Beyond the CpG: an integrative approach to decoding DNA methylation in immunometabolic health

Sinke, L.J.

CHAPTER SEVEN

Discussion

Aims of this thesis

A substantial body of epidemiological work positioned DNA methylation (DNAm) as a central component of ageing and age-related disease[1,2]. Large-scale epigenome-wide association studies (EWAS) robustly connected CpGs at metabolically relevant loci to a range of related factors, from dietary patterns to diabetes risk[2–4]. These findings cemented EWAS as a cornerstone of population epigenetics, and a compelling framework through which to dissect the molecular underpinnings of human health[5].

In parallel, circulating signalling molecules emerged as sensitive readouts of immunometabolic health across the lifespan[6]. These metabolic and inflammatory mediators, including interleukin-6 (IL-6), adiponectin, and leptin, were implicated in disease pathogenesis, predictive of diabetes onset and mortality, and could quantify intervention-driven health improvements[7–9]. Their responsiveness, prognostic value, and strong links to longevity rendered them prime candidates for capturing heterogeneity in healthy ageing[10–12]. Yet, the epigenetic architecture of such measures remained largely unexplored.

This thesis aimed to bridge the domains of epigenomics and immunometabolic health, uncovering how their intersection limits or leads to age-related disease. Specifically, it sought to map blood-based DNAm signatures of three related markers: IL-6, adiponectin, and leptin, and characterise multi-tissue epigenetic responses to a lifestyle intervention, in which older adults improved their immunometabolic health through a combination of increased physical activity and calorie restriction[13]. By assimilating multiple genomic layers, this work strived to trace causal pathways and reveal their regulatory points of control. Furthermore, by systematically embedding biological and clinical context into this integrative approach, it hoped to move beyond the nucleotide resolution of CpGs back up to higher-order insights with direct translational relevance.

To support these goals, this thesis endeavoured to develop and disseminate open-source computational tools. Equipped with clear and accessible documentation, this software set out to support epidemiological researchers throughout the analytical workflow, from study design to statistical inference. In sum, this work aspired to provide both empirical evidence and practical resources, taking meaningful strides towards realising the full potential of EWAS to understand, and ultimately improve, human health in ageing.

Summary of chapters

This thesis began by presenting EWAS meta-analyses of three blood-based markers of immunometabolic health. In a combined sample of over 4,000 individuals, **Chapter 2** identified 401 CpGs associated with IL-6 levels and mapped these sites to inflammatory and metabolic genes, such as *AIM2*, *MTOR*, and *IL6R*. IL-6 is secreted by circulating cells, meaning this methylation could have reflected direct epigenetic reprogramming of their transcriptional profiles and IL-6 production capabilities. In line with this, two-sample Mendelian randomisation (2SMR) implicated one locus (*NFATC2IP*) as plausibly influencing not only IL-6 production but also multiple immunometabolic traits, including body mass index (BMI) and type 2 diabetes (T2D).

Nevertheless, this CpG represented an exception to the trend supported by triangulation, in which DNAm was typically a consequence rather than a cause of IL-6. Follow-up mediation analysis uncovered two sites at *SOCS3* that not only responded to IL-6, but also partially transmitted its effects on C-reactive protein (CRP), BMI, and inflammatory bowel disease. This Chapter therefore broadly concluded that blood-based DNAm reflects immune cell exposure to IL-6 in the circulation, and that these responses can in turn propagate effects on inflammatory risk factors and disease.

Chapter 3 extended our approach to adiponectin and leptin, two adipokines secreted into the circulation by adipocytes. In this context, three causal scenarios were evaluated: (i) reverse causation, where immune cell DNAm responds to adipokine levels, (ii) leukocyte-mediated regulation of adipokine production, and (iii) shared upstream factors driving correlated epigenetic patterns in both blood and adipose tissues. Meta-analysis identified 73 CpGs associated with adiponectin and 211 sites linked to leptin, with triangulation revealing that, in contrast to IL-6, DNAm may play a role in regulating adiponectin production.

Two loci stood out: a CpG near *ADIPOQ*, the gene encoding adiponectin itself, and another connected to *SREBF1* expression, a transcription factor (TF) central to lipid metabolism that promotes pro-inflammatory polarisation of macrophages[14,15]. While the *ADIPOQ* CpG was discovered in blood, its links to *ADIPOQ* expression could not be examined here as immune cells do not produce adiponectin. Instead, we connected this CpG to *ADIPOQ* expression in adipocytes, and hypothesized that a shared upstream factor, such as diet, may be driving co-ordinated DNAm in blood and adipose tissues.

While such a causal structure was also plausible at the *SREBF1* site, an alternative model supported DNAm as epigenetically priming immune cells to amplify adipose tissue inflammation upon immune infiltration. This Chapter therefore showcased how large-scale blood-based studies can uncover biologically meaningful loci for traits non-hematopoietic in origin, and exemplified their validation in a targeted follow-up analysis.

The next section of this thesis shifted focus to software development, setting out to facilitate evaluation of DNAm in a novel experimental setting. **Chapter 4** introduced *Omixer*, an open-source Bioconductor package that proactively mitigated technical confounding in genomic studies. Designed to accommodate paired and complex study designs, *Omixer* outperformed standard randomization approaches, reducing associations between technical and biological factors, and thereby enhancing confidence in downstream results.

In **Chapter 5**, we presented *DNAmArray*, a flexible and scalable computational workflow for processing array-based DNAm data. Building on tools developed by our research group, including *methylAid*[16], *omicsPrint*[17], and *bacon*[18], this Chapter showcased streamlined and modular steps for quality control, normalisation, and analysis, providing high quality data and reproducible results in subsequent studies.

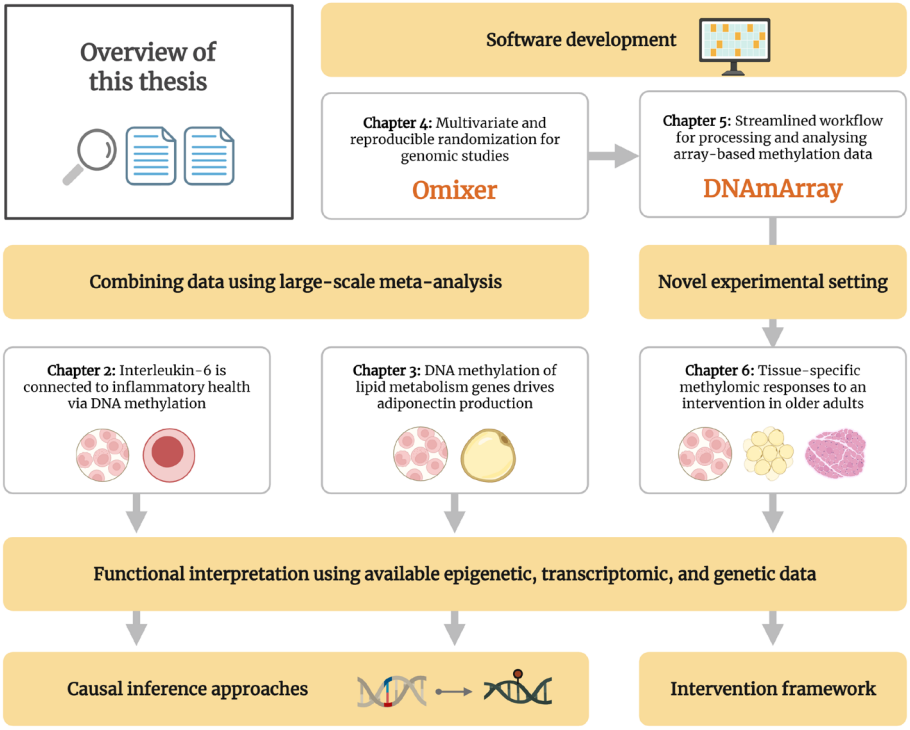


Figure 1 | Overview of this thesis. *Chapters 2 and 3* combined data using large-scale meta-analysis and evaluated DNAm signatures of IL-6, adiponectin, and leptin. *Chapters 4 and 5* developed software for use in *Chapter 6*, a novel experimental setting that interrogated tissue-specific methylomic responses to a lifestyle intervention in older adults. Tissues and approaches used are illustrated. *Figure created in BioRender.*

We made effective use of our integrative framework and software in **Chapter 6**, profiling DNAm before and after the Growing Old TOgether (GOTO) lifestyle intervention, where older adults improved their immunometabolic health through a combination of increased physical activity and calorie restriction. Three metabolically relevant tissues were investigated, revealing 162 differentially methylated CpGs in muscle, 230 sensitive sites in adipose, and 441 loci responsive to the intervention in blood. Yet, despite high tissue-specificity in these epigenetic signatures, all three could be connected to immunometabolic and physiological health changes, including reductions in BMI and leptin levels.

This last Chapter also applied epigenetic clock algorithms, uncovering consistent decreases in biological age (bAge) following the GOTO intervention. GrimAge, in particular, was associated with alterations in health, including circulating adipokine levels. Collectively, this Chapter concluded the work of this thesis, showcasing the value of our computational toolkit and integrative approach for dissecting the epigenetics of immunometabolic health (**Fig. 1**).

From association to interpretation

Common CpGs

In this thesis, we conducted EWAS meta-analyses of IL-6, adiponectin, and leptin, and evaluated DNAm responses to a lifestyle intervention in three tissues. Across these investigations, 44 CpGs were associated with more than one outcome, thereby representing loci whose biological relevance was supported by converging lines of evidence (**Fig. 2a**). The majority of these (n = 26) were shared between the IL-6 and leptin EWAS, with a further five overlapping between each of (i) IL-6 and adiponectin, (ii) the adipokines, and (iii) all three blood-based EWAS of *Chapters 2 and 3*.

Notably, all shared CpGs exhibited concordant effects for IL-6 and leptin, whereas their adiponectin associations were in the opposite direction. This pattern was biologically coherent: IL-6 and leptin are both pro-inflammatory and metabolically unfavourable signals, often increasing with adiposity, while adiponectin is anti-inflammatory and typically diminished in immunometabolic disease. The persistence of this directional signature across multiple loci pointed to common mechanisms connecting these biochemical markers and highlighted these sites as potential nodes of regulatory control.

Three loci emerged as particularly compelling exemplars within the context of established literature: cg06500161, cg17901584, and cg00574958. Each of these CpGs had been robustly reported in over 25 previous EWAS, spanning a spectrum of immunometabolic phenotypes including BMI[19–22], lipid levels, and T2D[23]. This thesis linked all three sites to adipokine levels, while also associating cg17901584 with IL-6 (*Chapters 2 and 3*).

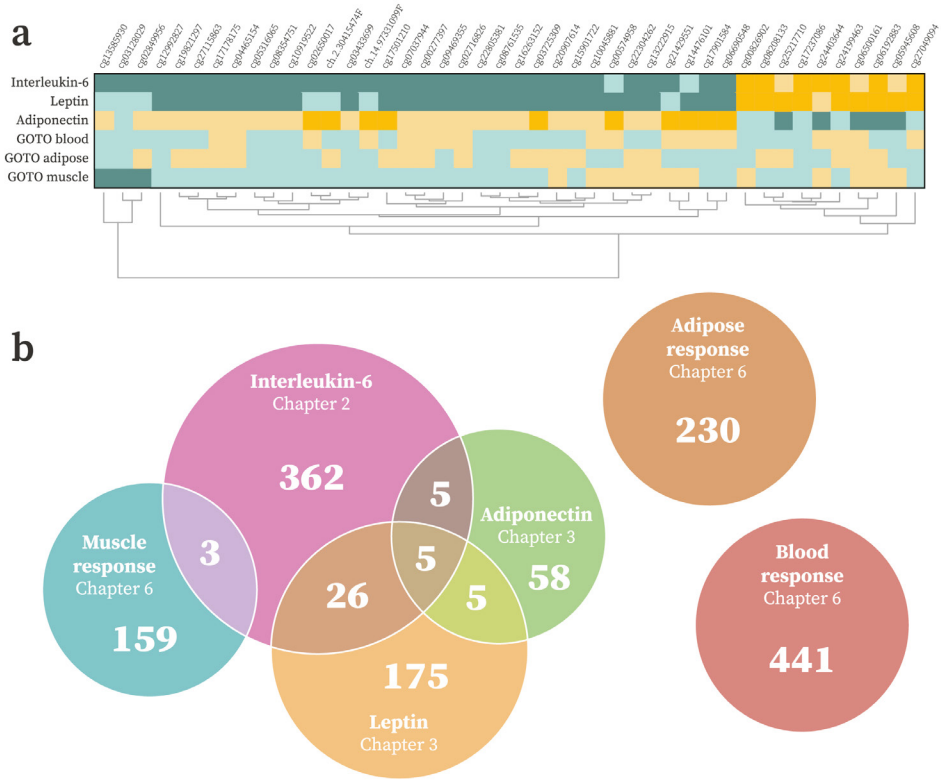


Figure 2 | CpG-level findings across this thesis. a) Heatmap of 44 CpGs identified in multiple investigations, clustered and coloured by effect size (positive: yellow, negative: blue). CpGs not reaching significance are shown in muted tones. **b)** Venn diagram illustrating number and overlap of CpGs found in each Chapter.

Functional annotation connected DNAm at these CpGs to genes critical for lipid transport (*ABCG1*)[24], biosynthesis (*DHCR24*)[25], and metabolism (*CPT1A*)[26], reinforcing methylation at lipid-related loci as a robust molecular readout of immunometabolic health.

The overlap between investigations in this thesis varied markedly (**Fig. 2b**). The pronounced tissue-specificity of DNAm responses to GOTO underscored how critical cellular context is to epigenetic patterns indicating that, in this setting, changes in blood do not mirror adaptations in the other tissues. Nevertheless, DNAm patterns did align with expression findings from this study, with the adipose and muscle methylome responding substantially, alongside smaller changes in blood[27]. Furthermore, the limited concordance between cross-sectional and experimental Chapters suggested that certain CpGs may capture long-term physiology, potentially shaped by cumulative lifetime exposure, whereas others reflect dynamically modifiable loci. This thesis, therefore, contributed to disentangling stable from dynamic sites, a necessary distinction for understanding heterogeneity in human health and intervention responsiveness.

Shared regulatory architecture

While DNAm was traditionally regarded as a repressive mark, particularly when located in promoters or enhancers, its regulatory consequences are now recognized as strongly context dependent. To characterise the chromatin landscape surrounding identified CpGs, we leveraged cell-type and tissue-specific reference epigenomes from the Roadmap Epigenomics Consortium[28]. These integrated epigenetic features to define eight active and seven repressive states, thereby positioning CpGs within their broader genomic environment and evaluating their functional potential (**Fig. 3a**).

In the adipocytokine EWAS (*Chapters 2 and 3*), as well as the muscle response to the GOTO intervention (*Chapter 6*), identified CpGs resided disproportionately within enhancers, consistent with a potential to disrupt TF binding and modulate nearby gene expression in relevant cell types. By contrast, the other tissues examined in the GOTO intervention exhibited alternate patterns: blood-derived CpGs showed no substantial deviation from the background distribution, suggesting that DNAm responses in this tissue reflect diffuse or systemic effects rather than focal transcriptional regulation. CpGs underlying the adipose tissue response were enriched in polycomb-repressed regions, a chromatin state linked to developmental gene silencing that accumulates DNAm as polycomb repression wanes with age[29]. This pattern raised intriguing questions about whether intervention-induced weight loss may have reinforced mature adipocyte identity through silencing of developmental transcriptional programs.

To gain complementary insights into the regulatory context of our DNAm associations, we also performed TF binding site (TFBS) enrichment analyses. Across Chapters, we identified twelve TFs whose binding sites were enriched within 50bp of CpG sets from multiple investigations (**Fig. 3b**). Notably, eleven of these belonged to the basic leucine

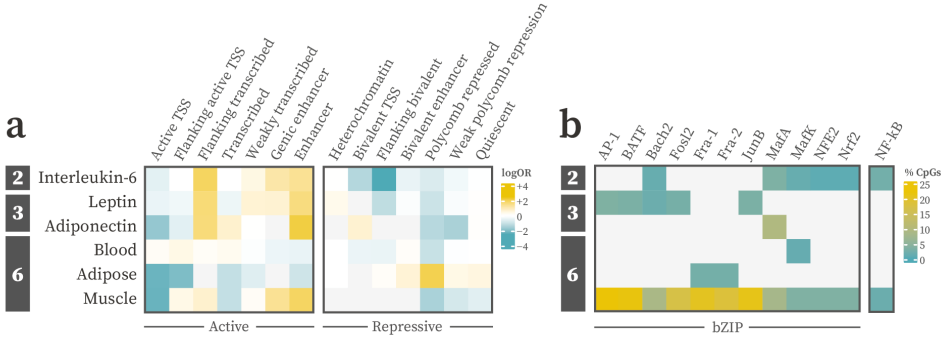


Figure 3 | Regulatory context of results across this thesis. a) Heatmap of chromatin state enrichment, separated into active and repressive states. Colours represent the log-transformed odds ratios (ORs). **b)** TFs whose binding sites were enriched within 50bp of identified CpGs in each Chapter. Colours indicate the percentage of CpGs underlying the enrichment, with all shown significant at the 5% level after adjusting for multiple testing.

zipper (bZIP) family, a group of dimeric TFs that recognise short palindromic motifs through a basic DNA-contact region (*b*) coupled to a leucine zipper dimerization domain (*ZIP*). Members of the bZIP family are established as regulators of cellular stress responses and metabolic flexibility[30], and their recurrent enrichment across this thesis pointed to a potentially conserved regulatory architecture underpinning investigated immunometabolic traits.

Previous work has demonstrated that most bZIP TFs exhibit reduced or abolished binding at methylated motifs, including several identified in our analysis including MafA, JunB, and Bach2[31]. Such sensitivity may have particular relevance at enhancers, where decreased DNAm could thereby facilitate TF binding and modulate target gene expression[32]. Across this thesis, traits associated with CpGs enriched for enhancers and bZIP TFBS were also predominantly negatively correlated with DNAm (70.3% of IL-6 CpGs in *Chapter 2*; 64.4% of adiponectin CpGs and 67.8% of leptin CpGs in *Chapter 3*; and 87.7% of CpGs that responded in muscle in *Chapter 6*). Our results therefore supported a model where elevated adipocytokine levels and intervention effects in muscle were accompanied by reduced DNAm at enhancers, increasing bZIP TF binding propensity and expression of stress-related and metabolic target genes.

Convergence on immunometabolism

Annotating CpGs to their target genes remains a major challenge in epigenetic research, reflecting both the context-dependent nature of DNAm and the crude nature of nearest gene pairings[33,34]. A more evidence-based strategy is expression quantitative trait methylation (eQTM) analysis, which directly correlates expression and methylation in matched samples. In the adipocytokine EWAS, we calculated eQTMs using the BIOS consortium, a large-scale resource comprising DNAm and RNA-seq profiles from over 3,000 individuals. This approach revealed several biologically plausible examples where CpGs were linked to their non-nearest genes, including the IL-6 receptor (*IL6R; Chapter 2*), and key lipid homeostatic genes such as *ABCG1* and *CPT1A* (*Chapter 3*). We adapted this framework to the tissue-specific experimental setting in *Chapter 6*, conducting paired analyses to assess whether genes responsive to the GOTO intervention in each tissue were associated with local DNAm adaptations, directly linking epigenetic changes to observed transcriptional responses.

To further refine CpG–gene links, we incorporated genetic association data in colocalization analyses, combining methylation quantitative trait loci (QTLs) from the Genetics of DNA Methylation Consortium (GoDMC)[35] with expression QTLs from eQTLGen (*Chapter 2*)[36]. This specific approach evaluated whether two traits shared causal variants, revealing common underlying regulatory architecture. This was particularly valuable given that DNAm can reflect transcriptional potential even for genes with low or absent expression at the time of sampling. Collectively, integration of large-scale population

data, dynamic tissue-specific eQTMs, and colocalisation analyses provided a more comprehensive and functionally grounded map of CpG-gene relationships across this thesis, and these strategies could serve as a starting point for data-driven annotation initiatives moving forward.

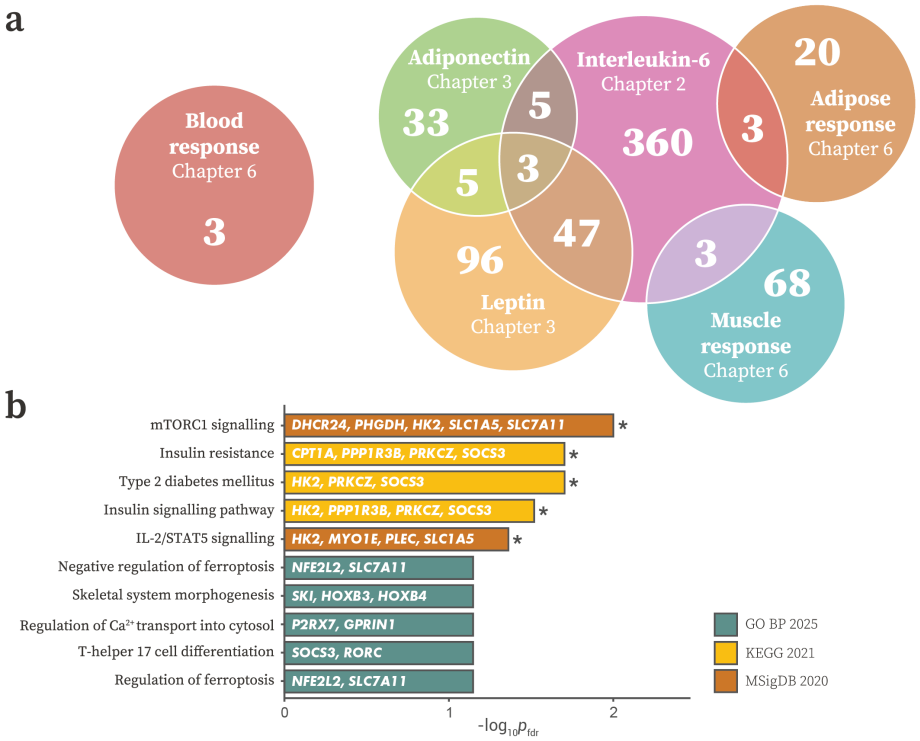


Figure 4 | Gene-based results across this thesis. a) Venn diagram showing number and overlap of genes identified in each Chapter. **b)** Top ten terms from an over-representation analysis of the 66 genes implicated in more than one investigation. Stars indicate significance at the 5% level after adjusting for multiple testing.

To explore common biological pathways implicated across Chapters, we performed an over-representation analysis of the 66 genes detected in multiple investigations (**Fig. 4a,b**). Enriched terms included insulin resistance (p_{FDR} = 0.021) and signalling (p_{FDR} = 0.032), mTORC1 signalling (p_{FDR} = 0.011), and T2D (p_{FDR} = 0.021). Insulin resistance represents a critical driver of immunometabolic dysfunction, contributing to the development of T2D, while mTORC1 functions as a nutrient-sensing hub with strong links to longevity and whose inhibition is emerging as a multimodal therapeutic[37–39]. Notably, many of the overlapping genes mapped to canonical mTORC1-driven effects, including translational upregulation, lipid accumulation, oxidative stress, and inflammation (**Supplementary Figures 1-4**). The convergence of multiple analyses on these pathways therefore highlighted immunometabolic regulation as a biological theme unifying this thesis.

Chapter-specific findings corroborated this pattern, with systematic interrogation of gene set databases and previous EWAS revealing enrichments for immunometabolic pathways and risk factors, including BMI, high-density lipoprotein (HDL) cholesterol, and fasting insulin, as well as T2D itself (**Fig. 5a,b**). Adiponectin-associated CpGs in particular displayed consistent and pronounced enrichment patterns, reinforcing findings that these sites may hold translational relevance for diabetes pathogenesis (*Chapter 3*). These coordinated relationships thereby consolidated the broader conclusion that epigenetic signatures across this thesis converged onto core immunometabolic phenotypes.

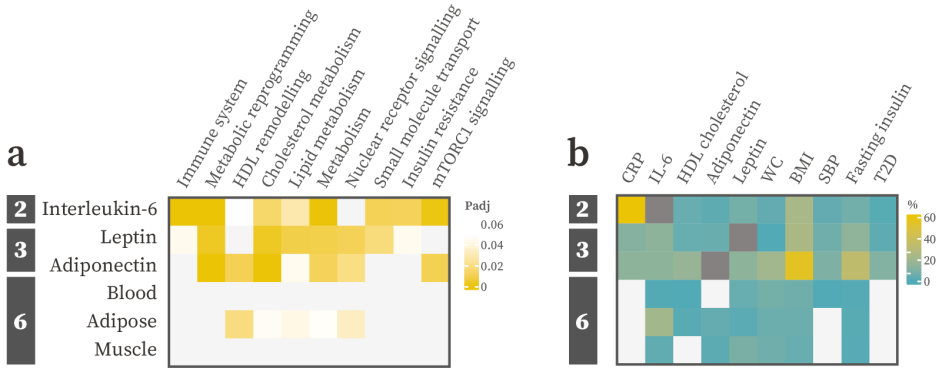


Figure 5 | Higher-order enrichments across this thesis. a) Heatmap showing pathway-level enrichments, coloured by FDR-adjusted *p*-value. **b)** Heatmap of trait-level enrichments (*Chapter 2 and 3*) and longitudinal associations (*Chapter 6*) across immunometabolic phenotypes. Colour represents the percentage of identified CpGs additionally associated with the labelled trait. Dark grey indicates index traits. All enrichments are significant at the 5% level after adjusting for multiple testing.

Considering causality

By altering chromatin accessibility and DNA interactions with regulatory proteins, CpG methylation can modulate local gene expression[40]. Yet, this regulatory *potential* by no means guarantees that DNAm-trait associations identified in EWAS arise from direct, functional effects within studied cells. A central challenge in epigenetic epidemiology is distinguishing DNAm that actively influences the investigated trait (**Fig. 6a,b**) from changes driven by reverse causation (**Fig. 6c**)[5,41,42]. While study design can assist by delineating parts of the causal pathway, as showcased by the GOTO intervention which preceded DNAm by design (*Chapter 6*), proposing plausible chains of events in EWAS requires integration of causal inference approaches with functional genomics and targeted follow-up.

A central component of this strategy was 2SMR, an instrumental variable framework that exploits the fact that genotypes are fixed at conception and almost entirely unaffected by exposures or phenotypic variation[43]. For each identified CpG, we leveraged

independent single nucleotide polymorphisms (SNPs) strongly associated with DNAm as genetic proxies. By combining these with large-scale genome-wide association study (GWAS) data, we interrogated not only if DNAm was likely driving the relevant adipocytokine, but also whether it might influence additional health-related phenotypes. These analyses identified one CpG that plausibly modulated IL-6 levels, and two loci that may alter adiponectin production, with evidence supporting all three sites as consequential for immunometabolic disease (*Chapters 2 and 3*). This thesis therefore highlighted both the value of embedding causal inference within EWAS, and the benefits of extending such analyses to encompass a range of clinically relevant outcomes.

This work was further strengthened by the emergence of QTL databases cataloguing associations between SNPs across the genome and molecular traits, including DNAm (GoDMC) and expression (eQTLGen)[35,36]. These resources provided unprecedented access to genetic association information, and we exploited both *cis*- and *trans*- methylation-QTL effect estimates to apply 2SMR bidirectionally. By comparatively inferring which direction of effect was most strongly supported by the data, we improved confidence in our 2SMR-based conclusions (*Chapters 2 and 3*).

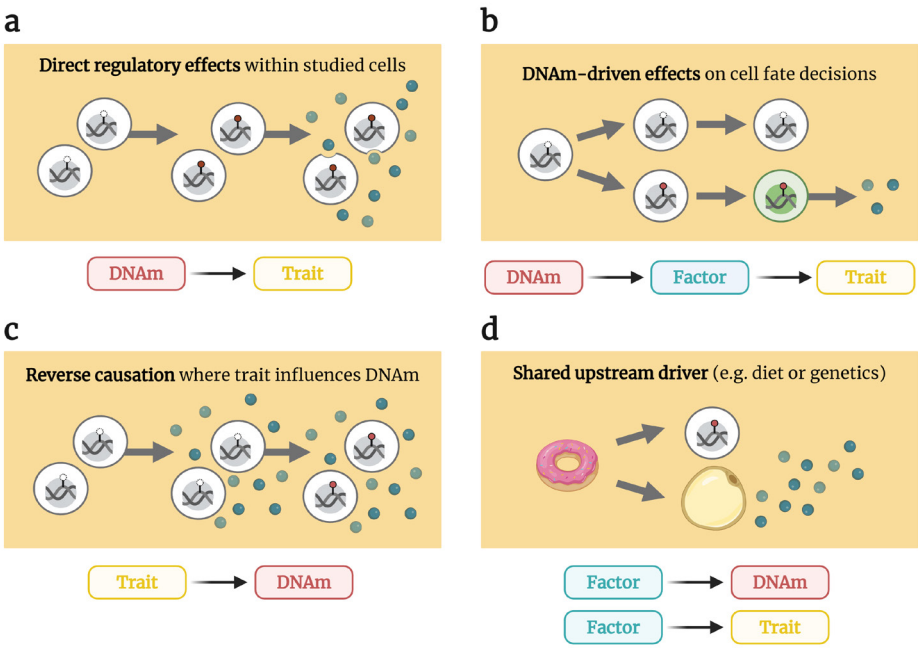


Figure 6 | Four potential causal structures underlying DNAm-trait associations in EWAS. a) Direct regulatory effects within studied cells alters expression and thereby influences the investigated trait. **b)** DNAm directed cell fate decisions lead to changes in sample composition that subsequently alter the trait. **c)** The trait influences DNAm in studied cells. **d)** Shared upstream factors like diet cause independent effects on DNAm in studied cells and the trait, including through DNAm in a non-studied cell type. *Figure created in BioRender.*

To complement 2SMR, this thesis also employed triangulation; a framework that adopts a more holistic view of causality. This approach posits that if an exposure is generally causal for an outcome, then genetically determined outcome levels should correlate with those predicted from genetically driven exposure levels and outcome-exposure associations. In this manner, a direction of effect can be inferred even if statistical power is limited for individual CpGs[21,44]. Bidirectional triangulation served to cross-validate our 2SMR findings. In the IL-6 EWAS, the methods diverged with triangulation concluding that DNAm most often reflected circulating IL-6 levels rather than driving them. As such, the CpG identified as influencing IL-6 production by 2SMR represented an exception to this general rule (*Chapter 2*). By contrast, triangulation corroborated 2SMR findings for adiponectin, indicating that DNAm typically drives adiponectin production and thus strengthening evidence for the two plausibly causal loci implicated by 2SMR (*Chapter 3*).

Since triangulation analyses indicated that DNAm was more often a consequence rather than a cause of circulating IL-6 levels, we employed mediation analysis to position DNAm between IL-6 and its downstream health effects, assessing whether methylation might act as a mediating factor. For two CpGs, both linked to *SOCS3* expression, the evidence supported this sequence of events. This suggested that these sites are more than markers of health status, and play a functional role in mediating IL-6–driven disease risk (*Chapter 2*). As GWAS and QTL resources continue to expand, the power of approaches such as bidirectional 2SMR, triangulation, and mediation analysis will increase, enabling more precise disentanglement of the causal relationships underlying complex molecular pathways.

To further strengthen our directional conclusions, we incorporated functional genomics throughout this thesis. This overlaid identified CpGs onto chromatin state and TFBS annotations and assessed enrichment for regulatory marks. In this manner, evidence for a genuine functional role was strengthened when CpGs implicated as causal fell within active chromatin close to relevant TFBS in studied cells (**Fig. 6a,b**). Sites in repressive regions were deemed more likely to reflect reverse causation or confounding (**Fig. 6c,d**). For example, 2SMR identified two CpGs plausibly influencing adiponectin: one close to *ADIPOQ* and the other associated with *SREBF1*. Functional genomic annotation revealed that the *SREBF1* site resided in a regulatory region within circulating cells, supporting a direct functional role in blood. In contrast, the *ADIPOQ* CpG lay in a repressive chromatin state in leukocytes, making such an effect less likely. In both cases, functional genomics provided essential information for interpreting causality and guiding follow-up analyses, highlighting the importance of considering the context-specific nature of DNAm in causal inference (*Chapter 3*).

Leveraging the reference epigenomes generated by the Roadmap Epigenomics Consortium[28], we examined the *ADIPOQ* site in greater detail. Since adiponectin is almost

exclusively produced by adipocytes, a direct effect of DNAm on its production in immune cells was biologically implausible; a view corroborated by this CpG's location in a repressive chromatin region in circulating cells. In adipocytes, however, this same locus was situated within an enhancer. This observation led us to hypothesize that a shared upstream driver, such as diet, could induce correlated epigenetic responses in a tissue-agnostic manner, generating DNAm-trait associations in blood that mirror functional effects in adipose tissues (**Fig. 6d**). To evaluate this hypothesis, we performed a targeted follow-up in adipocyte data, uncovering connections between *ADIPOQ* expression and methylation in these cells[45]. This finding not only strengthened support for our hypothesis at this CpG but also illustrated the utility of large-scale blood-based data to detect meaningful loci even for traits non-hematopoietic in origin (*Chapter 3*).

Countering confounding

As epidemiological investigations, EWAS are bound by the traditional rules of observational study design. As such, they are vulnerable to multiple forms of confounding (**Fig. 7a,b**)[46]. Variations in sample handling, array position, or environmental conditions during processing can be associated with both the trait of interest and DNAm, even when standard randomisation protocols are implemented[47]. Such technical artifacts do not reflect underlying biology in the organism or test system sampled and, if not properly addressed, can obscure signals or generate misleading results[48,49].

Although addressing technical confounding proactively is an established improvement over *post hoc* approaches, accessible tools for sample layout optimisation were limited[48,50]. To fill this gap, this thesis introduced *Omixer*, a Bioconductor package developed to minimise associations between biological variables and technical factors at the experimental design stage (*Chapter 4*). In the GOTO intervention study, *Omixer* ensured that all investigated correlations were both non-significant and had absolute values below 0.1 (**Fig. 7c**). *Omixer* also handled the paired design of this study, in which the intervention response was our primary focus, by positioning pre- and post-intervention samples from the same tissue and individual adjacently on the EPIC array (**Fig. 7d**). This arrangement largely eliminated concerns about technical confounding for intervention effects (*Chapter 6*). By developing user-friendly software to proactively control confounding and by demonstrating its practical application, this thesis showcased a strategy for routinely implementing best-practice recommendations to manage technical effects in high-throughput genomic studies moving forward.

Beyond technical factors, EWAS are also sensitive to biological confounding, where differences in cellular composition across samples are associated with both methylation and the trait of interest[46,51]. Blood, the most accessible and widely studied tissue in epigenetic epidemiology, offers exceptional opportunities for powerful large-scale analyses, but its inherent heterogeneity complicates interpretation of results.

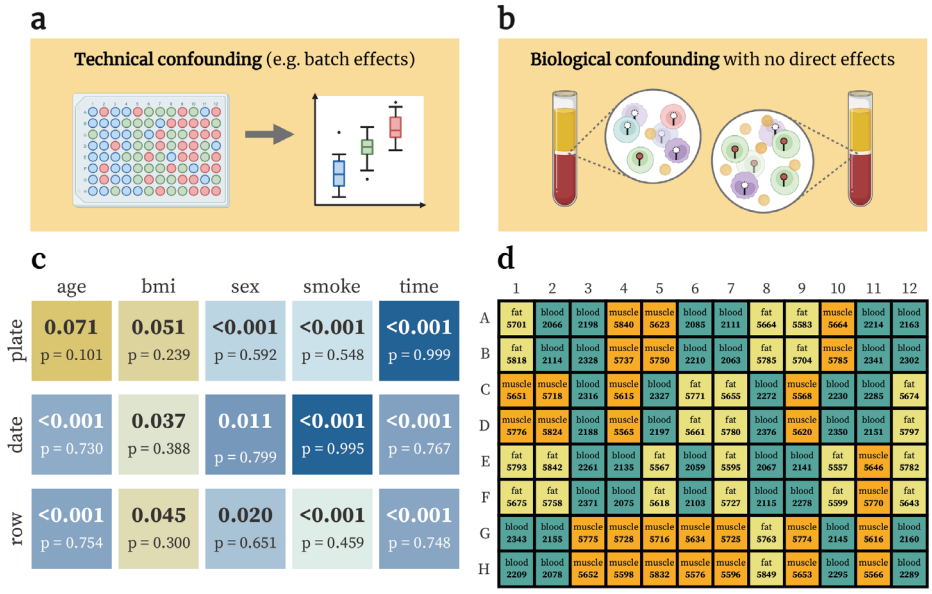


Figure 7 | Sources of and solutions to confounding in EWAS. a) Associations between the trait and technical factors induce systematic DNAm differences. **b)** Associations between sample composition and trait levels (represented by yellow circles) leads to apparent DNAm effects at cell-type specific CpGs. **c)** Omixer-output associations between biological variables and technical factors for DNAm samples profiled before and after the GOTO intervention in *Chapter 6*. **d)** Example of a lab-friendly plate layout generated by *Omixer* showing adjacent sample pairs along columns which were transferred to the eight vertical wells of the EPIC array. *Figure partly made in BioRender.*

Recent advances in algorithmic deconvolution, including within the IDOL-ext and UniLIFE frameworks, allow estimation of a wide range of immune cell subtypes from DNAm data[52,53]. This thesis leveraged these modern prediction tools to perform rigorous sensitivity analyses excluding CpGs whose effects were no longer significant after adjustment for estimated cell-type proportions, thereby ensuring focus remained on sites independently associated with the relevant adipocytokine (*Chapters 2 and 3*). While deconvolution methods continue to improve, they do not yet reach the resolution of purified cell assays. Future advances in single-cell DNAm profiling may provide the granularity needed to fully resolve the cellular sources of identified signals in EWAS.

In the GOTO intervention study, we had access to both these DNAm-derived cell-type proportion estimates, and cell counts measured by routine clinical testing. Using these, we were able to conclude, at multiple resolutions, that the immune cell proportions of participants was stable. Furthermore, in this multi-tissue setting, we extended deconvolution by integrating bulk RNA-sequencing profiles with single-cell and single-nuclei reference transcriptomes[54,55]. This enabled prediction of a distinct set of immune cell subtypes and allowed estimation of cell-type proportions in adipose and muscle tissues. Here, we detected an increase in endothelial cells in muscle post-intervention, consistent

with angiogenic adaptations to physical activity (*Chapter 6*)[56]. Taken together, these examples demonstrated the value of deploying deconvolution strategies in EWAS and illustrated how genomic data can derive meaningful readouts in complex and heterogenous tissues.

Leveraging genomic resources

The impact of integrative DNAm analyses is greatly amplified when paired with large-scale resources, collaborative initiatives, and rich genomic information. In this thesis, we combined in-depth functional annotations (**Fig. 8**) with a broad portfolio of available and novel data (**Table 2**) to strengthen our results. This included leveraging large-scale population cohorts and consortia for eQTM and meta-analyses (BBMRI-BIOS, TwinsUK, SHIP-TREND, and KORA F4; *Chapters 2 and 3*). By aggregating data across populations in this manner, we pooled statistical power and revealed modest but consistent DNAm associations that may have remained undetected in individual cohorts.

The Gene Expression Omnibus (GEO) proved particularly valuable as an accessible data repository. From GEO, we sourced not only the adipocyte and immune cell follow-up data (*Chapters 2 and 3*) but also single-cell and single-nuclei reference transcriptomes (*Chapter 6*)[57,58], and EPIC array data to test the external validity and applicability of *DNAmArray* (*Chapter 5*)[59]. This open-source workflow also incorporated publicly available probe information, excluding probes that were cross-reactive or contained common SNPs (from the Zhou-lab GitHub[60]), and flagging unreliable genomic regions (ENCODE blacklist[61]). This ensured that generated data was of high-quality with reduced risk of unreliable signals. Collectively, the combined utility of these resources showcased the value of accessible repositories such as GEO and Bioconductor and highlighted the importance of data and software democratisation in science.

A detailed understanding of positional context is essential for disentangling molecular architecture. Across this work, genomic locations and ranges were combined with cell-type reference epigenomes (from the Roadmap Epigenomics Project[28]) and TFBS motifs (accessed via HOMER[62]) to situate DNAm findings within regulatory landscapes. To explore causal directions, we extended this further, integrating SNP associations with DNAm (from GoDMC[35]), expression (from eQTLgen[36]), phenotypes (from GWAS catalog and IEU GWAS databases), and each other (via LDlink[62]), in instrumental variable based causal inference approaches. Overall, by systematically layering genetic, epigenetic, transcriptomic, and phenotypic resources in complementary ways, we strengthened the biological interpretation of our findings and ensured their relevance beyond the immediate scope of this thesis.

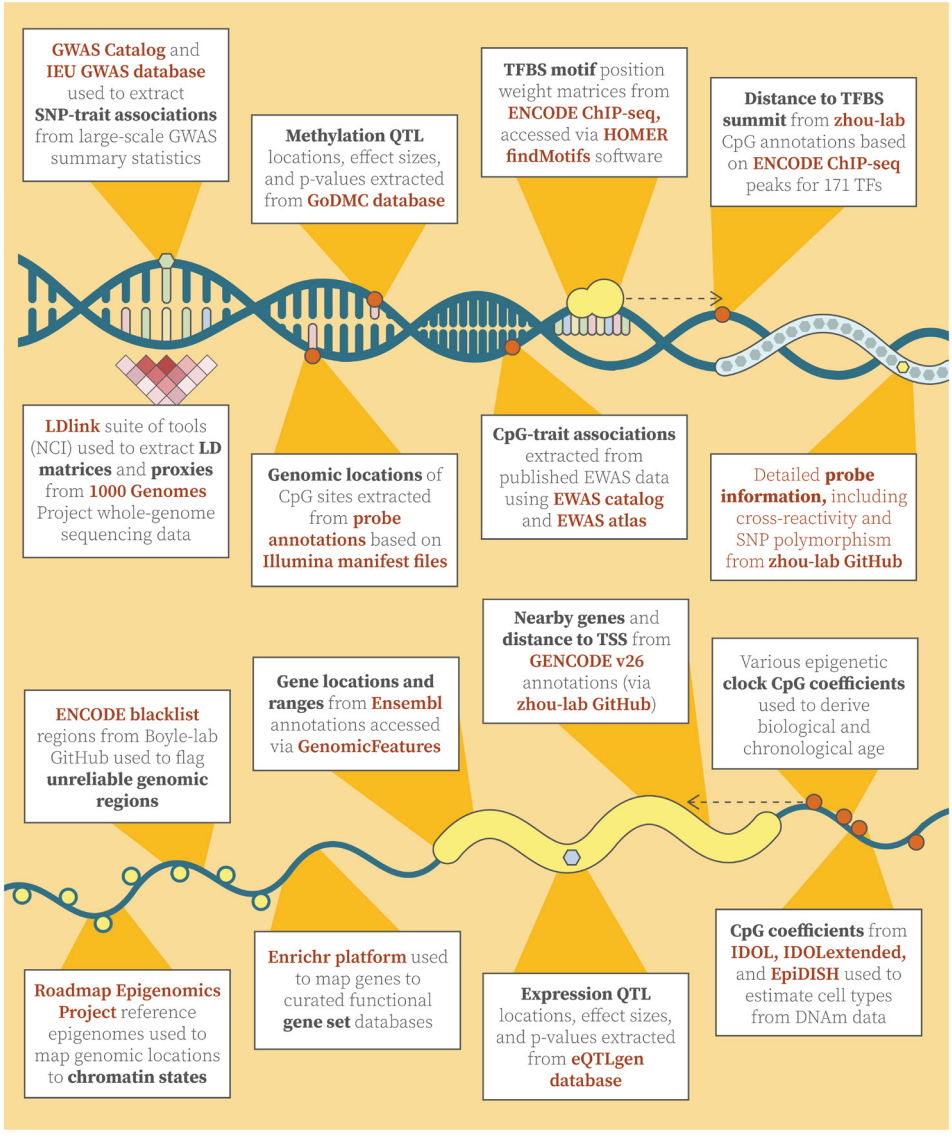


Figure 8 | Annotation information used throughout this thesis describing SNP-associations (LDlink, GWAS Catalogues, GoDMC, and eQTLgen), CpG associations (EWAS catalogues, CpG coefficients for clocks and cell counts), genomic co-ordinates (TFBS motifs, CpGs, genes, ENCODE blacklist, and chromatin states) and distances (CpG to TFBS summit, CpG to gene TSS), as well as detailed probe and gene set information.

Data	DNAm	Expression	Phenotypes	Chapters
BBMRI-BIOS	450k array blood	RNA-seq blood	age, sex, smoking, hsCRP, BMI, IL-6, adipokines	2, 3
KORA F4	450k array blood	-	age, sex, smoking, hsCRP, BMI, IL-6, adipokines	2, 3
SHIP-TREND	EPIC array blood	-	age, sex, smoking, BMI, adipokines	3
TwinsUK	450k array blood	-	age, sex, smoking, BMI, adipokines	3
GSE119593 GSE119539	450k array adipocytes	microarray adipocytes	-	3
GSE116339	EPIC array blood	-	age, sex, exposure status	5
Growing Old TOgether data	EPIC array 3 tissues	RNA-seq 3 tissues	age, sex, smoking, multiple immunometabolic traits	6
GSE149938	-	scRNA-seq blood	-	6
GSE167186	-	scRNA-seq muscle	-	6

Table 2 | DNAm and expression datasets used across this thesis. (WC: waist circumference, SBP: systolic blood pressure)

Drawing on our experience with novel EPIC data and large-scale population cohort analyses, we developed a modular and transparent pipeline providing current guidelines for processing and analysing DNAm data from 450k, EPIC, and the latest EPICv2 arrays. *DNAmArray* integrated several Bioconductor packages from our research group, including *MethylAid* for interactive quality control[16], *omicsPrint* for sample identity verification[17], and *bacon* for bias and inflation correction[18]. This workflow was designed to be flexible and scalable, enabling researchers to tailor analyses to match their specific study designs and aims (*Chapter 5*). By combining best-practice computational tools with clear, accessible documentation, *DNAmArray* facilitated data processing and analysis for the GOTO intervention study (*Chapter 6*), and now serves to exemplify of how well-curated, open-source pipelines can keep pace with evolving platforms and analytical challenges.

7

Healthy ageing

Investigating the epigenetic component of age-related traits was a central aim of this thesis. While chronological age remains the single strongest risk factor for most common chronic diseases, individuals exhibit marked heterogeneity in ageing trajectories[64,65]. DNAm has emerged as a critical molecular layer that both reflects and potentially influences this inter-individual variability. Certain age-associated CpGs undergo predictable remodelling across the lifespan, with a subset of these forming the basis of epigenetic clocks[66,67]. These DNAm-based models have been refined not only to capture calendar

age but also biological age (bAge), a measure more intricately linked to health and physiological function. In the GOTO intervention, we applied four such bAge predictors with all registering intervention-driven reductions. This demonstrated that even modest, short-term lifestyle changes can recalibrate molecular aging processes across tissues. Crucially, these bAge shifts correlated with improvements in clinical health parameters, including BMI, total body fat percentage, and leptin levels. Yet, despite compelling findings, challenges remain to uncover the specific biological processes captured by these algorithms, with emerging evidence suggesting that they may be driven at least in part by leukocyte population remodelling[68,69].

The application of muscle and blood-based epigenetic clocks exemplified another core tenet of this work: that both the progression and modulation of health can be tracked across multiple tissues. In this thesis, we characterised blood-based signals (*Chapters 2, 3, and 6*); explored targeted adipocyte-focused follow-up and adipose tissue signatures (*Chapters 3 and 6*); and analysed skeletal muscle, including by evaluating of the muscle-specific MEAT deconvolution algorithm (*Chapter 6*). Such a multi-tissue perspective is critical to build a comprehensive understanding of ageing, as physiological decline manifests not only in the circulation but also within less accessible tissues, where it can impair metabolic function and homeostasis. Furthermore, both adipocytes and myocytes are far from passive targets of signalling molecules; they actively secrete pro- and anti-inflammatory mediators shaping systemic immunity and organismal energy balance[70,71]. By presenting findings across distinct tissues sitting at the crossroads of immune surveillance and metabolic regulation, we emphasised the physiological heterogeneity of age-related phenotypes, advocating for continued, tissue-targeted strategies to uncover the molecular mechanisms that limit or lead to age-related disease.

By interrogating both inter-individual differences via large-scale EWAS meta-analyses (*Chapters 2 and 3*), and intra-individual dynamics through a multi-tissue intervention study (*Chapter 6*), this thesis identified molecular features that reflected cross-sectional physiological states distinct from those responsive to short-term lifestyle changes. Taken together, these findings established both a conceptual framework and a practical toolkit for distinguishing and dissecting an individual's intrinsic predisposition to a healthier later life from the epigenetic plasticity required to improve it through behavioural change.

Future perspectives

As the great regulatory filter that governs when, where, and how our DNA is expressed, epigenetic mechanisms promise to function as a molecular babel fish, translating the dialogue between our environment, our genes, and ultimately our physiology[38]. This strategic position at the genomic interface endows epigenetic research with immense potential to illuminate, and perhaps reshape, human health trajectories. Yet, the epigenome is inherently complex and heterogenous, making clear signals difficult. Much progress to date, including in this thesis, has relied on large-scale consortia pooling data, resources, expertise, and crucially, statistical power. Only within such collaborative frameworks can we detect the modest but consistent signals required to chart the DNAm landscape across the lifecourse and in disease. The future of epigenetic epidemiology therefore hinges on preserving and nurturing these partnerships, recognising that the task of understanding human health is a collective scientific pursuit.

Numbers alone will be insufficient. The field must also prioritise longitudinal initiatives. Despite originally being hailed as a stable archive of lifelong exposure and long-term physiological condition, this thesis and other recent research has uncovered dynamic subsections of the methylome responsive to even mild and short-term perturbations. This duality between stability and plasticity raises fundamental questions about the biological character of DNAm: is it a sturdy barometer of past exposures, a sensitive readout of environmental influence, or both? Integration of observations from longitudinal, experimental, and cross-sectional findings will be needed to disentangle the temporal variability of different sets of CpGs.

Transitioning from statistical associations to mechanistic insight will require more precise mapping of CpGs to plausible target genes. Simple proximity-based mapping is inadequate for a regulatory system as context-dependent and tissue-specific as DNAm. The field requires a rigorous, evidence-based framework that assigns graded confidence scores to CpG-gene links proportional to the weight of supporting data. Core evidence pillars could encompass longitudinal, tissue-specific, large-scale, and colocalisation analyses. Taken together, the result would provide a regulatory atlas, capable of connecting DNAm signals to their most likely functional targets and enhancing the biological interpretation of EWAS results past, present, and future.

By annotating CpGs more precisely, we will also be better equipped to tackle one of the field's most enduring challenges: causality. Instrumental variable-based methods, such as 2SMR and triangulation, can suggest directionality, but their genetic basis makes them vulnerable to tissue pleiotropy. Embedding these approaches within functional genomics, mediation frameworks, and tissue-specific analyses offers more refined dissection of underlying causal structures. Such integrative strategies can help clarify whether loci identified in blood represent direct epigenetic regulation inside immune cells, or instead reflect pleiotropic effects in other, less accessible tissues.

Concrete causal inference will ultimately require functional experiments and technological innovation, however. *In vitro* affinity-based assays have been invaluable for uncovering how TFs differentially recognise methylated motifs, but they have clear limitations: by operating outside the context of the chromatin landscape, they fail to capture the dependencies that govern TF binding *in vivo*[72]. In contrast, emerging CRISPR-based epigenetic editing platforms achieve targeted, single-CpG modifications within the native genomic environment[73]. These versatile systems support combinatorial perturbation of multiple loci, and their effects can be profiled by other molecular readouts[74]. Although the application of such cutting-edge tools remains restricted to cellular models, they provide a cornerstone for causality. Together with epigenetic epidemiology of human cohorts, they can reveal not just that DNAm matters, but precisely how and why.

With momentum behind epigenetic epidemiology, the field must also orient itself towards clinical translation. This thesis indicated the capacity of epigenetic clocks to track bAge reductions following lifestyle modifications, underscoring their potential to serve as dynamic biomarkers for monitoring intervention efficacy. Expanding these algorithms across multiple tissues could reveal organ-specific ageing processes that contribute to systemic decline, offering a more granular view of physiology. Moreover, integrating DNAm with genetic data to derive polygenic epigenetic risk scores could enhance patient stratification by capturing both inherited predisposition and environmental responsiveness. Such composite metrics may detect the subtle subclinical shifts that underlie heterogeneity in immunometabolic health and healthy ageing.

The future of epigenetic research, then, is not a list of incremental improvements but an integrated vision combining complementary strengths: large-scale consortia providing power and diversity, longitudinal studies capturing temporal dynamics, functional genomics revealing mechanistic insight, and experimental validation of causality. When woven together and translated to the clinic, these threads promise to create a comprehensive framework for understanding how environment and genome co-author human health. Within this vision, epigenetic signals can serve as critical focal points for mechanistic inquiry and generate testable hypotheses about how molecular phenotypes interact. The challenge, then, lies in bridging the gap between discovery and delivery, not only refining our science but translating it into better health for our ageing populations.

Conclusions

This General Discussion has synthesised the main findings of this thesis within the broader framework of immunometabolic health and ageing, emphasizing translational relevance, causal inference, and functional interpretation. By tracing results from the nucleotide resolution of CpGs back up to higher-order biology, it underscored the value of integrative approaches in EWAS, shifting focus from static associations to a multidimensional understanding of epigenetics in human health and disease. Cross-Chapter comparisons revealed convergent methylation signatures linked to immunometabolic risk and highlighted clear focal points for targeted experimental validation.

This thesis set out to interrogate relationships between DNAm and three adipocytokines, whilst also identifying CpG sites sensitive to lifestyle changes in older adults. By expanding EWAS to evaluate both intra- and inter-individual variation, and by employing longitudinal, multi-tissue, and large-scale analyses, it illuminated molecular markers of the subtle shifts in health that accompany ageing and disease pathogenesis. Recognising the context-dependent nature of DNAm, results were situated within a broader regulatory landscape encompassing TF binding, chromatin accessibility, and underlying genetic architecture, enabling biologically grounded interpretations.

Beyond foundational insights, this thesis also developed and disseminated computational tools, supporting researchers from design to data analysis. Leveraging this software in conjunction with the wealth of available genomic resources, it demonstrated that when DNAm is processed with care and rigorously contextualised, it can reveal critical biological pathways and prioritise promising loci. In sum, this thesis made meaningful strides towards realising the potential of EWAS to deepen our understanding of immunometabolic health in ageing.

References

1. Jones, M. J., Goodman, S. J. & Kobor, M. S. DNA methylation and healthy human aging. Aging Cell 14, 924 (2015).
2. Walaszczyk, E. 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 61, 354 (2017).
3. Ling, C. & Rönn, T. Epigenetics in Human Obesity and Type 2 Diabetes. Cell Metab 29, 1028 (2019).
4. Szarc Vel Szic, K., Declerck, K., Vidaković, M. & Vanden Berghe, W. From inflammaging to healthy aging by dietary lifestyle choices: Is epigenetics the key to personalized nutrition? Clin Epigenetics 7, (2015).
5. Bock, C. Analysing and interpreting DNA methylation data. Nature Reviews Genetics 2012 13:10 13, 705–719 (2012).
6. Wagner, K. H., Cameron-Smith, D., Wessner, B. & Franzke, B. Biomarkers of Aging: From Function to Molecular Biology. Nutrients 8, 338 (2016).
7. Hirata, T. et al. Associations of cardiovascular biomarkers and plasma albumin with exceptional survival to the highest ages. Nat Commun 11, 1–17 (2020).
8. Mossmann, M. et al. Increased serum IL-6 is predictive of long-term cardiovascular events in high-risk patients submitted to coronary angiography: an observational study. Diabetol Metab Syndr 14, (2022).
9. Yadav, R., Yadav, R. K., Khadgawat, R. & Pandey, R. M. Comparative efficacy of a 12 week yoga-based lifestyle intervention and dietary intervention on adipokines, inflammation, and oxidative stress in adults with metabolic syndrome: A randomized controlled trial. Transl Behav Med 9, 594–604 (2019).
10. Zeng, Y. et al. Novel loci and pathways significantly associated with longevity. Sci Rep 6, (2016).
11. Zhou, L. et al. Centenarians Alleviate Inflammaging by Changing the Ratio and Secretory Phenotypes of T Helper 17 and Regulatory T Cells. Front Pharmacol 13, (2022).
12. Minciullo, P. L. et al. Inflammaging and Anti-Inflammaging: The Role of Cytokines in Extreme Longevity. Arch Immunol Ther Exp (Warsz) 64, 111–126 (2016).
13. van de Rest, O. et al. Metabolic effects of a 13-weeks lifestyle intervention in older adults: The Growing Old Together Study. Aging 8, 111–126 (2016).
14. Shimano, H. & Sato, R. SREBP-regulated lipid metabolism: Convergent physiology-divergent pathophysiology. Nat Rev Endocrinol 13, 710–730 (2017).
15. Seo, J. B. et al. Adipocyte determination- and differentiation-dependent factor 1/sterol regulatory element-binding protein 1c regulates mouse adiponectin expression. Journal of Biological Chemistry 279, 22108–22117 (2004).
16. Van Iterson, M. et al. MethylAid: visual and interactive quality control of large Illumina 450k datasets. Bioinformatics 30, 3435–3437 (2014).
17. Van Iterson, M., Cats, D., Hop, P. & Heijmans, B. T. omicsPrint: detection of data linkage errors in multiple omics studies. Bioinformatics 34, 2142–2143 (2018).
18. van Iterson, M. et al. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. Genome Biol 18, 1–13 (2017).
19. Sun, D. et al. Body Mass Index Drives Changes in DNA Methylation: A Longitudinal Study. Circ Res 125, 824–833 (2019).
20. Mendelson, M. M. et al. Association of Body Mass Index with DNA Methylation and Gene Expression in Blood Cells and Relations to Cardiometabolic Disease: A Mendelian Randomization Approach. PLoS Med 14, (2017).
21. Wahl, S. et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature 2016 541:7635 541, 81–86 (2016).
22. Demerath, E. W. et al. Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. Hum Mol Genet 24, 4464–4479 (2015).
23. Hillary, R. F. et al. Blood-based epigenome-wide analyses of 19 common disease states: A longitudinal, population-based linked cohort study of 18,413 Scottish individuals. PLoS Med 20, (2023).
24. Kennedy, M. A. et al. ABCG1 has a critical role in mediating cholesterol efflux to HDL and preventing cellular lipid accumulation. Cell Metab 1, 121–131 (2005).
25. Zerenturk, E. J., Sharpe, L. J., Ikonen, E. & Brown, A. J. Desmosterol and DHCR24: unexpected new directions for a terminal step in cholesterol synthesis. Prog Lipid Res 52, 666–680 (2013).
26. Schlaepfer, I. R. & Joshi, M. CPT1A-mediated Fat Oxidation, Mechanisms, and Therapeutic Potential. Endocrinology 161, (2020).
27. Bogaards, F. A. et al. Secondary integrated analysis of multi-tissue transcriptomic responses to a combined lifestyle intervention in older adults from the GOTO nonrandomized trial. Nature Communications 2024 15:1 15, 1–18 (2024).
28. Roadmap Epigenomics Consortium et al. Integrative analysis of 111 reference human epigenomes. Nature 2015 518:7539 518, 317–330 (2015).

29. Dozmorov, M. G. Polycomb repressive complex 2 epigenomic signature defines age-associated hypermethylation and gene expression changes. Epigenetics 10, 484–495 (2015).
30. Hai, T. et al. ATF3, a Hub of the Cellular Adaptive-Response Network, in the Pathogenesis of Diseases: Is Modulation of Inflammation a Unifying Component?
31. Yin, Y. et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science (1979) 356, (2017).
32. Llorca, C. M., Potschin, M. & Zentgraf, U. bZIPs and WRKYs: Two large transcription factor families executing two different functional strategies. Front Plant Sci 5, (2014).
33. Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how and why? Mol Cell 49, 10.1016/j.molcel.2013.01.038 (2013).
34. Hannon, E. et al. Leveraging DNA-Methylation Quantitative-Trait Loci to Characterize the Relationship between Methylomic Variation, Gene Expression, and Complex Traits. Am J Hum Genet 103, 654–665 (2018).
35. Min, J. L. et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. Nature Genetics 2021 53:9 53, 1311–1321 (2021).
36. Võsa, U. et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nature Genetics 2021 53:9 53, 1300–1310 (2021).
37. Mannick, J. B. & Lamming, D. W. Targeting the biology of aging with mTOR inhibitors. Nat Aging 3, 642–660 (2023).
38. Johnson, S. C., Rabinovitch, P. S. & Kaeberlein, M. MTOR is a key modulator of ageing and age-related disease. Nature 493, 338–345 (2013).
39. Passtoors, W. M. et al. Gene expression analysis of mTOR pathway: Association with human longevity. Aging Cell 12, 24–31 (2013).
40. Moore, L. D., Le, T. & Fan, G. DNA Methylation and Its Basic Function. Neuropsychopharmacology 38, 23 (2012).
41. Birney, E., Smith, G. D. & Greally, J. M. Epigenome-wide Association Studies and the Interpretation of Disease -Omics. PLoS Genet 12, e1006105 (2016).
42. Mill, J. & Heijmans, B. T. From promises to practical strategies in epigenetic epidemiology. Nat Rev Genet 14, 585–594 (2013).
43. Teschendorff, A. E. & Relton, C. L. Statistical and integrative system-level analysis of DNA methylation data. Nature Reviews Genetics 2017 19:3 19, 129–147 (2017).
44. Wielscher, M. et al. DNA methylation signature of chronic low-grade inflammation and its role in cardio-respiratory diseases. Nature Communications 2022 13:1 13, 1–14 (2022).
45. Tini, G. et al. DNA methylation during human adipogenesis and the impact of fructose. Genes Nutr 15, 21 (2020).
46. Teschendorff, A. E. & Zheng, S. C. Cell-type deconvolution in epigenome-wide association studies: a review and recommendations. Epigenomics 9, 757–768 (2017).
47. Leek, J. T. et al. Tackling the widespread and critical impact of batch effects in high-throughput data. Nat Rev Genet 11, 733–739 (2010).
48. Harper, K. N., Peters, B. A. & Gamble, M. V. Batch effects and pathway analysis: Two potential perils in cancer studies involving DNA methylation array analysis. Cancer Epidemiology Biomarkers and Prevention 22, 1052–1060 (2013).
49. Zindler, T., Frieling, H., Neyazi, A., Bleich, S. & Friedel, E. Simulating ComBat: how batch correction can lead to the systematic introduction of false positive results in DNA methylation microarray studies. doi:10.1186/s12859-020-03559-6.
50. Price, E. M. & Robinson, W. P. Adjusting for batch effects in DNA methylation microarray data, a lesson learned. Front Genet 9, (2018).
51. Loyfer, N. et al. A DNA methylation atlas of normal human cell types. Nature 613, 355–364 (2023).
52. Salas, L. A. et al. Enhanced cell deconvolution of peripheral blood using DNA methylation for high-resolution immune profiling. Nature Communications 2022 13:1 13, 1–13 (2022).
53. Guo, X. et al. Unified high-resolution immune cell fraction estimation in blood tissue from birth to old age. Genome Med 17, 1–26 (2025).
54. Wang, X., Park, J., Susztak, K., Zhang, N. R. & Li, M. Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. Nat Commun 10, 1–9 (2019).
55. Newman, A. M. et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol 37, 773–782 (2019).
56. Kwak, S.-E., Lee, J.-H., Zhang, D. & Song, W. Angiogenesis: focusing on the effects of exercise in aging and cancer. J Exerc Nutrition Biochem 22, 21–26 (2018).
57. Xie, X. et al. Single-cell transcriptomic landscape of human blood cells. Natl Sci Rev 8, (2021).
58. Perez, K. et al. Single nuclei profiling identifies cell specific markers of skeletal muscle aging, frailty, and senescence. Aging 14, 9393–9422 (2022).
59. Curtis, S. W. et al. Exposure to polybrominated biphenyl (PBB) associates with genome-wide DNA methylation differences in peripheral blood. Epigenetics 14, 52–66 (2019).

60. Zhou, W., Laird, P. W. & Shen, H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. Nucleic Acids Res 45, e22 (2017).

61. Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. Scientific Reports 2019 9:1 9, 1–5 (2019).

62. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576–589 (2010).

63. Machiela, M. J. & Chanock, S. J. LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics 31, 3555–3557 (2015).

64. Niccoli, T. & Partridge, L. Ageing as a risk factor for disease. Current Biology 22, (2012).

65. Partridge, L., Deelen, J. & Slagboom, P. E. Facing up to the global challenges of ageing. Nature 2018 561:7721 561, 45–56 (2018).

66. Bernabeu, E. et al. Refining epigenetic prediction of chronological and biological age. Genome Med 15, 1–15 (2023).

67. Horvath, S. DNA methylation age of human tissues and cell types. Genome Biol 14, (2013).

68. Jonkman, T. H., Consortium, B., Zwet, E. W. van & Heijmans, B. T. Probing epigenetic clocks as a rational markers of biological age using blood cell counts. medRxiv 2025.05.12.25327213 (2025).

69. Guo, X. et al. Variations in Innate Immune Cell Subtypes Correlate with Epigenetic Clocks, Inflammaging and Health Outcomes. Advanced Science e05922 (2025).

70. Maury, E. & Brichard, S. M. Adipokine dysregulation, adipose tissue inflammation and metabolic syndrome. Mol Cell Endocrinol 314, 1–16 (2010).

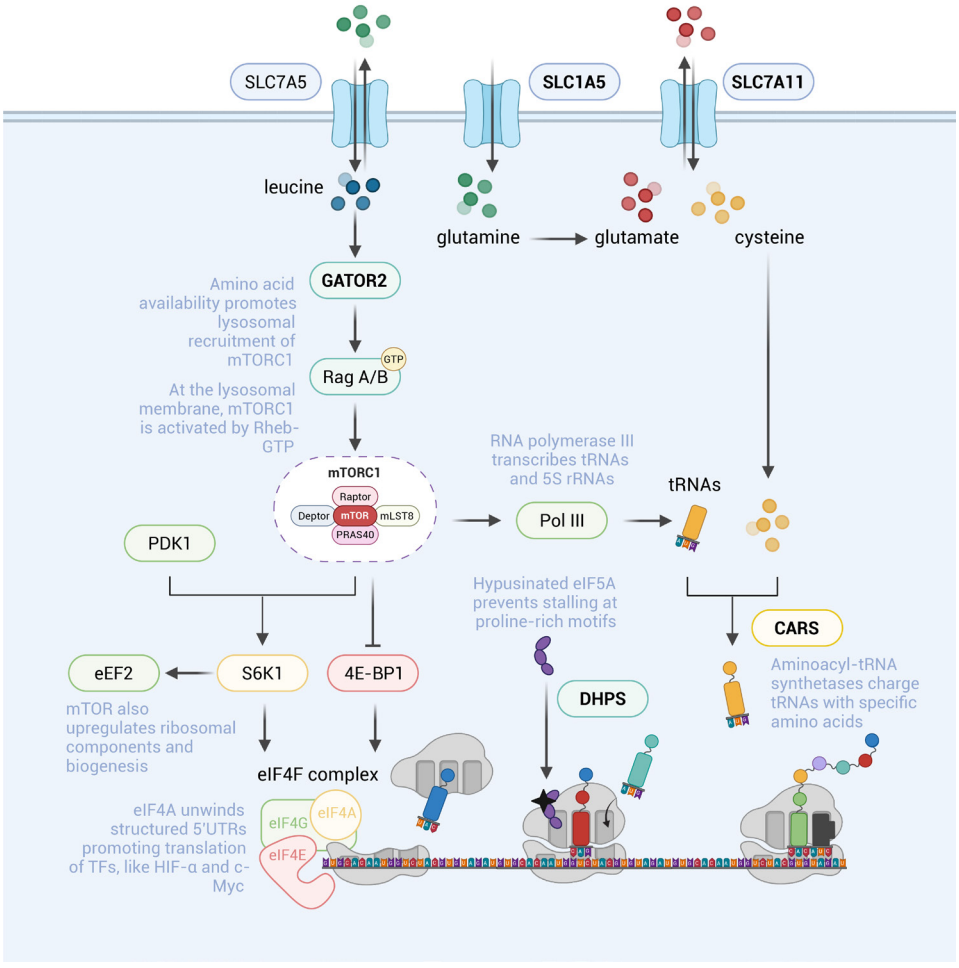
71. Leal, L. G., Lopes, M. A. & Batista, M. L. Physical exercise-induced myokines and muscle-adipose tissue crosstalk: A review of current knowledge and the implications for health and metabolic diseases. Front Physiol 9, (2018).

72. Liesenfelder, S. et al. Epigenetic editing at individual age-associated CpGs affects the genome-wide epigenetic aging landscape. Nature Aging 2025 5:6 5, 997–1009 (2025).

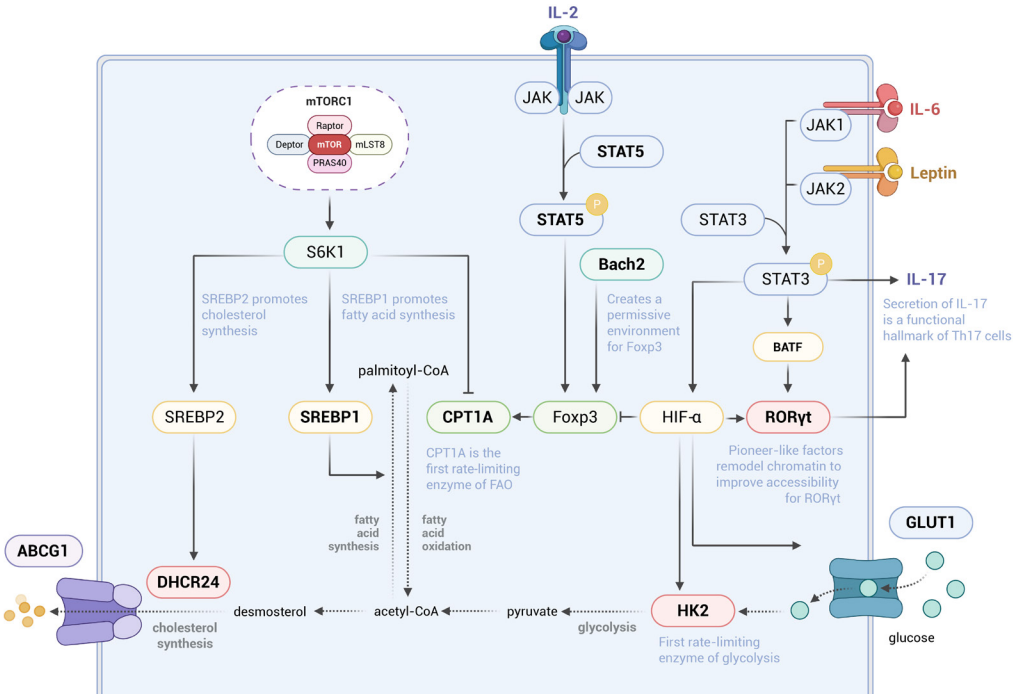
73. Zsido, B. and Hetenyi, C. Molecular structure, binding affinity, and biological activity in the epigenome. Int J Mol Sci. 21(11):4134 (2020).

74. de Groote, M., Vershure P., and Rots, M. Epigenetic editing: targeted rewriting of epigenetic marks to modulate expression of selected target genes. Nucleic Acids Res. 40(21):10596-613 (2012).

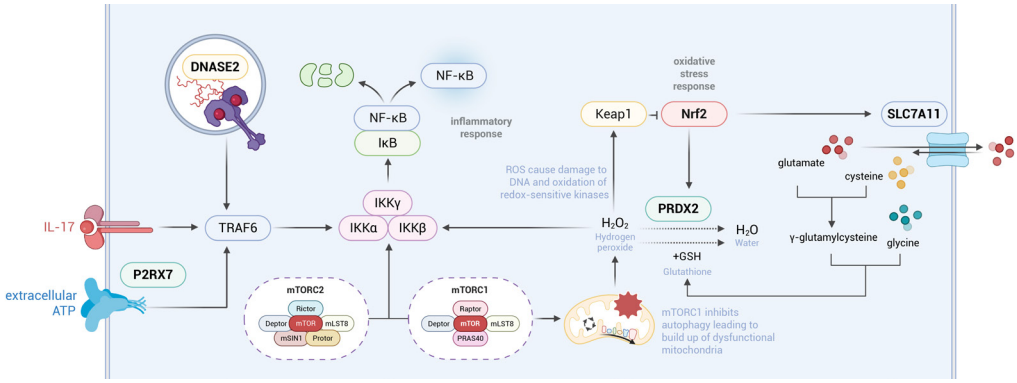
Supplementary Material



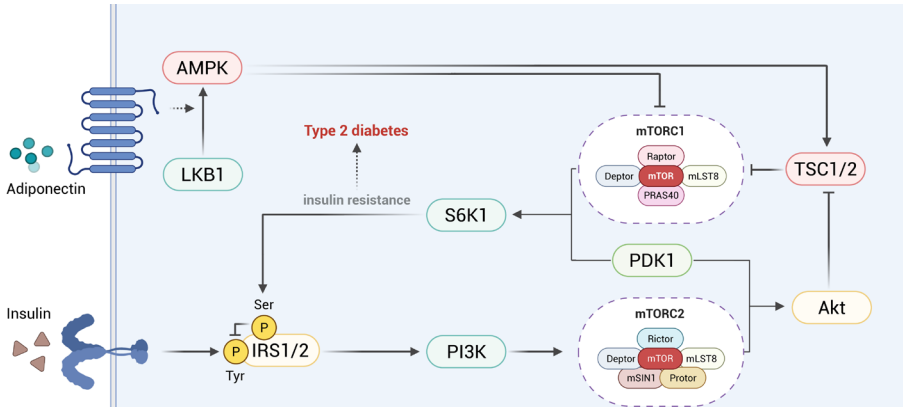
Supplementary Figure 1 | Amino acid supply promotes mTORC1 activation and downstream protein synthesis. Amino acids enter cells via transporters, such as SLC7A5, SLC1A5, and SLC67A11, where they trigger GATOR2 and RagA/B•GTP to position the mTORC1 complex at the lysosomal membrane ready for activation. Activated mTORC1 upregulates translation, fuelled by amino acids ligated to tRNAs by aminoacyl-tRNA synthetases, such as CARS, GARS, and VARS2, ribosomal biogenesis, and factors that prevent elongation stalling, such as hypusinated eIF5A generated via DHPS. By stimulating the eIF4F complex, mTORC1 preferentially upregulates the translation of mRNAs with structured 5'UTRs, including HIF-a and c-Myc. *Genes or proteins implicated in this thesis shown in the Figure in bold. Figure created in BioRender.*



Supplementary Figure 2 | Crosstalk between mTORC1, IL-6, and leptin in T-cell fate regulation. Th17 differentiation is promoted by IL-6 and leptin signalling, and TFs BATF, RORγt, and HIF-a. These upregulate glycolytic machinery, including GLUT1 and HK2 and suppress Treg pathways via inhibition of Foxp3. IL-2/STAT5 signalling and Bach2 promote Treg phenotypes by upregulating Foxp3 and fatty acid oxidation, including via CPT1A. mTORC1 actively undermines this by directly inhibiting CPT1A and further altering lipid handling through fatty acid synthesis (SREBP1) and cholesterol synthesis (SREBP2 and DHCR24). *Genes and proteins implicated in this thesis shown in the Figure in bold. Figure created in BioRender.*



Supplementary Figure 3 | mTOR drives both oxidative stress and inflammation. By inhibiting autophagy, mTORC1 promotes a build-up of dysfunctional mitochondria fuelling reactive oxygen species formation. This is characterised by Nrf2 activation of target genes, including the peroxiredoxin PRDX2 and cysteine/glutamate anti-porter SLC7A11. Oxidative stress drives DNA damage and inflammatory responses, the latter through oxidation of IKKB. NF-kB and inflammatory signalling is also triggered by misplaced self-DNA, whose clearance is impaired by reduced DNASE2, and extracellular ATP, detected by P2RX7. These highlight a role for the "garbaging" model of inflammaging in our results. *Genes and proteins identified in this thesis shown in the Figure in bold. Figure created in BioRender.*



Supplementary Figure 4 | Adiponectin dampens mTOR activity through AMPK activation. AMPK inhibits mTORC1 both through TSC1/2 activation and direct inhibitory phosphorylation of mTORC1 itself. This also has potential to counteract insulin resistance resulting from hyperactive mTORC1, mediated by S6K1 phosphorylation at inhibitory serine residues that blocks tyrosine phosphorylation essential for insulin signal transduction. *Figure created in BioRender.*