

Aus der Ludwig-Maximilians-Universität München, Lehrstuhl für Epidemiologie
am universitären Zentrum für Gesundheitswissenschaften am Klinikum Augsburg
(UNIKA-T)¹

Lehrstuhlinhaber: Prof. Dr. Jakob Linseisen



***Assoziation zwischen Lebensmittelverzehr und DNA-
Methylierung in Blutzellen (PBMC)***

Dissertation

zum Erwerb des Doktorgrades der Humanbiologie
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

vorgelegt von

Fabian Hellbach

aus

Frankfurt am Main

Jahr

2023

1) Seit 01.01.2021 Lehrstuhl für Epidemiologie der Universität Augsburg, am Universitätsklinikum Augsburg

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

Erster Gutachter: Prof. Dr. Jakob Linseisen
Zweiter Gutachter: Prof. Dr. Hans Hauner
Dritter Gutachter: Prof. Dr. med Berthold Koletzko
ggf. weitere Gutachter:

Mitbetreuung durch den
promovierten Mitarbeiter:

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 5.12.2023

Affidavit



Eidesstattliche Versicherung

Hellbach, Fabian

Name, Vorname

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

Assoziation zwischen Lebensmittelverzehr und DNA-Methylierung in Blutzellen (PBMC)

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Augsburg, 12.12.2023
Ort, Datum
Doktorand

Fabian Hellbach
Unterschrift Doktorandin bzw.

Inhaltsverzeichnis

Affidavit	3
Inhaltsverzeichnis	4
Abkürzungsverzeichnis	5
Publikationsliste.....	6
1. Beitrag zu den Veröffentlichungen	7
1.1 Beitrag zur Publikation I	7
1.2 Beitrag zur Publikation II	7
2. Einleitung.....	8
2.1 DNA-Methylierung.....	8
2.1.1 Ursprung und Lokalisation von DNA-Methylierung.....	8
2.1.2 Wirkungsweise von DNA-Methylierung	8
2.1.3 Dynamik der DNA-Methylierung	9
2.1.4 Analyse und Interpretation von DNA-Methylierungsdaten.....	10
2.2 Einflussfaktoren	11
2.2.1 Ernährung und DNA-Methylierung.....	11
2.2.2 Metabolischer Status und DNA-Methylierung.....	13
2.3 Aufgaben und Ziele der Dissertation	14
2.4 Beschreibung der Analyse und Beitrag zum beschriebenen Problem	14
3. Zusammenfassung	18
4. Abstract (English)	20
5. Publikation I.....	22
6. Publikation II.....	23
7. Literaturverzeichnis	24
Danksagung	28
Lebenslauf	Fehler! Textmarke nicht definiert.

Abkürzungsverzeichnis

AHEI	Alternate Healthy Eating Index
DNMT	DNA-Methyltransferase
EWAS	Epigenome-wide association study
FFQ	Food Frequency Questionnaire
MDS	Alternate Mediterranean Diet Score
PBMC	Menschliche periphere mononukleäre Blutzellen
5hmC	5-hydroxymethylcytosin
5mC	5-methylcytosin

Publikationsliste

1. Hellbach, Fabian; Baumeister, Sebastian-Edgar; Wilson, Rory; Wawro, Nina; Dahal, Chetana; Freuer, Dennis; Hauner, Hans; Peters, Annette; Winkelmann, Juliane; Schwettmann, Lars; Rathmann, Wolfgang; Kronenberg, Florian; Koenig, Wolfgang; Meisinger, Christa; Waldenberger, Melanie; Linseisen, Jakob. (2022): Association between Usual Dietary Intake of Food Groups and DNA Methylation and Effect Modification by Metabotype in the KORA FF4 Cohort. *Life* 12(7):1064 (2022). <https://doi.org/10.3390/life12071064>.
2. Hellbach, Fabian; Sinke, Lucy; Costeira, Ricardo; Baumeister, Sebastian-Edgar; Beekman, Marian; Louca Panayiotis; Leeming R., Emily; Mompeo, Olatz; Berry, Sarah; Wilson, Rory; Wawro, Nina; Freuer, Dennis; Hauner, Hans; Peters, Annette; Winkelmann, Juliane; Koenig, Wolfgang; Meisinger, Christa; Waldenberger, Melanie; Heijmans T., Bastiaan; Slagboom, P. Eline; Bell T., Jordana; Linseisen, Jakob. Pooled analysis of epigenome-wide association studies of food consumption in KORA, TwinsUK and LLS. *Eur J Nutr* 62, 1357-1375 (2023). <https://doi.org/10.1007/s00394-022-03074-9>.

1. Beitrag zu den Veröffentlichungen

1.1 Beitrag zur Publikation I

Fabian Hellbach hat wesentlich zur Konzeptualisierung des Manuskripts „*Association between Usual Dietary Intake of Food Groups and DNA methylation and Effect Modification by Metabotype in the KORA FF4 Cohort*“ beigetragen. Er hat die Methodik entschieden und war für die Ausführung der Analyse und Visualisierung zuständig. Weiterhin war er maßgeblich für die Interpretation der Ergebnisse, den Entwurf des Manuskripts, das Zirkulieren des Manuskripts, die Einarbeitung der Kommentare der Ko-Autoren, das Einreichen bei einer wissenschaftlichen Zeitschrift, sowie die Bearbeitung der Reviewer-Kommentare verantwortlich. Er war an jedem Schritt des Papiers mit Ausnahme der Datensammlung maßgeblich beteiligt.

1.2 Beitrag zur Publikation II

Fabian Hellbach und sein Doktorvater Jakob Linseisen entwarfen das Studiendesign im Kontext des DIMENSION Konsortiums für das Papier „*Pooled analysis of epigenome-wide association studies of food consumption in KORA, TwinsUK and LLS*“. Er war für die Wahl der Meta-Analyse, die Datenharmonisierung der drei Kohorten im Vorfeld, sowohl als auch die Ausführung zuständig. Weiterhin war er maßgeblich für die Interpretation der Ergebnisse, den Entwurf des Manuskripts, das Zirkulieren des Manuskripts, die Einarbeitung der Kommentare der Ko-Autoren, das Einreichen bei einer wissenschaftlichen Zeitschrift, sowie die Bearbeitung der Reviewer-Kommentare verantwortlich.

2. Einleitung

2.1 DNA-Methylierung

2.1.1 Ursprung und Lokalisation von DNA-Methylierung

Der Begriff Epigenetik setzt sich aus dem altgriechischen *epi*: „danach, dazu, außerdem“, und dem ebenfalls altgriechischen Wort „Genetik“ zusammen [1]. Dieses Forschungsgebiet beschäftigt sich mit den Mechanismen, die die Expression von Genen und daraus folgend die Entwicklung einer Zelle beeinflussen. Im Gegensatz zu der Genetik, ist in der Epigenetik nicht die DNA-Sequenz das primäre Forschungsziel, sondern verschiedene Modifikationen des DNA-Apparates. Die am besten untersuchten epigenetischen Modifikationen sind DNA-Methylierung und Histon-Acetylierung, -Methylierung und -Phosphorylierung [2]. Epigenetische Modifikationen sind verantwortlich für die Differenzierung von Stammzellen zu unterschiedlichen Zelltypen unseres Körpers, da allen Zellen die gleiche DNA-Sequenz zugrunde liegt. DNA-Methylierung ist ebenfalls verantwortlich für das Ausschalten des X-Chromosoms (*Imprinting*) beim weiblichen Geschlecht [3]. DNA-Methylierung beschreibt das Anhängen einer Methylgruppe an das 5. Kohlenstoff-Atom einer Cytosin-Restgruppe (als Teil des DNA-Strangs), um 5-Methylcytosin (5mC) zu bilden. Dieser Prozess wird katalysiert durch DNA-Methyltransferasen (DNMTs), welche Methylgruppen von S-Adenosyl-Methionin auf Cytosin transferieren [3]. Es ist wichtig anzumerken, dass beim Abbau von 5mC das Zwischenprodukt 5-Hydroxymethylcytosin (5hmC) entsteht, welchem ebenfalls epigenetische Charakteristika zugesprochen werden [4]. In allen erwähnten Arbeiten wird von 5mC ausgegangen. Obwohl die meistbenutzte Methode zur Messung von DNA-Methylierung (Bisulfit-sequenzierung) nicht zwischen 5mC und 5hmC unterscheiden kann, kann, aufgrund des geringen Vorkommens von 5hmC, von 5mC ausgegangen werden. Am häufigsten tritt DNA-Methylierung an Cytosinen auf, welchen Guanin-Nukleotide vorangehen, weshalb sie auch *CpG sites* genannt werden. *CpG sites* sind im menschlichen Genom nicht sehr weitverbreitet, sie treten vor allem in Clustern auf und bilden häufig so genannte *CpG islands*. *CpG sites* liegen meistens methyliert vor, außer innerhalb von *CpG islands*. Es existieren einige unterschiedliche Definitionen von *CpG islands* [5], weshalb hier nicht näher auf die genaue Charakterisierung eingegangen wird. Es wird geschätzt, dass ca. 70% aller Gen-Promotoren innerhalb von *CpG islands* liegen [3].

2.1.2 Wirkungsweise von DNA-Methylierung

Wie bereits erwähnt, kann DNA-Methylierung maßgeblichen Einfluss auf die Expression von Genen haben. Dies wird durch folgende Mechanismen erklärt: Splicing-Regulation [6], Nukleosom-Lokalisation [7] und Rekrutierung von Transkriptionsfaktoren [8]. DNA-Methylierung innerhalb einer Promotorsequenz wird häufig mit der Repression der Genexpression in Verbindung gebracht [4]. Allerdings gibt es einige weitere Faktoren, die bei der Beurteilung der Auswirkung der DNA-Methylierung in Betracht gezogen werden sollten. In einer Studie konnte nur bei 16% der *CpG*

sites nahe eines Transkriptionsstartpunkts eine negative Korrelation mit dem Methylierungsstatus festgestellt werden [9], wohingegen DNA-Methylierung innerhalb von *CpG* Islands in einer konsistenteren Unterdrückung der Genexpression resultiert [10]. Etwa 2 kb entfernt von *CpG islands* liegen *CpG shores*, wie bei *CpG islands* korreliert Methylierung hier ebenfalls mit Gensuppression [3]. Wie sich Methylierung innerhalb von Gensequenzen auswirkt, ist kontext-abhängig, da sowohl Erhöhung als auch Reduktion der Genexpression beobachtet wurden [3, 11]. Methylierung des ersten Exons wurde ebenfalls mit Suppression der Genexpression in Verbindung gebracht [12].

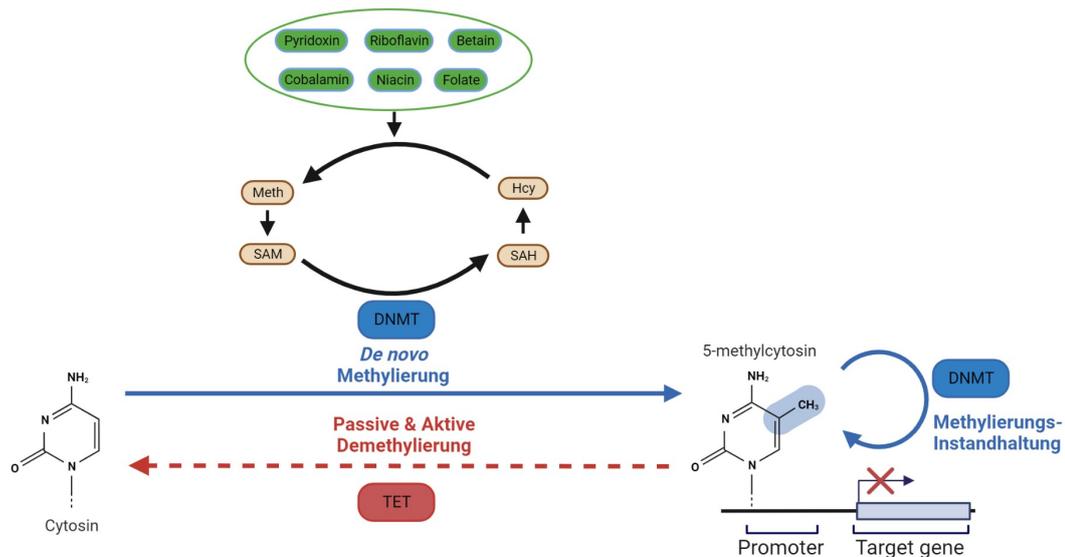


Abbildung 1: Dynamik der DNA-Methylierung. DNA-Methyltransferasen (DNMT) steuern die DNA-Methylierung über *de novo* und als Instandhaltungs-Methylierung. Die benötigten Methylgruppen stammen von S-Adenosylmethionin (SAM). Nach der Abgabe der Methylgruppe wird S-Adenosylmethionin zu S-Adenosylhomocystein. Das daraus abgespaltene Homocystein wird über Involvement des C1-Stoffwechsels (mit einigen Co-Faktoren, welche exogen durch unsere Ernährung zugeführt werden) zu Methionin. Dies wird schlussendlich wieder zu dem aktiven Methylgruppenüberträger S-Adenosylmethionin. Demethylierung kann entweder passiv vonstattengehen oder aktiv über die Enzymfamilie *Ten-eleven-translocation* (TET). Methylierung, die in der DNA enthaltenen Cytosine in Promotoren, kann die Genexpression hemmen, indem es die Zugänglichkeit von Transkriptionsfaktoren zur DNA hemmt.

2.1.3 Dynamik der DNA-Methylierung

Um eine Dynamik von DNA-Methylierung zu ermöglichen (siehe **Abbildung 1**), muss es aktive Methylierungs- und Demethylierungsmechanismen geben. DNMTs sind sowohl an der *de novo*- und Instandhaltungs-Methylierung, also intragenerationell, als auch bei der Fortpflanzung, also der Weitergabe von epigenetischen Merkmalen von einer Generation an die Nächste, beteiligt [3]. Passive Demethylierung findet bei teilenden Zellen statt, beispielsweise durch Inhibition von DNMTs oder defekten DNMTs [3]. Aktive enzymatische Demethylierung erfolgt durch die Enzymfamilie TET, welche 5mC zu 5hmC oxidiert und weiterhin 5-formylcytosin und 5-carboxycytosin bildet [4]. Die Dynamik der DNA-Methylierung zeigt sich in vielen Umwelteinflüssen, die erwiesenermaßen Einfluss darauf nehmen können: Rauchen [13], Umweltschadstoffe [14] oder verschiedenste Substanzen unserer Ernährung [15, 16]. Die genannten Einflussfaktoren wirken entweder indirekt über Mechanismen, die die Methylierung beeinflussen können, wie Inflammation,

oder direkt über die Beeinflussung der Enzyme, die an der Instandhaltung oder der *de novo*-Methylierung beteiligt sind.

2.1.4 Analyse und Interpretation von DNA-Methylierungsdaten

Die vorliegende Arbeit befasst sich mit der Analyse von Einflussfaktoren auf epigenomweite Methylierung. Diese Methode wird *epigenome-wide association study* (EWAS) genannt, wobei über das ganze Genom verteilt DNA Methylierung auf Nukleotid-Ebene einzeln gemessen wird. Unter den am weitesten verbreiteten Chips sind Illumina Infinium HumanMethylation450k BeadChip (Illumina 450k) und Illumina EPIC Human methylation microarray; diese messen die DNA-Methylierung an ~450.000, bzw. ~850.000 *CpG sites*. Diese enorme Menge an Daten resultiert in technischen und statistischen Herausforderungen. Die statistischen Herausforderungen sollen nun in Kürze erläutert werden.

Die Messwerte der Chips werden meist als *beta-values* in EWAS implementiert. Diese Werte können als durchschnittliche Methylierung an diesem Locus interpretiert werden und liegen deshalb zwischen 0 % und 100 % [17]. Die enorme Anzahl an statistischen Tests macht es unpraktikabel, die Grundannahmen einer linearen Regression für jeden Test zu überprüfen. Nichtsdestotrotz sind die Ergebnisse eines statistischen Tests nur valide, wenn die Annahmen dieses Tests erfüllt sind. Mansel et al. [18] haben in ihrer Analyse überprüft, ob Verletzungen der Annahmen zu Bias führen und kamen zu dem Schluss, dass die vorhandene Heteroskedastizität, nicht normalverteilte Residuen und Linearitätsverletzungen zu keiner relevanten Verzerrung der Ergebnisse führen. Weiterhin ist anzumerken, dass einige Autoren sich dafür aussprechen, *beta-values* zu *M-values* zu konvertieren [19]. Die Begründung liegt in der größeren statistischen Robustheit der *M-values*, allerdings ist, wie zuvor angesprochen, von keiner statistischen Verzerrung bei *beta-values* auszugehen, sodass die biologische Interpretationsfähigkeit der *beta-values* für deren Verwendung spricht. Aufgrund der Menge an statistischen Tests und die daraus resultierende hohe Anzahl falsch-positiver Ergebnisse, ist die Korrektur für die *family-wise error rate* Pflicht, um Ergebnisse zu erhalten, die die Chance haben sollen, reproduziert werden zu können. Es existieren verschiedene Ansätze dafür. Die häufig genutzte Bonferroni-Korrektur ist konservativ und geht von unabhängigen Tests aus [20]. Davon ist in EWAS aber nicht auszugehen [21], deshalb wurde in den vorliegenden Arbeiten, die weniger konservative und folglich mit einer höheren Power ausgestatteten, Benjamini-Hochberg *false-discovery rate* (FDR)-Prozedur angewandt [18, 22]. Eine letzte Besonderheit der Analyse von EWAS stellt die sogenannte genomische Inflation dar, welche charakterisiert ist durch die Abweichung von einer Gleichverteilung der resultierenden p-Werte. Dies wurde von Iterson et al. beschrieben [23], die auch gleichzeitig ein statistisches Modul (Software R) für deren Kontrolle entwickelten und bereitstellten. Die Nutzung dieses Verfahrens ist in Publikation I dargestellt.

Zusätzlich kommen biologische Herausforderungen in der Interpretation der Ergebnisse hinzu. In einigen Studien, ebenso wie in den hier vorgelegten Arbeiten, wird Vollblut zur Messung der DNA-Methylierung eingesetzt. Vollblut ist relativ einfach zu gewinnen, allerdings sind die Ergebnisse

nicht immer generalisierbar, da gezeigt wurde, dass DNA-Methylierungsmuster gewebsspezifisch sein können [24]. Im Vollblut sind außerdem verschiedene Anteile mononukleärer Zelltypen (Leukozyten, Granulozyten) enthalten, die dazu führen könnten, dass Unterschiede in der DNA-Methylierung in einer unterschiedlichen Zellverteilung (bspw. Infektion) begründet sein können. Aus diesem Grund wurden in den folgenden Arbeiten die gemessenen Zellzahlen mit in das Modell der linearen Regression aufgenommen. Für eine weiterführende Interpretation der Ergebnisse einer EWAS, ist es notwendig, Schlussfolgerungen zur Richtung, der dadurch veränderten Genexpression, ziehen zu können. Am besten ist dies bei einer gleichzeitigen Messung des Transkriptoms möglich. Falls dies nicht möglich ist, helfen Daten, die einen Interpretationsspielraum ermöglichen, wie die Spezifikation der Position des Locus. Wichtig hierbei sind die Position des *CpG site* hinsichtlich *CpG islands* und die Lokalisation relativ zum Gen, z.B. in der Promotorsequenz, nahe Promotor oder innerhalb der Gensequenz [4]. Aus diesen Informationen ist nur die Richtung der Genexpression potentiell ableitbar, das Ausmaß allerdings nicht.

2.2 Einflussfaktoren

2.2.1 Ernährung und DNA-Methylierung

Aufgrund der anhaltenden Zellteilung ist für die Aufrechterhaltung existierender Methylierungsmuster eine exogene Zufuhr an Methylgruppen notwendig [25]. Cholin/Betain, Methionin oder Methyl-Folat gelten als Methylgruppenspender und tragen direkt zum Methylgruppenhaushalt bei (siehe Abbildung 1). Komponenten des C1-Metabolismus waren als erstes im Fokus der Wissenschaft, um eine Verbindung zwischen Ernährung und DNA-Methylierung herzustellen. Der C1-Metabolismus soll nur verkürzt folgend dargestellt werden. Unter Beteiligung der Vitamine Niacin, Pyridoxin und Riboflavin wird 5-Methyltetrahydrofolat gebildet, welches als Methylgruppenspender für Homocystein gilt. Folglich entsteht, unter Beteiligung von Cobalamin, Methionin. An dieser Stelle können auch externe Methylgruppenspender, wie Cholin oder Betain, eintreten (siehe Abbildung 1). Das endogen gebildete Methionin wird zu dem Schlüsselmolekül der DNMT's, S-Adenosylmethionin [16].

In Zell-Studien konnte der Zusammenhang von Folat und DNA-Methylierung klar nachgewiesen werden. Globale Hypomethylierung, also eine verringerte Methylierung des vollständigen Genoms, und gen-spezifische Hypermethylierung konnten gezielt durch Folat-Supplementierung und Folat-Erschöpfung stimuliert werden [26]. Ein Zusammenhang von Cholin und DNA-Methylierung könnte durch die maßgebliche Beeinflussung der S-Adenosylmethionin-Konzentration erklärt werden. Eine Cholin-arme Diät in Ratten führte zu einer verringerten Konzentration an S-Adenosylmethionin und einer erhöhten Konzentration an S-Adosylhomocystein in der Leber [27]. Die Kombination aus beiden resultiert in einer Inhibition von Methyltransferase-Aktivität [28]. Der Zusammenhang von diätetischem Methionin und DNA-Methylierung wurde in einer Gänse-Studie untersucht [29]. Die Manipulation der Methionin-Zufuhr resultierte in veränderten Expressionsleveln von Genen, kritisch für den Methionin-Metabolismus. Außerdem konnte eine verän-

derte DNA-Methylierung und dazugehörige Genexpression von einem Loci nachgewiesen werden. EWAS-Ergebnisse zur Methionin-Aufnahme gibt es derzeit nicht, ebenfalls nicht für Cholin oder Betain, allerdings für Folat. Mandaviya et al. [30] analysierte ca. 6.000 Personen, gepooled aus mehreren Kohorten. In einem kategoriellen statistischen Modell der Folat-Zufuhr haben sie einige signifikante CpGs entdecken können. Ein *Letter to the Editor* [31] wurde veröffentlicht, in der die Ergebnisse von Mandaviya et al. kritisiert wurden. Hinsichtlich der Literatur und Ergebnisse ihrer eigenen durchgeführten Analyse kamen die Autoren zur Schlussfolgerung, dass im besten Fall nur eine schwache Evidenz für die Assoziation zwischen Folsäure und DNA-Methylierung in EWAS beim Menschen besteht.

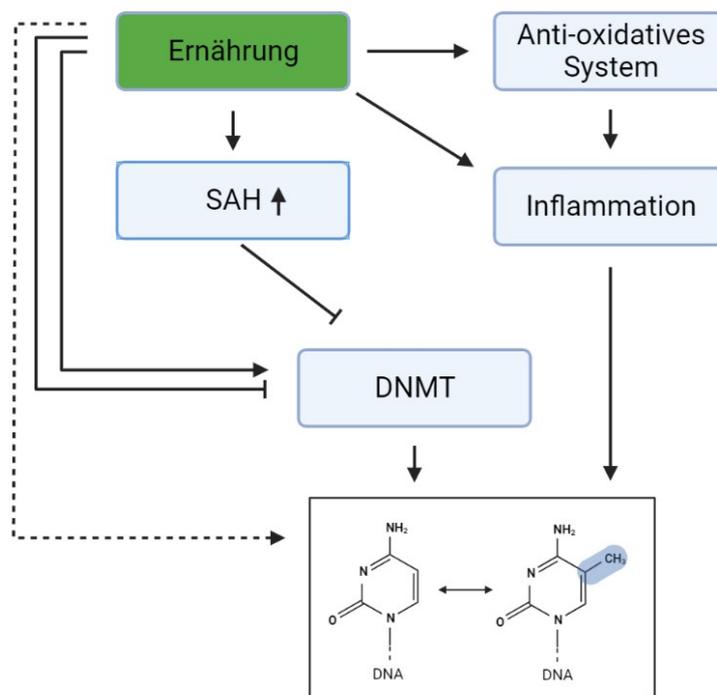


Abbildung 2: Mechanismen, wie Ernährung Einfluss auf DNA-Methylierung haben könnte. Unsere Ernährung kann z.B. via Konsum von Genistein zu einem Anstieg der S-Adenosylhomocystein (SAH) Konzentration führen, was hemmend auf DNMT's wirkt [15]. Nährstoffe des C1-Metabolismus können die Verfügbarkeit von SAM beeinflussen und dadurch die *de novo*-Methylierung beeinflussen [16]. Die Moleküle Curcumin und Epigallocatechingallat können direkt hemmend auf DNMTs wirken, indem die Thiol-Gruppe von DNMTs für Nucleotide blockiert wird [15, 32]. Einige antioxidative Inhaltsstoffe unserer Ernährung könnten über Beeinflussung der systemischen Inflammation ihren Einfluss auf die DNA-Methylierung ausüben. Weitere bisher nicht geklärte Mechanismen werden durch den gestrichelten Pfeil dargestellt.

Einige Pflanzeninhaltsstoffe wurden hinsichtlich des Mechanismus untersucht, wie sie die DNA-Methylierung beeinflussen können. Apigenin, Quercetin und Genistein verursachen bei ihrem Metabolismus durch unseren Körper einen Anstieg an S-Adenosylhomocystein, welches ein potenter Inhibitor von DNMT's ist [15, 32] (siehe **Abbildung 2**). Der gleiche Mechanismus wird vermutet bei den in Kaffee enthaltenen Stoffen Kaffeesäure und Chlorogensäure [32]. Apigenin, Curcumin und Sulforaphan könnten ihren Einfluss auf die DNA-Methylierung durch ihre bekannten anti-inflammatorischen und antioxidativen Effekte ausüben [15, 16]. Spezifischere Mechanismen

sind bekannt bei Curcumin und Epigallocatechingallat. Curcumin blockt die katalytische Thiol-Seite von DNMT, wohingegen Epigallocatechingallat eine Wasserstoffbrückenbindung in der katalytischen Region von DNMT eingeht, womit der Zugang von Nukleotiden blockiert wird [15, 32]. Ähnliche Mechanismen sind vorstellbar für die Summe und Kombination an Nährstoffen, die in Lebensmittelgruppen oder Ernährungsmustern vertreten sind.

EWAS, die Lebensmittel, Lebensmittelgruppen oder Ernährungsmuster analysieren, sind rar. Der Grundgedanke solcher Analysen ist das Erfassen komplexer zusammenspiele verschiedener Lebensmittel und deren Inhaltsstoffen. Einige Analysen gibt es dennoch. Zwei EWAS untersuchten den Zusammenhang von DNA-Methylierung und Kaffee und Tee. Die erste aus dem Jahr 2017 [33] fand nur zwei signifikante CpG sites in Verbindung mit Tee-Konsum bei einer stratifizierten Analyse in Frauen, jedoch keine in einer geschlechtsgemischten Analyse. Die Analyse der Assoziation mit Kaffee ergab weder stratifiziert noch gemischt ein signifikantes Ergebnis. Die zweite EWAS aus dem Jahr 2021 [34] hatte Daten einer wesentlich größeren Population zur Verfügung und fand elf signifikante CpG sites im Zusammenhang mit Kaffee-Konsum, allerdings ebenfalls keine Ergebnisse in Verbindung mit Tee-Konsum.

Die meistuntersuchten Ernährungsmuster sind der *Alternate Healthy Eating Index (AHEI)* und der *Alternate Mediterranean Diet Score (MDS)*. Um den Einfluss von Ernährungsmuster und DNA-Methylierung zu untersuchen, wählten Arpon et al. [35] eine Subpopulation ihrer Interventionsstudie aus, welche die mediterrane Ernährungsweise untersuchte. Sie fanden ca. 300 signifikante CpG sites und nutzten sie für eine *Pathway*-Analyse, in der sie fanden, dass die signifikanten CpGs sich vor allem in Genen befinden, welche in Entzündungsreaktionen oder intrazellulären Signalwegen involviert sind. Eine weitere Studie untersuchte AHEI und MDS in insgesamt 6.600 Personen europäischen Ursprungs und konnte für beide Ernährungsmuster einige signifikante CpG sites identifizieren [36].

2.2.2 Metabolischer Status und DNA-Methylierung

Abweichungen verschiedener metabolischer Parameter, wie Blutglucose [37] oder Blutlipide [38] von der physiologischen Norm, können ebenfalls Auswirkung auf die DNA-Methylierung haben und treten meist in Verbindung mit Krankheiten auf. Ein möglicher Mechanismus, der diese Assoziationen erklären kann ist, dass DNA-Methylierung durch Inflammation beeinflusst werden kann [16] und es ist charakteristisch, dass kardiometabolische Erkrankungen mit systemischer Inflammation einhergehen. Die pathologische Grundlage für kardiovaskuläre Erkrankungen ist Atherosklerose und geht mit Inflammation und oxidativem Stress einher. Risikofaktoren können Diabetes oder hohes LDL-Cholesterol sein [39]. Li et al. [40] zeigten in ihrer EWAS, dass die DNA-Methylierung von knapp 5.000 CpG sites mit Atherosklerose assoziiert ist. Weiterhin konnten sie sechs CpG sites identifizieren, die in einem prädiktiven logistischen Modell genutzt werden konnten um Atherosklerose zu klassifizieren. Adipositas ist charakterisiert durch einen BMI > 30 und geht ebenfalls mit systemischer Inflammation einher. Einige EWAS konnten bereits einen Zusammenhang zwischen BMI und DNA-Methylierung belegen [41]. Chen et al. [42] und Wahl et al. [43] untersuchten den Zusammenhang in EWAS und identifizierten in beiden Studien über 100

CpG sites, die mit DNA-Methylierung assoziiert sind. Eine andere Studie untersuchte die Effekt-Modifikation der metabolischen Gesundheit auf die Assoziation von BMI und DNA-Methylierung [44]. Sie fanden, dass 22 *CpG sites* signifikant mit dem Interaktionsterm assoziiert sind. Unser genutztes Metabotyp-Modell aus Publikation I basiert auf den Variablen Plasma-Glucose, HDL, nicht-HDL-Cholesterin, BMI und Harnsäure. Wie zuvor beschrieben sind einige dieser Variablen bekannt für eine Assoziation mit DNA-Methylierung. Es ist daher denkbar, dass durch die Zusammenführung dieser Variablen im Rahmen des Metabotyp-Modells, eine Effekt-Modifikation auf die Assoziation von Lebensmittelgruppen und DNA-Methylierung auftreten kann.

2.3 Aufgaben und Ziele der Dissertation

Die Aufgabe dieser Dissertation ist es, die Forschung von Ernährung und Epigenetik, welche in der Vergangenheit fast ausschließlich einzelne Nährstoffe untersuchte, um die Analyse mit Lebensmittelgruppen zu erweitern. Diese Vorgehensweise hat zwei maßgebliche Vorteile: (1) Lebensmittel sind eine Summe an Nährstoffen und Interaktionen zwischen ihnen können erfasst werden, (2) Ergebnisse von Lebensmittelgruppen-Analysen sind praxisorientierter, da potentielle Lebensmittelpfehlungen abgeleitet werden könnten. Folglich wurden zwei Projekte zur Analyse der Ernährung und der DNA-Methylierung in PBMC durchgeführt.

- (i) Im Rahmen des ersten Projektes wurde die Assoziation zwischen dem üblichen Lebensmittelverzehr und der DNA-Methylierung in der KORA FF4 Kohorte auf dem EPIC 850K Chip analysiert; dabei wurde außerdem eine mögliche Effekt-Modifikation durch die Stoffwechselsituation der StudienteilnehmerInnen, charakterisiert über die Zuordnung zu einem dreistufigen Metabotyp [45], untersucht.
- (ii) Die zweite Arbeit umfasste die Analyse der Assoziation von Lebensmittelgruppen-Verzehr und DNA-Methylierung auf dem 450K Chip in drei europäischen Kohortenstudien mittels Meta-Analyse.

Betrachtet wurden insbesondere Lebensmittelgruppen, (i) die Nährstoffe enthalten, die für den C1-Stoffwechsel relevant sind (z.B. Kohlgemüse), (ii) für die ein Einfluss auf systemische Inflammation in der Literatur beschrieben wurde (z.B. zuckergesüßte Getränke), (iii) die in der untersuchten bayrischen Bevölkerung üblicherweise verzehrt werden (z.B. Fleisch).

2.4 Beschreibung der Analyse und Beitrag zum beschriebenen Problem

Bei der ersten Publikation handelt es sich um eine EWAS, welche die Assoziation von üblichem Lebensmittelverzehr und DNA-Methylierung und der Effektmodifikation durch Metabotyp in der KORA FF4 Kohorte untersucht. KORA FF4 ist die zweite Folgestudie aus dem Jahr 2013/2014 des Surveys KORA S4, welche 2000/2001 in der Region Augsburg durchgeführt wurde. Ernährungsdaten stehen etwa für 1.400 Probanden zur Verfügung. Sie wurden als gewöhnlicher Lebensmittelverzehr modelliert. Die detaillierte Vorgehensweise kann der Publikation von Wawro et al. entnommen werden [46].

Tabelle 1: Übersicht über alle analysierten Expositionen in beiden Publikationen. Die EPIC-Soft Klassifizierung [47] teilt Lebensmittelgruppen in Haupt- und Subgruppen ein. Hauptgruppen beinhalten alle Sub-Gruppen der jeweiligen Lebensmittelgruppe, wovon nicht alle hier abgebildet sind. Fettgedruckte Lebensmittelgruppen stellen Hauptgruppen dar. Alternate Healthy Eating Index (AHEI); Alternate Mediterranean Diet Score (MDS)

Kartoffeln	Hülsenfrüchte	Vollkornprodukte	Zuckergesüßte Getränke	Spirituosen
Gemüse (Gesamt)	Früchte (Gesamt)	Fleisch (Gesamt)	Margarine	Alkohol (g/Tag)
Blattgemüse	Nüsse und Samen	Frisches rotes Fleisch	Süßigkeiten (Gesamt)	AHEI
Fruchtgemüse	Milch	Verarbeitetes Fleisch	Kuchen	MDS
Wurzelgemüse	Joghurt	Fisch (Gesamt)	Kaffee	Folsäure
Kohlgemüse	Käse	Eier	Tee	
Pilze	Sahne	Pflanzenöle	Wein	
Zwiebeln & Knoblauch	Getreideprodukte	Butter	Bier	

Daten von mindestens zwei 24-h Lebensmittellisten wurden als Grundlage für eine Verzehrswahrscheinlichkeit für das Individuum genutzt und multipliziert mit einer Verzehrsmenge, wenn das Lebensmittel konsumiert wurde. Die Verzehrsmenge wurde aus der Bayrischen Verzehrsstudie II ermittelt. Diese Modellierung der Ernährungsdaten reduziert den sonst prominenten Messfehler in der Ernährungserhebung. Die DNA-Methylierung wurde auf dem EPIC-Array von Illumina, mit 850.000 dargestellten CpG Loci, untersucht. Schlussendlich konnten 1.261 Teilnehmer in die Analyse miteinbezogen werden, welche alle nötigen Ernährungs-, Methylierungs- und Kovariaten aufwiesen. Das genutzte Metabotyp-Konzept [45] wurde von unserer Arbeitsgruppe entwickelt und nutzt Clustering-Methoden (u.a. *kmeans*), um metabolisch möglichst homogene Gruppen, auf Grundlage der Variablen: Plasma-Glukose, HDL-Cholesterol, nicht-HDL-Cholesterol, Harnsäure und BMI, zu bilden. Das Resultat dieses Prozesses sind drei Metabotyp-Klassifizierungen: Metabotyp 1 repräsentiert die metabolisch optimalsten und Metabotyp 3 repräsentiert die metabolisch ungünstigsten Parameter, während Metabotyp 2 ein intermediäres Cluster darstellt. Kovariaten für das lineare Regressionsmodell wurden anhand des *disjunctive-cause criteria* [48] und bisheriger Literatur gewählt. Für Ergebnisse der Analyse durch Effekt-Modifikation wurde, als Sensitivitätsanalyse, eine Methode angewandt, die die Inflation von niedrigen p-Werten in EWAS vermindert. In der Basis-Analyse (ohne Effekt-Modifikation) fanden wir vereinzelt signifikante CpG sites, die mit dem Verzehr der Lebensmittelgruppen Blattgemüse, Wurzelgemüse, Kohlgemüse, Zwiebeln und Knoblauch, Wein und Bier assoziiert sind. CpG sites die-

ser Analyse konnten Genen zugeordnet werden, welche Proteine für den C1-Metabolismus exprimieren. In der Analyse für Effektmodifikation durch Metabotyp fanden wir viele signifikante Ergebnisse nach Korrektur durch die *false discovery Rate* ($p < 0,1$). Die Lebensmittelgruppen mit absteigender Reihenfolge der Anzahl an signifikanten *CpG sites* sind: Käse, Kohlgemüse, Margarine, Eier, Vollkorn und Fleisch. Nach Korrektur der *genomic inflation* durch das R Modul *bacon* sind einige signifikante Ergebnisse verblieben, vor allem in den Lebensmittelgruppen Käse und Margarine. Das Gen *MTHFD1L*, *Methylenetetrahydrofolate dehydrogenase-1 similar*, konnte in den signifikanten *CpGs* verschiedener Lebensmittelgruppen identifiziert werden und ist Teil der Regeneration von Methionin aus Homocystein. Andere interessante Ergebnisse in der Lebensmittelgruppe Käse konnten Genen zugeordnet werden, welche mit Augengesundheit assoziiert sind (*RP1L1*, *EML1*, *PITPNC1*, *NRL*). Weitere signifikante *CpGs* waren mit Genen annotiert, welche mit dem antioxidativen System assoziiert sind, wie *GPX2* und *PON3*. Ergebnisse dieser Analyse durch Effekt-Modifikation können möglicherweise erklärt werden, durch die Auswirkung metabolischer Entgleisungen auf die Assoziation von Lebensmittelgruppen und DNA-Methylierung. Limitationen dieser Arbeit waren die geringe Stichprobengröße relativ zu der Anzahl an Tests und nicht vorhandene Genexpressionsdaten für eine effektivere Interpretation der Ergebnisse. Stärken hingegen waren, dass eine Effekt-Modifikation als statistisches Modell gewählt wurde, welches die statistische Power erhält, im Gegensatz zu einer stratifizierten Analyse und dass möglichst valide Verzehr Daten genutzt wurden.

Die zweite Publikation untersuchte die direkte Assoziation von Lebensmittelgruppen und DNA-Methylierung. Diese Analyse hat, aufgrund der gepoolten Daten der drei Kohorten, eine größere Power als die Analyse von Publikation I. Für die Meta-Analyse wurden Daten der deutschen Kohorte KORA FF4, der britischen Kohorte TwinsUK und der niederländischen Kohorte LLS, mit einer finalen Stichprobengröße von $n = 2.315$, herangezogen. Im Vorhinein wandten die Analysten eine Harmonisierungs-Prozedur an, die sicherstellen sollte, dass große Unterschiede in der Definition von Lebensmittelgruppen in den Kohorten ausgeräumt werden. Alle drei Kohorten klassifizierten ihre Lebensmittel nach dem EPIC-Soft Klassifizierungssystem [47]. Die Daten von TwinsUK und LLS basierten ausschließlich auf *Food Frequency Questionnaires* (FFQ) und wurden für diese Meta-Analyse in g/Tag umgerechnet. Da TwinsUK und LLS nur die Methylierungsdaten des *Illumina HumanMethylation450 array* vorlagen, musste die Meta-Analyse auf ~400.000 Loci beschränkt werden. Die Kovariaten für die Analyse waren in allen drei Kohorten gleich, mit der Ausnahme von TwinsUK, die einen kohortenspezifischen zufälligen Effekt zu ihrem Modell hinzufügten, da sie eine Zwillingskohorte sind. Eine *fixed-effects* Meta-Analyse wurde gewählt, da der Schätzer Tau, welcher Grundlage für eine *random-effects* Meta-Analyse ist, unpräzise sein kann, bei einer Meta-Analyse mit einer geringen Anzahl an Studien [49]. Um der Daten-Realität gerecht zu werden, haben wir eine *random-effects* Sensitivitätsanalyse, für die signifikanten *CpG sites* der *fixed-effects* Meta-Analyse, hinzugefügt. Die Ergebnisse wurden nach einer *false discovery rate* Adjustierung evaluiert. Die *fixed-effects* Analyse ergab zwei signifikante Assoziationen für Zwiebeln und Knoblauch, 18 für Nüsse, eine für Milch, elf für Sahne, 13 für Butter, vier für Pflanzenöle, fünf für Wein, 16 für Bier und sechs für Spirituosen. Die Sensitivitätsanalyse reproduzierte einige signifikante Assoziationen, allerdings nicht alle. Eine Hypothese, die in Betracht

der Ergebnisse, von uns aufgestellt wurde, war, dass der hohe Fett-Anteil dieser Lebensmittelgruppen mitursächlich für diese Assoziationen sein kann. Allerdings spricht dagegen, dass wir keine Ergebnisse für den Verzehr von Käse, Eier oder Margarine gesehen haben. Weiterhin gab es keine signifikanten Assoziationen für Lebensmittel, die charakteristisch für ihren Phytochemikalien-Gehalt sind, wie Kohlgemüse, Kaffee, Tee oder Blattgemüse. Ein interessantes Ergebnis war, dass eine CpG site in der Lebensmittelgruppe Sahne mit dem Gen *RPTOR* assoziiert ist. *RPTOR* ist in Signalwegen involviert, welche Zellwachstum, als Antwort auf Energie-Verfügbarkeit, regulieren. Das Gen *MYC*, als pro-fibrotischer Regulator, war mit einem signifikanten CpG der Lebensmittelgruppe Milch assoziiert. Eine wichtige Limitation dieser Studie war, dass wir in einigen Fällen eine hohe Heterogenität zwischen den Studien feststellen konnten, was auf Unterschiede in der Ernährungserfassung oder Unterschiede in der Population zurückzuführen sein kann. Eine Stärke dieser Arbeit war die vorangegangene Datenharmonisierung und eine vergleichbare Herangehensweise der Generierung der Methylierungsdaten. Es werden weitere umfassende Studien zu Lebensmittelgruppen und DNA-Methylierung benötigt.

3. Zusammenfassung

Epigenetik beschreibt Modifikationen des DNA-Apparates, welche nicht die DNA-Sequenz verändern. Die meistuntersuchten Modifikationen sind DNA-Methylierung und Histon-Acetylierung, -Methylierung und -Phosphorylierung. DNA-Methylierung beschreibt den Prozess der Bildung von 5-Methylcytosin, welches in den meisten Fällen Guanin-Nukleotiden vorangeht, daher werden sie auch *CpG sites* genannt. *Epigenome-wide association studies* (EWAS) wurden bisher insbesondere zu am C1-Metabolismus beteiligten Nährstoffen (vor allem Folsäure) und Alkohol (Ethanol) durchgeführt. Untersuchungen zu Lebensmittelverzehr und DNA-Methylierung liegen bisher kaum vor. Daher war es Ziel dieser Arbeit eine umfassende Analyse der Assoziation von Lebensmittelgruppen und DNA-Methylierung durchzuführen und folglich die Emergenz auf die Epigenetik zu untersuchen, die durch die vielen Charakteristika und Substanzen einer Lebensmittelgruppe gebildet wird.

Die erste Publikation untersuchte den Zusammenhang zwischen üblichem Lebensmittelverzehr und DNA-Methylierung bei 1.261 TeilnehmerInnen der KORA FF4-Studie. KORA FF4 ist die zweite Folgestudie aus dem Jahr 2013/2014 des Surveys KORA S4, welche 2000/2001 in der Region Augsburg durchgeführt wurde. Zur Reduktion des üblichen Messfehlers in Ernährungserhebungs-Daten wurde ein zweistufiges Modell angewandt, welches unter Nutzung der Daten von Verzehrswahrscheinlichkeit und Menge des Verzehrs die übliche Verzehrsmenge schätzte. Die DNA-Methylierungsdaten wurden mittels des EPIC-Array von Illumina, mit 850.000 dargestellten *CpG* Loci, gemessen. Ein wichtiges Ziel lag in der Untersuchung der Effekt-Modifikation von Ernährung auf DNA-Methylierung durch das Metabotyp-Konzept. Die Metabotypen (Clustering – *k-means* Methode) wurden, unter Einbeziehung der Parameter Plasma-Glukose, HDL-Cholesterol, nicht-HDL-Cholesterol, Harnsäure und BMI, gebildet, um möglichst homogene Gruppen darzustellen. Kovariaten wurden anhand des *disjunctive-cause criterion* gewählt und es wurde die *false discovery rate* (FDR) Korrektur angewandt mit einem Alpha von 0,1. Eine Sensitivitätsanalyse überprüfte die Inflation von niedrigen p-Werten (*genomic inflation*) unter Anwendung des R Moduls *bacon*. In der Analyse ohne Berücksichtigung des Metabotyps konnten wenige signifikante Assoziationen zwischen Lebensmittelgruppen und DNA-Methylierung beobachtet werden. Die Ergebnisse der Analyse der Effekt-Modifikation hingegen enthüllte viele signifikante Interaktionen zwischen dem Effekt des Verzehrs der Lebensmittelgruppen Käse, Kohlgemüse, Margarine, Eier, Vollkorn und Fleisch sowie dem metabolischen Status auf die DNA-Methylierung. Der Verzehr mehrerer Lebensmittelgruppen (Käse, Kohlgemüse) war mit *CpGs* assoziiert, welche mit dem Gen *MTHFD1L* annotiert sind. *MTHFD1L* ist an der Regeneration von Methionin aus Homocystein innerhalb des C1-Stoffwechsels beteiligt. Weiterhin wurden signifikante Assoziationen mit *CpGs* identifiziert, welche mit Genen annotiert sind, die mit dem antioxidativen System oder Augengesundheit, in Verbindung stehen. Die Stärken dieser Arbeit liegen in der Wahl der Analyse durch Effekt-Modifikation im Vergleich zur Wahl einer stratifizierten Analyse nach Metabotyp, die Nutzung möglichst valider Verzehrdaten und die Nutzung der Daten von 850.000 *CpGs*.

Mit der zweiten Publikation wurde die Assoziation des Verzehrs der Lebensmittelgruppen und DNA-Methylierung in einer Meta-Analyse untersucht. Daten der deutschen Kohorte KORA FF4,

der britischen Kohorte TwinsUK und der niederländischen Kohorte LLS wurden dafür herangezogen, mit 2.315 geeigneten TeilnehmerInnen. Um die Heterogenität zwischen den Studien zu minimieren, harmonisierten die Analytisten (i) die Ernährungsdaten, indem sie sicherstellten, dass es zu einer einheitlichen Definition der Lebensmittelgruppen kam (EPIC-Soft Klassifizierung) (ii) die statistischen Modelle, indem sie die Kovariaten vereinheitlichten. Aufgrund unterschiedlicher Verfügbarkeit von DNA-Methylierungsdaten wurde die Analyse auf 400.000 *CpG sites* beschränkt. Es wurde eine *fixed-effects* Meta-Analyse als Primär-Analyse und eine *random-effects* Meta-Analyse als Sensitivitätsanalyse durchgeführt. Nach einer Korrektur für multiples testen mittels der *false discovery rate* Adjustierung wurden Ergebnisse für die Lebensmittelgruppen Zwiebeln und Knoblauch, Nüsse, Milch, Sahne, Butter, Pflanzenöle, Wein, Bier und Spirituosen beobachtet. Basierend auf den Ergebnissen wurde die Hypothese generiert, dass ein hoher Fett-Anteil von Lebensmitteln ein Mediator sein könnte für den Zusammenhang des Verzehrs der Lebensmittelgruppen und DNA-Methylierung. Daraufhin wurde analysiert, ob signifikante Ergebnisse des Verzehrs der Lebensmittelgruppe Nüsse ebenfalls in anderen fettreichen Lebensmitteln identifiziert werden können. Die Ergebnisse bestätigten die Hypothese nicht. In der Analyse der signifikanten *CpG sites* wurde das Gen *RPTOR* identifiziert, welches in Signalwegen involviert ist, welche Zellwachstum, als Antwort auf Nährstoff-Verfügbarkeit, regulieren. Weiterhin wurde das Gen *MYC* identifiziert, dessen Produkt als pro-fibrotischer Regulator agiert. Die Stärke dieser Arbeit liegt in der vorangegangenen Harmonisierung der Daten, wohingegen eine Limitation fehlende Genexpressionsdaten ist, ohne welche biologische Konsequenzen nicht klar benannt werden können.

Die vorgelegten Arbeiten haben den Bereich der Ernährungs-EWAS umfassend, um Lebensmittelanalysen in der erwachsenen Bevölkerung, erweitert. Dabei wurden Lebensmittelgruppen (i) typisch für eine westliche Ernährungsweise, (ii) mit Assoziationen zu Inflammation und (iii) Lebensmittelgruppen mit bekannten Phytochemikalien untersucht. Die Ergebnisse der Meta-Analyse deuteten vor allem auf die Assoziation von DNA-Methylierung und fettreicher Lebensmittel hin. Darüber hinaus zeigte die Analyse der Effekt-Modifikation eindeutige Interaktionen auf zwischen dem Verzehr einiger Lebensmittelgruppen und der metabolischen Situation von StudienteilnehmerInnen auf die DNA-Methylierung. Effektstärken der signifikanten *CpG sites* deuteten nicht in eine Richtung, sondern waren sowohl positiv als auch negativ mit der DNA-Methylierung assoziiert. Es wurden einige statistische und inhaltliche Herausforderungen (wie z.B. *genomic inflation*) in der Analyse von EWAS-Daten und Lebensmittelverzehrdaten bewältigt. Limitationen dieser Arbeiten wurden diskutiert und Vorschläge für wissenschaftliche Anschlussprojekte abgeleitet.

4. Abstract (English)

Epigenetics describes modifications of the DNA apparatus that do not change the DNA sequence. The most studied modifications are DNA methylation and histone acetylation, methylation and phosphorylation. DNA methylation describes the process of forming 5-methylcytosine, which in most cases precedes guanine nucleotides, hence they are also called *CpG sites*. Epigenome-wide association studies (EWAS) have been performed so far, especially on nutrients (mainly folic acid) and alcohol (ethanol) involved in C1 metabolism. Studies on food consumption and DNA methylation are hardly available so far. Therefore, the aim of this work was to perform a comprehensive analysis of the association of food groups and DNA methylation and consequently to investigate the emergence on epigenetics formed by the many characteristics and substances of a food group.

The first publication examined the association between usual food consumption and DNA methylation in 1,261 participants of the KORA FF4 study. KORA FF4 is the second follow-up study from 2013/2014 of the survey KORA S4, which was conducted in 2000/2001 in the Augsburg region. To reduce the usual measurement error in dietary survey data, a two-stage model was applied, which estimated the usual consumption amount using the data of probability of consumption and amount of consumption. DNA methylation data were measured using Illumina's EPIC array, with 850,000 *CpG* loci represented. An important objective was to investigate the effect modification of diet on DNA methylation by the metabotype concept. Metabotypes (clustering - *k-means* method) were formed, including plasma glucose, HDL cholesterol, non-HDL cholesterol, uric acid, and BMI parameters, to represent groups as homogeneous as possible. Covariates were selected using the *disjunctive-cause criterion*, and *false discovery rate* (FDR) correction was applied with an alpha of 0.1. A sensitivity analysis checked for inflation of low p-values (genomic inflation) using the R module *bacon*. In the analysis without considering the metabotype, few significant associations between food groups and DNA methylation were observed. In contrast, the results of the effect modification analysis revealed many significant interactions between the effect of consumption of the food groups cheese, collard greens, margarine, eggs, whole grains, and meat and metabolic status on DNA methylation. Consumption of several food groups (cheese, collard greens) was associated with *CpGs* annotated with the gene *MTHFD1L*. *MTHFD1L* is involved in the regeneration of methionine from homocysteine within C1 metabolism. Furthermore, significant associations were identified with *CpGs* annotated with genes related to the antioxidant system or eye health. The strengths of this work are the choice of analysis by effect modification compared with the choice of stratified analysis by metabotype, the use of the consumption data as valid as possible, and the use of data from 850,000 *CpGs*.

With the second publication, the association of food group consumption and DNA methylation was investigated in a meta-analysis. Data from the German cohort KORA FF4, the British cohort TwinsUK, and the Dutch cohort LLS were used with 2,315 eligible participants. To minimize heterogeneity between studies, analysts harmonized (i) the dietary data by ensuring that there was a consistent definition of food groups (EPIC-Soft classification) (ii) the statistical models by standardizing covariates. Due to varying availability of DNA methylation data, the analysis was limited to 400,000 *CpG* sites. A fixed-effects meta-analysis was performed as the primary analysis and a random-effects meta-analysis was performed as the sensitivity analysis. After correction for multiple testing using the false discovery rate adjustment, results were observed for the food groups onions and garlic, nuts, milk, cream, butter, vegetable oils, wine, beer, and spirits. Based on the results, it was hypothesized that high fat content of foods could be a mediator of the association of food group consumption and DNA methylation. Subsequently, it was analyzed whether

significant results of the consumption of the food group nuts could also be identified in other high-fat foods. The results did not confirm the hypothesis. In the analysis of significant *CpG sites*, the gene *RPTOR* was identified, which is involved in signaling pathways that regulate cell growth in response to nutrients. Furthermore, the gene *MYC* was identified, whose product acts as a pro-fibrotic regulator. The strength of this work lies in the previous harmonization of data, whereas a limitation is missing gene expression data, without which biological consequences cannot be clearly named.

The work presented has comprehensively extended the field of nutritional EWAS to include food analyses in the adult population. Food groups (i) typical of a Western diet, (ii) with associations to inflammation, and (iii) food groups with known phytochemicals were examined. The results of the meta-analysis mainly indicated the association of DNA methylation and high-fat foods. In addition, effect modification analysis revealed clear interactions between consumption of some food groups and metabolic situation of study participants on DNA methylation. Effect sizes of significant CpG sites did not point in one direction but were both positively and negatively associated with DNA methylation. Some statistical and content challenges (such as genomic inflation) in the analysis of EWAS data and food consumption data were overcome. Limitations of this work were discussed and suggestions for follow-up scientific projects were derived.

5. Publikation I

Publikation I: Association between Usual Dietary Intake of Food Groups and DNA Methylation and Effect Modification by Metabotype in the KORA FF4 Cohort.

1. Hellbach, Fabian; Baumeister, Sebastian-Edgar; Wilson, Rory; Wawro, Nina; Dahal, Chetana; Freuer, Dennis; Hauner, Hans; Peters, Annette; Winkelmann, Juliane; Schwettmann, Lars; Rathmann, Wolfgang; Kronenberg, Florian; Koenig, Wolfgang; Meisinger, Christa; Waldenberger, Melanie; Linseisen, Jakob. Association between Usual Dietary Intake of Food Groups and DNA Methylation and Effect Modification by Metabotype in the KORA FF4 Cohort. *Life* 12(7):1064 (2022). <https://doi.org/10.3390/life12071064>.

6. Publikation II

Publikation II: Pooled analysis of epigenome-wide association studies of food consumption in KORA, TwinsUK and LLS

Hellbach, Fabian; Sinke, Lucy; Costeira, Ricardo; Baumeister, Sebastian-Edgar; Beekman, Marian; Louca Panayiotis; Leeming R., Emily; Mompeo, Olatz; Berry, Sarah; Wilson, Rory; Wawro, Nina; Freuer, Dennis; Hauner, Hans; Peters, Annette; Winkelmann, Juliane; Koenig, Wolfgang; Meisinger, Christa; Waldenberger, Melanie; Heijmans T., Bastiaan; Slagboom, P. Eline; Bell T., Jordana; Linseisen, Jakob. Pooled analysis of epigenome-wide association studies of food consumption in KORA, TwinsUK and LLS. *Eur J Nutr* 62, 1357-1375 (2023). <https://doi.org/10.1007/s00394-022-03074-9>.

7. Literaturverzeichnis

- [1] Wiktionary-Bearbeiter. Epigenetik; 2021UTC [cited 2022 August 31] Available from: URL: <https://de.wiktionary.org/w/index.php?title=Epigenetik&oldid=8892583>.
- [2] Gjaltema RAF, Rots MG. Advances of epigenetic editing. *Curr Opin Chem Biol* 2020; 57: 75–81
[<https://doi.org/10.1016/j.cbpa.2020.04.020>][PMID: 32619853]
- [3] Moore LD, Le T, Fan G. DNA methylation and its basic function. *Neuropsychopharmacology* 2013; 38(1): 23–38
[<https://doi.org/10.1038/npp.2012.112>][PMID: 22781841]
- [4] Tirado-Magallanes R, Rebbani K, Lim R, Pradhan S, Benoukraf T. Whole genome DNA methylation: beyond genes silencing. *Oncotarget* 2017; 8(3): 5629–37
[<https://doi.org/10.18632/oncotarget.13562>][PMID: 27895318]
- [5] Illingworth RS, Bird AP. CpG islands--'a rough guide'. *FEBS Lett* 2009; 583(11): 1713–20
[<https://doi.org/10.1016/j.febslet.2009.04.012>][PMID: 19376112]
- [6] Shukla S, Kavak E, Gregory M, *et al.* CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* 2011; 479(7371): 74–9
[<https://doi.org/10.1038/nature10442>][PMID: 21964334]
- [7] Chodavarapu RK, Feng S, Bernatavichute YV, *et al.* Relationship between nucleosome positioning and DNA methylation. *Nature* 2010; 466(7304): 388–92
[<https://doi.org/10.1038/nature09147>][PMID: 20512117]
- [8] Zhu W-G, Srinivasan K, Dai Z, *et al.* Methylation of adjacent CpG sites affects Sp1/Sp3 binding and activity in the p21(Cip1) promoter. *Mol Cell Biol* 2003; 23(12): 4056–65
[<https://doi.org/10.1128/MCB.23.12.4056-4065.2003>][PMID: 12773551]
- [9] Medvedeva YA, Khamis AM, Kulakovskiy IV, *et al.* Effects of cytosine methylation on transcription factor binding sites. *BMC Genomics* 2014; 15: 119
[<https://doi.org/10.1186/1471-2164-15-119>][PMID: 24669864]
- [10] Mohn F, Weber M, Rebhan M, *et al.* Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* 2008; 30(6): 755–66
[<https://doi.org/10.1016/j.molcel.2008.05.007>][PMID: 18514006]
- [11] Jones PA. The DNA methylation paradox. *Trends in Genetics* 1999; 15(1): 34–7
[[https://doi.org/10.1016/s0168-9525\(98\)01636-9](https://doi.org/10.1016/s0168-9525(98)01636-9)]
- [12] Brenet F, Moh M, Funk P, *et al.* DNA methylation of the first exon is tightly linked to transcriptional silencing. *PLoS One* 2011; 6(1): e14524
[<https://doi.org/10.1371/journal.pone.0014524>][PMID: 21267076]
- [13] Wilson R, Wahl S, Pfeiffer L, *et al.* The dynamics of smoking-related disturbed methylation: a two time-point study of methylation change in smokers, non-smokers and former smokers. *BMC Genomics* 2017; 18(1): 805
[<https://doi.org/10.1186/s12864-017-4198-0>][PMID: 29047347]
- [14] Poursafa P, Kamali Z, Fraszczyk E, Boezen HM, Vaez A, Snieder H. DNA methylation: a potential mediator between air pollution and metabolic syndrome. *Clin Epigenetics* 2022;

- 14(1): 82
[<https://doi.org/10.1186/s13148-022-01301-y>][PMID: 35773726]
- [15] Shankar E, Kanwal R, Candamo M, Gupta S. Dietary phytochemicals as epigenetic modifiers in cancer: Promise and challenges. *Semin Cancer Biol* 2016; 40-41: 82–99
[<https://doi.org/10.1016/j.semcancer.2016.04.002>][PMID: 27117759]
- [16] Mahmoud AM, Ali MM. Methyl Donor Micronutrients that Modify DNA Methylation and Cancer Outcome. *Nutrients* 2019; 11(3)
[<https://doi.org/10.3390/nu11030608>][PMID: 30871166]
- [17] Wu MC, Kuan P-F. A Guide to Illumina BeadChip Data Analysis. *Methods Mol Biol* 2018; 1708: 303–30
[https://doi.org/10.1007/978-1-4939-7481-8_16][PMID: 29224151]
- [18] Mansell G, Gorrie-Stone TJ, Bao Y, *et al.* Guidance for DNA methylation studies: statistical insights from the Illumina EPIC array. *BMC Genomics* 2019; 20(1): 366
[<https://doi.org/10.1186/s12864-019-5761-7>][PMID: 31088362]
- [19] Du P, Zhang X, Huang C-C, *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 2010; 11: 587
[<https://doi.org/10.1186/1471-2105-11-587>][PMID: 21118553]
- [20] Perneger TV. What's wrong with Bonferroni adjustments. *BMJ* 1998; 316(7139): 1236–8
[<https://doi.org/10.1136/bmj.316.7139.1236>][PMID: 9553006]
- [21] Ong M-L, Holbrook JD. Novel region discovery method for Infinium 450K DNA methylation data reveals changes associated with aging in muscle and neuronal pathways. *Aging Cell* 2014; 13(1): 142–55
[<https://doi.org/10.1111/accel.12159>][PMID: 24112369]
- [22] Chen S-Y, Feng Z, Yi X. A general introduction to adjustment for multiple comparisons. *J Thorac Dis* 2017; 9(6): 1725–9
[<https://doi.org/10.21037/jtd.2017.05.34>][PMID: 28740688]
- [23] van Iterson M, van Zwet EW, Heijmans BT. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome Biol* 2017; 18(1): 19
[<https://doi.org/10.1186/s13059-016-1131-9>][PMID: 28129774]
- [24] Thompson RF, Atzmon G, Gheorghe C, *et al.* Tissue-specific dysregulation of DNA methylation in aging. *Aging Cell* 2010; 9(4): 506–18
[<https://doi.org/10.1111/j.1474-9726.2010.00577.x>][PMID: 20497131]
- [25] Wilson MJ, Shivapurkar N, Poirier LA. Hypomethylation of hepatic nuclear DNA in rats fed with a carcinogenic methyl-deficient diet. *Biochem J* 1984; 218(3): 987–90
[<https://doi.org/10.1042/bj2180987>][PMID: 6721844]
- [26] Wasson GR, McGlynn AP, McNulty H, *et al.* Global DNA and p53 region-specific hypomethylation in human colonic cells is induced by folate depletion and reversed by folate supplementation. *J Nutr* 2006; 136(11): 2748–53
[<https://doi.org/10.1093/jn/136.11.2748>][PMID: 17056795]

- [27] Shivapurkar N, Poirier LA. Tissue levels of S-adenosylmethionine and S-adenosylhomocysteine in rats fed methyl-deficient, amino acid-defined diets for one to five weeks. *Carcinogenesis* 1983; 4(8): 1051–7
[https://doi.org/10.1093/carcin/4.8.1051][PMID: 6872150]
- [28] Tehlivets O, Malanovic N, Visram M, Pavkov-Keller T, Keller W. S-adenosyl-L-homocysteine hydrolase and methylation disorders: yeast as a model system. *Biochim Biophys Acta* 2013; 1832(1): 204–15
[https://doi.org/10.1016/j.bbadis.2012.09.007][PMID: 23017368]
- [29] Yang Z, Yang HM, Gong DQ, *et al.* Transcriptome analysis of hepatic gene expression and DNA methylation in methionine- and betaine-supplemented geese (*Anser cygnoides domesticus*). *Poult Sci* 2018; 97(10): 3463–77
[https://doi.org/10.3382/ps/pey242][PMID: 29931118]
- [30] Mandaviya PR, Joehanes R, Brody J, *et al.* Association of dietary folate and vitamin B-12 intake with genome-wide DNA methylation in blood: a large-scale epigenome-wide association analysis in 5841 individuals. *Am J Clin Nutr* 2019; 110(2): 437–50
[https://doi.org/10.1093/ajcn/nqz031][PMID: 31165884]
- [31] Dugué P-A, Chamberlain JA, Bassett JK, *et al.* Overall lack of replication of associations between dietary intake of folate and vitamin B-12 and DNA methylation in peripheral blood. *Am J Clin Nutr* 2020; 111(1): 228–30
[https://doi.org/10.1093/ajcn/nqz253][PMID: 31907529]
- [32] Kadayifci FZ, Zheng S, Pan Y-X. Molecular Mechanisms Underlying the Link between Diet and DNA Methylation. *Int J Mol Sci* 2018; 19(12)
[https://doi.org/10.3390/ijms19124055][PMID: 30558203]
- [33] Ek WE, Tobi EW, Ahsan M, *et al.* Tea and coffee consumption in relation to DNA methylation in four European cohorts. *Hum Mol Genet* 2017; 26(16): 3221–31
[https://doi.org/10.1093/hmg/ddx194][PMID: 28535255]
- [34] Karabegović I, Portilla-Fernandez E, Li Y, *et al.* Epigenome-wide association meta-analysis of DNA methylation with coffee and tea consumption. *Nat Commun* 2021; 12(1): 2830
[https://doi.org/10.1038/s41467-021-22752-6][PMID: 33990564]
- [35] Arpón A, Milagro FI, Razquin C, *et al.* Impact of Consuming Extra-Virgin Olive Oil or Nuts within a Mediterranean Diet on DNA Methylation in Peripheral White Blood Cells within the PREDIMED-Navarra Randomized Controlled Trial: A Role for Dietary Lipids. *Nutrients* 2017; 10(1)
[https://doi.org/10.3390/nu10010015][PMID: 29295516]
- [36] Ma J, Rebholz CM, Braun KVE, *et al.* Whole Blood DNA Methylation Signatures of Diet Are Associated With Cardiovascular Disease Risk Factors and All-Cause Mortality. *Circ Genom Precis Med* 2020; 13(4): e002766
[https://doi.org/10.1161/CIRCGEN.119.002766][PMID: 32525743]
- [37] Walaszczyk E, Luijten M, Spijkerman AMW, *et al.* DNA methylation markers associated with type 2 diabetes, fasting glucose and HbA1c levels: a systematic review and replication in a case-control sample of the Lifelines study. *Diabetologia* 2018; 61(2): 354–68
[https://doi.org/10.1007/s00125-017-4497-7][PMID: 29164275]

- [38] Jones AC, Irvin MR, Claas SA, Arnett DK. Lipid Phenotypes and DNA Methylation: a Review of the Literature. *Curr Atheroscler Rep* 2021; 23(11): 71
[https://doi.org/10.1007/s11883-021-00965-w][PMID: 34468868]
- [39] Wang C, Wang F, Liu F, Chen S. Epigenetic Regulation in Atherosclerosis. *Discov Med* 2021; 31(162): 45–9
[PMID: 34965371]
- [40] Li X, Dong X, Lu W, Yang K, Li X. Integrated Analysis of Gene Expression and Methylation Data to Identify Potential Biomarkers Related to Atherosclerosis Onset. *Oxid Med Cell Longev* 2022; 2022: 5493051
[https://doi.org/10.1155/2022/5493051][PMID: 35915606]
- [41] Do WL, Gohar J, McCullough LE, Galaviz KI, Conneely KN, Narayan KMV. Examining the association between adiposity and DNA methylation: A systematic review and meta-analysis. *Obes Rev* 2021; 22(10): e13319
[https://doi.org/10.1111/obr.13319][PMID: 34278703]
- [42] Chen Y, Kassam I, Lau SH, *et al.* Impact of BMI and waist circumference on epigenome-wide DNA methylation and identification of epigenetic biomarkers in blood: an EWAS in multi-ethnic Asian individuals. *Clin Epigenetics* 2021; 13(1): 195
[https://doi.org/10.1186/s13148-021-01162-x][PMID: 34670603]
- [43] Wahl S, Drong A, Lehne B, *et al.* Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 2017; 541(7635): 81–6
[https://doi.org/10.1038/nature20784][PMID: 28002404]
- [44] Do WL, Nguyen S, Yao J, *et al.* Associations between DNA methylation and BMI vary by metabolic health status: a potential link to disparate cardiovascular outcomes. *Clin Epigenetics* 2021; 13(1): 230
[https://doi.org/10.1186/s13148-021-01194-3][PMID: 34937574]
- [45] Dahal C, Wawro N, Meisinger C, *et al.* Evaluation of the metabotype concept after intervention with oral glucose tolerance test and dietary fiber-enriched food: An enable study. *Nutr Metab Cardiovasc Dis* 2022; 32(10): 2399–409
[https://doi.org/10.1016/j.numecd.2022.06.007][PMID: 35850752]
- [46] Wawro N, Kleiser C, Himmerich S, *et al.* Estimating Usual Intake in the 2nd Bavarian Food Consumption Survey: Comparison of the Results Derived by the National Cancer Institute Method and a Basic Individual Means Approach. *Ann Nutr Metab* 2017; 71(3-4): 164–74
[https://doi.org/10.1159/000481148][PMID: 28930718]
- [47] Slimani N, Deharveng G, Charrondière RU, *et al.* Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. *Computer Methods and Programs in Biomedicine* 1999; 58(3): 251–66
[https://doi.org/10.1016/s0169-2607(98)00088-1]
- [48] VanderWeele TJ, Shpitser I. A new criterion for confounder selection. *Biometrics* 2011; 67(4): 1406–13
[https://doi.org/10.1111/j.1541-0420.2011.01619.x][PMID: 21627630]
- [49] Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods* 2010; 1(2): 97–111
[https://doi.org/10.1002/jrsm.12][PMID: 26061376]

Danksagung

An erster Stelle möchte ich meinem Betreuer Prof. Dr. Jakob Linseisen danken. Für seine Empathie in schwierigen Phasen des Doktor-Studiums und dass er mir überhaupt diese Möglichkeit geboten hat. Er hat immer versucht neue Ideen einzubringen und auf Innovation zu pochen und trotzdem Rücksicht auf mich als Mensch genommen. Die Kommunikation war immer gut, was ihn zu einem spitzenmäßigen Betreuer machte.

Weiterhin möchte ich dem DIMENSION Konsortium danken, dass sie dieses Projekt ins Leben gerufen haben und somit die finanzielle Unterstützung bereitstellten, damit dieser Weg überhaupt möglich war.

Ebenfalls geht ein wichtiger Dank an die Mitarbeiter des Lehrstuhls für Epidemiologie am Universitätsklinikum Augsburg. Ein ganz besonderer Dank geht an Dennis Freuer, der mich immer statistisch unterstützt hat und wir immer rege Diskussionen über R oder Statistik hatten. Ohne so einen Ansprechpartner wäre die Vollendung dieses für mich riesigen Projektes nicht möglich gewesen.

Ein großer Dank an meine beiden Brüder Domi und Lars, die mich mit ihrer Faszination für das, was ich angehe und mit ihrem Stolz, motiviert haben, das Ganze zu vollenden.

Ein weiterer und wichtiger finaler Dank geht an zwei Therapeuten, die mich in meinem Leben stabilisiert und unterstützt haben. Frau Dr. Röbe hat sich immer Zeit für mich genommen und mir vor allem darin geholfen die Arbeit besser zu strukturieren und sich nicht davor zu scheuen nach Hilfe zu fragen. Außerdem danke, wie sie mich als Mensch schätzt. Frau Gropper war ebenfalls eine wichtige Unterstützung, als es mir am schlechtesten ging. Sie war maßgeblich daran beteiligt mich wiederaufzubauen und meinen Weg in die Arbeit zurückzufinden und mir Motivation zu geben den Doktor fertig zu stellen.