



Universiteit
Leiden
The Netherlands

Single-cell dissection of immune system aging and acute myeloid leukemia

Karakaslar, E.O.

Citation

Karakaslar, E. O. (2026, June 16). *Single-cell dissection of immune system aging and acute myeloid leukemia*. Retrieved from <https://hdl.handle.net/1887/4306748>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4306748>

Note: To cite this publication please use the final published version (if applicable).



Chapter 7

Discussion

This dissertation set out to study immune aging and hematopoiesis, with the overarching aim of understanding how these processes converge upon leukemogenesis in acute myeloid leukemia (AML). In conducting studies, we have explored how molecular pathways drive immune senescence and inflammation, how epigenetic manipulations can preserve or enhance hematopoietic stem cell (HSC) function, and how dissecting molecular data from AML at bulk or single-cell resolution can characterize the disease heterogeneity.

7.1 Immune Aging and Hematopoietic Decline: Mechanisms and Therapeutic Opportunities

In Chapter 2, we systematically studied immune aging in mice. Using transcriptional profiling and chromatin accessibility data, we demonstrated that older immune cells exhibit increased AP-1 activity, correlating with enhanced accessibility of the AP-1 binding site in genes encoding pro-inflammatory cytokines such as IL-6. AP-1 activity and accessibility of AP-1 binding sites in pro-inflammatory cytokine genes were also increased with age in short-lived (NZO) and long-lived (B6) mouse strains. Human data also showed that the accessibility profiles of *JUN* and *FOS*, together encoding the dimeric AP-1 transcription factor, are similar to those of mice. This suggests that AP-1 activation might be an evolutionarily conserved immune signature of aging. Our data refine the term inflammation by placing AP-1 transcription factors as one of the central nodes orchestrating pro-inflammatory gene programs. Yet, further studies are warranted to grasp whether these changes are predominantly driven by metabolic states (e.g., obesity, T2D) or by intrinsic epigenetic drift in immune cells.

Many studies show that inflammation is related to specific immune subsets like monocytes or macrophages¹⁻³. In the context of aging, this myeloid bias is associated with age-related clonal hematopoiesis, which is often driven by recognizable somatic mutations but can also occur in the absence of currently identifiable canonical driver variants⁴⁻⁷. Extensive studies on aging of the immune system demonstrated that a clonal mutation is found in less than 1% of individuals below 50 years of age⁵, but in more than 30% of those above 80 years^{6,7}. For instance, recent single-cell multi-omics data from bone marrow samples of multiple myeloma patients with a *DNMT3A* hotspot mutation revealed that these patients have a clonal expansion in the myeloid lineage⁸, suggesting that increased presence of myeloid cells may affect the entire bone marrow microenvironment, including cells interacting with lymphoid-derived plasma cells. As we age, numbers of myeloid-biased HSCs may increase, and this might be the fundamental culprit that eventually leads to systemic deterioration of the immune system. This leads to the fundamental question of whether myeloid skewing at the HSC level might induce aging of the immune system, and how this can be prevented.

If myeloid skewing in HSCs is indeed causal for immune aging, two complementary strategies may hold therapeutic potential: depleting myeloid-biased HSCs or expanding HSCs with more balanced output of lymphoid- and myeloid-biased HSCs. For instance, Ross *et al.*⁴ recently demonstrated that antibody-mediated depletion of myeloid-biased HSCs in aged mice reduced their dominance in the bone marrow and restored a more balanced pool of HSCs. This shift promoted the expansion of common lymphoid progenitors and increased the production of naive CD4⁺ and CD8⁺ T cells as well as naive B cells. Functionally, these changes translated into stronger vaccine-induced T-cell and antibody responses, improved control of viral infections, and a reduction in exhausted or age-associated immune subsets. Plasma levels of pro-inflammatory cytokines and other proteins elevated with age also decreased toward more youthful levels. Together, these

findings show that restoring the HSC balance not only corrects lineage skewing but also rejuvenates adaptive immunity at both the cellular and functional level, suggesting a promising strategy to delay immune aging and its associated morbidities.

Another way in which output and skewing of HSCs can be influenced is through epigenetic modulation. In Chapter 3, we investigated epigenetic manipulation of HSCs *ex vivo*. Using a BET inhibitor (CPI203) and HDAC inhibitor (Quisinostat), we demonstrated improved HSC expansion and viral transduction efficiency, which is critical for the success of gene therapy protocols. scRNA-seq data along with machine learning models showed that HSC pools were enriched after both treatment conditions. Additionally, Quisinostat unexpectedly skewed HSC differentiation toward the erythroid lineage, suggesting a potential advantage of using this drug for gene therapies to restore red blood cell disorders. Taken together, these findings indicate that epigenetic regulators originally developed for cancer treatment can, in lower dosages, be used to modulate stemness and lineage potential. The next step then would be to apply these gene therapy protocols in preclinical models to evaluate the safety and long-term engraftment potential of epigenetically modified HSCs.

7.2 Recurring computational ambiguities in analysis of cancer data: Context and Method Layers

A myeloid bias in hematopoiesis can also be the result of oncogenesis, like in the case of acute myeloid leukemia (AML). AML is characterized by abnormal expansion of malignant myeloid cells in the bone marrow, which blocks normal hematopoiesis. In Chapter 4, we tried to characterize this extreme derailment of hematopoiesis by applying deconvolution of bulk RNA-seq from over a thousand AML patients using a healthy bone marrow single-cell atlas as reference. This approach allowed us to calculate estimated cell compositions (ECCs) that recapitulated classical AML differentiation stages and revealed lineage arrests consistent with FAB subtypes. Importantly, these ECCs were not only descriptive of disease heterogeneity but also predictive of treatment outcomes. For example, samples enriched for CD14⁺ monocytes displayed resistance to the BCL-2 inhibitor Venetoclax, a pattern that outperformed BCL-2 expression alone. At the same time, the study highlighted critical challenges: leukemic blasts often deviate transcriptionally from their healthy counterparts, and results can shift depending on the reference atlas or resolution of the deconvolution method. So, our analyses revealed two recurring layers of ambiguity in analyzing cancer datasets: **context** (disease-specific phenotypes diverging from canonical healthy states) and **methods** (algorithm choice, modality, resolution).

As a concrete example of methods-related ambiguities, there are many algorithms to perform deconvolution on bulk RNA-seq data. In our study, we employed an algorithm called MuSiC⁹, based on its benchmark performance¹⁰ and that it is open-sourced, which allowed us to build the algorithm into our already existing HAMLET pipeline¹¹. Nevertheless, algorithm choice can strongly affect estimated fractions. Similar to our study, Zeng *et al.*¹² also deconvoluted bulk RNA-seq AML samples and systematically compared different algorithms by assessing how well estimated cell compositions correlated with matched single-cell RNA-seq data. In their evaluation, CIBERSORTx¹³ achieved the highest concordance, with MuSiC ranking second. Together, these findings underscore that while deconvolution is a powerful tool to dissect cell type composition in bulk data, interpretation must be tempered by awareness of algorithm-specific biases and limitations.

As an example of context-related ambiguities, cancer samples add further complexity because

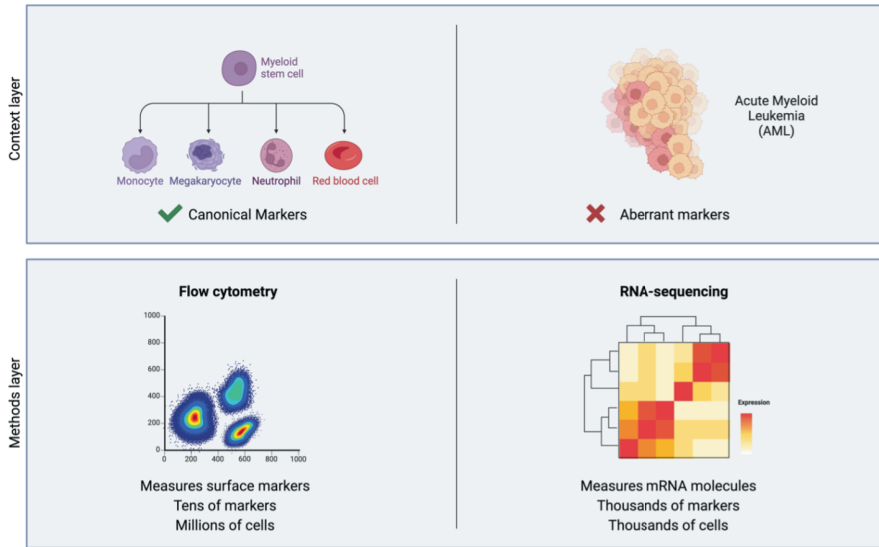


Figure 7.1: Examples for the two layers of ambiguity in analyzing cancer data.

aberrant phenotypes diverge from canonical healthy states¹⁰. Due to this complexity, AML samples are hard to phenotype by deconvolution. For instance, we used a detailed healthy bone marrow sample to deconvolute our in-house RNA-seq samples and correlated them with more coarse spectral flow cytometry phenotypes to understand the cell type compositions. However, although at a high level (e.g. monocytes) the methods agreed with each other, they do not perfectly match for detailed subsets due to technical differences, i.e. flow panels measure the presence of, in the order of ten, highly discriminative cell surface markers for millions of single cells, whereas bulk RNA-seq measures thousands of mRNA molecules for thousands of cells per sample (Figure 7.1).

These two layers also recurred in other chapters of the thesis while analyzing the different types of cancer datasets. For instance, in **Chapter 5**, we used bulk RNA-Seq from 1224 AML patients and demonstrated that the major AML subtypes in the World Health Organization (WHO) 2022 classification, which are defined by genetic aberrations, can also be successfully identified by gene expression profiling. However, we did experience ambiguities caused by the choice of methods when patients had to be assigned to robust clusters based on their transcriptional profiles. First, we identified either three major or five more refined transcriptional clusters of *NPM1*-mutated AML, which could be partly explained by co-mutational profiles. Secondly, the choice of clustering resolution was adhered to a balance between capturing the granularity of transcriptional heterogeneity and maintaining biological interpretability. A resolution that is too fine has the risk of overfitting to transcriptomic noise, while a resolution that is too coarse could obscure meaningful subgroups, particularly those driven by co-mutational patterns. Therefore, we evaluated clustering stability using consensus clustering and silhouette scores, ensuring that the resulting AML clusters were both robust and biologically relevant.

In **Chapter 6**, we used single cell RNA-seq and spectral flow cytometry to characterize the cellular

signatures of *NPM1*-mutated AML. Using 19 markers in flow cytometry, we were able to replicate the distinct transcriptional subtypes of *NPM1*-mutated AML, even though they were based on the expression of more than 20,000 genes measured by single cell RNA-Seq. However, methodological ambiguities also emerged. While the majority of markers showed good concordance between mRNA expression and protein abundance, some did not. For example, CD61 displayed robust protein staining despite weak or inconsistent transcript detection in the single-cell RNA-seq data. Such transcript-protein mismatches are well documented. Correlations between mRNA and protein are often modest, reflecting both technical limitations (e.g. mRNA drop-out, or degradation) and biological regulation (e.g. post-transcriptional control)^{14,15}. Another example of context layer in this chapter was a cluster of “High-mito” cells, enriched within one of the *NPM1*-mutated AML subtypes, which showed elevated mitochondrial gene expression yet remained below the conventional 15% threshold used to flag dying or low-quality cells. Their atypical profile, which was neither fully consistent with technical artifacts nor with canonical early progenitor phenotypes, illustrates how disease-specific metabolic states can complicate interpretation of cellular identities in AML.

Taken together, these findings indicate that analyses of cancer datasets are consistently influenced by two recurring and interacting sources of ambiguity: methodological and contextual. Methodological ambiguities arise from differences in algorithmic design, parameter choices, and modality-specific biases, while contextual ambiguities stem from malignant phenotypes that deviate from canonical healthy states. Careful recognition and transparent reporting of these limitations are essential to ensure that computational analyses yield results that are both robust and biologically meaningful.

7.3 Redefining Leukemia-Initiating Cells: Lessons from *NPM1* Mutations

A lot of the research in the field of AML is focused on the *leukemic stem cell* (LSC). A LSC is a malignant cell with self-renewal capacity that initiates the leukemia upon acquiring a genetic aberration. However, the true definition of a LSC is rather ambiguous. It was first shown by Dick *et al.* in mice with severe combined immunodeficiency (SCID) that the engraftment potential of leukemic human cells is much higher for CD34⁺CD38⁻ leukemic cells than CD34⁻ or CD34⁺CD38⁺ leukemic subsets¹⁶. Yet, as shown in **Chapter 4**, by performing single cell RNA-seq on human cord blood, CD34⁺CD38⁻ cells are still heterogeneous in HSC subsets, and this also holds for leukemic cells¹⁷. For instance, latest mice experiments have shown that knock-in of the *NPM1* hotspot mutation in more differentiated progenitor cell types (e.g. MPP, GMP) can also lead to engraftment and full-blown leukemia, albeit at a lower efficiency compared to earlier stem cell subsets¹⁸. These findings support that LSCs do not necessarily need to have a CD34⁺CD38⁻ phenotype, and that LSCs can also reside in AML cells with other phenotypes. Later, Dick *et al.* defined two gene signatures for LSCs commonly referred to as *stemness score*¹⁹. However, these signatures include *CD34* as one of the most important genes. In **Chapter 6**, we showed that *CD34* gene expression levels strongly correlate with stemness scores across the transcriptional clusters of *NPM1*-mutated AML²⁰, and that *NPM1* mutations are not confined to classical LSCs but detectable across all myeloid cell types. While this distribution does not imply uniform leukemia-initiating capacity, it suggests that leukemic potential is not exclusively restricted to immunophenotypically defined LSCs, but may also emerge from progenitors capable of retaining or re-acquiring stem-like features within a largely preserved differentiation hierarchy.

This principle of stemness plasticity is further exemplified by *NPM1*-mutated AML, where a single molecular lesion is sufficient to reprogram progenitors toward self-renewal. NPM1 mutations, most commonly a 4-bp TCTG insertion (type A) in exon 12, occur in ~30% of adult AML. The frameshift mutation disrupts a nuclear retention signal and creates a nuclear export signal, leading to mis-localization of the nucleolar chaperone to the cytoplasm. The mis-localization of NPM1 (NPM1c) represents both a defining biomarker and an unresolved mechanistic driver. Upon gaining the *NPM1* mutation, myeloid progenitor cells acquire long-term self-renewal. *NPM1*-mutated AML is associated with a distinctive transcriptional program characterized by high expression of *HOXA*, *HOXB*, and cofactors such as *MEIS1*²¹. Shuttling NPM1c back to the nucleus with XPO1 inhibitors extinguishes the *HOXA/MEIS1* program, drives maturation and extends survival in mouse models²².

Recent studies have further refined this mechanism by showing that NPM1c does not merely activate *HOX* genes indirectly, but can also directly regulate oncogenic transcription. In particular, a residual nuclear fraction of NPM1c has been shown to bind a restricted set of active chromatin regions, including the *HOXA* and *HOXB* clusters and *MEIS1*, where it cooperates with XPO1/CRM1, RNA polymerase II, and the MLL1–Menin transcriptional machinery to sustain the leukemic self-renewal program^{23,24}. These observations provide a more direct explanation for the characteristic *HOXA/HOXB*-high state of *NPM1*-mutated AML and also help explain the marked sensitivity of this leukemia subtype to Menin inhibition^{23,24}.

At the same time, indirect mechanisms likely contribute in parallel. Earlier studies showed that NPM1c-driven cytoplasmic mislocalization of key nuclear regulators such as PU.1 and CTCF perturbs differentiation-associated transcriptional control and chromatin architecture, thereby further reinforcing leukemic identity^{25,26}. Together, these findings indicate that NPM1c-mediated stemness is mechanistically multifactorial, reflecting both direct chromatin-level activation of leukemic transcriptional hubs and indirect rewiring of transcription factor and chromatin regulator localization.

Wild-type NPM1 transiently coats unduplicated centrosomes and is released after CDK2–cyclin E phosphorylates Thr199^{27,28}. In *NPM1*-mutated leukemic cells, the cytoplasmic mutant form exhibits elevated Thr199 phosphorylation and a heightened dependence on this modification for G₂/M progression²⁹.

While wild-type NPM1 participates in centrosome licensing, whether mislocalized NPM1c accumulates within centrosome-proximal compartments in primary blasts remains unclear; thus, any proposed centrosome-mediated link to *HOX* control should be viewed as a hypothesis that warrants testing. It would be interesting to address this hypothesis using advanced spatial proteomics (e.g. imaging mass cytometry) integrated with single-cell or spatial transcriptomics, to map centrosome-associated interactors and downstream kinases that could connect NPM1c export to leukemic identity. Taken together, these observations reinforce a central theme: leukemic potential is not limited to LSCs as initially thought but can also be conferred to myeloid progenitor cells other than HSCs through acquisition of mutations such as NPM1c that hijack stemness programs.

7.4 General Discussion and Future Perspective

7.4.1 Understanding the identity of a cell

A human cell is a product of 3.8 billion years of evolution. Arguably, the most complex machine in existence. Millions of proteins work seamlessly to establish intrinsic functions and accommodate to external conditions. A cell consists of different organic molecules, and particularly mRNAs and proteins play an essential role. As outlined in the **Introduction**, lineage identity is imposed not by changes in the DNA sequence but by reversible epigenetic programs such as DNA methylation, histone modifications and higher-order chromatin remodeling that modulate gene accessibility. This epigenetic regulatory layer has first been conceptualized by Waddington in 1942³⁰. Today, thousands of scientists around the world work on these epigenetic changes and try to further map the effects of these changes on cells. Perhaps the most striking example of these endeavors is reversing cell identity to stem cells by only 4 transcription factors³¹. This discovery propelled the epigenetics field, underscoring that cell identity can be reprogrammed by external cues, much like changing the software on fixed hardware. In other words, what cells do is “programmable” without altering the underlying genome. For instance, in **Chapter 3**, we used two different epigenetic modifiers to enhance the yield of HSCs when co-cultured with stem cell factor. Particularly Quisinostat significantly increased the HSC pool without interfering with the ability to transduce these cells with a lentivirus.

Building on this concept of programmability, totipotency represents the ultimate plasticity. It is a uniquely defining feature of cells in an early developing multi-cellular organism, which allows a single cell, the zygote, to create all the types of cells making a fully grown organism. In his classical view, Waddington used a landscape analogy to describe cell fates during development. He used the example of a marble rolling down from a hill. As it rolls down, the marble encounters valleys, which represent different developmental pathways. Once the marble enters a valley, it becomes more committed to a certain cell fate and the way back to a less differentiated state is much harder. This analogy captures the broad intuition, but modern single-cell data suggest refinements are needed. For instance, transcriptional entropy of a cell is correlated with its maturation stage³². High entropy, characterized by many genes that are expressed at similar levels, is associated with more stem cell phenotypes. In contrast, low entropy, characterized by a few genes that dominate expression, is associated with more differentiated cell phenotypes. This is similar to each fork taken by the rolling-marble in a valley which reduces the uncertainty of the outcome. However, rather than a simple continuum of stable fates, differentiation is better described as an attractor landscape, in which gene-regulatory dynamics stabilize discrete cell identities and intermediate states are usually transient. Along trajectories, cells can pass through intermediate configurations, but these are typically transient and low probability. Consistent with this, many surface markers, and their mRNA proxies, exhibit near on/off distributions created by feedback, cooperativity, and thresholds; stable intermediate expression is rare, usually confined to transient states at bifurcations.

Single-cell studies further show that differentiation often proceeds through groups of closely related cell states that form recognizable clusters. Within each cluster, cells have similar phenotypes, illustrating the top-down dynamics that happen in cellular differentiation. Beyond transcriptional and epigenetic states, physical and bioelectric properties also constrain differentiation trajectories. For instance, the cell surface is a thin layer with almost constant capacitance, but varying resistance. Interestingly, resting membrane potential (V_M) also correlates with differentiation³³, with depolarized cells (i.e. chondrocytes) being more prone to cellular division and hyperpolarized cells being less likely to undergo cell division, e.g. neurons, or skeletal muscle cells³⁴. A strik-

ing example of these bioelectric signals affecting cellular differentiation was shown by Levin et al. in planaria. These worms can be cut many times and still regenerate their entire body. Yet remarkably, altering their endogenous bioelectric signals can change the pattern of regeneration. Levin and colleagues demonstrated that by artificially modulating transmembrane voltage gradients, they could control the head size of planaria and could induce regeneration of multiple heads or, conversely, fail to regenerate a head at all^{35,36}. These findings support the view that bioelectric signals serve as a type of pattern memory, guiding cells toward specific morphological outcomes. In other words, the top-down influence of bioelectric states can override default molecular cues, reshaping an organism's anatomy despite a constant genomic blueprint. In the context of hematopoiesis and leukemogenesis, these perspectives suggest that stemness, lineage bias, and even clonal architecture may be governed not only by genetic and epigenetic alterations but also by emergent top-down constraints, from attractor dynamics to bioelectric patterning.

7.4.2 Towards a system-level understanding of biology

Hematopoiesis is a complex system that requires strict regulation. It is a dynamic process that changes throughout the lifespan of an animal. The disruption of this top-down process during hematopoiesis is most prominent in cancer settings. For instance, in **Chapter 4**, we show that AML with distinct morphology by microscopy are dominated by distinct cell types identified from molecular data (RNA-seq), indicating that internal changes in the leukemic blasts are reflected at system scale which can be seen with a microscope. Throughout the thesis, we showed that hematopoiesis can be deregulated and skewed towards the myeloid lineage as a consequence of aging or genetic abnormalities. This bias can be simply explained by competition between different myeloid- and lymphoid-skewed HSC clones. Clones with an evolutionary advantage (e.g. faster division rate) repopulate the hematopoietic system with more exponentially growing cells. This is supported by recent studies in humans with decades of genomics data^{37,38}. Clones with a growth advantage dominate the hematopoietic system and as a consequence, clonal diversity decreases significantly³⁹. These findings promote important questions in the field of systems immunology. As we age, the thymus involutes and naive T-cell output declines, while clonal hematopoiesis of indeterminate potential (CHIP) mutations, most commonly in *TET2* or *DNMT3A*, become more prevalent. *TET2*-mutant clones consistently show myeloid-biased differentiation in both humans and mice, whereas *DNMT3A* mutations primarily expand multipotent HSCs, with myeloid skewing occurring more specifically with hotspot mutations at sites such as R882. This shift toward myeloid output reduces the pool of common lymphoid progenitors⁸ (CLPs) and which in return may limit the number of HSCs available to seed the thymus. An open question is whether such CHIP-associated changes directly contribute to progressive thymic shrinkage during adulthood.

Another important question is how hematopoietic alterations contribute to other cancer types. For example, a recent paper revealed that in blastic plasmacytoid dendritic cell neoplasm (BPDCN), a rare but aggressive malignancy that often presents as skin lesions, the initial cells may originate from the bone marrow harboring only a CHIP mutation. Subsequently, these cells could acquire additional genetic mutations upon exposure to ultraviolet light in the skin⁴⁰. Another recent study revealed that interleukin-4 (IL-4) signaling within the bone marrow promotes the development of immunosuppressive myeloid cells, which facilitate tumor progression in non-small cell lung cancer (NSCLC)⁴¹. Both studies demonstrate that only a systems-level perspective can capture how hematopoietic alterations drive disease processes beyond the primary tissue, and how therapeutic strategies must therefore target the entire immune network.

This systemic understanding is important not only for cancers but also for other lethal diseases. For instance, the “Berlin Patient,” was the first individual who was accidentally cured of HIV after allogeneic hematopoietic stem cell transplantation (HSCT) for his AML. As of late 2024, seven patients have been cured from HIV by allogeneic HSCT⁴². These cases highlight the potential of bone marrow-derived interventions in treating even viral infections. Altogether, these examples highlight that hematopoiesis functions as a central hub influencing diverse systemic diseases. Understanding these interactions requires a systems-immunology perspective that integrates clonal dynamics, immune regulation, and organism-wide consequences.

7.4.3 What can we do for a true system understanding?

“Math is to physics what AI will be to biology.”

– Demis Hassabis

Although criticized by many experts, reductionist ideas can converge to reality. A great example to this is neural networks. Inspired by the neurons in the brain, a layer of simple perceptron can approximate any function⁴³. Cascades of these simple models are capable of understanding language, visual and audio clues and even creating new simulated worlds^{44,45}. Similarly, a cell can be seen as a “bag of information”, where the information can represent any molecule of the hardware such as DNA, RNA, or protein. With the digitization of the cells’ hardware, we can create computational models that are capable of understanding the statistical relations between these molecules and render the actual *algorithm*. Simulating molecular interactions with AI and creating a virtual cell to understand cell-cell interactions and eventually understanding the system seems to be the direction we are heading. One early example of these models is called Evo⁴⁶, where the aim is to predict the next base in a given bacterial genome via a hybrid transformer architecture. For instance, this foundation model is capable of predicting essentiality of the genes in knock-out experiments. Still, Evo is trained only on DNA sequence, which captures part of the biological “vocabulary” but not the full regulatory language of the cell.

Understanding how one cell reacts to certain stimuli will not be sufficient for system level understanding of multi-cellular life. To solve the language of cells, multi-omics single-cell data will be truly essential. We know that the molecular pathways are simply the means (or ‘sentences’) of the language, yet they are mostly undiscovered. These pathways are not yet truly understood at single cell level, mostly due to technical difficulties such as drop-out, sequencing biases or not having truly enormous single-cell multi-omics experiments. In the coming decade, a major goal will be to use artificial intelligence to understand cellular regulation from a systems-level perspective. Along the same trajectory, the same authors introduced Evo2⁴⁷, which was trained on a vastly larger corpus of genomes, spanning bacteria to apes and humans, and extended to RNA data. Evo2 can operate at the scale of entire chromosomes and has already been used not only for variant effect prediction but also for *de novo* generative design. Strikingly, a recent study demonstrated that Evo2-generated bacteriophage genomes, despite harboring hundreds of novel mutations, could assemble into fully functional infectious phages, some even outperforming wild-type strains in competition assays⁴⁸. This proof-of-concept shows that genome language models are not restricted to passive prediction but can actively design novel biological entities, pushing reductionist “bags of information” toward system-level engineering.

An analogy of infamous System I and II thinking⁴⁹ from Daniel Kahneman can also be applied to cells. A single cell is mostly reactive and does not have the capability of long-term thinking and planning and therefore remains at the System I level. However, multi-cellular or complex

Concluding Remarks

organisms can plan and act ahead, which suggests that multi-cellularity and top-down regulation is essential for System II level. Therefore, integration of single-cell insights with top-down, system-level approaches is critical for deciphering how multi-cellular organisms acquire System II-like abilities. We are at a point in time where we can control System I level of cells through molecular or cellular manipulations (Figure 7.2). For instance, we can edit the genome via gene therapy or CRISPR, or we can change the epigenetic landscape of cells and change their identity back to their stem cell state. Still, we are not fully at the level of system manipulation, although the first glimpses are already visible, e.g. designing organoids or rewiring electrical signals of multi-cellular organisms. By combining comprehensive multi-omics data with artificial intelligence and large-scale modeling, we can begin to bridge the gap between the reactive nature of individual cells and the emergent, forward-planning capabilities of entire tissues and organisms. This will be the challenge of this century.

Lastly, System III level is a new phenomenon that represents the next step in the evolutionary timeline, with the potential to create entirely new biological entities. While System II encompasses the ability to rewire existing multicellular organisms at the tissue or organ level and balancing back the homeostasis, System III goes one step further by integrating breakthroughs in synthetic biology, regenerative medicine, computational modeling, perhaps robotics, and AI to sculpt novel or profoundly re-engineered forms of life. Recent work with Evo2 demonstrates that genome models can already generate functional bacteriophages from scratch, hinting that System III is no longer theoretical but also an emerging reality. This entails orchestrating development, physiology, and even behavior in ways that fundamentally transcend current natural boundaries. For example, combining whole-organ engineering, dynamic bioelectrical patterning, and self-organizing organoids with advanced computational designs could allow us not only to repair complex tissues, but also to produce emergent biological systems.

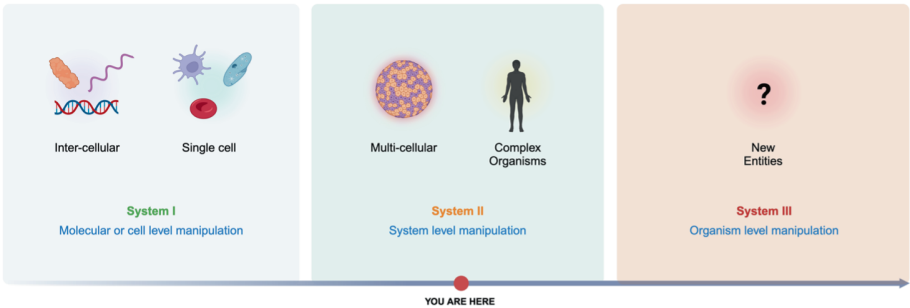


Figure 7.2: A simple schematic of manipulation levels.

7.5 Concluding Remarks

In conclusion, this dissertation has provided a comprehensive exploration of immune aging and hematopoiesis, revealing critical intersections that predispose to leukemogenesis in acute myeloid leukemia (AML). By integrating transcriptional profiling, chromatin accessibility analyses, and advanced computational techniques across murine and human models, this work helps our understanding of inflammaging through the potential of epigenetic interventions to rejuvenate

hematopoietic stem cell function. Moreover, the detailed investigation into the novel subtypes of AML using bulk and single-cell methodologies underscores the intrinsic challenges posed by methodological and contextual ambiguities, while offering new insights into the dynamic interplay between genetic aberrations and cellular identity.

In the last century, we began to understand how the hematopoietic system functions and how blood cells arise, operate, and sometimes fail. We are now recognizing how age-related changes in the hematopoietic system can drive skewing toward the myeloid lineage, thereby increasing the risk of cancer. Today, we are on the verge of understanding blood cells and the hematopoietic system from a more holistic perspective. These new insights, together with the ingenious tools that have been developed, may one day enable the full eradication of lethal blood cancers and perhaps eventually other cancer types as well. Once among the deadliest subtypes of AML, acute promyelocytic leukemia can now be effectively treated with differentiation-inducing drugs, and immunotherapy has revolutionized the treatment of B-cell lymphomas, with newer forms such as TCR-T cell therapy beginning to find their way into AML treatment.

Looking ahead, the future of hematology and cancer research is bright. Global collaborations, relentless scientific rigor, and continuous innovation will reshape our ability to understand the mechanisms underlying blood cancers. Curative treatments are no longer just a dream but becoming a real possibility. We are quickly entering an era where what once seemed impossible is turning into reality.

7.6 References

1. Auffray, C., Sieweke, M. H. & Geissmann, F. Blood Monocytes: Development, Heterogeneity, and Relationship with Dendritic Cells. *Annu. Rev. Immunol.* **27**, 669–692 (2009).
2. Dash, S. P., Chakraborty, P. & Sarangi, P. P. Inflammatory Monocytes and Subsets of Macrophages with Distinct Surface Phenotype Correlate with Specific Integrin Expression Profile during Murine Sepsis. *The Journal of Immunology* **207**, 2841–2855 (2021).
3. Shi, C. & Pamer, E. G. Monocyte recruitment during infection and inflammation. *Nat Rev Immunol* **11**, 762–774 (2011).
4. Ross, J. B. *et al.* Depleting myeloid-biased haematopoietic stem cells rejuvenates aged immunity. *Nature* **628**, 162–170 (2024).
5. Genovese, G. *et al.* Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N Engl J Med* **371**, 2477–2487 (2014).
6. Jaiswal, S. *et al.* Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *N Engl J Med* **371**, 2488–2498 (2014).
7. Van Zeveren, I. A. *et al.* Prevalence, predictors, and outcomes of clonal hematopoiesis in individuals aged ≥80 years. *Blood Advances* **5**, 2115–2122 (2021).
8. Nam, A. S. *et al.* Single-cell multi-omics of human clonal hematopoiesis reveals that DNMT3A R882 mutations perturb early progenitor states through selective hypomethylation. *Nat Genet* **54**, 1514–1526 (2022).
9. 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**, 380 (2019).
10. Avila Cobos, F., Alquicira-Hernandez, J., Powell, J. E., Mestdagh, P. & De Preter, K. Benchmarking of cell type deconvolution pipelines for transcriptomics data. *Nat Commun* **11**, 5650 (2020).
11. Arindrarto, W. *et al.* Comprehensive diagnostics of acute myeloid leukemia by whole transcriptome RNA sequencing. *Leukemia* **35**, 47–61 (2021).
12. Zeng, A. G. X. *et al.* A cellular hierarchy framework for understanding heterogeneity and predicting drug response in acute myeloid leukemia. *Nat Med* **28**, 1212–1223 (2022).
13. Newman, A. M. *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* **37**, 773–782 (2019).
14. Liu, Y., Beyer, A. & Aebersold, R. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**, 535–550 (2016).

References

15. Darmanis, S. *et al.* Simultaneous Multiplexed Measurement of RNA and Proteins in Single Cells. *Cell Reports* **14**, 380–389 (2016).
16. Lapidot, T. *et al.* A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367**, 645–648 (1994).
17. Tajer, P. *et al.* Utilizing epigenetic regulators to improve HSC-based lentiviral gene therapy. *Blood Advances* **8**, 4936–4947 (2024).
18. SanMiguel, J. M. *et al.* Cell origin–dependent cooperativity of mutant *Dnmt3a* and *Npm1* in clonal hematopoiesis and myeloid malignancy. *Blood Advances* **6**, 3666–3677 (2022).
19. Ng, S. W. K. *et al.* A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature* **540**, 433–437 (2016).
20. Karakaslar, E. O. *et al.* Resolving inter- and intra-patient heterogeneity in *NPM1*-mutated AML at single-cell resolution. Preprint at <https://doi.org/10.1101/2024.12.10.24317471> (2024).
21. Verhaak, R. G. W. *et al.* Mutations in nucleophosmin (*NPM1*) in acute myeloid leukemia (AML): association with other gene abnormalities and previously established gene expression signatures and their favorable prognostic significance. *Blood* **106**, 3747–3754 (2005).
22. Brunetti, L. *et al.* Mutant *NPM1* maintains the leukemic state through *HOX* expression. *Cancer Cell* **34**, 499–512.e9 (2018).
23. Uckelmann, H. J. *et al.* Mutant *NPM1* directly regulates oncogenic transcription in acute myeloid leukemia. *Cancer Discov.* **13**, 746–765 (2023).
24. Wang, X. Q. D. *et al.* Mutant *NPM1* hijacks transcriptional hubs to maintain pathogenic gene programs in acute myeloid leukemia. *Cancer Discov.* **13**, 724–745 (2023).
25. Gu, X. *et al.* Leukemogenic nucleophosmin mutation disrupts the transcription factor hub that regulates granulomonocytic fates. *J. Clin. Invest.* **128**, 4260–4279 (2018).
26. Wang, A. J. *et al.* *NPM1c* impedes CTCF functions through cytoplasmic mislocalization in acute myeloid leukemia. *Leukemia* **34**, 1278–1290 (2020).
27. Okuda, M. *et al.* Nucleophosmin/B23 is a target of CDK2/cyclin E in centrosome duplication. *Cell* **103**, 127–140 (2000).
28. Tokuyama, Y., Horn, H. F., Kawamura, K., Tarapore, P. & Fukasawa, K. Specific phosphorylation of nucleophosmin on Thr199 by cyclin-dependent kinase 2–cyclin E and its role in centrosome duplication. *J. Biol. Chem.* **276**, 21529–21537 (2001).
29. Chan, N. & Lim, T. M. Cytoplasmic nucleophosmin has elevated T199 phosphorylation upon which G₂/M phase progression is dependent. *Sci. Rep.* **5**, 11777 (2015).
30. Waddington, C. H. The Epigenotype. *Int. J. Epidemiol.* **41**, 10–13 (2012).
31. Takahashi, K. & Yamanaka, K. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, 663–676 (2006).
32. Teschendorff, A. E. & Enver, T. Single-cell entropy for accurate estimation of differentiation potency from a cell’s transcriptome. *Nat Commun* **8**, 15599 (2017).
33. Cone, Jr., C. D. Variation of the Transmembrane Potential Level as a Basic Mechanism of Mitosis Control. *Oncology* **24**, 438–470 (1970).
34. Yang, M. & Brackenbury, W. J. Membrane potential and cancer progression. *Front. Physiol.* **4**, (2013).
35. Beane, W. S., Morokuma, J., Lemire, J. M. & Levin, M. Bioelectric signaling regulates head and organ size during planarian regeneration. *Development* **140**, 313–322 (2013).
36. Durant, F. *et al.* Long-Term, Stochastic Editing of Regenerative Anatomy via Targeting Endogenous Bioelectric Gradients. *Biophysical Journal* **112**, 2231–2243 (2017).
37. Fabre, M. A. *et al.* The longitudinal dynamics and natural history of clonal haematopoiesis. *Nature* **606**, 335–342 (2022).
38. Mitchell, E. *et al.* Clonal dynamics of haematopoiesis across the human lifespan. *Nature* **606**, 343–350 (2022).
39. Weng, C. *et al.* Deciphering cell states and genealogies of human haematopoiesis. *Nature* **627**, 389–398 (2024).
40. Griffin, G. K. *et al.* Ultraviolet radiation shapes dendritic cell leukaemia transformation in the skin. *Nature* **618**, 834–841 (2023).
41. LaMarche, N. M. *et al.* An IL-4 signalling axis in bone marrow drives pro-tumorigenic myelopoiesis. *Nature* **625**, 166–174 (2024).
42. Mallapaty, S. Seventh patient ‘cured’ of HIV: why scientists are excited. *Nature* **632**, 235–236 (2024).
43. Hornik, K., Stinchcombe, M. & White, H. Multilayer feedforward networks are universal approximators. *Neural Networks* **2**, 359–366 (1989).
44. OpenAI *et al.* GPT-4 Technical Report. Preprint at <https://doi.org/10.48550/ARXIV.2303.08774> (2023).

45. Parker-Holder, J. *et al.* Genie 2: A Large-Scale Foundation World Model. (2024).
46. Nguyen, E. *et al.* Sequence modeling and design from molecular to genome scale with Evo. *Science* **386**, eado9336 (2024).
47. Brixi, G. *et al.* Genome modeling and design across all domains of life with Evo 2. Preprint at <https://doi.org/10.1101/2025.02.18.638918> (2025).
48. King, S. H. *et al.* Generative design of novel bacteriophages with genome language models. Preprint at <https://doi.org/10.1101/2025.09.12.675911> (2025).
49. Kahneman, D. *Thinking, Fast and Slow*. (macmillan, 2011).