Cre06.g278136	Cre13.g571050
Cre06.g278136	Cre13.g571200	Cre06.g278136	Cre13.g580450	Cre06.g278136	Cre13.g586450
Cre06.g278136	Cre13.g590350	Cre06.g278136	Cre14.g611000	Cre06.g278136	Cre14.g621172
Cre06.g278136	Cre14.g623800	Cre06.g278136	Cre16.g649433	Cre06.g278136	Cre16.g650400
Cre06.g278136	Cre16.g662650	Cre06.g278136	Cre16.g667900	Cre06.g278136	Cre16.g668200
Cre06.g278136	Cre16.g672400	Cre06.g278136	Cre16.g678900	Cre06.g278136	Cre16.g680232
Cre06.g278136	Cre16.g683953	Cre06.g278136	Cre16.g686250	Cre06.g278136	Cre16.g692500
Cre06.g278136	Cre17.g702650	Cre06.g278136	Cre17.g705950	Cre06.g278136	Cre17.g714229
Cre06.g278136	Cre17.g746547	Cre06.g278136	Cre14.g629700	Cre06.g278136	Cre12.g489700
Cre06.g278136	Cre12.g485150	Cre06.g278136	Cre10.g458450	Cre06.g278136	Cre03.g172000
Cre06.g278136	Cre16.g656400	Cre06.g278136	Cre16.g665250	Cre06.g278136	Cre12.g530600
Cre06.g278136	Cre06.g284200	Cre06.g278136	Cre12.g498750	Cre06.g278136	Cre07.g334150
Cre06.g278136	Cre16.g673250	Cre06.g278136	Cre03.g158900	Cre06.g278136	Cre03.g199800
Cre06.g278136	Cre03.g156900	Cre06.g278136	Cre12.g543400	Cre06.g282251	Cre03.g199800
Cre06.g283200	Cre03.g160600	Cre06.g283200	Cre08.g364050	Cre06.g283200	Cre09.g386753
Cre06.g283200	Cre16.g662650	Cre06.g283200	Cre16.g667900	Cre06.g283200	Cre10.g425100
Cre06.g283200	Cre16.g652150	Cre06.g283200	Cre06.g309000	Cre06.g283200	Cre11.g467350
Cre06.g284000	Cre07.g330250	Cre06.g284000	Cre16.g656400	Cre06.g284000	Cre03.g199800
Cre06.g284000	Cre03.g188650	Cre06.g285600	Cre03.g194200	Cre06.g285600	Cre12.g485950
Cre06.g285600	Cre16.g676421	Cre06.g285600	Cre06.g281200	Cre06.g285600	Cre08.g361400
Cre06.g285600	Cre12.g538550	Cre06.g285600	Cre05.g240650	Cre06.g285600	Cre07.g341800
Cre06.g285600	Cre16.g672400	Cre06.g285600	Cre12.g558950	Cre06.g285600	Cre01.g011150
Cre06.g285600	Cre06.g270850	Cre06.g285600	Cre16.g649433	Cre06.g285600	Cre06.g305200
Cre06.g285600	Cre11.g475950	Cre06.g285600	Cre08.g364050	Cre06.g285600	Cre01.g035350
Cre06.g285600	Cre12.g543400	Cre06.g285600	Cre12.g560950	Cre06.g285600	Cre12.g554250
Cre06.g285600	Cre11.g476550	Cre06.g285600	Cre10.g458450	Cre06.g285600	Cre17.g720250
Cre06.g285600	Cre16.g652150	Cre06.g286700	Cre09.g386753	Cre06.g286700	Cre16.g656400
Cre06.g286700	Cre07.g334150	Cre06.g286700	Cre03.g199800	Cre06.g286700	Cre12.g560950
Cre06.g289900	Cre10.g425100	Cre06.g293750	Cre01.g004600	Cre06.g293750	Cre01.g011150
Cre06.g293750	Cre03.g149400	Cre06.g293750	Cre03.g188200	Cre06.g293750	Cre06.g282251
Cre06.g293750	Cre06.g305200	Cre06.g293750	Cre07.g330050	Cre06.g293750	Cre09.g386753
Cre06.g293750	Cre13.g571050	Cre06.g293750	Cre13.g580450	Cre06.g293750	Cre16.g672400
Cre06.g293750	Cre16.g680232	Cre06.g297150	Cre10.g458450	Cre06.g297150	Cre09.g410750
Cre06.g305200	Cre01.g011150	Cre06.g305200	Cre01.g048800	Cre06.g305200	Cre03.g149350
Cre06.g305200	Cre06.g285600	Cre06.g305200	Cre13.g590350	Cre06.g305200	Cre14.g623800
Cre06.g305200	Cre16.g649433	Cre06.g305200	Cre16.g680232	Cre06.g305200	Cre12.g489700
Cre06.g305200	Cre16.g652150	Cre06.g305200	Cre10.g458450	Cre06.g305200	Cre03.g172000
Cre06.g305200	Cre16.g656400	Cre06.g305200	Cre12.g530600	Cre06.g305200	Cre06.g284200
Cre06.g305200	Cre12.g498750	Cre06.g305200	Cre07.g334150	Cre06.g305200	Cre03.g158900
Cre06.g305200	Cre03.g199800	Cre06.g305200	Cre03.g156900	Cre07.g319701	Cre14.g611700
Cre07.g319701	Cre07.g334150	Cre07.g321550	Cre01.g004600	Cre07.g321550	Cre01.g025050
Cre07.g321550	Cre01.g048501	Cre07.g321550	Cre02.g094150	Cre07.g321550	Cre02.g118250
Cre07.g321550	Cre03.g160600	Cre07.g321550	Cre03.g184600	Cre07.g321550	Cre04.g228400
Cre07.g321550	Cre05.g233551	Cre07.g321550	Cre06.g266850	Cre07.g321550	Cre06.g275600
Cre07.g321550	Cre06.g293750	Cre07.g321550	Cre06.g297904	Cre07.g321550	Cre07.g329900

Cre07.g321550	Cre09.g386753	Cre07.g321550	Cre10.g430950	Cre07.g321550	Cre10.g446850
Cre07.g321550	Cre12.g485950	Cre07.g321550	Cre13.g583150	Cre07.g321550	Cre14.g621050
Cre07.g321550	Cre16.g662650	Cre07.g321550	Cre16.g667900	Cre07.g321550	Cre17.g702650
Cre07.g321550	Cre17.g714229	Cre07.g321550	Cre17.g739450	Cre07.g321550	Cre17.g746547
Cre07.g321550	Cre10.g425100	Cre07.g321550	Cre16.g652150	Cre07.g321550	Cre10.g458450
Cre07.g321550	Cre02.g092350	Cre07.g321550	Cre01.g032650	Cre07.g321550	Cre12.g498750
Cre07.g321550	Cre11.g476550	Cre07.g321550	Cre16.g673250	Cre07.g321550	Cre05.g241850
Cre07.g335500	Cre03.g199800	Cre07.g341800	Cre12.g501050	Cre07.g344668	Cre01.g004600
Cre07.g344668	Cre01.g025050	Cre07.g344668	Cre01.g048501	Cre07.g344668	Cre02.g094150
Cre07.g344668	Cre02.g118250	Cre07.g344668	Cre03.g160600	Cre07.g344668	Cre03.g184600
Cre07.g344668	Cre03.g197350	Cre07.g344668	Cre04.g228400	Cre07.g344668	Cre05.g233551
Cre07.g344668	Cre06.g266850	Cre07.g344668	Cre06.g275600	Cre07.g344668	Cre06.g293750
Cre07.g344668	Cre06.g297904	Cre07.g344668	Cre07.g329900	Cre07.g344668	Cre09.g386753
Cre07.g344668	Cre10.g430950	Cre07.g344668	Cre10.g446850	Cre07.g344668	Cre12.g485950
Cre07.g344668	Cre13.g583150	Cre07.g344668	Cre14.g621050	Cre07.g344668	Cre16.g662650
Cre07.g344668	Cre16.g667900	Cre07.g344668	Cre17.g702650	Cre07.g344668	Cre17.g714229
Cre07.g344668	Cre17.g739450	Cre07.g344668	Cre17.g746547	Cre07.g344668	Cre16.g652150
Cre07.g344668	Cre10.g458450	Cre07.g344668	Cre02.g092350	Cre07.g344668	Cre16.g656400
Cre07.g344668	Cre01.g032650	Cre07.g344668	Cre01.g018500	Cre07.g344668	Cre11.g476550
Cre07.g344668	Cre16.g673250	Cre07.g344668	Cre06.g266200	Cre07.g345350	Cre14.g629700
Cre07.g345350	Cre07.g330250	Cre07.g345350	Cre08.g377300	Cre07.g345350	Cre02.g092350
Cre07.g352600	Cre12.g543400	Cre07.g353500	Cre08.g364050	Cre07.g353500	Cre09.g386753
Cre07.g353500	Cre07.g330250	Cre07.g353500	Cre03.g188650	Cre07.g353555	Cre09.g415700
Cre07.g354100	Cre12.g550850	Cre07.g354500	Cre01.g032650	Cre08.g358534	Cre10.g425100
Cre08.g358534	Cre12.g489700	Cre08.g364400	Cre08.g364050	Cre08.g364400	Cre03.g188650
Cre08.g364450	Cre12.g550850	Cre08.g378450	Cre12.g489700	Cre08.g378450	Cre14.g611700
Cre08.g378450	Cre16.g677500	Cre08.g378450	Cre07.g334150	Cre08.g378450	Cre06.g308500
Cre08.g378800	Cre07.g334150	Cre08.g378800	Cre06.g308500	Cre09.g386753	Cre10.g458450
Cre09.g386753	Cre03.g194200	Cre09.g386753	Cre12.g501050	Cre09.g390430	Cre14.g629700
Cre09.g392300	Cre07.g334150	Cre09.g392300	Cre03.g199800	Cre09.g392300	Cre12.g554250
Cre09.g393284	Cre03.g160600	Cre09.g393284	Cre08.g364050	Cre09.g393284	Cre09.g386753
Cre09.g393284	Cre16.g662650	Cre09.g393284	Cre16.g667900	Cre09.g393284	Cre12.g489700
Cre09.g393284	Cre17.g720250	Cre09.g393284	Cre10.g458450	Cre09.g393284	Cre02.g116450
Cre09.g393284	Cre03.g158900	Cre09.g393284	Cre03.g199800	Cre09.g393284	Cre12.g543400
Cre09.g394200	Cre16.g673250	Cre09.g399552	Cre10.g425100	Cre09.g399552	Cre07.g330250
Cre09.g399552	Cre16.g652150	Cre09.g401219	Cre02.g078831	Cre09.g401219	Cre07.g330250
Cre09.g401219	Cre07.g328200	Cre09.g401219	Cre16.g665250	Cre09.g401219	Cre07.g334150
Cre09.g401219	Cre09.g410750	Cre09.g401219	Cre03.g158900	Cre09.g401219	Cre03.g199800
Cre09.g402350	Cre04.g231124	Cre09.g410450	Cre04.g216200	Cre09.g410450	Cre06.g270850
Cre09.g410450	Cre07.g351550	Cre09.g410450	Cre08.g364050	Cre09.g410450	Cre09.g386753
Cre09.g410450	Cre13.g583200	Cre09.g410450	Cre02.g113200	Cre09.g410450	Cre01.g038900
Cre09.g410450	Cre10.g425100	Cre09.g410450	Cre12.g489700	Cre09.g410450	Cre07.g330250
Cre09.g410450	Cre16.g652150	Cre09.g410450	Cre10.g458450	Cre09.g410450	Cre03.g194200
Cre09.g410450	Cre03.g182050	Cre09.g410450	Cre06.g309000	Cre09.g410450	Cre05.g241850
Cre10.g429880	Cre01.g022950	Cre10.g430750	Cre07.g334150	Cre10.g438850	Cre01.g025050
Cre10.g438850	Cre02.g094150	Cre10.g438850	Cre04.g228400	Cre10.g438850	Cre07.g329900
Cre10.g438850	Cre10.g430950	Cre10.g438850	Cre10.g446850	Cre10.g438850	Cre12.g485950
Cre10.g438850	Cre17.g702650	Cre10.g438850	Cre17.g746547	Cre10.g438850	Cre16.g652150
Cre10.g438850	Cre02.g092350	Cre10.g438850	Cre01.g032650	Cre10.g438850	Cre11.g476550
Cre10.g438850	Cre05.g241850	Cre10.g446450	Cre16.g673250	Cre10.g446450	Cre03.g199800
Cre10.g450500	Cre04.g216200	Cre10.g450500	Cre06.g270850	Cre10.g450500	Cre07.g351550
Cre10.g450500	Cre08.g364050	Cre10.g450500	Cre09.g386753	Cre10.g450500	Cre13.g583200
Cre10.g450500	Cre02.g113200	Cre10.g450500	Cre01.g038900	Cre10.g450500	Cre10.g425100
Cre10.g450500	Cre12.g489700	Cre10.g450500	Cre07.g330250	Cre10.g450500	Cre14.g611450
Cre10.g450500	Cre16.g652150	Cre10.g450500	Cre12.g485150	Cre10.g450500	Cre17.g720250
Cre10.g450500	Cre10.g458450	Cre10.g450500	Cre12.g530950	Cre10.g450500	Cre03.g182050

Cre10.g450500	Cre08.g377300	Cre10.g450500	Cre02.g092350	Cre10.g450500	Cre06.g309000
Cre10.g450500	Cre03.g172000	Cre10.g450500	Cre12.g546150	Cre10.g450500	Cre12.g484200
Cre10.g450500	Cre07.g328200	Cre10.g450500	Cre12.g550850	Cre10.g450500	Cre16.g677500
Cre10.g450500	Cre06.g284200	Cre10.g450500	Cre07.g334150	Cre10.g450500	Cre11.g467350
Cre10.g450500	Cre11.g476550	Cre10.g450500	Cre12.g554250	Cre10.g450500	Cre03.g188650
Cre10.g450500	Cre08.g365950	Cre10.g450500	Cre09.g415700	Cre10.g450500	Cre01.g038600
Cre10.g450500	Cre05.g241850	Cre10.g460400	Cre01.g004600	Cre10.g460400	Cre01.g011150
Cre10.g460400	Cre01.g022950	Cre10.g460400	Cre01.g025050	Cre10.g460400	Cre01.g043550
Cre10.g460400	Cre01.g048501	Cre10.g460400	Cre01.g048800	Cre10.g460400	Cre02.g082550
Cre10.g460400	Cre02.g083750	Cre10.g460400	Cre02.g093950	Cre10.g460400	Cre02.g101150
Cre10.g460400	Cre02.g101500	Cre10.g460400	Cre02.g104700	Cre10.g460400	Cre02.g108450
Cre10.g460400	Cre02.g145550	Cre10.g460400	Cre03.g144747	Cre10.g460400	Cre03.g149350
Cre10.g460400	Cre03.g149400	Cre10.g460400	Cre03.g158550	Cre10.g460400	Cre03.g160600
Cre10.g460400	Cre03.g173350	Cre10.g460400	Cre03.g184600	Cre10.g460400	Cre03.g188200
Cre10.g460400	Cre03.g189650	Cre10.g460400	Cre03.g193400	Cre10.g460400	Cre03.g200431
Cre10.g460400	Cre03.g209953	Cre10.g460400	Cre03.g212641	Cre10.g460400	Cre04.g221200
Cre10.g460400	Cre04.g221950	Cre10.g460400	Cre04.g226400	Cre10.g460400	Cre04.g231124
Cre10.g460400	Cre05.g233551	Cre10.g460400	Cre06.g257700	Cre10.g460400	Cre06.g262800
Cre10.g460400	Cre06.g266850	Cre10.g460400	Cre06.g266950	Cre10.g460400	Cre06.g270850
Cre10.g460400	Cre06.g275350	Cre10.g460400	Cre06.g275550	Cre10.g460400	Cre06.g282251
Cre10.g460400	Cre06.g285600	Cre10.g460400	Cre06.g297904	Cre10.g460400	Cre06.g305200
Cre10.g460400	Cre07.g319701	Cre10.g460400	Cre07.g321550	Cre10.g460400	Cre07.g330050
Cre10.g460400	Cre07.g332250	Cre10.g460400	Cre07.g341800	Cre10.g460400	Cre07.g351500
Cre10.g460400	Cre07.g351550	Cre10.g460400	Cre07.g354100	Cre10.g460400	Cre07.g354500
Cre10.g460400	Cre07.g355700	Cre10.g460400	Cre08.g361400	Cre10.g460400	Cre08.g364050
Cre10.g460400	Cre08.g371052	Cre10.g460400	Cre09.g386753	Cre10.g460400	Cre09.g390023
Cre10.g460400	Cre09.g411600	Cre10.g460400	Cre10.g429880	Cre10.g460400	Cre10.g436800
Cre10.g460400	Cre10.g466550	Cre10.g460400	Cre11.g467577	Cre10.g460400	Cre11.g479800
Cre10.g460400	Cre12.g485950	Cre10.g460400	Cre12.g502251	Cre10.g460400	Cre12.g503800
Cre10.g460400	Cre12.g507500	Cre10.g460400	Cre12.g517350	Cre10.g460400	Cre12.g520850
Cre10.g460400	Cre12.g538550	Cre10.g460400	Cre12.g540400	Cre10.g460400	Cre13.g571050
Cre10.g460400	Cre13.g571200	Cre10.g460400	Cre13.g580450	Cre10.g460400	Cre13.g586450
Cre10.g460400	Cre13.g590350	Cre10.g460400	Cre14.g611000	Cre10.g460400	Cre14.g621172
Cre10.g460400	Cre14.g623800	Cre10.g460400	Cre16.g649433	Cre10.g460400	Cre16.g650400
Cre10.g460400	Cre16.g662650	Cre10.g460400	Cre16.g667900	Cre10.g460400	Cre16.g668200
Cre10.g460400	Cre16.g672400	Cre10.g460400	Cre16.g678900	Cre10.g460400	Cre16.g680232
Cre10.g460400	Cre16.g683953	Cre10.g460400	Cre16.g686250	Cre10.g460400	Cre16.g692500
Cre10.g460400	Cre17.g702650	Cre10.g460400	Cre17.g705950	Cre10.g460400	Cre17.g714229
Cre10.g460400	Cre17.g746547	Cre10.g460400	Cre12.g489700	Cre10.g460400	Cre16.g652150
Cre10.g460400	Cre12.g485150	Cre10.g460400	Cre10.g458450	Cre10.g460400	Cre03.g172000
Cre10.g460400	Cre16.g656400	Cre10.g460400	Cre16.g665250	Cre10.g460400	Cre12.g530600
Cre10.g460400	Cre06.g284200	Cre10.g460400	Cre12.g498750	Cre10.g460400	Cre01.g018500
Cre10.g460400	Cre07.g334150	Cre10.g460400	Cre16.g673250	Cre10.g460400	Cre09.g410750
Cre10.g460400	Cre03.g199800	Cre10.g464750	Cre03.g199800	Cre10.g464750	Cre12.g554250
Cre11.g476250	Cre16.g673250	Cre11.g479800	Cre09.g410750	Cre11.g480950	Cre08.g377300
Cre12.g485950	Cre16.g656400	Cre12.g485950	Cre01.g032650	Cre12.g485950	Cre01.g038600
Cre12.g501600	Cre01.g025050	Cre12.g501600	Cre07.g329900	Cre12.g501600	Cre09.g386753
Cre12.g501600	Cre17.g702650	Cre12.g501600	Cre17.g714229	Cre12.g501600	Cre17.g746547
Cre12.g501600	Cre12.g484200	Cre12.g501600	Cre01.g032650	Cre12.g501600	Cre12.g498750
Cre12.g508150	Cre02.g078831	Cre12.g508150	Cre13.g567250	Cre12.g508150	Cre07.g328200
Cre12.g508150	Cre12.g530600	Cre12.g508150	Cre07.g334150	Cre12.g508150	Cre09.g410750
Cre12.g511400	Cre16.g673250	Cre12.g516050	Cre03.g158550	Cre12.g516050	Cre04.g215450
Cre12.g516050	Cre13.g567250	Cre12.g516050	Cre14.g629700	Cre12.g516050	Cre01.g038900
Cre12.g516050	Cre10.g425100	Cre12.g516050	Cre12.g489700	Cre12.g516050	Cre07.g330250
Cre12.g516050	Cre03.g188650	Cre12.g516050	Cre05.g241850	Cre12.g517350	Cre02.g087551
Cre12.g517350	Cre09.g386753	Cre12.g517350	Cre03.g158900	Cre12.g520650	Cre16.g677500

Cre12.g520650	Cre07.g334150	Cre12.g520650	Cre06.g308500	Cre12.g534450	Cre01.g011150
Cre12.g534450	Cre01.g022950	Cre12.g534450	Cre01.g048501	Cre12.g534450	Cre01.g048800
Cre12.g534450	Cre02.g087600	Cre12.g534450	Cre02.g091550	Cre12.g534450	Cre02.g101150
Cre12.g534450	Cre02.g108450	Cre12.g534450	Cre02.g146629	Cre12.g534450	Cre03.g158550
Cre12.g534450	Cre03.g160600	Cre12.g534450	Cre03.g180800	Cre12.g534450	Cre04.g215450
Cre12.g534450	Cre04.g221950	Cre12.g534450	Cre06.g266850	Cre12.g534450	Cre06.g275350
Cre12.g534450	Cre06.g286700	Cre12.g534450	Cre06.g297904	Cre12.g534450	Cre06.g305200
Cre12.g534450	Cre07.g341800	Cre12.g534450	Cre08.g361400	Cre12.g534450	Cre08.g364050
Cre12.g534450	Cre09.g386753	Cre12.g534450	Cre13.g567250	Cre12.g534450	Cre13.g571050
Cre12.g534450	Cre14.g621050	Cre12.g534450	Cre16.g662650	Cre12.g534450	Cre16.g667900
Cre12.g534450	Cre16.g672400	Cre12.g534450	Cre16.g673250	Cre12.g534450	Cre16.g676421
Cre12.g534450	Cre17.g739450	Cre12.g534450	Cre14.g629700	Cre12.g534450	Cre01.g038900
Cre12.g534450	Cre10.g425100	Cre12.g534450	Cre12.g489700	Cre12.g534450	Cre07.g330250
Cre12.g534450	Cre03.g198950	Cre12.g534450	Cre10.g458450	Cre12.g534450	Cre14.g615050
Cre12.g534450	Cre08.g377300	Cre12.g534450	Cre02.g092350	Cre12.g534450	Cre03.g172000
Cre12.g534450	Cre09.g412100	Cre12.g534450	Cre12.g550850	Cre12.g534450	Cre02.g082750
Cre12.g534450	Cre16.g656400	Cre12.g534450	Cre16.g665250	Cre12.g534450	Cre17.g723650
Cre12.g534450	Cre06.g284200	Cre12.g534450	Cre07.g334150	Cre12.g534450	Cre11.g467350
Cre12.g534450	Cre11.g476550	Cre12.g534450	Cre09.g410750	Cre12.g534450	Cre03.g199800
Cre12.g534450	Cre03.g156900	Cre12.g534450	Cre12.g554250	Cre12.g534450	Cre03.g188650
Cre12.g534450	Cre12.g543400	Cre12.g534450	Cre05.g241850	Cre12.g538801	Cre01.g004600
Cre12.g538801	Cre01.g022950	Cre12.g538801	Cre01.g025050	Cre12.g538801	Cre01.g043550
Cre12.g538801	Cre01.g048800	Cre12.g538801	Cre02.g101150	Cre12.g538801	Cre02.g101500
Cre12.g538801	Cre02.g104700	Cre12.g538801	Cre02.g108450	Cre12.g538801	Cre02.g145550
Cre12.g538801	Cre03.g144747	Cre12.g538801	Cre03.g149350	Cre12.g538801	Cre03.g149400
Cre12.g538801	Cre03.g173350	Cre12.g538801	Cre03.g184600	Cre12.g538801	Cre03.g188200
Cre12.g538801	Cre03.g189650	Cre12.g538801	Cre03.g193400	Cre12.g538801	Cre03.g200431
Cre12.g538801	Cre03.g209953	Cre12.g538801	Cre03.g212641	Cre12.g538801	Cre04.g221200
Cre12.g538801	Cre04.g221950	Cre12.g538801	Cre04.g226400	Cre12.g538801	Cre05.g233551
Cre12.g538801	Cre06.g262800	Cre12.g538801	Cre06.g266850	Cre12.g538801	Cre06.g266950
Cre12.g538801	Cre06.g270850	Cre12.g538801	Cre06.g275350	Cre12.g538801	Cre06.g275550
Cre12.g538801	Cre06.g282251	Cre12.g538801	Cre06.g297904	Cre12.g538801	Cre06.g305200
Cre12.g538801	Cre07.g319701	Cre12.g538801	Cre07.g330050	Cre12.g538801	Cre07.g332250
Cre12.g538801	Cre07.g341800	Cre12.g538801	Cre07.g351550	Cre12.g538801	Cre07.g354100
Cre12.g538801	Cre07.g354500	Cre12.g538801	Cre08.g361400	Cre12.g538801	Cre08.g364050
Cre12.g538801	Cre08.g371052	Cre12.g538801	Cre09.g386753	Cre12.g538801	Cre09.g411600
Cre12.g538801	Cre10.g429880	Cre12.g538801	Cre10.g436800	Cre12.g538801	Cre11.g467577
Cre12.g538801	Cre12.g503800	Cre12.g538801	Cre12.g507500	Cre12.g538801	Cre12.g517350
Cre12.g538801	Cre12.g538550	Cre12.g538801	Cre12.g540400	Cre12.g538801	Cre13.g571050
Cre12.g538801	Cre13.g580450	Cre12.g538801	Cre13.g586450	Cre12.g538801	Cre13.g590350
Cre12.g538801	Cre14.g621172	Cre12.g538801	Cre16.g649433	Cre12.g538801	Cre16.g650400
Cre12.g538801	Cre16.g662650	Cre12.g538801	Cre16.g668200	Cre12.g538801	Cre16.g672400
Cre12.g538801	Cre16.g678900	Cre12.g538801	Cre16.g680232	Cre12.g538801	Cre16.g683953
Cre12.g538801	Cre16.g686250	Cre12.g538801	Cre17.g702650	Cre12.g538801	Cre17.g705950
Cre12.g538801	Cre16.g665250	Cre12.g538801	Cre06.g284200	Cre12.g538801	Cre09.g410750
Cre12.g538801	Cre03.g199800	Cre12.g538801	Cre12.g543400	Cre12.g540350	Cre12.g508550
Cre12.g540400	Cre10.g425100	Cre12.g540400	Cre16.g652150	Cre12.g540400	Cre11.g467350
Cre12.g540400	Cre16.g679500	Cre12.g548250	Cre04.g231124	Cre12.g550250	Cre09.g386753
Cre12.g550250	Cre16.g652150	Cre12.g550250	Cre10.g458450	Cre12.g550250	Cre03.g188650
Cre13.g562400	Cre02.g145550	Cre13.g562400	Cre03.g174150	Cre13.g562400	Cre12.g558950
Cre13.g562400	Cre16.g662650	Cre13.g562400	Cre16.g672400	Cre13.g562400	Cre17.g720250
Cre13.g562400	Cre16.g673250	Cre13.g562400	Cre03.g199800	Cre13.g572150	Cre12.g554250
Cre13.g583150	Cre09.g386753	Cre13.g583150	Cre16.g656400	Cre13.g583150	Cre07.g334150
Cre13.g583150	Cre03.g199800	Cre13.g583200	Cre03.g160600	Cre13.g583200	Cre08.g364050
Cre13.g583200	Cre16.g662650	Cre13.g583200	Cre16.g667900	Cre13.g583200	Cre03.g198950
Cre13.g583200	Cre03.g194200	Cre13.g583200	Cre12.g484200	Cre13.g583200	Cre01.g018500

Cre13.g583200	Cre07.g334150	Cre13.g583200	Cre16.g679500	Cre13.g583200	Cre12.g554250
Cre13.g590350	Cre01.g004600	Cre13.g590350	Cre01.g025050	Cre13.g590350	Cre01.g048501
Cre13.g590350	Cre02.g094150	Cre13.g590350	Cre02.g118250	Cre13.g590350	Cre03.g160600
Cre13.g590350	Cre03.g184600	Cre13.g590350	Cre03.g197350	Cre13.g590350	Cre04.g228400
Cre13.g590350	Cre05.g233551	Cre13.g590350	Cre06.g266850	Cre13.g590350	Cre06.g275600
Cre13.g590350	Cre06.g293750	Cre13.g590350	Cre06.g297904	Cre13.g590350	Cre07.g329900
Cre13.g590350	Cre09.g386753	Cre13.g590350	Cre10.g430950	Cre13.g590350	Cre10.g446850
Cre13.g590350	Cre12.g485950	Cre13.g590350	Cre13.g583150	Cre13.g590350	Cre14.g621050
Cre13.g590350	Cre16.g662650	Cre13.g590350	Cre16.g667900	Cre13.g590350	Cre17.g702650
Cre13.g590350	Cre17.g714229	Cre13.g590350	Cre17.g739450	Cre13.g590350	Cre17.g746547
Cre13.g590350	Cre10.g425100	Cre13.g590350	Cre07.g330250	Cre13.g590350	Cre16.g652150
Cre13.g590350	Cre10.g458450	Cre13.g590350	Cre02.g092350	Cre13.g590350	Cre12.g484200
Cre13.g590350	Cre01.g032650	Cre13.g590350	Cre12.g498750	Cre13.g590350	Cre01.g018500
Cre13.g590350	Cre16.g673250	Cre13.g602750	Cre01.g022950	Cre13.g602750	Cre16.g673250
Cre14.g610550	Cre09.g386753	Cre14.g610550	Cre16.g656400	Cre14.g610550	Cre07.g334150
Cre14.g610550	Cre03.g199800	Cre14.g610550	Cre03.g188650	Cre14.g610582	Cre16.g656400
Cre14.g610582	Cre03.g199800	Cre14.g610582	Cre03.g188650	Cre14.g610631	Cre09.g386753
Cre14.g610631	Cre07.g330250	Cre14.g610631	Cre16.g656400	Cre14.g614400	Cre02.g078831
Cre14.g614400	Cre13.g567250	Cre14.g614400	Cre07.g330250	Cre14.g614400	Cre10.g458450
Cre14.g614400	Cre07.g328200	Cre14.g614400	Cre16.g665250	Cre14.g614400	Cre12.g530600
Cre14.g614400	Cre07.g334150	Cre14.g614400	Cre09.g410750	Cre14.g614400	Cre03.g158900
Cre14.g614400	Cre03.g199800	Cre14.g623800	Cre12.g554250	Cre15.g641200	Cre01.g022950
Cre15.g641200	Cre16.g673250	Cre16.g653258	Cre02.g078831	Cre16.g653258	Cre16.g665250
Cre16.g653258	Cre07.g334150	Cre16.g662650	Cre12.g554250	Cre16.g667150	Cre01.g022950
Cre16.g667150	Cre03.g160600	Cre16.g667150	Cre09.g386753	Cre16.g667150	Cre11.g467581
Cre16.g667150	Cre13.g572150	Cre16.g667150	Cre16.g649433	Cre16.g667150	Cre16.g654550
Cre16.g667150	Cre17.g705950	Cre16.g667150	Cre14.g629700	Cre16.g667150	Cre07.g330250
Cre16.g667150	Cre12.g530950	Cre16.g667150	Cre06.g270200	Cre16.g667150	Cre03.g172000
Cre16.g667150	Cre03.g165100	Cre16.g667150	Cre02.g082750	Cre16.g667150	Cre02.g116450
Cre16.g667150	Cre01.g032650	Cre16.g667150	Cre16.g665250	Cre16.g667150	Cre16.g677500
Cre16.g667150	Cre12.g530600	Cre16.g667150	Cre12.g536050	Cre16.g667150	Cre12.g498750
Cre16.g667150	Cre07.g334150	Cre16.g667150	Cre17.g699100	Cre16.g667150	Cre12.g501050
Cre16.g667150	Cre16.g673250	Cre16.g667150	Cre09.g410750	Cre16.g667150	Cre03.g199800
Cre16.g667150	Cre12.g560950	Cre16.g667150	Cre09.g415700	Cre16.g667150	Cre06.g308500
Cre16.g667150	Cre01.g038600	Cre16.g667150	Cre05.g241850	Cre16.g668200	Cre10.g425100
Cre16.g668200	Cre16.g652150	Cre16.g668200	Cre11.g467350	Cre16.g668200	Cre16.g679500
Cre16.g671900	Cre01.g048800	Cre16.g671900	Cre04.g215450	Cre16.g671900	Cre17.g739450
Cre16.g671900	Cre11.g467350	Cre16.g671900	Cre11.g476550	Cre16.g671900	Cre05.g241850
Cre16.g672300	Cre11.g476750	Cre16.g672300	Cre02.g113200	Cre16.g672300	Cre14.g629700
Cre16.g672300	Cre10.g425100	Cre16.g672300	Cre12.g489700	Cre16.g672300	Cre14.g611700
Cre16.g672300	Cre16.g652150	Cre16.g672300	Cre17.g720250	Cre16.g672300	Cre10.g458450
Cre16.g672300	Cre03.g172000	Cre16.g672300	Cre12.g484200	Cre16.g672300	Cre16.g656400
Cre16.g672300	Cre16.g677500	Cre16.g672300	Cre17.g723650	Cre16.g672300	Cre06.g284200
Cre16.g672300	Cre12.g498750	Cre16.g672300	Cre01.g018500	Cre16.g672300	Cre07.g334150
Cre16.g672300	Cre12.g501050	Cre16.g672300	Cre16.g673250	Cre16.g672300	Cre12.g543400
Cre16.g672300	Cre08.g365950	Cre16.g672300	Cre06.g308500	Cre16.g672300	Cre05.g241850
Cre16.g672400	Cre10.g425100	Cre16.g672400	Cre16.g652150	Cre16.g674890	Cre09.g386753
Cre16.g674890	Cre03.g188650	Cre16.g675246	Cre03.g160600	Cre16.g675246	Cre08.g364050
Cre16.g675246	Cre09.g386753	Cre16.g675246	Cre16.g662650	Cre16.g675246	Cre16.g667900
Cre16.g675246	Cre10.g425100	Cre16.g675246	Cre16.g652150	Cre16.g677382	Cre10.g425100
Cre16.g677382	Cre07.g330250	Cre16.g677382	Cre07.g334150	Cre16.g680588	Cre03.g172000
Cre16.g686250	Cre07.g330250	Cre16.g686250	Cre16.g652150	Cre16.g686250	Cre08.g377300
Cre16.g695600	Cre07.g330250	Cre16.g695600	Cre12.g546150	Cre16.g695600	Cre01.g037850
Cre17.g696300	Cre01.g004600	Cre17.g696300	Cre01.g011150	Cre17.g696300	Cre01.g022950
Cre17.g696300	Cre01.g025050	Cre17.g696300	Cre01.g043550	Cre17.g696300	Cre01.g048501
Cre17.g696300	Cre01.g048800	Cre17.g696300	Cre02.g082550	Cre17.g696300	Cre02.g083750

Cre17.g696300	Cre02.g093950	Cre17.g696300	Cre02.g101150	Cre17.g696300	Cre02.g101500
Cre17.g696300	Cre02.g104700	Cre17.g696300	Cre02.g108450	Cre17.g696300	Cre02.g145550
Cre17.g696300	Cre03.g144747	Cre17.g696300	Cre03.g149350	Cre17.g696300	Cre03.g149400
Cre17.g696300	Cre03.g158550	Cre17.g696300	Cre03.g160600	Cre17.g696300	Cre03.g173350
Cre17.g696300	Cre03.g184600	Cre17.g696300	Cre03.g188200	Cre17.g696300	Cre03.g189650
Cre17.g696300	Cre03.g193400	Cre17.g696300	Cre03.g200431	Cre17.g696300	Cre03.g209953
Cre17.g696300	Cre03.g212641	Cre17.g696300	Cre04.g221200	Cre17.g696300	Cre04.g221950
Cre17.g696300	Cre04.g226400	Cre17.g696300	Cre05.g233551	Cre17.g696300	Cre06.g257700
Cre17.g696300	Cre06.g262800	Cre17.g696300	Cre06.g266850	Cre17.g696300	Cre06.g266950
Cre17.g696300	Cre06.g270850	Cre17.g696300	Cre06.g275350	Cre17.g696300	Cre06.g275550
Cre17.g696300	Cre06.g282251	Cre17.g696300	Cre06.g285600	Cre17.g696300	Cre06.g297904
Cre17.g696300	Cre06.g305200	Cre17.g696300	Cre07.g319701	Cre17.g696300	Cre07.g321550
Cre17.g696300	Cre07.g330050	Cre17.g696300	Cre07.g332250	Cre17.g696300	Cre07.g341800
Cre17.g696300	Cre07.g351500	Cre17.g696300	Cre07.g351550	Cre17.g696300	Cre07.g354100
Cre17.g696300	Cre07.g354500	Cre17.g696300	Cre07.g355700	Cre17.g696300	Cre08.g361400
Cre17.g696300	Cre08.g364050	Cre17.g696300	Cre08.g371052	Cre17.g696300	Cre09.g386753
Cre17.g696300	Cre09.g411600	Cre17.g696300	Cre10.g429880	Cre17.g696300	Cre10.g436800
Cre17.g696300	Cre10.g466550	Cre17.g696300	Cre11.g467577	Cre17.g696300	Cre11.g479800
Cre17.g696300	Cre12.g485950	Cre17.g696300	Cre12.g503800	Cre17.g696300	Cre12.g507500
Cre17.g696300	Cre12.g517350	Cre17.g696300	Cre12.g520850	Cre17.g696300	Cre12.g538550
Cre17.g696300	Cre12.g540400	Cre17.g696300	Cre13.g571050	Cre17.g696300	Cre13.g571200
Cre17.g696300	Cre13.g580450	Cre17.g696300	Cre13.g586450	Cre17.g696300	Cre13.g590350
Cre17.g696300	Cre14.g611000	Cre17.g696300	Cre14.g621172	Cre17.g696300	Cre14.g623800
Cre17.g696300	Cre16.g649433	Cre17.g696300	Cre16.g650400	Cre17.g696300	Cre16.g662650
Cre17.g696300	Cre16.g667900	Cre17.g696300	Cre16.g668200	Cre17.g696300	Cre16.g672400
Cre17.g696300	Cre16.g678900	Cre17.g696300	Cre16.g680232	Cre17.g696300	Cre16.g683953
Cre17.g696300	Cre16.g686250	Cre17.g696300	Cre16.g692500	Cre17.g696300	Cre17.g702650
Cre17.g696300	Cre17.g705950	Cre17.g696300	Cre17.g714229	Cre17.g696300	Cre17.g746547
Cre17.g696300	Cre12.g489700	Cre17.g696300	Cre16.g652150	Cre17.g696300	Cre12.g485150
Cre17.g696300	Cre10.g458450	Cre17.g696300	Cre03.g172000	Cre17.g696300	Cre16.g656400
Cre17.g696300	Cre16.g665250	Cre17.g696300	Cre12.g530600	Cre17.g696300	Cre06.g284200
Cre17.g696300	Cre12.g498750	Cre17.g696300	Cre01.g018500	Cre17.g696300	Cre07.g334150
Cre17.g696300	Cre11.g467350	Cre17.g696300	Cre16.g673250	Cre17.g696300	Cre09.g410750
Cre17.g696300	Cre03.g158900	Cre17.g696300	Cre03.g199800	Cre17.g696300	Cre03.g156900
Cre17.g700600	Cre16.g656400	Cre17.g705950	Cre01.g018500	Cre17.g729750	Cre03.g160600
Cre17.g729750	Cre08.g364050	Cre17.g729750	Cre16.g662650	Cre17.g729750	Cre16.g667900
Cre17.g729750	Cre03.g198950	Cre17.g729750	Cre06.g270200	Cre17.g729750	Cre03.g194200
Cre17.g729750	Cre12.g546150	Cre17.g729750	Cre12.g484200	Cre17.g729750	Cre01.g018500
Cre17.g729750	Cre07.g334150	Cre17.g729750	Cre16.g679500	Cre17.g729750	Cre03.g199800
Cre17.g729750	Cre12.g554250	Cre17.g729750	Cre08.g365950	Cre17.g736850	Cre01.g022950
Cre17.g738600	Cre02.g113200	Cre17.g738600	Cre07.g334150	Cre17.g738632	Cre12.g558950
Cre17.g739450	Cre09.g410750				

Table A.9 shows the numbers of the PDI overlap between Y1H and PampDAP-seq. Table A.10 contains all detected PDIs from two biological replicates for six TFs. Table A.2 contains all forward and reverse primers used for the amplification of ORFs and promoters of the Y1H screen.

Table A.9: Overlap of found PDIs in Y1H and PampDAP-seq

Gene ID	TF name	Overlap	Y1H interactions	Pamp-DAP-seq interactions
Cre04.g231124	C3H_5	7	86	16
Cre12.g534450	CGL86	26	55	111
Cre06.g305200	C2H2_6	2	18	15
Cre02.g145602	CSB17	5	12	110
Cre12.g516050	FHA_8	2	10	6
Cre03.g197100	LRL1	7	26	62

Table A.10: Detected interactions in PampDAP-seq

C2H2_6	Cre06.g305200	Cre03.g148250, Cre03.g174500, Cre03.g177700, Cre07.g318050, Cre08.g377300, Cre10.g430950, Cre12.g485950, Cre13.g583150, Cre01.g025050, Cre03.g199800, Cre04.g218050, Cre12.g502251, Cre14.g623800, Cre16.g671900, Cre17.g714229
CGL86	Cre12.g534450	Cre01.g018500, Cre01.g038900, Cre01.g048501, Cre02.g082550, Cre02.g083750, Cre02.g087551, Cre02.g091550, Cre02.g092350, Cre02.g101850, Cre02.g108450, Cre02.g109650, Cre02.g112750, Cre02.g113200, Cre02.g116450, Cre02.g118250, Cre02.g146629, Cre03.g149400, Cre03.g152150, Cre03.g153050, Cre03.g160700, Cre03.g168050, Cre03.g188650, Cre03.g193400, Cre03.g197350, Cre03.g198950, Cre04.g215450, Cre04.g221950, Cre04.g224600, Cre04.g226400, Cre05.g242650, Cre06.g257700, Cre06.g266200, Cre06.g266850, Cre06.g270200, Cre06.g278108, Cre06.g281200, Cre06.g282800, Cre06.g284000, Cre06.g297904, Cre06.g305200, Cre06.g305850, Cre06.g307150, Cre06.g308500, Cre06.g309000, Cre07.g318050, Cre07.g329900, Cre07.g332250, Cre07.g334150, Cre07.g345050, Cre07.g349152, Cre07.g351500, Cre07.g354500, Cre07.g355700, Cre08.g358534, Cre08.g361400, Cre08.g362900, Cre08.g364450, Cre08.g371052, Cre08.g373346, Cre08.g377300, Cre09.g389750, Cre09.g397475, Cre09.g399552, Cre09.g412100, Cre09.g415700, Cre10.g417650, Cre10.g421550, Cre10.g425100, Cre10.g430950, Cre10.g446450, Cre11.g467350, Cre11.g467577, Cre11.g475100, Cre11.g476550, Cre11.g479800, Cre11.g483250, Cre12.g484000, Cre12.g484200, Cre12.g485950, Cre12.g498750, Cre12.g502251, Cre12.g507500, Cre12.g508550, Cre12.g523000, Cre12.g530600, Cre12.g530950, Cre12.g536050, Cre12.g538550, Cre12.g546150, Cre13.g567250, Cre13.g570150, Cre13.g571050, Cre13.g573000, Cre13.g583150, Cre13.g583972, Cre13.g586450, Cre14.g610550, Cre14.g617151, Cre14.g621050, Cre14.g623800, Cre14.g631750, Cre15.g635100, Cre16.g653250, Cre16.g656400, Cre16.g662650, Cre16.g667450, Cre16.g668200, Cre16.g671900, Cre16.g672400, Cre16.g675100, Cre16.g677500, Cre16.g678900, Cre16.g680232, Cre16.g683050, Cre16.g683953, Cre16.g686100, Cre17.g702451, Cre17.g709800, Cre17.g714229, Cre17.g715250, Cre17.g720250, Cre17.g729750

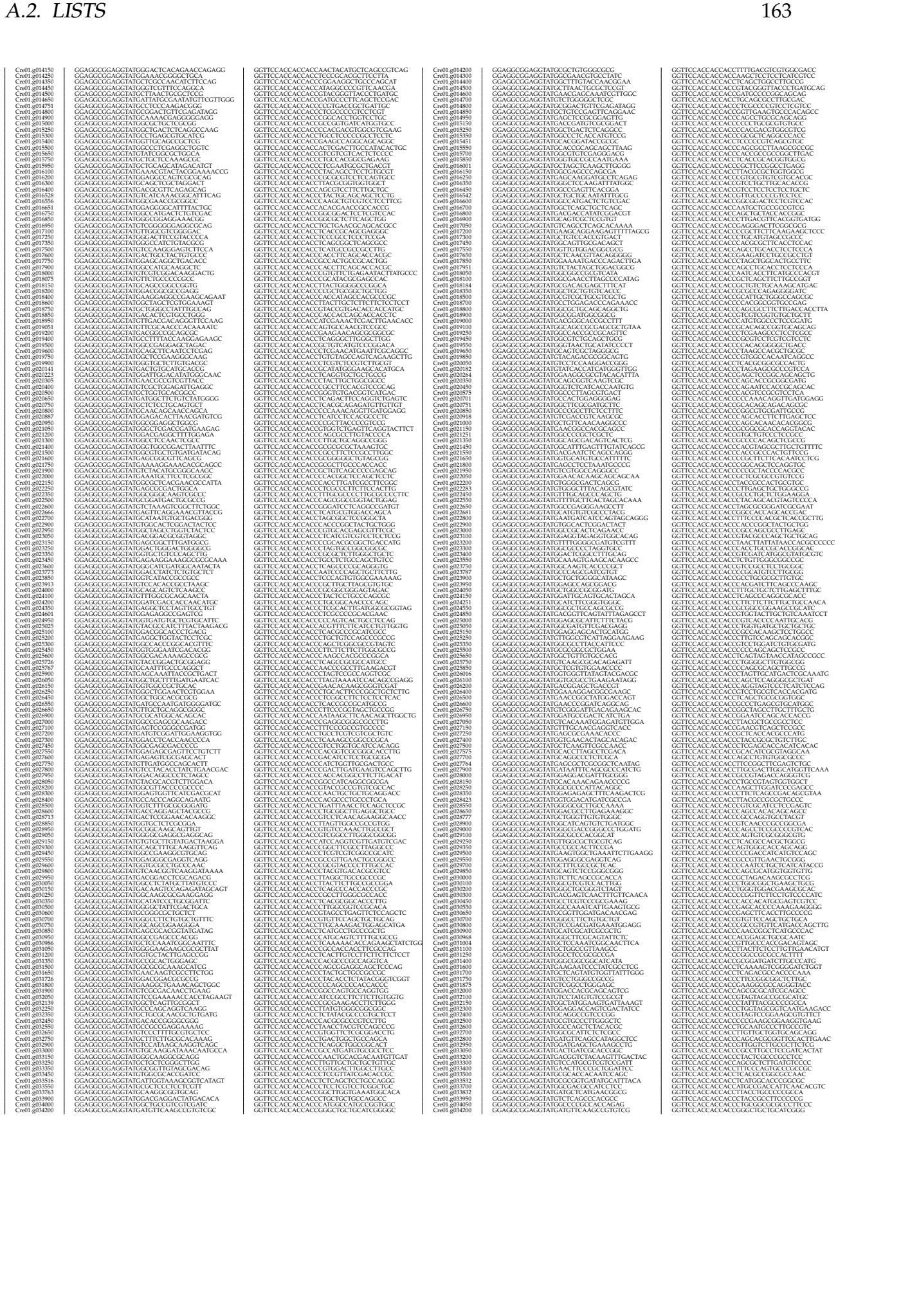
LRL1	Cre03.g197100	Cre01.g004600, Cre01.g037850, Cre01.g038050, Cre01.g043550, Cre02.g082750, Cre02.g087600, Cre02.g091550, Cre02.g101150, Cre02.g109650, Cre02.g113200, Cre03.g144424, Cre03.g153050, Cre03.g172000, Cre03.g177700, Cre03.g182050, Cre03.g188650, Cre03.g194200, Cre03.g197350, Cre03.g199800, Cre03.g207700, Cre03.g210625, Cre04.g213650, Cre04.g216950, Cre05.g241636, Cre05.g241850, Cre05.g242600, Cre05.g242650, Cre06.g257700, Cre06.g275600, Cre06.g278108, Cre06.g281200, Cre06.g284000, Cre06.g285600, Cre06.g293750, Cre06.g297150, Cre06.g307150, Cre07.g355700, Cre08.g361400, Cre08.g362900, Cre08.g364450, Cre08.g365950, Cre09.g393654, Cre09.g397475, Cre09.g412100, Cre10.g417650, Cre10.g418600, Cre10.g430950, Cre10.g446850, Cre10.g463150, Cre11.g475950, Cre11.g476550, Cre11.g476750, Cre12.g502251, Cre12.g506600, Cre12.g508550, Cre12.g523000, Cre12.g530950, Cre12.g534100, Cre12.g557300, Cre12.g560950, Cre13.g583150, Cre13.g583972, Cre14.g611000, Cre14.g629700, Cre14.g631750, Cre16.g656400, Cre16.g667450, Cre16.g673250, Cre16.g679151, Cre16.g683953, Cre17.g702451, Cre17.g739450
FHA_8	Cre12.g516050	Cre01.g038900, Cre06.g308500, Cre07.g323000, Cre10.g430950, Cre03.g188650, Cre13.g583972
CSB17	Cre02.g145602	Cre01.g018500, Cre01.g037850, Cre01.g043550, Cre01.g048501, Cre01.g062902, Cre02.g082550, Cre02.g091550, Cre02.g101150, Cre02.g101500, Cre02.g108450, Cre02.g109650, Cre02.g112750, Cre02.g116450, Cre02.g118250, Cre03.g145407, Cre03.g148250, Cre03.g149400, Cre03.g152450, Cre03.g158550, Cre03.g160700, Cre03.g168050, Cre03.g188650, Cre03.g193400, Cre03.g198950, Cre04.g221200, Cre04.g226400, Cre04.g229948, Cre05.g241850, Cre06.g257700, Cre06.g266850, Cre06.g270200, Cre06.g270350, Cre06.g275600, Cre06.g278108, Cre06.g284000, Cre06.g293750, Cre06.g305200, Cre06.g308500, Cre07.g316600, Cre07.g318050, Cre07.g334150, Cre07.g340450, Cre07.g349152, Cre07.g354500, Cre08.g362900, Cre08.g363500, Cre08.g364450, Cre08.g375400, Cre09.g389750, Cre09.g399552, Cre09.g410450, Cre10.g421550, Cre10.g429880, Cre10.g430950, Cre10.g446450, Cre10.g458450, Cre10.g463150, Cre11.g467577, Cre11.g475100, Cre12.g484200, Cre12.g485950, Cre12.g489700, Cre12.g498750, Cre12.g502251, Cre12.g503800, Cre12.g507500, Cre12.g508550, Cre12.g530950, Cre12.g534100, Cre12.g536050, Cre12.g538550, Cre12.g538801, Cre12.g554250, Cre12.g557300, Cre13.g567250, Cre13.g570150, Cre13.g571050, Cre13.g572150, Cre13.g573000, Cre13.g583972, Cre13.g586450, Cre14.g610550, Cre14.g611700, Cre14.g617151, Cre16.g650400, Cre16.g652150, Cre16.g654550, Cre16.g656400, Cre16.g662650, Cre16.g667450, Cre16.g671900, Cre16.g672400, Cre16.g675100, Cre16.g678900, Cre16.g680232, Cre16.g683953, Cre17.g702451, Cre17.g714229, Cre17.g739450, Cre02.g120100, Cre02.g146629, Cre03.g153050, Cre03.g174200, Cre03.g197100, Cre03.g201250, Cre03.g212977, Cre04.g216200, Cre04.g221950, Cre05.g240650, Cre05.g242650, Cre06.g266200, Cre06.g284200, Cre07.g332250, Cre07.g335150, Cre07.g353700, Cre08.g377300, Cre09.g396600, Cre09.g397475, Cre10.g425100, Cre11.g479800, Cre12.g506600, Cre12.g516050, Cre13.g583200, Cre14.g631750, Cre16.g683050, Cre17.g705950

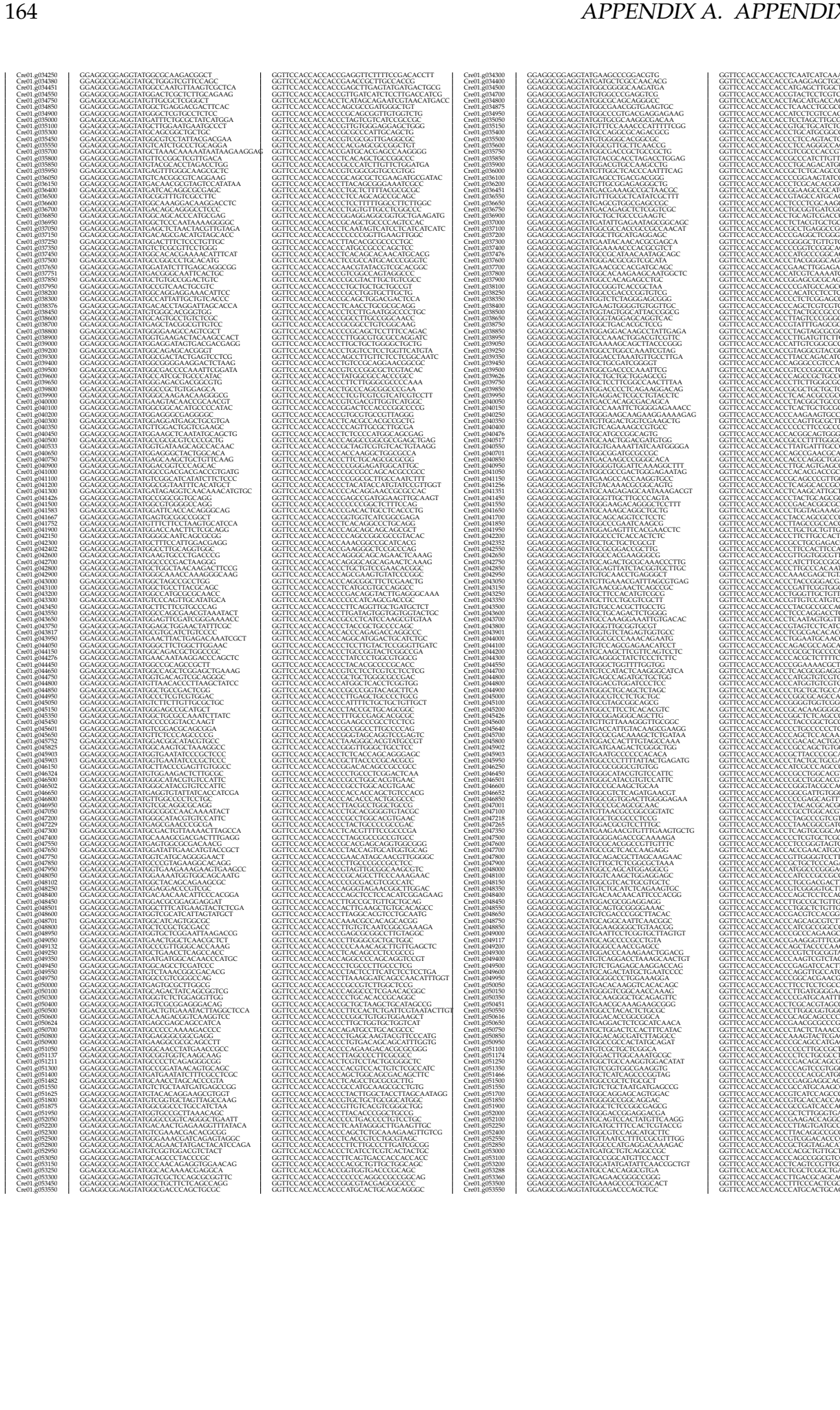
C3H_5	Cre04.g231124	Cre02.g108450, Cre03.g160700, Cre03.g168050, Cre04.g226400, Cre04.g231124, Cre10.g430950, Cre11.g467577, Cre12.g498750, Cre12.g538801, Cre16.g656400, Cre16.g667450, Cre16.g680232, Cre02.g082750, Cre03.g148250, Cre03.g197700, Cre08.g363500, Cre16.g677500

Table A.11: List of primers used for ORF and promoter amplification

Cre ID	Forward primer	Reverse primer
Cre01.g000017	GGAGGCGGAGGTATGATACTCATTCAAGGTGGTTAT	GGTTCCACCACCACCTTAATAGGAAGGCTGGGACG
Cre01.g000150	GGAGGCGGAGGTATGGCGGCGAACGCGTTC	GGTTCCACCACCACCCGCCCACTGAGCCAGCAC
Cre01.g000250	GGAGGCGGAGGTATGCTGCCGGCAAGCG	GGTTCCACCACCACCTTCCTGCTGTTGCTTGCC
Cre01.g000350	GGAGGCGGAGGTATGATGGGCCCCACACCC	GGTTCCACCACCACCGAAGTTGAGCAGGATCTTCTCG
Cre01.g000450	GGAGGCGGAGGTATGGGCGCCCAGGGAATAA	GGTTCCACCACCACCGAATATCTGTCCGGGCATGCG
Cre01.g000550	GGAGGCGGAGGTATGTCGGGCACGTCCAGT	GGTTCCACCACCACCCTACCCGACGCCCGCC
Cre01.g000650	GGAGGCGGAGGTATGGCGACAGCTTTCAATATGAC	GGTTCCACCACCACCCAGCTTGGACTTGGGCAGC
Cre01.g000750	GGAGGCGGAGGTATGAAGCTGGACGGCGG	GGTTCCACCACCACCCTTGACCACCGCCACTG
Cre01.g000850	GGAGGCGGAGGTATGGCGACCGCAATGTT	GGTTCCACCACCACCCTGCTGCTGCTGCTGTG
Cre01.g000950	GGAGGCGGAGGTATGCCGGCCACCACGAAAG	GGTTCCACCACCACCGGCATCGCTATCGCCGTATC
Cre01.g001100	GGAGGCGGAGGTATGTCCAATGCCAACGCGTTC	GGTTCCACCACCACCGGCCTTGCGGCGGTTGTTCTTC
Cre01.g001400	GGAGGCGGAGGTATGCATAGACGTAGAGGTGCC	GGTTCCACCACCACCCCGGTACGACGCATATTGGC
Cre01.g001501	GGAGGCGGAGGTATGGACATGTGCCCATCATCA	GGTTCCACCACCACCGTACCGCTCCCCCGCC
Cre01.g001600	GGAGGCGGAGGTATGCTGCGGGCTAAGGAGC	GGTTCCACCACCACCCTAGCACCGCGGCCGC
Cre01.g001657	GGAGGCGGAGGTATGGACGAACATAACGTGCA	GGTTCCACCACCACCCAGCTGCTCCCAGCACC
Cre01.g001671	GGAGGCGGAGGTATGAGTAGACAGGGACGTGAC	GGTTCCACCACCACCCTAGGCACCAGCCGGC
Cre01.g001685	GGAGGCGGAGGTATGTACGCCACCCCGC	GGTTCCACCACCACCACAGATGGACATATCGCTGC
Cre01.g001800	GGAGGCGGAGGTATGGACAAGTACATTCGCGGC	GGTTCCACCACCACCCGCGCCCTCCATGTGC
Cre01.g001983	GGAGGCGGAGGTATGAAGGGCGGCCACCCG	GGTTCCACCACCACCGAACTTGAAGCCAAAGCCCTTG
Cre01.g002100	GGAGGCGGAGGTATGCGTAGAGGTAGAGCTGG	GGTTCCACCACCACCCCTGCCGGCCTTGCTG
Cre01.g002200	GGAGGCGGAGGTATGGCTGACGACTATGATGAGG	GGTTCCACCACCACCGTCCGGCACCACCAGCTC
Cre01.g002203	GGAGGCGGAGGTATGTCCCCGCGGCGAC	GGTTCCACCACCACCCAAGCGCCGCAACTGC
Cre01.g002234	GGAGGCGGAGGTATGCTGCTTGCAAATTCCAAG	GGTTCCACCACCACCCTACTGCCCGGCACCC
Cre01.g002300	GGAGGCGGAGGTATGCCGAACCCCATTGTCTTC	GGTTCCACCACCACCGGCCAGCTGGCCGCAGTC
Cre01.g002400	GGAGGCGGAGGTATGGTTCACAAGGGCTTCC	GGTTCCACCACCACCCTTCTTCACCGCGCCC
Cre01.g002500	GGAGGCGGAGGTATGGACTTCAAGAGCATCAGC	GGTTCCACCACCACCCACGTCCAGGTGCAGGC
Cre01.g002600	GGAGGCGGAGGTATGTCAAGTACCGCGGTTG	GGTTCCACCACCACCCTCCATGAGCCCCATCGC
Cre01.g002700	GGAGGCGGAGGTATGGACTCGCTGCCATACG	GGTTCCACCACCACCCCGGTCCCGCTCCCTG
Cre01.g002824	GGAGGCGGAGGTATGCTCAACCTGGAGTCGGTC	GGTTCCACCACCACCTGCGGCCGCGGACAAAGAG
Cre01.g002950	GGAGGCGGAGGTATGCTCGCGGATAATTGCCCG	GGTTCCACCACCACCCTGCGCCTGCAGGCGG
Cre01.g003050	GGAGGCGGAGGTATGGCCGTCCAGCCTCCG	GGTTCCACCACCACCCTCCCCGCCTGCCCCTTGTTG
Cre01.g003100	GGAGGCGGAGGTATGGCTCGCCAAGACAGC	GGTTCCACCACCACCTACCTCGGACGACCTCGC
Cre01.g003200	GGAGGCGGAGGTATGCCGACGACGCTCCCC	GGTTCCACCACCACCCAGCGTGGGCCCAGCCGC
Cre01.g003376	GGAGGCGGAGGTATGGGTTTTCAGCTATCGCTC	GGTTCCACCACCACCCTCGTCGCCGTCGTCG
Cre01.g003463	GGAGGCGGAGGTATGGGCGGCAAGAGCAC	GGTTCCACCACCACCTCAGCACGCCGCCACC
Cre01.g003463	GGAGGCGGAGGTATGCCACCCAAGGCGGC	GGTTCCACCACCACCCCCCGGCGCGATGCAG
Cre01.g003475	GGAGGCGGAGGTATGTGGAGGGCGGCATTT	GGTTCCACCACCACCTCACACCAGGCGGCGG
Cre01.g003508	GGAGGCGGAGGTATGGCGAGCAAGCTGAAG	GGTTCCACCACCACCCACTCCTCCAGCACCTCC
Cre01.g003524	GGAGGCGGAGGTATGGGCTGCAGCAGCAGTAAG	GGTTCCACCACCACCCTCGTAGCCGTCGCCGTC
Cre01.g003550	GGAGGCGGAGGTATGTCCGCCCTGAACCT	GGTTCCACCACCACCCTTGCGGCCACCAAAGG
Cre01.g003650	GGAGGCGGAGGTATGAGCGAACTAGAATTTCGAGA	GGTTCCACCACCACCCTATTCCTCCCAGGGCCG
Cre01.g003850	GGAGGCGGAGGTATGAAGAGCGCACTTAGCGCAT	GGTTCCACCACCACCCTGCTGCTGCCCGCGCTG
Cre01.g004000	GGAGGCGGAGGTATGTATGAGCGGCTCTTTAACA	GGTTCCACCACCACCCTACTGCCCCTTCTGCGT
Cre01.g004200	GGAGGCGGAGGTATGCGGCGAGCATCCG	GGTTCCACCACCACCCCCATGCCTGACCGATTTAT
Cre01.g004300	GGAGGCGGAGGTATGTGCGGCATTCTTGCC	GGTTCCACCACCACCCTTCTTGGCCTTCTTGGCAG
Cre01.g004400	GGAGGCGGAGGTATGGATTCGCCGCTTGAAGTT	GGTTCCACCACCACCTTACGCGCGCGGGGTG
Cre01.g004500	GGAGGCGGAGGTATGTCTCTGCAGGCCACCAT	GGTTCCACCACCACCCATGTACTCGCGGGGGTC
Cre01.g004651	GGAGGCGGAGGTATGTCAGATTGGGCCCTTCG	GGTTCCACCACCACCCTACTTCATTATTTCGCGCCTC
Cre01.g004750	GGAGGCGGAGGTATGGAAGCAGGCGGAGG	GGTTCCACCACCACCGAGGTTCCGTGGCGGG
Cre01.g004900	GGAGGCGGAGGTATGGCGAACAAGATTTGCATTAT	GGTTCCACCACCACCTCAGCTGCTAAGTGCCTCC
Cre01.g005050	GGAGGCGGAGGTATGAGGGCAACTATTCTGGC	GGTTCCACCACCACCTGCGTCGCCCGACTTG
Cre01.g005150	GGAGGCGGAGGTATGGCCCTTCGCGAGGTG	GGTTCCACCACCACCCTTCAGGAGGCTCTCGCG
Cre01.g005250	GGAGGCGGAGGTATGCAGCGGCTGCTGC	GGTTCCACCACCACCCTAGCCGTAGGCGCCG
Cre01.g005350	GGAGGCGGAGGTATGCGACAAAAGAGCCGC	GGTTCCACCACCACCCTAGCCGTAGGCGCCG
Cre01.g005450	GGAGGCGGAGGTATGGCTGACGACGAGCTTC	GGTTCCACCACCACCTGCCACAGAGAACTCCTTGC
Cre01.g005534	GGAGGCGGAGGTATGCCCAGCGCATGTAGC	GGTTCCACCACCACCGCTATCGTGCACTTTGGTGG
Cre01.g005550	GGAGGCGGAGGTATGCGCGTTCTCATGGTCG	GGTTCCACCACCACCCGAGAACAGGTAAATGCGGCT
Cre01.g005651	GGAGGCGGAGGTATGGTGTATGCGTCACGC	GGTTCCACCACCACCCCGTTTCCCGAACTCCAAT
Cre01.g005813	GGAGGCGGAGGTATGGCTGCCATGGAATCCG	GGTTCCACCACCACCCTCGTCGTCACTGTCATACACT
Cre01.g005900	GGAGGCGGAGGTATGGAGCTCACGCCGCAG	GGTTCCACCACCACCGAAAATGAAGGGCGGGACCTC
Cre01.g006000	GGAGGCGGAGGTATGGAAGCGCTTGCCCA	GGTTCCACCACCACCGCTCGAGCCTTCTGACCA
Cre01.g006150	GGAGGCGGAGGTATGCTGCCCAGTCGCGCT	GGTTCCACCACCACCCACGTCGGGCGGTAGGGCGTAG
Cre01.g006250	GGAGGCGGAGGTATGTATCTTGTGCGTCGTTGT	GGTTCCACCACCACCTTGCTCCTCAGCAGTTTCAG
Cre01.g006402	GGAGGCGGAGGTATGACCGTATTGTCCGGAAGA	GGTTCCACCACCACCAGCAACCTCTGCCGTGG
Cre01.g006500	GGAGGCGGAGGTATGAGCAGGTGCCCATCT	GGTTCCACCACCACCCTGGTGTGCATGGTGTCC
Cre01.g006576	GGAGGCGGAGGTATGCCCTACTTGGCAGCG	GGTTCCACCACCACCTTATGAGCCCCCGCCG
Cre01.g006670	GGAGGCGGAGGTATGTTGCGCCAGATTCGCG	GGTTCCACCACCACCCGCGCGAAGCACGAGC
Cre01.g006733	GGAGGCGGAGGTATGTCCGAGTCCTACTACTATGC	GGTTCCACCACCACCCTAGCCGGTGTCTTGCCC
Cre01.g006800	GGAGGCGGAGGTATGGCTACAGCGGCGC	GGTTCCACCACCACCCACTCGCTTGGCGGTG
Cre01.g006876	GGAGGCGGAGGTATGAGGCTAGTATCGGAACCT	GGTTCCACCACCACCGTGGCTGTCGCGTGTG
Cre01.g006950	GGAGGCGGAGGTATGCAGGCCAAGATGCAG	GGTTCCACCACCACCTTACTTGCCGGCGCCG
Cre01.g007050	GGAGGCGGAGGTATGCAGCGAGCTTGCG	GGTTCCACCACCACCGATCCACTCGGCGCTGC
Cre01.g007150	GGAGGCGGAGGTATGGCTGGTACTTCGGGC	GGTTCCACCACCACCCATTGACATGGACTCGGTTCCA
Cre01.g007251	GGAGGCGGAGGTATGAGCTTGTGGAGCTGC	GGTTCCACCACCACCGCGTCTGACACCGGCA
Cre01.g007350	GGAGGCGGAGGTATGACGGCAGGAACAGAGCAAT	GGTTCCACCACCACCCGCAAAATACAGGCCCTGAAG
Cre01.g007450	GGAGGCGGAGGTATGGAGAATGGCACGTGGC	GGTTCCACCACCACCCCAGTGCAATGGGTGTGC
Cre01.g007600	GGAGGCGGAGGTATGGTCGGCTTTGAGACCC	GGTTCCACCACCACCCTCCTTGAGGGCTGCAGC
Cre01.g007700	GGAGGCGGAGGTATGGCCCAGCTGCTTCGC	GGTTCCACCACCACCCTTGGCCTTGGCCGCCTGTTG
Cre01.g007774	GGAGGCGGAGGTATGTCCGTCGTCGCGG	GGTTCCACCACCACCCTGGGAGAGGATGCTCACG
Cre01.g007850	GGAGGCGGAGGTATGTCGCTGAACAGCCTG	GGTTCCACCACCACCGTACATCGGCGTCCAGTAGA
Cre01.g007950	GGAGGCGGAGGTATGGACAAGGACTTCCCCTTCT	GGTTCCACCACCACCCCAGGTGACGGGCAGGGC
Cre01.g008051	GGAGGCGGAGGTATGCGGCTGTCAACGG	GGTTCCACCACCACCCGTGAACACGGACTCGG
Cre01.g008250	GGAGGCGGAGGTATGCCAGCGCTTAACCCG	GGTTCCACCACCACCCGCAGGCAGCAGGCCT
Cre01.g008350	GGAGGCGGAGGTATGGACCGCCATACAGCCG	GGTTCCACCACCACCCCGTGCGGCCTCCGAG
Cre01.g008500	GGAGGCGGAGGTATGCCTGCAGCTGATTTTGT	GGTTCCACCACCACCCATGACGCGCACCCCA
Cre01.g008600	GGAGGCGGAGGTATGCGGGTTCAGTCCCG	GGTTCCACCACCACCCTTCTGCAGCAGCGGG
Cre01.g008750	GGAGGCGGAGGTATGTCGGCAGCAGCTGT	GGTTCCACCACCACCCGCCACCCAGCGGTGTA
Cre01.g008850	GGAGGCGGAGGTATGGACCAGACGACGTCG	GGTTCCACCACCACCCAGCTCCACCTCCCGC
Cre01.g008900	GGAGGCGGAGGTATGAGGCGGAGCTACGCC	GGTTCCACCACCACCTCACCGACGCGCACGC
Cre01.g008976	GGAGGCGGAGGTATGCACTCCCACTCCCG	GGTTCCACCACCACCTTACGGCCGCGGGTGT
Cre01.g009025	GGAGGCGGAGGTATGCCAGAAGGTAGTGGCATTAC	GGTTCCACCACCACCTCACCGCGCGCCATCG
Cre01.g009101	GGAGGCGGAGGTATGAACACCCTTCAGCAGC	GGTTCCACCACCACCCGCCATGCTCTCAGCAC
Cre01.g009200	GGAGGCGGAGGTATGGGAGGGCGCGAGGAG	GGTTCCACCACCACCAAACGGTGCTGTCCCGAC
Cre01.g009300	GGAGGCGGAGGTATGGCGGTTCCAGGGAAGATTT	GGTTCCACCACCACCCAACAGCTCCGAGTCCAGC
Cre01.g009350	GGAGGCGGAGGTATGCTGAAGACTGCCAAGGC	GGTTCCACCACCACCCTAGCCGCGCGGAGCG
Cre01.g009450	GGAGGCGGAGGTATGTTGTCCCAAGCCCACTC	GGTTCCACCACCACCTCACTCAGGCGGCGGC
Cre01.g009550	GGAGGCGGAGGTATGTTGTTGTCTGCTGGAGCG	GGTTCCACCACCACCATCGCCGCAGTCACCACC
Cre01.g009650	GGAGGCGGAGGTATGGCGTCACGCACGCAT	GGTTCCACCACCACCCACGCCTCCAGCCGCCAC
Cre01.g009731	GGAGGCGGAGGTATGGCGCTACCTCGCG	GGTTCCACCACCACCCTCCCGCTCACGCTTCA
Cre01.g009800	GGAGGCGGAGGTATGGCGCTCACCGTGGAA	GGTTCCACCACCACCCGCATCCGAATCCTCAGCGAAC
Cre01.g009900	GGAGGCGGAGGTATGGCACAGCGGCGAG	GGTTCCACCACCACCTCAGCGCAGGGCGATG
Cre01.g010000	GGAGGCGGAGGTATGGCGGACCCTGGGGCC	GGTTCCACCACCACCGGTGTGCAGCTCGGCCAT
Cre01.g010150	GGAGGCGGAGGTATGGCCAAGGCGCTTCGC	GGTTCCACCACCACCCAGCAGTCGGCGCCGCCG
Cre01.g010300	GGAGGCGGAGGTATGGCTGCGGTCTTCCTCTA	GGTTCCACCACCACCGTAGCCGTACGACGACGG
Cre01.g010400	GGAGGCGGAGGTATGAACTCCATGCTCGTTGC	GGTTCCACCACCACCGCGCAGCAGGAACAGC
Cre01.g010500	GGAGGCGGAGGTATGTCATCGATCTCCGATAACC	GGTTCCACCACCACCCTAGTGTTTCCCGCAGTGC
Cre01.g010650	GGAGGCGGAGGTATGGCCGCAAATGATGTTCT	GGTTCCACCACCACCCTTGGAAGCCGTGCCC
Cre01.g010750	GGAGGCGGAGGTATGCAGCAGCCCGCGT	GGTTCCACCACCACCAAGGCCGTACGCGCCG
Cre01.g010816	GGAGGCGGAGGTATGGAGGTTGACAACTTGCC	GGTTCCACCACCACCCACGCGCACCAGCTCC
Cre01.g010848	GGAGGCGGAGGTATGGCGGACAAGCCCA	GGTTCCACCACCACCCTCCTTCTTTGGCGAAGCG
Cre01.g011000	GGAGGCGGAGGTATGGCCAAGCCCGCCA	GGTTCCACCACCACCGAACTTCATCAGGTGCGGGC
Cre01.g011100	GGAGGCGGAGGTATGTGGAAATTCATTTCTGCCGG	GGTTCCACCACCACCGCTGTTGATCTTGATCTCCTCCT
Cre01.g011300	GGAGGCGGAGGTATGCGGAGCCTCGGCCTTTC	GGTTCCACCACCACCAGAAAACTTTGCGCCCTCCAG
Cre01.g011376	GGAGGCGGAGGTATGCTTTTACTGCAGCGGG	GGTTCCACCACCACCGGTGCACGTACTTGTCCAC
Cre01.g011450	GGAGGCGGAGGTATGGGCCCAAGCTTCTGG	GGTTCCACCACCACCATCCGCCTTCTGCTTCATG
Cre01.g011550	GGAGGCGGAGGTATGCCGCCGAAATGGTTCG	GGTTCCACCACCACCTGACCCCTCGGCCTCCGC
Cre01.g011630	GGAGGCGGAGGTATGCGCACTGGGCGTG	GGTTCCACCACCACCGACGCGAGGCTTCACATT
Cre01.g011750	GGAGGCGGAGGTATGGACCTAGAAGGCCGG	GGTTCCACCACCACCCAGCTCGATGTTGCCCTG
Cre01.g011850	GGAGGCGGAGGTATGGCGGTATTGCTGCTG	GGTTCCACCACCACCCAGCTGCTGCATCATGATCG
Cre01.g012050	GGAGGCGGAGGTATGTCCACAAGCATCGTCGT	GGTTCCACCACCACCTACCATGGACTTGCCGTTCG
Cre01.g012126	GGAGGCGGAGGTATGCCAGCGCTCAACCGG	GGTTCCACCACCACCCCCCGATGTGCCGGTCCC
Cre01.g012175	GGAGGCGGAGGTATGTCACAGGCTGGGATGC	GGTTCCACCACCACCCCCGTAACAGTGGGGCT
Cre01.g012244	GGAGGCGGAGGTATGGACGCTACATCACGAGA	GGTTCCACCACCACCCTTGGCAACCTGGCCG
Cre01.g012350	GGAGGCGGAGGTATGGCTGCCGACAGCGGC	GGTTCCACCACCACCCCGGCGGGCCTTCTTGGC
Cre01.g012450	GGAGGCGGAGGTATGCGTGTGAAGCTCAAGAA	GGTTCCACCACCACCCCCGAACCCGAATCCCA
Cre01.g012550	GGAGGCGGAGGTATGACTCCCGGGGCCTA	GGTTCCACCACCACCCAAGTCCTCCTCCTCCAGT
Cre01.g012800	GGAGGCGGAGGTATGAGCAAGTCTGTCCATGTCA	GGTTCCACCACCACCCAACATGATGTCCGCTGCTTG
Cre01.g012900	GGAGGCGGAGGTATGAACGCCATTCTGCTGCG	GGTTCCACCACCACCCACCGACTCGGGCGCG
Cre01.g013000	GGAGGCGGAGGTATGAGTGTTGGCAGAATCGCA	GGTTCCACCACCACCCGCGTTGCGCTGGTGC
Cre01.g013100	GGAGGCGGAGGTATGCGTCCGGCATCACTAATAC	GGTTCCACCACCACCCACCTTCGGCAGGACGTTTT
Cre01.g013250	GGAGGCGGAGGTATGACCAATGGCGGTGGC	GGTTCCACCACCACCCCGCATGCTGTGCCACCT
Cre01.g013350	GGAGGCGGAGGTATGCTTGCTCAGGCTCGTA	GGTTCCACCACCACCCGCCAGCAGCAGCATG
Cre01.g013450	GGAGGCGGAGGTATGTCTTCCGCTTCTGCCG	GGTTCCACCACCACCAGCCGCGAAGCAGCTG
Cre01.g013550	GGAGGCGGAGGTATGACAGAGACTGATGCCGAC	GGTTCCACCACCACCGCCGTACATGGTCACGGAC
Cre01.g013650	GGAGGCGGAGGTATGCCGCGTGGTGGCAGAG	GGTTCCACCACCACCACCTGTAGCGGGGGCCGC
Cre01.g013750	GGAGGCGGAGGTATGTCGGCGCTTGCGCAG	GGTTCCACCACCACCGAAGCCCACCATGTCCTCC
Cre01.g013800	GGAGGCGGAGGTATGTCCACAGACAGCGACG	GGTTCCACCACCACCCCAGCGCGGCGTCTTA
Cre01.g013900	GGAGGCGGAGGTATGGTGGACTCTAGGTTGGAG	GGTTCCACCACCACCCTTCAAGATGGCTGCCCC
Cre01.g014050	GGAGGCGGAGGTATGGCGCCGAAGGCAG	GGTTCCACCACCACCTTGCCCTGCTGCCCCG

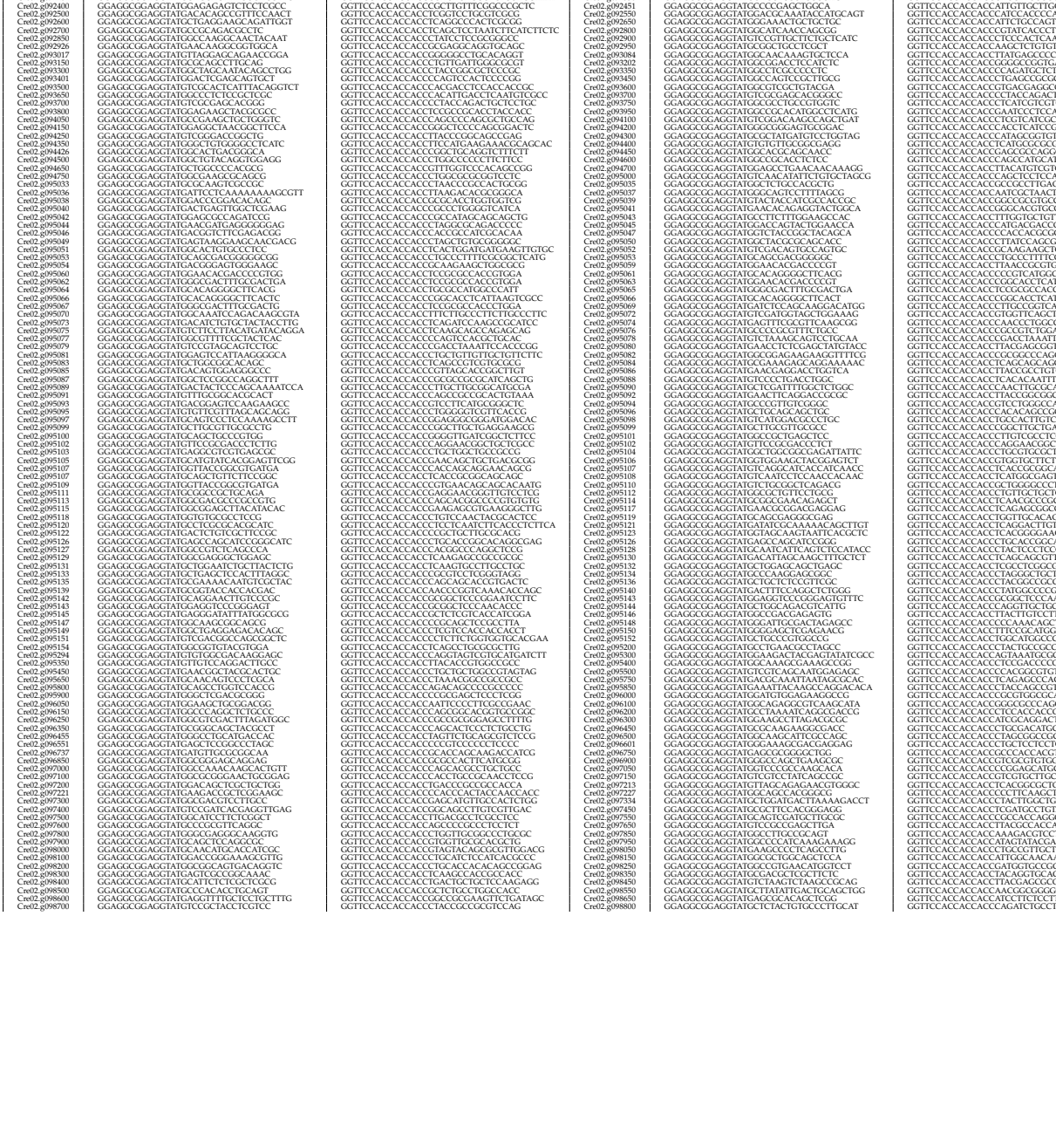
Gene ID	Forward primer	Reverse primer
Cre01.g000033	GGAGGCGGAGGTATGTGCGGTGCGCCTG	GGTTCCACCACCACCACCCCGGCACGACTGC
Cre01.g000150	GGAGGCGGAGGTATGGCGGCGAACGCGT	GGTTCCACCACCACCCGCCCACTGAGCCAGC
Cre01.g000300	GGAGGCGGAGGTATGAGCAGAAGTTTCTACACACA	GGTTCCACCACCACCTTACCGCACCCGACCA
Cre01.g000400	GGAGGCGGAGGTATGTCAATCACTGTCCTTGTTTTGT	GGTTCCACCACCACCCGTCCCTTTGCTGAGCCC
Cre01.g000500	GGAGGCGGAGGTATGGGCGCCCAGGGTATC	GGTTCCACCACCACCGCCGCCCGCATCGTCC
Cre01.g000600	GGAGGCGGAGGTATGACTGAGCCGCCACA	GGTTCCACCACCACCGTGCCTCCACAGCTTAAGC
Cre01.g000700	GGAGGCGGAGGTATGAGGCGACCCCGGG	GGTTCCACCACCACCCAGGTACGGCACCGCC
Cre01.g000800	GGAGGCGGAGGTATGTCGCTAGCTGTTGCAG	GGTTCCACCACCACCCTACCGGCGGCTCTGG
Cre01.g000900	GGAGGCGGAGGTATGCTTGGCCAGCGGG	GGTTCCACCACCACCCAGCCCCGAGGGCGAG
Cre01.g001000	GGAGGCGGAGGTATGGGTAAAGTGACTGCTGCG	GGTTCCACCACCACCTTAGCCCGGACCCGCC
Cre01.g001350	GGAGGCGGAGGTATGGGCGCTGGAGCTGCT	GGTTCCACCACCACCGTGGTGGTGGTGCTGCTG
Cre01.g001450	GGAGGCGGAGGTATGACGGTGAAATACAGGCG	GGTTCCACCACCACCCTTCGCACCCTTCATGCC
Cre01.g001550	GGAGGCGGAGGTATGCTAGCTCGGGGCGGC	GGTTCCACCACCACCCTTGGCTGCTGCGGTGGC
Cre01.g001650	GGAGGCGGAGGTATGGCGGCACTTCGCG	GGTTCCACCACCACCTCACTCGCAAGGCGCG
Cre01.g001664	GGAGGCGGAGGTATGTTCAACACCGACGGGA	GGTTCCACCACCACCCTAGCGCAGCCGCAGC
Cre01.g001678	GGAGGCGGAGGTATGTCGTGCTGCTTTGGTGG	GGTTCCACCACCACCGATGACCAGCAGGCTGGG
Cre01.g001750	GGAGGCGGAGGTATGATGCTGACATTGAAGAACCA	GGTTCCACCACCACCTTACACCGCCACGGGT
Cre01.g001950	GGAGGCGGAGGTATGCAGCAGCACATTCCG	GGTTCCACCACCACCTTACGTGAGCTGCGCCC
Cre01.g002050	GGAGGCGGAGGTATGTCTCGGCTACCGGTG	GGTTCCACCACCACCCTTGGCAGCTGCCGCC
Cre01.g002150	GGAGGCGGAGGTATGGTCTCGTGCCAAAACA	GGTTCCACCACCACCTCCCGACTCCATGGCC
Cre01.g002201	GGAGGCGGAGGTATGTCTGGAAGGCCGGT	GGTTCCACCACCACCCAGCAGGAATCCAATGCTACC
Cre01.g002227	GGAGGCGGAGGTATGTCCACAGCAGCGC	GGTTCCACCACCACCTCTGTGGTCCAGCAGCG
Cre01.g002250	GGAGGCGGAGGTATGCTGAACCACAAAGTCCA	GGTTCCACCACCACCCTACACCCGCAGCAGC
Cre01.g002350	GGAGGCGGAGGTATGGCGCCCGTAGCTGAC	GGTTCCACCACCACCCAGCTCCAGCACGGTGGT
Cre01.g002451	GGAGGCGGAGGTATGGCGTGGAAAACTCGTAG	GGTTCCACCACCACCGACCTTGGCCTTCTCGC
Cre01.g002550	GGAGGCGGAGGTATGCAGCAGGGCTCCC	GGTTCCACCACCACCTAGGTGCTTGTCGCCGC
Cre01.g002650	GGAGGCGGAGGTATGTGCAGCGTGGACG	GGTTCCACCACCACCGATCAGCGACAGCCACTC
Cre01.g002750	GGAGGCGGAGGTATGGACGAAACGCCTCGC	GGTTCCACCACCACCCATGGTGTTGCGGCTGT
Cre01.g002900	GGAGGCGGAGGTATGTGCAAGTTCTATTGCATGTC	GGTTCCACCACCACCCTAGTTGCCAGTGTTGTCGG
Cre01.g003000	GGAGGCGGAGGTATGGTGCCGCGCTCGG	GGTTCCACCACCACCTCATGATGCCGGCGCG
Cre01.g003050	GGAGGCGGAGGTATGGCCGTCCAGCCTCC	GGTTCCACCACCACCCTCCCCGCCTGCCCCT
Cre01.g003150	GGAGGCGGAGGTATGTGGCGCTGGCAGCGCATC	GGTTCCACCACCACCTAGGTCGCCGTCCCCCAG
Cre01.g003300	GGAGGCGGAGGTATGAAGGCTCCCATGCTTG	GGTTCCACCACCACCGGCAGCAATCGGCTCG
Cre01.g003450	GGAGGCGGAGGTATGGGCCACAGCAGGAG	GGTTCCACCACCACCTGACAGCAGCGACTCCA
Cre01.g003463	GGAGGCGGAGGTATGCTGGGCCCGGACG	GGTTCCACCACCACCTCAGCACGCCGCCACC
Cre01.g003463	GGAGGCGGAGGTATGCCACCCAAGGCGGC	GGTTCCACCACCACCTCAGCACGCCGCCACC
Cre01.g003487	GGAGGCGGAGGTATGTCTTTTGCTGTGCGGC	GGTTCCACCACCACCCATGTCGCCCATGCTCC
Cre01.g003516	GGAGGCGGAGGTATGATCCGGGCGCAGC	GGTTCCACCACCACCGCTCCACTTCTTCATGGCG
Cre01.g003541	GGAGGCGGAGGTATGCCCTTGACCTCGGAG	GGTTCCACCACCACCAGCCGTGTGGTACCCAT
Cre01.g003600	GGAGGCGGAGGTATGGCTTCGCGATGCTTCTG	GGTTCCACCACCACCGTCGGCCTGCGCCAGC
Cre01.g003826	GGAGGCGGAGGTATGCATGGGCGATGGCG	GGTTCCACCACCACCCAGCGCCGCAGCTGTC
Cre01.g003950	GGAGGCGGAGGTATGTCGAGCGTGGCGG	GGTTCCACCACCACCGGCCACATTGTCGGCG
Cre01.g004050	GGAGGCGGAGGTATGCCTGGCCAGATCACT	GGTTCCACCACCACCGTCGCTGTCGCACAGG
Cre01.g004250	GGAGGCGGAGGTATGGAGGGGGTAGATCCCG	GGTTCCACCACCACCGATCGCCAGACCAAACACCG
Cre01.g004350	GGAGGCGGAGGTATGCCGACCGCTCAATCG	GGTTCCACCACCACCCTGCTGGTAATACGGCTGC
Cre01.g004450	GGAGGCGGAGGTATGATGCTGGCCCAGAAGA	GGTTCCACCACCACCCAGCAGCGCCAGGGTG
Cre01.g004550	GGAGGCGGAGGTATGGAGGTGGAGGAGCCG	GGTTCCACCACCACCCTACCGCGCCGATCGC
Cre01.g004700	GGAGGCGGAGGTATGGCGCTCGTGCCTCTCTTG	GGTTCCACCACCACCGTTGGCCTTTGCCTTGGC
Cre01.g004800	GGAGGCGGAGGTATGGAGGTTGCGACCATAGG	GGTTCCACCACCACCGCTCATGATGGTGGCAGC
Cre01.g005001	GGAGGCGGAGGTATGCAATCGCTGCGGACG	GGTTCCACCACCACCTTATGCCGCGGCCGGC
Cre01.g005100	GGAGGCGGAGGTATGGACAGCAATAAGCGTCG	GGTTCCACCACCACCGGTGTTGAACAGCCGTTTG
Cre01.g005200	GGAGGCGGAGGTATGCGAAGGTCAACACGCC	GGTTCCACCACCACCCGTGCACTTCTTGAACAGCG
Cre01.g005300	GGAGGCGGAGGTATGGCAATCGTGTACGACG	GGTTCCACCACCACCCTAGCCGTAGGCGCCG
Cre01.g005400	GGAGGCGGAGGTATGGCGTCAGTGCAACAGG	GGTTCCACCACCACCGGCGCCCACGCTCATCATCATG
Cre01.g005500	GGAGGCGGAGGTATGAAATTCAAGCGTTATCTAGCA	GGTTCCACCACCACCTCACGCTCTCCCCAGC
Cre01.g005543	GGAGGCGGAGGTATGGCCAACCCTCCTAATAAGC	GGTTCCACCACCACCTCATGCTTGGCGGGGT
Cre01.g005600	GGAGGCGGAGGTATGATGCAAAGGCTTCGTGC	GGTTCCACCACCACCCTGCTGCCGCCGCTCT
Cre01.g005701	GGAGGCGGAGGTATGCTCGTAGTTAGACGGTTC	GGTTCCACCACCACCCCCTTCATAGTCCTGCGC
Cre01.g005850	GGAGGCGGAGGTATGAAGATCTACACGCGATCTGG	GGTTCCACCACCACCGCTGCGCGCCTTTTGGTAG
Cre01.g005950	GGAGGCGGAGGTATGCTGGGTTATAAGCTGAGGG	GGTTCCACCACCACCTTAGCGCTTGGCCGCG
Cre01.g006100	GGAGGCGGAGGTATGGTAGTCGCCTCCGCC	GGTTCCACCACCACCCGTGTCCCAGCCATCCTG
Cre01.g006202	GGAGGCGGAGGTATGGGCATTACCATTTTTTAGGC	GGTTCCACCACCACCTCAGGCAGCAGGGGGG
Cre01.g006301	GGAGGCGGAGGTATGCGGAGCAGCTCGT	GGTTCCACCACCACCGCTGCCCGACATCTTAAACC
Cre01.g006450	GGAGGCGGAGGTATGGAGGCTGGGGTGC	GGTTCCACCACCACCAGCCTCCAGCTCCATAGC
Cre01.g006550	GGAGGCGGAGGTATGCCCGTCTCCGAGC	GGTTCCACCACCACCCTAGCCCTGCAGCTCGC
Cre01.g006600	GGAGGCGGAGGTATGGCCAACTTGGCGGT	GGTTCCACCACCACCCGCCGGCGTGTTCTGC
Cre01.g006700	GGAGGCGGAGGTATGGTGTCCAAGGCCGT	GGTTCCACCACCACCGCCCTGAACCTCCTCCA
Cre01.g006766	GGAGGCGGAGGTATGGCCCCTCTTCCTTGC	GGTTCCACCACCACCGTTCTCAAACCCATGGCCG
Cre01.g006850	GGAGGCGGAGGTATGTCCTTAAGCGCCACG	GGTTCCACCACCACCCGCATCCGAGTCACTGC
Cre01.g006900	GGAGGCGGAGGTATGGGGCTGTACCTGCAG	GGTTCCACCACCACCGCGCGCACGTGAAGCT
Cre01.g007000	GGAGGCGGAGGTATGAAAATAAGAATCGAACCCGGG	GGTTCCACCACCACCGCGCTTCTGGTGGCGCAC
Cre01.g007051	GGAGGCGGAGGTATGCAAATCTTCGTGAAGACCC	GGTTCCACCACCACCTTACTTGAGCTTCTTCTTGGGG
Cre01.g007200	GGAGGCGGAGGTATGGCCAGCCTAAACGTTGAG	GGTTCCACCACCACCCGCTACATGGGGGCGCAC
Cre01.g007300	GGAGGCGGAGGTATGACCGTTGAAGGCTCAGC	GGTTCCACCACCACCTGTCGCCGCCTCCACTTT
Cre01.g007400	GGAGGCGGAGGTATGTTGCCGACACGGC	GGTTCCACCACCACCGCACCCACCAGCCGAC
Cre01.g007550	GGAGGCGGAGGTATGATACAGCCCAGTACAATTGG	GGTTCCACCACCACCCCTCCGCTGCTGTCTCCG
Cre01.g007651	GGAGGCGGAGGTATGGCTGCTTACTCTGACAGG	GGTTCCACCACCACCCGTCTGCATCGGCTCAG
Cre01.g007737	GGAGGCGGAGGTATGCGGACGCGTGTTC	GGTTCCACCACCACCTAGCCGCCCCTTCTCCT
Cre01.g007811	GGAGGCGGAGGTATGTCGACGAATGGAGTTACTG	GGTTCCACCACCACCTCACGCCGTCTTGCGC
Cre01.g007901	GGAGGCGGAGGTATGGAGGTTAGGGCGACACAAA	GGTTCCACCACCACCCCGCTCATCAAGCTCACCCAAT
Cre01.g008000	GGAGGCGGAGGTATGACGGGGCCAGACGT	GGTTCCACCACCACCTCACTCGTCCGCGCCC
Cre01.g008100	GGAGGCGGAGGTATGGCGCACTGCCTCAAG	GGTTCCACCACCACCCATGCCGGTACCCAAGGTT
Cre01.g008300	GGAGGCGGAGGTATGCAGCAACAACACCTGC	GGTTCCACCACCACCACACCCGATGGCAGCTACAA
Cre01.g008450	GGAGGCGGAGGTATGTCTTCTTTCAGCTTTCCGA	GGTTCCACCACCACCCTACAGCTGCATCGCCAG
Cre01.g008550	GGAGGCGGAGGTATGCGGCAAGCCACTCT	GGTTCCACCACCACCCTTGGGCGAGCGGTAGC
Cre01.g008700	GGAGGCGGAGGTATGGAGCCAATACTAGACGCG	GGTTCCACCACCACCCTACAGCCGTAGCGGCG
Cre01.g008800	GGAGGCGGAGGTATGACACAAGCTAATGCTGTTG	GGTTCCACCACCACCCTATACTCCGCCGCCGT
Cre01.g008891	GGAGGCGGAGGTATGATTCGTCCCAAGGTGTTGC	GGTTCCACCACCACCCAAAGAATCGCACACTGCCCAT
Cre01.g008950	GGAGGCGGAGGTATGGCCGCGCAACCTT	GGTTCCACCACCACCCAGGCCCGTCACGCCA
Cre01.g009000	GGAGGCGGAGGTATGACTCTTGAGGTCATTGGGA	GGTTCCACCACCACCTCACCGACGGTAACGCC
Cre01.g009050	GGAGGCGGAGGTATGAGGCGGAGCTACGC	GGTTCCACCACCACCTTAAACAAAGAACAACCGGAACAA
Cre01.g009150	GGAGGCGGAGGTATGGCGGACTACGTGCT	GGTTCCACCACCACCCGCCTGCTTCTTCACGC
Cre01.g009250	GGAGGCGGAGGTATGGAGATCGATATTATGGAGGACG	GGTTCCACCACCACCGTCCGACAGCACCTCGAAG
Cre01.g009300	GGAGGCGGAGGTATGGCGGTTCCAGGGAA	GGTTCCACCACCACCCAACAGCTCCGAGTCCAGC
Cre01.g009400	GGAGGCGGAGGTATGAGCGATCAAACCTACTGC	GGTTCCACCACCACCTCACACGCCCTCGACG
Cre01.g009500	GGAGGCGGAGGTATGGGTTGCTCGTCTTCAAAG	GGTTCCACCACCACCCGCCCGATGCGACTGC
Cre01.g009601	GGAGGCGGAGGTATGATGGACGATTGGCAGC	GGTTCCACCACCACCGTTCCTGCTGCTGGGC
Cre01.g009676	GGAGGCGGAGGTATGCTCATAACGAGGTTTTTAAGG	GGTTCCACCACCACCCTATTGCATCAGTCCATGCTTAC
Cre01.g009765	GGAGGCGGAGGTATGAGGACTGCCTCGCA	GGTTCCACCACCACCCTACTCCCCAGCCGCC
Cre01.g009850	GGAGGCGGAGGTATGTCGCTGGAAGAAGAGTTG	GGTTCCACCACCACCCGCGAAGGGCATATTCCG
Cre01.g009950	GGAGGCGGAGGTATGGTCGCTGGAGGGTT	GGTTCCACCACCACCTTACACGATCTTCCACTTCATCTC
Cre01.g010100	GGAGGCGGAGGTATGGACTCCAGAGGCAGC	GGTTCCACCACCACCCTAGGCGGCGTTGCCC
Cre01.g010296	GGAGGCGGAGGTATGGAGAACACCCCGTCC	GGTTCCACCACCACCGAGTCCCCAGCGCGTG
Cre01.g010350	GGAGGCGGAGGTATGCATGTTCGCCGTCCG	GGTTCCACCACCACCAGCAGCGCGCGGTTCGAAG
Cre01.g010450	GGAGGCGGAGGTATGAAATCGCGGAGAGCC	GGTTCCACCACCACCCTGGTCCTTTTTCTCCTCCG
Cre01.g010550	GGAGGCGGAGGTATGCGCAACACCCGGC	GGTTCCACCACCACCCTACCGCGCCTGCGGC
Cre01.g010700	GGAGGCGGAGGTATGAGCTTCCAGGACCTGG	GGTTCCACCACCACCCGTGAGGATGATGATGAGCAG
Cre01.g010800	GGAGGCGGAGGTATGGATCGCGGAAGGCGGTTC	GGTTCCACCACCACCCAGCAGTAGCGAGCCGGA
Cre01.g010832	GGAGGCGGAGGTATGCGCGCGCGTGTTA	GGTTCCACCACCACCCACCGCCTGATGGTGCA
Cre01.g010900	GGAGGCGGAGGTATGGCCGCCATGATGCAGAAG	GGTTCCACCACCACCCGCCACCCACTTCTTGGC
Cre01.g011050	GGAGGCGGAGGTATGGGCCATGACCCCG	GGTTCCACCACCACCTTATGGTACAGTTTGCAATCGTAC
Cre01.g011150	GGAGGCGGAGGTATGGCGTCCCTCGAATTTTG	GGTTCCACCACCACCGCGATGGGTCCGGGAG
Cre01.g011350	GGAGGCGGAGGTATGGACGTAGACGAGCAGTTG	GGTTCCACCACCACCCCGGAACATGCGCCTGCC
Cre01.g011400	GGAGGCGGAGGTATGCCGAAAAGCTCTGATGG	GGTTCCACCACCACCTTCCCACTTCAGCAGCGC
Cre01.g011500	GGAGGCGGAGGTATGGACGGCTTGCAGCG	GGTTCCACCACCACCGAAGACCACCGTGTCCAGC
Cre01.g011600	GGAGGCGGAGGTATGTCAAACCATACTTCCAGCAG	GGTTCCACCACCACCCGCATCGAACAGCCGTATG
Cre01.g011660	GGAGGCGGAGGTATGCTGGCGGCTGCTG	GGTTCCACCACCACCGACGCGTGGCAGGACG
Cre01.g011800	GGAGGCGGAGGTATGCTGCCTCTGCCGC	GGTTCCACCACCACCCTACATCTTTAGCTGCTGCATCA
Cre01.g011901	GGAGGCGGAGGTATGAGTGCCACGCCCC	GGTTCCACCACCACCTGGACCAGCGGCAAACT
Cre01.g012100	GGAGGCGGAGGTATGTTCATCACTTTAGCCTGTAGC	GGTTCCACCACCACCTCCACGTCCACGTGTGCG
Cre01.g012150	GGAGGCGGAGGTATGGCGCTTCGTGCTG	GGTTCCACCACCACCGCCGTAGCAGCGGATG
Cre01.g012222	GGAGGCGGAGGTATGGTGCCGCCCGTTG	GGTTCCACCACCACCCTACTGCTGCTGCGGGC
Cre01.g012300	GGAGGCGGAGGTATGCTGCTAGCTGCCCG	GGTTCCACCACCACCGAAGGTCTCCTGCGGCG
Cre01.g012400	GGAGGCGGAGGTATGACAACTGGCGAGCTCG	GGTTCCACCACCACCCGTCGCCGTGTCCATCGA
Cre01.g012500	GGAGGCGGAGGTATGAGCGCGCCAGTGC	GGTTCCACCACCACCTACTGCCGTGGAAACGGT
Cre01.g012600	GGAGGCGGAGGTATGCAGGCCCCCCGCCTTGTC	GGTTCCACCACCACCGTGCGGCGTCGCCGGCCGCTTC
Cre01.g012850	GGAGGCGGAGGTATGCGTGTGTCAGACCTCCTT	GGTTCCACCACCACCGGCGTTTCGGTCCTGCGAA
Cre01.g012950	GGAGGCGGAGGTATGCTGCAACTAGATCATGGGA	GGTTCCACCACCACCGCAAATCGTGCAGCTGCAG
Cre01.g013050	GGAGGCGGAGGTATGGCAAGCGACAATGCTCC	GGTTCCACCACCACCCTCATCGGCCGCCTCTTC
Cre01.g013200	GGAGGCGGAGGTATGGCCAAGCACGTAATCC	GGTTCCACCACCACCAACCTTGCCCAGCCGC
Cre01.g013300	GGAGGCGGAGGTATGCCGCTACAAGGCCG	GGTTCCACCACCACCTCATTGCTGCGCGGCC
Cre01.g013400	GGAGGCGGAGGTATGGCGGAGTCTGAGCGA	GGTTCCACCACCACCGCTCTCGCTGCCCTTGAC
Cre01.g013500	GGAGGCGGAGGTATGGCCGAGGTCGCTG	GGTTCCACCACCACCCGCGGCGAAGCACAGG
Cre01.g013600	GGAGGCGGAGGTATGTCGTCTCAAACTTCTAGCA	GGTTCCACCACCACCCGACTTGCTCTTGTCGCCATC
Cre01.g013700	GGAGGCGGAGGTATGCCGGCTGCCGTCGCTTTTG	GGTTCCACCACCACCGGCCAGGTCCAGGGCAAAG
Cre01.g013769	GGAGGCGGAGGTATGAGCTCGTTTCCAGGTTTTGT	GGTTCCACCACCACCTCACGCCGCCGCACAG
Cre01.g013850	GGAGGCGGAGGTATGGAGGATGGCACAATCGCTT	GGTTCCACCACCACCGTAACCCTCCTCTTCGTCCTC
Cre01.g014000	GGAGGCGGAGGTATGCAGCGTGCCTGTAGG	GGTTCCACCACCACCATCGATGCCGGCAATGC
Cre01.g014100	GGAGGCGGAGGTATGATAGGAGCAGACGAAGAGG	GGTTCCACCACCACCCACCATCACCTCCTCCCAGTC

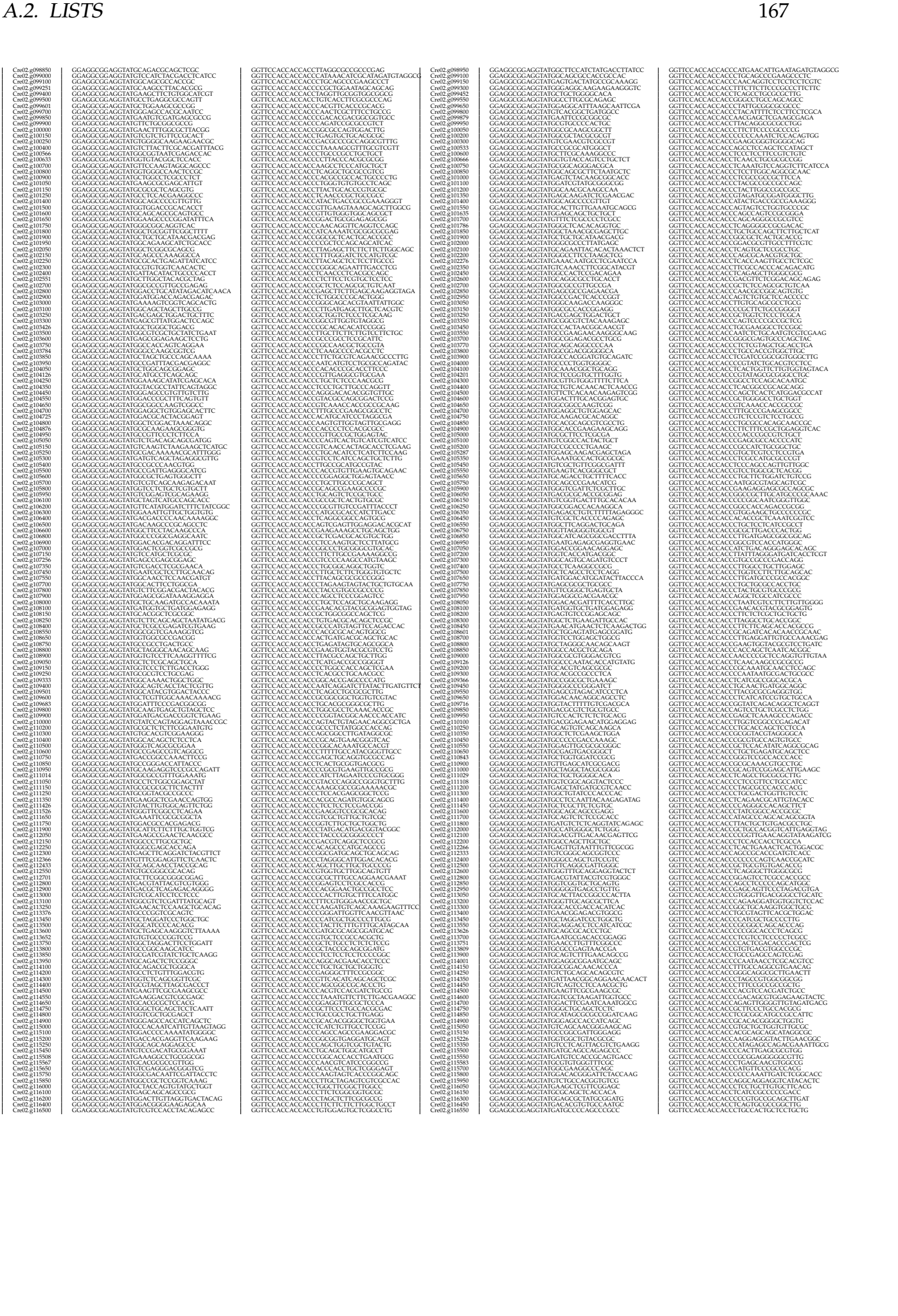


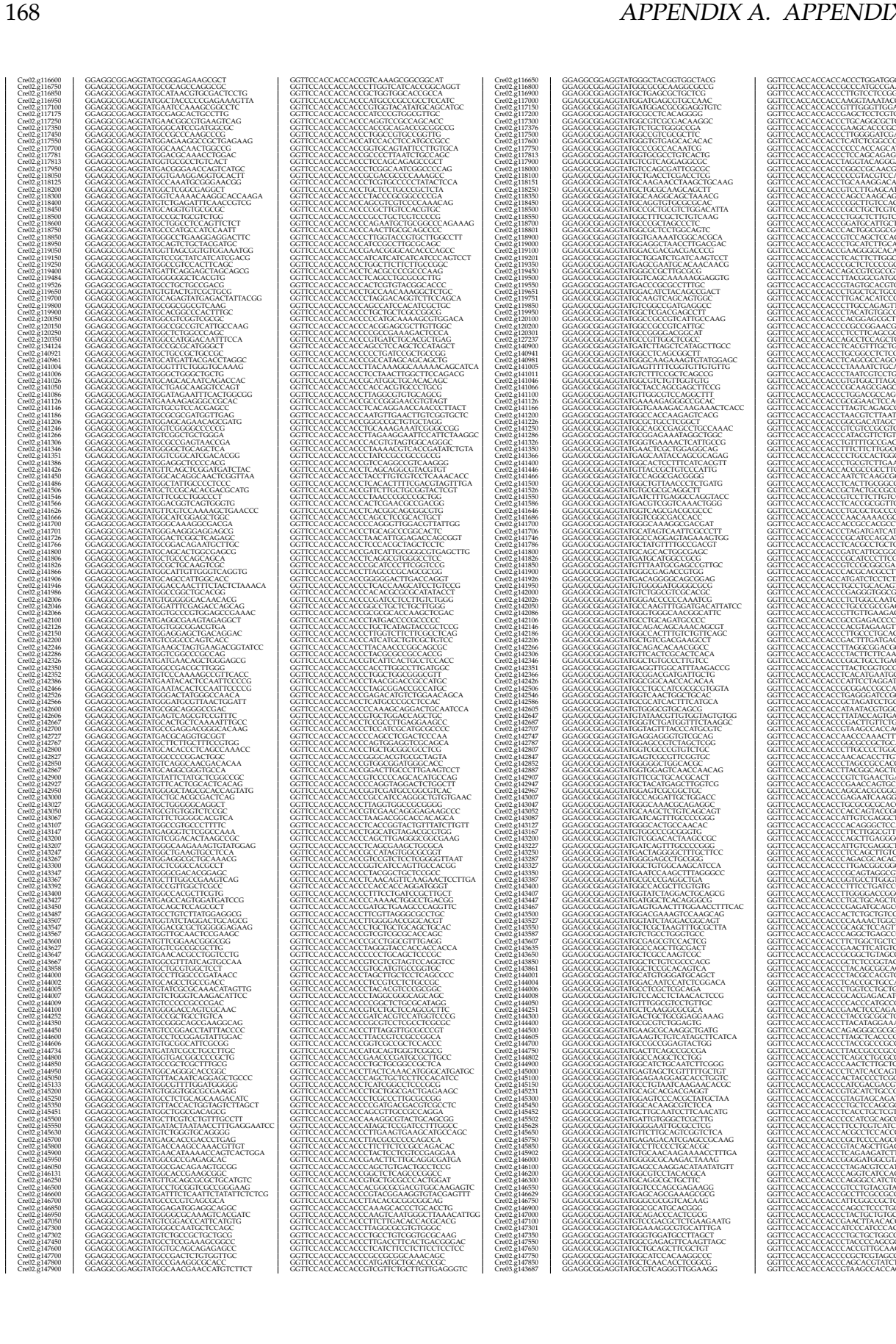


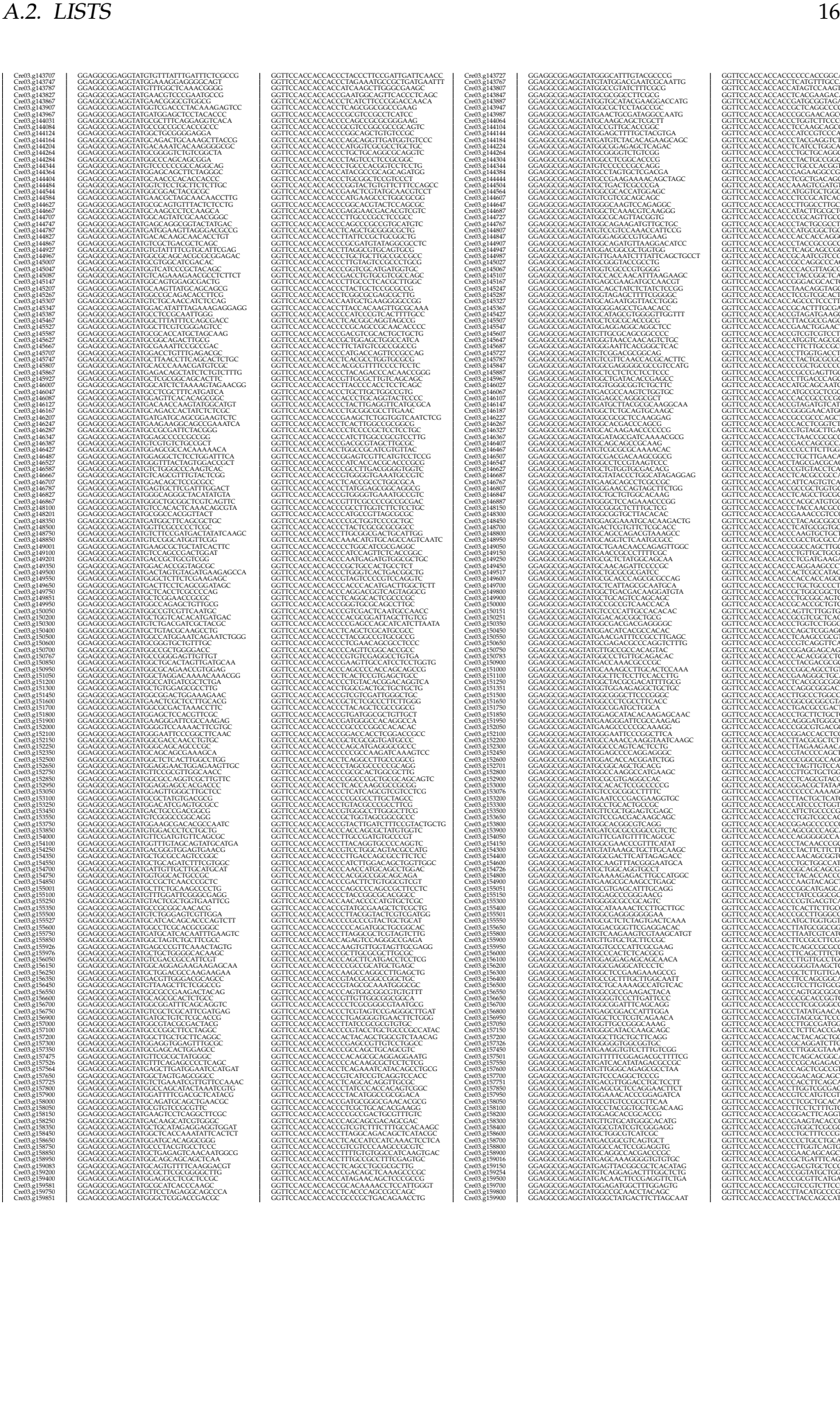
Cre01.g053600	GGAGGCGGAGGTATGAAGCTCACGGCTGGC	GGTTCCACCACCACCCGAAGCCTTCTTTGCTGCT
Cre01.g053800	GGAGGCGGAGGTATGATGTCTTTCGGCCGAC	GGTTCCACCACCACCCCGCCAGACTGTAGGTAGG
Cre01.g053900	GGAGGCGGAGGTATGGACGACGAGGCTCTCG	GGTTCCACCACCACCCTACCGCGCCAGCGGC
Cre01.g054000	GGAGGCGGAGGTATGAAACTCCCGCCAGGA	GGTTCCACCACCACCTCATGCCTCACCGCCC
Cre01.g054100	GGAGGCGGAGGTATGGCGGGACACACGTC	GGTTCCACCACCACCTGCGGGGATGCGGTCC
Cre01.g054200	GGAGGCGGAGGTATGTTCCGGCGGAGCC	GGTTCCACCACCACCGTCCCAGACTTCGTCCTTGA
Cre01.g054300	GGAGGCGGAGGTATGGCCCTTGGACTTTCGC	GGTTCCACCACCACCCAGCCCGTTCCACTCGC
Cre01.g054400	GGAGGCGGAGGTATGCCGCTTGGCGCCG	GGTTCCACCACCACCCAGGTGCGCCATGTCCCA
Cre01.g054500	GGAGGCGGAGGTATGCTATCCATGCTCGGTCGTA	GGTTCCACCACCACCGGCCTTGATGCCCAGGGTC
Cre01.g054600	GGAGGCGGAGGTATGAGGGTGGTGGTGCA	GGTTCCACCACCACCCTGGCTCTCCTCGTGCC
Cre01.g054700	GGAGGCGGAGGTATGCAGGCACTCGGGT	GGTTCCACCACCACCCTGTTGCCGAGCATCCG
Cre01.g054800	GGAGGCGGAGGTATGAGTACGCCCGGGC	GGTTCCACCACCACCCAGGTCCAGTGTGAGGTGG
Cre01.g054900	GGAGGCGGAGGTATGAACAGCACCATGAGCTG	GGTTCCACCACCACCCAGCACGTCCATCTCGAC
Cre01.g055000	GGAGGCGGAGGTATGTCGGCCAGCGCTCC	GGTTCCACCACCACCCGATGGAGGCGCCGGC
Cre01.g055100	GGAGGCGGAGGTATGGACGGGCCTCCTCG	GGTTCCACCACCACCCATCACCCTCCGTGCTGA
Cre01.g055200	GGAGGCGGAGGTATGGCGGCAGCCTCAAAA	GGTTCCACCACCACCCTGCCCGGCTAGGTTCAG
Cre01.g055300	GGAGGCGGAGGTATGGTGGGGATGAGCGCA	GGTTCCACCACCACCGAGCCCTTGCCCGCGC
Cre01.g055400	GGAGGCGGAGGTATGGCCAAGGACAAGCGC	GGTTCCACCACCACCCTCGTCGTCGCTCTCAAACTT
Cre01.g055408	GGAGGCGGAGGTATGCAAACAAAGTCTTTCAGCCC	GGTTCCACCACCACCCTTGCCGCGCAGCTCCAC
Cre01.g055416	GGAGGCGGAGGTATGCAGATGGACGAGGTGC	GGTTCCACCACCACCGTAGCGGTACTTGGTCGAGTA
Cre01.g055424	GGAGGCGGAGGTATGGCGGGAGATGTGGAA	GGTTCCACCACCACCGCGGCTAGCGCTGGTG
Cre01.g055432	GGAGGCGGAGGTATGAAGCAGGCATGCCTGC	GGTTCCACCACCACCGCCCTCAAAGATGACGCTCG
Cre01.g055436	GGAGGCGGAGGTATGTCAAAGCGCGGGGC	GGTTCCACCACCACCTCACGACCCAGCGCCG
Cre01.g055444	GGAGGCGGAGGTATGCTGGACGTCTATCGCC	GGTTCCACCACCACCCTCCTCGTCCTCCTCCTCC
Cre01.g055457	GGAGGCGGAGGTATGAGCGTTAGAAATCGAACGA	GGTTCCACCACCACCCTGCTCCGTCATCCACCG
Cre01.g055465	GGAGGCGGAGGTATGACAAAAAACATCGTGGACGC	GGTTCCACCACCACCGGCCACCACCTGCACGCTG
Cre01.g055477	GGAGGCGGAGGTATGTCCGGCAGGGGTGGAG	GGTTCCACCACCACCTGAACGGACACTCCAGTACAC
Cre01.g056696	GGAGGCGGAGGTATGTGCAGCAGGTCTGAGC	GGTTCCACCACCACCGTCTCTGTGTCCTCCCAGC
Cre01.g057426	GGAGGCGGAGGTATGACGGGTGTTGGTGCT	GGTTCCACCACCACCACGAAGTGCTTGCGAGTG
Cre01.g058156	GGAGGCGGAGGTATGTGCGAAGGACGTGACG	GGTTCCACCACCACCTCAGCCTGCGCGCTTG
Cre01.g059252	GGAGGCGGAGGTATGTTGCTGACGCGGCTG	GGTTCCACCACCACCCTGCTGCCGGCTCGGCAC
Cre01.g059982	GGAGGCGGAGGTATGTTGCTGACGCGGCT	GGTTCCACCACCACCTCACTGCTGCCGGCCC
Cre01.g061077	GGAGGCGGAGGTATGGCCGCAACTCTTGTC	GGTTCCACCACCACCGTAGGACTCGTTGCTGCTG
Cre01.g061807	GGAGGCGGAGGTATGGACGCTGTAGAGAGGCTC	GGTTCCACCACCACCGTCGCGGCGCTGCTTGTT
Cre01.g062537	GGAGGCGGAGGTATGGGCGGGCACGACG	GGTTCCACCACCACCTTCGAACGGCGCCCTG
Cre01.g063632	GGAGGCGGAGGTATGGCGGACCCGCCCACG	GGTTCCACCACCACCCGAGCCGACACCTGCGCC
Cre01.g064727	GGAGGCGGAGGTATGCAAAGCGGACTGCTTG	GGTTCCACCACCACCGCTCACAGCAGACAGCTGCATG
Cre01.g065457	GGAGGCGGAGGTATGCAGCCACACTTGGAGC	GGTTCCACCACCACCGGCCAGAGAGGGAGGCCG
Cre01.g066187	GGAGGCGGAGGTATGGCGTCTAGCGGAGAT	GGTTCCACCACCACCTGTGGCGCTGAACAGC
Cre01.g066917	GGAGGCGGAGGTATGGCCTTCGCCCTTGC	GGTTCCACCACCACCCTGGGGGGTGAACTTGGT
Cre01.g068012	GGAGGCGGAGGTATGGGGAGCTTGGGCCATGAAG	GGTTCCACCACCACCCCGCAAACCGCCCGGCCG
Cre01.g068742	GGAGGCGGAGGTATGGCAGCTGAGCCAGAC	GGTTCCACCACCACCCTTCAGGTCCTGCCCGCC
Cre01.g069472	GGAGGCGGAGGTATGGCGCACCGGCCTGTG	GGTTCCACCACCACCCAGGCTGGCCGCGCTGCCTG
Cre01.g069837	GGAGGCGGAGGTATGGGCACGCTCGTGG	GGTTCCACCACCACCTGCCTGCCACTGCGGT
Cre01.g070567	GGAGGCGGAGGTATGTCGTCAAAACGTCAGCG	GGTTCCACCACCACCCTTCTTGCCACCCGCCTT
Cre01.g071662	GGAGGCGGAGGTATGTACGAGCGCAGCATCA	GGTTCCACCACCACCCTTGCCGCGCAGCTCCAC
Cre02.g073150	GGAGGCGGAGGTATGCAGGTACCTGCCTCC	GGTTCCACCACCACCCTGCTGGCTGGAGTGCT
Cre02.g073250	GGAGGCGGAGGTATGGCGCAGCGCGAAAGG	GGTTCCACCACCACCCCTCCGGTCCCTGTCGCG
Cre02.g073350	GGAGGCGGAGGTATGCCAGAAGGTAGTGGCA	GGTTCCACCACCACCCTGCACGGGGTGTACGC
Cre02.g073500	GGAGGCGGAGGTATGATGACAGATAGGCCTCCAAC	GGTTCCACCACCACCTTTGCTCATGCTACTTGCCG
Cre02.g073600	GGAGGCGGAGGTATGCAGCCTGCCTACCAAG	GGTTCCACCACCACCGGCCTTCAGATGCCACTCTG
Cre02.g073700	GGAGGCGGAGGTATGCAGAGCTCCCTGACCC	GGTTCCACCACCACCCACCAGCACGGGCTGGCG
Cre02.g073800	GGAGGCGGAGGTATGGAGGAGGCAAGGGCA	GGTTCCACCACCACCAAAGTCGATGCACTTGCCCTT
Cre02.g073850	GGAGGCGGAGGTATGACCTTGTCACAGGCTGC	GGTTCCACCACCACCGAACAGCCCAAGCACCTTG
Cre02.g073950	GGAGGCGGAGGTATGCTGAAGCGGTACAGTGC	GGTTCCACCACCACCTTTCTGCTGTACGAATGTGCCG
Cre02.g074100	GGAGGCGGAGGTATGAACAGGCTCACAAGTAAG	GGTTCCACCACCACCTCAGTTAATCTTAATGCCCTTGT
Cre02.g074200	GGAGGCGGAGGTATGTCACGTATCTTGCTAGGAAC	GGTTCCACCACCACCTCAGAAGTCAAACATGCCGC
Cre02.g074350	GGAGGCGGAGGTATGGGGCCAGAAGAGCG	GGTTCCACCACCACCTGCATCTTTCTCCTTGCGAA
Cre02.g074437	GGAGGCGGAGGTATGGTCACATCAAACCACAGC	GGTTCCACCACCACCGGACAGTGCTCTGCCAGA
Cre02.g074511	GGAGGCGGAGGTATGCGCACGCAAAACTCT	GGTTCCACCACCACCGCCCAGCCAGTTCTTGG
Cre02.g074600	GGAGGCGGAGGTATGGGGTTCACGGTGCTG	GGTTCCACCACCACCCATGTCCAGCAGGTGGCC
Cre02.g074700	GGAGGCGGAGGTATGGCTTGCAGCCTTGCC	GGTTCCACCACCACCGATGAGGTACCGCGCGCG
Cre02.g074737	GGAGGCGGAGGTATGGCCGACTTCGGCA	GGTTCCACCACCACCGCTCCAAGCGGATGGC
Cre02.g074758	GGAGGCGGAGGTATGCAGCAGCGGCAAA	GGTTCCACCACCACCGCCGTGCGTATCAAACAGT
Cre02.g074850	GGAGGCGGAGGTATGAATATATTGCCTCAAGAAACTGGT	GGTTCCACCACCACCAACAGGCGGTTCGGTGGAG
Cre02.g075000	GGAGGCGGAGGTATGGCCGTCGCTCCCG	GGTTCCACCACCACCGGTTCCACCATTGCGGATCA
Cre02.g075100	GGAGGCGGAGGTATGACCGACTTGGCCTCTCA	GGTTCCACCACCACCCCGCTCGAAGCGGATGCTGTAG
Cre02.g075200	GGAGGCGGAGGTATGCTGCGTCGAGGGC	GGTTCCACCACCACCTGCCGCACCAACCCAC
Cre02.g075350	GGAGGCGGAGGTATGAGCGCCGGCATGGCATAC	GGTTCCACCACCACCTGAATACGACACTCGGCGC
Cre02.g075500	GGAGGCGGAGGTATGACTTCTCCCTCTGCTGTTCT	GGTTCCACCACCACCTCAGCGCATGCGGGGC
Cre02.g075650	GGAGGCGGAGGTATGGCGGCCCGGCTCCTG	GGTTCCACCACCACCCTGGTGCTGCTGCTGCGG
Cre02.g075700	GGAGGCGGAGGTATGGTGTCCCTCAAGCTGC	GGTTCCACCACCACCCTTCTTCTGGGCGGGGG
Cre02.g075800	GGAGGCGGAGGTATGGCCGTGGGTTCGC	GGTTCCACCACCACCTATGCACGGGCAGCAGC
Cre02.g075900	GGAGGCGGAGGTATGGCTTCAGCACCGTCAGC	GGTTCCACCACCACCTCGTTGCTGCGGCGCC
Cre02.g076000	GGAGGCGGAGGTATGGCGAATCAGATTTCATTTAGCG	GGTTCCACCACCACCCTCCAGCTGGAACTCGCTG
Cre02.g076150	GGAGGCGGAGGTATGTATGCTCGCATCGGC	GGTTCCACCACCACCCTACCCGGCGTTTGGC
Cre02.g076300	GGAGGCGGAGGTATGAACAAGAGCGTCGGTG	GGTTCCACCACCACCTCCATTACGCCCAGTAAGGA
Cre02.g076400	GGAGGCGGAGGTATGGGCAGCTTCGCCTTT	GGTTCCACCACCACCAATGAACTTGACACCGAGTTCC
Cre02.g076466	GGAGGCGGAGGTATGGCCACCTTTGCCAAC	GGTTCCACCACCACCAACCTTCAGGATGTACCACGC
Cre02.g076550	GGAGGCGGAGGTATGTTGCGGCTGAGGCA	GGTTCCACCACCACCCTACATTGAGCCTGGAAACATGTA
Cre02.g076625	GGAGGCGGAGGTATGGCTGAAGCGCCCGC	GGTTCCACCACCACCTCAACCTGCGCGCCGG
Cre02.g076700	GGAGGCGGAGGTATGGCAGCACAGGTTTCAGAC	GGTTCCACCACCACCTCACCGCCCCTGCCCT
Cre02.g076850	GGAGGCGGAGGTATGGCGGAAGTGGAGCCC	GGTTCCACCACCACCTGCATCACTACTGCGACCG
Cre02.g076950	GGAGGCGGAGGTATGGGCAGGGCGTGGGGG	GGTTCCACCACCACCGATGCGCGTCAGGCGGAA
Cre02.g076987	GGAGGCGGAGGTATGCTTGCGCTACTTGCGC	GGTTCCACCACCACCTACCAGCCGCACGCGG
Cre02.g077061	GGAGGCGGAGGTATGCCGGAGGTGCCCG	GGTTCCACCACCACCAGACGGCGAGCCCCTA
Cre02.g077150	GGAGGCGGAGGTATGAATAACAGCGCATTACGGC	GGTTCCACCACCACCCGCAATGTGCCTGAGCGC
Cre02.g077250	GGAGGCGGAGGTATGCCTAGAAGCGCCACC	GGTTCCACCACCACCCCCTAGCCGGGCTGGC
Cre02.g077350	GGAGGCGGAGGTATGGAGGCCCCGTCTACG	GGTTCCACCACCACCCTGCATGCCCAAACGCAG
Cre02.g077451	GGAGGCGGAGGTATGACAGGCAGCTGGTTTAC	GGTTCCACCACCACCCGTGCTGGGAAGGGCC
Cre02.g077550	GGAGGCGGAGGTATGAGGATCCGCGCGCT	GGTTCCACCACCACCTTACCGCCCCCCGGGC
Cre02.g077600	GGAGGCGGAGGTATGGCCCAGCAGCTGA	GGTTCCACCACCACCTTTCCCCCCGGTTGAATCA
Cre02.g077750	GGAGGCGGAGGTATGCGCTCATCGAGCGCG	GGTTCCACCACCACCCTTGCGGGCGCGGTCCAG
Cre02.g077850	GGAGGCGGAGGTATGCTGTTGCTTTGTGGGCTT	GGTTCCACCACCACCGAGCTTCTTGCGGCCGAA
Cre02.g077925	GGAGGCGGAGGTATGGCGTGTGGCGGGG	GGTTCCACCACCACCCCGTGTGGCGGCAACG
Cre02.g078050	GGAGGCGGAGGTATGTCCTCTGAGGAGGTGCT	GGTTCCACCACCACCCTACCGCCCACCGGCG
Cre02.g078100	GGAGGCGGAGGTATGGGAAAGAAGAAGGCAGC	GGTTCCACCACCACCGGTGACGTTGAAGGTCTTCTT
Cre02.g078226	GGAGGCGGAGGTATGGTGACCATTCGCGAAG	GGTTCCACCACCACCACCAACCGCATGCCAATC
Cre02.g078300	GGAGGCGGAGGTATGGACAAGGTGGACGTGA	GGTTCCACCACCACCGCGGCTGTTGATGAGCT
Cre02.g078350	GGAGGCGGAGGTATGAGCTGGCTGAAGCGTA	GGTTCCACCACCACCCATCGGTGGCGAGTGC
Cre02.g078450	GGAGGCGGAGGTATGCCCACCCCAGCAG	GGTTCCACCACCACCATAACCCCACGCGGCC
Cre02.g078500	GGAGGCGGAGGTATGTTCCAGAACAACTTCAGCG	GGTTCCACCACCACCAACCCGGTACTCGTAAGCCA
Cre02.g078550	GGAGGCGGAGGTATGGCAACTTACTCTGTAATAAGCA	GGTTCCACCACCACCCCGCATCGTCTCCAGGTTG
Cre02.g078650	GGAGGCGGAGGTATGGGCGTGCAACCAAAACTAT	GGTTCCACCACCACCCAGTTCGTCGTGCGCCCC
Cre02.g078777	GGAGGCGGAGGTATGGCGGAATGTCACGCG	GGTTCCACCACCACCGTTTACTATCGTCGCAGGGCAG
Cre02.g078831	GGAGGCGGAGGTATGGACGGTCGGGATGATGG	GGTTCCACCACCACCTCAAGCGGGCCTGGGC
Cre02.g078885	GGAGGCGGAGGTATGTCACAACCTCTGACTCAGC	GGTTCCACCACCACCCATACTATCCTGTCCCCCACCA
Cre02.g078939	GGAGGCGGAGGTATGCACGCTCGACTTTTGC	GGTTCCACCACCACCCTTGTACCGGTCAAACTCCG
Cre02.g079000	GGAGGCGGAGGTATGGCGGGCAGCCCTA	GGTTCCACCACCACCGTCGTACATCTGGGCGTAC
Cre02.g079004	GGAGGCGGAGGTATGGCGCAGCCACCAT	GGTTCCACCACCACCCGCCCGCGCAAAACCT
Cre02.g079050	GGAGGCGGAGGTATGGTGGCCAGACACCAGGATC	GGTTCCACCACCACCGCACACCTCCAGCGGCCG
Cre02.g079150	GGAGGCGGAGGTATGCTGCTCGTACTAGTGGG	GGTTCCACCACCACCCCTCCCACCGCCTTTCA
Cre02.g079250	GGAGGCGGAGGTATGGAAAACACCACATACATGTTT	GGTTCCACCACCACCTCACGCCTCCTGAGCC
Cre02.g079400	GGAGGCGGAGGTATGCTGCGCCTCGCTG	GGTTCCACCACCACCCTGGCTAGGCTGCGCC
Cre02.g079500	GGAGGCGGAGGTATGGGCGCGGGAACAA	GGTTCCACCACCACCGTACGCGGCGACGTCC
Cre02.g079600	GGAGGCGGAGGTATGTCTCTTTTGAGGGCCCC	GGTTCCACCACCACCCCGCAGCTCCACCTTGAG
Cre02.g079700	GGAGGCGGAGGTATGCAGACCGCGCTTCAC	GGTTCCACCACCACCCTTGTTCAGGTCATCAGCGGTC
Cre02.g079850	GGAGGCGGAGGTATGGCGGGGGGCTACACTC	GGTTCCACCACCACCGGTTAGGGTCCGGGATGAGC
Cre02.g080000	GGAGGCGGAGGTATGTACTCAAGGCCACGTACC	GGTTCCACCACCACCAGTGTTGGTATGCTCGTCG
Cre02.g080100	GGAGGCGGAGGTATGGCGTCGGCAGTAGAAGTTT	GGTTCCACCACCACCCGTGGTCTGCGCATCACC
Cre02.g080150	GGAGGCGGAGGTATGGTGCCACATAAAACCGC	GGTTCCACCACCACCTCGCTGGCCTTGCGAG
Cre02.g080250	GGAGGCGGAGGTATGGCGACCTGCGCCG	GGTTCCACCACCACCGTGGCCGTCAAAGTAGCACA
Cre02.g080350	GGAGGCGGAGGTATGTCGGTGGTGGACGAGG	GGTTCCACCACCACCCTGCTGTGCGGGCAGGCG
Cre02.g080500	GGAGGCGGAGGTATGACCACATTATCAATCCTGGGG	GGTTCCACCACCACCCGCGTCCTTGTCCTTCGAGTTC
Cre02.g080650	GGAGGCGGAGGTATGGACCGGAGGGCCATC	GGTTCCACCACCACCCAGCTCGTCCTTCTCCTCC
Cre02.g080750	GGAGGCGGAGGTATGTCGGAGCCAACCGTTAG	GGTTCCACCACCACCCTCCAGCTTCTTCGCCGGAAA
Cre02.g080850	GGAGGCGGAGGTATGCCAGTATTTTTGTCGTCATGG	GGTTCCACCACCACCTTCGCTCAAGTCCAACGCC
Cre02.g080950	GGAGGCGGAGGTATGGTGCAGCTAGACCAAGT	GGTTCCACCACCACCTGCACACTCCGCATCTGG
Cre02.g081100	GGAGGCGGAGGTATGGCCGGAACAGAGCCC	GGTTCCACCACCACCGTTCCCCTCCATGAGGTGC
Cre02.g081176	GGAGGCGGAGGTATGCGCGCGCACGTTTTAAG	GGTTCCACCACCACCTGCCGACGACGGCGCCAC
Cre02.g081250	GGAGGCGGAGGTATGCAGCTCCTTGCAGTG	GGTTCCACCACCACCCTTGCGCTTCTGGGCG
Cre02.g081350	GGAGGCGGAGGTATGGAGCAGAGCCATTCAGC	GGTTCCACCACCACCCTGGATCGGGTGGGGC
Cre02.g081450	GGAGGCGGAGGTATGCGGTGCGCGACTG	GGTTCCACCACCACCGTCCTTGCCCACATCAGTG
Cre02.g081600	GGAGGCGGAGGTATGGCGCTACCGCGCGTG	GGTTCCACCACCACCCACCATGGGCTTGAAGGGC
Cre02.g081750	GGAGGCGGAGGTATGAACCCTCAGGGCATACTA	GGTTCCACCACCACCCATGCCACGACCCGCA
Cre02.g081900	GGAGGCGGAGGTATGCCGCAGCGCGCATGG	GGTTCCACCACCACCCCGCAGCAGCTCCTGTGC
Cre02.g081950	GGAGGCGGAGGTATGGGTATCCTTTCATGCCTG	GGTTCCACCACCACCGGTGGTGTTGAGCAGCTG
Cre02.g082050	GGAGGCGGAGGTATGTCGAGTTTGGCTGCTG	GGTTCCACCACCACCTCACATCGGCGCGCAC
Cre02.g082150	GGAGGCGGAGGTATGTCGACTGGATGTGGCAG	GGTTCCACCACCACCTCACCGGGAGGGGGGC
Cre02.g082250	GGAGGCGGAGGTATGGCCATCCGCCAGC	GGTTCCACCACCACCCACAAACTTGGCGCCCT
Cre02.g082350	GGAGGCGGAGGTATGTCCTCCCGCATTGCG	GGTTCCACCACCACCCTACTGCGCCTGCTGC
Cre02.g082450	GGAGGCGGAGGTATGAGCACCTACAAGGGGTTC	GGTTCCACCACCACCGAACAGGCGCAGCTCCAG
Cre02.g082550	GGAGGCGGAGGTATGCTCGCGAGCACTTACA	GGTTCCACCACCACCAGCCGTCACCAGGCGG
Cre02.g082600	GGAGGCGGAGGTATGAGGTCGCGGTCCC	GGTTCCACCACCACCTCAATCGTCGCGGCCG
Cre02.g082700	GGAGGCGGAGGTATGTCGGGCAGCCAGGG	GGTTCCACCACCACCTTACTCCGCCGCCGCC
Cre02.g082825	GGAGGCGGAGGTATGGCCGGCAAGGGGC	GGTTCCACCACCACCATGCTTCTTGGCGGGCG
Cre02.g082900	GGAGGCGGAGGTATGGATGCTGACGTGGGAAG	GGTTCCACCACCACCCACCAACCACAGCCAGCC
Cre02.g083000	GGAGGCGGAGGTATGAGCGGCCACGAGCC	GGTTCCACCACCACCGCCCGACGTCACCGCC
Cre02.g083065	GGAGGCGGAGGTATGCACAAAGTGGGAGTGAGC	GGTTCCACCACCACCTCAGCCCACCACGCGC
Cre02.g083259	GGAGGCGGAGGTATGCCCCAAACTCAGAGCT	GGTTCCACCACCACCCCTGGCGAATGCACGAA
Cre02.g083350	GGAGGCGGAGGTATGTCTCTGCGCACTTACAGC	GGTTCCACCACCACCACGTTTCTTCATCCTGTTGGTG
Cre02.g083400	GGAGGCGGAGGTATGGCGGCTACTACGGC	GGTTCCACCACCACCCCGCCCTGTCAGCCAC
Cre02.g083500	GGAGGCGGAGGTATGGGCGCCACGGCTGAG	GGTTCCACCACCACCGGACGAGGTGGCCGAGGACTTG
Cre02.g083600	GGAGGCGGAGGTATGGCTCCCAAGAAGGAGAAG	GGTTCCACCACCACCCTTCTTCGCGGACGCCGT
Cre02.g083800	GGAGGCGGAGGTATGGCGACCAGCAACGGCAAC	GGTTCCACCACCACCGTTGGCCAGGAATGGCCC
Cre02.g083935	GGAGGCGGAGGTATGCCCAAAGTCCGCGA	GGTTCCACCACCACCCAAGCACCGCCTCCATG
Cre02.g084000	GGAGGCGGAGGTATGCTGGCTACTCCCCG	GGTTCCACCACCACCCTTGACGCTGCGAGCG
Cre02.g084100	GGAGGCGGAGGTATGGCTCCGTCAGGGCG	GGTTCCACCACCACCCTAGCGCATGCGGCCG
Cre02.g084250	GGAGGCGGAGGTATGCTAGGTCTTTGGCCTCA	GGTTCCACCACCACCTCATGAGGCCCCGCCC
Cre02.g084350	GGAGGCGGAGGTATGGCGCTGGCGGCATCG	GGTTCCACCACCACCCACGAACAGGTCCACGATGG
Cre02.g084450	GGAGGCGGAGGTATGCGTCTTCTATCAACGCTCC	GGTTCCACCACCACCGAGCTCCGGCGACATGTC
Cre02.g084550	GGAGGCGGAGGTATGCTAGCGCGCAGCG	GGTTCCACCACCACCTCACGCAAGTTCCTTTGCA
Cre02.g084700	GGAGGCGGAGGTATGCAGCAGCGCCCGC	GGTTCCACCACCACCCTGGTTCTTGCGCGCCG
Cre02.g084800	GGAGGCGGAGGTATGTCGAAGAGGGGTGGTG	GGTTCCACCACCACCGATCCAGTCTCCCTCCTGCAG
Cre02.g084873	GGAGGCGGAGGTATGAACCACCGGCAGTTTTTA	GGTTCCACCACCACCTCAGCCTGCGCGCTTG
Cre02.g084950	GGAGGCGGAGGTATGGCACATCGCACGCTT	GGTTCCACCACCACCAGGCGCCATCTGGTTCTG
Cre02.g085000	GGAGGCGGAGGTATGGCGGAAGAGTTGCTCA	GGTTCCACCACCACCCGACCTCCCGCGGATC
Cre02.g085050	GGAGGCGGAGGTATGTCAGTCACCACTCGTGT	GGTTCCACCACCACCCAGCTGCAGCCCTAGCA
Cre02.g085150	GGAGGCGGAGGTATGGCCCGGCTCGAGGCG	GGTTCCACCACCACCGCGGGCATGCGCAGATCC
Cre02.g085257	GGAGGCGGAGGTATGGATTATGGAGGCGCCCT	GGTTCCACCACCACCCTCGCCAAGCTCGCGCTTC
Cre02.g085300	GGAGGCGGAGGTATGTCCGTCAAGGGCTCC	GGTTCCACCACCACCCACCGTGAAGTGAGTGATGC
Cre02.g085350	GGAGGCGGAGGTATGCGCGAGCTGAAAACG	GGTTCCACCACCACCCTGCACGGGGTGTACGC
Cre02.g085450	GGAGGCGGAGGTATGGCACTGCAAGCCTCAA	GGTTCCACCACCACCCACCCAGACGCGGGGGTT
Cre02.g085550	GGAGGCGGAGGTATGCAGCAGCCTCGGC	GGTTCCACCACCACCCGAGGCCCGCTGCTTC
Cre02.g085650	GGAGGCGGAGGTATGCGGTTCATCGACATCGG	GGTTCCACCACCACCCGCCGTAGGAAAGAACATTCG

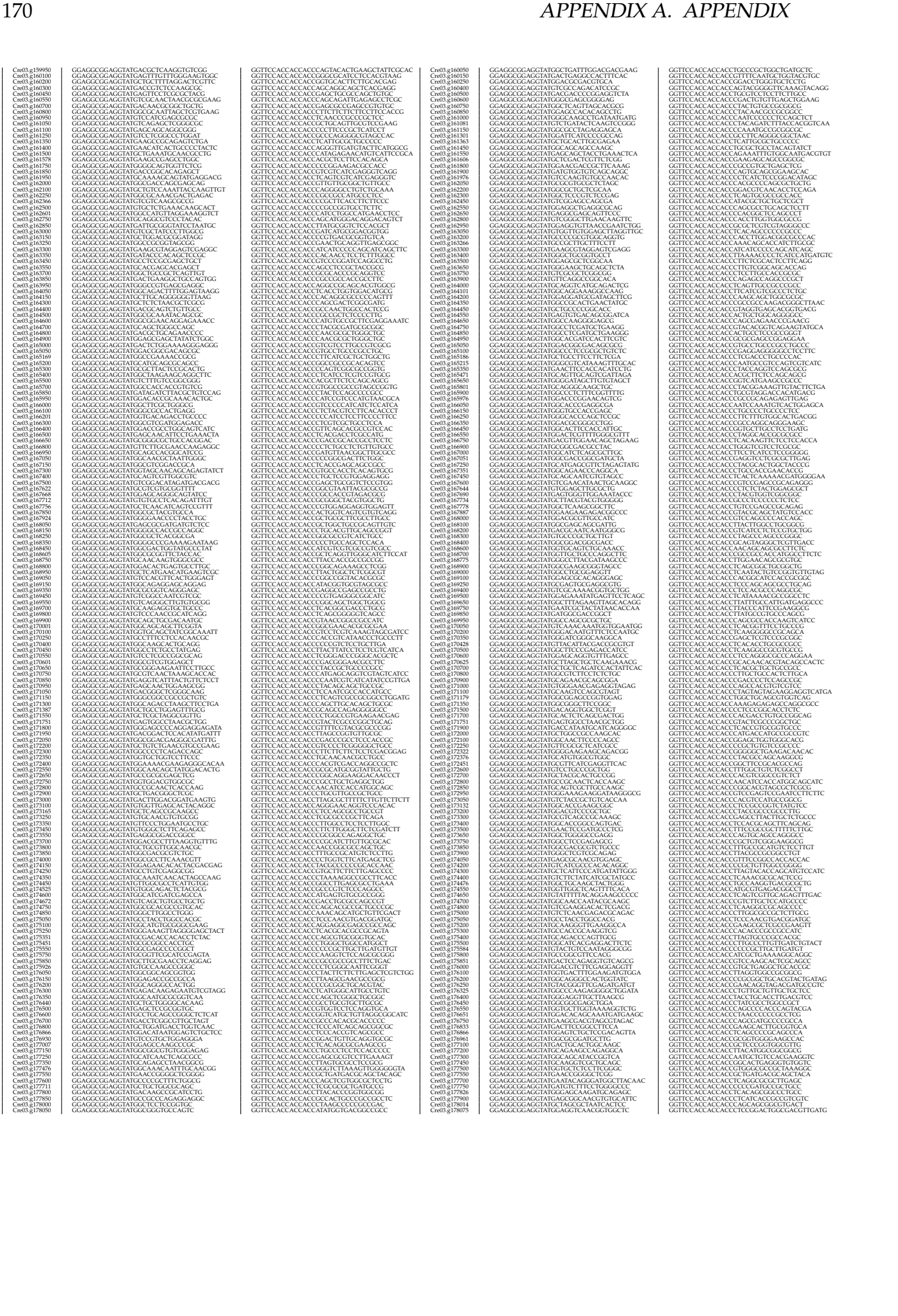
Cre01.g053750	GGAGGCGGAGGTATGGCAGTCGGCGCAG	GGTTCCACCACCACCCTGAAGGAACTCCCGCAGG
Cre01.g053850	GGAGGCGGAGGTATGAATCCGCTGCCCTGT	GGTTCCACCACCACCTTACAAGCCCCGCGCC
Cre01.g053950	GGAGGCGGAGGTATGTTCGAGGCTGCAACC	GGTTCCACCACCACCCGCCAGGTCGACGCTG
Cre01.g054050	GGAGGCGGAGGTATGCAGAACGTGTTGTTGCA	GGTTCCACCACCACCGAACAGCGTGACCAGTATCC
Cre01.g054150	GGAGGCGGAGGTATGATACAGCTACCTGCAACAGTT	GGTTCCACCACCACCTCACGCGGACACGGCG
Cre01.g054250	GGAGGCGGAGGTATGCTGCGTGCTTTGAGCT	GGTTCCACCACCACCCTCCGCCCTGGCAGCC
Cre01.g054350	GGAGGCGGAGGTATGCCGCTTGCTGAGGG	GGTTCCACCACCACCCAGCTTGGCCAGCTTGG
Cre01.g054433	GGAGGCGGAGGTATGAACGTGAAACGCCGC	GGTTCCACCACCACCGCCTTCAGGGCGGCTA
Cre01.g054550	GGAGGCGGAGGTATGCATGGAAAACAAGATACGGG	GGTTCCACCACCACCCTACTGGCTCTCCTCGTGC
Cre01.g054650	GGAGGCGGAGGTATGGACGCGGTGGAACT	GGTTCCACCACCACCCTCGGCGTCCAGCCAC
Cre01.g054750	GGAGGCGGAGGTATGAAGGCCCTTTGCATCGC	GGTTCCACCACCACCCGGCTCCGTCATCCACCG
Cre01.g054850	GGAGGCGGAGGTATGGCTCAGCTGCAGTCC	GGTTCCACCACCACCCTGCGCGTCGTCCACG
Cre01.g054950	GGAGGCGGAGGTATGGAGTTTGTGCACAAGCT	GGTTCCACCACCACCCTGCCGCTCCTTGCTTG
Cre01.g055050	GGAGGCGGAGGTATGAAGAAGGGGGAGCGCAAG	GGTTCCACCACCACCCTCCCCCGCCTTCTCGCGAATC
Cre01.g055151	GGAGGCGGAGGTATGCCACGCCACATCATC	GGTTCCACCACCACCGCTGAGCGCCAGGAAG
Cre01.g055250	GGAGGCGGAGGTATGCTGCTTCAGTACTCGC	GGTTCCACCACCACCGTTGCCCCCCTCCTCC
Cre01.g055316	GGAGGCGGAGGTATGCGGGCTTCATCAGTGC	GGTTCCACCACCACCCTGCTGCCTCCCCGCC
Cre01.g055404	GGAGGCGGAGGTATGCGTGCAGGCAGTCG	GGTTCCACCACCACCCTGACGACCGCCACGC
Cre01.g055412	GGAGGCGGAGGTATGAAGACTCTTAAATGCGGCC	GGTTCCACCACCACCCTACACATCCGCGCAGC
Cre01.g055420	GGAGGCGGAGGTATGGCGTCCTCTTCAACCTC	GGTTCCACCACCACCCGCGCAGTACATGTAGAGCG
Cre01.g055428	GGAGGCGGAGGTATGGATCAGGACGGGTCTGG	GGTTCCACCACCACCTCAGCTGCGCGCCACG
Cre01.g055432	GGAGGCGGAGGTATGAAGCAGGCATGCCTG	GGTTCCACCACCACCGCCCTCAAAGATGACGCTC
Cre01.g055440	GGAGGCGGAGGTATGAAGAACCCCTACCTGGC	GGTTCCACCACCACCCCCCCTAAGCTTGCCGC
Cre01.g055448	GGAGGCGGAGGTATGTACGGCCACACCGTCA	GGTTCCACCACCACCCTAGCGGCGGCAGCCC
Cre01.g055461	GGAGGCGGAGGTATGGTGTACGGCCTGCG	GGTTCCACCACCACCGGCCAGTTCACTTGAGTCC
Cre01.g055469	GGAGGCGGAGGTATGGGGGACCATGCGG	GGTTCCACCACCACCCAGCCGTGTTCCGTGC
Cre01.g056331	GGAGGCGGAGGTATGGCTTCTGCTGGCGT	GGTTCCACCACCACCCTACACAAAGTACTGGAACATGAC
Cre01.g057061	GGAGGCGGAGGTATGAGTATTGAAGCGGCAGC	GGTTCCACCACCACCCTCCTCATCGCTGCTGTC
Cre01.g057791	GGAGGCGGAGGTATGGGACCGCCTAAACTGC	GGTTCCACCACCACCCAGTCCCTGCTGCAGCT
Cre01.g058521	GGAGGCGGAGGTATGTATACGGACGCGTTAGATTC	GGTTCCACCACCACCCTAGGGCGGCAGCACA
Cre01.g059617	GGAGGCGGAGGTATGGCGCCTGCCTTGG	GGTTCCACCACCACCTCACTGCTGCCGGCCC
Cre01.g060347	GGAGGCGGAGGTATGCTGACCAGCCGAGGC	GGTTCCACCACCACCTCACTGCTGCCGGCCC
Cre01.g061442	GGAGGCGGAGGTATGCCGAAGCAAACGGTGAAG	GGTTCCACCACCACCCGTGTCCAGGAAGTTTGCG
Cre01.g062172	GGAGGCGGAGGTATGGCTCCCAAGGCTGC	GGTTCCACCACCACCGGTGCTGGTGAACTTGGTC
Cre01.g063267	GGAGGCGGAGGTATGCCTGGCAAAAAACGAATTC	GGTTCCACCACCACCGTCCTCCATCATGAGCACCG
Cre01.g063997	GGAGGCGGAGGTATGAGCACATTCGCTGCGAAG	GGTTCCACCACCACCCTCCATGTCGCGGGGCCC
Cre01.g065092	GGAGGCGGAGGTATGGCCAAGCAGCGCGAA	GGTTCCACCACCACCCATCTCCTCCAGGTCCTCAGG
Cre01.g065822	GGAGGCGGAGGTATGCGGACCGTCGTTGCTT	GGTTCCACCACCACCCACCAGCGGCCGGCCCTG
Cre01.g066552	GGAGGCGGAGGTATGGCAACTCCTTGCGTCC	GGTTCCACCACCACCGAGGTGCTTGCCGATCAGG
Cre01.g067647	GGAGGCGGAGGTATGGAGGTGGAGACGGC	GGTTCCACCACCACCGAAAGGCCCCGCCCAG
Cre01.g068377	GGAGGCGGAGGTATGGCCGAGCAGGGCG	GGTTCCACCACCACCGCAAGAAGTCACGCCGC
Cre01.g069107	GGAGGCGGAGGTATGTCGTGTAAGGGTCTGCTG	GGTTCCACCACCACCATAGCCCTTGTTGCCCTGGATC
Cre01.g069837	GGAGGCGGAGGTATGGGCACGCTCGTGGGC	GGTTCCACCACCACCTGCCTGCCACTGCGGTTT
Cre01.g070202	GGAGGCGGAGGTATGGCTGGGTGCGCAA	GGTTCCACCACCACCAGCGCACTCCAGTGGC
Cre01.g071297	GGAGGCGGAGGTATGCTGCTGGGGCACAAG	GGTTCCACCACCACCGCCTGTGCGCTTGCGC
Cre01.g072027	GGAGGCGGAGGTATGAACGTGAAACGCCGC	GGTTCCACCACCACCCTCCGCCGCCGAGACC
Cre02.g073200	GGAGGCGGAGGTATGGATAGGTCGATGCTCGGAA	GGTTCCACCACCACCCTGGATGAACTGATCAAACACCT
Cre02.g073300	GGAGGCGGAGGTATGTCGTTATCAGGGCGGC	GGTTCCACCACCACCACCACTCGCCTCACCTCC
Cre02.g073400	GGAGGCGGAGGTATGTACCAGCGGTTGTCGG	GGTTCCACCACCACCGGCCGCAATGCCGGTC
Cre02.g073550	GGAGGCGGAGGTATGGCACTAGCTCTGCTCACTC	GGTTCCACCACCACCGTACTCGCCCTCCTCGTACTC
Cre02.g073650	GGAGGCGGAGGTATGTTTTGCTCTATCAGCGGAAC	GGTTCCACCACCACCAGCGGAGGGGCTGGAGAA
Cre02.g073750	GGAGGCGGAGGTATGGTGAAGCAAGCTGTGAAGG	GGTTCCACCACCACCTCACCGCGCCAGGCCT
Cre02.g073826	GGAGGCGGAGGTATGGCCAGGCTCCTGG	GGTTCCACCACCACCTCAGCCTGCGCGCTTG
Cre02.g073900	GGAGGCGGAGGTATGCAAGCACTACATTATCACTTGT	GGTTCCACCACCACCTTTCTGCTGTACGAATGTGCCG
Cre02.g074050	GGAGGCGGAGGTATGCCATGGGCCAGCTGTT	GGTTCCACCACCACCCACCGCCTTGTGTTTCGC
Cre02.g074150	GGAGGCGGAGGTATGGCACCCACCGGCAGC	GGTTCCACCACCACCCATGAACAGCGCCTGCGC
Cre02.g074217	GGAGGCGGAGGTATGACCCTGGTGCCATCA	GGTTCCACCACCACCCTCCTCGCTGTCGCGG
Cre02.g074400	GGAGGCGGAGGTATGGACCTGGGCACATGG	GGTTCCACCACCACCGGCCCCCAGCGGACTTA
Cre02.g074474	GGAGGCGGAGGTATGCTCGGCAGCAGCAG	GGTTCCACCACCACCTCAACCTGCGCGCCGG
Cre02.g074550	GGAGGCGGAGGTATGTCCACACCGCCTAAGC	GGTTCCACCACCACCCTCTTGCTTAGCGTGCGC
Cre02.g074650	GGAGGCGGAGGTATGTCCTGCCCACTCTCATA	GGTTCCACCACCACCTCATGAGCTGCCGCTGC
Cre02.g074720	GGAGGCGGAGGTATGGACCAGTGTTTCTCACTG	GGTTCCACCACCACCCCTGGTATTGGACACGACTAC
Cre02.g074758	GGAGGCGGAGGTATGCAGCAGCGGCAAACC	GGTTCCACCACCACCGCCGTGCGTATCAAACAGTG
Cre02.g074800	GGAGGCGGAGGTATGGGCAATCGCATAAGCA	GGTTCCACCACCACCCAGCAGGTCAGCGGACG
Cre02.g074950	GGAGGCGGAGGTATGAGATGGTTCCAACGCAA	GGTTCCACCACCACCGCCCGACATGCAGCTG
Cre02.g075050	GGAGGCGGAGGTATGGCGACTCCGCCGC	GGTTCCACCACCACCCTTTGCGGGCTGCTCAATG
Cre02.g075150	GGAGGCGGAGGTATGTACGGAGCTGATGAGGACC	GGTTCCACCACCACCGTGCACGGGGATTGCCGC
Cre02.g075301	GGAGGCGGAGGTATGGAAGGGCCGCACGC	GGTTCCACCACCACCCCTGCAGCCCCGCACC
Cre02.g075400	GGAGGCGGAGGTATGTCGGCCGACTGGAA	GGTTCCACCACCACCCTCCGCCTTGGCCTCG
Cre02.g075550	GGAGGCGGAGGTATGTCCGATGGCGTAGATGAC	GGTTCCACCACCACCGAAGGTCAGCTCCGCAGG
Cre02.g075676	GGAGGCGGAGGTATGTCAGGGATTTCGGGACC	GGTTCCACCACCACCAGCATTAGCTACAGTGCCG
Cre02.g075750	GGAGGCGGAGGTATGGCGCGACTGGTAGC	GGTTCCACCACCACCCGACTGCAGCTGCCGC
Cre02.g075850	GGAGGCGGAGGTATGGTCAAGCGTGAAGTGCGAA	GGTTCCACCACCACCGTCGTTCATGCCGAACTGGG
Cre02.g075994	GGAGGCGGAGGTATGAACACGCCGGGTCC	GGTTCCACCACCACCACGCCGCTTGGCAGCT
Cre02.g076050	GGAGGCGGAGGTATGAGCCACCGGCTGG	GGTTCCACCACCACCCTTCTTGGCCGGTGTGG
Cre02.g076200	GGAGGCGGAGGTATGTCATTTAACACGGAGTTGGT	GGTTCCACCACCACCCGTCACGCGCTGCTTGAG
Cre02.g076350	GGAGGCGGAGGTATGGCCGTTGACGCGC	GGTTCCACCACCACCGTGGTCCTCCTTGCGCT
Cre02.g076433	GGAGGCGGAGGTATGGTGTTCAAAGTAAGCCTGA	GGTTCCACCACCACCTCATGCCCTCCCAAAGCG
Cre02.g076500	GGAGGCGGAGGTATGGAGGAGGGGGATAAGCC	GGTTCCACCACCACCCCCGCTCTCGGGGTCC
Cre02.g076600	GGAGGCGGAGGTATGCCCAAACGCACAGCG	GGTTCCACCACCACCGATACGCTGTCCCGAGGC
Cre02.g076650	GGAGGCGGAGGTATGGCCCTGCGCCGTC	GGTTCCACCACCACCGCGGTACAAAGCCTGCTTG
Cre02.g076800	GGAGGCGGAGGTATGGGGAAAGGAGGAACTCAG	GGTTCCACCACCACCGTACAGGTACGGCACGATGC
Cre02.g076900	GGAGGCGGAGGTATGGGGTGTGCACAGAGTA	GGTTCCACCACCACCCTTGCTGTCGTCGATCCAG
Cre02.g076950	GGAGGCGGAGGTATGGGCAGGGCGTGGG	GGTTCCACCACCACCGATGCGCGTCAGGCGG
Cre02.g077024	GGAGGCGGAGGTATGGACAGCGGCCTCG	GGTTCCACCACCACCCAGCCCCTGCCACTCTA
Cre02.g077100	GGAGGCGGAGGTATGGCTCTCGCCTCAGGC	GGTTCCACCACCACCGTACATGAACTCCTTGTACAGCG
Cre02.g077200	GGAGGCGGAGGTATGAAGGAGCTTCTCAGCATGG	GGTTCCACCACCACCTTTCCCCTTGGCTGCTCG
Cre02.g077300	GGAGGCGGAGGTATGGGCTTCGGTGACCGC	GGTTCCACCACCACCCTCCTTCTTCTTGCCGGCG
Cre02.g077401	GGAGGCGGAGGTATGGCTGAAGACAAAGCCAGG	GGTTCCACCACCACCTGGCTGCGGTGCAGCC
Cre02.g077500	GGAGGCGGAGGTATGTCCGTGTACTCGCTCTAC	GGTTCCACCACCACCCATGGCCTGCTGTAAGTATCTC
Cre02.g077600	GGAGGCGGAGGTATGGCCCAGCAGCTGACGATG	GGTTCCACCACCACCTTTCCCCCGGTTGAATCAAAT
Cre02.g077650	GGAGGCGGAGGTATGTCGTGCAAGGCGGAC	GGTTCCACCACCACCAGCGCTAGCCCCCAGGAATC
Cre02.g077800	GGAGGCGGAGGTATGCGCTTTAACTGCGCG	GGTTCCACCACCACCCAGCTTCTTGCGGCCG
Cre02.g077900	GGAGGCGGAGGTATGGAGTCCGCCAAGGAG	GGTTCCACCACCACCGCTCTGTTCGTCCCAGGC
Cre02.g078000	GGAGGCGGAGGTATGCTGCCAAGAGGGTGGC	GGTTCCACCACCACCTCACGTGCGGCGCAGG
Cre02.g078100	GGAGGCGGAGGTATGGGAAAGAAGAAGGCAGCA	GGTTCCACCACCACCGGTGACGTTGAAGGTCTTCTTG
Cre02.g078200	GGAGGCGGAGGTATGGCGGAGCTCATGAACC	GGTTCCACCACCACCTCAGCCTGCGCGCTTG
Cre02.g078251	GGAGGCGGAGGTATGGCACAGGCGGCAC	GGTTCCACCACCACCGTTCTCACTGCCCCCCG
Cre02.g078316	GGAGGCGGAGGTATGGCCGGGCGACAGC	GGTTCCACCACCACCGCCCCACACCACACGG
Cre02.g078400	GGAGGCGGAGGTATGGAGGCCGAGGGTAC	GGTTCCACCACCACCATAAAGGGGCGGCGACT
Cre02.g078476	GGAGGCGGAGGTATGTTCTTCCTGGGCGCA	GGTTCCACCACCACCTCAATGGCTATGCTTTAAGCAAG
Cre02.g078507	GGAGGCGGAGGTATGGCTCTCGCCACTCG	GGTTCCACCACCACCTTACTTGCCAATCTCAAGGAAGT
Cre02.g078600	GGAGGCGGAGGTATGCAGAGCCACTCTAGTAGA	GGTTCCACCACCACCCTACTTGGCCTGCGGC
Cre02.g078700	GGAGGCGGAGGTATGGCCGCGAGCAACA	GGTTCCACCACCACCCTCGTGCTTGGCACCC
Cre02.g078804	GGAGGCGGAGGTATGGAATTGACTGAATTATCACGC	GGTTCCACCACCACCTTAGGGAGCAGCCTCCTG
Cre02.g078858	GGAGGCGGAGGTATGGCCACCCCATCCTT	GGTTCCACCACCACCGAAGAATGCCAGGTTGGTGG
Cre02.g078912	GGAGGCGGAGGTATGTCAGAAATGGAGTCTCTTGG	GGTTCCACCACCACCCTACGGCTGCTGCTGC
Cre02.g078950	GGAGGCGGAGGTATGCGAGATGGCCCCGCTTC	GGTTCCACCACCACCGGCGCAGTCGGCACACGAG
Cre02.g079003	GGAGGCGGAGGTATGGTGGCCAGACACCA	GGTTCCACCACCACCTTACACGCCACGCCACG
Cre02.g079005	GGAGGCGGAGGTATGGCGGGCAGCCCTA	GGTTCCACCACCACCGTCGTACATCTGGGCGTAC
Cre02.g079100	GGAGGCGGAGGTATGAAGCTGTACCTCAGAAGC	GGTTCCACCACCACCAAACTCCTCCACGTCTGCC
Cre02.g079200	GGAGGCGGAGGTATGAGCGGGGACGAGGGC	GGTTCCACCACCACCCTTGGGGGGCTCCTGCTT
Cre02.g079300	GGAGGCGGAGGTATGGCCGTATATATTGGATTCAA	GGTTCCACCACCACCTCACGACGCCTCCTCG
Cre02.g079450	GGAGGCGGAGGTATGCTTTACAAAGGCCTTCTAGC	GGTTCCACCACCACCCTATTTAGCTGCCCCCATCTC
Cre02.g079550	GGAGGCGGAGGTATGGCAGCTCTAGGGACGC	GGTTCCACCACCACCCTACCGCGTGCCCGGT
Cre02.g079650	GGAGGCGGAGGTATGGCCACCATTGCCACG	GGTTCCACCACCACCATGTGGGTGTGGGTGTGC
Cre02.g079800	GGAGGCGGAGGTATGCGCCTTGCTGCTCAG	GGTTCCACCACCACCCGGCTTCAGCGCCACAGC
Cre02.g079850	GGAGGCGGAGGTATGGCGGGGGGCTACA	GGTTCCACCACCACCGGTTAGGGTCCGGGATGA
Cre02.g080050	GGAGGCGGAGGTATGGGCGACAAGAACGC	GGTTCCACCACCACCCGCGACAAAGTCAGGGTC
Cre02.g080100	GGAGGCGGAGGTATGGCGTCGGCAGTAGA	GGTTCCACCACCACCCGTGGTCTGCGCATCAC
Cre02.g080200	GGAGGCGGAGGTATGCAGACCATGCTGAAGCAG	GGTTCCACCACCACCGTGCTGCAGGGTGGCCTT
Cre02.g080300	GGAGGCGGAGGTATGCTCGGCGAGGGCGCTG	GGTTCCACCACCACCTGGGCCCGCCCGGCGTGG
Cre02.g080400	GGAGGCGGAGGTATGAGGAGCCTGTTGGGT	GGTTCCACCACCACCTCAGGCCGGCACCATG
Cre02.g080600	GGAGGCGGAGGTATGGCCAATAGCTTGAAGACTC	GGTTCCACCACCACCTCAAAGCTCGTCGTGGTCG
Cre02.g080700	GGAGGCGGAGGTATGGCGCAGTGGAAGGCT	GGTTCCACCACCACCGAGCTCGTCGTGGTCGCC
Cre02.g080800	GGAGGCGGAGGTATGGCATCAATGATGACGCATTC	GGTTCCACCACCACCCTACCGCCCCGTGCCG
Cre02.g080900	GGAGGCGGAGGTATGCAGCAGTCTGCGTCG	GGTTCCACCACCACCGTTGCAGGAGTGCGGGTAGAT
Cre02.g081050	GGAGGCGGAGGTATGGCGTCTCGCGGCC	GGTTCCACCACCACCCAGGCCCCAGCCCAGC
Cre02.g081150	GGAGGCGGAGGTATGCGCGAGGAGGGCA	GGTTCCACCACCACCAGACGCCTGCCAGAACA
Cre02.g081200	GGAGGCGGAGGTATGCCGTTTGGCCAGGT	GGTTCCACCACCACCCTGAGAACCGCCTTGCC
Cre02.g081300	GGAGGCGGAGGTATGCACGCGCTGCCCCGG	GGTTCCACCACCACCTGCCGAGCCCGACGCCGCGATG
Cre02.g081400	GGAGGCGGAGGTATGAGCCCCTCCATCTGCC	GGTTCCACCACCACCTTCCAGGTTAGGGTCAGCGAAA
Cre02.g081550	GGAGGCGGAGGTATGAGCGGTCAAACTCAACGT	GGTTCCACCACCACCCCGGTGCTCCGCCATGGC
Cre02.g081650	GGAGGCGGAGGTATGCGCTGGCGGGAGA	GGTTCCACCACCACCTCAATAGATCTTATCATACCGCGC
Cre02.g081850	GGAGGCGGAGGTATGACTCGGCCGTCACC	GGTTCCACCACCACCCTGATCTTCCTCGCCTTCATC
Cre02.g081950	GGAGGCGGAGGTATGGGTATCCTTTCATGCCTGT	GGTTCCACCACCACCGGTGGTGTTGAGCAGCTGG
Cre02.g082000	GGAGGCGGAGGTATGCGTGCGGGCGCTAAG	GGTTCCACCACCACCCTTAGCCTTAGCCTTGCCTTTT
Cre02.g082100	GGAGGCGGAGGTATGTCCTCCTCGTCGTCAG	GGTTCCACCACCACCCGACCTGATTGCACGTGT
Cre02.g082200	GGAGGCGGAGGTATGTCGTCTGGATGTGGCA	GGTTCCACCACCACCCACAAACTTGGCGCCCT
Cre02.g082300	GGAGGCGGAGGTATGGCCGTCGACGTCG	GGTTCCACCACCACCCTTGGGTGCCGCATTGA
Cre02.g082400	GGAGGCGGAGGTATGCCACCAGCATCGGA	GGTTCCACCACCACCCTACTGCCGCAGCCGC
Cre02.g082500	GGAGGCGGAGGTATGGCCATCTCTGCTCGC	GGTTCCACCACCACCGTTGGAGCGCAGGGTG
Cre02.g082576	GGAGGCGGAGGTATGTCTGCCCTGGACTTTGG	GGTTCCACCACCACCCGGCGAGGGAGACGGG
Cre02.g082651	GGAGGCGGAGGTATGCCAGGCGAAATATCTGC	GGTTCCACCACCACCCTACCCGCGGCCAAGC
Cre02.g082800	GGAGGCGGAGGTATGTCTCTGCGCACTTACAG	GGTTCCACCACCACCACGTTTCTTCATCCTGTTGGT
Cre02.g082877	GGAGGCGGAGGTATGGTTGCCAAAAACTGCAAGG	GGTTCCACCACCACCACGTTTCTTCATCCTGTTGGTG
Cre02.g082950	GGAGGCGGAGGTATGGATACGCAGGGCGTTTTA	GGTTCCACCACCACCTCGGCGGTGTCTATGCCG
Cre02.g083050	GGAGGCGGAGGTATGGCCTTCCCCTTGCA	GGTTCCACCACCACCCAAGCCACCCACTCCGG
Cre02.g083100	GGAGGCGGAGGTATGCTGACCATTCCAAACCTAT	GGTTCCACCACCACCTTAGCGCGTTTGCTGGAC
Cre02.g083273	GGAGGCGGAGGTATGACAGGAGCTTGCGACG	GGTTCCACCACCACCCACGCCATGACTTACTGCT
Cre02.g083354	GGAGGCGGAGGTATGGCCCTCGCAGCAT	GGTTCCACCACCACCCCCCCGGGTTGAGTCC
Cre02.g083450	GGAGGCGGAGGTATGCGGGGCAGTTTGGT	GGTTCCACCACCACCACGGAACGACCAGCTCC
Cre02.g083550	GGAGGCGGAGGTATGCTCCAACGCAAGTTTCATG	GGTTCCACCACCACCACGCACGAAGTACGTGCC
Cre02.g083650	GGAGGCGGAGGTATGGCTCGCCTTCACATCC	GGTTCCACCACCACCGAAGTTGGAGGGCTGGTAGG
Cre02.g083850	GGAGGCGGAGGTATGCACATCGTTGCGACT	GGTTCCACCACCACCGTAGTGCACCACCTGCTC
Cre02.g083950	GGAGGCGGAGGTATGATGCTTGTGCGGCAG	GGTTCCACCACCACCGGCGATGGCGAAGGCG
Cre02.g084050	GGAGGCGGAGGTATGGGGAGCGAAGACAAGGA	GGTTCCACCACCACCCAAGTCCTGCTTGGCGCTA
Cre02.g084200	GGAGGCGGAGGTATGTCTTCAACTACAGCTGCTC	GGTTCCACCACCACCCTACCAGCGCATGTTGCC
Cre02.g084300	GGAGGCGGAGGTATGAGGTGCCTAGAGTCTTCAC	GGTTCCACCACCACCTCGGCCCTTGGTTGCCAC
Cre02.g084400	GGAGGCGGAGGTATGCGTTTCGTCGCCCTC	GGTTCCACCACCACCAAGCTCCGGCGACATGTC
Cre02.g084500	GGAGGCGGAGGTATGTACTTCTCCGAGCTTGATC	GGTTCCACCACCACCTCATGTCGCGGCGGTT
Cre02.g084650	GGAGGCGGAGGTATGTTGATGGGCCTGGGGG	GGTTCCACCACCACCCTCGGGCGTGGAGCCG
Cre02.g084750	GGAGGCGGAGGTATGTCTGAACAGGCGCCTA	GGTTCCACCACCACCTTAGTAGGGCGCCGCG
Cre02.g084850	GGAGGCGGAGGTATGGGCGGGCAGCAAG	GGTTCCACCACCACCCACTCCGCCCAGCAGC
Cre02.g084900	GGAGGCGGAGGTATGGACGTGCAATTCCAAATCC	GGTTCCACCACCACCCGCGCACCGGTACTTATGC
Cre02.g085000	GGAGGCGGAGGTATGGCGGAAGAGTTGCTCATC	GGTTCCACCACCACCCGACCTCCCGCGGATCGC
Cre02.g085050	GGAGGCGGAGGTATGTCAGTCACCACTCGTGTTA	GGTTCCACCACCACCCAGCTGCAGCCCTAGCAG
Cre02.g085100	GGAGGCGGAGGTATGAGCTTCACTGCGCG	GGTTCCACCACCACCCTAGGCCGCCAGACGC
Cre02.g085200	GGAGGCGGAGGTATGGACCTGGCTGGGCCC	GGTTCCACCACCACCGTCGCCGATGGGGTCCAG
Cre02.g085279	GGAGGCGGAGGTATGTATGTTGGAAAACGCGGA	GGTTCCACCACCACCCCCGTTGTCTAGGATGCTG
Cre02.g085326	GGAGGCGGAGGTATGGTGGACAAGGTCCAGA	GGTTCCACCACCACCTTAGCCCGCAAAGCCG
Cre02.g085400	GGAGGCGGAGGTATGGACGCTGCAGTCGT	GGTTCCACCACCACCTGCTACTGCAACACAACGA
Cre02.g085500	GGAGGCGGAGGTATGCTGCTGCTGTCCAAC	GGTTCCACCACCACCCAGCAGGATGCAGCGG
Cre02.g085600	GGAGGCGGAGGTATGGCCGCGGTTACCA	GGTTCCACCACCACCATGACGCTTCCCGTCCG
Cre02.g085701	GGAGGCGGAGGTATGGAGACCGCGCGCGAAG	GGTTCCACCACCACCCACCACCATCTGCCATCGC

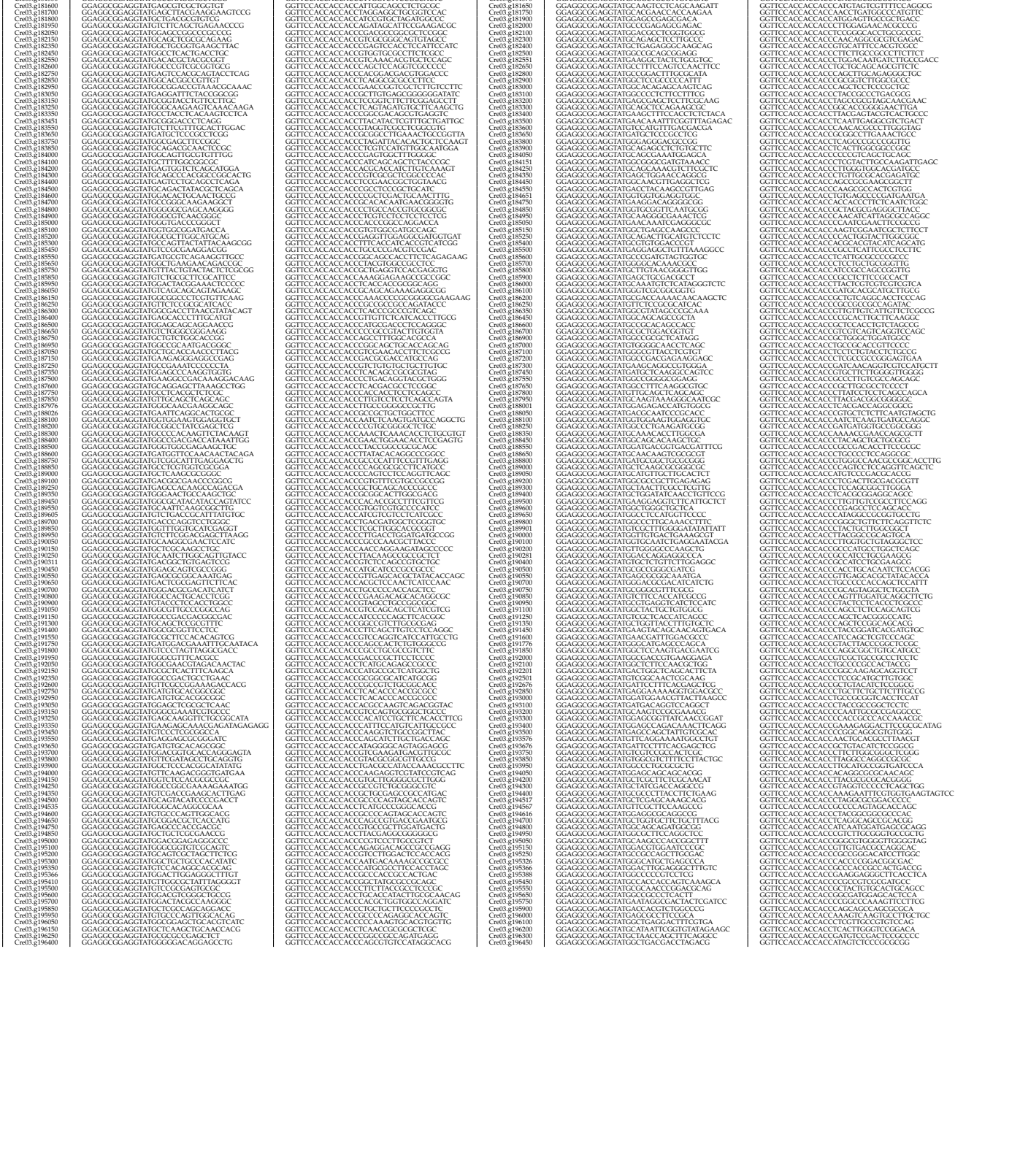


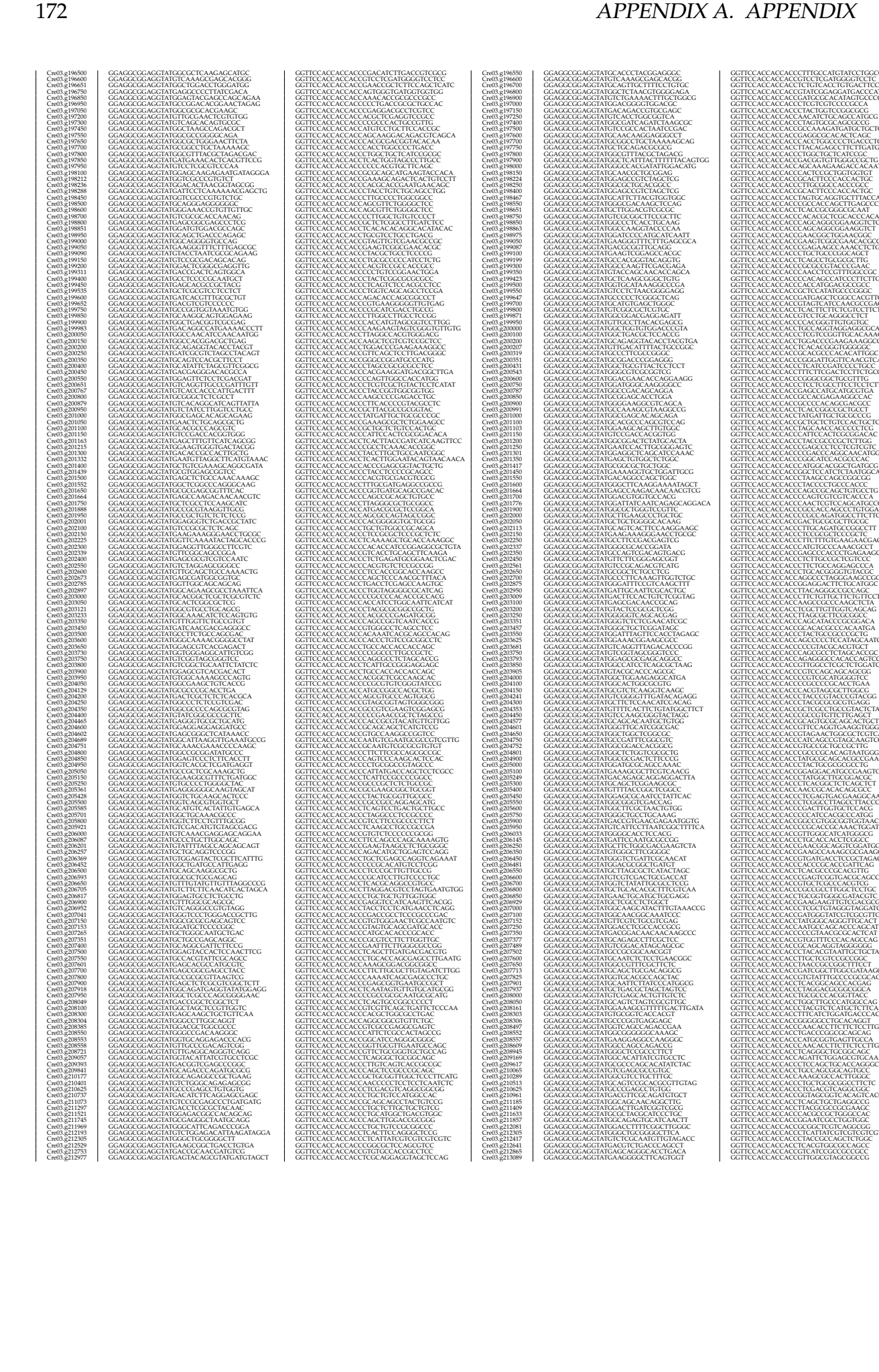


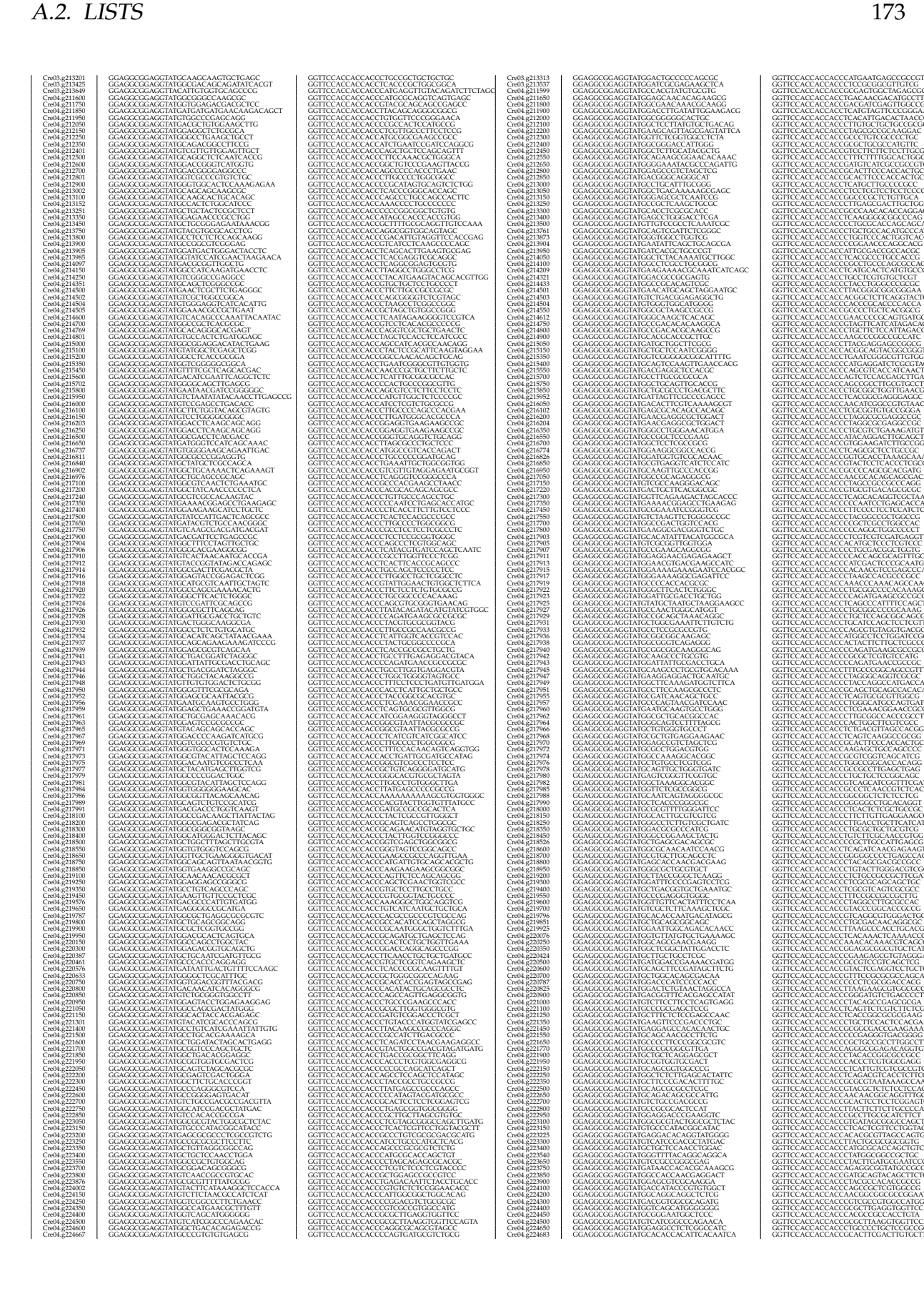


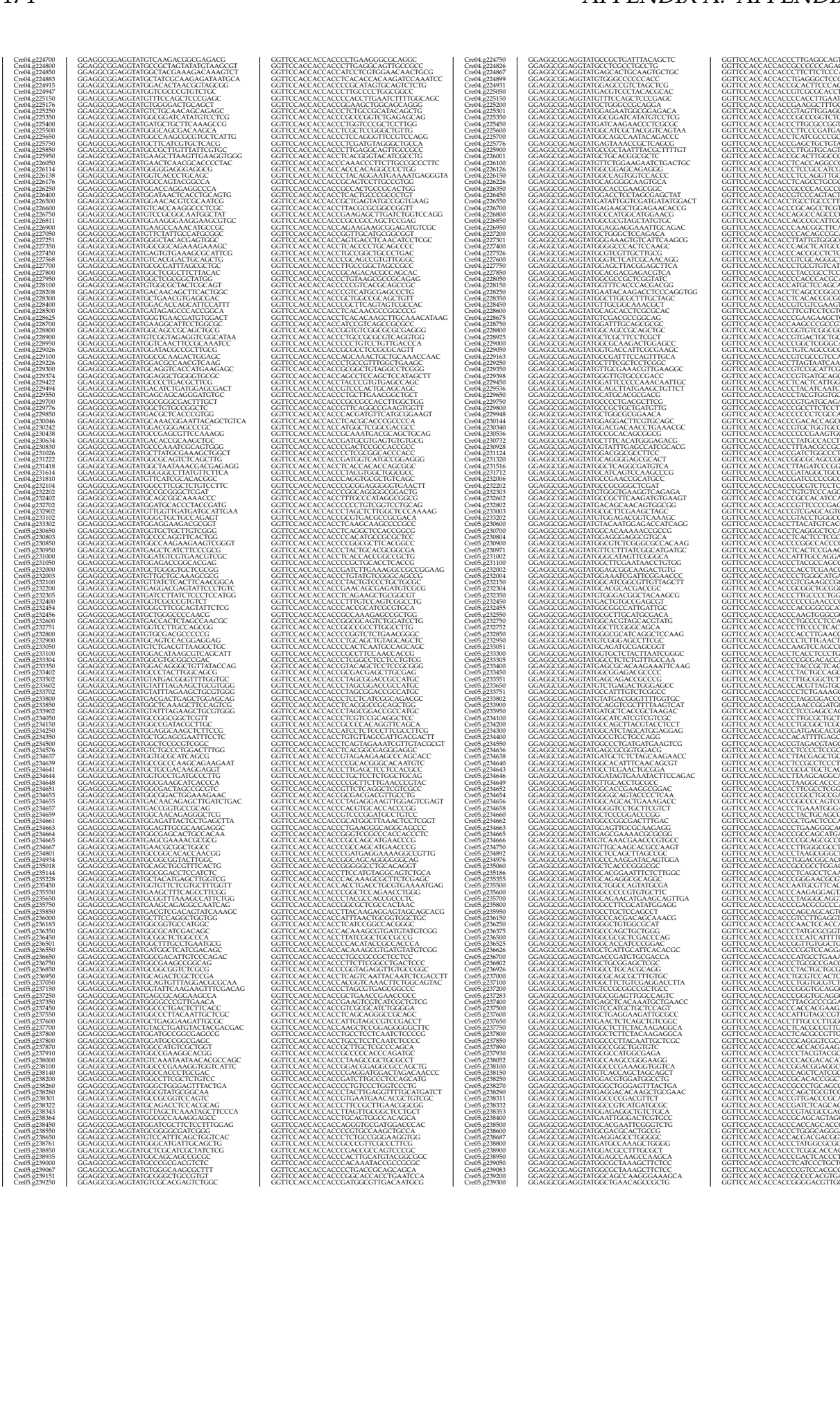




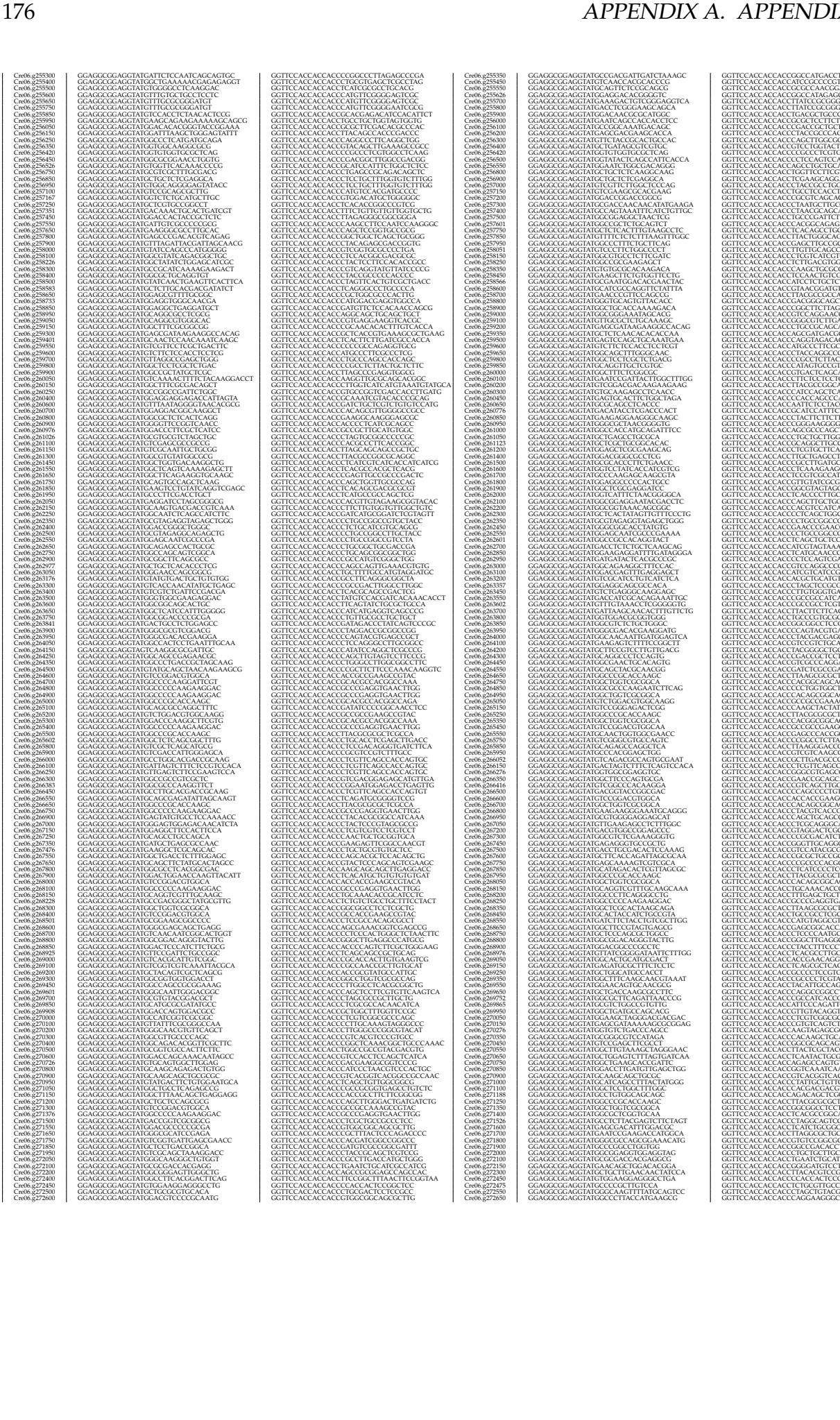


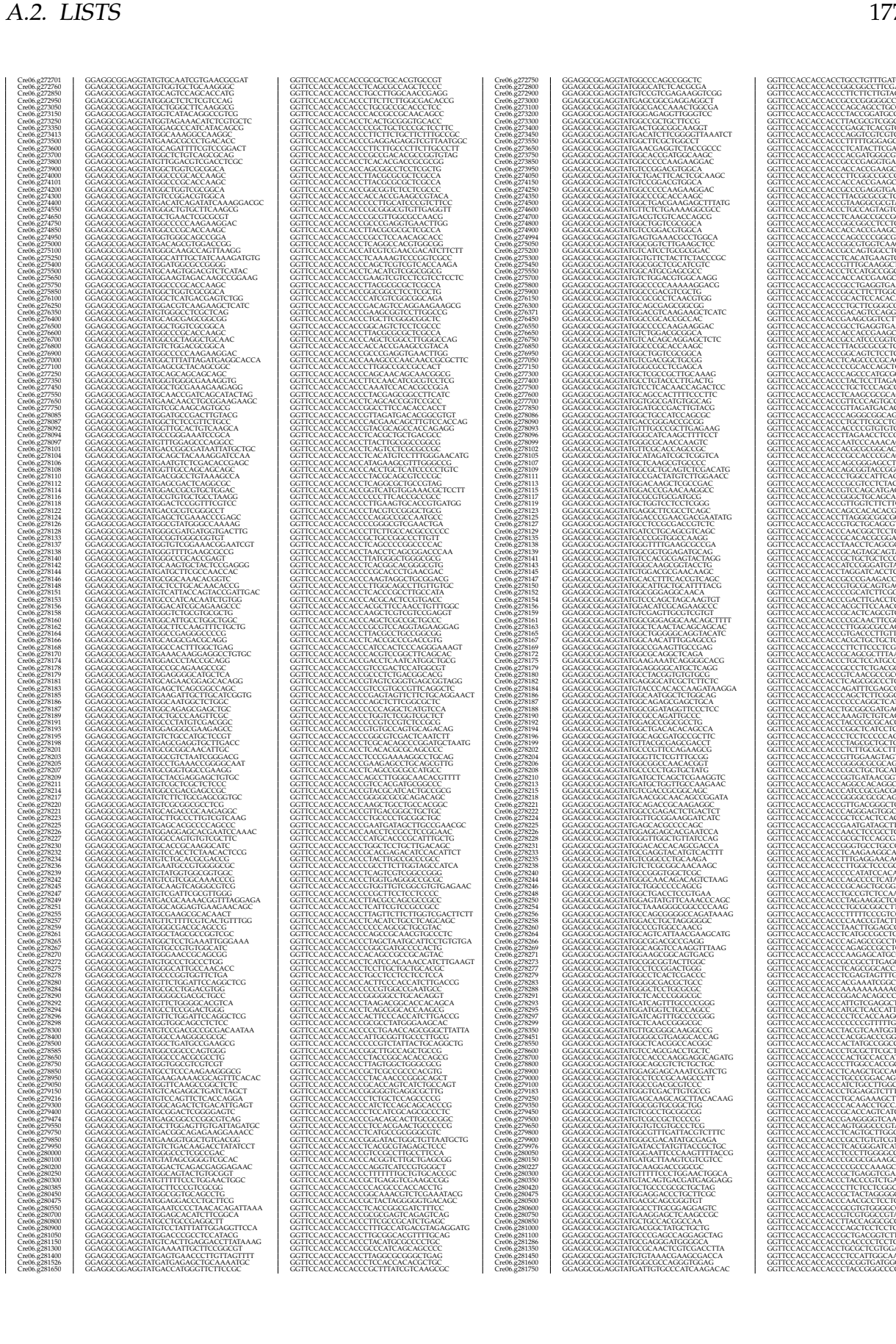


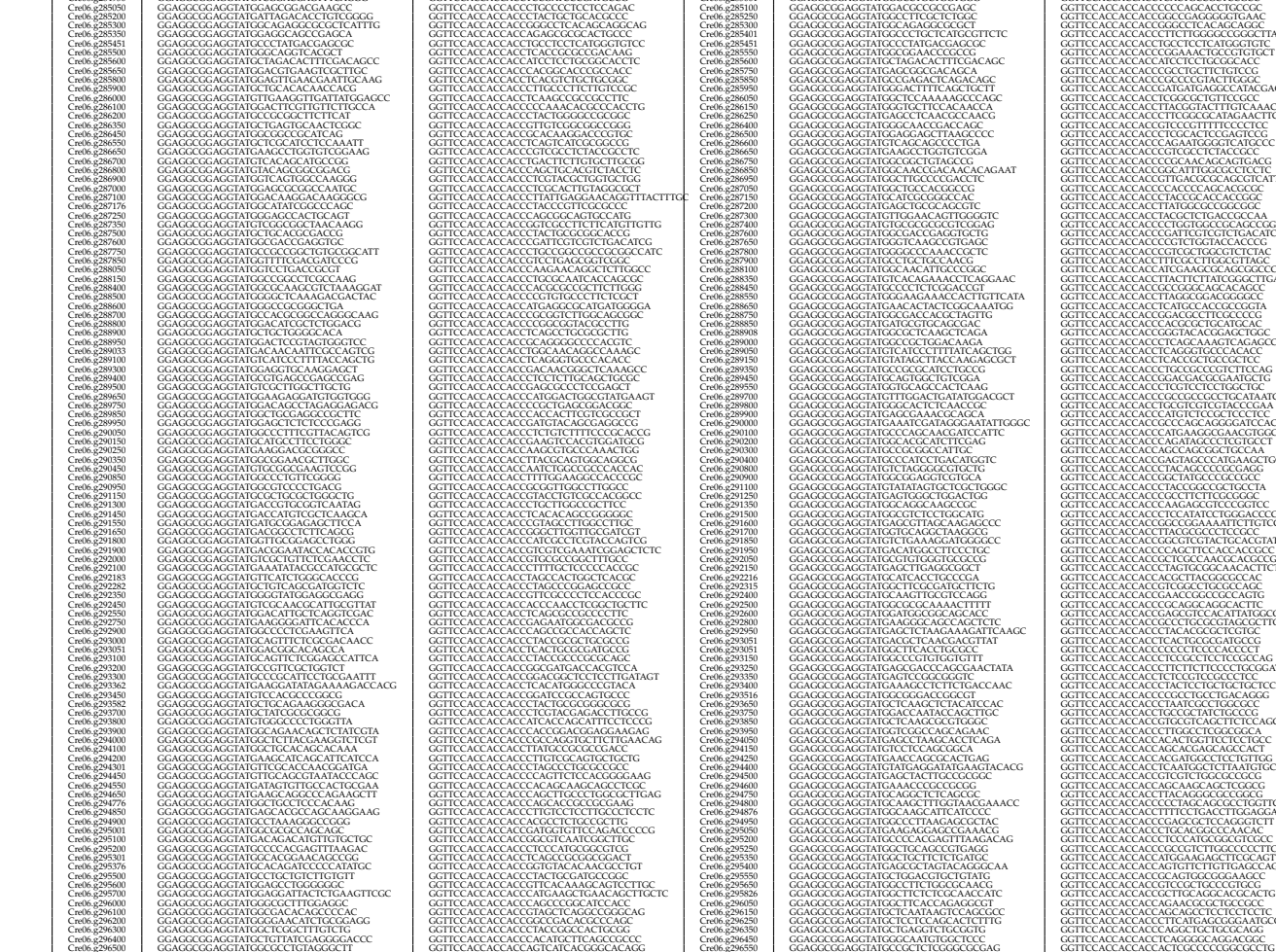


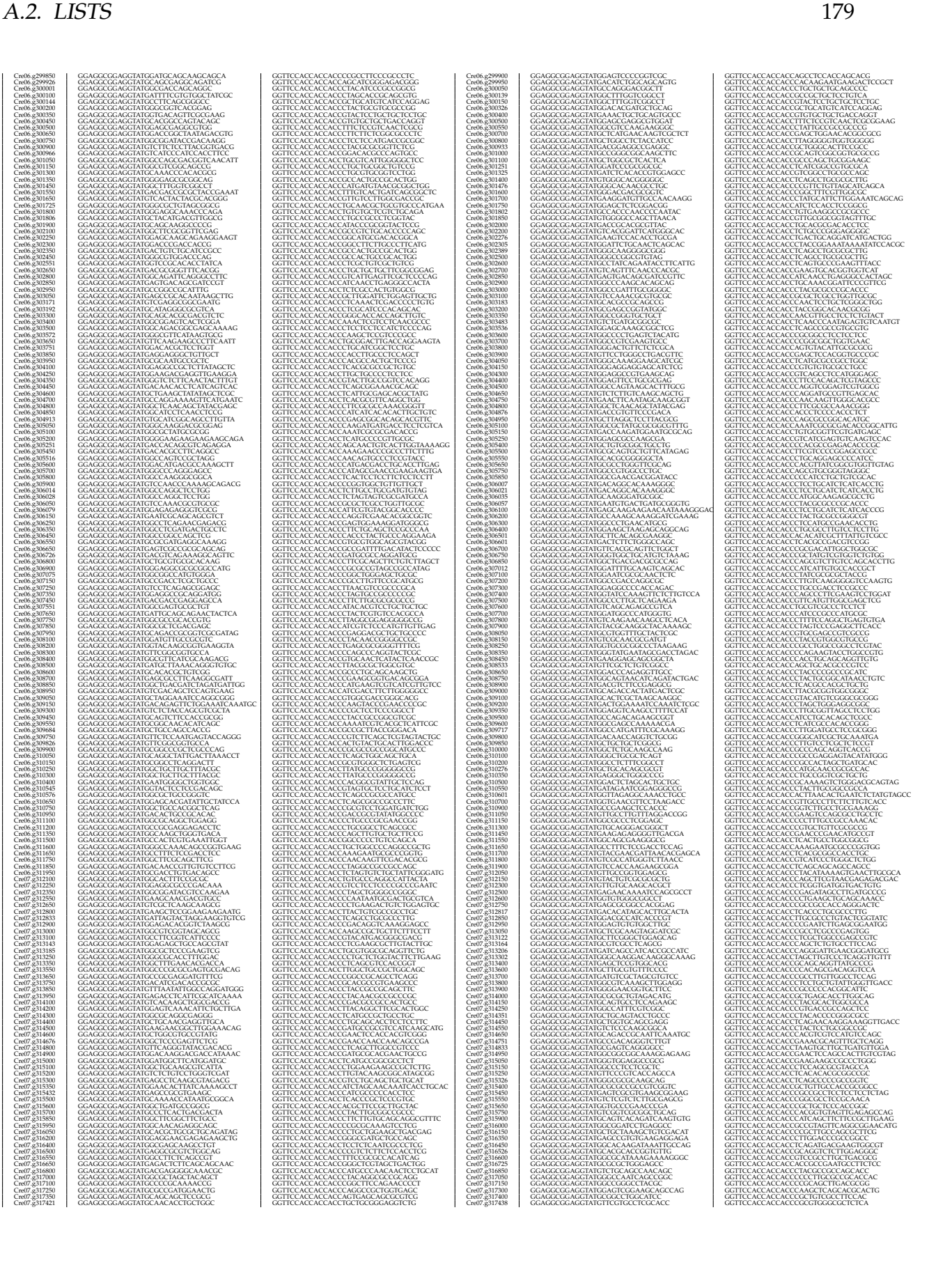


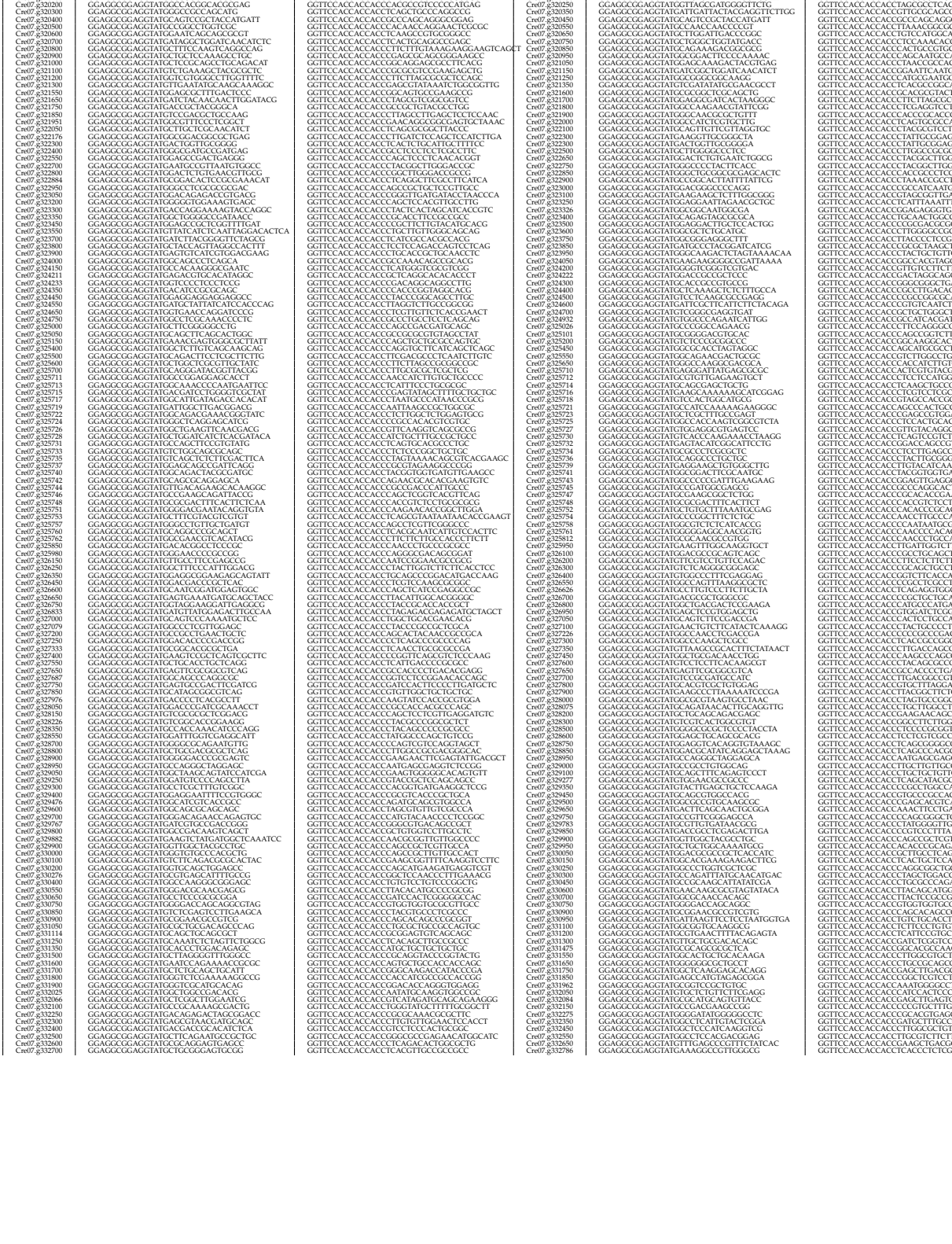
Cre05.g239400	GGAGGCGGAGGTATGGACTTACTACTGGACACGG	GGTTCCACCACCACCTTAGCAGGGGCGCGTG
Cre05.g239500	GGAGGCGGAGGTATGGGTGCCGCAGACC	GGTTCCACCACCACCGCCGCAGAAGCCCTTC
Cre05.g239650	GGAGGCGGAGGTATGTCGACGGAGCCAGGC	GGTTCCACCACCACCCGCGCTGAGCACGCCCTTTG
Cre05.g239700	GGAGGCGGAGGTATGGCAACTGACATAACAATGC	GGTTCCACCACCACCTCAATCAAACACCAGCTGTCC
Cre05.g239800	GGAGGCGGAGGTATGAGGTCATCAAAGCTGGGC	GGTTCCACCACCACCGCTGCTGCTGGTGCTGCC
Cre05.g239900	GGAGGCGGAGGTATGTGCACTACAAACTGTAGTCTG	GGTTCCACCACCACCGGACGACGTCCACTCGCG
Cre05.g239978	GGAGGCGGAGGTATGAACCGGCGCGAAAC	GGTTCCACCACCACCCGCCCTCGGCCAGAAC
Cre05.g240150	GGAGGCGGAGGTATGGCCACTGCTCGGCAA	GGTTCCACCACCACCCCGGGCGTAGTCCAGTGTGTAC
Cre05.g240225	GGAGGCGGAGGTATGCCCTCCGGACTCG	GGTTCCACCACCACCGCTGGTCGGTGGGAA
Cre05.g240300	GGAGGCGGAGGTATGATGCTTTCGCACAAGAGT	GGTTCCACCACCACCCAAGCCCTGCGCGAAG
Cre05.g240400	GGAGGCGGAGGTATGATGACATGGGGCGGC	GGTTCCACCACCACCTTTGTACGCCACCGCCGC
Cre05.g240500	GGAGGCGGAGGTATGGAGGAATGCCCCGTTTG	GGTTCCACCACCACCGGGCGCTCCGTAGCGCTC
Cre05.g240550	GGAGGCGGAGGTATGACTTTCAGCTGCTCCG	GGTTCCACCACCACCTTACAGCGCGTTGCCG
Cre05.g240700	GGAGGCGGAGGTATGCATAGCAGACAGCGC	GGTTCCACCACCACCAAGCAACGTCCGGCGC
Cre05.g240850	GGAGGCGGAGGTATGCAGCTGTTCGCTCGTTC	GGTTCCACCACCACCGTGGTACAGCTGGGCGCC
Cre05.g240950	GGAGGCGGAGGTATGCAGTCCGCCGCTGA	GGTTCCACCACCACCCTGGCGGGGCCAGGGT
Cre05.g241050	GGAGGCGGAGGTATGGAATGGAATCTGCGGATAG	GGTTCCACCACCACCCTACCGTTTCGACATGGCC
Cre05.g241202	GGAGGCGGAGGTATGTCCTTCTTGGACCCCGAC	GGTTCCACCACCACCCCGCCGACGAGCCAGC
Cre05.g241250	GGAGGCGGAGGTATGACGTCGCTCACTGCC	GGTTCCACCACCACCTCCGCTCACAAAGTCTGACA
Cre05.g241350	GGAGGCGGAGGTATGCGCCCAGCGTCAGCG	GGTTCCACCACCACCCACGTTGGTGGCCGAGCCAGAG
Cre05.g241400	GGAGGCGGAGGTATGACCTACTCGGCTCGCATTC	GGTTCCACCACCACCAGTCGTGAAGCTGAAGCCG
Cre05.g241450	GGAGGCGGAGGTATGCAGACGACCGTGGGG	GGTTCCACCACCACCCTTGGTGCCGGCAGTGGC
Cre05.g241550	GGAGGCGGAGGTATGGTGCCCAAGCTTCGTGG	GGTTCCACCACCACCTGGCGTTGTCGCGGCG
Cre05.g241629	GGAGGCGGAGGTATGCCCTTCGACACTTCTCT	GGTTCCACCACCACCTAAGGGCACGTATGCCATCC
Cre05.g241631	GGAGGCGGAGGTATGATAGCTCGCGTTCTTGGC	GGTTCCACCACCACCATCGACAATGCCCTTCGGC
Cre05.g241634	GGAGGCGGAGGTATGGACAGCCACGACCC	GGTTCCACCACCACCCTTGCCCGTCGCCTCC
Cre05.g241635	GGAGGCGGAGGTATGCTGCCGATGCCCTA	GGTTCCACCACCACCTACCATCCGCAGGTACTCG
Cre05.g241637	GGAGGCGGAGGTATGGATGACTCTATGGACTACCC	GGTTCCACCACCACCTCACAAGCCCAGCACATC
Cre05.g241639	GGAGGCGGAGGTATGCAATCTCAGCAGCTGC	GGTTCCACCACCACCGAACAGCCCGAACAGCC
Cre05.g241641	GGAGGCGGAGGTATGGACCCGGCATTCGGAC	GGTTCCACCACCACCACGGCGGCAACGCCAC
Cre05.g241643	GGAGGCGGAGGTATGGGCGTCTTAAAGCCG	GGTTCCACCACCACCAGTCCTGCTTTCTGCTGTTC
Cre05.g241645	GGAGGCGGAGGTATGCAGCGCGCAGCAG	GGTTCCACCACCACCCTACTTCCACGCGGCGC
Cre05.g241647	GGAGGCGGAGGTATGCCCGCCTTGGACT	GGTTCCACCACCACCCTACTCGTCACTGTCATAGACATC
Cre05.g241649	GGAGGCGGAGGTATGCCCGCCTTGGACT	GGTTCCACCACCACCCTACTCGTCACTGTCATAGACATC
Cre05.g241651	GGAGGCGGAGGTATGCGCTTATTGTGCACCG	GGTTCCACCACCACCCGTGAACTGCCATGAACCC
Cre05.g241653	GGAGGCGGAGGTATGGCCCGACAGAAAAGTAAA	GGTTCCACCACCACCCTACTTCTTCAACATCTCAGGGAT
Cre05.g241655	GGAGGCGGAGGTATGTCGATGTGCATCGCG	GGTTCCACCACCACCGCCGTACGATACGTAGCG
Cre05.g241751	GGAGGCGGAGGTATGATAGGGCGTGCCAGA	GGTTCCACCACCACCCACACCATGGCGGAACAA
Cre05.g241900	GGAGGCGGAGGTATGTCTTTGTGTTCGGCAGG	GGTTCCACCACCACCATCCTCATTGAACTGTGCGTC
Cre05.g242000	GGAGGCGGAGGTATGAAGTCTCTCTGCCATGAGC	GGTTCCACCACCACCCTAGTAGCCGCCCTTGGC
Cre05.g242100	GGAGGCGGAGGTATGCGCCCACGGTTACTA	GGTTCCACCACCACCAGACCTCCCACGCTTCTT
Cre05.g242179	GGAGGCGGAGGTATGCCGCTGTTGGAGCT	GGTTCCACCACCACCTCAAGACCGTGCGGCG
Cre05.g242300	GGAGGCGGAGGTATGTCAGCAGCATTTCCAGT	GGTTCCACCACCACCTTACGCGTCGTCCGCG
Cre05.g242350	GGAGGCGGAGGTATGGCGGACGACGGGTTT	GGTTCCACCACCACCTTCCGCCTTCTTGGCGGC
Cre05.g242400	GGAGGCGGAGGTATGCTGGCCTCCAAGCCC	GGTTCCACCACCACCAGCCAGGAAGCCCAGCTT
Cre05.g242501	GGAGGCGGAGGTATGGCCAAGTCATGCAACA	GGTTCCACCACCACCTCCAAGCGTAGCTCCGC
Cre05.g242550	GGAGGCGGAGGTATGGCCGCGGATGTGG	GGTTCCACCACCACCCTCGTCCTCCTCCTCCG
Cre05.g242602	GGAGGCGGAGGTATGGCGCCATCCTTCTCAAC	GGTTCCACCACCACCCTACGCCCCGCCTCGC
Cre05.g242750	GGAGGCGGAGGTATGTCGGTTTCCATAGGAACG	GGTTCCACCACCACCCGCCTGGATCTTTGCAGG
Cre05.g242856	GGAGGCGGAGGTATGCAAGCCACGGCTCTG	GGTTCCACCACCACCGAGGACTTGGATGGAGCCG
Cre05.g242859	GGAGGCGGAGGTATGGTGGAAGAGTCGGTGA	GGTTCCACCACCACCGTCCGCGTCGTCTGAGA
Cre05.g242950	GGAGGCGGAGGTATGTCTGGTGCCGAGACG	GGTTCCACCACCACCTACCTCGGTCCGGTACGC
Cre05.g243050	GGAGGCGGAGGTATGGCGCTGCCTTTGCGTC	GGTTCCACCACCACCGTTCTTGGGCGGCTGGTG
Cre05.g243351	GGAGGCGGAGGTATGCCGACAGCTATAGAAGACG	GGTTCCACCACCACCTTACTTCATGCACAGCACCA
Cre05.g243358	GGAGGCGGAGGTATGCCGCCAGCCTGGA	GGTTCCACCACCACCGTGCCTGCGGCCTCCT
Cre05.g243451	GGAGGCGGAGGTATGATTCTTCAGCACACGCA	GGTTCCACCACCACCCAGCTTCAGCTTGCCCTC
Cre05.g243453	GGAGGCGGAGGTATGGCCGAAAGAGCATCG	GGTTCCACCACCACCCCACGCGGTCCTGAAC
Cre05.g243455	GGAGGCGGAGGTATGGAGACCCGGGACAGT	GGTTCCACCACCACCCTCCATGGCCCAGAACGC
Cre05.g243550	GGAGGCGGAGGTATGAGAACCATTCGCGAAACC	GGTTCCACCACCACCAGCGCCGGAGTCTGAC
Cre05.g243601	GGAGGCGGAGGTATGCTACATGAGCTCGGTCG	GGTTCCACCACCACCCGAACCAGTGCAGCCG
Cre05.g243700	GGAGGCGGAGGTATGAAGTCAAGGGCGCTG	GGTTCCACCACCACCGTCCTCATGCTGCTGCTC
Cre05.g243802	GGAGGCGGAGGTATGTCCATGGACAGCGCT	GGTTCCACCACCACCCTGGACTTGTGCCTCCAC
Cre05.g243850	GGAGGCGGAGGTATGAAGGACCTGCAGGCC	GGTTCCACCACCACCCCAGGGGCAGCATCTTTG
Cre05.g243950	GGAGGCGGAGGTATGTCTAACCGCGACAAGC	GGTTCCACCACCACCTGGCTTTAAGCCTCCTGCT
Cre05.g244050	GGAGGCGGAGGTATGCCTTTGCGAGGAAGAGA	GGTTCCACCACCACCCTGGCGGCGAGTGGGT
Cre05.g244150	GGAGGCGGAGGTATGGAAAGTGGTTCAAAAGAAGA	GGTTCCACCACCACCCCTAGGTGGACCTCCACAGAG
Cre05.g244236	GGAGGCGGAGGTATGAGCAACTTTCTAGCATGCA	GGTTCCACCACCACCCCCGCACACGAACTTGCG
Cre05.g244350	GGAGGCGGAGGTATGCGCGGTGCCAATGT	GGTTCCACCACCACCTCAGGCTTCGCCGGCG
Cre05.g244450	GGAGGCGGAGGTATGGATGCAGTCAACCAGC	GGTTCCACCACCACCCGCCCTCCTTCTTTCCCC
Cre05.g244650	GGAGGCGGAGGTATGGGCTGGGAGGCCG	GGTTCCACCACCACCGCCCTGCACCGGCTCC
Cre05.g244750	GGAGGCGGAGGTATGACGAGGCCGCCGC	GGTTCCACCACCACCCAGCTCCACGCCCTCG
Cre05.g244850	GGAGGCGGAGGTATGGCTATTTCAGGCTCGTTACA	GGTTCCACCACCACCCCCGAAGAGCTCCCACGA
Cre05.g244901	GGAGGCGGAGGTATGCCTAAAGTCGTTGGTTTCC	GGTTCCACCACCACCGTGGTGGTCCTCCTTTGGC
Cre05.g245100	GGAGGCGGAGGTATGAACAACAAGGATGTGGATCG	GGTTCCACCACCACCTCACTGCTGCGTGCCT
Cre05.g245158	GGAGGCGGAGGTCCACACCCGCGCCTGC	GGTTCCACCACCACCCTTCCACACCACGTTGCCG
Cre05.g245351	GGAGGCGGAGGTATGCTCGACAACACCACAC	GGTTCCACCACCACCGTTGCTGCTGCTGCCG
Cre05.g245451	GGAGGCGGAGGTATGCATCTGCGCTTCGTGA	GGTTCCACCACCACCCCGCCTGCTGCGGAACAG
Cre05.g245701	GGAGGCGGAGGTATGGACTTTGCGGAAAGCTG	GGTTCCACCACCACCCACGCCCGCAGTGGTG
Cre05.g245900	GGAGGCGGAGGTATGAGCTTGGGCTTCCGC	GGTTCCACCACCACCCTTGACGGGGTAGACCCA
Cre05.g246000	GGAGGCGGAGGTATGACTGGTGCTGCGC	GGTTCCACCACCACCCTCCAGCGTGTGGTCCT
Cre05.g246100	GGAGGCGGAGGTATGCTGCTAATCGGCTATTCC	GGTTCCACCACCACCGCTGGTGGTGCCTCTCA
Cre05.g246200	GGAGGCGGAGGTATGGCTCATTGGGCTGCTT	GGTTCCACCACCACCGCTACTGCCGAATGCGCT
Cre05.g246376	GGAGGCGGAGGTATGCTGCTTGGTAGAGCGC	GGTTCCACCACCACCACCTGCGCGCTGGCTT
Cre05.g246400	GGAGGCGGAGGTATGGGCGGCGAGGTGTGG	GGTTCCACCACCACCCGGTTGCAGCTGGGGCTG
Cre05.g246551	GGAGGCGGAGGTATGGCGCGGGGGTCAG	GGTTCCACCACCACCTCGCGACACACCTATGCG
Cre05.g246553	GGAGGCGGAGGTATGACGGCCACCGAGG	GGTTCCACCACCACCGGCCACACTGGCCATG
Cre05.g246650	GGAGGCGGAGGTATGCCTCGGCGTTCTCT	GGTTCCACCACCACCCTCATCGTCGCTGTCGC
Cre05.g246800	GGAGGCGGAGGTATGCTGGCCCAAACACACA	GGTTCCACCACCACCGAACAGCGACTGTGTCCGC
Cre05.g246900	GGAGGCGGAGGTATGCCGACGAGCGGCG	GGTTCCACCACCACCGGCCTTGTTGACGATGATGGT
Cre05.g247050	GGAGGCGGAGGTATGGGGTTGCCCGCACTC	GGTTCCACCACCACCGCTGCCGTGGCACGCCTC
Cre05.g247150	GGAGGCGGAGGTATGCTTGGCCTCTGCCCC	GGTTCCACCACCACCCAATGCGGCCACGCGC
Cre05.g247350	GGAGGCGGAGGTATGGCAGTGGTGGTATGCA	GGTTCCACCACCACCGTGCACCCTGCTGCCT
Cre05.g247450	GGAGGCGGAGGTATGCCGCTATCGCTCCA	GGTTCCACCACCACCGTCGCTGGTGCTGGCG
Cre05.g247550	GGAGGCGGAGGTATGCCGGGCGAGGGAGCAG	GGTTCCACCACCACCCAGCTTGCTCTGCGCCGC
Cre05.g247650	GGAGGCGGAGGTATGCCAGGCTCCGGCT	GGTTCCACCACCACCGGCCGCTCGTACGGCA
Cre05.g247850	GGAGGCGGAGGTATGTCGTCAACAACATCCAGA	GGTTCCACCACCACCGAACTCCACGTCTGGGAAC
Cre05.g247900	GGAGGCGGAGGTATGGGTCTGGACAGTTTGCCG	GGTTCCACCACCACCCGCCGCCCACTCCAGC
Cre05.g248000	GGAGGCGGAGGTATGAAGCTAACAAAGAGGTGCT	GGTTCCACCACCACCTCACAGGCGACCCAGC
Cre05.g248150	GGAGGCGGAGGTATGGCGGCATCCACTTCG	GGTTCCACCACCACCGGCGGTGAGGCCATGCTT
Cre05.g248250	GGAGGCGGAGGTATGCAAGTAAACGTGCGCC	GGTTCCACCACCACCCTTCTCCACCACAAACTCCTC
Cre05.g248400	GGAGGCGGAGGTATGTCGTCGCGGAATGTC	GGTTCCACCACCACCGGCGATCTCGGTCACCT
Cre05.g248450	GGAGGCGGAGGTATGTCGTCGCGGAATGTCG	GGTTCCACCACCACCGGCAATCTCGGTCACCTTGC
Cre05.g248550	GGAGGCGGAGGTATGGCTGGCGCAGCAG	GGTTCCACCACCACCCAGCACGCGAGTGTTGTA
Cre05.g248650	GGAGGCGGAGGTATGTCGCGCATCGGCC	GGTTCCACCACCACCGCTCCCGTACCCGCCA
Cre06.g246958	GGAGGCGGAGGTATGCCACATCGCTTCACAGG	GGTTCCACCACCACCCAGGCGGGAGATGGCCCAAAG
Cre06.g248800	GGAGGCGGAGGTATGGTGAAGAAGAAGGTTGGAG	GGTTCCACCACCACCTCATTCCTGCATTTTGCTGGA
Cre06.g248900	GGAGGCGGAGGTATGGACCACGCAGACCT	GGTTCCACCACCACCCGTTGCCTCTTCTCCGGG
Cre06.g249050	GGAGGCGGAGGTATGGGCCAGGAGGGCAAG	GGTTCCACCACCACCGTCTTCCTTGCCGCGGTT
Cre06.g249150	GGAGGCGGAGGTATGGCAAGCATTCGCATGAG	GGTTCCACCACCACCCTGCTTGAGGAGCCGGAAG
Cre06.g249250	GGAGGCGGAGGTATGAGTGAGGCTGTGAACCC	GGTTCCACCACCACCGATGAGCAGCTTCTCGATGG
Cre06.g249350	GGAGGCGGAGGTATGTGCAATCTGCAGGCTCG	GGTTCCACCACCACCCACGCCCTCCTCCTCCAC
Cre06.g249400	GGAGGCGGAGGTATGTTGCGAGCGAGTAGTC	GGTTCCACCACCACCTCATTCGGAATTTGACTCATCAGC
Cre06.g249500	GGAGGCGGAGGTATGCTCGGGATCCAGGGCTT	GGTTCCACCACCACCACCACATGCCTCCTGCAC
Cre06.g249550	GGAGGCGGAGGTATGTGGCTGCTTGCTGCG	GGTTCCACCACCACCCCCTGCGATACCGCCCTC
Cre06.g249600	GGAGGCGGAGGTATGCGCCTTGCGGACCTT	GGTTCCACCACCACCCTCTGCGCTACCGCCCTC
Cre06.g249650	GGAGGCGGAGGTATGGCTGGGCCGCTCC	GGTTCCACCACCACCATTGCACACCAAGAACAGCG
Cre06.g249750	GGAGGCGGAGGTATGGCACCCGCTTCCA	GGTTCCACCACCACCCTTGCGGCCCTTCTTGC
Cre06.g249850	GGAGGCGGAGGTATGTCATCCTTCGCGACAAA	GGTTCCACCACCACCTAGCGCCACGCTGCTG
Cre06.g249900	GGAGGCGGAGGTATGGCAAGCACCGCCAA	GGTTCCACCACCACCCGTCGCCGCACTGGCT
Cre06.g250000	GGAGGCGGAGGTATGGCGGCTTTTCTGAAGC	GGTTCCACCACCACCCCTCCCGTTCTTGGCCC
Cre06.g250100	GGAGGCGGAGGTATGCCGGTTCAGCAGATGACTT	GGTTCCACCACCACCCTTCTTGTCAGTGAACTCGGCG
Cre06.g250200	GGAGGCGGAGGTATGGCCGCCCTGGAGT	GGTTCCACCACCACCCTCCAGCTTCTTCACGGTC
Cre06.g250276	GGAGGCGGAGGTATGCCGCGCATGGGCT	GGTTCCACCACCACCTGGGGCACGCCGAGCT
Cre06.g250350	GGAGGCGGAGGTATGTTGCCCCAGGCGACC	GGTTCCACCACCACCCCCGAAGTACACGCTCAGC
Cre06.g250500	GGAGGCGGAGGTATGGCCGCGCGCACTGGG	GGTTCCACCACCACCCGAGCCGCCGCCCCCGCC
Cre06.g250567	GGAGGCGGAGGTATGACATCCACCCACAGATGG	GGTTCCACCACCACCTTACGGCCCCTGTGTGC
Cre06.g250650	GGAGGCGGAGGTATGGAATCGACCTTTCTACCAC	GGTTCCACCACCACCCTACAGCTGGGCGGCG
Cre06.g250750	GGAGGCGGAGGTATGGGCGCCAATAGCAGC	GGTTCCACCACCACCCCCTGCCTTCTCGTAGAACAG
Cre06.g250850	GGAGGCGGAGGTATGGCTGCCATCGAAAAGCC	GGTTCCACCACCACCCTCGGCCTTACGCTTGCT
Cre06.g250950	GGAGGCGGAGGTATGGCCGAGCACTTGGC	GGTTCCACCACCACCCCACCCACCACCGCCG
Cre06.g251050	GGAGGCGGAGGTATGTCGTCGGCCAAGGCCAAG	GGTTCCACCACCACCCCGAGCCGCTACGTTCCC
Cre06.g251150	GGAGGCGGAGGTATGTCGATTGCTGCACTCC	GGTTCCACCACCACCGTCCAGGTCCACGATGCC
Cre06.g251400	GGAGGCGGAGGTATGCTCCGAGCTCGGC	GGTTCCACCACCACCCCTGTACCGCCTGTACTGC
Cre06.g251500	GGAGGCGGAGGTATGGCCAACTACGCCAAGC	GGTTCCACCACCACCCGCCTTGACGCCGGCTCC
Cre06.g251600	GGAGGCGGAGGTATGCAGAATATCGGAGCCGC	GGTTCCACCACCACCCTCGTCGTCCTCGTCCTCC
Cre06.g251683	GGAGGCGGAGGTATGCTGCTTGGTAGAACGCG	GGTTCCACCACCACCCTGCTGCGTGCGGCCTA
Cre06.g251800	GGAGGCGGAGGTATGAGCACGCAGCCATG	GGTTCCACCACCACCCTACGCCGCTGCCAGT
Cre06.g251900	GGAGGCGGAGGTATGTACGTGCTGGCCGA	GGTTCCACCACCACCTCAAGCCGACGCACCG
Cre06.g252000	GGAGGCGGAGGTATGTCAATACAACTCTCTGAAGATGG	GGTTCCACCACCACCGGCGGTGTCTGCACTGGG
Cre06.g252100	GGAGGCGGAGGTATGGGGAAGGATACGAAAGGG	GGTTCCACCACCACCCCAGAGTGCCTGGCAGT
Cre06.g252250	GGAGGCGGAGGTATGCCTCGGGCTACGACG	GGTTCCACCACCACCCCGCGTCAGGCCTCGC
Cre06.g252350	GGAGGCGGAGGTATGCTCGGCGGCGCTG	GGTTCCACCACCACCGCCCTGCCTTTTGACGCG
Cre06.g252450	GGAGGCGGAGGTATGGGGCACCTGGGGG	GGTTCCACCACCACCGATGCGCTTCACAATCTGC
Cre06.g252576	GGAGGCGGAGGTATGCGCAGGATGGGGC	GGTTCCACCACCACCGAAGGGCCCCGCCCAG
Cre06.g252650	GGAGGCGGAGGTATGTTGGCGGCGCAGAAG	GGTTCCACCACCACCGGCAGTTTTGATCATGCCCTG
Cre06.g252743	GGAGGCGGAGGTATGATTTGCGGCCCCG	GGTTCCACCACCACCCTAGTGCCATACTATCAAACGCT
Cre06.g252800	GGAGGCGGAGGTATGACATTTTACGTCCCCAAGGT	GGTTCCACCACCACCGTTTCCACCCTGAAACCCCAG
Cre06.g252871	GGAGGCGGAGGTATGCAGCTGAGCGACCGG	GGTTCCACCACCACCTCCTGCCGCAGGGCAGAG
Cre06.g252892	GGAGGCGGAGGTATGCGGTGGTGGAGCATG	GGTTCCACCACCACCAGTCAGGTGAAGCCGCCC
Cre06.g253000	GGAGGCGGAGGTATGAGGACTTCGCTGTGGG	GGTTCCACCACCACCCCTCCGCAGGCCCCTC
Cre06.g253102	GGAGGCGGAGGTATGAACCCTAGCCCCAGC	GGTTCCACCACCACCGTGAATGCGAATGCAGGACG
Cre06.g253300	GGAGGCGGAGGTATGGAGTGGACGACAATAGAATCG	GGTTCCACCACCACCCTGCTTCGTACCATTGCCTCC
Cre06.g253404	GGAGGCGGAGGTATGCTTAAAGAGCAGATCAGCC	GGTTCCACCACCACCCTATCCACGCAGCCCTCC
Cre06.g253566	GGAGGCGGAGGTATGCTGCTTTATACACCCCG	GGTTCCACCACCACCTTATATAATTGCCGCCACTAGTT
Cre06.g253679	GGAGGCGGAGGTATGGCTGTTGCCTGCG	GGTTCCACCACCACCGCCCACCTTGCTGGTAAC
Cre06.g253753	GGAGGCGGAGGTATGGCCTGTAGCGCGG	GGTTCCACCACCACCCTGCAACTCTCCAACTCCG
Cre06.g253757	GGAGGCGGAGGTATGGCTGTTGCCTGCG	GGTTCCACCACCACCGCCCACCTTGCTGGTAAC
Cre06.g253759	GGAGGCGGAGGTATGGTGTCCGGTAAGATTTGGC	GGTTCCACCACCACCTCAGCGGGGTTGCCGG
Cre06.g253800	GGAGGCGGAGGTATGGCTGTTGCCTGCGC	GGTTCCACCACCACCTCAGCGGGGTTGCCGG
Cre06.g253851	GGAGGCGGAGGTATGCTGCTTTATACACCCCGC	GGTTCCACCACCACCTATAATTGCCGCCACTAGTTGC
Cre06.g253900	GGAGGCGGAGGTATGCGATGGAGCTTCTACC	GGTTCCACCACCACCTTATATAATTGCCGCCACTAGTT
Cre06.g253975	GGAGGCGGAGGTATGTCCTGGGCATTCGGC	GGTTCCACCACCACCGACGAGATGACGGGTCATCT
Cre06.g254077	GGAGGCGGAGGTATGGCTGTTGCCTGCGC	GGTTCCACCACCACCTCAGCGGGGTTGCCGG
Cre06.g254150	GGAGGCGGAGGTATGAACTACGTGTTTTCCAACCT	GGTTCCACCACCACCATCAGAAGCCTCCTCCTCCTC
Cre06.g254200	GGAGGCGGAGGTATGGCCACTTCATGCCTCC	GGTTCCACCACCACCGGTGTTGACGACCGTGGC
Cre06.g254224	GGAGGCGGAGGTATGATTGAGGATGCGGCG	GGTTCCACCACCACCCTTACTGCCGGCGGGG
Cre06.g254240	GGAGGCGGAGGTATGGCTCTGCAACACTACCG	GGTTCCACCACCACCCTTGCGTGTCTGCTGGT
Cre06.g254275	GGAGGCGGAGGTATGACGCGCTTCCTGCTG	GGTTCCACCACCACCCCCCCGACACTTTCCAGACTG
Cre06.g254350	GGAGGCGGAGGTATGTCGCCAGGCGCAT	GGTTCCACCACCACCCACGCGCATCTGCACC
Cre06.g254450	GGAGGCGGAGGTATGCTTGTGAAGATATTCTCCCAG	GGTTCCACCACCACCCACCAGCATGTAGTCCATCTGG
Cre06.g254500	GGAGGCGGAGGTATGGCGGCTCGGCTGG	GGTTCCACCACCACCCGGCGGTGAGTCCCAC
Cre06.g254600	GGAGGCGGAGGTATGCGCGCAAGGCTATT	GGTTCCACCACCACCGTTGAGCTCGGGCAATGT
Cre06.g254850	GGAGGCGGAGGTATGCCGACCGGAGGAGGC	GGTTCCACCACCACCTTGCCGCCGCTGCGTG
Cre06.g255100	GGAGGCGGAGGTATGCTTGCGTGCTCAACTC	GGTTCCACCACCACCCGCCTGCTGCTGCGTC
Cre06.g255200	GGAGGCGGAGGTATGCCTCCCAGCCCGTTG	GGTTCCACCACCACCGTGGATCTCATGGATCTTGCGC
Cre05.g239450	GGAGGCGGAGGTATGGCGCCCAATTCGC	GGTTCCACCACCACCCATCAGCTGCTCCTCCGC
Cre05.g239550	GGAGGCGGAGGTATGTCTGCGGCCACTGA	GGTTCCACCACCACCTTCACACTCTTCCGCATCTC
Cre05.g239652	GGAGGCGGAGGTATGGAATCGCCGTTGGAGTG	GGTTCCACCACCACCCAGCCCCATCGGCCGG
Cre05.g239750	GGAGGCGGAGGTATGCTTGCTACAGTAACGCTGT	GGTTCCACCACCACCTGCCAGCCATGTGCGGTA
Cre05.g239850	GGAGGCGGAGGTATGGCGACGAGCTCAGTTG	GGTTCCACCACCACCAAGCGCATACACCGGCAC
Cre05.g239950	GGAGGCGGAGGTATGCTGCAGCCGCGTG	GGTTCCACCACCACCCTCCTGTAGCGGCCCTG
Cre05.g240000	GGAGGCGGAGGTATGGTCAATTGGAATAGTATTGGTCAC	GGTTCCACCACCACCGTATCGCCCACTGACAGCG
Cre05.g240200	GGAGGCGGAGGTATGTCCGGAACGAGCGC	GGTTCCACCACCACCGCCACATACGGCGGCC
Cre05.g240251	GGAGGCGGAGGTATGTCTTTCCGCTCAGCCG	GGTTCCACCACCACCCTAATCGTCCTGGAAGTCTAGGTA
Cre05.g240350	GGAGGCGGAGGTATGGTCATTGACCCGACGC	GGTTCCACCACCACCGCTGCCCAGCAGCTGC
Cre05.g240450	GGAGGCGGAGGTATGGCTTCTAATGAGCAACAACC	GGTTCCACCACCACCTCATTTCGGTGTGTCCTCG
Cre05.g240533	GGAGGCGGAGGTATGTCTGGCGAGCGCGCAAG	GGTTCCACCACCACCGGCGGTGGCCCCACCTCG
Cre05.g240650	GGAGGCGGAGGTATGGCGCCGAAACAGCGC	GGTTCCACCACCACCGCCCTTAAGCGCGGCGTG
Cre05.g240800	GGAGGCGGAGGTATGTCAAACGCCCTGCG	GGTTCCACCACCACCCGCCTCCAGCTCCTCA
Cre05.g240900	GGAGGCGGAGGTATGGCGCGGTGTCAGG	GGTTCCACCACCACCCTGGTAGCCAGTCGGATACA
Cre05.g241000	GGAGGCGGAGGTATGATGCTGTCTCGCAACAC	GGTTCCACCACCACCCTGCTTGTCCACCACCAG
Cre05.g241100	GGAGGCGGAGGTATGGGGGCTTGTGGGCGAG	GGTTCCACCACCACCCTTGGGCATAAACTCCTCCACC
Cre05.g241250	GGAGGCGGAGGTATGACGTCGCTCACTGCC	GGTTCCACCACCACCTCCGCTCACAAAGTCTGACAC
Cre05.g241300	GGAGGCGGAGGTATGGCGGCTTTCGTGC	GGTTCCACCACCACCTGACATGGGCTCGTCCTC
Cre05.g241350	GGAGGCGGAGGTATGCGCCCAGCGTCAG	GGTTCCACCACCACCCACGTTGGTGGCCGAG
Cre05.g241400	GGAGGCGGAGGTATGACCTACTCGGCTCGC	GGTTCCACCACCACCAGTCGTGAAGCTGAAGCCG
Cre05.g241500	GGAGGCGGAGGTATGTGCCCCGTGGGTG	GGTTCCACCACCACCCCTGGGGCACCGGGTC
Cre05.g241600	GGAGGCGGAGGTATGTGGCTGACGAAATTAAGAGC	GGTTCCACCACCACCTTAGACAGGGAGGGCCAGG
Cre05.g241630	GGAGGCGGAGGTATGAGCGCCACCGCTG	GGTTCCACCACCACCTGTCCGGGCGATGAGG
Cre05.g241633	GGAGGCGGAGGTATGCTCGCGGCCTCCG	GGTTCCACCACCACCCAGCCGCCTCAGCAGC
Cre05.g241635	GGAGGCGGAGGTATGCTGCCGATGCCCTAC	GGTTCCACCACCACCTACCATCCGCAGGTACTCGTC
Cre05.g241636	GGAGGCGGAGGTATGGACGCTGTTGTTGGTGG	GGTTCCACCACCACCCTAAGCGGCGGGCGGC
Cre05.g241638	GGAGGCGGAGGTATGCATTTGACGCTGCCAC	GGTTCCACCACCACCCTCGTCGTCGTCCTCCTCA
Cre05.g241640	GGAGGCGGAGGTATGGCCACGGGCACAC	GGTTCCACCACCACCGCCGTAGTACTCATCCAGCT
Cre05.g241642	GGAGGCGGAGGTATGGAATCCCGGACCTGTC	GGTTCCACCACCACCGGTTACTTTCATCGCCTTGCG
Cre05.g241644	GGAGGCGGAGGTATGGGCCGACAAGGGCC	GGTTCCACCACCACCCTACGTGCGGCCCCGG
Cre05.g241646	GGAGGCGGAGGTATGGACGTGGTGGGAGC	GGTTCCACCACCACCCTACTTCCACGCGGCG
Cre05.g241648	GGAGGCGGAGGTATGGGCAAACGAGGGCG	GGTTCCACCACCACCTTCCAGGGGTTGGGGGT
Cre05.g241650	GGAGGCGGAGGTATGAGCCTATTCTGCCGGC	GGTTCCACCACCACCGGCAGTAGAGAAGCGCTTCAG
Cre05.g241652	GGAGGCGGAGGTATGGGCGCTTCAGCAGCG	GGTTCCACCACCACCCCTGCAAGCAATGTCGTACCA
Cre05.g241654	GGAGGCGGAGGTATGAACAGAACGCGCAACC	GGTTCCACCACCACCCCTGCCGTGCCTCCTG
Cre05.g241750	GGAGGCGGAGGTATGTTGGCAAGCCTTGCGC	GGTTCCACCACCACCGGGCAGCGGCTCCTGG
Cre05.g241850	GGAGGCGGAGGTATGAAGCGTCTAGCAGGGATCG	GGTTCCACCACCACCTCATGCCGCAGCGCCG
Cre05.g241950	GGAGGCGGAGGTATGCCCGTCGTCAGCTTC	GGTTCCACCACCACCCGCCAGGTCCACTGCGAAG
Cre05.g242050	GGAGGCGGAGGTATGCTGGCACCATGGCG	GGTTCCACCACCACCACGTAGTTCGGCAGACAGT
Cre05.g242178	GGAGGCGGAGGTATGACATTCATTGGCAAGCTAGTA	GGTTCCACCACCACCCTACTGGTGGTGCTGTTGC
Cre05.g242180	GGAGGCGGAGGTATGTCGTCCTTGTTGTCCTG	GGTTCCACCACCACCGTTGCGGGCTACACAGG
Cre05.g242301	GGAGGCGGAGGTATGGACGTGGACCGGG	GGTTCCACCACCACCTTACGAAGGAAGAACACGAAGAA
Cre05.g242350	GGAGGCGGAGGTATGGCGGACGACGGGT	GGTTCCACCACCACCTTCCGCCTTCTTGGCGG
Cre05.g242500	GGAGGCGGAGGTATGGCGGCAGTGGACAG	GGTTCCACCACCACCCTCGTCCGCCTCGGCC
Cre05.g242502	GGAGGCGGAGGTATGTGGGCTGCTGTGCT	GGTTCCACCACCACCTCAGCTCCAGGCCGCG
Cre05.g242600	GGAGGCGGAGGTATGGCCCCCGCCCCAGCTTTC	GGTTCCACCACCACCGTAGCTCGCACGGTTGCGTTTG
Cre05.g242700	GGAGGCGGAGGTATGTTTGGCACTTTTGTCCCG	GGTTCCACCACCACCTGTACCCTTAATGCTCCGCT
Cre05.g242850	GGAGGCGGAGGTATGGCATCTAATAACACGGTGT	GGTTCCACCACCACCCTATGACTGCTTGTCGTGAGC
Cre05.g242857	GGAGGCGGAGGTATGGGCGCAGGAAGCT	GGTTCCACCACCACCCTACCCCAGCCCCAGC
Cre05.g242860	GGAGGCGGAGGTATGGCTTGGGTTCTGCCA	GGTTCCACCACCACCTACCCTTGGCGCTGCA
Cre05.g243000	GGAGGCGGAGGTATGGACTCCCAAGGCGTGAAA	GGTTCCACCACCACCGTTTTTCAGCAGGAAGTTGAAGA
Cre05.g243151	GGAGGCGGAGGTATGGGTGGAGGAGCGTCAA	GGTTCCACCACCACCGGACGCCTCGAGCGCC
Cre05.g243354	GGAGGCGGAGGTATGAACCGTGTCGCCAAC	GGTTCCACCACCACCCCACATTCCGCTGCAAGG
Cre05.g243450	GGAGGCGGAGGTATGCTATTGCCGCTACTCTTG	GGTTCCACCACCACCCTATTGCGGTCGCGCC
Cre05.g243452	GGAGGCGGAGGTATGAGCCGCTCGAGAGC	GGTTCCACCACCACCCTGCCGCTGCTGCTGG
Cre05.g243454	GGAGGCGGAGGTATGTTTAGTTCTCATGGAAGTGGT	GGTTCCACCACCACCCTATGGATTCCGAATGGTGCG
Cre05.g243472	GGAGGCGGAGGTATGATGGCTCGACTCTCGG	GGTTCCACCACCACCCTCTTCCCACGGTGAAACTC
Cre05.g243600	GGAGGCGGAGGTATGGGGTACGGCACCACC	GGTTCCACCACCACCCACCGCACTGCTGTACACG
Cre05.g243650	GGAGGCGGAGGTATGAAAACCATTCGCGAAACCC	GGTTCCACCACCACCAGCGCCGGAGTCTGACATT
Cre05.g243800	GGAGGCGGAGGTATGGCGTCTATTACCTGCAGC	GGTTCCACCACCACCGCGGTTGCGGGTCAGCAG
Cre05.g243803	GGAGGCGGAGGTATGCACCTGCCACGTCC	GGTTCCACCACCACCGGAGCCGGCCGTACAC
Cre05.g243900	GGAGGCGGAGGTATGAGTCCCTGGGGCAC	GGTTCCACCACCACCGTCGGCAAGGCCACCC
Cre05.g244000	GGAGGCGGAGGTATGGCGCCACCAAGCG	GGTTCCACCACCACCCAGGCCGGAGCTGTACG
Cre05.g244052	GGAGGCGGAGGTATGGCGACGAGGGCAG	GGTTCCACCACCACCTCACGGCCCCCAGAGC
Cre05.g244200	GGAGGCGGAGGTATGGCGGCCGATGCAG	GGTTCCACCACCACCCCGCTTTCGGGTGTTGC
Cre05.g244250	GGAGGCGGAGGTATGGACTGGGATGACAGCA	GGTTCCACCACCACCATCCGCCAGCGACAGC
Cre05.g244400	GGAGGCGGAGGTATGGCGGTGCAGCTCGGG	GGTTCCACCACCACCCTTGGTCGCTGCCAGCATTTC
Cre05.g244550	GGAGGCGGAGGTATGGCCGACTGGTCCACA	GGTTCCACCACCACCCGACGGCGATCCGGCG
Cre05.g244700	GGAGGCGGAGGTATGGAAACGGAGGAGCTGTTG	GGTTCCACCACCACCACCGTACGGCGTGTTGTG
Cre05.g244800	GGAGGCGGAGGTATGCGCGCGTTGCACA	GGTTCCACCACCACCTCAAGCCGCCGCGGTG
Cre05.g244900	GGAGGCGGAGGTATGGTGCTGGGCACCTTCA	GGTTCCACCACCACCGCTGCCCTGCAGGGCG
Cre05.g244950	GGAGGCGGAGGTATGAAGTCCACGATGATCTCGC	GGTTCCACCACCACCGTCGTCCTTGGTGAGCTCCTT
Cre05.g245150	GGAGGCGGAGGTATGGGCAAGAACATGCCTG	GGTTCCACCACCACCCCGCCTCGTTGGCGTC
Cre05.g245350	GGAGGCGGAGGTATGTGGCGAAATGCGCTG	GGTTCCACCACCACCTGTGCGACCACGTCTCCG
Cre05.g245450	GGAGGCGGAGGTATGGCGCGCAGGCTAGCTC	GGTTCCACCACCACCCAGCTTGAACCCAGAGTGCAG
Cre05.g245700	GGAGGCGGAGGTATGGGCTGCGGCGCAT	GGTTCCACCACCACCCCGCACGTTTACGTCGGC
Cre05.g245800	GGAGGCGGAGGTATGGTGGTAGACGTGGCG	GGTTCCACCACCACCCTGCATGTACATCAGCGTGC
Cre05.g245950	GGAGGCGGAGGTATGGAGCGCGTGATCGGG	GGTTCCACCACCACCCATGAACGCCGCGATCTCCTTG
Cre05.g246050	GGAGGCGGAGGTATGGCGAAGAAAGAGAAGATTACA	GGTTCCACCACCACCAATCACCAGCGCTGCCAC
Cre05.g246150	GGAGGCGGAGGTATGGGGCGCTTTGTTGG	GGTTCCACCACCACCTGGCTTGGGGACTGCA
Cre05.g246300	GGAGGCGGAGGTATGGCTCCCGCAGCTG	GGTTCCACCACCACCTCAAGCCCCGCCACCG
Cre05.g246377	GGAGGCGGAGGTATGCATTGCCAAGCAGCA	GGTTCCACCACCACCCTGCCGCCGCTTCTCC
Cre05.g246550	GGAGGCGGAGGTATGGCCCTAGTCGCAGTCC	GGTTCCACCACCACCCTACTGCCGCCGCCGC
Cre05.g246552	GGAGGCGGAGGTATGCCCGCTCTGGGCT	GGTTCCACCACCACCAAACGCACTCCTCAACCCA
Cre05.g246600	GGAGGCGGAGGTATGTTTGGGAAGTTCCGCTG	GGTTCCACCACCACCTTACTTCCTTGACGAGGGAAAAA
Cre05.g246752	GGAGGCGGAGGTATGTCGAACTCGTGCTCG	GGTTCCACCACCACCCGCGGGGGCCTTAAAC
Cre05.g246850	GGAGGCGGAGGTATGGGCCTCATCACCGG	GGTTCCACCACCACCGAATATGATTGCCAGCAGCAG
Cre05.g246950	GGAGGCGGAGGTATGCCGAGCCTCAACCT	GGTTCCACCACCACCGATGCCCATGGGCATGC
Cre05.g247100	GGAGGCGGAGGTATGGCGGGGCTTGGCG	GGTTCCACCACCACCCTACCGCAGCCCCAGCC
Cre05.g247200	GGAGGCGGAGGTATGGGGCGCTACACCTGG	GGTTCCACCACCACCCCCCACGCCATGGCCC
Cre05.g247400	GGAGGCGGAGGTATGCAAGCACTTCTCGTCCC	GGTTCCACCACCACCGTGCGCTGGCGGCTTC
Cre05.g247500	GGAGGCGGAGGTATGGCAACGAATGCCGCC	GGTTCCACCACCACCTCAGCTGCGCCGAGGC
Cre05.g247600	GGAGGCGGAGGTATGTCGGTGCCTGTGCCG	GGTTCCACCACCACCGAACATGGTGCCTTCGGGC
Cre05.g247700	GGAGGCGGAGGTATGGACAGAGGGAGGCTG	GGTTCCACCACCACCCGCCAGGTGCGCCTTT
Cre05.g247851	GGAGGCGGAGGTATGCAGGCCAAGTTTCTTGC	GGTTCCACCACCACCCTGGGGAGCAATGCGCTCAATG
Cre05.g247950	GGAGGCGGAGGTATGGCGCTCTCCAACTCT	GGTTCCACCACCACCCTTCTTCGCACCGGCC
Cre05.g248100	GGAGGCGGAGGTATGTCTAACCCGCTGGCACTA	GGTTCCACCACCACCGCTGAGCAGCTCCTTGAGC
Cre05.g248200	GGAGGCGGAGGTATGTTCGCTACCGTTCAGC	GGTTCCACCACCACCCCAGCGGCCCGATAGG
Cre05.g248300	GGAGGCGGAGGTATGGGGGGACCTGGAGC	GGTTCCACCACCACCCTATTGCCGGCGGCCG
Cre05.g248401	GGAGGCGGAGGTATGGCCCGCGGGTCTT	GGTTCCACCACCACCTCAGCAGTCGCCGTGC
Cre05.g248500	GGAGGCGGAGGTATGCTTTCGACCATCGGCC	GGTTCCACCACCACCCTTTGCGTGCTTCTGAATGAGG
Cre05.g248600	GGAGGCGGAGGTATGCTCGCATCTACAATCCGC	GGTTCCACCACCACCCTGCGCCAGCTTTGCCTCAAAG
Cre05.g248700	GGAGGCGGAGGTATGGCTGGCGAGGAGG	GGTTCCACCACCACCGGCGTGCCTTGCAGTT
Cre06.g248750	GGAGGCGGAGGTATGGGGCAAGCGCTAGCC	GGTTCCACCACCACCGTACCTCCGCACCAACATAAAA
Cre06.g248850	GGAGGCGGAGGTATGGAAGACGAAGCTAGGAAGC	GGTTCCACCACCACCATCGTCCTGCCGTTGGCC
Cre06.g248950	GGAGGCGGAGGTATGCTTGCGTTGCTCGGATT	GGTTCCACCACCACCATGCAGCCTCAACGGCAC
Cre06.g249101	GGAGGCGGAGGTATGCTGCTGGGGCACA	GGTTCCACCACCACCTCAGCCTGCGCGCTTG
Cre06.g249200	GGAGGCGGAGGTATGTCATCCAATATCTACGGCG	GGTTCCACCACCACCTCAGCCGATTATCAGCTCCC
Cre06.g249300	GGAGGCGGAGGTATGGTTGCGTCCGCCG	GGTTCCACCACCACCCGGGCACTTTCGGTTGAT
Cre06.g249350	GGAGGCGGAGGTATGTGCAATCTGCAGGCTC	GGTTCCACCACCACCCACGCCCTCCTCCTCC
Cre06.g249450	GGAGGCGGAGGTATGCCGAAGAGAACACCCA	GGTTCCACCACCACCGTGTGCGGCCTCGGAC
Cre06.g249500	GGAGGCGGAGGTATGCTCGGGATCCAGGG	GGTTCCACCACCACCACCACATGCCTCCTGCA
Cre06.g249555	GGAGGCGGAGGTATGCGGCTGCGAGGTGCC	GGTTCCACCACCACCCTGCAACGGCTCCTCAGC
Cre06.g249650	GGAGGCGGAGGTATGGCTGGGCCGGCTCCTG	GGTTCCACCACCACCATTGCACACCAAGAACAGCGAT
Cre06.g249700	GGAGGCGGAGGTATGCAGCTGCAAATCCTTCA	GGTTCCACCACCACCCGCCTTGTGCTCGCTC
Cre06.g249800	GGAGGCGGAGGTATGGAGGAAAAGGCGCCG	GGTTCCACCACCACCGCTGCGCTGCGAGAAGAA
Cre06.g249900	GGAGGCGGAGGTATGGCAAGCACCGCCAAAAA	GGTTCCACCACCACCCGTCGCCGCACTGGCTGTC
Cre06.g249950	GGAGGCGGAGGTATGTGCGGTGTGGGCACG	GGTTCCACCACCACCGGGCACAAACACGCGCTG
Cre06.g250050	GGAGGCGGAGGTATGATTGCACTGCGGCC	GGTTCCACCACCACCAACAGCGCCAAGAGCAG
Cre06.g250150	GGAGGCGGAGGTATGGTGCTGGAGCACACG	GGTTCCACCACCACCTCAGCCTGCGCGCTTG
Cre06.g250250	GGAGGCGGAGGTATGGTATATTGGCTGGTCTCCC	GGTTCCACCACCACCCGCGCTGTAGTCGGTGCTTAAG
Cre06.g250300	GGAGGCGGAGGTATGTCCAGCGATTCGCG	GGTTCCACCACCACCTACTGACAGGAACAGCGACA
Cre06.g250450	GGAGGCGGAGGTATGGTGGAGGTCGCGG	GGTTCCACCACCACCCCGGCTGGTCATTTGGAG
Cre06.g250550	GGAGGCGGAGGTATGGACTGGAATGAGGCATG	GGTTCCACCACCACCCCGCGTCAGACCTCGC
Cre06.g250600	GGAGGCGGAGGTATGGTGGCCCCCGACG	GGTTCCACCACCACCCCGTAGCACCTTGTACACCA
Cre06.g250700	GGAGGCGGAGGTATGATTGACCGCATTAAGTCAAAG	GGTTCCACCACCACCTCAGGGCGGGATGACG
Cre06.g250800	GGAGGCGGAGGTATGGAGACAATTATGTGCCCCTC	GGTTCCACCACCACCTTACGGCATGTCCGCTGG
Cre06.g250850	GGAGGCGGAGGTATGGCTGCCATCGAAAAGC	GGTTCCACCACCACCCTCGGCCTTACGCTTGC
Cre06.g250976	GGAGGCGGAGGTATGCGGCACGCGCTGA	GGTTCCACCACCACCTGCTCGCCAGCAAAGACC
Cre06.g251100	GGAGGCGGAGGTATGGACAGCTTTAAAGGACAAGC	GGTTCCACCACCACCCATTAGCACCAGCCCCACG
Cre06.g251200	GGAGGCGGAGGTATGGATGATGACGCACCAAAGG	GGTTCCACCACCACCCAGCTTGATGGGCATGGGCATG
Cre06.g251450	GGAGGCGGAGGTATGCGCCTATTCGCAAACA	GGTTCCACCACCACCCTACCGGGCGGTGTGG
Cre06.g251550	GGAGGCGGAGGTATGGCCACTGGGTTCTCA	GGTTCCACCACCACCCACCTGTGTCACCAGCG
Cre06.g251650	GGAGGCGGAGGTATGAAGTTTACGCACCAGCTGA	GGTTCCACCACCACCGAAGCCCACCGCCAGCAT
Cre06.g251716	GGAGGCGGAGGTATGCGCCGAAGTGCCA	GGTTCCACCACCACCATCCAGCGGCCGTGAT
Cre06.g251850	GGAGGCGGAGGTATGACGGCAAAAGATGTGCT	GGTTCCACCACCACCATCCTGGCTCACTGCGG
Cre06.g251951	GGAGGCGGAGGTATGTCACAGGCACTTGCTG	GGTTCCACCACCACCCTAGCGGAAACCCGGCG
Cre06.g252050	GGAGGCGGAGGTATGGCTGAAGCGCAGGA	GGTTCCACCACCACCGGGCATCAGGTGCTTGC
Cre06.g252200	GGAGGCGGAGGTATGGCGCAGCCTCCTCGG	GGTTCCACCACCACCGTCGTCGTCGTACATCTGCTC
Cre06.g252300	GGAGGCGGAGGTATGCAAAGCATTGCACATCGC	GGTTCCACCACCACCGTAAGGAGCCATGTTAATGGGC
Cre06.g252401	GGAGGCGGAGGTATGGCCAGCGCATTCGAG	GGTTCCACCACCACCATTCCGGTTGGCCTCGTTG
Cre06.g252550	GGAGGCGGAGGTATGCAGTTGCCCGGGGTTCTAC	GGTTCCACCACCACCGCTGCTGGTGGCGCTCAG
Cre06.g252600	GGAGGCGGAGGTATGAACCCGCGTGGCG	GGTTCCACCACCACCTACTCCAACCTCGGCATCC
Cre06.g252700	GGAGGCGGAGGTATGTCATTCATCGCGGCTCG	GGTTCCACCACCACCAAACCCGGGTGGTGGCGG
Cre06.g252750	GGAGGCGGAGGTATGTATCGTTTTCAGGTGTACTT	GGTTCCACCACCACCCTATGGCTTGTGTCTGAGC
Cre06.g252850	GGAGGCGGAGGTATGTCAGAGCGGCCGC	GGTTCCACCACCACCGTACCGGTCCTGTTGACCT
Cre06.g252871	GGAGGCGGAGGTATGCAGCTGAGCGACCG	GGTTCCACCACCACCTCCTGCCGCAGGGCAG
Cre06.g252913	GGAGGCGGAGGTATGGACTTGCCCCCGCC	GGTTCCACCACCACCCACGCGGCCCTCCCCT
Cre06.g253051	GGAGGCGGAGGTATGGACCCAGGCAAACCATGG	GGTTCCACCACCACCGGACGCGACGGCGTGG
Cre06.g253150	GGAGGCGGAGGTATGCTGGACGTCTTCCGC	GGTTCCACCACCACCGCTATACGTGCGCTGCC
Cre06.g253350	GGAGGCGGAGGTATGGCCCTTCGCGCATTC	GGTTCCACCACCACCGTGCTTCTCGCATTCCTTGG
Cre06.g253550	GGAGGCGGAGGTATGGTGCCCCGTCAGAC	GGTTCCACCACCACCGACGAGATGACGGGTCATCT
Cre06.g253600	GGAGGCGGAGGTATGGCTGTTGCCTGCGC	GGTTCCACCACCACCTCAGCGGGGTTGCCGG
Cre06.g253750	GGAGGCGGAGGTATGGCTGTTGCCTGCGC	GGTTCCACCACCACCTCAGCGGGGTTGCCGG
Cre06.g253754	GGAGGCGGAGGTATGCTTAAAGAGCAGATCAGCC	GGTTCCACCACCACCCTATCCACGCAGCCCTCC
Cre06.g253758	GGAGGCGGAGGTATGTCCGAGTCCTACTACTATGC	GGTTCCACCACCACCCTAGCCGGTGTCTTGCCC
Cre06.g253789	GGAGGCGGAGGTATGGTGTCCGGTAAGATTTGG	GGTTCCACCACCACCGCCCACCTTGCTGGTAAC
Cre06.g253850	GGAGGCGGAGGTATGTCCTGGGCATTCGGC	GGTTCCACCACCACCGACGAGATGACGGGTCATCT
Cre06.g253852	GGAGGCGGAGGTATGGCGTGTAAGCTTCAGG	GGTTCCACCACCACCTTATATAATTGCCGCCACTAGTT
Cre06.g253903	GGAGGCGGAGGTATGAAGGCGTGGAAGCATG	GGTTCCACCACCACCGGTCGTCCATAGGCGCG
Cre06.g254002	GGAGGCGGAGGTATGCTGCTTTATACACCCCG	GGTTCCACCACCACCTTATATAATTGCCGCCACTAGTT
Cre06.g254100	GGAGGCGGAGGTATGCCGCCGACACGCCTG	GGTTCCACCACCACCGTTCTCACGCAGGCCGCGCTTA
Cre06.g254175	GGAGGCGGAGGTATGGCTGATGGGGCCTG	GGTTCCACCACCACCGTCGTCCCCGACACAGC
Cre06.g254213	GGAGGCGGAGGTATGGCTTCGCGATGCTTCTG	GGTTCCACCACCACCGTCGGCCTGCGCCAGC
Cre06.g254232	GGAGGCGGAGGTATGAAGTCCGAGTTTGCGTC	GGTTCCACCACCACCCTTGGTCAGTCGCCTGGT
Cre06.g254250	GGAGGCGGAGGTATGCCCGTGCCCGGTC	GGTTCCACCACCACCTCAGGGAAAGTTCAGCATTTTTCC
Cre06.g254300	GGAGGCGGAGGTATGATGAAGACTTCCACCATGAA	GGTTCCACCACCACCTCACAGGAGCCACAGCA
Cre06.g254400	GGAGGCGGAGGTATGGAGCGGCTGGCGG	GGTTCCACCACCACCCACCTGCCACTTCTGGAAGA
Cre06.g254500	GGAGGCGGAGGTATGGCGGCTCGGCTGGGAC	GGTTCCACCACCACCCGGCGGTGAGTCCCACGG
Cre06.g254550	GGAGGCGGAGGTATGGATGTCCAGCAACCGA	GGTTCCACCACCACCCAGCGGCTGAGCGAGTA
Cre06.g254650	GGAGGCGGAGGTATGAAGGACGACGCGGCAG	GGTTCCACCACCACCCCCCACCACGCCCATGCG
Cre06.g255050	GGAGGCGGAGGTATGGGAGCCTCGGCTTCG	GGTTCCACCACCACCGTACACGCCCAGCCACCG
Cre06.g255150	GGAGGCGGAGGTATGCGCGCCAGTGGGG	GGTTCCACCACCACCATAAATGCTTGGGCTGTCGCT
Cre06.g255250	GGAGGCGGAGGTATGGGCAAGTCCGACAAAG	GGTTCCACCACCACCCGTCCGCTTGGCAATGG

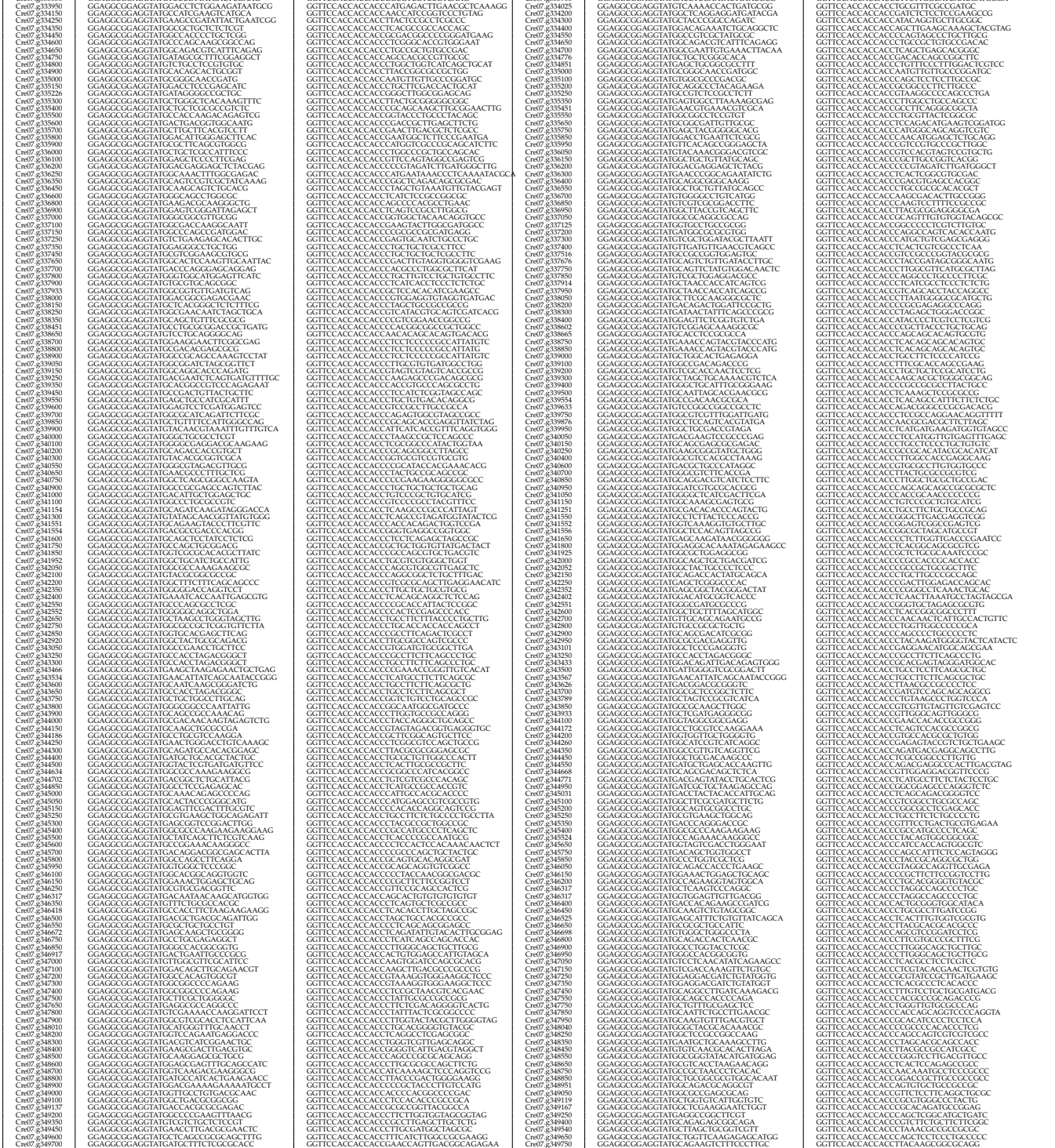


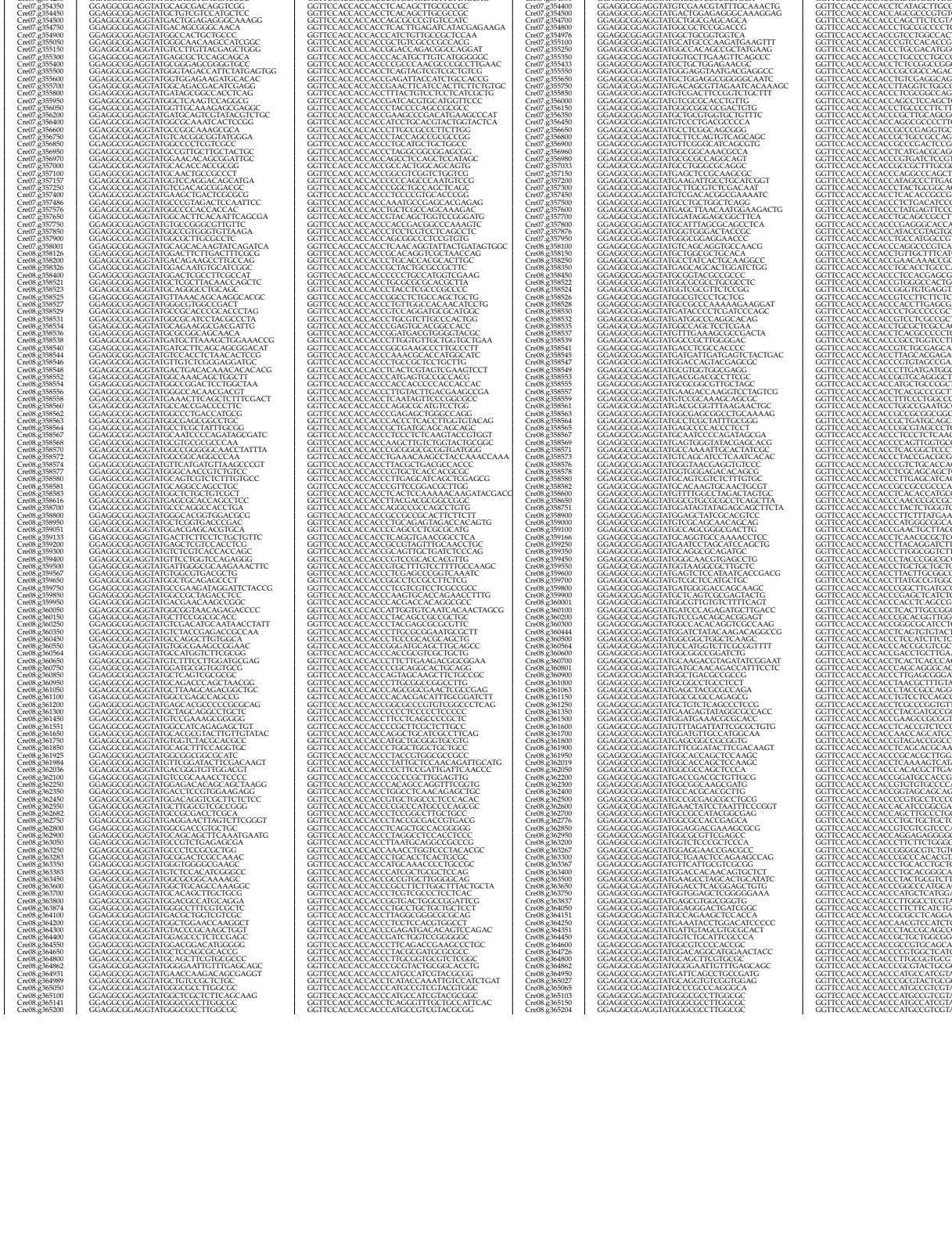


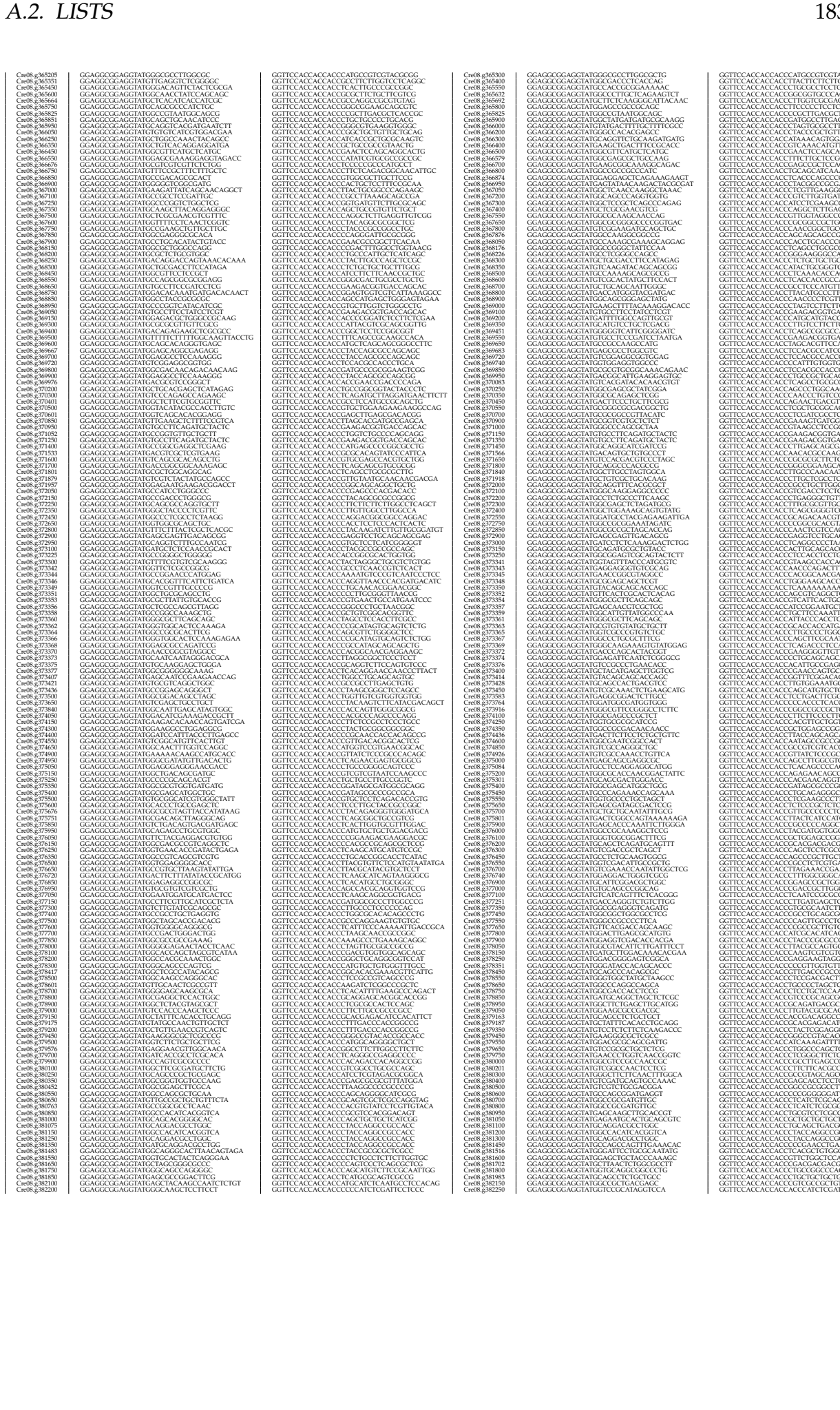


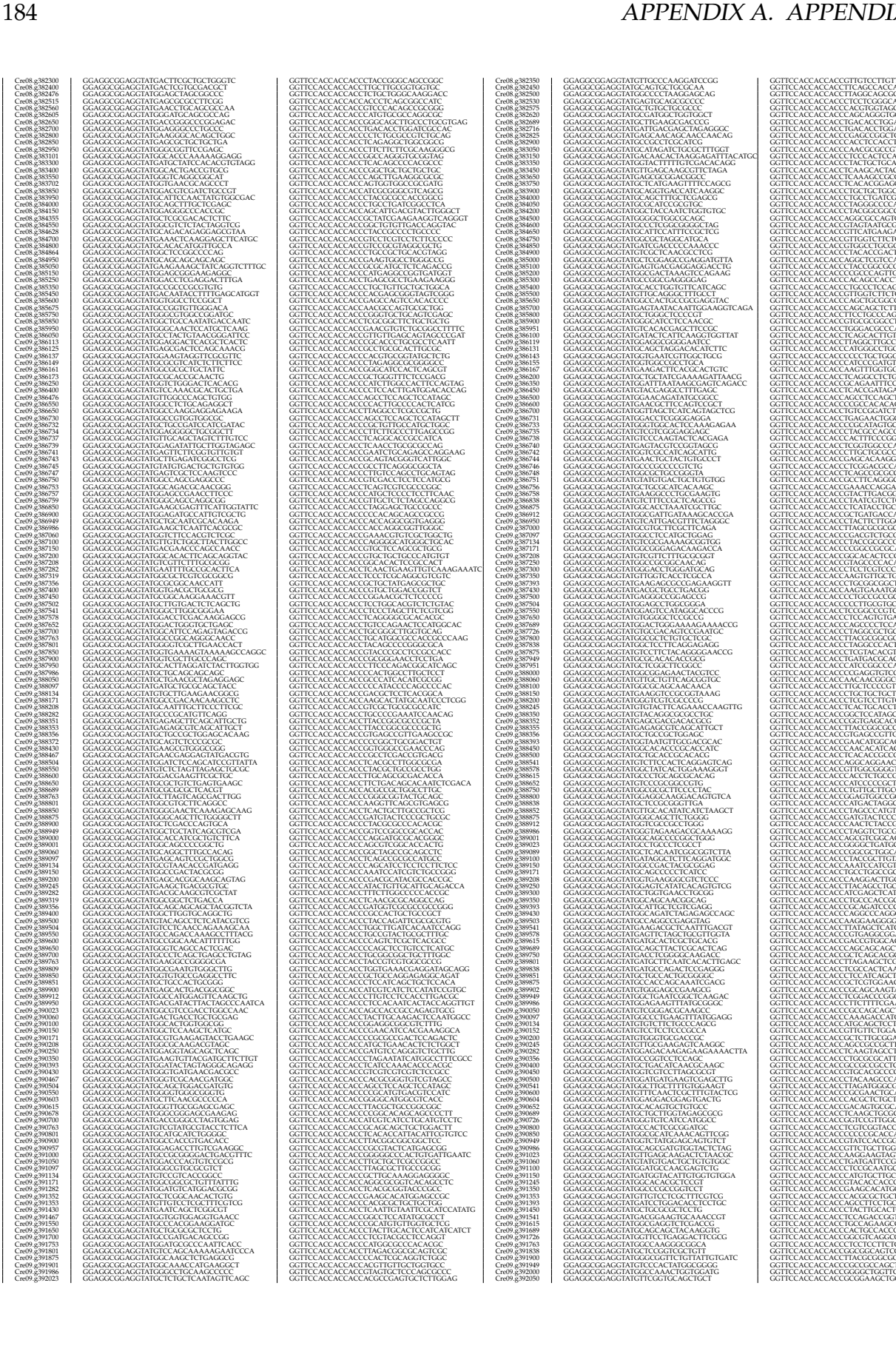






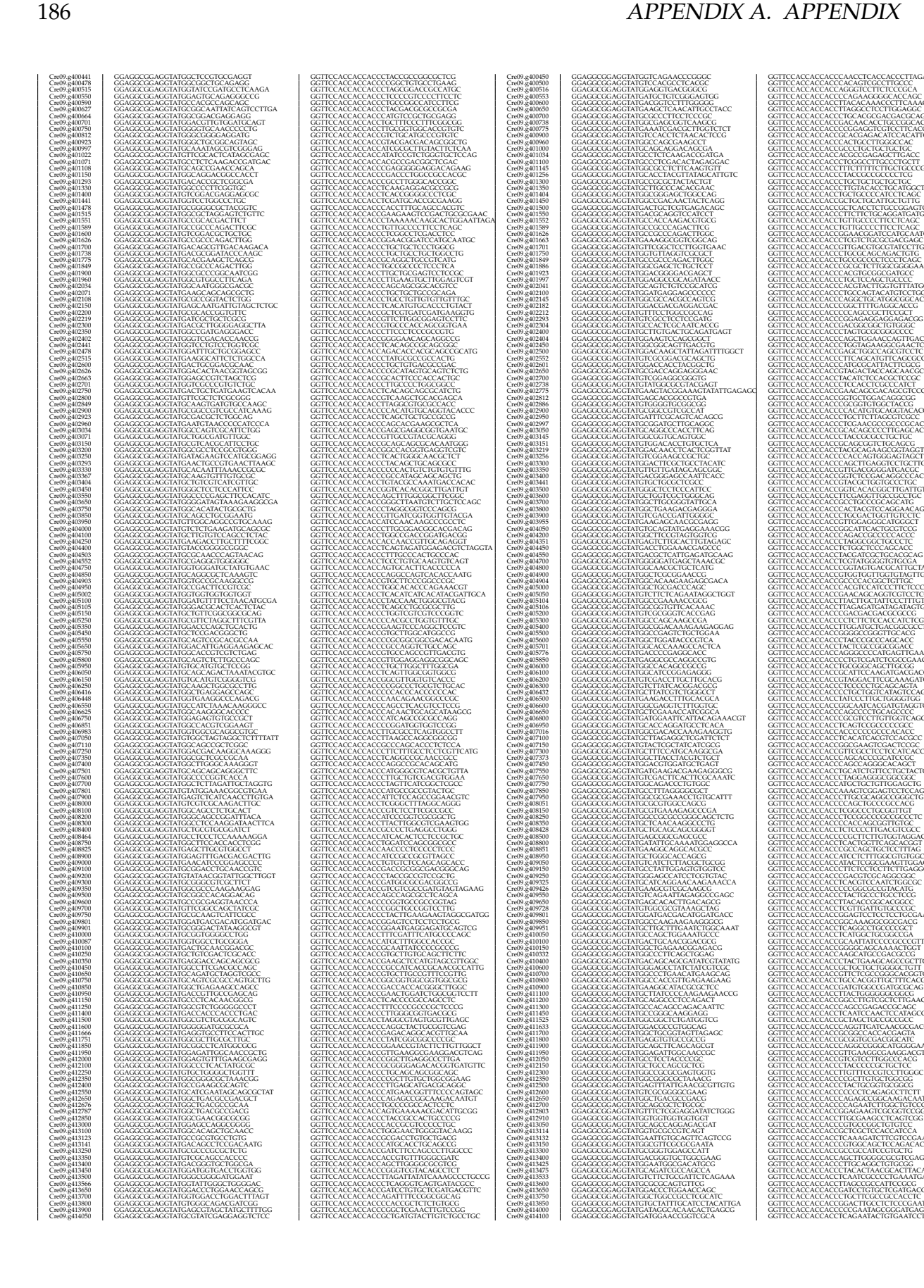


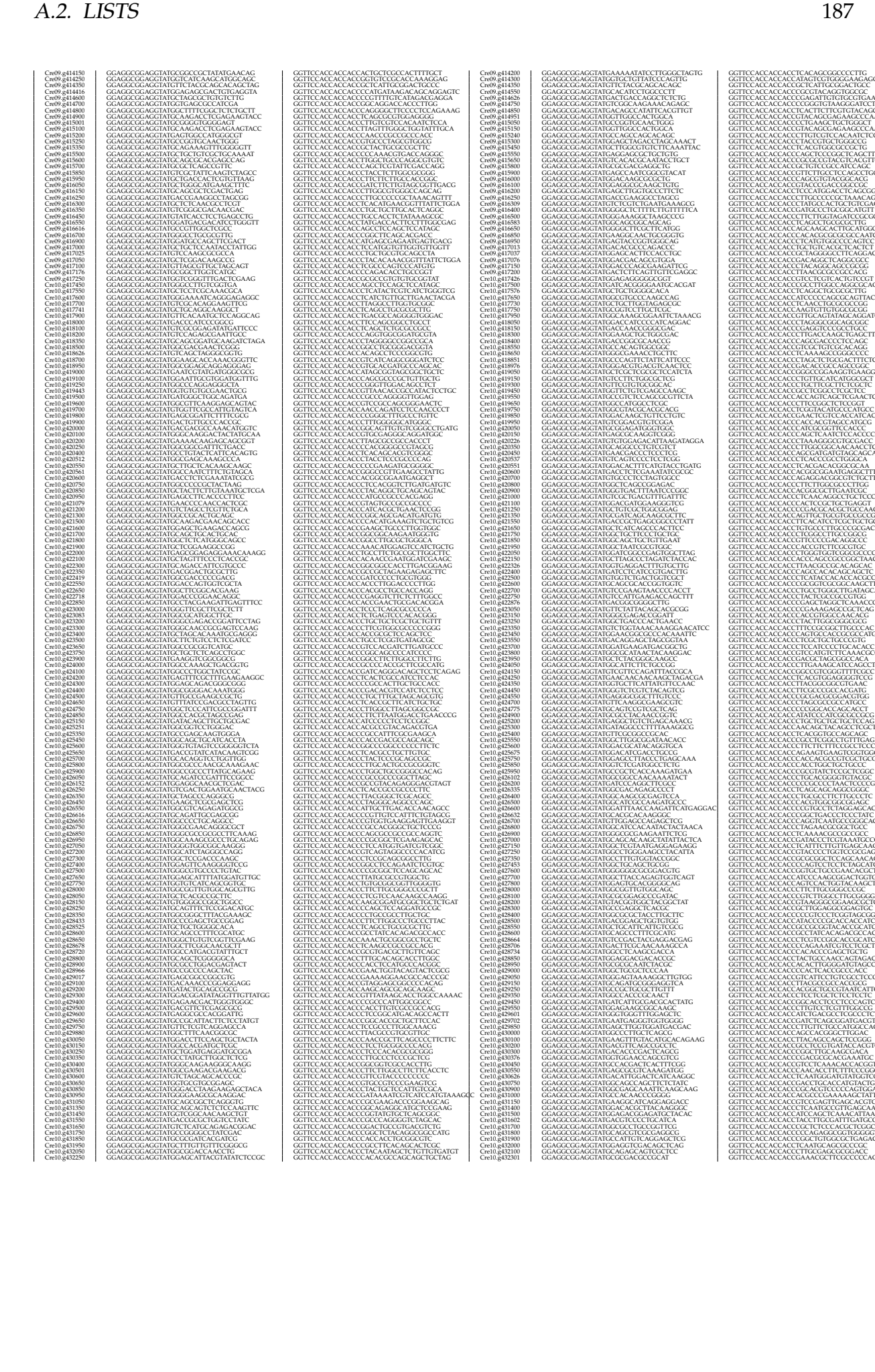


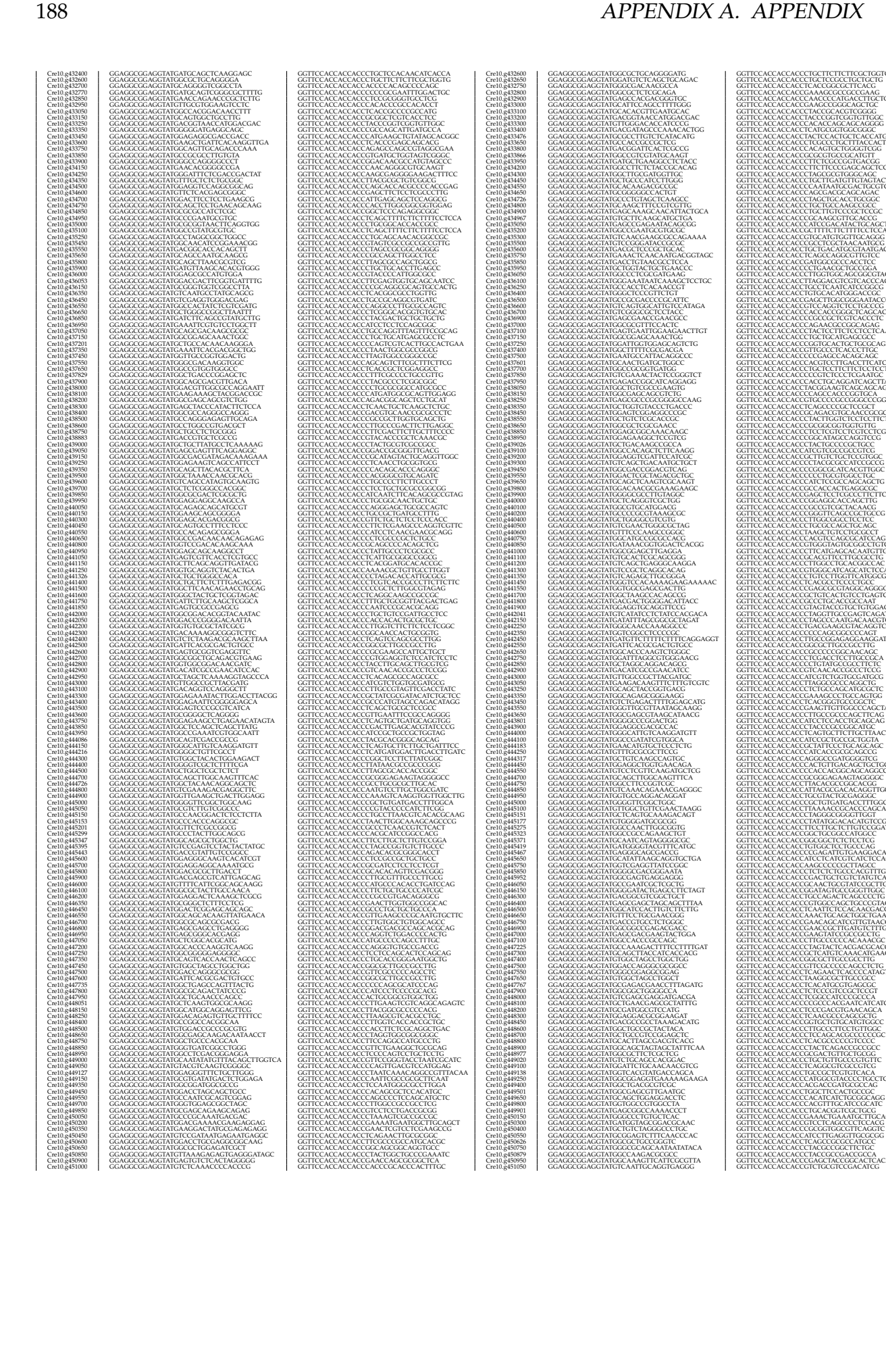


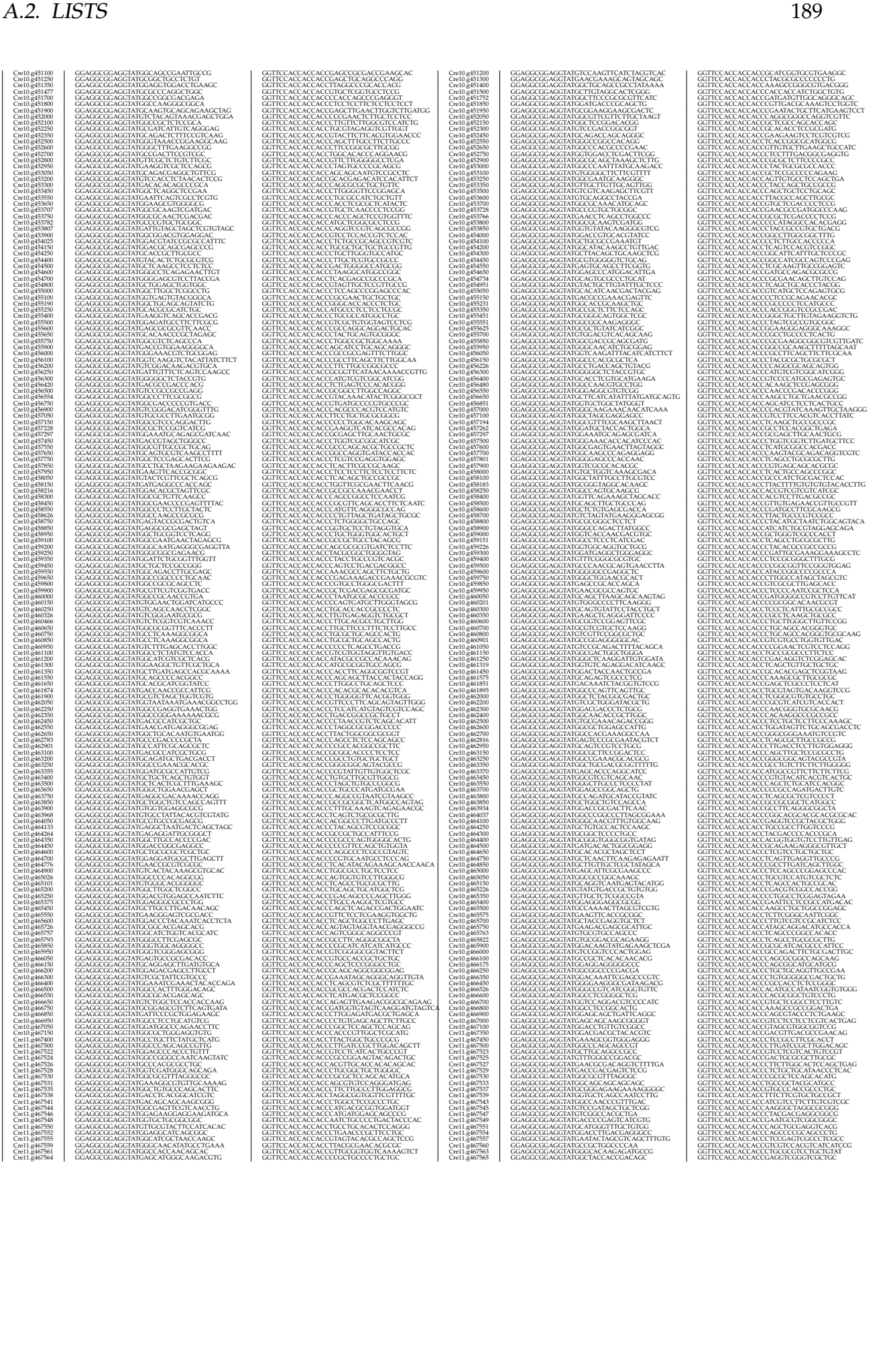
Cre09.g392060	GGAGGCGGAGGTATGGCCGCGTTGCTGGGTC	GGTTCCACCACCACCCCCCTCCACTAGCAGCGT
Cre09.g392105	GGAGGCGGAGGTATGTGGCACGCGGCAG	GGTTCCACCACCACCCTTCACCCCGCCTACCG
Cre09.g392134	GGAGGCGGAGGTATGCTGACATACAGCCTGCC	GGTTCCACCACCACCCTGCTGCTGGCCCTCCTG
Cre09.g392208	GGAGGCGGAGGTATGACGCGTATGCAGGC	GGTTCCACCACCACCCAGCCCCATGCTGGCC
Cre09.g392251	GGAGGCGGAGGTATGGCTGGCAACGGAGGG	GGTTCCACCACCACCTCAGCAGCTGGGCGCG
Cre09.g392282	GGAGGCGGAGGTATGCCGGGCGTAGCAG	GGTTCCACCACCACCAAACGCGAGCCCAGTGA
Cre09.g392319	GGAGGCGGAGGTATGGCCAAAAAGCAAAATTACCA	GGTTCCACCACCACCCTTGCAAGCACACGTGTAGG
Cre09.g392356	GGAGGCGGAGGTATGAGAGGGGTAAAGCGGGAG	GGTTCCACCACCACCGCCCTTCATCTTCTTCTCGAGC
Cre09.g392400	GGAGGCGGAGGTATGACCAATACGTATTCTTGGCA	GGTTCCACCACCACCTCACATGAGGTCCTCTGGC
Cre09.g392467	GGAGGCGGAGGTATGTCTACCGAGGTGGTCCTTA	GGTTCCACCACCACCCGAGGACACGAGCTCGGC
Cre09.g392503	GGAGGCGGAGGTATGAAGGTGGACTTCACACTC	GGTTCCACCACCACCAAGCGCATAGCAGTGCG
Cre09.g392505	GGAGGCGGAGGTATGATGCACGGTAGCCTGTAC	GGTTCCACCACCACCCTCAGCGCCCTCCTCAAC
Cre09.g392542	GGAGGCGGAGGTATGGGAAGGCGACGGAG	GGTTCCACCACCACCGCTCCCCACCGCCATC
Cre09.g392616	GGAGGCGGAGGTATGCGTGCCGCTCTCG	GGTTCCACCACCACCGTTCTGCGGCCCAGCG
Cre09.g392652	GGAGGCGGAGGTATGTTCGCCCGCAGCG	GGTTCCACCACCACCTCACGCATCCAGTGCAAG
Cre09.g392654	GGAGGCGGAGGTATGTATGCTGCAGCCTCTCG	GGTTCCACCACCACCAATAATGCCGGCACACGT
Cre09.g392700	GGAGGCGGAGGTATGGTTGCGGCCAGGG	GGTTCCACCACCACCCTTCCCACTGCCGCGC
Cre09.g392750	GGAGGCGGAGGTATGGCAAAAGTCTGCGATTCC	GGTTCCACCACCACCGTGCTCGTTGACGCCACC
Cre09.g392800	GGAGGCGGAGGTATGTCGGGTTCCTACGCTA	GGTTCCACCACCACCTCAGCGGAAGGGCCTC
Cre09.g392840	GGAGGCGGAGGTATGCTGTCGCTGCACAAG	GGTTCCACCACCACCCGCTGTAGTCACTTCACGC
Cre09.g392950	GGAGGCGGAGGTATGGATACGGTTGCAGAGGA	GGTTCCACCACCACCCTACATCGCGCCGCCG
Cre09.g392988	GGAGGCGGAGGTATGGACCCCGCTGCTGAA	GGTTCCACCACCACCCGCGGCAAACTTGGGGTT
Cre09.g393025	GGAGGCGGAGGTATGGCGGCTTTGGACTG	GGTTCCACCACCACCTGGTTCCTTGTGTCGCG
Cre09.g393062	GGAGGCGGAGGTATGACGCTGTGCCACCTG	GGTTCCACCACCACCCAGCTGCCGCCGCATCTG
Cre09.g393099	GGAGGCGGAGGTATGGTGCTACGCACTTCCG	GGTTCCACCACCACCGGCCTGAATGCTGCCC
Cre09.g393136	GGAGGCGGAGGTATGGATCCGATGCAAAAGCA	GGTTCCACCACCACCTCAGAACAGCGCGCCG
Cre09.g393173	GGAGGCGGAGGTATGCAGTCCCTTGCATCGC	GGTTCCACCACCACCGAAGAAAGCCTCGCCGCCCTTA
Cre09.g393210	GGAGGCGGAGGTATGTTTCGTTCAGCAGCGC	GGTTCCACCACCACCCAAACCATCGTCATCTTCCATCC
Cre09.g393250	GGAGGCGGAGGTATGGAGCATCGACTGCTTGC	GGTTCCACCACCACCCTTGGCCCCGGGCTCCTG
Cre09.g393300	GGAGGCGGAGGTATGACTACTCTTGTGCTCGTGGC	GGTTCCACCACCACCTCACTCCCCGGCCGCC
Cre09.g393350	GGAGGCGGAGGTATGTCAGACGGTCGCCTC	GGTTCCACCACCACCCAACGCCTCCGCCACG
Cre09.g393395	GGAGGCGGAGGTATGGAGCAGGCGAGCG	GGTTCCACCACCACCTCCTAATGGCGAGCCGG
Cre09.g393432	GGAGGCGGAGGTATGCAGTGCCGTGCGGAC	GGTTCCACCACCACCGACCAGAACGCTATCGGGTC
Cre09.g393550	GGAGGCGGAGGTATGGCTCCAAAGAAGCGCAAG	GGTTCCACCACCACCCTCCCAGCCCCGCTCTGC
Cre09.g393600	GGAGGCGGAGGTATGGCTGACGTGCTGGG	GGTTCCACCACCACCGATGAGCACTGCAATGTCCT
Cre09.g393650	GGAGGCGGAGGTATGGGAAATTGCGTCTCCG	GGTTCCACCACCACCCTCCAGGAAGCCCGCG
Cre09.g393700	GGAGGCGGAGGTATGTCAAATCTGCGTGCGC	GGTTCCACCACCACCCTGGCAGTCGCCGTCC
Cre09.g393750	GGAGGCGGAGGTATGAGCGGCAACGACCGG	GGTTCCACCACCACCGCACCACCCTGCGCCTCC
Cre09.g393802	GGAGGCGGAGGTATGCTCAGAGCCGCTTCG	GGTTCCACCACCACCCTTACTGCCCTGCTTGGC
Cre09.g393876	GGAGGCGGAGGTATGCAAAGTTCTTTGCGCGC	GGTTCCACCACCACCCTCCTGCTTCCCTGCCGTC
Cre09.g393913	GGAGGCGGAGGTATGGCACCGTGTAGAGACA	GGTTCCACCACCACCCTGCTTCTCCCGGAAAAGC
Cre09.g393953	GGAGGCGGAGGTATGTCCCGCTACGCCG	GGTTCCACCACCACCCTTCTGCCACTGGTACATGAA
Cre09.g393991	GGAGGCGGAGGTATGGACGCGCAACGGC	GGTTCCACCACCACCGCCTTGCCTTTCTCTACCTGG
Cre09.g394050	GGAGGCGGAGGTATGCCCCCTCTTCAGAGC	GGTTCCACCACCACCCACCTGTGACGAGGACGA
Cre09.g394102	GGAGGCGGAGGTATGGGCGCGGAGCAGAGC	GGTTCCACCACCACCGATGAGCTGCTCAGTGGGG
Cre09.g394176	GGAGGCGGAGGTATGGAGGAGGCAAGTGTAGC	GGTTCCACCACCACCTTCGCACAATGCCTGGC
Cre09.g394200	GGAGGCGGAGGTATGGGGCGTCTTTCAGTCG	GGTTCCACCACCACCCGGGCAGCCAGAGAACG
Cre09.g394251	GGAGGCGGAGGTATGGCGACCCGAGCAGTTC	GGTTCCACCACCACCGTACGCCGTAGTGGTTGCG
Cre09.g394325	GGAGGCGGAGGTATGGCCCTCTCGATGACTCT	GGTTCCACCACCACCGATGAGGGGCTTGCCCATG
Cre09.g394362	GGAGGCGGAGGTATGGCCCCAAAATCAGACC	GGTTCCACCACCACCTCGGCTCCGCTGCATC
Cre09.g394436	GGAGGCGGAGGTATGGCCGCTCCGGACG	GGTTCCACCACCACCCGCAATGCCGAAGAAGGAG
Cre09.g394473	GGAGGCGGAGGTATGGCTGCGTCAATGAAATGG	GGTTCCACCACCACCCTTGGCCAACTGCATCTCG
Cre09.g394547	GGAGGCGGAGGTATGCTTCTCAAATGCTCACGA	GGTTCCACCACCACCTTAGCGCCAGCTAGCACG
Cre09.g394584	GGAGGCGGAGGTATGGCGTTGCTAGCAGACAAAA	GGTTCCACCACCACCCAGCGGCTTCTTGTTCGG
Cre09.g394621	GGAGGCGGAGGTATGGCGTCCACTGGCAGT	GGTTCCACCACCACCCAGTGCGTCACGCCAGTT
Cre09.g394695	GGAGGCGGAGGTATGGCGAGCCCTGGTGC	GGTTCCACCACCACCCTACGCCGGCAGCCCG
Cre09.g394700	GGAGGCGGAGGTATGCGGCCCGTCAGTG	GGTTCCACCACCACCCTCCTCCTGCGGCTCC
Cre09.g394750	GGAGGCGGAGGTATGCGCCTCTCTGCTCGC	GGTTCCACCACCACCGTAGACGAAGGGCTGGCC
Cre09.g394806	GGAGGCGGAGGTATGCTATGTCGAAATACAGCAGT	GGTTCCACCACCACCTGCACATGAGGCGTCCCAAATC
Cre09.g394900	GGAGGCGGAGGTATGCCAACAAGTTCGGCG	GGTTCCACCACCACCCGGCGGCTCCAATGCG
Cre09.g394954	GGAGGCGGAGGTATGCCTTGCCGCCTTTC	GGTTCCACCACCACCTGCAAAGTTGGGCTTGCC
Cre09.g395028	GGAGGCGGAGGTATGCGCGCGAGTTGGGTG	GGTTCCACCACCACCGCCCTGCATGGGCCCGCG
Cre09.g395065	GGAGGCGGAGGTATGACACACCTGCGGTCG	GGTTCCACCACCACCGTCACGGAGCAATGCCA
Cre09.g395102	GGAGGCGGAGGTATGCAAGACTTGCAGGACTG	GGTTCCACCACCACCTGGCAGCTGCGTCTGC
Cre09.g395150	GGAGGCGGAGGTATGGAGGTAGAGCAGTACTGC	GGTTCCACCACCACCGTGCACCACCTGCGCC
Cre09.g395200	GGAGGCGGAGGTATGCGCCAGCTCGTGGT	GGTTCCACCACCACCCGCCACCTTGCCCGCC
Cre09.g395250	GGAGGCGGAGGTATGTATATCAACCCGTTGATGGG	GGTTCCACCACCACCGTTTGCTCCCATGATGTTTTGC
Cre09.g395288	GGAGGCGGAGGTATGCCATGTTCTCGCGCG	GGTTCCACCACCACCAGACACCAACACTGCCGC
Cre09.g395325	GGAGGCGGAGGTATGGCAAAGAAAAAATCCAAGACG	GGTTCCACCACCACCTCAGAAGCAGCACTCGCA
Cre09.g395362	GGAGGCGGAGGTATGACGGTCGGAGTTGGGA	GGTTCCACCACCACCCTTCTTCTTCGGATCCTGCTGC
Cre09.g395400	GGAGGCGGAGGTATGGCGTCGGACGCGGAC	GGTTCCACCACCACCGTAATGCTGCACTCCCTGCT
Cre09.g395436	GGAGGCGGAGGTATGTCGCTCTGGGGCTGC	GGTTCCACCACCACCCGATGCGCTGTAGTACCCC
Cre09.g395473	GGAGGCGGAGGTATGTTCAAGGTGCTTGCTTTG	GGTTCCACCACCACCCTAGGAATCCTCTGCGGCT
Cre09.g395547	GGAGGCGGAGGTATGGCTCTCGCCTCGG	GGTTCCACCACCACCCACGAGCCGCCCGAAG
Cre09.g395621	GGAGGCGGAGGTATGGAGCAACTGCGCACG	GGTTCCACCACCACCCCCGCATGGCCGCAGCTTG
Cre09.g395650	GGAGGCGGAGGTATGAAGACTATCGCCCACAAGT	GGTTCCACCACCACCGCTGAGGCGGACCATGAC
Cre09.g395658	GGAGGCGGAGGTATGCAGGCCCACGGAC	GGTTCCACCACCACCCTCCGCAGGCAGGCTT
Cre09.g395732	GGAGGCGGAGGTATGAGGGCGTGGACGC	GGTTCCACCACCACCAGGCGAAAAGAACGCCC
Cre09.g395769	GGAGGCGGAGGTATGGAGGGGGGCCTCT	GGTTCCACCACCACCGGAGTCGTATGTGGCCATG
Cre09.g395806	GGAGGCGGAGGTATGTGGGCCCGGGGCACGAG	GGTTCCACCACCACCCCGGCCACCCGCCACCTC
Cre09.g395850	GGAGGCGGAGGTATGCCGCGTCCACTCTT	GGTTCCACCACCACCCCAAAGCTCCTGCCGGC
Cre09.g395900	GGAGGCGGAGGTATGCAGTCCCACGCAGC	GGTTCCACCACCACCCTCGTCCTCGCCCTCG
Cre09.g395925	GGAGGCGGAGGTATGCTTCGTCACCAGCGG	GGTTCCACCACCACCGTGCTGCTGCCGCCAC
Cre09.g395954	GGAGGCGGAGGTATGAACGTTCTCCGCGC	GGTTCCACCACCACCTGGCGCTGTTCTCGCC
Cre09.g396000	GGAGGCGGAGGTATGTCTACAGGCGTCGACTACT	GGTTCCACCACCACCCGACACCTGCACACGCGC
Cre09.g396050	GGAGGCGGAGGTATGACTCGACAGCAATTATTCACC	GGTTCCACCACCACCCTACAGCCGCACGAACTG
Cre09.g396100	GGAGGCGGAGGTATGGCGCTGCTATGCCT	GGTTCCACCACCACCCACCTTGTAGACGGGGCA
Cre09.g396139	GGAGGCGGAGGTATGCTCGCTCCCTTGACG	GGTTCCACCACCACCAACCACCATCTCCACATCCTC
Cre09.g396176	GGAGGCGGAGGTATGTCGAGCGCATGCAC	GGTTCCACCACCACCTTGTTGGCCCACGCATG
Cre09.g396213	GGAGGCGGAGGTATGGCCCTCCGCGCTG	GGTTCCACCACCACCCTTCAGCTGGGCGTACCAC
Cre09.g396252	GGAGGCGGAGGTATGAACGTGGCCATGAACCTC	GGTTCCACCACCACCGGCGGGCTTGCGGACC
Cre09.g396300	GGAGGCGGAGGTATGATGTTGACCCAGACTCCTG	GGTTCCACCACCACCGGCCTTGACTGCGGCCTT
Cre09.g396350	GGAGGCGGAGGTATGTCCACACTCATTCTAGACGG	GGTTCCACCACCACCAGCGTCATGCAGCTCAGC
Cre09.g396400	GGAGGCGGAGGTATGCAGATTTTCGTCAAGACTCTC	GGTTCCACCACCACCTCAGAGCCCGCCACGC
Cre09.g396451	GGAGGCGGAGGTATGTTTCGAGAGCCGCGATTTA	GGTTCCACCACCACCACTGACGGGCGCCACGAC
Cre09.g396500	GGAGGCGGAGGTATGGCAGATAATACTAAGCGGA	GGTTCCACCACCACCGTTCAACTTGTAGTGGGCCATC
Cre09.g396549	GGAGGCGGAGGTATGACGCCCCTTCACCT	GGTTCCACCACCACCGCAAGACACAAAGCCCGC
Cre09.g396586	GGAGGCGGAGGTATGCGGCACGCGCTGA	GGTTCCACCACCACCCGCACATGCACGCTGC
Cre09.g396600	GGAGGCGGAGGTATGGCGCTTGGTCTTCTTGC	GGTTCCACCACCACCAGCATCGGCCTCGGCG
Cre09.g396650	GGAGGCGGAGGTATGTCTCTGAACAGTAGCACTATG	GGTTCCACCACCACCTTAAACGGCCGCGACG
Cre09.g396698	GGAGGCGGAGGTATGCATAGTAGGCGCTTATGC	GGTTCCACCACCACCGGCTTGTCGGGATGCG
Cre09.g396735	GGAGGCGGAGGTATGGCTACAGTCGCCGCT	GGTTCCACCACCACCTCTCCGCTTCACGGCCGTGATG
Cre09.g396800	GGAGGCGGAGGTATGCGTTTAATACGTGTGCCA	GGTTCCACCACCACCCATGGAGTACTCGGACGCG
Cre09.g396846	GGAGGCGGAGGTATGGCCGTCGTTGCTCGC	GGTTCCACCACCACCCATCCCACGCTCCGACGC
Cre09.g396883	GGAGGCGGAGGTATGCTCAGAGGCTGCGG	GGTTCCACCACCACCTGCAGCTGGTCCCAGG
Cre09.g396920	GGAGGCGGAGGTATGAAGAAGGCCGCCGT	GGTTCCACCACCACCGGCTACGGGCTGCTGG
Cre09.g396957	GGAGGCGGAGGTATGGAGCCGCGGTCGATTTC	GGTTCCACCACCACCCTGCTCCGACCGCAGCCG
Cre09.g397002	GGAGGCGGAGGTATGTCTAGGACGTTGACATGGA	GGTTCCACCACCACCTCACCACCGGGGCAGA
Cre09.g397052	GGAGGCGGAGGTATGCTGCTTGGTAGAGCGC	GGTTCCACCACCACCTCAACCTGCGCGCCGG
Cre09.g397095	GGAGGCGGAGGTATGGCCCTGCTTGACCGC	GGTTCCACCACCACCGCCTGTGCGGCCCTCC
Cre09.g397100	GGAGGCGGAGGTATGGACTCTGCAACTCAGCACC	GGTTCCACCACCACCTCAGCGCCCCCGGTCC
Cre09.g397150	GGAGGCGGAGGTATGCTCGCCAGTGCGCCT	GGTTCCACCACCACCCGGAAAGTCCAACTCAAAGTCC
Cre09.g397200	GGAGGCGGAGGTATGTCGTCTGTCAAGACGCTC	GGTTCCACCACCACCCTGCTTGCGCATGTTGATGC
Cre09.g397250	GGAGGCGGAGGTATGGCGTCTCTGCAGCT	GGTTCCACCACCACCTTATGCAGCCGCCTCGG
Cre09.g397290	GGAGGCGGAGGTATGTCAGCTCCTGCTGCAA	GGTTCCACCACCACCCGCCAACCGTCGCAGC
Cre09.g397327	GGAGGCGGAGGTATGGAAGCGTTTCGCTCAA	GGTTCCACCACCACCCAGGTAGGGCTTCCACCC
Cre09.g397364	GGAGGCGGAGGTATGGCTGAAGCGCCCG	GGTTCCACCACCACCAAGTGTAAGCGCATGCTTCA
Cre09.g397401	GGAGGCGGAGGTATGCACTCCAACGATGTTGC	GGTTCCACCACCACCGGCACCTACCTTCCGAGC
Cre09.g397450	GGAGGCGGAGGTATGCCTCTGCGCGAGG	GGTTCCACCACCACCCTTTCGTTGGGTATGCCGG
Cre09.g397500	GGAGGCGGAGGTATGAAAGGTTGGCGGCC	GGTTCCACCACCACCGGGATAGCTCATGCAACATGC
Cre09.g397586	GGAGGCGGAGGTATGGCAGAAGCAGACATATGG	GGTTCCACCACCACCTGCATCTGAAGAGGGAGAGAG
Cre09.g397660	GGAGGCGGAGGTATGGCTCGCCAGCTTTC	GGTTCCACCACCACCTTAGTTGCCGGCGGCG
Cre09.g397734	GGAGGCGGAGGTATGTACGCAGCGTTGGCG	GGTTCCACCACCACCGCGCTCCGTCAGCGTCAG
Cre09.g397808	GGAGGCGGAGGTATGCTGGGCTGGAAAGAAG	GGTTCCACCACCACCGTTGCCACTCAGCTGCG
Cre09.g397850	GGAGGCGGAGGTATGCTGCCTTGCACCCT	GGTTCCACCACCACCCGCAGGGCCCACCTTG
Cre09.g397919	GGAGGCGGAGGTATGCTCGCGTGGCAGGG	GGTTCCACCACCACCCTACTGGCCGCCGCCG
Cre09.g397993	GGAGGCGGAGGTATGCAGGTGTTTGCGCT	GGTTCCACCACCACCCTCCACCTCCTGCAGTGC
Cre09.g398104	GGAGGCGGAGGTATGGGGAAAGGCGAACGC	GGTTCCACCACCACCGACGACGCGCTTCCGCTT
Cre09.g398200	GGAGGCGGAGGTATGGGCAGTGGTCGAGATA	GGTTCCACCACCACCCTTCTTGCTCGGCCCCG
Cre09.g398250	GGAGGCGGAGGTATGGGTAATTGCCAGGTCA	GGTTCCACCACCACCCTAGTTGTAGGCATAATTTGCAG
Cre09.g398289	GGAGGCGGAGGTATGGCGCGTAAAAGCAGTTTG	GGTTCCACCACCACCCTGCTCATCCGGCGCCATCTC
Cre09.g398326	GGAGGCGGAGGTATGGAGCTCCTTAGCATTCTGA	GGTTCCACCACCACCTCACACCACGAAGCCAGG
Cre09.g398363	GGAGGCGGAGGTATGAGGGCTGCAGGCG	GGTTCCACCACCACCATGCTGAGCTTCTCGGCA
Cre09.g398401	GGAGGCGGAGGTATGGAGCTCCTTAGCATTCTGA	GGTTCCACCACCACCTCACACCACGAAGCCAGG
Cre09.g398450	GGAGGCGGAGGTATGAGCGCTCAGATGCAGATG	GGTTCCACCACCACCGCCGTCGCGGTAGTACAG
Cre09.g398500	GGAGGCGGAGGTATGGGCCAGACGACGCGACAAC	GGTTCCACCACCACCCTGCTGCTGCGGCTGGGG
Cre09.g398549	GGAGGCGGAGGTATGGGGGTGTCGCGCG	GGTTCCACCACCACCCGCCGCCGCTACACTG
Cre09.g398586	GGAGGCGGAGGTATGGCCGAGCCCACGG	GGTTCCACCACCACCCACCACGAAGCCAGGCA
Cre09.g398650	GGAGGCGGAGGTATGGGCAAGACGGACAAAAAAG	GGTTCCACCACCACCACGCCTGTAGCGTGTGCG
Cre09.g398660	GGAGGCGGAGGTATGGAGCTCCTTAGCATTATGATT	GGTTCCACCACCACCTTACTGTTTCTTACACCAAGTGCG
Cre09.g398702	GGAGGCGGAGGTATGCCCACCCACGCGC	GGTTCCACCACCACCGCGGGTGACGCAGGTG
Cre09.g398750	GGAGGCGGAGGTATGCAGCTCAACCGAGCC	GGTTCCACCACCACCTCACCCGCGAGACGGG
Cre09.g398845	GGAGGCGGAGGTATGGCTATAGTCGTCGGCC	GGTTCCACCACCACCTCACTCTTTTGAATAAGCGACAAG
Cre09.g398882	GGAGGCGGAGGTATGGCCGCACCAGCTATG	GGTTCCACCACCACCCTGGTTTCCGGCATTGGAGG
Cre09.g398919	GGAGGCGGAGGTATGCGAGTTGGCTCTTCAGC	GGTTCCACCACCACCCCGGTCACTCGCCGCC
Cre09.g398993	GGAGGCGGAGGTATGGGCGCTGGTGGCG	GGTTCCACCACCACCCTACACATTTAGGGTGCGCTCA
Cre09.g399030	GGAGGCGGAGGTATGGGAGCGGGTGGTG	GGTTCCACCACCACCTTAAACATTTAAGGTGCGCTCA
Cre09.g399067	GGAGGCGGAGGTATGCCTGCTGACCCTGATAGG	GGTTCCACCACCACCGTAGCTGCCGGGGTCCAC
Cre09.g399104	GGAGGCGGAGGTATGGCTGAACAAAGGTGGAC	GGTTCCACCACCACCCTACAGCCGGCGCACG
Cre09.g399150	GGAGGCGGAGGTATGGCGAGCTCCTTGCG	GGTTCCACCACCACCCGCCTTCGAGGCCTTG
Cre09.g399200	GGAGGCGGAGGTATGGGGCTGCGTGAGAAGATC	GGTTCCACCACCACCGGCGCTCACAGGTGTGCC
Cre09.g399252	GGAGGCGGAGGTATGCGCCTACAGGCAGG	GGTTCCACCACCACCCTCCTCGCGTTCGTCCT
Cre09.g399300	GGAGGCGGAGGTATGCGCACCTCCATCATCG	GGTTCCACCACCACCCAGGAACTCCGCCAGACTGAT
Cre09.g399350	GGAGGCGGAGGTATGGAGGCGCTGCTATCG	GGTTCCACCACCACCGAAATACAGACGAGGTGGCGC
Cre09.g399402	GGAGGCGGAGGTATGGGCTTGCTCGATACATTCT	GGTTCCACCACCACCGTCGTCGCGGCGCTGAAG
Cre09.g399450	GGAGGCGGAGGTATGGGCGTTTTGGCTCAA	GGTTCCACCACCACCCTTGCCTCCTTCAATCATGCC
Cre09.g399513	GGAGGCGGAGGTATGTCGTCCTTGGACTATGATAGT	GGTTCCACCACCACCTCAAAAAGGCATGTAGAAGCGG
Cre09.g399551	GGAGGCGGAGGTATGGAAGCCGCACACAG	GGTTCCACCACCACCAGGCCCTGTGTGTACCTG
Cre09.g399589	GGAGGCGGAGGTATGCAGTGCTCGCCCCAAG	GGTTCCACCACCACCGAAGATGTCTTGCGGAAAGCTC
Cre09.g399601	GGAGGCGGAGGTATGTCTAAGCGACCCGCC	GGTTCCACCACCACCCTGCGCCAGAACCGGTA
Cre09.g399650	GGAGGCGGAGGTATGGACCAAACGCAGCG	GGTTCCACCACCACCGCGGAAACGCAGCCAC
Cre09.g399738	GGAGGCGGAGGTATGGGCGCAGGCAGCATTAA	GGTTCCACCACCACCAACTTTGATGCCCACACGCTC
Cre09.g399900	GGAGGCGGAGGTATGACAGCTCCTCTAATGACAGC	GGTTCCACCACCACCCTACTCCTCCCCTCTCGC
Cre09.g399909	GGAGGCGGAGGTATGCCACCCCTCCCCG	GGTTCCACCACCACCGAGATCATCTGCACGCCC
Cre09.g399911	GGAGGCGGAGGTATGTCGTCGACGACGCCAG	GGTTCCACCACCACCCCGGATGCTGCGCAGGCC
Cre09.g399915	GGAGGCGGAGGTATGGGCACAAGCTCTAGCAA	GGTTCCACCACCACCCTACCACCGCCGCGGT
Cre09.g399950	GGAGGCGGAGGTATGCTCAGCTGCCGCC	GGTTCCACCACCACCGCCCTGGAACGCCTTG
Cre09.g400000	GGAGGCGGAGGTATGGCGGGTGTTCAGGACTC	GGTTCCACCACCACCCTGCCCGCCGGCAACCTC
Cre09.g400100	GGAGGCGGAGGTATGGAGCGCGAGCAGGCC	GGTTCCACCACCACCGATCTTCTGCACCTTGCTGGC
Cre09.g400145	GGAGGCGGAGGTATGTGGGGCGTGGAACA	GGTTCCACCACCACCCTAAGCGGGCTCCAGCC
Cre09.g400200	GGAGGCGGAGGTATGCACTCAGAGAAGCCCG	GGTTCCACCACCACCCGCAAAAATTACGCGCGG
Cre09.g400250	GGAGGCGGAGGTATGCTGGTGCTGGACGTCATC	GGTTCCACCACCACCTCGGCCGCTGGAAGCGTC
Cre09.g400293	GGAGGCGGAGGTATGAACTCCACACCCTCGG	GGTTCCACCACCACCCATGAAGCCCAGTTCGCG
Cre09.g400330	GGAGGCGGAGGTATGGCCGACTCGCAGC	GGTTCCACCACCACCGTCGTCCATGTCCTCATCG
Cre09.g400400	GGAGGCGGAGGTATGGCGGCAGGCCACTCT	GGTTCCACCACCACCCAGCTTCTTGGCCTTGAGCT

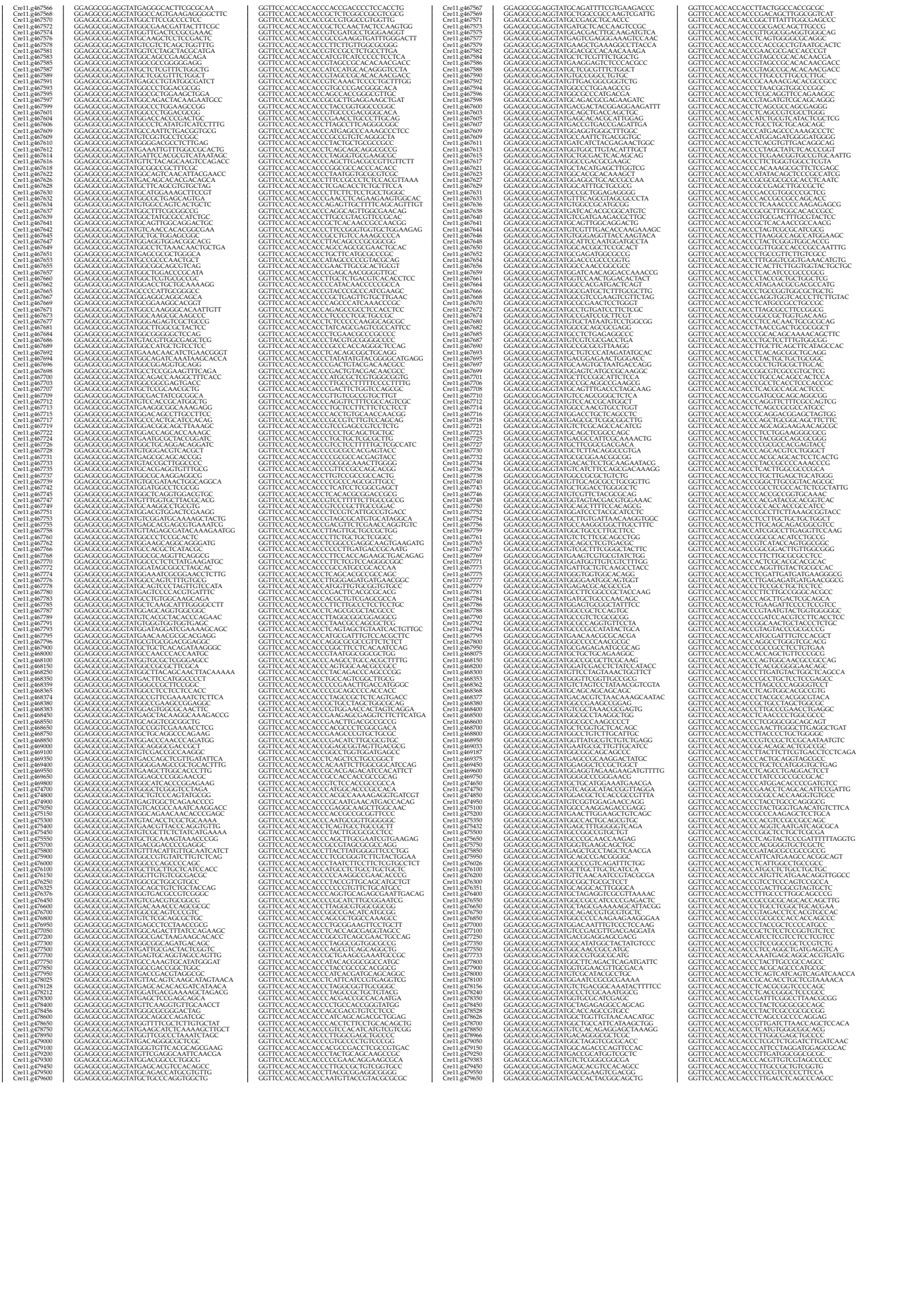
Cre09.g392097	GGAGGCGGAGGTATGCCTGCGGCGGTCG	GGTTCCACCACCACCCAAGGACTCCGCGCCATC
Cre09.g392106	GGAGGCGGAGGTATGGGCGTGGAGCCCTA	GGTTCCACCACCACCATGTGCCACGCATGCC
Cre09.g392171	GGAGGCGGAGGTATGAGGCGCTGGCGAG	GGTTCCACCACCACCTTAGGCTGCGCCGCCT
Cre09.g392245	GGAGGCGGAGGTATGAAGCGGTTCCACTACATT	GGTTCCACCACCACCTCAGCCCGTTGACACGC
Cre09.g392252	GGAGGCGGAGGTATGGAGGGGCGGCTGA	GGTTCCACCACCACCCGAGTTGGGGCACTGGC
Cre09.g392300	GGAGGCGGAGGTATGCATGAAGGACAAAACACATG	GGTTCCACCACCACCTCACCGCTTGCCGTGC
Cre09.g392350	GGAGGCGGAGGTATGTCGACCAAGCTCTACGTC	GGTTCCACCACCACCGTAGCCGCCGCCACCGTAG
Cre09.g392393	GGAGGCGGAGGTATGGTCAAGGCCGAGCC	GGTTCCACCACCACCGCCCTTCATCCTCTTCTCGA
Cre09.g392430	GGAGGCGGAGGTATGAGTTTCTGGTACTATTTGTTCCTG	GGTTCCACCACCACCGTGCGTCTTCTTAACACCG
Cre09.g392500	GGAGGCGGAGGTATGGGCTTGTGGAAGCG	GGTTCCACCACCACCCGACTCGTCGTCGCTG
Cre09.g392504	GGAGGCGGAGGTATGGGGCCGCCGGAAC	GGTTCCACCACCACCCTTCCCGTTCGTACCCGAC
Cre09.g392542	GGAGGCGGAGGTATGGGAAGGCGACGGAGTTA	GGTTCCACCACCACCGCTCCCCACCGCCATCAC
Cre09.g392579	GGAGGCGGAGGTATGTCAATCGCGTTGCTCT	GGTTCCACCACCACCAGCCGTAGCCACAGCT
Cre09.g392650	GGAGGCGGAGGTATGAAGCGAGACAAGGACAG	GGTTCCACCACCACCAGTGAAGCAGCGCCGG
Cre09.g392653	GGAGGCGGAGGTATGACAGGGCTGCTGATGC	GGTTCCACCACCACCGCTACTCACGTCTGACTTGAC
Cre09.g392692	GGAGGCGGAGGTATGGCCGATGACTTTTACCTT	GGTTCCACCACCACCCTTCTGAATGGGCTGGATCTT
Cre09.g392729	GGAGGCGGAGGTATGGGCGGGCCCGTTCAG	GGTTCCACCACCACCACAGTTTGCACCGGCGCC
Cre09.g392766	GGAGGCGGAGGTATGCAGCCTACAGAGCAGAGC	GGTTCCACCACCACCCTAGCGGCCGCTGGCG
Cre09.g392803	GGAGGCGGAGGTATGTTACTTGTCGCGATTGGG	GGTTCCACCACCACCCTTCACCATGCCGCTGAC
Cre09.g392914	GGAGGCGGAGGTATGCGAGGTGACGCCA	GGTTCCACCACCACCGCCAGTCCCTCCACCG
Cre09.g392951	GGAGGCGGAGGTATGGGAGGCCTACACGC	GGTTCCACCACCACCAAGGTGCCACCGGCCA
Cre09.g393000	GGAGGCGGAGGTATGATGCGGCGCCGGG	GGTTCCACCACCACCTGCGTACGTTGACTTCCAGG
Cre09.g393050	GGAGGCGGAGGTATGAATATCACCAATTCCGACGC	GGTTCCACCACCACCCTCATCCTCCCCGCTGCT
Cre09.g393062	GGAGGCGGAGGTATGACGCTGTGCCACCT	GGTTCCACCACCACCCAGCTGCCGCCGCATC
Cre09.g393100	GGAGGCGGAGGTATGTCTGGCACCACCAAGA	GGTTCCACCACCACCCTCCGTTTCTTCCTGCTCC
Cre09.g393150	GGAGGCGGAGGTATGGACAGCAAGGAAGCG	GGTTCCACCACCACCTTACGCGGCCGTGACG
Cre09.g393200	GGAGGCGGAGGTATGCTGCAGCGCTCGGGAG	GGTTCCACCACCACCCTTCTTCTCCTCAGAGGAGCTG
Cre09.g393247	GGAGGCGGAGGTATGCTCGGACGCCGAG	GGTTCCACCACCACCTTACGCCCCACGACCG
Cre09.g393284	GGAGGCGGAGGTATGTTCGGCTGCGGCG	GGTTCCACCACCACCCAGCGCTGCACGCTGC
Cre09.g393321	GGAGGCGGAGGTATGCTCGGCAATGTTGCAG	GGTTCCACCACCACCCTGCCAGCCGCTGATGAA
Cre09.g393358	GGAGGCGGAGGTATGGACCAGAACGTGTTTGCTC	GGTTCCACCACCACCGGACAGAGCCGGGAAGTG
Cre09.g393400	GGAGGCGGAGGTATGCCCAGCACTGCGCTG	GGTTCCACCACCACCCTGCTGCAGCTTGCGGCC
Cre09.g393469	GGAGGCGGAGGTATGGCGCTTGGCCTTCTG	GGTTCCACCACCACCGTGCAGCATGCGGCGG
Cre09.g393580	GGAGGCGGAGGTATGCACCCTCAGCTTAAACAG	GGTTCCACCACCACCCATGACGGCACAGCCG
Cre09.g393617	GGAGGCGGAGGTATGGAGGCTAACGAGCATCT	GGTTCCACCACCACCGCTGCCGCGTTGCATG
Cre09.g393691	GGAGGCGGAGGTATGCCCCGGTTGAAGCGTGTAG	GGTTCCACCACCACCCCGATGCTGCCGCGCCGAC
Cre09.g393728	GGAGGCGGAGGTATGGCTGAAGCCTTTGAAGAATG	GGTTCCACCACCACCCTCGCTGGCCTTGGGTGC
Cre09.g393765	GGAGGCGGAGGTATGCAGTCGCTCCGCG	GGTTCCACCACCACCGTCGGTGGCCTGCATG
Cre09.g393850	GGAGGCGGAGGTATGAACAACTCTATGCAGGGG	GGTTCCACCACCACCCTGCCCAGCAGATTCGC
Cre09.g393900	GGAGGCGGAGGTATGGCGAGCCTGGAGCGC	GGTTCCACCACCACCCGACGAGCCCGCATCGCTG
Cre09.g393950	GGAGGCGGAGGTATGAACTCAGACGATGAGGACGC	GGTTCCACCACCACCCTACTGGCCGCCGCCG
Cre09.g393954	GGAGGCGGAGGTATGCAGGAGGAGTCTGTGTA	GGTTCCACCACCACCATCGTCCACCACCAGCA
Cre09.g394000	GGAGGCGGAGGTATGATGGTCAGCGCGATAG	GGTTCCACCACCACCATCCACCAGCCCCATGC
Cre09.g394065	GGAGGCGGAGGTATGATGACTCCCGCTGAGG	GGTTCCACCACCACCGCCGAAGTGAATCTGGGAAG
Cre09.g394150	GGAGGCGGAGGTATGCCAGGTATGCCACTAGC	GGTTCCACCACCACCGCTCCAGCGCACCAGG
Cre09.g394200	GGAGGCGGAGGTATGGGGCGTCTTTCAGTCG	GGTTCCACCACCACCCGGGCAGCCAGAGAACG
Cre09.g394213	GGAGGCGGAGGTATGGCGGAAGCCGCTGACTTC	GGTTCCACCACCACCGTACAGGTACACGCCGAAG
Cre09.g394288	GGAGGCGGAGGTATGAGAGCTCTGTGCTGCA	GGTTCCACCACCACCAGCGCCCACAAAGCAC
Cre09.g394350	GGAGGCGGAGGTATGGCAGCGCCTGCTTGG	GGTTCCACCACCACCCCGCACCAGTGCCTCGGC
Cre09.g394399	GGAGGCGGAGGTATGGTGAAGCCGCAGCG	GGTTCCACCACCACCCAGCCGCCTGCTCTTG
Cre09.g394450	GGAGGCGGAGGTATGTCTGGCGAGTCTGAAGC	GGTTCCACCACCACCCTCGTCCTTGGTGGTCGC
Cre09.g394510	GGAGGCGGAGGTATGCTGAACCAATGCATTAACAG	GGTTCCACCACCACCCCGCCCATTGGTGGCGTT
Cre09.g394550	GGAGGCGGAGGTATGGCTGCTCAAAACCTTGC	GGTTCCACCACCACCGATCTGCAGCGCGGCG
Cre09.g394600	GGAGGCGGAGGTATGTCGCAGAACCCCTTCTCTC	GGTTCCACCACCACCAGCTCCCCAGTCCATATCTC
Cre09.g394658	GGAGGCGGAGGTATGGCCAACACTGTGCCGATTC	GGTTCCACCACCACCCTCGTCCCACAGCTCGGC
Cre09.g394700	GGAGGCGGAGGTATGCGGCCCGTCAGTGTG	GGTTCCACCACCACCCTCCTCCTGCGGCTCCAC
Cre09.g394701	GGAGGCGGAGGTATGGCCTCGCTCTTTGC	GGTTCCACCACCACCCTTGGCCTCCTCGGGG
Cre09.g394800	GGAGGCGGAGGTATGGGGCAGGTGGAGCTG	GGTTCCACCACCACCAAGCGGTGACGACACGCC
Cre09.g394843	GGAGGCGGAGGTATGTCGTCCGTGTATGAGCG	GGTTCCACCACCACCCAGCTTGCGGATGGCC
Cre09.g394917	GGAGGCGGAGGTATGCATACTCGCTGGCAGAAT	GGTTCCACCACCACCCGAGTGCTTGCCCACGAT
Cre09.g395000	GGAGGCGGAGGTATGCTCCTGATCCTCACGC	GGTTCCACCACCACCCGGCTTTGAACCGCCG
Cre09.g395050	GGAGGCGGAGGTATGCATGCCACCAACACC	GGTTCCACCACCACCTGCCTGGGTAGCACCAG
Cre09.g395100	GGAGGCGGAGGTATGGCAAAAAGTCCTACGCAGT	GGTTCCACCACCACCGCTGATGCAGCAGAAGCG
Cre09.g395139	GGAGGCGGAGGTATGACGGACGTCGGCAGTTTC	GGTTCCACCACCACCGCTGTGCGGGAACACGCC
Cre09.g395176	GGAGGCGGAGGTATGCTGGGACGGATGGC	GGTTCCACCACCACCCGGCGACTGCGGATATG
Cre09.g395213	GGAGGCGGAGGTATGACGGGCAAGGGCC	GGTTCCACCACCACCCTTGTGGGAGCGGGTGT
Cre09.g395251	GGAGGCGGAGGTATGGCGCTGTGCGTCG	GGTTCCACCACCACCCTGCGGCATGCCTCCG
Cre09.g395300	GGAGGCGGAGGTATGAGCGATGACGGTGATG	GGTTCCACCACCACCCTGCTGGTCAGCATCCG
Cre09.g395350	GGAGGCGGAGGTATGTCCGCTGCTATCGGG	GGTTCCACCACCACCATGGTGCGCCTCCACC
Cre09.g395399	GGAGGCGGAGGTATGTCTCAAGAATCAGATGAAACT	GGTTCCACCACCACCCTACCCGCCTCCTTCAC
Cre09.g395400	GGAGGCGGAGGTATGGCGTCGGACGCGG	GGTTCCACCACCACCGTAATGCTGCACTCCCTGCT
Cre09.g395450	GGAGGCGGAGGTATGCCCACGCTCGCCAAG	GGTTCCACCACCACCCAGGTCACCATGGGTATCGTTG
Cre09.g395500	GGAGGCGGAGGTATGAGCGGCACAAAAAAGTC	GGTTCCACCACCACCCTGCTGCTGCTTGTCGC
Cre09.g395600	GGAGGCGGAGGTATGGCCACAAAACAGCCG	GGTTCCACCACCACCCAGCAGCCCCTTCGGC
Cre09.g395646	GGAGGCGGAGGTATGCCGTGGCCGTCGC	GGTTCCACCACCACCTCATCCCCACGCCAGCG
Cre09.g395658	GGAGGCGGAGGTATGCAGGCCCACGGACTA	GGTTCCACCACCACCCTCCGCAGGCAGGCTTCC
Cre09.g395700	GGAGGCGGAGGTATGTGCGCCGGCATTCGG	GGTTCCACCACCACCGGCCGCCTTGTTGGCTTTGATG
Cre09.g395750	GGAGGCGGAGGTATGCCGCGCGTTCGATCG	GGTTCCACCACCACCGCGCTTCTGGAAGTTAAGCT
Cre09.g395800	GGAGGCGGAGGTATGCCTACCTTCCTTTTATGCA	GGTTCCACCACCACCTCACCCCTTGTGCAGCT
Cre09.g395843	GGAGGCGGAGGTATGGCAAGCCGGCTGGCAAG	GGTTCCACCACCACCAAACGGCGGCGGCGGGAG
Cre09.g395880	GGAGGCGGAGGTATGGCCCAGGTGACGGGC	GGTTCCACCACCACCCACCTGCGCAAGGTAGTCATC
Cre09.g395917	GGAGGCGGAGGTATGACTTTGCAATCTTACAAAGGGC	GGTTCCACCACCACCCCGCTTCTTTGTGAGGTCGTAG
Cre09.g395950	GGAGGCGGAGGTATGCTTCAGACCGCACCTA	GGTTCCACCACCACCCTAGGGCAGCTGGCTGG
Cre09.g395991	GGAGGCGGAGGTATGTACCAAGTGCTGGTGGG	GGTTCCACCACCACCGCACACACTCACCGGTG
Cre09.g396028	GGAGGCGGAGGTATGGCAACGCTTTCGCGG	GGTTCCACCACCACCCACAACACCTGCTGTCCGG
Cre09.g396065	GGAGGCGGAGGTATGGGCCTAGGACGACTC	GGTTCCACCACCACCATCATCTTCACCGTCCGGC
Cre09.g396102	GGAGGCGGAGGTATGGGCTCGCTCCAGAAA	GGTTCCACCACCACCCTGTGGCGGCATGGGC
Cre09.g396150	GGAGGCGGAGGTATGCCGCGGTCGGTATCC	GGTTCCACCACCACCCTCGCGGTACCCGGCG
Cre09.g396200	GGAGGCGGAGGTATGGCTGCTCTGAAGTGGC	GGTTCCACCACCACCCATGCGCTCGTCGCCA
Cre09.g396250	GGAGGCGGAGGTATGCCGGCGACCTACCTC	GGTTCCACCACCACCCACCGACAGGCTGATCGAG
Cre09.g396289	GGAGGCGGAGGTATGAACAGCCACAGCTGGC	GGTTCCACCACCACCTGCGGCCATCAGCGGG
Cre09.g396326	GGAGGCGGAGGTATGAGCGGCACGGGGC	GGTTCCACCACCACCATCGTCGTCCTCGGCGA
Cre09.g396363	GGAGGCGGAGGTATGGCGACTCCTCTTGTGG	GGTTCCACCACCACCGAAGGGTGACTTTCCGACCTTC
Cre09.g396450	GGAGGCGGAGGTATGGATAAGCGCTACTTTCCTC	GGTTCCACCACCACCTCACAACGGGTCCCCAA
Cre09.g396475	GGAGGCGGAGGTATGGAGGTGCTGGGGG	GGTTCCACCACCACCGCCGTACACGCGAATCC
Cre09.g396512	GGAGGCGGAGGTATGGCGCTCAGGCACG	GGTTCCACCACCACCTCAAGCACTCGCGCCG
Cre09.g396550	GGAGGCGGAGGTATGGTGTGGGTGTCGGC	GGTTCCACCACCACCCATGATGGCGCTGCTGA
Cre09.g396587	GGAGGCGGAGGTATGACGCCCGGCGTCG	GGTTCCACCACCACCCACCGACGCCGCTGCT
Cre09.g396624	GGAGGCGGAGGTATGGGCAACACGGTCACG	GGTTCCACCACCACCCTGCTTCTTGACGGCGAT
Cre09.g396661	GGAGGCGGAGGTATGGAGCACTTTGAGGTGGA	GGTTCCACCACCACCCGCGAGTGCCATCATTCG
Cre09.g396700	GGAGGCGGAGGTATGCTGGCTGGCAAAGTGC	GGTTCCACCACCACCTCAAGCGGCCAGCGCC
Cre09.g396772	GGAGGCGGAGGTATGCGGCAGACAATTGCTCC	GGTTCCACCACCACCTGACGGAGGAGGCCGAGG
Cre09.g396809	GGAGGCGGAGGTATGGTGCGCATGGTTATTG	GGTTCCACCACCACCTTAAAGCATGACGTATTCATCAAC
Cre09.g396850	GGAGGCGGAGGTATGCTAAGCAACAAGCGACA	GGTTCCACCACCACCTTAATCGCGCGCAGGTG
Cre09.g396900	GGAGGCGGAGGTATGGACGCGGGTCTTCCG	GGTTCCACCACCACCCAACTTGGCCGCTGCACC
Cre09.g396950	GGAGGCGGAGGTATGCTGTGGCGTGGGCATC	GGTTCCACCACCACCGTACTCGCGGGGCTGGCC
Cre09.g396994	GGAGGCGGAGGTATGGCACTGTCATCAGCATGG	GGTTCCACCACCACCCCGCTCCGTCAGCGTCAC
Cre09.g397031	GGAGGCGGAGGTATGATGGAGAAGTACGGGCGC	GGTTCCACCACCACCCGGCACAGCCACCGCC
Cre09.g397068	GGAGGCGGAGGTATGTTGGCAACTGGACATGT	GGTTCCACCACCACCTGGCTCGCTTCCCTTCG
Cre09.g397098	GGAGGCGGAGGTATGAGCGGCGAAACCG	GGTTCCACCACCACCTTGTGCGCGATAAATGCCT
Cre09.g397142	GGAGGCGGAGGTATGCGGAGACCTCCGATAC	GGTTCCACCACCACCCTCAGGGTCCCCGCCG
Cre09.g397179	GGAGGCGGAGGTATGTCGGCCCGTAGCC	GGTTCCACCACCACCCGCCTCCCCACCACCA
Cre09.g397216	GGAGGCGGAGGTATGGAGCTCGTGTCAGGC	GGTTCCACCACCACCCCGCGCCTTGACCACCAC
Cre09.g397253	GGAGGCGGAGGTATGCCACAAACACCTCAGGC	GGTTCCACCACCACCCAGCGCCCTCAGCAGCCC
Cre09.g397300	GGAGGCGGAGGTATGCTCGGGCAGACGCTG	GGTTCCACCACCACCCGGCACGGGTGACACGCC
Cre09.g397350	GGAGGCGGAGGTATGCCCCAATACAGCATCCG	GGTTCCACCACCACCCACAAAGTCGTCATCGTCTCG
Cre09.g397400	GGAGGCGGAGGTATGGCCCCAGGACACATGG	GGTTCCACCACCACCTTGTGCGCCTGCCGGC
Cre09.g397438	GGAGGCGGAGGTATGTCAGCGCCTGCCCT	GGTTCCACCACCACCGCCGTCTGCCTGCGCA
Cre09.g397475	GGAGGCGGAGGTATGAACAGGCCGGACACG	GGTTCCACCACCACCGTCTAACATGAAAACTGGGCCTG
Cre09.g397549	GGAGGCGGAGGTATGAGCGACAAACAGTTCCCG	GGTTCCACCACCACCCTGGCGGTTGTAGCCGCCAAGC
Cre09.g397623	GGAGGCGGAGGTATGGCACGGCGGTCCC	GGTTCCACCACCACCCGCCCAGATGATCTTGCA
Cre09.g397697	GGAGGCGGAGGTATGGCCGCCGCCATTC	GGTTCCACCACCACCGGCGGTCTTGGTCTGCT
Cre09.g397800	GGAGGCGGAGGTATGGAGCAGCTTTCAAGTTGGA	GGTTCCACCACCACCTGTACCTCCGCCGCTATG
Cre09.g397845	GGAGGCGGAGGTATGGCCCTGGGCCCCA	GGTTCCACCACCACCCAGAGCCAGGGACCGCC
Cre09.g397900	GGAGGCGGAGGTATGCCCGGAGAGGTTGCC	GGTTCCACCACCACCTTTCCCCTTTTTCCCGCCC
Cre09.g397950	GGAGGCGGAGGTATGGCTTATCATAGTCACTACAACA	GGTTCCACCACCACCCATGGTGCGGTCCAAAAAGC
Cre09.g398067	GGAGGCGGAGGTATGGCCACCGACGAGGTC	GGTTCCACCACCACCAGCCGACGCCACCACCTC
Cre09.g398178	GGAGGCGGAGGTATGTTGTTGGTGCGTGAGG	GGTTCCACCACCACCCGCAGGTGCAGGAGACGAG
Cre09.g398215	GGAGGCGGAGGTATGTCGCTGCAGCTGTTGC	GGTTCCACCACCACCCTATGCGGCCGCCGCG
Cre09.g398252	GGAGGCGGAGGTATGATGCTCAGCACGAAGGC	GGTTCCACCACCACCCGCCGAAACGGTGGCGGC
Cre09.g398300	GGAGGCGGAGGTATGCAGTATCAGCATGGATACG	GGTTCCACCACCACCTCAAATGTGCGTGGAGAG
Cre09.g398350	GGAGGCGGAGGTATGGGGGACGCCGAGC	GGTTCCACCACCACCCTTGCCCGCGCCACTG
Cre09.g398400	GGAGGCGGAGGTATGGATGCGCAGCAAAATATCC	GGTTCCACCACCACCCATGTCGTCATCCAGCCTCTC
Cre09.g398438	GGAGGCGGAGGTATGCGGCGCAGGTCAA	GGTTCCACCACCACCCCTGTGCAGCCGCTGG
Cre09.g398475	GGAGGCGGAGGTATGGAGCTCCTTAGCATTCTGA	GGTTCCACCACCACCTCACACCACGAAGCCAGG
Cre09.g398512	GGAGGCGGAGGTATGGAGCTCCTTAGCATACTGA	GGTTCCACCACCACCTCACACCACGAAGCCAGG
Cre09.g398550	GGAGGCGGAGGTATGAAAAAGAAGCATGGCATTATG	GGTTCCACCACCACCCTAGCCGTCCGCCATC
Cre09.g398623	GGAGGCGGAGGTATGCGGCGCAGGTCAA	GGTTCCACCACCACCCCGTACTGTCGTTCGCTTG
Cre09.g398652	GGAGGCGGAGGTATGTCGACCAAGTTCAGCGG	GGTTCCACCACCACCACGGCACGGGTGGGTG
Cre09.g398697	GGAGGCGGAGGTATGGAGCTCCTTAGCATTATGATT	GGTTCCACCACCACCTCAGGTCTCCAGAGCCCG
Cre09.g398734	GGAGGCGGAGGTATGCAGCAGGGGGCAG	GGTTCCACCACCACCCACCACGAAGCCAGGC
Cre09.g398771	GGAGGCGGAGGTATGGAGCTCCTTAGCATTATGATT	GGTTCCACCACCACCTCACACCACGAAGCCAGG
Cre09.g398850	GGAGGCGGAGGTATGGACGCCATCAAAGCAGC	GGTTCCACCACCACCGTGCCACTCGGCCGTCTCC
Cre09.g398900	GGAGGCGGAGGTATGATGCGGCGGCAACAC	GGTTCCACCACCACCCTTCAGGTTGAGGCCGACC
Cre09.g398983	GGAGGCGGAGGTATGCAATCAGTAGGGACGC	GGTTCCACCACCACCGGCGGCTCTCTCCTCG
Cre09.g399000	GGAGGCGGAGGTATGGAGCGGGTTAAGGAGGC	GGTTCCACCACCACCGGCGCTCACAGGAGTTCCTC
Cre09.g399050	GGAGGCGGAGGTATGGCGGCTAACTACACGC	GGTTCCACCACCACCGTCTCCTGGCGCCTGC
Cre09.g399073	GGAGGCGGAGGTATGCTAGCATCGGTGACCC	GGTTCCACCACCACCCGCCTTGGCGCCGTTG
Cre09.g399141	GGAGGCGGAGGTATGACGGTGGCGACGC	GGTTCCACCACCACCATCGAACAGCCTCTCCCC
Cre09.g399178	GGAGGCGGAGGTATGGCTGGGGAGTCTAGCA	GGTTCCACCACCACCCGTCTCCATCGGCACCGC
Cre09.g399250	GGAGGCGGAGGTATGTCGGCCCTTGCAGC	GGTTCCACCACCACCCTTGGGGAAGGGCACGC
Cre09.g399289	GGAGGCGGAGGTATGCCAATACTGCAACTGGACC	GGTTCCACCACCACCCAGGTCTCGCGGCTGACTTCTA
Cre09.g399326	GGAGGCGGAGGTATGGGCGAGCTTGCAACG	GGTTCCACCACCACCTTAGGCGGCGCTGGCG
Cre09.g399400	GGAGGCGGAGGTATGCCTCAGAACGGGCAC	GGTTCCACCACCACCCGCCTTCAGGTCAGCCGC
Cre09.g399439	GGAGGCGGAGGTATGCTTGCAGAGCGTACCG	GGTTCCACCACCACCCTGCACGGTGACAGTGGTC
Cre09.g399476	GGAGGCGGAGGTATGGAGGAAGCGCGATCAG	GGTTCCACCACCACCAGCCCGCTTCGGCCATTG
Cre09.g399550	GGAGGCGGAGGTATGGGCACAAGCTCTAGCAAG	GGTTCCACCACCACCCGCCAGGCCGTTCTTCTG
Cre09.g399552	GGAGGCGGAGGTATGACGGAGACCGACCAC	GGTTCCACCACCACCGCCAAGCAGGATGCGC
Cre09.g399600	GGAGGCGGAGGTATGGCCGATTGGGCCT	GGTTCCACCACCACCCCGCTTGCACTGCACC
Cre09.g399626	GGAGGCGGAGGTATGTTGAGCACGCGCG	GGTTCCACCACCACCGTTTCCGCGCAGCTCC
Cre09.g399701	GGAGGCGGAGGTATGTCTAAGAAGGATGGAGATGC	GGTTCCACCACCACCTTATTGCGCAGCACGCG
Cre09.g399849	GGAGGCGGAGGTATGTACTGGAAGAAGGAGCCA	GGTTCCACCACCACCAACCCAGACTTTGTCGCC
Cre09.g399907	GGAGGCGGAGGTATGGCTGACTTTAGCGACG	GGTTCCACCACCACCGTTTGGGCTGGTATTCACAAC
Cre09.g399910	GGAGGCGGAGGTATGAAGAAACCTGCATATGAAGGC	GGTTCCACCACCACCTTACGCGCCGACAACC
Cre09.g399914	GGAGGCGGAGGTATGTCTAAGCGACCCGCC	GGTTCCACCACCACCCTTTGCACAGGGCCGC
Cre09.g399916	GGAGGCGGAGGTATGCTCCCCGGTGCAC	GGTTCCACCACCACCAATCTGGGGTAGCGGCA
Cre09.g399997	GGAGGCGGAGGTATGGTAGGGGGCCCTGT	GGTTCCACCACCACCTCAACCTTGTAAGGATGATTCCTC
Cre09.g400050	GGAGGCGGAGGTATGTATGCATCTTTGAAGCCGT	GGTTCCACCACCACCCTACCGCTGCTGGCCG
Cre09.g400108	GGAGGCGGAGGTATGCATTCTATACCGCAAGCG	GGTTCCACCACCACCTCAACCTTGTAAGGATGATTCCTC
Cre09.g400150	GGAGGCGGAGGTATGTCTGCCGCCGGCT	GGTTCCACCACCACCCACCTCGCTCACGGTGA
Cre09.g400219	GGAGGCGGAGGTATGCAGTTCCACACTGACC	GGTTCCACCACCACCAGGCAAAAGCTTTCCACC
Cre09.g400256	GGAGGCGGAGGTATGATGGCTGGCTGCTTTG	GGTTCCACCACCACCCACCTCTTGTGCGCCC
Cre09.g400300	GGAGGCGGAGGTATGCTGGTGCTGGACGTCATC	GGTTCCACCACCACCCTGCCGCCCACTAGCGGG
Cre09.g400367	GGAGGCGGAGGTATGCTTCGCTTTAAAGGCAGT	GGTTCCACCACCACCCTACGTGACGCGGCGG
Cre09.g400404	GGAGGCGGAGGTATGGTCTTCGGCTGGGGC	GGTTCCACCACCACCCCCGTCGCGCTGCTCCTC

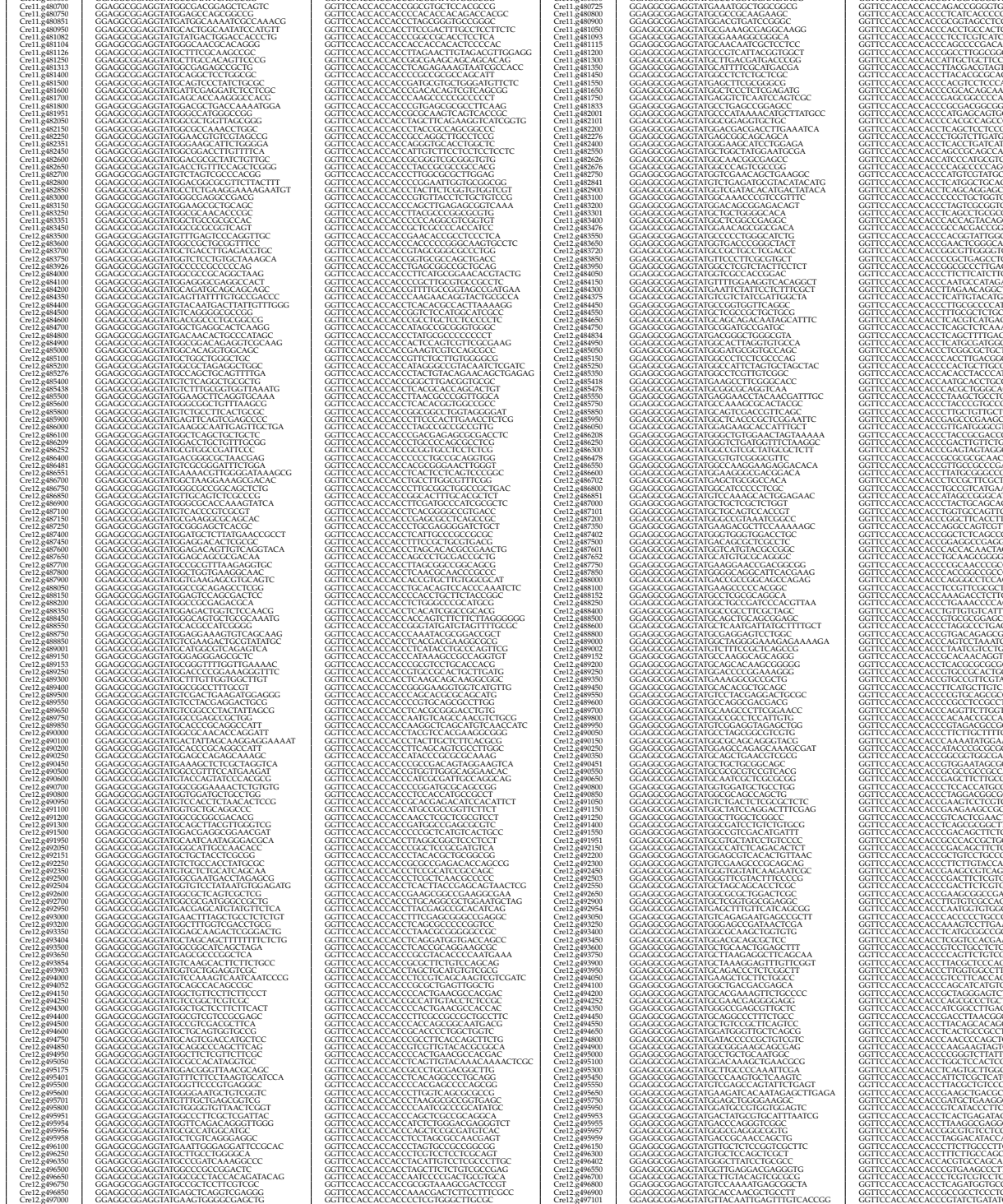


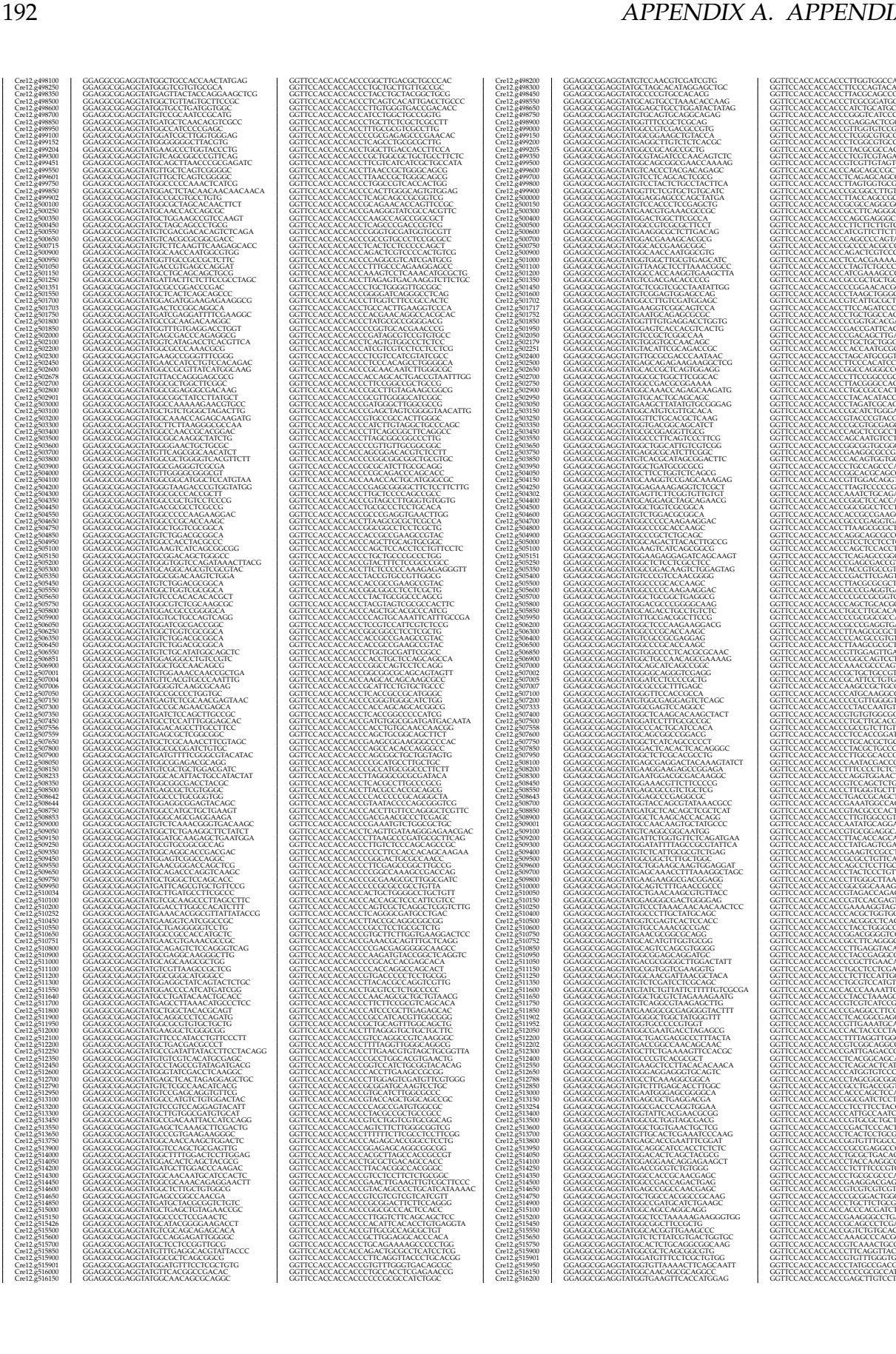


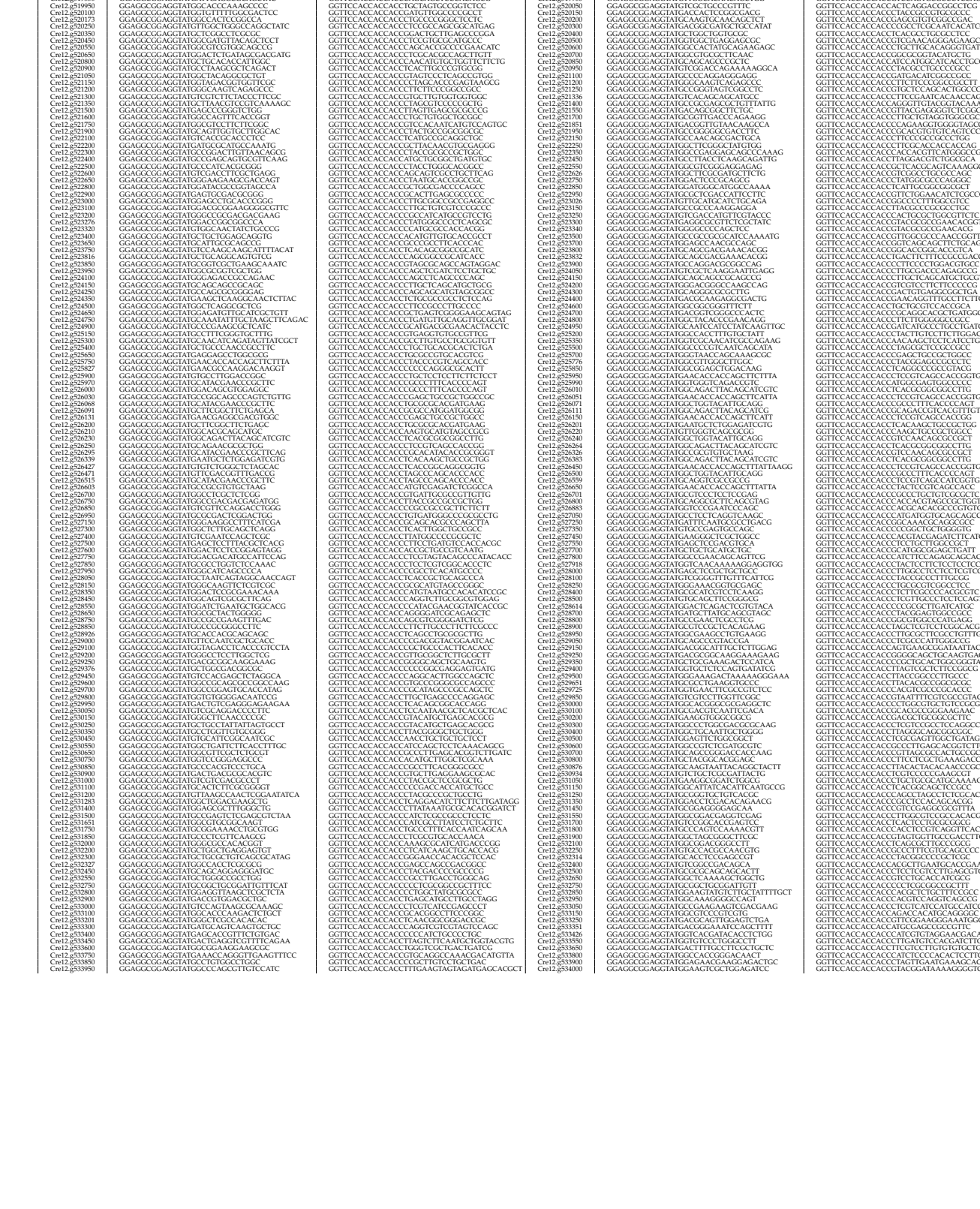


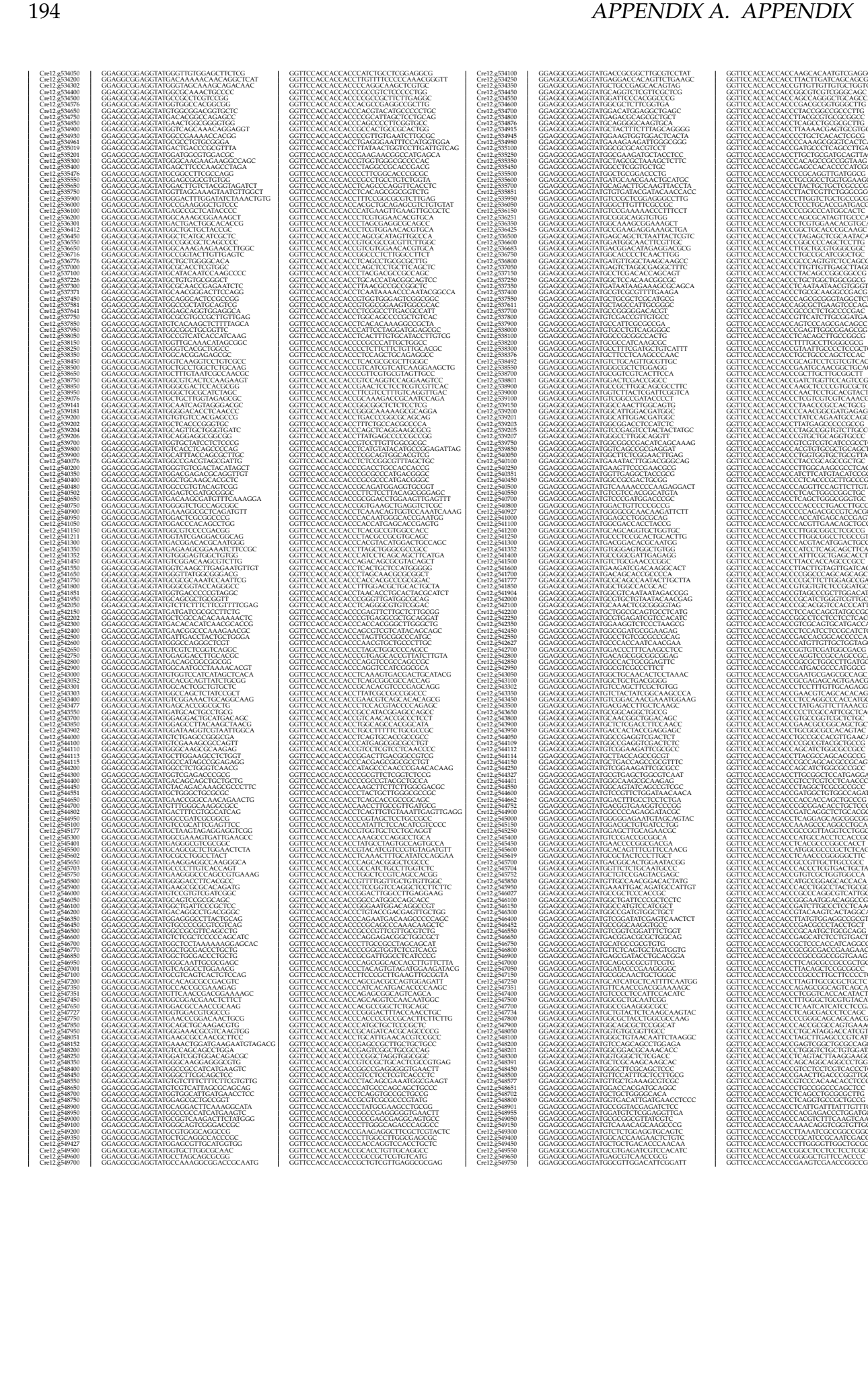


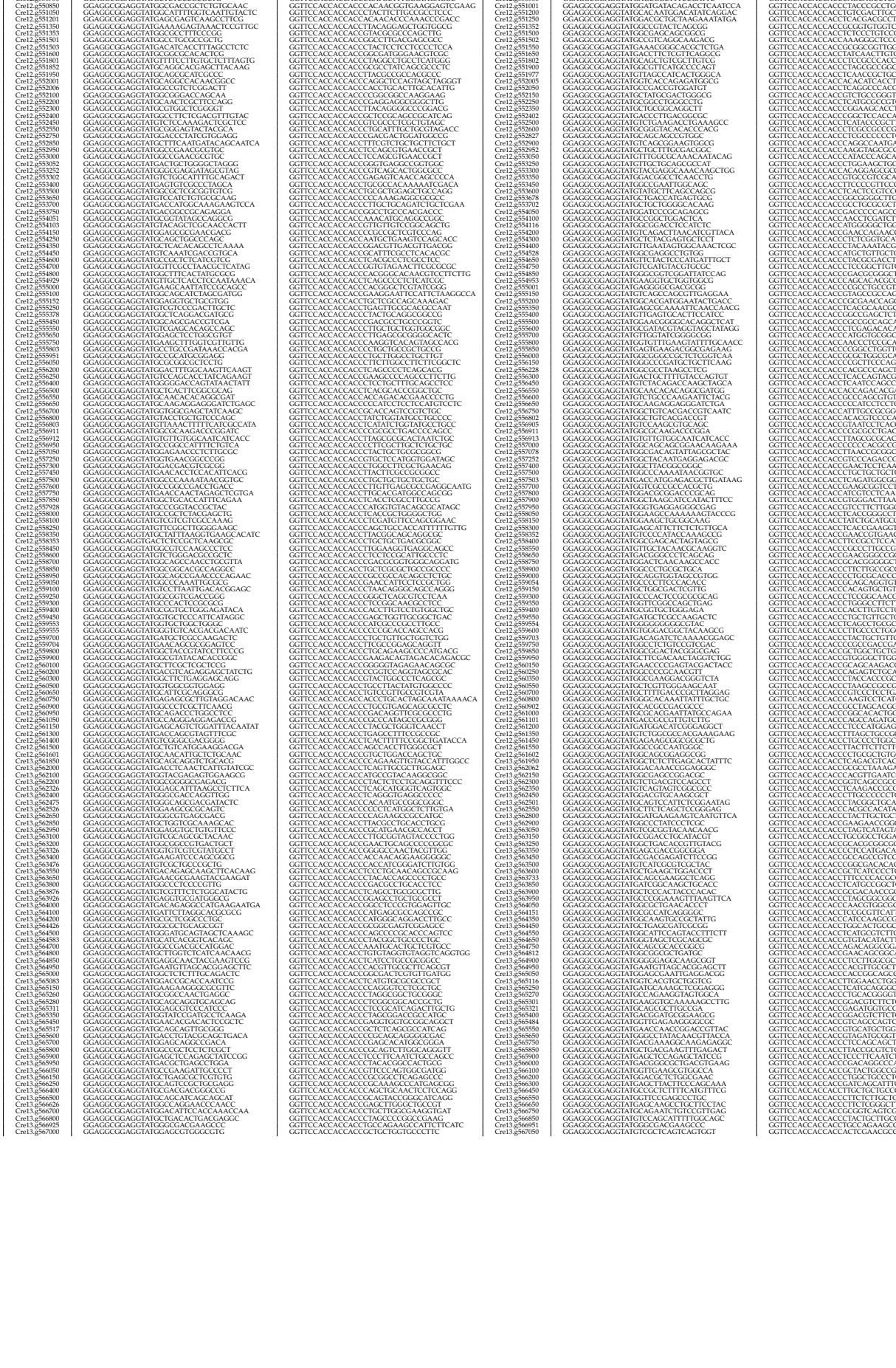


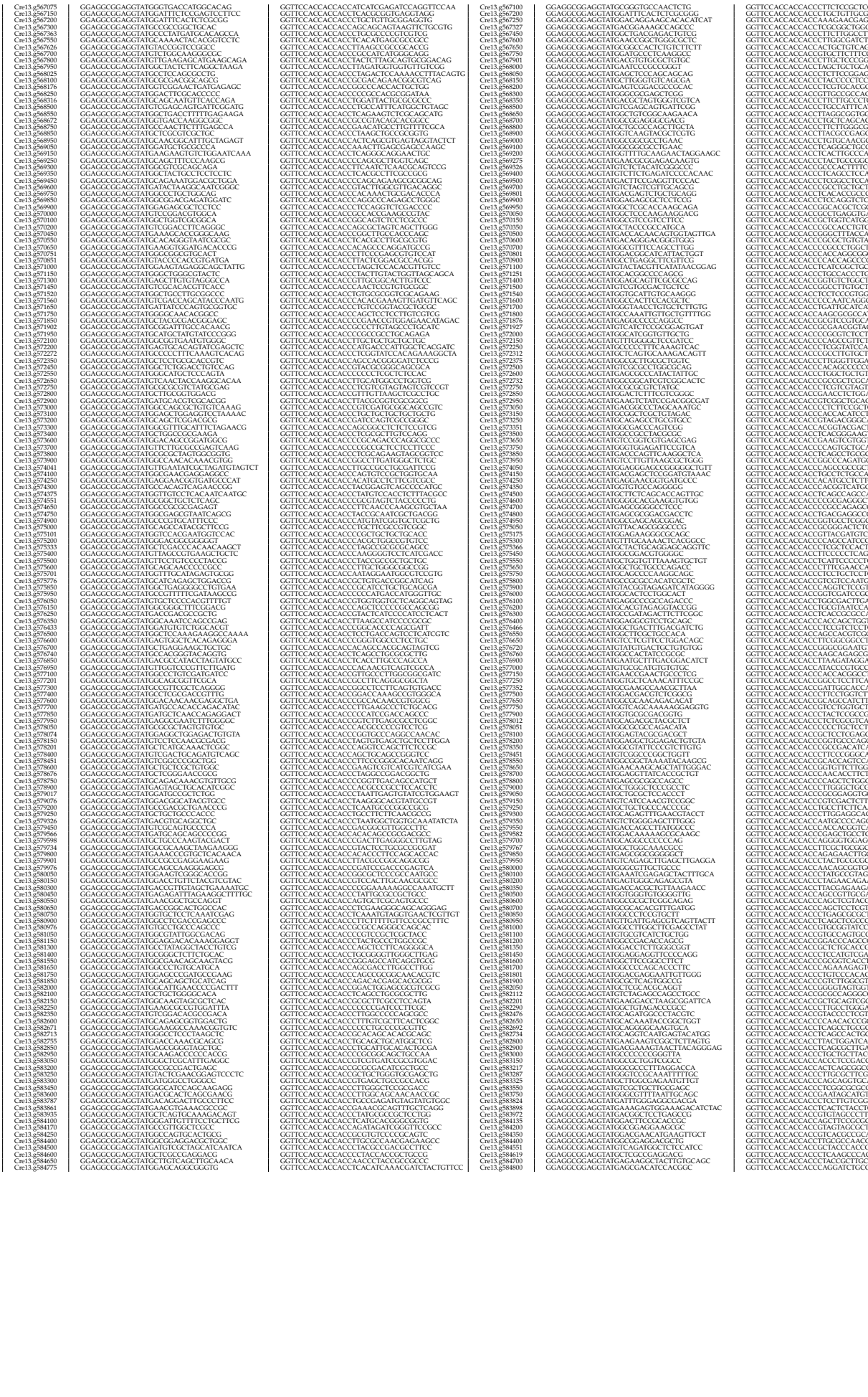


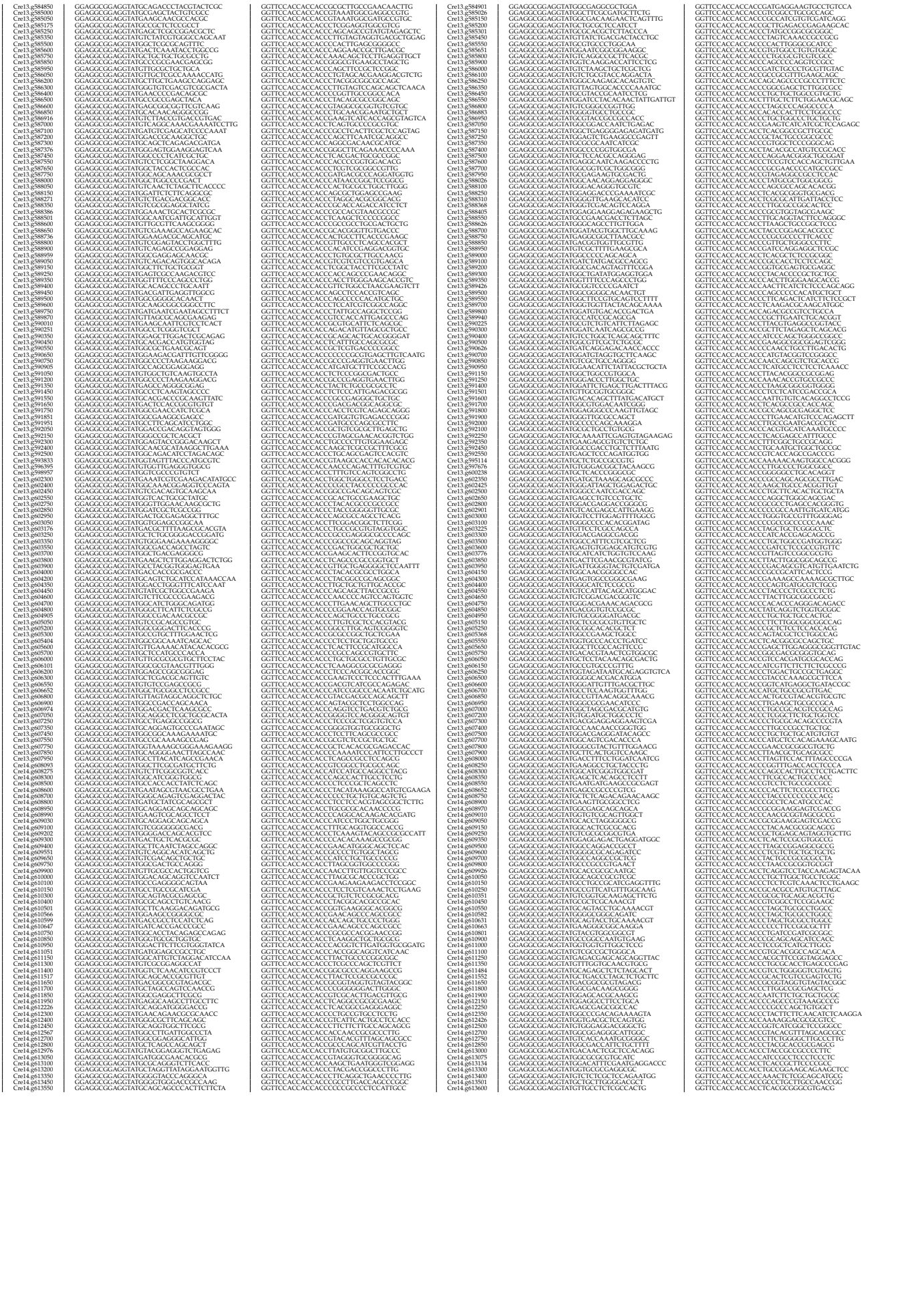


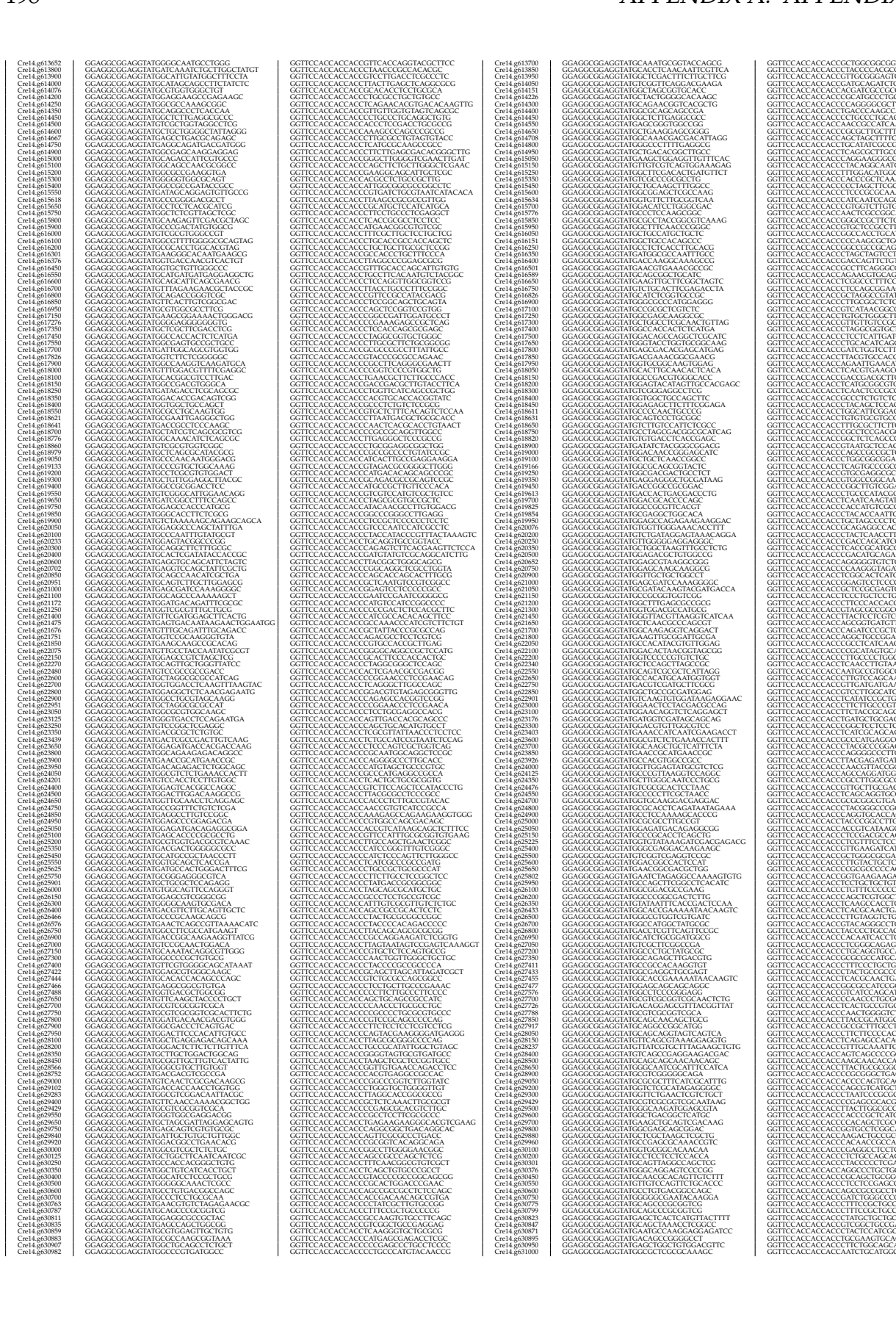


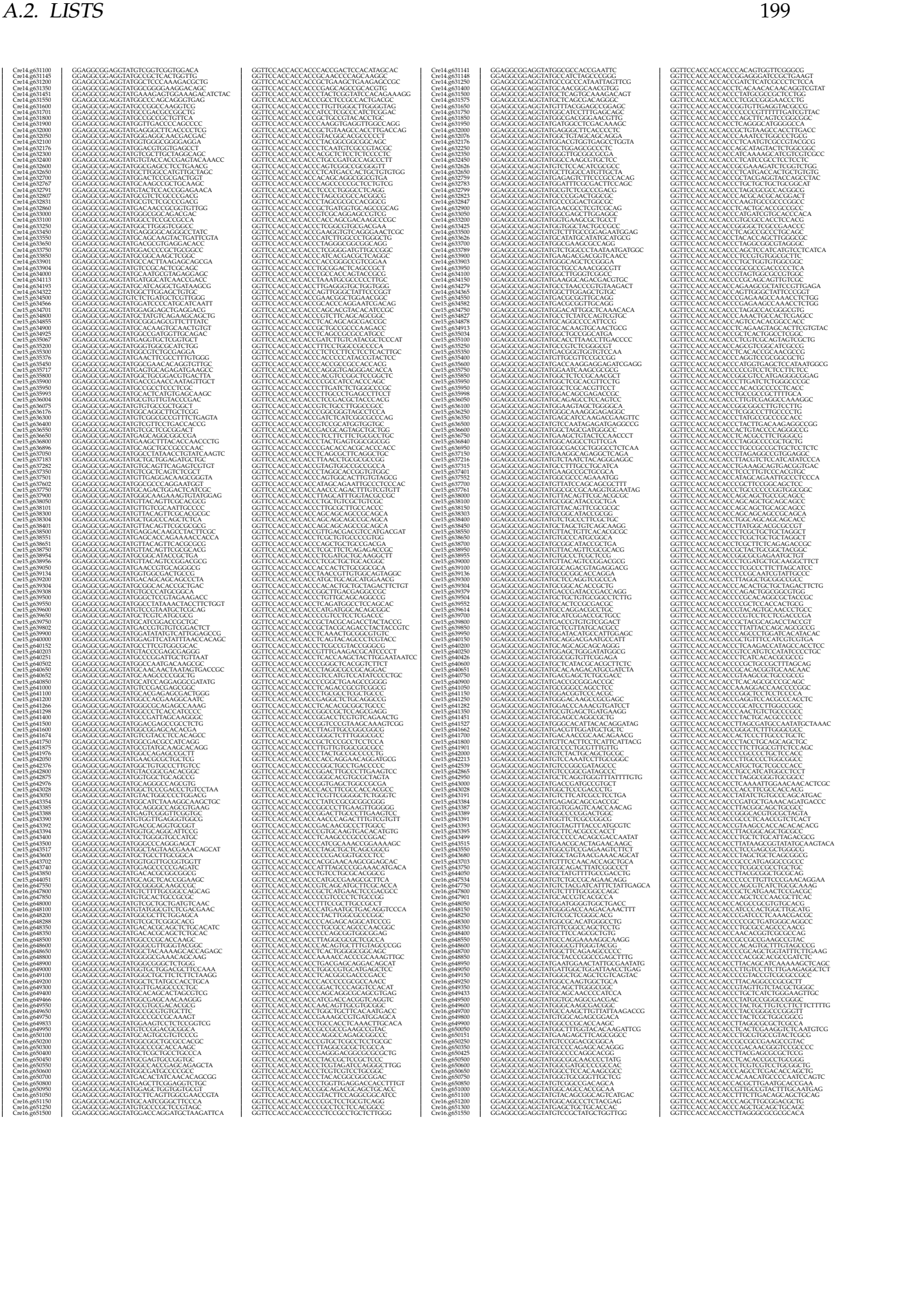


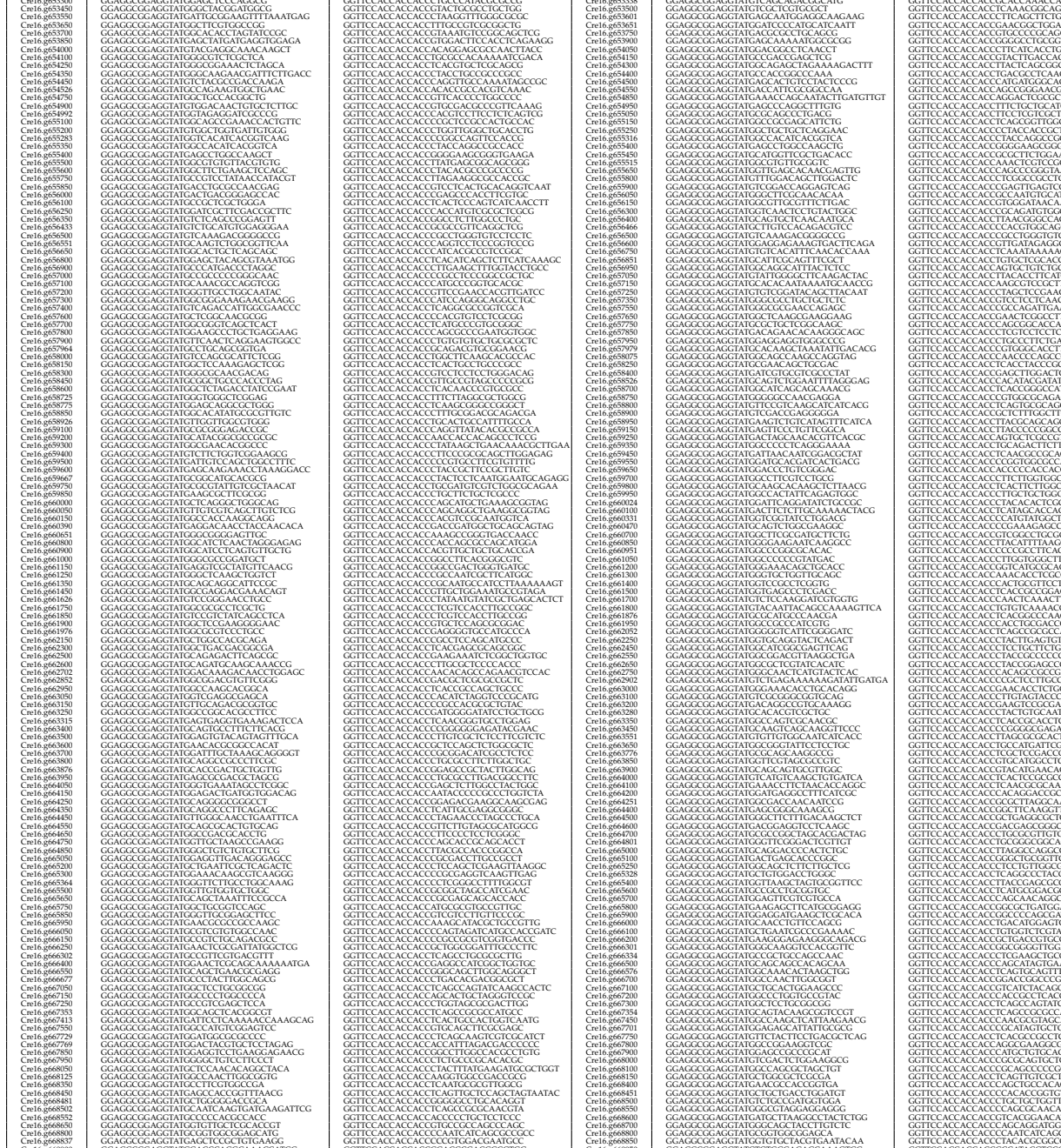


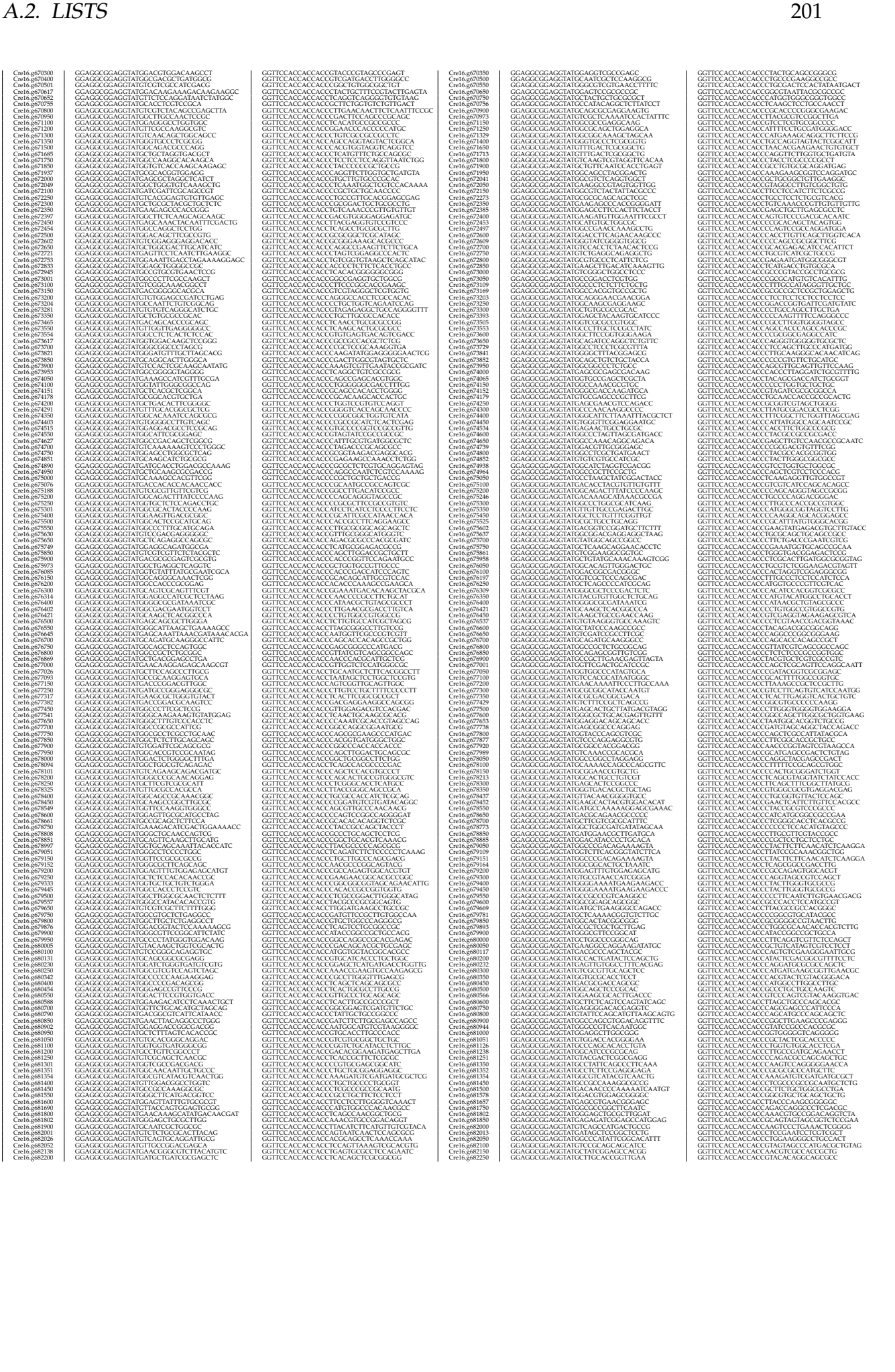


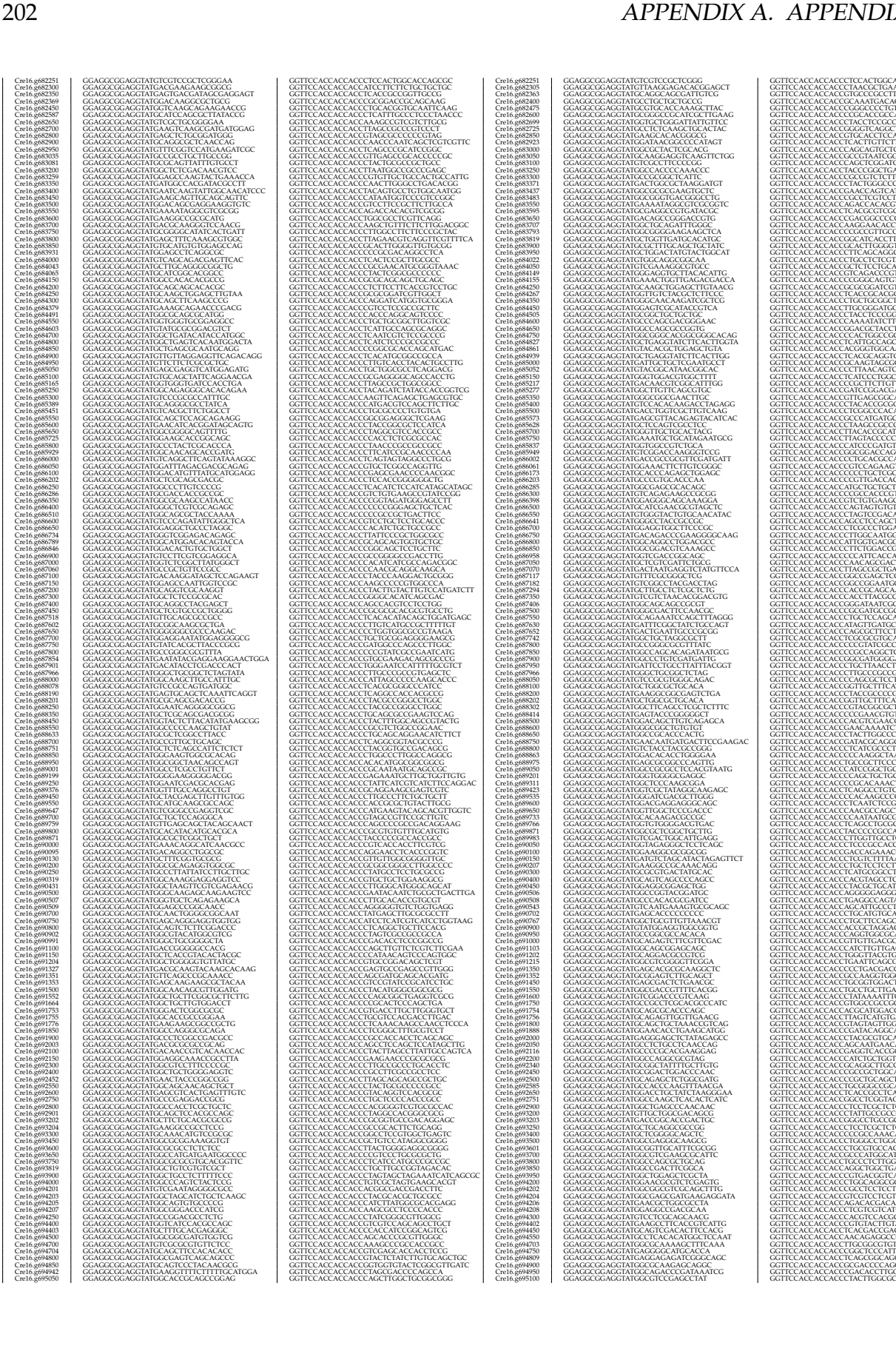












Cre16.g695200	GGAGGCGGAGGTATGTGCACTGCGCAGAC	GGTTCCACCACCACCCTCACCGATACCCCACAGC
Cre16.g695218	GGAGGCGGAGGTATGAGCGGGGAGAGGATGC	GGTTCCACCACCACCATAAAATTTGGGTGTGTCGCTGG
Cre16.g695350	GGAGGCGGAGGTATGGCAGTGCAACAACAAACG	GGTTCCACCACCACCTGCAGCGGGTCCTCGG
Cre16.g695500	GGAGGCGGAGGTATGGAGCAAGTTTACTGGGTC	GGTTCCACCACCACCGGCCTTCTCGGCGGTG
Cre16.g695600	GGAGGCGGAGGTATGTCAAGATGTTCGTTGGCG	GGTTCCACCACCACCGTCGTCGCCGTCATCTCC
Cre16.g695700	GGAGGCGGAGGTATGGGTCATGAGAAGTACCCT	GGTTCCACCACCACCCTTGGCGGGGCTGACC
Cre16.g695850	GGAGGCGGAGGTATGTCGCGCGTTTCGTGT	GGTTCCACCACCACCTCACTGCCGTGCCGCG
Cre16.g695900	GGAGGCGGAGGTATGCAGTCGCGCTATGAG	GGTTCCACCACCACCGGTAATGCCCCTGCGC
Cre16.g696000	GGAGGCGGAGGTATGGAGCTTGCTTTCACCG	GGTTCCACCACCACCCTTGCCCTTGACGACCTG
Cre17.g692676	GGAGGCGGAGGTATGGAGGGGTTTCAACCCC	GGTTCCACCACCACCTCAGCCTGCGCGCTTG
Cre17.g696250	GGAGGCGGAGGTATGGCCCGCAGCGAGG	GGTTCCACCACCACCAGACGACAGCCAGCTGTATG
Cre17.g696350	GGAGGCGGAGGTATGGCCGAAGACTTGGTCG	GGTTCCACCACCACCTCACTTCTTCTTCTTGCTCTCC
Cre17.g696450	GGAGGCGGAGGTATGGCGTCTTGTGCGCACAATG	GGTTCCACCACCACCAAACCCGCTCAGCCGCAA
Cre17.g696550	GGAGGCGGAGGTATGGCCGCTCCACTCCT	GGTTCCACCACCACCTCTCCGCCTGCGTGGC
Cre17.g696626	GGAGGCGGAGGTATGTACCCGAGCTGTCCA	GGTTCCACCACCACCTTGGTTCCTTGAAGGTAGCTC
Cre17.g696800	GGAGGCGGAGGTATGCGCCCACGACTGA	GGTTCCACCACCACCTCGCACAACAATGGCCG
Cre17.g696900	GGAGGCGGAGGTATGCAGCGCGACGGTGGG	GGTTCCACCACCACCGAGCGCCGCCGCTGCGTG
Cre17.g697050	GGAGGCGGAGGTATGCTTCAAATGCGTTTGCAGC	GGTTCCACCACCACCCTGCACCTCCGCCTCCAG
Cre17.g697250	GGAGGCGGAGGTATGGCATATGTATATAGCTGGGC	GGTTCCACCACCACCTCACCGCCAGTGCCAG
Cre17.g697350	GGAGGCGGAGGTATGGCTGCCAAGACGTCAC	GGTTCCACCACCACCCGACTCTGCCCCGGCC
Cre17.g697500	GGAGGCGGAGGTATGGGTCTATTTTCATGCTTCGC	GGTTCCACCACCACCTCAGTTCTTGCCGGGCG
Cre17.g697600	GGAGGCGGAGGTATGACGGTAAACCCTGAAATTACT	GGTTCCACCACCACCCTACTGCACGACACACCTC
Cre17.g697650	GGAGGCGGAGGTATGGTTTCGGCCTCCCA	GGTTCCACCACCACCTCACTTGGCAGCGGGC
Cre17.g697750	GGAGGCGGAGGTATGACGCGCGAGCTGACCAAAG	GGTTCCACCACCACCGAAGTCCGCGACTGCGCGAG
Cre17.g697850	GGAGGCGGAGGTATGATGATTGATGAGTCTACTGAC	GGTTCCACCACCACCTTAGCACGAGACATCCACA
Cre17.g697934	GGAGGCGGAGGTATGGCGGATTCTGGTGCC	GGTTCCACCACCACCTCACGGCCGTGCCACC
Cre17.g698000	GGAGGCGGAGGTATGTTGTCCAGCGTGCGC	GGTTCCACCACCACCCTTCTTGGGCAGGGGCAC
Cre17.g698150	GGAGGCGGAGGTATGACGCCCGACGTCGA	GGTTCCACCACCACCTGGCACAGGAGTCCCCG
Cre17.g698233	GGAGGCGGAGGTATGGCGGCTCAGTGCTA	GGTTCCACCACCACCTCAGTGCATGCCGCCG
Cre17.g698299	GGAGGCGGAGGTATGGATAGCGATCAAAACGCG	GGTTCCACCACCACCCTAACCGCGCCCCTCC
Cre17.g698365	GGAGGCGGAGGTATGGCCGACCCGATGGAC	GGTTCCACCACCACCTCTCCGCCACCACGTCAA
Cre17.g698450	GGAGGCGGAGGTATGTCGACTCCTGCGACCA	GGTTCCACCACCACCCCGGCGTCCTGCGGTC
Cre17.g698516	GGAGGCGGAGGTATGCTCGTGCCACGACG	GGTTCCACCACCACCCTGCTTGGAGCTGGAAACG
Cre17.g698600	GGAGGCGGAGGTATGGTAATATGCACGACGTCGT	GGTTCCACCACCACCCTGCTCTTGCCCACCCAG
Cre17.g698683	GGAGGCGGAGGTATGCTCGCTGCGAGGG	GGTTCCACCACCACCCCCCTGCCTCGGTGCC
Cre17.g698800	GGAGGCGGAGGTATGACCGGGCAGGCAG	GGTTCCACCACCACCCTGCCTAGCAACCGCC
Cre17.g698900	GGAGGCGGAGGTATGGCCGCGGATGGGC	GGTTCCACCACCACCCCTCATCGCCCGGACATAGC
Cre17.g699050	GGAGGCGGAGGTATGCCCCACGCAGACG	GGTTCCACCACCACCGATCCAGCCCCCGCCC
Cre17.g699200	GGAGGCGGAGGTATGCGGCTGTTCGTCAAG	GGTTCCACCACCACCCTTGGCAGTAGCAGCAGC
Cre17.g699350	GGAGGCGGAGGTATGGCGGGCAATCCGGTC	GGTTCCACCACCACCCACACTTGGGCCGCTGGGTAC
Cre17.g699450	GGAGGCGGAGGTATGGAGCATACCGATAGCAGG	GGTTCCACCACCACCCACTCCAAGCCTCGCCA
Cre17.g699600	GGAGGCGGAGGTATGATGGACGCACAGCACCTA	GGTTCCACCACCACCTGCCGCCTGCATGGCG
Cre17.g699700	GGAGGCGGAGGTATGGAGCCCTCCGACAC	GGTTCCACCACCACCATGAGCAACAGGGCCGC
Cre17.g699776	GGAGGCGGAGGTATGGACGAGAACGAATTTGACG	GGTTCCACCACCACCCTAGGACCGGGGGGAATC
Cre17.g699850	GGAGGCGGAGGTATGGCCGTGCGCATCGTG	GGTTCCACCACCACCGATCTTGAGGGCGGGCCAG
Cre17.g699926	GGAGGCGGAGGTATGCTGCGCCGACGCG	GGTTCCACCACCACCTGTGCGTATGAACGGTGACGC
Cre17.g700000	GGAGGCGGAGGTATGGCTAGCTTCGCCGC	GGTTCCACCACCACCGTTCCACTCAGCCGCC
Cre17.g700133	GGAGGCGGAGGTATGACGCAGCAGAGCGTG	GGTTCCACCACCACCCATGATGCGAGACACACTGGC
Cre17.g700250	GGAGGCGGAGGTATGGTGCCAGAGGTACCC	GGTTCCACCACCACCCTCATCATCCTCCCCATCCG
Cre17.g700425	GGAGGCGGAGGTATGGCTATCACGCCCCGC	GGTTCCACCACCACCCGACAGCTGTCGCGCC
Cre17.g700500	GGAGGCGGAGGTATGCTGGACGGGACTCAA	GGTTCCACCACCACCCACGGTGAACAGTGGCTC
Cre17.g700600	GGAGGCGGAGGTATGGCTAGGACACGGTCTT	GGTTCCACCACCACCGTGCGTTCGCATCCGC
Cre17.g700750	GGAGGCGGAGGTATGTCGCCTGTCGCCG	GGTTCCACCACCACCCAGGAAAGCGACGGGGC
Cre17.g700850	GGAGGCGGAGGTATGCCCAGCCGGGAGAC	GGTTCCACCACCACCTCAACCAGGCGCGCCC
Cre17.g700950	GGAGGCGGAGGTATGCTGTGCGCGCGCTCC	GGTTCCACCACCACCCTGGTGCTTGCCGTACTCG
Cre17.g701050	GGAGGCGGAGGTATGGTGTGGTTTCACTGCG	GGTTCCACCACCACCTTATTTCTTCTTGGTCTTCTTCGC
Cre17.g701200	GGAGGCGGAGGTATGCCGTTCAAGCGCTAC	GGTTCCACCACCACCCTTCTTCTTGGGGGCGGC
Cre17.g701250	GGAGGCGGAGGTATGGCGAACATTGACCCC	GGTTCCACCACCACCCGCCCCCAGCTGGATC
Cre17.g701350	GGAGGCGGAGGTATGTCTGAACCAGCGTCGG	GGTTCCACCACCACCGAACGCCTCCGCTATCCG
Cre17.g701450	GGAGGCGGAGGTATGGACAGCGCCAACAC	GGTTCCACCACCACCGAACTCGCCACCATCAAAATC
Cre17.g701551	GGAGGCGGAGGTATGGTGGCGGCACAGG	GGTTCCACCACCACCTCCTGACACTGCCTCAATGA
Cre17.g701700	GGAGGCGGAGGTATGGCTCTGGGCCAGCA	GGTTCCACCACCACCCAGGGCCACCTCGCGG
Cre17.g701800	GGAGGCGGAGGTATGGGCAACTGGCACACA	GGTTCCACCACCACCCTCGCCGTAGTACAGCGG
Cre17.g701850	GGAGGCGGAGGTATGGGTGGCACTCCAAAGAGAA	GGTTCCACCACCACCCTTTCCACAACAGTCAGGTGGC
Cre17.g701917	GGAGGCGGAGGTATGGTCGCCCGTGTCT	GGTTCCACCACCACCCTTTCCACAACAGTCAGGTGG
Cre17.g702002	GGAGGCGGAGGTATGGTCGCCCGTGTCTGC	GGTTCCACCACCACCCTTGCCCCTGGCGGCC
Cre17.g702102	GGAGGCGGAGGTATGCTGAAAGCCAGCCG	GGTTCCACCACCACCCGGAACGAGATGGGGTATGC
Cre17.g702200	GGAGGCGGAGGTATGTTGCGCTCCGCTTGC	GGTTCCACCACCACCCAGCTCCAGGCGCTTAATGAG
Cre17.g702300	GGAGGCGGAGGTATGCGGAGGACCCTGTT	GGTTCCACCACCACCGATGATCTGGAGGCCGGG
Cre17.g702451	GGAGGCGGAGGTATGGCGACGCCTCCCAGT	GGTTCCACCACCACCTGCTGGGTCCGCTGGCATGAAA
Cre17.g702550	GGAGGCGGAGGTATGGGCAAGCGCAAGG	GGTTCCACCACCACCCTTGTTGGCCTCCTCGC
Cre17.g702650	GGAGGCGGAGGTATGTTGCGCCAGCTTTGC	GGTTCCACCACCACCCGCCCGCTTGGCCTGA
Cre17.g702800	GGAGGCGGAGGTATGGCGCAGCGAGCTAGC	GGTTCCACCACCACCCTCGCACTGCCACGCCCG
Cre17.g702900	GGAGGCGGAGGTATGGCCCAGACTCATCATACT	GGTTCCACCACCACCCAGGAAAGCGACGGGGC
Cre17.g702976	GGAGGCGGAGGTATGCCTCCAATGCGTGG	GGTTCCACCACCACCGCTAATGCCACGCCCG
Cre17.g703050	GGAGGCGGAGGTATGAGACCTGCGCTGTCA	GGTTCCACCACCACCCGGCCGGCCAAACTCA
Cre17.g703101	GGAGGCGGAGGTATGCCGCCCACCGAGC	GGTTCCACCACCACCCTGGCGGAGGCGGTGC
Cre17.g703176	GGAGGCGGAGGTATGTGGCATCATACCCCGC	GGTTCCACCACCACCGATGTTGTAGCTGAGGTCCCC
Cre17.g703250	GGAGGCGGAGGTATGGCCGACCGAAACGA	GGTTCCACCACCACCCTACGCCGCAGCATGG
Cre17.g703350	GGAGGCGGAGGTATGGCCGACTTCGGCA	GGTTCCACCACCACCAGGCTGGCTGACCCAG
Cre17.g703473	GGAGGCGGAGGTATGTTCGGAACGGACCCG	GGTTCCACCACCACCCACATCCAGCAGCCCG
Cre17.g703550	GGAGGCGGAGGTATGACAAAATCAAATGAAGGGGAC	GGTTCCACCACCACCTCACTCATACCGTGCCTGC
Cre17.g703650	GGAGGCGGAGGTATGACAGCGCTATCACGC	GGTTCCACCACCACCGCACCTACCCAGACGCC
Cre17.g703800	GGAGGCGGAGGTATGTCCAACGCGACGTGTAC	GGTTCCACCACCACCTGCCCGGTTGGGCTGCAC
Cre17.g703932	GGAGGCGGAGGTATGCCTCCCAACGAGCT	GGTTCCACCACCACCAGCCTCCAGCTCCATAGC
Cre17.g703976	GGAGGCGGAGGTATGTGCTCGGACCGCC	GGTTCCACCACCACCAGGCTGGCCAGCTCTC
Cre17.g704050	GGAGGCGGAGGTATGCTGGCAACATCTGCC	GGTTCCACCACCACCGCTCGGTTCATACTGCCC
Cre17.g704150	GGAGGCGGAGGTATGGCATACCGAGGAGCATC	GGTTCCACCACCACCCGAATAAGCCTGCTCACCACAT
Cre17.g704251	GGAGGCGGAGGTATGACGGAAGCGGAAGCG	GGTTCCACCACCACCGGACGTTTTCAGCTTGATGTC
Cre17.g704400	GGAGGCGGAGGTATGCCACCGCGTGCCAGG	GGTTCCACCACCACCCGGATTCAAAGGTGCCGGT
Cre17.g704500	GGAGGCGGAGGTATGAGCAGAGGCGTCGC	GGTTCCACCACCACCAGGCCCAAAATCTTCACCAG
Cre17.g704600	GGAGGCGGAGGTATGGCAGCTCGCGTACCT	GGTTCCACCACCACCAAACTGCTTGATGGCGAGCAG
Cre17.g704700	GGAGGCGGAGGTATGGGGTTGCAGAGGAGCTT	GGTTCCACCACCACCCGGATGCAACGGTGCGGG
Cre17.g704802	GGAGGCGGAGGTATGTATGTGGCTGCTGTGTG	GGTTCCACCACCACCACTGTGGAGCCATTTGTAGC
Cre17.g704901	GGAGGCGGAGGTATGCTTGCAGCCTTGTCC	GGTTCCACCACCACCCGCCTCACTCATTGATCGC
Cre17.g705050	GGAGGCGGAGGTATGCCAACGCGGCACCCG	GGTTCCACCACCACCCATGTTGCCGGGGCGGTC
Cre17.g705150	GGAGGCGGAGGTATGCACGACGACATCGGCAAG	GGTTCCACCACCACCCACAAACACATGCCCCTTACAC
Cre17.g705250	GGAGGCGGAGGTATGCGCTTGGAGCCAGCT	GGTTCCACCACCACCCTAGTGGCCGGCGCGC
Cre17.g705350	GGAGGCGGAGGTATGGGGGCCTGCTGTT	GGTTCCACCACCACCAGCGTTGTAGATGTCCGC
Cre17.g705450	GGAGGCGGAGGTATGGACGCTGATGACCCG	GGTTCCACCACCACCCGCCGCCACACCCTCA
Cre17.g705550	GGAGGCGGAGGTATGGAGGACTGTGCTGAGGC	GGTTCCACCACCACCCGCCAGCGGCGTAAGCAC
Cre17.g705700	GGAGGCGGAGGTATGACCGCGCCATCGCTA	GGTTCCACCACCACCCGCCGCCGGCAGTGTT
Cre17.g705836	GGAGGCGGAGGTATGCGCTGTGCCACGC	GGTTCCACCACCACCCCACGGCCACCACCGC
Cre17.g705900	GGAGGCGGAGGTATGGGCTGGATGAGCTGG	GGTTCCACCACCACCCTTCTTGCTTGATGTGGGCG
Cre17.g706000	GGAGGCGGAGGTATGGCAGCTGTTGTCGG	GGTTCCACCACCACCCACGCCCGTAGCTCCAA
Cre17.g706150	GGAGGCGGAGGTATGAATCTCGAGGCGCGA	GGTTCCACCACCACCCACGTGCAGCAGCGTG
Cre17.g706250	GGAGGCGGAGGTATGGGAAGCGGTTGTTCAACG	GGTTCCACCACCACCGCTCCTACAGCACGACAGC
Cre17.g706400	GGAGGCGGAGGTATGTCCAACGCTAAGTTCGCTC	GGTTCCACCACCACCTCACGCGAACCGCGCC
Cre17.g706500	GGAGGCGGAGGTATGGCTGAGGACCTGCTG	GGTTCCACCACCACCCACGCATGCCGCCTTA
Cre17.g706600	GGAGGCGGAGGTATGGCCGAAGCCGCTG	GGTTCCACCACCACCCTACTGGCCGGGCGGC
Cre17.g706700	GGAGGCGGAGGTATGTCGCGAGTCACCACG	GGTTCCACCACCACCCTGGTTTGTCGAGCGGC
Cre17.g706850	GGAGGCGGAGGTATGGACTTAAAAATTACGCGTCTGT	GGTTCCACCACCACCCGAGTTGGCGGTGGTGGA
Cre17.g707050	GGAGGCGGAGGTATGAGCACCGGCTGGT	GGTTCCACCACCACCCGACTCCATCGGCTGCC
Cre17.g707150	GGAGGCGGAGGTATGGCGGAGGTTGACGCC	GGTTCCACCACCACCCAGGCACACAAACTGCCG
Cre17.g707200	GGAGGCGGAGGTATGACGAGGGCCGCGA	GGTTCCACCACCACCCATATTTCCCGGCCGGTC
Cre17.g707350	GGAGGCGGAGGTATGGACATCGTGGACCTAACT	GGTTCCACCACCACCGTTGCCGAAGCCTGATGT
Cre17.g707450	GGAGGCGGAGGTATGTACGAGGCCTCGCT	GGTTCCACCACCACCCAGTGATGCGTCTGCG
Cre17.g707600	GGAGGCGGAGGTATGATCATTCGCATCAGCCTG	GGTTCCACCACCACCGTCGTCTAGCGCCGCC
Cre17.g707750	GGAGGCGGAGGTATGCTGCGGCAGGCTT	GGTTCCACCACCACCCATGTTCTTGCGCATCATGG
Cre17.g707850	GGAGGCGGAGGTATGGGTTCTGACAACAGGGA	GGTTCCACCACCACCGACACGCCGCTTCTTGG
Cre17.g707950	GGAGGCGGAGGTATGCTACAAACCAAGTCTTGGG	GGTTCCACCACCACCCTGCTTGTTCGGCTCCGC
Cre17.g708013	GGAGGCGGAGGTATGGTGACCGGCAGCA	GGTTCCACCACCACCGGACGCATCAGCGCGT
Cre17.g708050	GGAGGCGGAGGTATGCATGGCGAGGAGGGC	GGTTCCACCACCACCATGTGCAACAGCAACGACGTT
Cre17.g708100	GGAGGCGGAGGTATGGTCTCTGTCGCTGCG	GGTTCCACCACCACCCTCGTCGCCGCTCTCGCC
Cre17.g708200	GGAGGCGGAGGTATGTCCGGACGTGGCA	GGTTCCACCACCACCGCCACCGAAGCCGTAC
Cre17.g708300	GGAGGCGGAGGTATGGCCCTTCAGGAGGCC	GGTTCCACCACCACCTACTATCCGCTCGAGCTCCTT
Cre17.g708400	GGAGGCGGAGGTATGGACCAAGCAGCCAGGT	GGTTCCACCACCACCGCTGAGCGTCGCCGGC
Cre17.g708550	GGAGGCGGAGGTATGGCTGGTCGCGGCA	GGTTCCACCACCACCGGCGGTCTCCTCGCCC
Cre17.g708650	GGAGGCGGAGGTATGTCCGGACGTGGCA	GGTTCCACCACCACCGCCACCGAAGCCGTAC
Cre17.g708750	GGAGGCGGAGGTATGGCTCGCTCGCTTAGT	GGTTCCACCACCACCTTAGTTGGAAGTTGTGGAGTTGTA
Cre17.g708850	GGAGGCGGAGGTATGGCCGGGGTCGGGG	GGTTCCACCACCACCGGCGTTTGCACGCAGCT
Cre17.g708950	GGAGGCGGAGGTATGCGAGTCGCAGGAGG	GGTTCCACCACCACCCAGCACTTGCCAGACCG
Cre17.g709100	GGAGGCGGAGGTATGTCCGGACGTGGCA	GGTTCCACCACCACCGCCACCGAAGCCGTAC
Cre17.g709200	GGAGGCGGAGGTATGGCTGGTCGCGGCA	GGTTCCACCACCACCGGCGGTCTCCTCGCCC
Cre17.g709300	GGAGGCGGAGGTATGGGCCGACTAAATTGCG	GGTTCCACCACCACCCACGCAAAGCTTGTTCTGG
Cre17.g709400	GGAGGCGGAGGTATGCAAGAGAACACAGGGGGAA	GGTTCCACCACCACCTTGGTGCTGCTGCTGCCG
Cre17.g709500	GGAGGCGGAGGTATGGCCGAGGACCTCGGG	GGTTCCACCACCACCTGCGTCGGCACCACCGGC
Cre17.g709550	GGAGGCGGAGGTATGAGTAGCTCTTATGAAGCGGG	GGTTCCACCACCACCCTCCCGCCTCAACTCGAC
Cre17.g709651	GGAGGCGGAGGTATGTCTGCATGTCGCGGC	GGTTCCACCACCACCTCAGCAAGCGGCGGCG
Cre17.g709750	GGAGGCGGAGGTATGCGAATAGCACTCCGCG	GGTTCCACCACCACCAGACTGCAGAAGCTTCTGAAAC
Cre17.g709800	GGAGGCGGAGGTATGGCAACCGTCACGC	GGTTCCACCACCACCCAGCCTGTCACCGCTG
Cre17.g709876	GGAGGCGGAGGTATGAAGCTTCCACGGCG	GGTTCCACCACCACCCACGCATGAGCACTCAGAA
Cre17.g710000	GGAGGCGGAGGTATGCCGCGGCTTCAAGCCATC	GGTTCCACCACCACCCTGGTGTTGCATGGGCGA
Cre17.g710100	GGAGGCGGAGGTATGGCGGAAGAGACTCAGGTTC	GGTTCCACCACCACCCGACACGATGTAGTCGCTCTC
Cre17.g710200	GGAGGCGGAGGTATGCACGGGCGGCAACCCATG	GGTTCCACCACCACCCGGCGGCACCAGCAAGTG
Cre17.g710300	GGAGGCGGAGGTATGCTCCTCGCCAAGGG	GGTTCCACCACCACCGTGCCTGGCGCCAGTA
Cre17.g710400	GGAGGCGGAGGTATGGCTGGTCGCGGCA	GGTTCCACCACCACCGGCGGTCTCCTCGCCC
Cre17.g710500	GGAGGCGGAGGTATGTCCGGACGTGGCA	GGTTCCACCACCACCGCCACCGAAGCCGTAC
Cre17.g710650	GGAGGCGGAGGTATGTCCAGTCGTCACAACGAAG	GGTTCCACCACCACCGTCCCCAACGCTGCTGCAT
Cre17.g710750	GGAGGCGGAGGTATGTCAATGAACCCCTTCGAG	GGTTCCACCACCACCTTACCGCCGCTCCATGG
Cre17.g710850	GGAGGCGGAGGTATGGGGCAGCTGCTGACG	GGTTCCACCACCACCCCGGTGCGCGTGGTGG
Cre17.g710879	GGAGGCGGAGGTATGCAACCACAAGCCGC	GGTTCCACCACCACCACATCCCCCTGTGCGC
Cre17.g710881	GGAGGCGGAGGTATGGCGCAGCAGGACG	GGTTCCACCACCACCAGCCGTCTTCTTGTGAGCG
Cre17.g710900	GGAGGCGGAGGTATGCGCGCGGCTTTGG	GGTTCCACCACCACCCTAGCACGCCGCCACC
Cre17.g711000	GGAGGCGGAGGTATGCCTGGCCCGAACT	GGTTCCACCACCACCCTAGCACGCCGCCACC
Cre17.g711100	GGAGGCGGAGGTATGCACAGTAGGCATTCCG	GGTTCCACCACCACCGGATCCACGGGGACGC
Cre17.g711200	GGAGGCGGAGGTATGAAGTCTCAATTGTCGACGC	GGTTCCACCACCACCCTGCTGCTGCTGCTGCTG
Cre17.g711300	GGAGGCGGAGGTATGCCCCAGAAGCCGA	GGTTCCACCACCACCTGTCAGCGGCAGCAGC
Cre17.g711401	GGAGGCGGAGGTATGACCTCAGCGTTCCTGC	GGTTCCACCACCACCGCCCCGAGCACTGACGGC
Cre17.g711480	GGAGGCGGAGGTATGCGGCACGCGCTGA	GGTTCCACCACCACCGACAACACCGGCAGGTCTC
Cre17.g711526	GGAGGCGGAGGTATGAGCGGCGAATACGC	GGTTCCACCACCACCTCAGCCTGCGCGCTTG
Cre17.g711600	GGAGGCGGAGGTATGCGGGGGGGCAGGA	GGTTCCACCACCACCGGCGTGGTACTGGACGC
Cre17.g711700	GGAGGCGGAGGTATGGCTGGTCGCGGCA	GGTTCCACCACCACCGGCGGTCTCCTCGCCC
Cre17.g711800	GGAGGCGGAGGTATGTCCGGACGTGGCA	GGTTCCACCACCACCGCCACCGAAGCCGTAC
Cre17.g711900	GGAGGCGGAGGTATGAAGCAACGATACGGGC	GGTTCCACCACCACCAAGCTCCTCATACACGTCGTA
Cre17.g712000	GGAGGCGGAGGTATGGCGCGTGTGTATGTC	GGTTCCACCACCACCGTCGCGTACAGGCGAGC
Cre17.g712100	GGAGGCGGAGGTATGCCCGCCACCTACAAG	GGTTCCACCACCACCCAGCTTGGCCTTGATGGCC
Cre17.g712200	GGAGGCGGAGGTATGCCCACCCGCCGGGAC	GGTTCCACCACCACCCGCCTCCGCAGCCTTGCC
Cre17.g712300	GGAGGCGGAGGTATGAGGTCTGCGCTGCTCAAG	GGTTCCACCACCACCCCAGAGCCGGCCCTTGGCAATC
Cre17.g712400	GGAGGCGGAGGTATGGCAGGCTGTCCGC	GGTTCCACCACCACCGTCGTCCTTGCTCTCCGG
Cre17.g712500	GGAGGCGGAGGTATGGGTCAATGCACAAGCA	GGTTCCACCACCACCGATACCCGCCACACCCT
Cre17.g712600	GGAGGCGGAGGTATGACGAAGCGTTCTCGTC	GGTTCCACCACCACCTCGGGACCCATACACCG
Cre17.g712742	GGAGGCGGAGGTATGTGCGAGGCAGACCA	GGTTCCACCACCACCGATCCGCAATGATCCCTGTG
Cre17.g712800	GGAGGCGGAGGTATGAGACTAACTCCCGAGCTC	GGTTCCACCACCACCGGCCGAATGGTCGCCG
Cre17.g713000	GGAGGCGGAGGTATGGATGACTCGAAGCGGA	GGTTCCACCACCACCCTTATCAGCCGCCTTGCC

Cre16.g695202	GGAGGCGGAGGTATGGACGCGTCCTCGC	GGTTCCACCACCACCCGTCGCGGCATAGCGC
Cre16.g695250	GGAGGCGGAGGTATGAGCGCCTACGCGG	GGTTCCACCACCACCCTATGCCGGCCCTGCT
Cre16.g695450	GGAGGCGGAGGTATGGCTGACAAACGGGACC	GGTTCCACCACCACCGTTCTCCGGCACCGCG
Cre16.g695550	GGAGGCGGAGGTATGGAGAACGAGTTTGGAGC	GGTTCCACCACCACCATAAATGGCCTTGCCGCC
Cre16.g695650	GGAGGCGGAGGTATGACGCGGGTGTCGGTC	GGTTCCACCACCACCCTGGGGCGTGTAGAATGAGG
Cre16.g695800	GGAGGCGGAGGTATGCACGTCGCTTCCAC	GGTTCCACCACCACCGCTGCGCTTCCCCTGG
Cre16.g695851	GGAGGCGGAGGTATGGAAGGGGGCCCCGCT	GGTTCCACCACCACCCACCACCTGTTCCCCCAAC
Cre16.g695950	GGAGGCGGAGGTATGTGCGCGGCACCCCGC	GGTTCCACCACCACCGTGCTGTCCGCTGCCGCTG
Cre17.g692564	GGAGGCGGAGGTATGAGGCAGCGGGGTC	GGTTCCACCACCACCGGGGTAGCTGCTGTGCG
Cre17.g696200	GGAGGCGGAGGTATGGCAGGGGCGTTGT	GGTTCCACCACCACCGTCTTCCTTGTCGTCTCCAG
Cre17.g696300	GGAGGCGGAGGTATGGACGCTTCAGCACCT	GGTTCCACCACCACCGTCCTCGTCCCATGTGGG
Cre17.g696400	GGAGGCGGAGGTATGGCCGCTTCCTGCCTTC	GGTTCCACCACCACCGCAGCCACCCGGCTCAAT
Cre17.g696500	GGAGGCGGAGGTATGCTCCTACGAGTAACCGC	GGTTCCACCACCACCCCGCGACTCGGGTGTCGT
Cre17.g696600	GGAGGCGGAGGTATGTCGGGACAGCAGGGC	GGTTCCACCACCACCAATCCTAGACGTAGACGCCACA
Cre17.g696700	GGAGGCGGAGGTATGGCGCGACACTGTATACTG	GGTTCCACCACCACCTACGCTGGGGAAGCTGGT
Cre17.g696850	GGAGGCGGAGGTATGGCAGAACAAGTTGCGGA	GGTTCCACCACCACCCGCAGCGTGCGGCTTC
Cre17.g697000	GGAGGCGGAGGTATGGAGGAGCAAACATACAGTCT	GGTTCCACCACCACCCTCGGACTCGACCACACG
Cre17.g697200	GGAGGCGGAGGTATGACGACTGCGGCGACG	GGTTCCACCACCACCCAAGTCACGCTGCACCCC
Cre17.g697334	GGAGGCGGAGGTATGGCATTGCTAGGTGGCAAG	GGTTCCACCACCACCCGAGTACATGTAGTCCTGGCC
Cre17.g697406	GGAGGCGGAGGTATGTGCGCTGCGACATC	GGTTCCACCACCACCCTACTGCTGGGCCGGC
Cre17.g697550	GGAGGCGGAGGTATGCAGAGTTTGAACGTACCT	GGTTCCACCACCACCGCAGCTCGACGTCGTTG
Cre17.g697624	GGAGGCGGAGGTATGGTCGGCGTGTTTGGTC	GGTTCCACCACCACCTGCCGGCATTCTCGGCTC
Cre17.g697701	GGAGGCGGAGGTATGACGCTGGCGTGTAGG	GGTTCCACCACCACCATGCACCACATGCACAACC
Cre17.g697800	GGAGGCGGAGGTATGGACCGAAATGGCGACAA	GGTTCCACCACCACCACAGTACCGGAACTTGGCG
Cre17.g697900	GGAGGCGGAGGTATGCGTTTGTCTGCAAGGGC	GGTTCCACCACCACCCTGCTCCGCCTGCGGC
Cre17.g697950	GGAGGCGGAGGTATGGCTTGCGTTGGCCT	GGTTCCACCACCACCTCATGCCGCCGCAGCA
Cre17.g698100	GGAGGCGGAGGTATGGCGCCAGTGGTGCTC	GGTTCCACCACCACCAATGCGTGTGCCGCCGCTG
Cre17.g698200	GGAGGCGGAGGTATGCGAGAAAATGGACTTATAACG	GGTTCCACCACCACCTCACACATCAATGATGTCGTCC
Cre17.g698266	GGAGGCGGAGGTATGGACGACAGCGTGAAGTA	GGTTCCACCACCACCCGCAAACAGCTGCCTCA
Cre17.g698332	GGAGGCGGAGGTATGGAGGGGAGTTGCGT	GGTTCCACCACCACCTGTCAGCGGCAGCAGC
Cre17.g698400	GGAGGCGGAGGTATGCCCACGAACACGAAG	GGTTCCACCACCACCATCGGCCTCCTCCGAGG
Cre17.g698500	GGAGGCGGAGGTATGTCCCAGAGGAGCTCTAG	GGTTCCACCACCACCGTCCCAGACTTCGTCCTTGA
Cre17.g698532	GGAGGCGGAGGTATGTGTCCAATCCAAAAGCAGC	GGTTCCACCACCACCCTATAGCTGGGGAGCCCG
Cre17.g698650	GGAGGCGGAGGTATGGATGTAGAGGCTTCAGCAGA	GGTTCCACCACCACCCTAGCCTGCCGCCGCC
Cre17.g698750	GGAGGCGGAGGTATGCGCGTTGCCAGCTC	GGTTCCACCACCACCCCGCCTCCTGCGTGCG
Cre17.g698850	GGAGGCGGAGGTATGATGATACAAGGACACGTCC	GGTTCCACCACCACCTCAGTGCTTCCGGGCC
Cre17.g699000	GGAGGCGGAGGTATGGCCTTCGCGAGCAGC	GGTTCCACCACCACCCTTGGGCGCTGCTGCCGC
Cre17.g699100	GGAGGCGGAGGTATGGACGCCCGCCACAGA	GGTTCCACCACCACCGTACGCAAACACGTCCAGAG
Cre17.g699250	GGAGGCGGAGGTATGTATGCTAGGCGCGGC	GGTTCCACCACCACCGCAGCACTTCCCGCCG
Cre17.g699400	GGAGGCGGAGGTATGGCCAGCGAGGATGCG	GGTTCCACCACCACCAGGCCGGTATCGTATCCGC
Cre17.g699500	GGAGGCGGAGGTATGCTTGGGGCGCCTCGCAAC	GGTTCCACCACCACCGGCGTTGCCCGGCTTGGC
Cre17.g699650	GGAGGCGGAGGTATGGAAGACCTAACTGTTCCAAG	GGTTCCACCACCACCCTAGCTCCCAAAGCGCTG
Cre17.g699750	GGAGGCGGAGGTATGGCGTCGCCCAGCC	GGTTCCACCACCACCTCAACCCGTCAGGGCCG
Cre17.g699800	GGAGGCGGAGGTATGTGCAAGCCGGCGA	GGTTCCACCACCACCCTCGTTCAGGTGCCGGC
Cre17.g699900	GGAGGCGGAGGTATGGCTGCGATGCGGAGG	GGTTCCACCACCACCGTCCAGGTCGTTGGTGTTCTTC
Cre17.g699950	GGAGGCGGAGGTATGGCCAGCGGCGAGG	GGTTCCACCACCACCGCGACCGAACCACTCCAT
Cre17.g700100	GGAGGCGGAGGTATGCCTACCCTCAATATTATCACG	GGTTCCACCACCACCTTAAGCAAAGGTGGAGCCG
Cre17.g700200	GGAGGCGGAGGTATGTCTGCTCTGCTGCCG	GGTTCCACCACCACCGTCAGCATTGCCGGCTC
Cre17.g700300	GGAGGCGGAGGTATGACGCGCGCGGACGGC	GGTTCCACCACCACCAGCCGCCGCCACGCCGCC
Cre17.g700450	GGAGGCGGAGGTATGTCCAACCTAAACGGCG	GGTTCCACCACCACCGATGTAGGCGACGGGCC
Cre17.g700550	GGAGGCGGAGGTATGGTTTGTTGTCCAGGCTG	GGTTCCACCACCACCGTGCGTTCGCATCCGC
Cre17.g700650	GGAGGCGGAGGTATGCTCCCACCGCCTC	GGTTCCACCACCACCGATGAAGTGCACAGGGCG
Cre17.g700800	GGAGGCGGAGGTATGCAACGCCCTCAAAGC	GGTTCCACCACCACCCGTGATAAAGATGCGCTCGC
Cre17.g700900	GGAGGCGGAGGTATGCCCAACGCTGACG	GGTTCCACCACCACCGCGCCACAACCCACTAC
Cre17.g701000	GGAGGCGGAGGTATGGCTCAACACGCAACA	GGTTCCACCACCACCGTGCTGCGACTGAGCC
Cre17.g701100	GGAGGCGGAGGTATGGCAGGTGCAGAACAGC	GGTTCCACCACCACCTCACGACACCGCCGCC
Cre17.g701250	GGAGGCGGAGGTATGGCGAACATTGACCCCTATC	GGTTCCACCACCACCCGCCCCCAGCTGGATCGTC
Cre17.g701300	GGAGGCGGAGGTATGGGCTTGGTGGTAATCTCC	GGTTCCACCACCACCGCTGTTCGCACTCCCTG
Cre17.g701400	GGAGGCGGAGGTATGGCTCTCTTGAGCACTATTC	GGTTCCACCACCACCTCAGACGTCACCTCTTCCA
Cre17.g701500	GGAGGCGGAGGTATGTTTGGAGGGTTCCCGTTC	GGTTCCACCACCACCCTGCTGCGCGCACTGCAC
Cre17.g701650	GGAGGCGGAGGTATGGTGAAGTTCCTGAAGCC	GGTTCCACCACCACCGAAGCGGAGCTTGCTGAA
Cre17.g701751	GGAGGCGGAGGTATGGAAACGTTCGGCCGC	GGTTCCACCACCACCTTTTTTGACGGCGCGCGC
Cre17.g701809	GGAGGCGGAGGTATGGCCACCATCAAATTCTACC	GGTTCCACCACCACCTTAGCGTTCGCCCGTCT
Cre17.g701884	GGAGGCGGAGGTATGGAGCCGTCTAGCTCG	GGTTCCACCACCACCGCACTTCCCACCACTGC
Cre17.g701950	GGAGGCGGAGGTATGGAGCCGTCTAGCTCG	GGTTCCACCACCACCGCACTTCCCACCACTGC
Cre17.g702052	GGAGGCGGAGGTATGGGTGTAGGTTTTGGAAGAG	GGTTCCACCACCACCTCAGCCATTCGCACCCC
Cre17.g702150	GGAGGCGGAGGTATGCAGACGAGTGTGTCAAC	GGTTCCACCACCACCTTACGCGTGGTCGCGG
Cre17.g702250	GGAGGCGGAGGTATGGATTCCGCCCGGAC	GGTTCCACCACCACCACCCCTCTTGGCCGCA
Cre17.g702400	GGAGGCGGAGGTATGGAAGTTGAGCGCGACG	GGTTCCACCACCACCTTAGGCCAGCGCCCCG
Cre17.g702500	GGAGGCGGAGGTATGTCGCTGCTCCTAGGAAAC	GGTTCCACCACCACCGATGGTCGGGGGCTCGCAG
Cre17.g702600	GGAGGCGGAGGTATGGAGGGTCTGGCCAAAGT	GGTTCCACCACCACCGGGCGCGGCGGTGTAG
Cre17.g702750	GGAGGCGGAGGTATGTTCGGCAAGCGACTGG	GGTTCCACCACCACCCAGCGCTGCCCCCGAC
Cre17.g702850	GGAGGCGGAGGTATGCCAGAAGCGGGGG	GGTTCCACCACCACCAACGCCGCTAGCTGACG
Cre17.g702950	GGAGGCGGAGGTATGGCTCACGCTCATCATAC	GGTTCCACCACCACCGATGTAGGCGACGGGGC
Cre17.g703000	GGAGGCGGAGGTATGAAGCTTCAGCAACACCT	GGTTCCACCACCACCGCTCGAAGGAGGATCGCG
Cre17.g703075	GGAGGCGGAGGTATGCCGACGCGGCGAA	GGTTCCACCACCACCAAACTGTCCGTGCGTGCG
Cre17.g703126	GGAGGCGGAGGTATGGGGCCTGGCTGGG	GGTTCCACCACCACCTGCCATCATGGCCTCCTG
Cre17.g703200	GGAGGCGGAGGTATGCACTTCTTCAGCCACGC	GGTTCCACCACCACCGTGAGCGGCGATGATGCC
Cre17.g703300	GGAGGCGGAGGTATGTCAGTGGCATCTGCGATG	GGTTCCACCACCACCACCTGCCTTCCCCACCTT
Cre17.g703400	GGAGGCGGAGGTATGAATGCGGTCGCATGC	GGTTCCACCACCACCCTGCTTTGACTGCTTGAAGGC
Cre17.g703495	GGAGGCGGAGGTATGTACGCTGTGGCCACC	GGTTCCACCACCACCCGTCCCCACACCGTCC
Cre17.g703600	GGAGGCGGAGGTATGATAGCGGACTACATGCGC	GGTTCCACCACCACCAAAGACGCTGCCTGCCGAGAAC
Cre17.g703700	GGAGGCGGAGGTATGCTTCCAGCGCTGAC	GGTTCCACCACCACCGGCGATGGAGGCCACG
Cre17.g703850	GGAGGCGGAGGTATGGCTCAGAAGAGCACCCTTA	GGTTCCACCACCACCCACCTTGATGGCGCGCTG
Cre17.g703950	GGAGGCGGAGGTATGGCCTTGATGCTGGTCC	GGTTCCACCACCACCCCCCTGCTTGTTGTCCTTG
Cre17.g704000	GGAGGCGGAGGTATGGATCAGTCGCAGGTGC	GGTTCCACCACCACCCAGGCGGAACGCGGTC
Cre17.g704100	GGAGGCGGAGGTATGCATTTGGAGAGCTTCCTC	GGTTCCACCACCACCCAAAAGCTGTTTCCGCCAC
Cre17.g704200	GGAGGCGGAGGTATGGGTGACGCAGGGG	GGTTCCACCACCACCCTAGCTGGCGGCCAGG
Cre17.g704350	GGAGGCGGAGGTATGGGGACGAAGCCACATC	GGTTCCACCACCACCCACGCCGTTGGCTGCCAT
Cre17.g704450	GGAGGCGGAGGTATGAGCCCTAATGCTCAGCC	GGTTCCACCACCACCAACCCTTAAGAAACCAAAGTGCA
Cre17.g704550	GGAGGCGGAGGTATGTACCAGCCCGGCCCG	GGTTCCACCACCACCCGGCCTCACTCCCAGGTAGAAG
Cre17.g704650	GGAGGCGGAGGTATGCCGCGGATCGAGG	GGTTCCACCACCACCCTGCGCAGAGTCCTGACG
Cre17.g704750	GGAGGCGGAGGTATGCCACCTCGAAGGCA	GGTTCCACCACCACCAGCGCCACAGTACCGG
Cre17.g704850	GGAGGCGGAGGTATGGCTGACGTTGAGGCC	GGTTCCACCACCACCCAGGCCCTCCTTCTCCAC
Cre17.g705000	GGAGGCGGAGGTATGGCGAAGGTTGGAACTCTG	GGTTCCACCACCACCCAACCCGCGCATCATTGC
Cre17.g705100	GGAGGCGGAGGTATGCACAGCATTAAGTCGTCT	GGTTCCACCACCACCCTGCCGCCGGAACTTATC
Cre17.g705200	GGAGGCGGAGGTATGAGGGCCCGGAAAACG	GGTTCCACCACCACCGGCCTCGTCATCGCCTCC
Cre17.g705300	GGAGGCGGAGGTATGGCACGCTCAGCAGT	GGTTCCACCACCACCACCCCGGCTGGGCTTA
Cre17.g705400	GGAGGCGGAGGTATGAGTCGAGGAGGCAGCG	GGTTCCACCACCACCATCGCGCTCGCTGATGATG
Cre17.g705500	GGAGGCGGAGGTATGGGGACCAGCAAGCC	GGTTCCACCACCACCTACGCCCGTGACGTCG
Cre17.g705600	GGAGGCGGAGGTATGTTCGCTTACTGCCGG	GGTTCCACCACCACCCTAAGACGGCGCCGCA
Cre17.g705806	GGAGGCGGAGGTATGCTGTGTGCCGCGC	GGTTCCACCACCACCAAGCTTGCGCCACACCC
Cre17.g705850	GGAGGCGGAGGTATGATGGAGCAGCATAAAGGC	GGTTCCACCACCACCCAGACGACGCAACGGACA
Cre17.g705950	GGAGGCGGAGGTATGGCCTATGAATGCGAGTATA	GGTTCCACCACCACCGCAGCCCAGGTAGCTATCCTC
Cre17.g706050	GGAGGCGGAGGTATGCCGTGCCACGCCA	GGTTCCACCACCACCCTGGCGCTCCTGCTGC
Cre17.g706200	GGAGGCGGAGGTATGGGCGTTGACTGCTTAG	GGTTCCACCACCACCTCACGCCGACGACCAC
Cre17.g706350	GGAGGCGGAGGTATGCGACCTGGCGCAG	GGTTCCACCACCACCCAAGGCTGGCGGTGGC
Cre17.g706450	GGAGGCGGAGGTATGAACACTGTGGCTGTCATTA	GGTTCCACCACCACCTCACATGTTGCCGGGCC
Cre17.g706550	GGAGGCGGAGGTATGCCCTGCATCCTGTTCGT	GGTTCCACCACCACCCTACGCGGGCGCCTGC
Cre17.g706650	GGAGGCGGAGGTATGCGGCGAGGTCCGG	GGTTCCACCACCACCTCACCCCGGTGCCGGG
Cre17.g706800	GGAGGCGGAGGTATGACCGGTCCGGACCCC	GGTTCCACCACCACCCTTGCCAGCCTCAAACTTGG
Cre17.g706900	GGAGGCGGAGGTATGGCGGAAGCTGAAACGAAG	GGTTCCACCACCACCCATGCCGCCCCGTGAAGTC
Cre17.g707137	GGAGGCGGAGGTATGCTCGACGAGTCTACGCAG	GGTTCCACCACCACCCTTAAACAAGATTGCGTCCAGCT
Cre17.g707200	GGAGGCGGAGGTATGACGAGGGCCGCGAGG	GGTTCCACCACCACCCATATTTCCCGGCCGGTCG
Cre17.g707300	GGAGGCGGAGGTATGTCTTCAGATAACGGCTGGC	GGTTCCACCACCACCTGACGCCTTGGCGTGCAC
Cre17.g707400	GGAGGCGGAGGTATGGGCGGGGAGACGA	GGTTCCACCACCACCGCGCGTGGGATAGGCA
Cre17.g707500	GGAGGCGGAGGTATGCCGAAACTGCCGAAT	GGTTCCACCACCACCCTTCTTGGGGCGAGCG
Cre17.g707650	GGAGGCGGAGGTATGGGCGGCTGCCTGAGC	GGTTCCACCACCACCTGCCGACACCTTGATGGC
Cre17.g707800	GGAGGCGGAGGTATGGGGCAGTGCGTCT	GGTTCCACCACCACCGTGTTGCATAGGGTCGTATCC
Cre17.g707900	GGAGGCGGAGGTATGAGGTCTGCAGCGTACCT	GGTTCCACCACCACCGCCCTGCTTCGCCTTGCC
Cre17.g708000	GGAGGCGGAGGTATGAGCGACCATCTCAGCC	GGTTCCACCACCACCATTCATGCTGGAGGCTACAC
Cre17.g708038	GGAGGCGGAGGTATGTGGCGCTTGCGCT	GGTTCCACCACCACCGCAAAGCGCAGGACACT
Cre17.g708076	GGAGGCGGAGGTATGACTGACCTGCCCGC	GGTTCCACCACCACCCAAGCGGGCCAAGGTG
Cre17.g708150	GGAGGCGGAGGTATGGCCCGCACCAAGC	GGTTCCACCACCACCTTAAGCGCGCTCGCCA
Cre17.g708250	GGAGGCGGAGGTATGCCACCAACTAGCGGCATC	GGTTCCACCACCACCGGACTTGGCGTGCTTGGT
Cre17.g708350	GGAGGCGGAGGTATGAAGTTTTCCCTACTTTTGCCG	GGTTCCACCACCACCTTACGAGTCCTTTTTCGTGGAG
Cre17.g708450	GGAGGCGGAGGTATGCTCCGATATCGCGCG	GGTTCCACCACCACCACCCCGCTTCTTAGGAGGT
Cre17.g708600	GGAGGCGGAGGTATGGCCCCCAAGAAGGAC	GGTTCCACCACCACCGCCCGAGGTGAACTTGG
Cre17.g708700	GGAGGCGGAGGTATGGCCCGCACCAAGCA	GGTTCCACCACCACCGGCGCGCTCTCCACGG
Cre17.g708800	GGAGGCGGAGGTATGCAGCACGAAATTCCTCGAG	GGTTCCACCACCACCCTCCACCAGGTAAGGGCTATCC
Cre17.g708901	GGAGGCGGAGGTATGATAAGGCTGTGCTCGTTAA	GGTTCCACCACCACCTTACTGCCCTGCTGGCG
Cre17.g709050	GGAGGCGGAGGTATGGCCCGCACCAAGCA	GGTTCCACCACCACCGGCACGCTCGCCACGG
Cre17.g709150	GGAGGCGGAGGTATGGCCCCCAAGAAGGAC	GGTTCCACCACCACCGCCCGAGGTGAACTTGG
Cre17.g709250	GGAGGCGGAGGTATGGAGAGCCTATCCAAACTCT	GGTTCCACCACCACCCTTGGCCACCATGGCGTA
Cre17.g709350	GGAGGCGGAGGTATGCCATCAGGCTTTCGGTTAG	GGTTCCACCACCACCCAGGAACCGCGCCAGCTT
Cre17.g709450	GGAGGCGGAGGTATGGTCTGTCCTATGCGCT	GGTTCCACCACCACCTTATGCAGCGCGGGCC
Cre17.g709500	GGAGGCGGAGGTATGGCCGAGGACCTCGG	GGTTCCACCACCACCTGCGTCGGCACCACCG
Cre17.g709600	GGAGGCGGAGGTATGGGCGCTCAGCGTGCGTTG	GGTTCCACCACCACCCCCGCGTGTGCGCTTGAG
Cre17.g709700	GGAGGCGGAGGTATGTATAGGTTCCGAATTGCTGT	GGTTCCACCACCACCTCACGGCTGTGGCTGC
Cre17.g709800	GGAGGCGGAGGTATGGCAACCGTCACGCTG	GGTTCCACCACCACCCAGCCTGTCACCGCTGGG
Cre17.g709850	GGAGGCGGAGGTATGCAACGCGTCTTGGTC	GGTTCCACCACCACCTTATGCCGCTGCGGCT
Cre17.g709950	GGAGGCGGAGGTATGCCCGCTCCGGCCT	GGTTCCACCACCACCTGACCCCTGCGGTGGC
Cre17.g710050	GGAGGCGGAGGTATGTGGCAGTTCCAGGGT	GGTTCCACCACCACCCAGACCCGCGCCACTG
Cre17.g710150	GGAGGCGGAGGTATGTCCTCAGAGGCGGAGAAG	GGTTCCACCACCACCCTTCTGGCCGCCGTCCCC
Cre17.g710254	GGAGGCGGAGGTATGCTCCCGGCGACTG	GGTTCCACCACCACCCGAGCCAGTGACGTGGA
Cre17.g710351	GGAGGCGGAGGTATGGCGGCTGTAGCAGG	GGTTCCACCACCACCGTGAACCATGCCACGCC
Cre17.g710450	GGAGGCGGAGGTATGGCCCCCAAGAAGGAC	GGTTCCACCACCACCGCCCGAGGTGAACTTGG
Cre17.g710550	GGAGGCGGAGGTATGGCCCGCACCAAGC	GGTTCCACCACCACCTTAGGCGCGCTCGCCA
Cre17.g710700	GGAGGCGGAGGTATGGCGTTCAAGACCAACA	GGTTCCACCACCACCTCAGCGGCACCACCTG
Cre17.g710800	GGAGGCGGAGGTATGTTGCTGCGCAACAAGAC	GGTTCCACCACCACCGGCCTCCACCAGCCGCAC
Cre17.g710871	GGAGGCGGAGGTATGCGCCACAGGGCAG	GGTTCCACCACCACCGCCGCTTTGGACCCCA
Cre17.g710881	GGAGGCGGAGGTATGGCTCCCGAGGACAG	GGTTCCACCACCACCTCCGGCACCATCCCAG
Cre17.g710881	GGAGGCGGAGGTATGGCGCAGCAGGACG	GGTTCCACCACCACCCTAGCCTGGGGCGCCG
Cre17.g710950	GGAGGCGGAGGTATGGCGCAGCAGGACG	GGTTCCACCACCACCAGCCGTCTTCTTGTGAGCG
Cre17.g711050	GGAGGCGGAGGTATGCCGCTGGTAGTCATTTGC	GGTTCCACCACCACCCACGTATTCGCCCGCTGC
Cre17.g711150	GGAGGCGGAGGTATGACGGTCACTCGGCGCAAG	GGTTCCACCACCACCGCGGTGGTACCAGTACACC
Cre17.g711250	GGAGGCGGAGGTATGCGCAGCATTCGAATTCC	GGTTCCACCACCACCCACGTAGCCTCCCGCACT
Cre17.g711350	GGAGGCGGAGGTATGGATAGCGATCAAAACGCG	GGTTCCACCACCACCCTAACCGCGCCCCTCC
Cre17.g711450	GGAGGCGGAGGTATGACAGACCGAACACAAGACC	GGTTCCACCACCACCGCTGTTGAACCAGTCGAAGAAC
Cre17.g711503	GGAGGCGGAGGTATGTGGTTGAGGGTGGCG	GGTTCCACCACCACCAACCCAGACTTTGTCGTGC
Cre17.g711550	GGAGGCGGAGGTATGTTGGCAGCGGTGGG	GGTTCCACCACCACCGTTACTTGACAGCAATGGCG
Cre17.g711650	GGAGGCGGAGGTATGCCGCCGCACTGCGTTATC	GGTTCCACCACCACCCATGATCTTGGCCCCGGAAAAG
Cre17.g711750	GGAGGCGGAGGTATGGCCCCCAAGAAGGAC	GGTTCCACCACCACCGCCCGAGGTGAACTTGG
Cre17.g711850	GGAGGCGGAGGTATGGCCCGCACCAAGC	GGTTCCACCACCACCTTAGGCGCGCTCGCCA
Cre17.g711950	GGAGGCGGAGGTATGCAGCGGCCTGCGGAAAG	GGTTCCACCACCACCCACCGTCGCGCCCTCTTTCAAT
Cre17.g712050	GGAGGCGGAGGTATGGCCGACGTGCTCG	GGTTCCACCACCACCCTCCAGGATTGTCCGCGT
Cre17.g712150	GGAGGCGGAGGTATGGCGGCCAACAAAATACT	GGTTCCACCACCACCCTGCACGGGGTGTACGC
Cre17.g712250	GGAGGCGGAGGTATGGGCGTGAAGTGTGC	GGTTCCACCACCACCCATTATGCTCCGCAATGACAG
Cre17.g712350	GGAGGCGGAGGTATGGTGCTGTGGCTGCTG	GGTTCCACCACCACCCCCTGGCCGCAGCACC
Cre17.g712450	GGAGGCGGAGGTATGGACTGGGGCCGCA	GGTTCCACCACCACCCCCCTGCCCCTCCTCT
Cre17.g712550	GGAGGCGGAGGTATGACGTCCCGCATGTCC	GGTTCCACCACCACCCAGCATCATAGGGGCCGC
Cre17.g712700	GGAGGCGGAGGTATGCGGGCTGAGCACG	GGTTCCACCACCACCGCCGTTCTTGGCGTGT
Cre17.g712771	GGAGGCGGAGGTATGGCCTTCTGGGCTGA	GGTTCCACCACCACCGGTCCCGTGCTCGTGC
Cre17.g712850	GGAGGCGGAGGTATGACCTCGTTTGCTGGGAG	GGTTCCACCACCACCCTCCTGGTGGTCGTGCTC
Cre17.g713025	GGAGGCGGAGGTATGCTCCAGTATCGCACAGC	GGTTCCACCACCACCCAGCGAGTCATATACGCCTGC