Comprehensive Classification of Pioneer Transcription Factors:

Insights into Binding Characteristics and Chromatin Interactions

by Sofiia Shapoval

A Thesis submitted to the School of Graduate Studies in partial fulfillment of the

requirements for the degree of Master of Science in Medicine (Human Genetics)

Division of Biomedical Sciences, Faculty of Medicine

Memorial University of Newfoundland

October, 2024

St. John's, Newfoundland and Labrador

Abstract

This study focuses on classifying pioneer transcription factors (pTFs) and understanding their crucial roles in gene regulation across various biological processes like development, differentiation, and responses to environmental cues. Insight into pTFs is valuable for therapeutic interventions, particularly in regenerative medicine, disease treatment, and tissue engineering. They are essential for unpacking chromatin, enabling access to target genes and reprogramming cells into induced pluripotent stem cells (iPSCs), impacting stem cell research and medical applications. Additionally, pTFs are key in cellular responses to environmental changes, especially in cancer biology and immunology, making them significant in industrial and medical contexts. Dysregulation of pTFs can lead to diseases like cancer, underscoring the importance of their classification in identifying potential therapeutic targets.

The study outlines methodologies for comprehensive data collection, feature grouping, and categorization related to DNA binding, chromatin accessibility, cell fate alterations, and cancer roles. Results encompass pTFs classification, gene ontology enrichment analysis, pathway analysis, and gene expression analysis under baseline and differential conditions. The classification identified distinct groups of pTFs based on their binding preferences and interactions with chromatin. This detailed understanding highlights the maintenance of cellular identity, differentiation mechanisms, and potential therapeutic interventions targeting specific pTFs. Additionally, the study performed Gene Ontology (GO) enrichment analysis and pathway analysis, offering insights into the functional implications of different pTF groups across cellular components, biological processes, and molecular functions. These analyses enhance our understanding of the roles played by the newly defined groups of pTFs in various cellular processes and regulatory mechanisms. Additionally, gene expression analysis categorized TFs based on their baseline and differential expression patterns across different tissues and processes, revealing distinct functional peculiarities within TF families. This highlights the

importance of integrating expression data and functional characteristics for accurate classification and understanding of pTFs' regulatory impact. In conclusion, this comprehensive research significantly advances our understanding of pTFs, unraveling their diverse roles, regulatory mechanisms, and implications across various cellular processes. The findings presented offer a multi-faceted view of pTFs, enriching our knowledge in the field of transcriptional regulation and paving the way for further studies to deepen our understanding of their molecular mechanisms and functional implications.

General Summary

In this research, we delve into the world of pioneer transcription factors (pTFs), which are essential players in controlling how genes are expressed in living organisms, affecting everything from growth to how cells respond to their environment. We've worked to classify these TFs comprehensively, shedding light on their crucial roles in kickstarting cell specialization and maintaining cellular identity, particularly during embryonic development and tissue repair. Understanding these TFs could revolutionize regenerative medicine, allowing precise manipulation of cell behavior for therapeutic purposes. These TFs also help unwind tightly wound DNA, making genes accessible for regulation, influencing everything from stem cell research to cancer treatment. By categorizing these TFs systematically, we've laid out methods for gathering and analyzing data, revealing insights into their functions and potential roles in diseases like cancer. This research not only advances scientific understanding but also offers promising avenues for medical and industrial applications.

Co-authorship statement

I, Sofiia Shapoval, acknowledge my third author status for the version of the manuscript incorporated in Sections 2-3 of this thesis, which includes the newly created pioneer transcription factors database. This manuscript was collaboratively authored alongside my supervisor and colleagues.

Dyer M, Lin QXX, Shapoval S, Thieffry D, Benoukraf T. MethMotif.Org 2024: a database integrating context-specific transcription factor-binding motifs with DNA methylation patterns. Nucleic Acids Res. 2024 Jan 5;52(D1):D222-D228. doi: 10.1093/nar/gkad894. PMID: 37850642; PMCID: PMC10767921.

Authors Contributions: M.J.D., Q.X.X.L. and T.B. designed the platform; M.J.D. and Q.X.X.L. programmed the website and performed all analysis; M.J.D., Q.X.X.L., S.S., D.T. and T.B. interpreted the data; M.J.D., S.S., D.T., and T.B. wrote the manuscript. T.B. directed the study.

Acknowledgments

I would like to express my deepest gratitude to several individuals who played pivotal roles in the completion of this thesis, without whom this journey would not have been possible.

First and foremost, I extend my sincere appreciation to my supervisor, Dr. Benoukraf, for his unwavering guidance, invaluable insights, and continuous support throughout the entire research process. His expertise and encouragement have been instrumental in shaping the trajectory of my work.

I am also indebted to my co-supervisors, Dr. Woods and Dr. Esseltine, for their collaborative efforts, constructive feedback, and scholarly mentorship. Their collective wisdom significantly enriched the quality of this thesis.

I would like to acknowledge the assistance and guidance provided by Dr. Akinola Aalafiyatao in the lab. His expertise and willingness to share knowledge were invaluable, and his supervision greatly enhanced the experimental aspects of my research.

Special recognition is due to my classmates, Nafiza Hague and Alecia Robichaud, whose support and collaborative efforts in navigating challenging coursework played a crucial role in my academic journey.

I express my gratitude for the generous funding provided by the Faculty of Medicine and the School of Graduate Studies at Memorial University of Newfoundland for supporting my master's thesis research.

A heartfelt thank you is extended to Asmaa Mahmoud for her constant moral support and infectious positivity, which brought joy and cheerfulness to my life, especially during challenging times.

Lastly, my deepest gratitude goes to my family and friends for their unwavering love and support. Despite the hardships stemming from the war in my home country, Ukraine, they provided a foundation of strength and encouragement that sustained me throughout this academic endeavor. Their resilience and unconditional support have been a beacon of light in my journey.

To everyone mentioned and countless others who may not be named, your contributions have left an indelible mark on this thesis, and for that, I am sincerely grateful.

Glory to Ukraine

Table of Contents

Figures

Tables

Abbreviations

5mC	5-Methylcytosine
AML	Acute Myeloid Leukemia
AR	Androgen Receptor
ARMS	Alveolar Rhabdomyosarcoma
Ascl1	Achaete-scute homolog 1
ATF	Activating Transcription Factor
ATP	Adenosine Triphosphate
bHLH	Basic Helix–Loop–Helix
BIRC5	Baculoviral IAP Repeat Containing 5
Brn3b	Brain-specific homeobox/POU domain protein 3B
bZIP	Basic Leucine-Zipper
Cbfβ	Core Binding Factor Beta
CDKs	Cyclin-Dependent Kinases
CEBP	CCAAT/enhancer-binding protein
CGI	CpG Island
CHD	Chromodomain Helicase DNA-binding
Chx10	Ceh-10 homeodomain-containing homolog
CLL	Chronic Lymphocytic Leukemia
CRC	Colorectal Cancer
CREB	cAMP Response Element-Binding Protein
Crx	Cone-rod homeobox
CSC	Cancer Stem Cell
CTCF	CCCTC-binding Factor
DCT	Dopachrome Tautomerase
DLBCL	Diffuse Large B-Cell Lymphoma
DMR	Differentially Methylated Region

DMRT	Doublesex and Mab-3-Related Transcription Factor
DNA	Deoxyribonucleic Acid
DNMT	DNA Methyltransferase
DS-AMKL	Down Syndrome-Related Acute Megakaryoblastic Leukemia
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin ligase
EMT	Epithelial-Mesenchymal Transition
ER	Estrogen Receptor
ESCs	Embryonic Stem Cells
FHRE	Forkhead Response Element
FOX	Forkhead Box
GATA	GATA Binding Protein
GREs	Glucocorticoid Response Elements
GSC	Glioblastoma Stem Cell
HAT	Histone Acetyltransferase
HDAC	Histone Deacetylase
HIF-1α	Hypoxia-Inducible Factor 1-alpha
HIFs	Hypoxia-Inducible Factors
HMGA2	High Mobility Group A2
HMT	Histone Methyltransferase
HNF4α	Hepatocyte Nuclear Factor 4 Alpha
HREs	Hypoxia Response Elements
HSPs	Heat Shock Proteins
HTH	Helix-Turn-Helix
IAPs	Inhibitors of Apoptosis
ICRs	Imprinting Control Regions

Ig Fold	Immunoglobulin-Like Fold
IP3	Inositol Triphosphate
iPSC	Induced Pluripotent Stem Cell
KLF	Krüppel-Like Factor
LBDs	Ligand-Binding Domains
LHX8	LIM homeobox 8
LSD1	Lysine-Specific Demethylase 1
MBD	Methyl Binding Domain
MeDIP	Methylated DNA Immunoprecipitation
Mef2	Myocyte Enhancer Factor 2
MITF	Microphthalmia-Associated Transcription Factor
MMP	Matrix Metalloproteinase
MUC2	Mucin 2
MWNT	Multiwalled Carbon Nanotube
Myf5	Myogenic Factor 5
MyoD	Myogenic Differentiation 1
Nanog	Homeobox Protein Nanog
NES	Nuclear Export Signals
NFIA	Nuclear Factor I A
NF-κB	Nuclear Factor-kappa B
Ngn1/Ngn2	Neurogenin 1/Neurogenin 2
NHRs	Nuclear Hormone Receptors
NLS	Nuclear Localization Signals
Olig2	Oligodendrocyte transcription factor 2
Osterix	Sp7 transcription factor
Otx2	Orthodenticle homeobox 2
Pax	Paired Box Transcription Factor
PDX1	Pancreatic and Duodenal Homeobox 1

PGCs	Primordial Germ Cells
PML	Promyelocytic Leukemia Protein
PNET	Primitive Neuroectodermal Tumor
PO4	Phosphate group
PPARγ	Peroxisome Proliferator-Activated Receptor Gamma
PPREs	PPAR Response Elements
PPyox	Overoxidized Polypyrrole
PRC	Polycomb Repressive Complex
PRC1	Polycomb Repressive Complex 1
PSA	Prostate-Specific Antigen
pTF	Pioneer Transcription Factor
qPCR	Quantitative Polymerase Chain Reaction
RAR	Retinoic Acid Receptor
RCC	Renal Cell Carcinoma
RHH	Ribbon-Helix-Helix
RNA	Ribonucleic Acid
RNAP II	RNA polymerase II
ROS	Reactive Oxygen Species
RPE	Retinal Pigment Epithelium
Runx	Runt-related transcription factor
Rx	Retinal homeobox protein
RXR	Retinoid X Receptor
SATB1	Special AT-rich Sequence-Binding Protein 1
SENPs	Sentrin-specific proteases
Six3	Sine oculis homeobox 3
SMCHD1	Structural Maintenance of Chromosomes Hinge Domain Containing 1
SOX	SRY-Box Transcription Factor

SRF	Serum Response Factor
SS18-SSX	Synovial Sarcoma Fusion Proteins
STAT	Signal Transducer and Activator of Transcription
SUMO	Small Ubiquitin-like Modifier
SWI/SNF	SWItch/Sucrose Non-Fermentable
T-ALL	T-Cell Acute Lymphoblastic Leukemia
TF	Transcription Factor
TFIID	Transcription Factor IID
TFIIE	Transcription Factor IIE
TFIIH	Transcription Factor IIH
THR	Thyroid Hormone Receptor
TMPRSS2	Transmembrane Protease Serine 2
TSS	Transcription Start Site
VEGF	Vascular Endothelial Growth Factor
VHL	von Hippel-Lindau
WGBS	Whole-Genome Bisulfite Sequencing
wHTH	Winged Helix-Turn-Helix
XCI	X-Chromosome Inactivation
XIST	X-Inactive Specific Transcript
YY1	Yin Yang 1
Zfp423	Zinc Finger Protein 423
Zn2+	Zinc Ion

I. Introduction

1.1 Historical development of the concept of transcriptional factors

The concept of TFs began with the discovery that DNA is the genetic material responsible for carrying genetic information. This foundational work was initiated by Friedrich Miescher in 1869 (Dahm 2005) and culminated in the elucidation of the DNA double helix structure by James Watson and Francis Crick in 1953 (Watson and Crick 1953). In the 1950s, RNA polymerase was identified as the enzyme responsible for transcribing DNA into RNA (Hurwitz 2005). This discovery laid the groundwork for understanding the process of transcription. In 1961 François Jacob and Jacques Monod proposed the operon model, which described the regulation of gene expression in bacteria (Voorhees et al. 1976). They introduced the concept of regulatory proteins (later identified as transcription factors) that control the transcription of genes by binding to specific DNA sequences. Later researchers Robert Roeder and Richard Tjian began identifying transcriptional activator proteins (Roeder 2005; Falender et al. 2005). These proteins enhance gene transcription by binding to specific DNA sequences known as enhancers or promoter-proximal elements. The first eukaryotic transcription factor to be cloned, the yeast protein Gal4, was shown to activate transcription of genes involved in galactose metabolism (Johnston and Hopper 1982). In the 1980s, the development of techniques like DNA footprinting allowed to identify the precise binding sites of transcription factors on DNA (Schmitz and Galas 1980). During that era, transcription factors were frequently categorized into families based on their shared DNA-binding domains, i.e. helix-turn-helix, zinc finger, and basic leucine zipper (bZIP) domains (Blumberg et al. 1987; Schuh et al. 1986; Kouzarides and Ziff 1989). After the 2000s, advances in genomics and high-throughput sequencing technologies have enabled the comprehensive identification of TF binding sites and the mapping of entire regulatory networks in various organisms (Bartlett et al. 2017; L. A. Liu and

Bader 2009). Computational methods and bioinformatics tools have become increasingly important for predicting TF binding sites and understanding their roles in gene regulation (Tognon, Giugno, and Pinello 2023; T. Huang et al. 2022). In the late 1980s, TRANSFAC emerged as a pioneering bioinformatics database for TF binding sites, utilizing manual curation and computational methods for motif analysis (Matys et al. 2003). The MEME Suite, originating in the mid-1990s, expanded into a comprehensive motif analysis toolkit, reflecting broader trends in bioinformatics towards algorithm development and software integration (Bailey et al. 2015). JASPAR, introduced in the early 2000s, provided an alternative with curated binding profiles derived from high-throughput experiments, aligning with the rise of next-generation sequencing (Castro-Mondragon et al. 2022). Meanwhile, ChIP-Seq analysis tools like MACS and SICER, developed in the mid-2000s, standardized ChIP-seq data analysis, enabling genome-wide studies of TF binding dynamics (Zhang et al. 2008; Zang et al. 2009). These tools exemplify the evolution of computational methods in TF research, driven by advancements in experimental techniques and collaborative efforts within the scientific community. The concept of TFs has evolved significantly over time, from early observations of gene regulation in bacteria to the detailed molecular understanding of eukaryotic gene regulation. Today, TFs play a central role in our understanding of how genes are turned on or off in response to various cellular signals and environmental cues, and they are key players in the field of molecular biology and genetics.

Table 1. Evolution of Transcription Factors: Key Milestones, Contributors, and Concepts

Milestone	Year	Key Contributors	Key Concepts
Discovery of DNA as genetic material	1869	Friedrich Miescher	DNA carries genetic information
DNA double helix structure elucidation	1953	James Watson and Francis Crick	DNA structure revealed
Identification of RNA polymerase	1950s	Hurwitz	RNA synthesis from DNA
Proposal of the operon model	1961	François Jacob and Jacques Monod	Introduction of regulatory proteins (transcription factors) in gene expression
Identification of transcriptional activator proteins	1970s	Robert Roeder and Richard Tjian	Transcriptional activators enhance gene transcription
Cloning of the first eukaryotic transcription factor (Gal4)	1982	Johnston and Hopper	Gal4 activates genes in galactose metabolism
Development of DNA footprinting techniques	1980s	Schmitz and Galas	Precise identification of TF binding sites
Categorization of TFs into families	1980s	Blumberg, Schuh, Kouzarides, Ziff	Classification based on DNA-binding domains
Advances in genomics and high-throughput sequencing technologies	2000s	Bartlett, Liu, Bader	Comprehensive identification of TF binding sites and regulatory networks
Rise of computational methods and bioinformatics tools	Post-2000s	Tognon, Huang	Prediction of TF binding sites and understanding their roles
Evolution of the concept of TFs	Ongoing	N/A	Central role in gene regulation, responsive to cellular signals and environmental cues

1.2 The role of transcriptional factors in gene expression

A defining characteristic of TFs lies in their capacity to discern and selectively bind to specific DNA sequences, called TF binding sites (TFBSs) (Sela and Lukatsky 2011). This unique ability

3

empowers them to orchestrate the dynamic modulation of chromatin structure, thereby exerting precise control over transcriptional processes. This control is achieved through their adeptness at recruiting a constellation of additional proteins and facilitating their assembly into intricate complexes (Aref, Sanad, and Schüller 2021; Soni et al. 2014), to exert context-specific control over gene expression, making TFs central players in the regulation of cellular processes (Francois, Donovan, and Fontaine 2020).

TFs are defined by several key features, including:

1. DNA Binding Specificity: TFs possess the ability to recognize and bind to specific DNA sequences, typically located in the regulatory regions of genes. This specificity enables them to target particular genes for regulation (Bouhlel, Lambert, and David-Cordonnier 2015; Sönmezer et al. 2021; Damante et al. 1994).

2. Modulation of Transcription: TFs can act as activators or repressors of gene transcription. Activator TFs enhance gene expression by promoting the recruitment of the transcriptional machinery, while repressor TFs inhibit gene expression by preventing transcriptional initiation (Willy, Kobayashi, and Kadonaga 2000; Fukaya 2023; Frejtag et al. 2001; Stephanou et al. 1999).

3. Protein-Protein Interactions: TFs often interact with other proteins to form complexes that influence gene expression. These interactions can involve co-activators or co-repressors, which further modulate transcription (Francois, Donovan, and Fontaine 2020; Johnsen et al. 1996; Liang and Hai 1997; Adams, Chandru, and Cowley 2018).

4. Context-Specific Function: TFs often exert their regulatory effects in a context-dependent manner. They may activate one set of genes in a particular cellular condition while repressing others, allowing for precise control of gene expression (Marchal, Defossez, and Miotto 2022; Mony et al. 2021; Stone et al. 2019; Kribelbauer et al. 2020).

5. Response to Environmental Signals: Many TFs are responsive to external signals or environmental cues. They can integrate information from the cell's surroundings to trigger appropriate gene expression changes in response to these signals (Marinho et al. 2014; X. Wang, Niu, and Zheng 2021; Obata et al. 2020; Qian et al. 2022).

6. Conservation Across Species: Some TFs exhibit a high degree of evolutionary conservation, indicating their fundamental roles in gene regulation across diverse organisms (Hemberg and Kreiman 2011; Nitta et al. 2015; Diehl and Boyle 2018).

7. DNA-Binding Domains: TFs typically contain specific structural motifs, such as DNA-binding domains, that facilitate their interaction with DNA sequences (Ling Zhu and Huq 2011; Inukai, Kock, and Bulyk 2017; Zamanighomi et al. 2017; Dror et al. 2015).

8. Transactivation Domains: Activator TFs often possess transactivation domains that facilitate the recruitment of RNA polymerase and other transcriptional machinery components (Du, McConnell, and Yang 2010; Fukushima et al. 1998; Nioi et al. 2005).

9. DNA-Binding Motifs: These motifs serve as the molecular machinery responsible for the discerning interaction between TFs and their target DNA sequences (Stegmaier et al. 2013; Rhee and Pugh 2011; Kazemian et al. 2013; Gheorghe et al. 2019).

10. Regulation of Cellular Processes: TFs play a critical role in regulating various cellular processes, including development, differentiation, response to stress, and immune responses (Belo et al. 2013; Jackson, Nutt, and McCormack 2023; Almalki and Agrawal 2016; X. Wang, Niu, and Zheng 2021; Yuan et al. 2022).

As previously discussed, TFs are primarily recognized for their pivotal involvement in governing gene expression. They play essential roles in:

1. Initiation of Transcription. The complex of general TFs (TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH) (Koch et al. 2011) along with RNA polymerase II directly bind to the core promoter region (Haberle and Stark 2018). Concurrently, TFIIA plays a crucial

role in stabilizing the binding interaction between TFIID and DNA (Coleman et al. 1999). TFIIB interacts with the DNA by identifying specific elements near the promoter region, helping locate the exact starting point of transcription (TSS) (Buratowski and Zhou 1993). TFIIH contains helicase enzymes that unwind the DNA strands (Hoeijmakers, Egly, and Vermeulen 1996; Rimel and Taatjes 2018). Additionally, TFIIE recruits TFIIH to the recently formed initiation complex (Compe et al. 2019; Ohkuma 1997), while TFIIF brings in RNA polymerase II (Pol II) (Luse 2012; Henry et al. 1994).

2. Enhancing of Transcription. Certain TFs have been established as predominantly binding to enhancers that play a pivotal role in activating the expression of target genes. (Bensimhon et al. 1983) Some TFs can recruit co-activator proteins to the promoter region, which have enzymatic activities, such as histone acetyltransferases (HATs), which modify histones by adding acetyl groups (Ortega et al. 2018; Imhof et al. 1997). This modification loosens the chromatin structure, making it more accessible for transcriptional machinery and resulting in enhanced gene transcription (Wapenaar and Dekker 2016).

3. Repression of Transcription. TFs have the ability to suppress transcription through a variety of mechanisms. They may physically interfere with the binding of RNA polymerase or recruit co-repressors that inhibit the transcriptional machinery (N. Liu et al. 2018). In a competitive binding mechanism, TFs engage in a competition with activator proteins for the occupation of shared DNA regulatory elements (Meyer, Gustafsson, and Carlstedt-Duke 1997). TFs can recruit co-repressor proteins to the promoter region (Shapiro and Shapiro 2011). Co-repressors often have enzymatic activity, such as histone deacetylases (HDACs) (Schoch and Abel 2014), which modify chromatin structure to make it less accessible to the transcriptional machinery. In some

cases, TFs can recruit silencing complexes, such as Polycomb repressive complexes (PRCs) (Dobrinić, Szczurek, and Klose 2021), which are involved in maintaining gene silencing during development and differentiation (L. Wang et al. 2004). Some TFs can indirectly influence gene repression by recruiting DNA methyltransferases (Hervouet, Vallette, and Cartron 2010). DNA methylation often leads to gene silencing by preventing the binding of TFs or transcriptional machinery to the promoter region (Brenner et al. 2005).

4. Cell-specific regulation of transcription by TFs is a critical mechanism that allows different cell types within an organism to express distinct sets of genes (Engel et al. 1992). Certain TFs have been discovered to exhibit variations in their genome binding preferences contingent on whether they are expressed in isolation or within the framework of protein complexes (MacIsaac et al. 2010).

5. TFs play a crucial role in development and differentiation processes in multicellular organisms. TFs are responsible for determining the fate of embryonic cells during development (Yun Zhao, Wang, and Wang 2023). They activate or repress specific genes that drive cells toward specific lineages or cell types (Ng et al. 2021). During development, TFs also establish spatial patterns within tissues and organs. Homeobox genes, for example, control the segmentation of the body in many organisms (Cardoso 1995; Hoekman et al. 2006). In some cases, TFs can induce cellular plasticity, allowing cells to switch between different fates or become more versatile (Wessely et al. 2021).

1.3 Regulation of Transcriptional factors

TFs are subject to intricate regulation across multiple levels, ensuring precise control of gene expression in response to diverse internal and external signals. The initial level of regulation

involves post-translational modifications, which encompass covalent chemical alterations occurring on proteins after their synthesis during or after translation (J. M. Lee et al. 2023). Phosphorylation, a prevalent post-translational modification, is a pivotal mechanism in regulating gene expression through TFs. This alteration involves the addition of a phosphate group (PO_4) to specific serine, threonine, or tyrosine residues within the TF protein. The effects of phosphorylation on TFs are diverse, impacting their activity, stability, localization, and interactions with other proteins (W. J. Zhang et al. 2023). Phosphorylated residues often act as docking sites for other proteins, including co-activators or co-repressors, and interactions with other signaling components can modulate the TF's activity (Bohmann 1990). Several well-known TFs are regulated by phosphorylation, such as CREB (cAMP Response Element-Binding Protein) (Pulimood et al. 2021), NF-κB (Nuclear Factor-Kappa B) (Viatour et al. 2005), p53 (Jenkins et al. 2012), and STAT (Signal Transducer and Activator of Transcription) protein (Breit et al. 2015).

Acetylation of TFs constitutes a post-translational modification where specific lysine residues within the TF protein receive an acetyl group (CH_3CO) (Imhof et al. 1997). In many instances, acetylation enhances a TF's DNA binding ability (Park et al. 2015). Acetylation also plays a role in regulating TF stability by shielding them from degradation through the ubiquitin-proteasome system (Xu and Wan 2023). Additionally, it can impact the subcellular localization of TFs, promoting their retention in the nucleus or export to the cytoplasm. Several TFs are subject to acetylation regulation. For instance, acetylation of p53 enhances its DNA-binding and transcriptional activity, promoting processes like cell cycle arrest and apoptosis (Nagasaka et al. 2022).

Ubiquitination entails the covalent attachment of a small protein known as ubiquitin to particular lysine residues within the TF protein (Ran et al. 2023). One of the primary functions of ubiquitination is to mark proteins for degradation by the proteasome (Shaid et al. 2013).

Ubiquitination has also the capacity to impact the subcellular localization of TFs. While ubiquitination can tag a protein for degradation, it can also prompt its relocation within the cell. In specific instances, ubiquitination may cause the translocation of a TF from the nucleus to the cytoplasm (Gulshan, Thommandru, and Moye-Rowley 2012). For example, the ubiquitination of HIF-1α regulates its stability in response to oxygen levels (Rashid et al. 2021).

Another critical post-translational modification regulating TFs is SUMOylation, Small Ubiquitin-like Modifier. This process involves the covalent attachment of a SUMO protein to specific lysine residues within target proteins (Saldanha and Tollefsbol 2018). SUMOylation predominantly regulates the activity and function of the modified proteins (Yang et al. 2017). SUMOylation can significantly impact the subcellular localization of TFs, frequently favoring nuclear localization (Le et al. 2017). Additionally, SUMOylation plays a crucial role in protein stabilization, safeguarding them against ubiquitin-mediated degradation by the proteasome. For instance, SUMOylation of PML (Promyelocytic Leukemia Protein) contributes to the formation of nuclear bodies (Maroui et al. 2018).

The activity of numerous TFs predominantly occurs when they are situated within the cell's nucleus. TFs often contain NLSs, are specific sequences or motifs within a protein that signal its transport into the nucleus that allow them to be transported into the nucleus (Lu et al. 2021), where they can access DNA (Kosugi et al. 2009). Nuclear export signals (NES) are sequences found within proteins and are responsible for guiding proteins on their journey from the nucleus to the cytoplasm, contributing to the dynamic regulation of protein localization within the cell (Gerace 1995). Cytoplasmic sequestration is a regulatory mechanism that entails confining particular proteins, including TFs, within the cytoplasm of a cell, thereby preventing their entry into the nucleus. This sequestration serves to maintain these proteins in an inactive state until specific signals or conditions prompt their release and subsequent translocation into the nucleus, where they can execute their functions (Haller et al. 2010).

Epigenetic modifications encompass changes in gene expression that do not entail modifications to the underlying DNA sequence. Epigenetic modifications can exert a substantial influence on the regulation of TFs by affecting the accessibility of DNA and the structure of chromatin. These modifications can have consequences on how TFs engage with their target DNA sequences, ultimately leading to the modulation of gene expression (L. Sun, Zhang, and Gao 2022).

Table 2. Post-translational Modifications and Cellular Regulation of Transcription Factors

Modification	Definition	Regulation Mechanism	Effects on TFs	Examples
Phosphorylation	Addition of phosphate group (PO4) to specific residues	Context-dependent, catalyzed by kinases	Diverse impact on activity, stability, DNA binding, localization, and interactions with other proteins	CREB, NF-κB, p53, STAT protein
Acetylation	Addition of acetyl group to lysine residues	Reversible, regulated by HATs and HDACs	Enhances DNA binding, transcriptional activation, chromatin accessibility	p53, NF-κB
Ubiquitination	Attachment of ubiquitin to lysine residues	Regulated by E1, E2, and E3 ubiquitin ligases	Marks for proteasomal degradation, modulates transcriptional activity, affects DNA binding	p53, NF-κB, HIF-1α
SUMOylation	Attachment of SUMO protein to lysine residues	Enzymatic cascade involving E1, E2, and E3 enzymes	Regulates activity, subcellular localization, stability; competes with ubiquitination	PML, Sp1, PPARγ
Cellular Localization	Determines TF's position within the cell	Governed by nuclear localization signals (NLS)	Impact on TF activity based on access to DNA; regulated by NLS, NES, cytoplasmic sequestration	Multiple TFs, e.g., regulated during the cell cycle
Epigenetic Modifications	Changes in gene expression without altering DNA	Affects DNA accessibility and chromatin structure	Modulates how TFs engage with target DNA, influencing gene expression	Various TFs

DNA methylation is a heritable epigenetic modification entailing the addition of a methyl group to the C-5 position of cytosine within DNA. This modification is facilitated by DNA methyltransferases (DNMTs) and is vital for gene regulation and genome stability. In mammals, the three well-known DNA Methyltransferases (DNMTs) are DNMT1, DNMT3A, and

DNMT3B, each playing unique roles in the establishment and maintenance of DNA methylation patterns (Figure 1) (Lyko 2018). DNMT1, recognized as a maintenance enzyme, possesses a strong affinity for hemi-methylated DNA and assumes a critical role in restoring methylation patterns following DNA replication (Hermann, Goyal, and Jeltsch 2004). Conversely, DNMT3A and DNMT3B are regarded as *de novo* methylation enzymes (Okano, Xie, and Li 1998). They participate in the establishment of DNA methylation patterns during early embryonic development and target CpG sites in both hemi- and unmethylated contexts (D. Monk 2015; M. Monk, Boubelik, and Lehnert 1987). Dysregulation of DNMTs can result in aberrant DNA methylation patterns, which have been associated with the onset and progression of various diseases, encompassing cancer, neurological disorders, and developmental anomalies (Law and Jacobsen 2010).
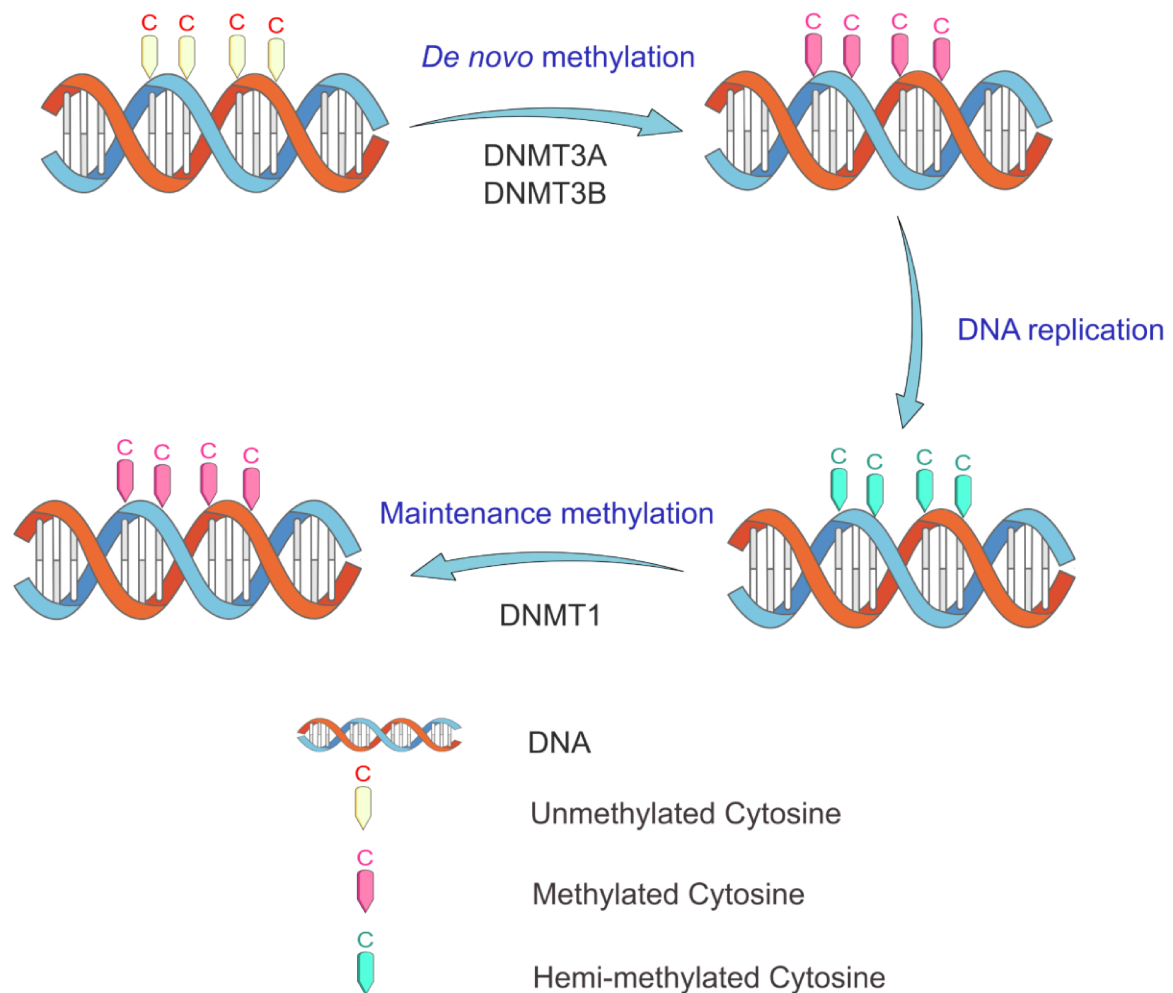
Figure 1. Dynamics of DNA Methylation

The figure was created in Inkscape (https://inkscape.org/) and SciDraw (https://scidraw.io/).

In mammals, 5-methylcytosine (5mC) is predominantly found in the symmetrical context of CG dinucleotides, enabling the inheritance of DNA methylation patterns during cell division (Song et al. 2011). Notably, a substantial proportion of CpG dinucleotides are located in regions known as CpG islands (CGIs) (Deaton and Bird 2011). CGIs are typically approximately 1000 base pairs (bp) long, characterized by a high GC base composition, positioned near Transcription Start Sites (TSS) in roughly 70% of vertebrate gene promoters and are mainly unmethylated. Approximately 80% of CpG sites exhibit methylation throughout the non-embryonic genome (Walsh and Bestor 1999; A. P. Bird 1986). Methylation of CpG islands

(CGIs) leads to transcriptional repression, playing a pivotal role in gene regulation (Meehan et al. 1992; Hsieh 1994). Methylation binding proteins possess specialized Methyl binding domains (MBD) that enable them to bind to 5mC in a non-sequence-specific manner. Higher CpG methylation density is strongly correlated with increased nucleosome occupancy (Collings and Anderson 2017). Methylation of CGIs plays a crucial role in several significant biological processes, including genome imprinting (Plass and Soloway 2002), transposon silencing (Walsh, Chaillet, and Bestor 1998), and X-chromosome inactivation (Cotton et al. 2011). The influence of DNA methylation on the binding dynamics of TFs has become a subject of significant interest in the field. Historically, TF classification has been based on their affinity for methylated DNA motifs. In a comprehensive study involving 519 TFs, it was discovered that 33% of these factors do not contain CpG sites within their binding motifs. Furthermore, a division was observed between those preferring unmethylated CpG sites (23%) and those exhibiting a preference for methylated CpG sites (24%). Notably, the remaining TFs displayed varying degrees of sensitivity to DNA methylation, with some showing minimal or negligible effects. It is important to note that these proportions are specific to *in vitro* studies, and it is expected that *in vivo* results may differ, necessitating further investigation (Y. Yin et al. 2017). Additionally, the interaction between TFs and methylated DNA is finely modulated by the presence of co-factors. For instance, experimental findings have shown that MAFF's binding to methylated DNA exclusively occurs when its cofactors NFE2 and NFE2L2 are absent. This underscores the intricate relationship between TFs and their associated co-factors in the context of DNA methylation binding (Lin et al. 2020).

Understanding the interplay between DNA methylation and TF binding reveals distinct scenarios (Figure 2) (H. Zhu, Wang, and Qian 2016):

1. Many TFs primarily bind to non-methylated DNA motifs in open chromatin regions. DNA methylation within their binding sites (BSs) hinders TF binding. Notable examples include AP-2, MYC, E2F, NF-kB, and ETS (Héberlé and Bardet 2019; Tate and Bird 1993).

2. Conversely, some TFs, such as KAISO/ZBTB33 and CEBPA/B, preferentially bind to DNA regions containing methylated cytosine (5mC) (Rishi et al. 2010).

3. Certain TFs can bind equally well to both methylated and non-methylated DNA, as they possess binding sites accommodating both types of sequences. One example is the CTCF protein (CCCTC-binding factor). *In vitro* experiments have demonstrated that CTCF can interact with DNA sequences containing methylated cytosines (Stadler et al. 2011). However, CTCF has a preference for binding to DNA sequences that are unmethylated (Hao Wang et al. 2012).

4. Some TFs exhibit different target sequences depending on DNA methylation status. For instance, KLF4 binds to two unmethylated sites (TACpGCC) and two methylated sites (CCmCpGCC) (S. Hu et al. 2013).
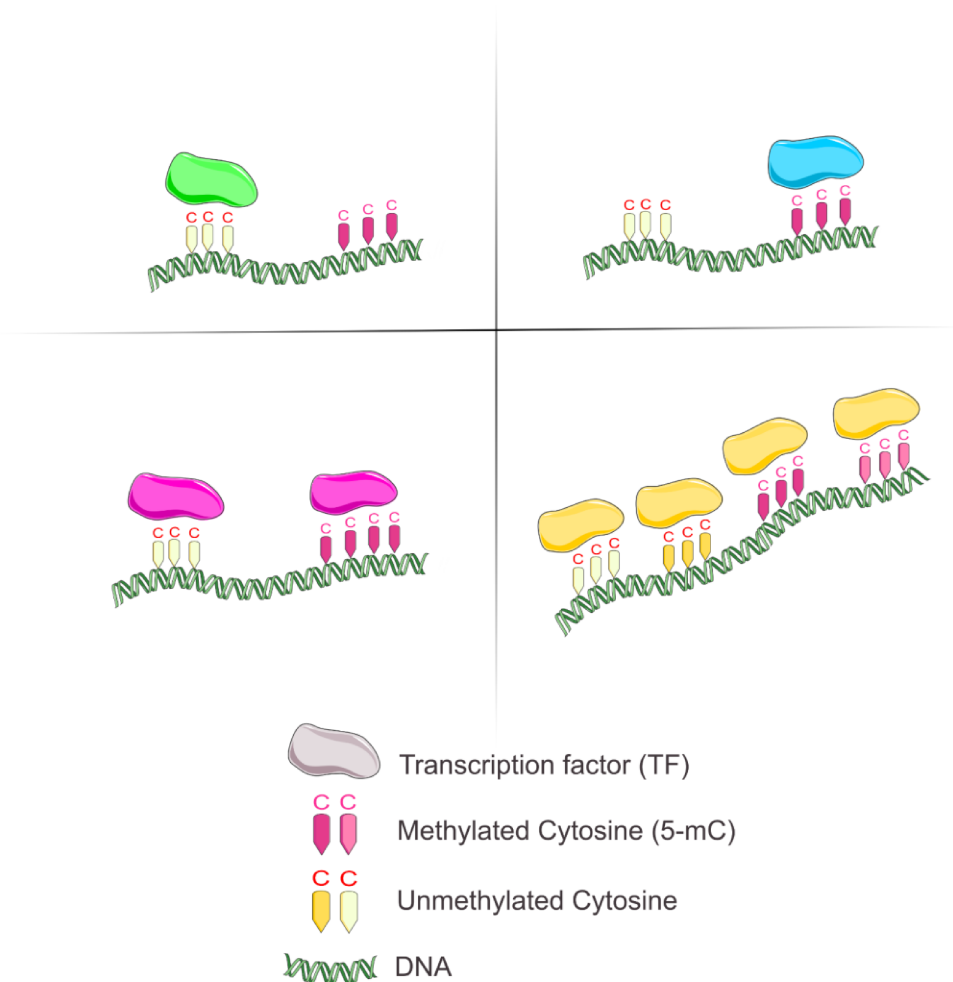
Figure 2. The Interplay of DNA Methylation and Transcription Factor Binding

The figure illustrates the categorization of TFs according to their affinity for binding methylated DNA. The figure was created in Inkscape (https://inkscape.org/) and SciDraw (https://scidraw.io/).

1.4 Pioneer Transcription Factors

Pioneer Transcription Factors (pTFs) represent an indispensable group of proteins that exert a critical influence on the intricate mechanisms governing gene expression and the orchestration of cellular destinies during the phases of development and differentiation (Mayran and Drouin 2018). The concept of pioneering transcription factors emerged as a distinct category of regulatory proteins in 2002 (Cirillo et al. 2002). The nomenclature "pioneer" is ascribed to these proteins due to their capacity to instigate alterations in chromatin configuration. This

16

transformation renders chromatin structures more accessible, thus facilitating the subsequent recruitment and activation or repression of target genes by other TFs and the transcriptional apparatus. The inherent pioneering prowess plays a central role in the pivotal transition from a condensed, inert chromatin state to an open, transcriptionally permissive condition, thereby affording the opportunity for the expression of genes that hold sway over cellular identity and functionality (Fernandez Garcia et al. 2019). The hallmark features of PTFs (PTFs) encompass their remarkable capacity for DNA binding. PTFs are endowed with specialized DNA-binding domains, enabling them to discern and engage with precise sequences within the genome. These sequences are often intricately linked to genes pivotal in cell differentiation and developmental processes (Soufi et al. 2015). One of the remarkable attributes of PTFs is their ability to recruit chromatin-remodeling complexes and induce alterations in chromatin structure. This can entail the relaxation of nucleosome positioning or the displacement of histones, thereby rendering DNA more amenable to interaction with other TFs and the transcriptional machinery (Wolf et al. 2023). During embryonic development and tissue-specific differentiation, PTFs emerge as pivotal orchestrators, decisively influencing cell fate and differentiation. They can either activate or repress specific target genes, ultimately governing the trajectory of diverse cell types and tissues (Aydin et al. 2022). Crucially, PTFs exert their influence in a context-dependent manner, their actions contingent upon the specific cellular milieu, signaling pathways, and developmental stage. This versatility enables them to finely calibrate gene expression in accordance with the needs of the cell. Notably, PTFs play an indispensable role in cellular reprogramming, a process wherein one cell type is transmuted into another, as exemplified in induced pluripotent stem cell (iPSC) generation (Xiao et al. 2016). In this context, PTFs serve as catalysts, initiating the unwinding of chromatin within the target cell type. This event paves the way for the introduction of factors capable of effecting a change in the cell's identity. Nonetheless, it is critical to underscore that the dysregulation of PTFs has been implicated in

various diseases, prominently including cancer. The misexpression of these transcriptional pioneers can lead to the emergence of aberrant gene expression patterns and the disruption of finely-tuned cell differentiation processes (Pavithran and Kumavath 2021).

1.5 DNA binding specificity of pioneer transcription factors

The DNA binding specificity of TFs is influenced by both the structural and chemical characteristics of these proteins. These DNA binding domains come in various types, including zinc fingers, helix-turn-helix motifs, leucine zippers, and homeodomains, each boasting unique structures that enable the recognition and binding to precise DNA sequences (Jolma et al. 2013). These motifs tend to be relatively short, usually spanning 6-12 base pairs, and are commonly located within the regulatory regions of target genes (Inukai, Kock, and Bulyk 2017). For instance, in zinc finger proteins, the coordination of zinc ions by cysteine and histidine residues within the zinc finger structure significantly contributes to DNA binding specificity. TFs often collaborate with other TFs to orchestrate gene expression regulation. In such instances, cooperative binding can heighten specificity (Paillard, Deremble, and Lavery 2004). Multiple TFs may simultaneously bind to adjacent or overlapping DNA sites, refining gene regulation with precision (Hai and Curran 1991). This synergistic control offers a vast array of gene regulatory possibilities and augments specificity (Todeschini, Georges, and Veitia 2014). Moreover, epigenetic modifications, such as DNA methylation and histone acetylation, can alter the accessibility of DNA regions and, consequently, affect the ability of TFs to bind to their target sites (Y. Yin et al. 2017). Another crucial determinant is the recognition of DNA shape. TFs don't solely rely on the DNA's specific sequence but also on its three-dimensional structure or conformation. This phenomenon, known as DNA shape readout (Schnepf et al. 2020), involves some TFs recognizing the minor groove of the DNA double helix (Rodríguez et al. 2015). While less common, some TFs can recognize the major groove of DNA, which

relies on the shape and chemical properties of the major groove (Takayama and Marius Clore 2012). DNA is subject to bending or twisting in specific ways due to the sequence and stacking of base pairs. Proteins interacting with DNA can detect and respond to these structural variations, allowing for sequence-specific binding based on the three-dimensional conformation of DNA (Love et al. 1995). By recognizing the shape of DNA, TFs can distinguish between different DNA sequences that may have similar base pair compositions but possess distinct three-dimensional structures (Stadhouders, Filion, and Graf 2019).

Table 3. Determinants of Transcription Factor DNA Binding Specificity

Determinant	Description	Characteristic
DNA Binding Domains	Structural motifs within TFs engaging with DNA	Unique structures for DNA recognition
Recognition Motifs	Specific DNA sequences recognized by TFs	Short motifs (6-12 base pairs); Often located in regulatory regions of target genes
Cooperative Binding	Collaboration of TFs for enhanced gene regulation	Multiple TFs bind adjacent or overlapping DNA sites; Increases gene regulatory possibilities
Epigenetic Modifications	DNA methylation, histone acetylation altering DNA	Alters DNA accessibility; Affects TFs' ability to bind to target sites
DNA Shape Readout	Recognition of DNA's three-dimensional structure	Involves TFs recognizing minor groove or major groove; Shape-based interaction with DNA
DNA Structural Variations and Dynamics	Bending, twisting, and fluctuations in DNA structure	TFs detect and respond to structural variations; Sequence-specific binding based on three- dimensional conformation
Dynamic Responses to DNA Changes	Recognition and response to dynamic aspects of DNA shape	TFs stabilize or induce specific DNA conformational changes; Facilitates transcription, replication, or repair

1.6 Interplay of chromatin remodeling and pioneer transcription factors

PTFs possess the ability to bind to regions of the genome tightly ensconced in heterochromatin, which is typically linked to gene repression (Swinstead et al. 2016). These pioneer factors wield specialized DNA binding domains designed to recognize their target DNA sequences, even when these sequences are partially obscured by nucleosomes or other repressive elements (Felipe, Shin, and Kolomeisky 2022). The transition to an open chromatin state serves as the gateway to the recruitment and assembly of the preinitiation complex. (Kubik, Bruzzone, and Shore 2017). In contrast, closed chromatin is frequently associated with DNA methylation and histone modifications known for their repressive influence (Mayran and Drouin 2018). Pioneer factors possess the capability to interact with enzymes responsible for the removal or modification of these epigenetic marks (Figure 3). For instance, they can recruit DNA demethylases or histone acetyltransferases (HATs), countering DNA methylation and histone deacetylation (Shakya et al. 2015; Fuglerud et al. 2018). This fosters a more permissive chromatin environment for TF binding. Some pTFs can directly engage with nucleosomes, inducing shifts, deformations, or repositions. Through pushing, pulling, or altering nucleosome positions, these pioneer factors create openings in chromatin structure, exposing DNA for binding by other TFs and the transcription machinery (Schiessel and Blossey 2020). Moreover, pTFs can enlist the aid of chromatin remodeling complexes, exemplified by SWI/SNF or CHD proteins (Wolf et al. 2023). These complexes harness the energy derived from ATP hydrolysis to modify the structure of nucleosomes, thereby enhancing DNA accessibility for TF binding (H. Zhang and Kuchroo 2019; Mivelaz 2021). Another pivotal mechanism involves the recruitment of histone chaperones, such as the H2A-H2B histone chaperone FACT (Facilitates Chromatin Transcription) (Echigoya et al. 2020). FACT interacts with nucleosomes and assist in the displacement of one or both H2A-H2B histone dimers (Fujimoto et al. 2012). This action destabilizes the nucleosome, effectively opening up the DNA for transcriptional access. The

recruitment of FACT or similar histone chaperones by pioneer factors contributes to the modification of chromatin structure, rendering the DNA more accessible to other TFs and the RNA polymerase machinery (Cheloufi and Hochedlinger 2017).
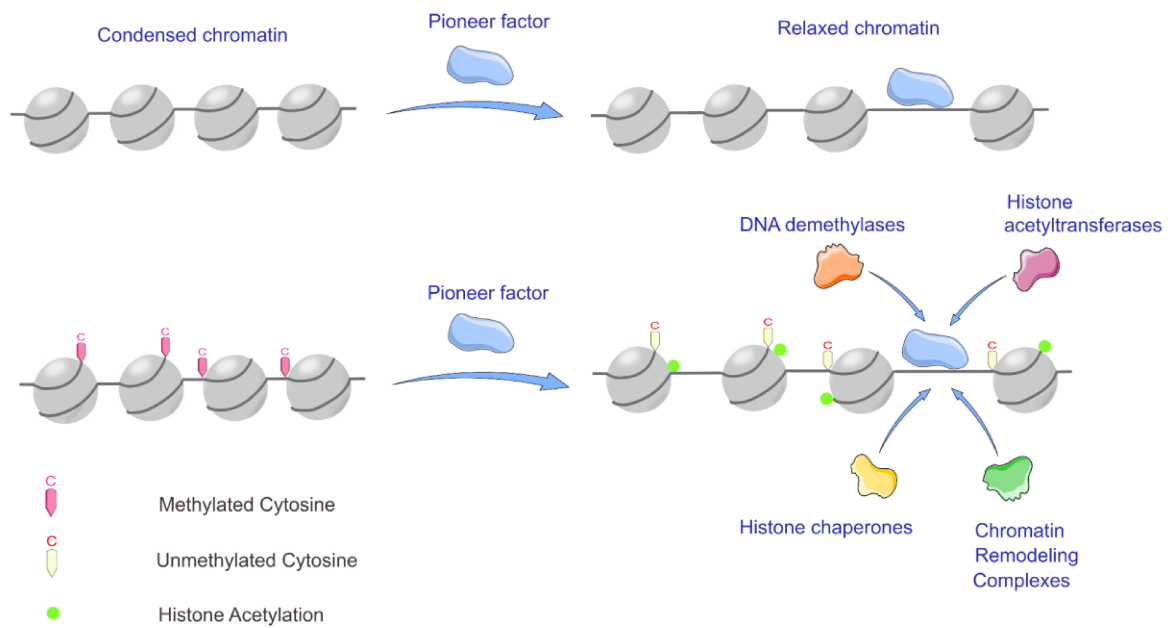
Figure 3. Pioneer transcription factors drive chromatin remodeling.

The figure was created in Inkscape (https://inkscape.org/) and SciDraw (https://scidraw.io/). The figure illustrates the process of chromatin remodeling by pTFs.

1.7 Classification of Pioneer Transcription Factors

A multitude of TF families exists, each distinguished by its unique characteristics and functions.

Here, we present an overview of some pTF families, offering brief descriptions and examples.

Table 4. Functional Classification of Transcription Factors

Functional Classification	Description	Examples
Transcription Activators	Amplify or activate transcription of target genes; bind to DNA sequences in promoter or enhancer regions; recruit co-activator proteins to enhance transcription initiation	Heat Shock Factors (HSFs), Hypoxia-Inducible Factors (HIFs), Nuclear Factor-kappa B (NF-κB)
Transcriptional Repressors	Suppress or inhibit transcription of specific target genes; bind to DNA sequences in regulatory regions; may obstruct RNA polymerase binding or recruit co-repressor proteins	REST, Rb, MeCP2
Stress Response TFs	Govern expression of genes aiding cells in adapting to stress; examples include Heat Shock Factors (HSFs) and Hypoxia-Inducible Factors (HIFs)	Heat Shock Factors (HSFs), Hypoxia-Inducible Factors (HIFs)
Cell Cycle Regulators	Govern cell cycle progression; oversee expression of genes involved in DNA replication, cell division, and checkpoints	E2F TFs
Homeotic TFs	Regulate development and patterning; establish body plan and determine segmental identity along the anterior-posterior axis of animals	MADS-Box
Nuclear Receptors	Ligand-activated regulators of gene expression; activated by specific ligands; possess DNA-binding domains for binding to hormone response elements (HREs) or nuclear receptor response elements (NRREs)	Estrogen Receptor (ER), Androgen Receptor (AR), Thyroid Hormone Receptor (THR), Retinoic Acid Receptor (RAR),
Master Regulator TFs	Central role in governing gene expression programs; determine cell fate or identity	MyoD, PAX6

Homeodomain pTFs feature a conserved 60-amino acid DNA-binding domain known as the homeodomain (Hankey, Chen, and Wang 2020). Notably, the Homeobox (Hox) genes encode a cluster of TFs instrumental in sculpting the body plan during embryonic development. Hox proteins serve as pioneer factors by participating in the establishment of anterior-posterior axis identity (Paul, Peraldi, and Kmita 2024). Oct4 (Octamer-binding TF 4) is a key regulator in

preserving pluripotency and self-renewal in embryonic stem cells. Its significance extends to early embryonic development, and it serves as a pioneer factor in the context of reprogramming somatic cells into induced pluripotent stem cells (iPSCs).

Basic Helix-Loop-Helix (bHLH) PTFs exerts a pivotal role in gene expression and cell differentiation (Donovan et al. 2023). A defining characteristic is their ability to form homo- or heterodimers with other bHLH proteins. Dimerization plays a critical role in enhancing their DNA-binding affinity and specificity (Nebert 2017). MyoD is instrumental in muscle development and differentiation. It plays a crucial role in the activation of muscle-specific genes by binding to closed chromatin regions and facilitating the recruitment of other regulatory factors (Battistelli, Garbo, and Maione 2022).

Zinc finger pTFs are characterized by their DNA-binding domains, which contain zinc finger motifs. They are often involved in recognizing specific DNA sequences and facilitating the initial binding of the pioneer factor to the chromatin (Soufi et al. 2015). GATA-1 plays a crucial role in the development of red blood cells (erythropoiesis) and is involved in the activation of genes required for this process (Imanishi et al. 2010). By binding to this motif, GATA-1 helps open the chromatin structure around key erythropoiesis-related genes, allowing other TFs and regulatory proteins to access these genes and promote their expression (Y. W. Kim et al. 2020). Nuclear Hormone Receptors (NHRs) represent a class of proteins with a distinctive role as pTFs, exerting a crucial influence on gene regulation (Aranda and Pascual 2001). Their primary function is to modulate the transcription of specific target genes in response to small, lipophilic molecules, which are often a wide range of compounds, including steroid hormones, thyroid hormones, retinoids, and fatty acids hormones or dietary metabolites (Beato and Klug 2000; Yaffe and Samuels 1984; Tao et al. 2020). One well-recognized example within the NHR family is the Estrogen Receptor (ER), which plays a central role in mediating the effects of estrogen hormones (Lindberg et al. 2003). ER has two isoforms, ERα and ERβ, and is

instrumental in regulating the expression of genes associated with cell proliferation, differentiation, and the development of female secondary sexual characteristics (Paige et al. 1999). ER operates as a pTF by engaging with estrogen response elements (EREs) situated in the target genes (Yaşar et al. 2017). In breast tissue, for instance, ER binds to EREs within gene promoters such as ESR1 (estrogen receptor alpha) and progesterone receptor (PR), effectively initiating their transcription (Murphy et al. 2000).

The Forkhead Box (FOX) pTFs, characterized by a conserved DNA-binding domain called the Forkhead box or winged-helix domain (Lalmansingh et al. 2012). Their influence spans an array of biological processes, including development, metabolism, and immune regulation. One example from this family is FOXA1, a versatile factor integral to diverse biological processes, ranging from development and metabolism to cancer. It wields the capability to engage with compacted chromatin, aiding in its unfolding and rendering it accessible for other TFs (Lupien et al. 2008). In the context of hormone-dependent breast cancer, FOXA1 collaborates with the estrogen receptor (ER). It homes in on enhancer regions in DNA, thereby facilitating the binding of ER to its target genes. This synergy holds paramount importance for the expression of genes responsive to estrogen and for the growth of cancer cells (Augello, Hickey, and Knudsen 2011).

The Signal Transducer and Activator of Transcription (STAT) pTFs serve as essential intermediaries in the transmission of extracellular signals, including cytokines and growth factors, from the cell surface to the nucleus (Singh et al. 2012). Once in the nucleus, they activate or repress the transcription of specific genes. One prime illustration is STAT1, which operates significantly in the immune response to interferons. It assumes the role of a pTF, initiating the transcription of genes that facilitate the antiviral response. An instance of this is STAT1's binding to the promoter regions of interferon-stimulated genes (ISGs), which are instrumental in executing antiviral functions (X. Ren et al. 2023; Ruvolo et al. 2003). STAT5

plays an indispensable role in mediating the effects of a variety of cytokines, including erythropoietin and prolactin and takes the lead in regulating genes associated with hematopoiesis and milk protein synthesis (Nosaka et al. 1999). Peroxisome Proliferator-Activated Receptors (PPARs) form a group of nuclear receptors with a pivotal role in orchestrating gene expression and a wide array of physiological processes, particularly those related to lipid and glucose metabolism (Chinetti, Fruchart, and Staels 2000). MYB pTFs constitute a vital group of regulatory proteins, wielding a significant impact on gene expression and regulation through their specific binding to DNA sequences, thus instigating alterations in chromatin structure (Lemma et al. 2021). The defining characteristic of MYB pTFs lies in their DNA-binding domain, typically composed of one to three MYB repeats. Each MYB repeat, consisting of approximately 50 amino acids, assumes a helix-turn-helix structure, endowing these factors with the capability to bind to particular DNA sequences (Ito 2005; Jin and Martin 1999). One example in this category is c-MYB, which plays an indispensable role in the development and upkeep of blood cells. By binding to the regulatory regions of genes necessary for blood cell formation, c-MYB can catalyze their expression, even when embedded within heterochromatin (Greig, Carotta, and Nutt 2008; Allen, Bender, and Siu 1999). In a more complex context, certain MYB pTFs have been implicated in cancer. In this scenario, MYB pTFs can contribute to the inception and progression of cancer by activating genes that fuel uncontrolled cell growth, thus playing a role in the cancer's pathogenesis (Yihao Li et al. 2016; Martinez and Dimaio 2011).

Table 5. Structural Classification of Transcription Factors

Structural Classification	Description
Helix-Turn-Helix (HTH) Motif	Consists of three helices; the recognition helix establishes contact with DNA bases; the second helix enhances stability; modifications involve zinc chelation and tetra-helical variants
Winged HTH (wHTH) Domains	Core tri-helical bundle with a C-terminal β-strand hairpin wing; wing provides additional interface for DNA contact, including interactions with methylated adenine
Basic Helix–Loop–Helix (bHLH) Motif	Consists of two amphipathic α helices separated by an intervening loop; forms a dimeric structure binding to E-boxes in DNA; various families with distinct functions
Forkhead Domain (Winged-Helix Domain)	Distinguished by a three-dimensional structure resembling a "winged forkhead"; comprised of three alpha helices and two loops; recognizes forkhead binding sites or forkhead response elements (FHREs)
Ribbon-Helix-Helix (RHH) Motif	Common in prokaryotes; utilizes an anti-parallel β-sheet to recognize nucleotide sequences; employs α-helices to secure the β-sheet within the DNA major groove
Basic Leucine-Zipper (bZIP) Domains	Positively charged region for interaction with DNA; leucine zipper promotes dimerization; recognizes short, inverted, repeat DNA sequences
Homeodomain	Conserved DNA-binding domain with around 60 amino acids; binds to homeoboxes or Hox binding sites; regulates transcription of target genes in development
C2H2 Zinc Finger Motif	Contains two cysteine and two histidine residues coordinating with a zinc ion; compact, finger-like structure; multiple motifs recognize specific DNA sequences
Beta-Scaffold (β-Sheet) Motif	Defined by a distinctive beta-sheet structure; interacts with DNA bases using convex and concave sides; stabilized by hydrogen bonds
Immunoglobulin-Like (Ig Fold) Domain	Beta-sheet structure composed of antiparallel beta-strands; varying numbers of domains; diverse domain composition expands DNA recognition range

1.8 The role of pTFs in embryonic development

Embryonic development encompasses the intricate sequence of events through which a

fertilized ovum, or zygote, undertakes a meticulously orchestrated transformation to

multicellular organism (Canse, Yildirim, and Yaba 2023). During fertilization, when a haploid spermatozoon unites with a haploid oocyte, a diploid zygote is created (Egozcue et al. 2002). Subsequent to fertilization, the zygote embarks on a swift series of cleavages. Though the embryo's size remains relatively unchanged, the number of cells amplifies significantly, manifesting as smaller units termed blastomeres (Z.-C. Wang, Zhang, and Li 2022). As cleavage persists, blastomeres unite to form a compact cell mass, the morula. The morula, in due course, undergoes transformation into a blastula. During gastrulation, the blastula's cellular reorganization occurs, which gives rise to the establishment of the three primary germ layers: the ectoderm, mesoderm, and endoderm. Cdx2, or Caudal-Type Homeobox 2, serves as a pivotal pTF during early embryonic development, specifically guiding the differentiation of trophoblast stem cells into trophectoderm cells (Blij et al. 2015). In this role, Cdx2 binds to closed chromatin regions, effectively triggering the expression of genes essential for trophectoderm formation (Dady et al. 2014).

Epigenetics plays a crucial role in the complex process of embryonic development, as it has impact on how genes are turned on or off without changing the underlying DNA code. Differentiating cells into specialized types, like muscle cells, neurons, or skin cells, heavily relies on the intricate web of DNA methylation regulation (Smith and Meissner 2013). In the realm of DNA methylation, two critical processes come into play: *de novo* methylation and reprogramming (He and Feng 2022). *De novo* methylation is all about setting up DNA methylation patterns during embryonic development. This process is crucial in primordial germ cells (PGCs), which serve as the precursors to sperm and eggs (Andrews et al. 2023). After the initial methylation of PGCs, there's a follow-up reprogramming process, which serves to wipe away many of the methylation marks, ensuring that germ cells become eggs or sperm with an epigenetic state suitable for fertilization and future embryonic development (Zeng and Chen 2019).

DNA methylation also significantly influences placental development (Bianco-Miotto et al. 2016; Apicella et al. 2019). For instance, the PEG1/MEST gene is crucial for proper placental development, and its methylation pattern is essential for healthy fetal growth (Hahn, Yang, and Chung 2005). When it comes to heart development, DNA methylation is also significant. Genes related to cardiac development, like NKX2-5, are under the regulation of DNA methylation (Akazawa and Komuro 2005; Tong 2016; P. Zhou et al. 2022). Lastly, DNA methylation patterns are pivotal for limb development. Genes involved in shaping and differentiating limbs, such as HOXD genes, are influenced by DNA methylation. Any alterations in DNA methylation can result in limb abnormalities (Fabre et al. 2018; Williamson et al. 2012).

Genomic imprinting is when specific genes are expressed in a manner determined by the parent from whom they were inherited (Bartolomei, Oakey, and Wutz 2020). This relies on whether a gene comes from the mother or father and entails distinct DNA methylation at imprinting control regions (ICRs) (Matsuzaki et al. 2018). For instance, the *Igf2* gene (insulin-like growth factor 2) is only expressed from the father's allele, while the mother's allele is silenced through DNA methylation (Toder et al. 1996). One crucial player in genome imprinting is the CCCTC-binding factor (CTCF). CTCF binds to DNA, organizing the chromatin structure by creating loops and boundaries that separate active and repressed regions of the genome (Weth and Renkawitz 2011).

X-chromosome inactivation (XCI) is a crucial epigenetic process in female mammals that serves to equalize gene expression between the sexes, as males have only one X chromosome (Galupa and Heard 2018). DNA methylation and the long non-coding RNA XIST, derived from one X chromosome, function as key players in gene silencing. XIST acts as a "silencer" by enveloping the chromosome and inhibiting active gene transcription. (Aguilar et al. 2022; Brockdorff 2019). In X-chromosome inactivation, pioneer factors, notably CTCF, play a crucial role in initiating the silencing of one X chromosome in female cells. CTCF binds to XIST RNA,

facilitating its spread across the inactive X chromosome and initiating the silencing process. (Kung et al. 2015; H. Fang et al. 2023).

Epigenetic memory is a crucial aspect of preserving and passing on epigenetic information during embryonic development. Epigenetic marks acquired early in development can influence gene expression patterns and cell fate decisions as the embryo progresses and transforms into different tissues and cell types (A. Bird 2002; Cheedipudi, Genolet, and Dobreva 2014). PTFs lead in creating and sustaining epigenetic memory by transmitting gene expression patterns through cell divisions without changing the DNA sequence. (Mayran et al. 2018; Pataskar et al. 2016; Reizel et al. 2021). One example of a pTF is Oct4, which is critical for maintaining the pluripotent state of embryonic stem cells (ESCs). Oct4 binds to the regulatory regions of key pluripotency genes and promotes an open chromatin structure by recruiting chromatin-remodeling complexes (Whyte et al. 2013). This open chromatin state allows other TFs, like Sox2 and Nanog, to bind and reinforce the pluripotent state of ESCs (Rizzino and Wuebben 2016; Swain et al. 2020). In summary, pTFs are central in establishing and preserving epigenetic memory. They initiate changes in chromatin structure that enable other TFs to access and regulate specific genes, ensuring the proper development and differentiation of cells.

Histone modifications are key players in the intricate process of embryonic development, as they regulate gene expression and cell differentiation at various stages of growth (Macrae, Fothergill-Robinson, and Ramalho-Santos 2023). Histones, proteins organizing DNA in the nucleus, undergo changes that impact DNA packing, influencing the accessibility of specific genes for transcription. Acetylation typically opens up the chromatin structure, making genes more accessible for transcription, often associated with gene activation (Shvedunova and Akhtar 2022). Histone acetylation of the Sox2 promoter region allows this gene to be expressed, promoting the development of the neural tube (Ura et al. 2011; Karadkhelkar et al. 2023). Methylation can either activate or repress gene expression, depending on the specific amino

acids modified and the extent of methylation (Matoba et al. 2014). Histone phosphorylation often occurs during the cell cycle's progression and influences the regulation of genes involved in cell division, differentiation, and DNA repair (Wendt and Shilatifard 2006; Murakami 2019). In embryonic stem cells, the ubiquitination of histone H2B is associated with gene activation, enabling the expression of pluripotency genes that maintain the undifferentiated state of these cells (Ooga, Suzuki, and Aoki 2015; M. K.-W. Ma et al. 2011). PTFs (TFs) play a crucial role in directly influencing gene transcription by interacting with histone-modifying enzymes. They often collaborate with other TFs to regulate gene expression. FoxA1, a pTFs essential in liver development, binds to silent chromatin regions. It works with other TFs to recruit histone acetyltransferases and methyltransferases, activating genes crucial for liver development and function (Kohler and Cirillo 2010; Bommi-Reddy et al. 2022).

In summary, pTFs are pivotal players in the realm of histone modification and chromatin remodeling during embryonic development. They enable precise control of gene expression, allowing cells to differentiate and acquire specific identities. Their interactions with histone-modifying enzymes shape the epigenetic landscape, which is crucial for gene regulation during development.

1.9 The role of pioneer transcription factors in cell fate determination

The role of pTFs in cell fate determination is to initiate the transcription of specific genes that are essential for the development and maintenance of a particular cell type or cell fate. They set the stage for other TFs to bind and work together to establish and maintain cell identity.

During muscle cell differentiation, pTFs collaborate to initiate and facilitate muscle development by unwinding chromatin structure. This regulatory network is crucial for generating functional muscle cells and forming muscle tissue. Mef2 (Myocyte Enhancer Factor 2), particularly Mef2c, bind to enhancer regions of muscle-specific genes, helping to maintain

an open chromatin structure required for gene expression in mature muscle cells (Taylor and Hughes 2017).

In the context of adipogenesis, pTFs are indispensable for orchestrating the transformation of precursor cells into mature adipocytes (Mueller 2014). PPARγ (Peroxisome Proliferator-Activated Receptor Gamma), often considered one of the master regulators of adipogenesis, plays a central role in converting preadipocytes into mature adipocytes (Garin-Shkolnik et al. 2014). It activates genes responsible for lipid metabolism and storage, including those encoding lipogenic enzymes such as fatty acid synthase and lipoprotein lipase (Nakamura, Yudell, and Loor 2014).

PTFs play a critical role in initiating and shaping gene expression programs essential for the differentiation of various cell types within the heart, including cardiomyocytes, cardiac fibroblasts, and endothelial cells. GATA4 binds to closed chromatin regions during early heart development, enabling access to cardiac-specific genes (Lili Zhu et al. 2022; P. Zhou et al. 2022). GATA4 activates genes related to cardiomyocyte differentiation, including those encoding cardiac structural proteins like myosin, troponin, and actin (Davidson et al. 2005; Wu et al. 2022). Additionally, GATA4 collaborates with other cardiac TFs, such as NKX2-5, to coordinate gene expression during cardiomyocyte differentiation (Tong 2016).

PTFs are pivotal in the differentiation of blood cells, as they initiate the activation of genes and guide hematopoietic stem cells toward specific blood cell lineages, such as erythrocytes, myeloid cells, and lymphoid cells. PU.1 is a key pTF in myeloid differentiation, regulating the development of granulocytes, monocytes, and dendritic cells (Nerlov and Graf 1998). It binds to enhancer and promoter regions of myeloid-specific genes, facilitating chromatin accessibility (Kueh et al. 2013). PU.1 also interacts with co-factors like C/EBPα and RUNX1 to aid in lineage commitment and myeloid cell differentiation (D. E. Zhang et al. 1996).

PTFs are central to neural cell differentiation, establishing the initial transcriptional landscape that dictates the fate of various neural cell types such as neurons, astrocytes, and oligodendrocytes. (Horisawa and Suzuki 2023). Neurogenin (Ngn1 and Ngn2) are essential for the differentiation of neural progenitor cells into neurons (Kele et al. 2006; Q. Ma et al. 1999). They bind to closed chromatin regions of target genes, including those encoding neural-specific adhesion molecules and ion channels (Satoh et al. 2010; Hulme et al. 2020). These factors play a critical role in initiating the neuronal differentiation program in neural progenitor cells.

PTFs are pivotal in the process of bone cell differentiation, initiating and maintaining the transcription of genes associated with osteoblast development. Runx2 is considered the master TF in osteoblast differentiation (Vimalraj et al. 2015; Komori 2019). Runx2 binds to specific chromatin regions, recruiting other TFs such as Osterix, and activating osteoblast-specific genes (Hesse et al. 2010; X. Yin et al. 2022).

PTFs are pivotal in the differentiation of sperm cells, enabling the access of other TFs to specific DNA regions and promoting the expression of genes necessary for the development and maturation of sperm cells. CREM (cAMP-Responsive Element Modulator) is a pTF that plays a critical role in post-meiotic spermatid development and sperm maturation (Peri and Serio 2000). It is involved in the process of histone displacement and chromatin remodeling in haploid spermatids (Hogeveen and Sassone-Corsi 2006). CREM binds to chromatin, opening it up to facilitate the expression of genes necessary for spermatid differentiation and the formation of functional spermatozoa (Martianov et al. 2010).

PTFs play a crucial role in the process of female egg cell (oocyte) differentiation. LHX8 is a pTF associated with female reproductive development. It is involved in the differentiation of primordial follicles, which house the oocytes (Y. Ren et al. 2015). LHX8 helps open up chromatin regions that regulate key genes in the folliculogenesis process. PTFs are critical in oocyte differentiation because they are responsible for reshaping the epigenetic landscape of

the genome, making it more permissive for subsequent stages of differentiation (Z. Wang et al. 2020; Jagarlamudi and Rajkovic 2011).

PTFs hold a pivotal role in both initiating and sustaining the differentiation of epithelial cells. Krüppel-like factor 4 (KLF4) is a noteworthy pioneer factor, particularly in the context of Epithelial-Mesenchymal Transition (EMT), is a process in which epithelial cells undergo a transition into mesenchymal cells (N. Tiwari et al. 2013). KLF4 activates genes responsible for the loss of epithelial traits and the acquisition of mesenchymal characteristics (Cui et al. 2013). It stimulates the activation of key EMT regulators such as Snail, Slug, and ZEB1 (A. Tiwari et al. 2017; Z. Li et al. 2018).

PTFs are indispensable for pancreatic cell differentiation including both endocrine and exocrine cells (Edlund 2001). FOXA2 stands out as a pivotal pTF, especially in the early stages of pancreatic development, where it exerts its influence on both endocrine and exocrine cells (C. S. Lee et al. 2005; K. Lee et al. 2019). FOXA2's primary function is to unravel the chromatin structure surrounding genes essential for pancreatic organogenesis. It also plays a central role in the differentiation of various cell types within the pancreas, including the insulin-producing beta cells (Campbell and Hoffman 2016).

In eye cell differentiation, pTFs, exemplified by Pax6, play a central role in regulating diverse cell types and establishing precise gene expression for functional eye structures. Pax6 is a master regulator crucial for differentiating multiple eye cell types, including the cornea, lens, and retina. (Ashery-Padan and Gruss 2001; Baker et al. 2018). Pax6 acts as a pTF by opening up the chromatin structure, granting other TFs access to target genes (Elvenes et al. 2010; Cvekl et al. 2004). In the lens, Pax6 triggers the expression of genes like crystallins, which are vital for lens transparency and refractive properties (Shaham et al. 2012).

In summary, pTFs guide cell fate determination by initiating gene expression and cellular differentiation. They bind to condensed chromatin, unlocking specific genes for transcription,

enabling other TFs to regulate target genes and impact cell fate decisions. This fundamental mechanism ensures proper tissue development and maintains tissue function.

1.10 Dysregulation of pioneer transcription factors in cancer

Transcription factor (TF) dysregulation is common in cancer, driving initiation, progression, and metastasis by influencing target gene transcription. In breast cancer, specific TFs, such as the Estrogen Receptor (ER), act as pTFs, exerting influence, particularly in hormone receptor-positive cases (Reese et al. 2022). This subtype constitutes a significant portion of breast cancer diagnoses. ER binds to estrogen response elements (EREs) within DNA, instigating the transcription of genes linked to cell growth and proliferation (Klinge 2001). Targeted therapies like tamoxifen and aromatase inhibitors effectively modulate ER activity, providing an effective treatment strategy for ER-positive breast cancers (Tonetti and Jordan 1997; Adhikari, Baidya, and Jha 2020; Qin et al. 2022). BRCA1 (Breast Cancer 1) is a tumor suppressor gene involved in DNA repair and maintaining genomic stability (Yoshida and Miki 2004). Mutations in BRCA1 are associated with an increased risk of hereditary breast cancer and ovarian cancer (Elstrodt et al. 2006). BRCA1 serves as a pTF by facilitating the recruitment of DNA repair complexes to sites of DNA damage (Timms et al. 2014).

In the realm of kidney cancer, also known as renal cell carcinoma (RCC), these TFs exert their impact on gene expression, contributing to the onset and advancement of the disease (Bahadoram et al. 2022). HIF-1α (Hypoxia-Inducible Factor 1α) is a well-recognized pTF in kidney cancer. The loss of the von Hippel-Lindau (VHL) gene leads to HIF-1α stabilization (Schödel et al. 2016). HIF-1α promotes the transcription of genes involved in angiogenesis, such as VEGF, facilitating tumor vascularization (Hoefflin et al. 2020). It also regulates glycolysis and shifts cell metabolism toward a glycolytic state, promoting cancer cell survival and proliferation (S.-H. Lee et al. 2020).

Sarcoma is a heterogeneous group of cancers that arise from the mesenchymal tissues, which include bone, muscle, fat, and connective tissues (Helman and Meltzer 2003). PTFs play a crucial role in the development and progression of sarcoma cancers. EWS-FLI1 is a fusion protein formed as a result of a translocation between the EWSR1 and FLI1 genes and is a hallmark of Ewing's sarcoma (Shi et al. 2020). EWS-FLI1 acts as a pTF by binding to closed chromatin regions and recruiting other TFs and co-activators (Sunkel et al. 2021; Grünewald et al. 2018). This fusion protein promotes the expression of genes associated with cell proliferation and tumorigenesis, such as MYC and CD99 (H. Sun et al. 2017; Rocchi et al. 2010). It also inhibits the expression of genes that promote differentiation.

PTFs are essential for initiating and regulating gene expression in bladder cancer. GATA3 is crucial for bladder cancer regulation. It plays a role in maintaining the differentiation of urothelial cells and preventing them from undergoing malignant transformation (Inoue et al. 2017). In bladder cancer, the loss of GATA3 expression is associated with more aggressive and invasive forms of the disease (Yi Li et al. 2014). GATA3 is involved in regulating genes like CDH1 (E-cadherin) and KRT20, both of which are important for maintaining the normal differentiated state of urothelial cells (Guo and Czerniak 2019; B. Kim et al. 2021).

PTFs play a crucial role in the development and progression of prostate cancer. Androgen Receptor (AR) is one of the most well-known pTFs in prostate cancer (Heinlein and Chang 2004). In normal prostate cells, AR regulates the expression of genes that control prostate development and function (Labbé and Brown 2018). In prostate cancer, AR can become hyperactive, leading to the uncontrolled growth of cancer cells (Tan et al. 2015). This phenomenon is called androgen receptor signaling, and it is a key driver of prostate cancer progression (Culig and Santer 2014). In prostate cancer, AR can facilitate the expression of genes involved in cell proliferation, survival, and metastasis, such as PSA (Prostate-Specific

Antigen) (J. Kim and Coetzee 2004) and TMPRSS2 (Transmembrane Protease, Serine 2) (Cai et al. 2009).

PTFs play a crucial role in the development and progression of leukemia cancer, which is characterized by the uncontrolled proliferation of white blood cells in the bone marrow. GATA-1 is a TF essential for normal hematopoiesis, particularly in the development of erythrocytes and megakaryocytes (Shimizu et al. 2004). Mutations in GATA-1 are associated with various leukemia subtypes, such as Down syndrome-related acute megakaryoblastic leukemia (DS-AMKL) (Hitzler et al. 2003). In DS-AMKL, GATA-1 mutations act as pioneer factors that bind to specific sites on DNA and alter the transcriptional program, promoting the transformation of hematopoietic cells into leukemia cells (Burda, Laslo, and Stopka 2010).

In melanoma, a highly aggressive and deadly form of skin cancer, pTFs play a crucial role in initiating and maintaining the cancerous phenotype. MITF (Microphthalmia-Associated Transcription Factor) is a key TF in melanoma, and it acts as a pioneer factor by regulating genes involved in melanocyte differentiation and pigment production (Levy, Khaled, and Fisher 2006). In melanoma, MITF is often amplified or mutated, leading to its dysregulated expression. This results in the activation of genes associated with melanoma proliferation and survival (Hartman and Czyz 2015). MITF's pioneer activity helps maintain the melanoma phenotype by binding to target genes, such as TYR (Tyrosinase) (Niu, Yin, and Aisa 2018) , TYRP1 (Tyrosinase-Related Protein 1) (D. Fang, Tsuji, and Setaluri 2002), and DCT (Dopachrome Tautomerase) (Jiao et al. 2004).

In lymphoma, a type of blood cancer primarily affecting lymphocytes, the influence of pTFs is considerable and can significantly impact disease progression and the formation of tumors (Mugnaini and Ghosh 2016). PU.1 (SPI1) is a pTF predominantly expressed in lymphoid and myeloid cells, crucial for B- and T-lymphocyte development and function (Ivascu et al. 2007). Dysregulation of PU.1 in lymphoma can alter chromatin accessibility, influencing the

expression of genes involved in lymphocyte differentiation and activation (Rosenbauer et al. 2006). For example, in chronic lymphocytic leukemia (CLL), a B-cell lymphoma, downregulation of PU.1 is associated with disease progression and a less favorable prognosis (Huskova et al. 2015; Okuno and Yuki 2012). PU.1 regulates numerous target genes critical for lymphocyte development and function, establishing its role as a pivotal pioneer factor (Desai et al. 2009).

In the context of colorectal cancer (CRC), pTFs are central to the dysregulation of key genes and pathways that contribute to the oncogenic process. FOXO1 (Forkhead Box O1) is a pTF that governs genes involved in apoptosis, DNA repair, and cell cycle control. In CRC, dysregulated FOXO1 activity can drive uncontrolled cell proliferation and inhibit apoptosis (D. Wang et al. 2019). For instance, the downregulation of FOXO1 can result in increased expression of oncogenes like MYC and cyclin D1, contributing to tumorigenesis (Shang et al. 2020; Jiang et al. 2018).

In the realm of brain cancer, pTFs play a pivotal role in influencing tumor development, progression, and responses to therapeutic interventions. SOX2 serves as a pTF in Glioblastoma (GBM), preserving the stemness of glioblastoma stem cells (GSCs) by opening closed chromatin regions. SOX2 overexpression in neural progenitor cells can induce gliomagenesis instigating tumor formation bearing GBM-like characteristics, thereby underscoring its role in tumor initiation (Caglar and Duzgun 2023). SOX2 is implicated in resistance against radiation therapy by regulating DNA repair mechanisms, thereby enhancing the repair of radiation-induced DNA damage in GBM cells and diminishing their susceptibility to treatment (Garros-Regulez et al. 2016). SOX2 is also implicated in chemotherapy resistance through its association with the upregulation of drug efflux pumps like ABCG2, which actively expel chemotherapy drugs from cancer cells, thereby diminishing their efficacy (Wee et al. 2016).

1.12 Experimental Techniques for Studying DNA Methylation

Accurate quantification and validation of 5-methylcytosine play a pivotal role in the field of epigenetic investigation (Khodadadi et al. 2021). Scientists use two main methodologies for understanding DNA methylation: molecular biology and bioinformatics. Molecular techniques, like bisulfite conversion, DNA sequencing, and PCR amplification, discern and quantify 5-methylcytosine. Bioinformatic methods analyze next-generation (next-gen) sequencing data for methylated cytosines. This comparative analysis explores their strengths and limitations, providing insights into optimal strategies for 5-methylcytosine quantification. Bisulfite treatment, a widely used technique, distinguishes methylated and non-methylated cytosines, allowing precise identification and quantification after PCR amplification (L. Zhang et al. 2015). Whole-genome bisulfite sequencing (WGBS) stands out as the method of choice for DNA methylation analysis due to its remarkable capacity to accurately detect every methylated cytosine within the genome, covering approximately 95% of all CpG sites in the human genome. WGBS uses treatment of extracted DNA with sodium bisulfite. This chemical treatment converts unmethylated cytosine residues to uracil, while methylated cytosines remain unchanged. This conversion creates a difference that can be detected during sequencing (Stirzaker et al. 2014). However, it's important to note that whole-genome bisulfite sequencing (WGBS) has limitations, including the requirement for relatively large input material quantities and the relatively recent development of protocols, making them less suitable for high-throughput applications and potentially limiting their use in clinical contexts. (Olova et al. 2018). Additionally, it's widely acknowledged that the amplification of bisulfite-treated DNA can potentially introduce unwanted biases (Bundo et al. 2012). Despite its precision and comprehensiveness, WGBS is essential for deciphering DNA methylation patterns and their implications for health and disease. Enrichment analysis, such as Methylated DNA Immunoprecipitation (MeDIP), selectively isolates methylated DNA using antibodies. MeDIP

results, analyzed through various methods, may have limitations, including the need to assess antibody specificity and potential bias towards hypermethylated regions, impacting result accuracy. (Weber et al. 2005). Methyl-CpG binding domain-based capture (MBDCap) is another potent technique in molecular biology for the isolation of methylated DNA fragments (Brinkman et al. 2010). This method leverages methylation-affinity proteins to capture methylated DNA fragments. Initially, the DNA is fragmented through sonication and subsequently incubated with specific MBD proteins labeled with tags. These tagged MBD proteins are then precipitated using antibody-conjugated beads that specifically target the protein tags (Nair et al. 2011). MBDCap, when combined with next-generation sequencing, has the potential to offer comprehensive coverage of 5mC in densely CpG-methylated and repeat regions. However, it is important to recognize certain limitations. MBDCap tends to show bias towards hypermethylated regions and does not achieve the same single-base resolution as whole-genome bisulfite sequencing (WGBS) (Bock et al. 2010).

Molecular methods for DNA methylation analysis involve bioinformatic analysis, which processes and interprets sequencing or microarray data. Despite being a powerful tool for understanding DNA methylation dynamics, bioinformatic analysis presents challenges, including the complexity of selecting suitable tools and the requirement for proficient data analysts experienced in processing and statistically analyzing large datasets. This underscores the importance of expertise in both molecular biology and computational analysis to effectively unravel DNA methylation intricacies in diverse biological contexts (Rauluseviciute, Drabløs, and Rye 2019). The choice of a molecular method for DNA methylation analysis significantly influences the quantitative nature of the data. Different methods, such as MeDIP-seq and WGBS, provide distinct levels of resolution in methylation information. MeDIP-seq focuses on comparing the relative abundance of methylated fragments, while WGBS offers individual cytosine base-level resolution, enabling more detailed analyses like statistical testing of

differential methylation, often focusing on differentially methylated regions (DMRs) [(Laird 2010)](). WGBS is valuable, but it requires bisulfite conversion to distinguish methylated and non-methylated cytosines. However, this process can lead to DNA degradation, necessitating meticulous purification steps to eliminate residual sodium bisulfite and minimize the risk of introducing artifacts during subsequent sequencing. [(Ortega-Recalde et al. 2021)](). Attention to detail in sample preparation is crucial for accurate and reliable methylation data from WGBS. A challenge in bisulfite sequencing involves converting unmethylated cytosines to thymidines, reducing sequence complexity and making accurate read alignment challenging. Ensuring precise alignment of bisulfite-converted reads is critical for accurately identifying methylated cytosines. Proper alignment algorithms and software tools are essential to mitigate this challenge and obtain reliable DNA methylation data in bisulfite sequencing experiments. [(Meissner et al. 2008)](). The popularity of WGBS has led to the emergence of numerous bioinformatics tools for managing large-scale sequencing data. The typical WGBS DNA methylation analysis pipeline involves critical steps such as read alignment, quality control, methylation information extraction, identification of DMRs), and downstream analyses integrating methylation data with gene expression, histone modification, or TF binding site annotation. Widely used tools, including FastQC, facilitate initial quality control by generating comprehensive reports to assess sequencing read quality [(Andrews and Others 2010)](). FastQ Screen is employed to check the contamination and origin of sequences in a fastq file. It helps identify potential issues such as contamination from other species or sample cross-contamination [(Wingett and Andrews 2018)](). Trim Galore is a tool designed for quality control and trimming of sequencing data. It can be particularly useful for removing low-quality sequences and adapter sequences from the raw data [("Babraham Bioinformatics - Trim Galore!")](). Trimmomatic is another popular tool for read trimming and quality control. It allows users to remove adapters and low-quality bases, ensuring that the input data is of high quality

before downstream analysis (Bolger, Lohse, and Usadel 2014). Tools like FastQC are crucial for maintaining the quality and integrity of DNA methylation data obtained through WGBS, ensuring result reliability in the bioinformatics pipeline for analysis. (Gong et al. 2022). Mapping bisulfite-treated sequences to a reference genome is computationally demanding due to altered DNA complexity and the challenge of decoding methylation states. Bisulfite treatment reduces sequence complexity by converting unmethylated cytosines to thymines. This increased degeneracy complicates mapping reads, and considering both Watson and Crick strands further adds complexity, resulting in up to four possible DNA strands for each genomic position. Accurate mapping requires considering all strands to determine the true origin of bisulfite-treated reads. (M. Liu and Xu 2021). Bisulfite-treated DNA sequences differ from standard DNA sequencing, representing a mix of methylated and unmethylated cytosines at specific positions. This variability complicates mapping, addressed by sophisticated algorithms and tools developed to accurately map bisulfite-treated sequences to a reference genome. These tools consider bisulfite treatment complexities and methylation variability, enabling reliable determination of methylation patterns across the genome. Advanced computational methods are crucial for studying DNA methylation in diverse biological processes (H. Li and Homer 2010). To align reads from WGBS to reference genomes, specialized mapping tools like Bismark and Bowtie2 are essential. These tools are designed to handle the unique characteristics of bisulfite-converted DNA, accurately processing converted unmethylated cytosine bases within reads. By using converted reference genomes accounting for bisulfite treatment, these tools precisely align WGBS reads, facilitating downstream analyses and comprehensive exploration of DNA methylation. Bismark, a widely used software for bisulfite-sequencing data analysis, distinguishes between methylated and unmethylated cytosines, providing essential functionalities like methylation extraction and downstream analysis of differentially methylated regions. (Krueger and Andrews 2011). Bowtie2, a versatile and efficient sequence alignment

tool, can be adapted for bisulfite sequencing analysis. Despite not being specifically designed for bisulfite-converted DNA, its settings can be adjusted to handle the unique nature of bisulfite-sequenced reads. Its speed and flexibility make it a valuable choice, especially for large-scale WGBS datasets (Langmead and Salzberg 2012). These tools, utilizing converted reference genome sequences, enable precise alignment of WGBS reads, forming the foundation for downstream analyses. Essential for researchers, they explore genome epigenetic modifications, revealing regulatory mechanisms influencing gene expression and cellular function (H. Kim et al. 2022).

In DNA methylation analysis, pivotal tools like Bismark, MethylDackel, methylKit, and BSmooth play a crucial role in methylation calling. These tools meticulously assess cytosine methylation status across the genome, calculating the degree of methylation for all mapped cytosines in sequenced data. MethylDackel, working in harmony with Bismark, aligns reads to the reference genome, determining the methylation status of each cytosine (Ryan 2017). MethylDackel excels in scrutinizing methylation patterns in large-scale bisulfite sequencing datasets. Going beyond mere calling, methylKit, an R package, broadens the scope of methylation analysis by enabling researchers to uncover differentially methylated cytosines and regions (DMCs and DMRs) (Akalin et al. 2012). Empowering scientists to identify significant methylation changes, methylKit, and BSmooth, both R packages, go beyond basic calling. They enable the detection of DMCs and DMRs, crucial for understanding the functional implications of DNA methylation fluctuations. (Hansen, Langmead, and Irizarry 2012). Tools like BSmooth analyze bisulfite sequencing datasets, detecting methylation pattern shifts and uncovering hypomethylated regions. Foundational in DNA methylation analysis, these tools rely on disparities between methylated and unmethylated cytosines. BSmooth's adaptability allows tailored analyses for specific objectives and domains, crucial in cancer research for customizing parameters to account for cancer-specific DNA methylation patterns. Researchers leverage

them for deeper insights into DNA methylation, advancing our understanding of complex phenomena, including cancer.

Table 6. Methods and Tools for 5-Methylcytosine Quantification in Epigenetic Research

Methodology	Techniques/Tools	Strengths	Limitations
Molecular Biology	Bisulfite Treatment	Distinguishes methylated from non-methylated cytosines	Potential DNA degradation, bias in amplification
Molecular Biology	WGBS	Accurate detection of methylated cytosines, comprehensive	Large input material, new protocols, potential bias in amplification
Molecular Biology	MeDIP	Selectively isolates methylated DNA, next-gen sequencing	Potential antibody bias, bias towards hypermethylated regions
Molecular Biology	MBDCap	Captures methylated DNA fragments, next-gen sequencing	Bias towards hypermethylated regions, lower resolution
Data Analysis	WGBS Analysis Tools	Precise alignment, methylation calling, identification of DMRs	Need for thorough sample preparation, potential DNA degradation
Alignment Tools for WGBS	Bismark, Bowtie2	Accurate alignment, distinguish methylated from unmethylated cytosines	Need for converted reference genome, computational demands

1.13 Databases for pioneer transcription factors

TF databases serve diverse molecular biology and genetics research needs. They fall into categories: some provide comprehensive information on TFs, including sequences, structures, modifications, and expressions. Others focus on documenting binding sites, offering insights into how TFs interact with DNA. Some catalog target genes, revealing downstream consequences and network complexities. Certain databases provide functional annotations, enriching understanding of TFs in various contexts. Some offer predictive tools for identifying

binding sites, aiding gene regulation investigations. Database choice depends on research needs, often using multiple for a comprehensive understanding of TFs, binding sites, and target genes.

Table 7. Transcription Factor Databases

Database Name	Purpose and Features	Special Features and Updates
MethMotif	Repository of two-dimensional TF motifs with CpG methylation data from ChIP-seq and WGBS datasets.	Over 700 PWMs, cell type-specific perspective, integrates pioneer factor information (2024 version).
JASPAR	Compilation of experimentally determined TF binding profiles (PWMs) with a broad scope covering various species.	Expansive collection of PWMs, inclusivity for a wide spectrum of species, accuracy and reliability from empirical data (ChIP-seq).
TcoF-DB	Specialized repository focused on human transcription co-factors, their interactions with TFs, and genomic information about TF binding sites.	Comprehensive information on human co-factors, detailed annotations, integration with external databases, expanded content in TcoF-DB v2 (2024 version).
TRANSFAC	Bioinformatics tool for systematic collection of data on TFs and their binding sites across various organisms.	Versatility with data from diverse organisms, matrix profiles for predicting binding sites, in-depth functional annotations, interconnected with external databases.
TFSyntax	Database dedicated to mapping TF binding sites within human and mouse genomes, offering comprehensive characterizations of TF binding syntax.	Mapping of binding motifs for 1299 human TFs and 890 mouse TFs across 382 cells and tissues, detailed information on motif positional preferences, density, and co-localization.

1.13.1 Methmotif

The MethMotif database serves as a repository of two-dimensional TF motifs, which encompass TF binding sites (TFBS) and associated position weight matrices (PWMs) (Xuan Lin et al. 2019). MethMotif, a unique database, combines CpG methylation data from ChIP-seq and whole-genome bisulfite sequencing datasets, offering a cell type-specific perspective. Integrating TFBS motifs with TFBS DNA methylation patterns, it provides a comprehensive

portrayal of DNA loci recognized by TFs. With over 700 PWMs spanning species like human, mouse, and Arabidopsis thaliana, MethMotif aids researchers in studying TF binding, gene regulation, and the impact of DNA methylation. Its focus on cell type specificity and epigenetic integration makes it invaluable for understanding transcriptional regulation in diverse biological contexts (Dyer et al. 2022), MethMotif has advanced by incorporating pioneer factor information from a comprehensive meta-analysis. This enhancement provides valuable insights into the role of pioneer factors in transcriptional regulation and their interactions with DNA methylation patterns.

1.13.2 JASPAR

JASPAR stands as a valuable resource, offering an extensive compilation of experimentally determined TF binding profiles, commonly referred to as PWMs (Castro-Mondragon et al. 2022). JASPAR's expansive collection of PWMs reveals the specific binding preferences of various TFs for distinct DNA sequences, providing detailed insights into gene regulation. Notably inclusive, JASPAR offers binding profiles for a broad range of species, making it a valuable resource for researching TF binding in diverse organisms. In essence, JASPAR stands as a comprehensive repository, supporting scientific inquiry into gene regulation across species with its extensive PWM collection.

1.13.3 TcoF-DB

The Transcription Co-Factors Database (TcoF-DB), initially developed by Schaefer and colleagues in 2011 (Schaefer, Schmeier, and Bajic 2011) and later updated to TcoF-DB v2 by Schmeier and team in 2017 (Schmeier et al. 2017), serves as a comprehensive and specialized repository of information specifically focused on human transcription co-factors and their interactions with TFs. TcoF-DB is a valuable resource that compiles comprehensive data on transcription co-factors, including co-activators, co-repressors, and regulatory proteins. It provides detailed information on the interactions between TFs and their co-factors, shedding

light on complex gene regulatory networks. The database includes annotations, descriptions, and genomic information on TF binding sites, crucial for understanding gene expression regulation. TcoF-DB also serves as a potential resource for investigating links between co-factors and diseases. In its updated version, TcoF-DB v2 offers expanded content, integrating data from Gene Ontology, biological pathways, diseases, and molecular signatures. With 958 human TcoFs and additional entries for mouse co-factors, TcoF-DB is a centralized repository enhancing research on gene regulation mechanisms and their implications in health and disease.

1.13.4 TRANSFAC

TRANSFAC (Matys et al. 2003), an acronym for "TRANScription FACtor database," serves as a specialized repository for the systematic collection and arrangement of data pertaining to transcription factors (TFs) and their corresponding binding sites across a wide spectrum of organisms. Within TRANSFAC, researchers access comprehensive information about TFs, including details about their structural characteristics, functional attributes, and DNA-binding domains (Matys et al. 2006). TRANSFAC is a versatile database with comprehensive data from diverse organisms, ideal for comparative genomics. Researchers use TRANSFAC to uncover regulatory motifs, predict binding sites with matrix profiles, and explore functional annotations for TFs, including their roles in gene regulation and associations with diseases. Its interconnectivity with other databases allows seamless data integration, facilitating exploration of complex networks in transcriptional regulation and gene function, advancing knowledge in molecular biology, genetics, and genomics.

1.13.5 TFSyntax

TFSyntax is an extensive database that is dedicated to the intricate arrangement of TF binding sites within the human and mouse genomes (Yongbing Zhao 2023). TFSyntax plays a pivotal role in mapping the binding motifs of 1299 human TFs and 890 mouse TFs across 382 distinct cells and tissues. This comprehensive effort positions it as one of the most exhaustive TF

binding resources available. Beyond identifying TF binding sites, TFSyntax defines motif positional preferences, density, and co-localization within accessible genomic elements. Its user-friendly interface and functional modules allow researchers to search, browse, analyze, and download data, making it invaluable for studying transcriptional regulation and the impact of regulatory DNA variants on diseases. TFSyntax serves as a pivotal resource for investigating transcriptional regulation mechanisms and understanding how variations in regulatory DNA elements contribute to diseases.

1.14 The Need for Classification

PTFs play a crucial role in regulating gene expression and directing cellular behavior across various biological processes, including development, differentiation, and responses to environmental stimuli. The classification of these pioneering factors holds significant importance for several reasons. Firstly, pTFs act as initiator of cellular differentiation and maintaining cellular identity. By categorizing these factors, we gain valuable insights into the mechanisms involved in establishing and preserving distinct cell types during embryonic development and tissue regeneration. Moreover, understanding the classification and functions of pTFs provides us with opportunities for therapeutic interventions. This knowledge enables the development of strategies to precisely control and manipulate cell fate, offering potential applications in regenerative medicine, disease therapeutics, and tissue engineering. Notably, pTFs play a critical role in unpacking tightly condensed chromatin, making target genes accessible to other regulatory factors. Deepening our understanding of their classification helps unravel the complex landscape of gene regulation and the coordinated action of specific TFs. PTFs are also pivotal in reprogramming somatic cells into induced pluripotent stem cells (iPSCs), thereby advancing stem cell research and its medical applications. Furthermore, these factors are closely linked to cellular responses to environmental changes, including stressors

and signaling molecules. Developing the classification of pTFs enhances our comprehension of how cells adapt and respond to their surroundings, which is crucial in fields such as cancer biology and immunology. Additionally, comparative analyses of these factors across species shed light on the evolution of gene regulation and its role in species-specific traits. In biotechnology, pTFs can be utilized to produce custom proteins or metabolites. A detailed understanding of their classification and functions can accelerate the development of more efficient bioprocesses for industrial and medical purposes. Researchers can design pTFs to precisely control the timing and level of gene expression. This customization allows for the fine-tuning of metabolic pathways and the production of proteins or metabolites in specific quantities. Conversely, disruptions in pTF activity can lead to various diseases, including cancer. Therefore, classifying these factors is essential for uncovering disease mechanisms and identifying potential therapeutic targets.

In summary, the classification of pTFs plays a pivotal role in unraveling the complexities of gene regulation, cellular differentiation, and responses to environmental cues.

II. Materials and Methods

2.1 Data Collection and Compilation

Data Collection

Known Pioneer TF

YES N/A

NO

Excluded from analysis

Database Compilation

- TF's Name
- TF's Family
- Ability to bind methylated DNA
- Ability to bind heterochromatin
- Pioneer status
- Promotes DNA accessibility
- Cancer participation status
- Cell fate alterations ability

YES/NA/NO

Heatmap clustering

Ability to bind methylated DNA

NO

YES

GROUP **1**

GROUP **2**

Ability to bind heterochromatin

NO

YES

GROUP 1**A**

GROUP 1**B**

GROUP 2**A**

GROUP 1**B**

Molecular Functions

Biological Processes

Cellular Components

Gene Ontology Enrichment Analysis

DAVID

Pathway analysis

Reactome database

Gene expression analysis

Baseline Differential

Figure 4. Systematic Methodology Overview

This formal flowchart outlines our methodology, utilizing dark blue squares for main analysis steps, light blue figures for decision points and supplementary information, and grey squares for newly created TF clusters. Positive outcomes are denoted by green circles, while negative results are represented by red circles. The figure was created in Inkscape (https://inkscape.org/) and SciDraw (https://scidraw.io/).

In order to create a comprehensive classification system for TFs, this research project underwent an extensive data collection process. The initial step in this data collection process involved a thorough review of existing literature. This review aimed to identify relevant studies, research articles, and databases containing information about TFs and their associated features. A paper manager tool, Paperpile (https://paperpile.com/app), was utilized to track literature specifically mentioning the pioneer status of TFs. Key sources included Google Scholar, PubMed, and Google Search. We utilized specific search terms such as "TFs," "DNA methylation," "TF binding sites," and "methylated DNA". Our collected data points included the TF's name or symbol, NCBI gene symbol or protein ID, the organism or species it is associated with, details about its DNA-binding domain, its regulatory functions and target genes, its ability to bind to heterochromatin, its interaction with methylated DNA, its capacity to enhance DNA accessibility, its potential influence on cell fate, its classification as a pioneer factor, and any established roles in cancer. Data for these criteria was recorded using one of the following options: 'YES,' 'NO,' or 'N/A'. Any conflicted data was excluded from analysis. To streamline the extraction process, a structured data extraction template was designed and refined based on the data points identified. Data were extracted from both primary and secondary sources to ensure the comprehensiveness of the dataset. Data preprocessing was a crucial step in ensuring the quality and consistency of the collected information. During this phase, we carried out data validation, ensuring the accuracy of the information through cross-

referencing and verification of TF names and associated data. After this, we cleaned the data by identifying and rectifying errors, inconsistencies, and missing values in the dataset, which were excluded from analysis. We next standardized the format and representation of data to ensure uniformity. Next, a data integration step included merging information from various sources to create a single, cohesive dataset. Following data preprocessing, the curated dataset was subjected to further curation to enhance its utility. This involved the elimination of duplicate entries, the consolidation of data with synonyms or different naming conventions, and the assignment of unique identifiers to each TF. The aim was to create a clean, well-organized dataset that could serve as the basis for classification. The collected and curated data were integrated into a comprehensive TF database, which served as the primary resource for developing the classification model. This integrated dataset was designed to be easily accessible and queryable, allowing for efficient retrieval and analysis of TF information. The data collection and preprocessing phases were fundamental in creating a robust foundation for the pTF classification system. By meticulously compiling and organizing TF data from existing literature, we established a comprehensive dataset that would support subsequent stages of this research. The next section will delve into the methodology for feature selection and model development, leveraging this well-prepared TF dataset to construct a novel classification system.

2.2 Feature Grouping and Categories

2.2.1 DNA Methylation Binding

The interaction between TFs and DNA, particularly in the context of DNA methylation, plays a crucial role in gene regulation. To ensure that our research was both comprehensive and current, we adopted a dual-pronged approach, combining a thorough literature review with cross-referencing information from the MethMotif database. We harnessed the MethMotif

database to augment our comprehension of TF-DNA methylation interactions. MethMotif is a specialized database that houses a comprehensive collection of experimentally validated binding motifs for TFs, with a particular focus on their preferences for methylated or unmethylated DNA sequences. To corroborate our findings from the literature review, we extracted data from MethMotif, emphasizing TFs known to exhibit specific binding preferences for methylated DNA. Additionally, we gathered information regarding the DNA sequences to which these TFs bind when DNA is methylated. The data sourced from the literature review and MethMotif was integrated into our database. This enabled us to conduct a comprehensive analysis of the collective evidence from both sources, allowing us to identify commonalities, discrepancies, and emerging trends.

2.2.2 Heterochromatin Binding

Heterochromatin binding refers to the ability of a TF to interact with and regulate genes within heterochromatic regions of the genome. Through an extensive literature review, we have evaluated each TF's capacity to bind heterochromatin and classified them as either "yes" if they exhibit this function, "no" if they do not, or "n/a" if data on their heterochromatin binding is lacking.

2.2.3 DNA Accessibility Promotion

DNA accessibility promotion represents TF's role in enhancing chromatin accessibility, allowing other regulatory elements to bind. The database includes the assessment of each TF for this function, with a "yes" indicating its promotion of DNA accessibility, "no" if it does not, and "n/a" for unreported cases.

2.2.4 Cell Fate Alterations

Cell fate alterations involve TFs driving changes in cellular identity and differentiation. The literature review categorizes TFs as "yes" if they are associated with cell fate alterations, "no" if there is no such association, and "n/a" for instances with insufficient data.

2.2.5 Cancer

The cancer-related function of TFs is a critical aspect of this analysis. We have classified TFs as "yes" if they are known to play a role in cancer development, "no" if they do not, and "n/a" when data is lacking.

2.2.6 Identification of PTFs

For each PTFs identified, specific information regarding the genes they regulate and the context in which their PTFs status was established was recorded. The compilation of PTFs in the database allowed for a comprehensive examination of their characteristics and the evidence supporting their PTFs status. PTFs are essential in understanding gene regulation in various cellular contexts, including development, differentiation, and disease.

2.4 Heatmap Analysis

To visually represent our comprehensive analysis of multidimensional data, we used the pheatmap package (Kolde 2019), version 1.0.12, in the R Studio software, version 1.1.463. The clustering and arrangement of TFs within the heatmap considered a diverse set of criteria. These criteria encompassed the TF's propensity to bind methylated DNA, its interaction capabilities with closed chromatin regions, its role in modulating DNA accessibility, its influence on cell programming or fate alterations, its association with cancer, and its potential function as a pioneer factor in cellular processes. This systematic evaluation enabled the extraction of

meaningful patterns and relationships embedded within the intricate landscape of TF characteristics.

2.5 Gene Ontology (GO) Enrichment Analysis

Gene Ontology (GO) is a standardized system for the functional annotation of genes and their products. To gain a deeper insight into the distinctions among recently identified TF groups, we conducted a comprehensive analysis employing the Gene Ontology (GO) enrichment approach. Our primary objective was to unveil the specific functionalities associated with each TF group, elucidating the molecular intricacies that render them distinct.

For this detailed analysis, we utilized the bioinformatics tool Database for Annotation, Visualization, and Integrated Discovery (DAVID), version 2023q2 (Sherman et al. 2022). For the DAVID GO analysis, we employed a curated list of TFs of interest, specifically utilizing their NCBI gene symbols. Our search was tailored to the species of interest, Homo sapiens, with the 9606 accession code specified.

The GO term enrichment analysis yielded a comprehensive list of the most relevant GO terms associated with the provided gene list. This list includes essential information such as "Count," indicating the number of genes encompassed within the specific term out of the total searched genes. Additionally, the derived "Percentage" represents the proportion of both the overall searched genes and those involved in the respective GO term. This percentage offers insight into the relative significance of each GO term within the context of the entire gene set under consideration. Furthermore, the result list featured Fold Enrichment values, representing the ratio of genes within our list to those associated with a specific pathway. This Fold Enrichment metric provides insight into the significance of the genes in our list relative to the broader context of the pathway (T. Zhou, Yao, and Liu 2017).

The statistical values supporting the GO analysis encompassed several crucial metrics. The EASE Score, denoting the modified one-tail Fisher Exact P-value, was employed to assess gene

enrichment in annotation terms. Additionally, the adjusted p-values were provided, including the Bonferroni, Benjamini, and FDR (false discovery rate) values (Akey et al. 2007). The FDR, or false discovery rate, in DAVID employs adaptive linear step-up adjusted p-values for approximate control over the false discovery rate (Reiner, Yekutieli, and Benjamini 2003). Significantly, we carried out distinct GO enrichment analyses for each TF group and its subgroups. In the final step of our work with DAVID, we exported the data resulting from our search. This data exportation serves as a crucial step in consolidating the outcomes of our analyses, facilitating further examination, interpretation, and integration with additional research or datasets.

The outcomes of this analysis were visually represented using the R Studio software, version 1.1.463, taking advantage of its robust data visualization capabilities, facilitating the intuitive interpretation of the enriched GO terms associated with each TF group. This not only helped us interpret the findings comprehensively but also generated visually engaging representations of the enriched GO terms.

Following the acquisition of separate datasets for each group and subgroup, we consolidated them into a compiled dataset. This integrated dataset was then visualized through a scatter plot using the ggplot2 package, version 3.4.2 (Wickham, Chang, and Wickham 2016). The figures presented two key visual representations of the results. Firstly, the size of the dots conveyed the percentage of DAVID genes in the list associated with a specific annotation term for each category. Larger dots indicated a higher percentage of genes linked to the respective annotation term. Secondly, a color gradient ranging from dark blue to light yellow was employed to represent the -log10 false discovery rate (FDR). This color-coded approach provided a visual indication of the statistical significance of the results, with darker colors denoting lower -log10 FDR values and, therefore, lower significance.

2.6 Pathway Analysis

Pathway analysis is a bioinformatics method used in the interpretation of high-throughput biological data, such as gene expression or protein interaction data. The goal of pathway analysis is to gain insights into the biological processes and molecular interactions that are relevant to a particular set of genes or proteins. Pathway analysis was conducted employing the Reactome database (Gillespie et al. 2022). The ensuing results are meticulously presented in dedicated tables for each group and subgroup, with a deliberate focus on spotlighting the top 25 pathways deemed most significant. These pathways are thoughtfully organized according to their False Discovery Rate (FDR), offering enhanced clarity and a strategic prioritization of relevance.

The decision to employ FDR as a sorting criterion was intentional, serving as a cornerstone for a robust and statistically informed presentation of the pathway analysis outcomes. This meticulous approach contributes significantly to the precision and interpretability of the results, ensuring that the identified pathways not only bear relevance but also carry a statistically validated weight, thereby enriching our comprehension of the biological contexts intertwined with the analyzed gene or protein sets.

Following the creation of distinct datasets for each designated group and subgroup, we integrated them to form a consolidated dataset. Subsequently, this unified dataset underwent visualization using a scatter plot implemented in the ggplot2 package, version 3.4.2 (Wickham, Chang, and Wickham 2016).

2.7 Gene Expression Analysis

Pathway analysis was conducted utilizing the Expression Atlas database (Papatheodorou et al. 2020). Through the meticulous querying of our predefined group/subgroup pTFs subset, we compiled an exhaustive list detailing the specific experiments, tissues, and cell types wherein

the gene of interest is expressed. Furthermore, we have discerned the conditions under which this gene functions as a marker for both baseline and differential expression profiles.

Baseline expression, denoting the standard level of gene or protein expression under specific conditions, serves as the default observed in a standard or control group. This definition of baseline expression establishes a reference point against which changes in expression can be effectively compared. Our baseline expression data is sourced from three distinct projects, employing a 0.5 Transcripts Per Million (TPM) cutoff (Yingdong Zhao et al. 2021). TPM, representing the relative abundance of each transcript in a sample, ensures comparability across diverse samples. The datasets were derived from three projects: first, RNA-Seq data from individual human tissues and a composite of 16 tissues, known as the Illumina Body Map (Goldstein et al. 2016); second, the Genotype-Tissue Expression (GTEx) project encompassing RNA-seq data from 53 human tissue samples (GTEx Consortium 2015); and third, the RIKEN FANTOM5 project employing RNA-Seq CAGE (Cap Analysis of Gene Expression) analysis of human tissues (Lizio et al. 2019). All results obtained were visualized in R using the ggplot2 package, version 3.4.2 (Wickham, Chang, and Wickham 2016), with a specific focus on illustrating the TPM scale.

Differential expression, signifying alterations in gene or protein expression levels between two or more experimental conditions or groups, brings attention to those genes or proteins experiencing significant changes in response to specific factors like diseases, treatments, or environmental variations. To investigate the distinctive patterns of differential expression for our genes of interest, we utilized the 'differential expression' function within the Expression Atlas. This function provides pairwise comparisons where the genes of interest exhibit significant differential expression with an adjusted p-value < 0.05 and a log2 fold-change > 1. The obtained results were visually represented in R using the ggplot2 package, version 3.4.2, (Wickham, Chang, and Wickham 2016).

III. Results

3.1 Pioneer Transcription Factor Classification

3.1.1. Data Collection and Compilation

Our data collection process focused on gathering the TF's name or symbol, its affinity for heterochromatin binding, its interaction with methylated DNA, its capability to enhance DNA accessibility, its potential impact on cell fate, its classification as a pioneer factor, and any established roles in cancer. The data for these criteria were documented using one of the following response options: 'YES,' 'NO,' or 'NA.' Through a rigorous compilation and organization of TF data extracted from pertinent literature sources, we successfully refined our database to encompass 61 TFs. This curation was designed specifically to identify the pioneer status of a TF, as indicated by the responses 'YES' or 'NA' (Table 8). The resulting dataset provides a comprehensive overview of TFs with notable implications for their role in influencing DNA dynamics, cellular fate, and potential involvement in cancer, thus contributing to a more nuanced understanding of their functional relevance.

Table 8. Database of 61 Transcription Factors

Name of TF	Family	Binds methylated DNA	Binds to closed chromatin	Promotes DNA accessibility	Cell fate alterations	Cancer	Pioneer
Ascl1	Basic helix-loop-helix factors	NO	YES	YES	YES	YES	YES
CEBPA	Basic leucine zipper factors	YES	YES	YES	YES	YES	YES
CEBPB	Basic leucine zipper factors	YES	YES	YES	YES	YES	YES
CEBPD	Basic leucine zipper factors	YES	YES	YES	YES	YES	YES
CEBPG	Basic leucine zipper factors	YES	YES	YES	YES	YES	YES
DUX4	Homeobox	NO	YES	YES	YES	YES	YES
EBF1	Rel homology region factors	YES	NO	YES	YES	YES	YES
FLI1	Tryptophan cluster factors	NA	NO	YES	YES	YES	NA
FoxA1	Fork head factors	NO	YES	YES	YES	YES	YES
FoxA2	Fork head factors	NO	NO	YES	YES	YES	YES
FoxA3	Fork head factors	NO	NO	YES	YES	YES	YES
FoxC1	Fork head factors	NO	NO	YES	YES	YES	YES
FOXD3	Fork head factors	NO	YES	YES	YES	YES	YES
FoxL1	Fork head factors	NO	NA	YES	YES	YES	NA

FoxM1	Fork head factors	NO	NA	YES	YES	YES	NA
FoxO1	Fork head factors	NO	YES	YES	YES	YES	YES
FoxO3	Fork head factors	NO	NA	YES	YES	YES	NA
FoxH1	Fork head factors	YES	YES	YES	YES	YES	YES
FoxP1	Fork head factors	NO	NA	YES	YES	YES	NA
FoxK1	Fork head factors	NO	NA	YES	YES	YES	NA
FoxK2	Fork head factors	YES	YES	YES	YES	YES	YES
FoxQ1	Fork head factors	NO	NA	YES	YES	YES	NA
GATA1	Other C4 zinc finger-type factors	NO	YES	YES	YES	YES	YES
GATA2	Other C4 zinc finger-type factors	NO	YES	YES	YES	YES	YES
GATA3	Other C4 zinc finger-type factors	NO	YES	YES	YES	YES	YES
GATA4	Other C4 zinc finger-type factors	YES	YES	YES	YES	YES	YES
GATA5	Other C4 zinc finger-type factors	NO	NA	NA	YES	YES	NA
GATA6	Other C4 zinc finger-type factors	NO	NA	YES	YES	YES	YES
GRHL2	Grainyhead domain factors	YES	YES	YES	YES	YES	YES

Klf1	C2H2 zinc finger factors	YES	NO	NA	YES	YES	NA
Klf4	C2H2 zinc finger factors	YES	YES	YES	YES	YES	YES
Klf5	C2H2 zinc finger factors	YES	NA	YES	YES	YES	NA
KLf9	C2H2 zinc finger factors	YES	NA	NA	YES	YES	YES
Klf10	C2H2 zinc finger factors	YES	NA	YES	YES	YES	NA
Klf13	C2H2 zinc finger factors	YES	NA	NA	YES	YES	NA
MAFK	Basic leucine zipper factors	YES	NA	YES	YES	YES	NA
MEF2A	MADS box factors	NO	NO	YES	YES	YES	NA
Myc	Basic helix-loop-helix factors	NO	NO	YES	YES	YES	YES
NANOG	Homeo domain factors	NO	NO	YES	YES	YES	YES
Oct4	POU-homeodomain family	NO	YES	YES	YES	YES	YES
Pax5	Paired Box	NO	NA	YES	YES	YES	NA
Pax7	Paired Box	NO	YES	YES	YES	YES	YES
PBX1	Homeo domain factors	NA	YES	YES	YES	YES	YES
PBX2	Homeo domain factors	NA	NA	NA	YES	YES	NA
PBX3	Homeo domain factors	NA	NA	NA	YES	YES	NA

|							
RUNX1	Runt domain factors	NO	YES	YES	YES	YES	YES
RUNX3	Runt domain factors	NO	YES	YES	YES	YES	YES
Sox2	High-mobility group	NO	YES	YES	YES	YES	YES
Sox5	High-mobility group	NO	NA	YES	YES	YES	NA
Sox6	High-mobility group	NO	NA	YES	YES	YES	NA
Sox9	High-mobility group	NO	YES	YES	YES	YES	YES
SPI1	Tryptophan cluster factors	NA	YES	YES	YES	YES	YES
TCF1	Basic helix-loop-helix factors	YESs	YES	YES	YES	YES	YES
TCF3	Basic helix-loop-helix factors	NO	NA	YES	YES	YES	YES
TFAP2 A/B/C	Basic helix-span-helix factors	YES	NA	YES	YES	YES	YES
HOXB13	Homeobox	NO	NA	YES	YES	YES	YES
NEUROD1	Basic helix-loop-helix factors	NO	YES	YES	YES	YES	YES
Etv2	Ets-related	NA	NA	YES	YES	YES	YES
NF-Y A/B/C	NFY	NA	YES	YES	YES	YES	YES
DMRT1	DMRT	NO	YES	YES	YES	YES	YES

Table 9. Database of 61 Transcription Factors

The database includes the transcription factor's name or symbol, ability to bind to heterochromatin, interaction with methylated DNA, capacity to enhance DNA accessibility, potential influence on cell fate, classification as a pioneer factor, and established roles in cancer. The data for the last six criteria is visually represented using the options 'YES,' 'NO,' or 'NA'.

3.1.2. Heatmap Analysis

Our heatmap analysis resulted in the identification of two distinct groups of TFs, each further delineated into two subgroups (Figure 5). This classification was based on two crucial categories: 'DNA methylation binding' and 'heterochromatin binding.' Our heatmap analysis helped us to discerned two distinct clusters of TFs, each further stratified into two subgroups. This taxonomic categorization was derived from the critical examination of two pivotal attributes: 'DNA methylation binding' and 'heterochromatin binding.' In Group 1, we identified transcription factors that exhibited a conspicuous absence of binding to 5-methylcytosine (5-mC). Subsequently, we subdivided Group 1 into two discernible classes: 1A, comprising pTFs that do not bind 5-mC and also do not interact with closed chromatin, and 1B, comprising pTFs that do not bind 5-mC but engage with closed chromatin. On the other hand, Group 2 encompassed pTFs that demonstrated a clear propensity for binding to 5-mC. Further refinement of Group 2 yielded two distinctive subclasses: 2A, characterized by pTFs binding to 5-mC but refraining from interaction with heterochromatin, and 2B, featuring pTFs that bind both 5-mC and heterochromatin. This comprehensive classification provides a nuanced understanding of the varied binding patterns exhibited by transcription factors in relation to DNA methylation and heterochromatin, shedding light on the intricate regulatory mechanisms at play.
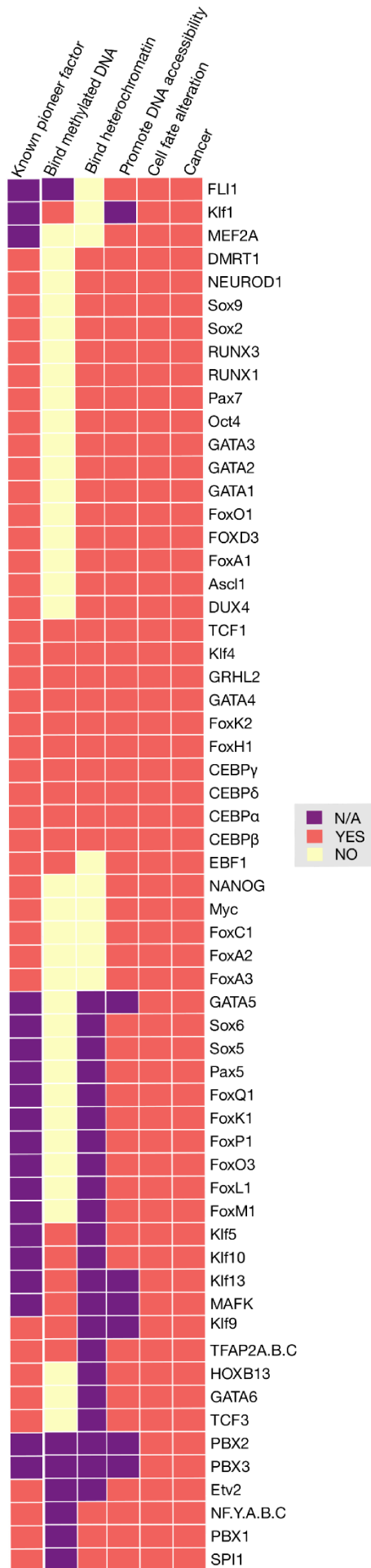
**Figure 5. Comprehensive Heatmap Analysis of Transcription Factors**

The figure depicts the clustering of TFs exhibiting pioneer features. This clustering was determined by evaluating various criteria, such as the TF's ability to bind methylated DNA, its capacity to bind closed chromatin regions, its role in enhancing DNA accessibility, its involvement in cell programming or fate alterations, its association with cancer, and whether it functions as a pioneer factor. In the figure, red represents a positive response ('yes') for a particular criterion, while yellow represents a negative response ('no'). When there is no available information in the literature regarding a specific criterion, the corresponding cell is depicted as violet.

3.2. Gene Ontology (GO) Enrichment Analysis

3.2.1. Analysis of Cellular Component (CC)

The visualization of Gene Ontology (GO) analysis focusing on the cellular component revealed common terms across all analyzed gene groups, including 'nucleus,' 'chromatin,' and 'TF complex.' This consistency is in line with the known characterization of the analyzed genes as pioneer TFs. Specifically, the term 'C/EBP complex' emerged as particularly relevant to group 2, aligning with the presence of TFs from the C/EBP family, as illustrated in Figure 6. The terms 'Golgi apparatus' and 'nucleolus,' were specific to group 1.
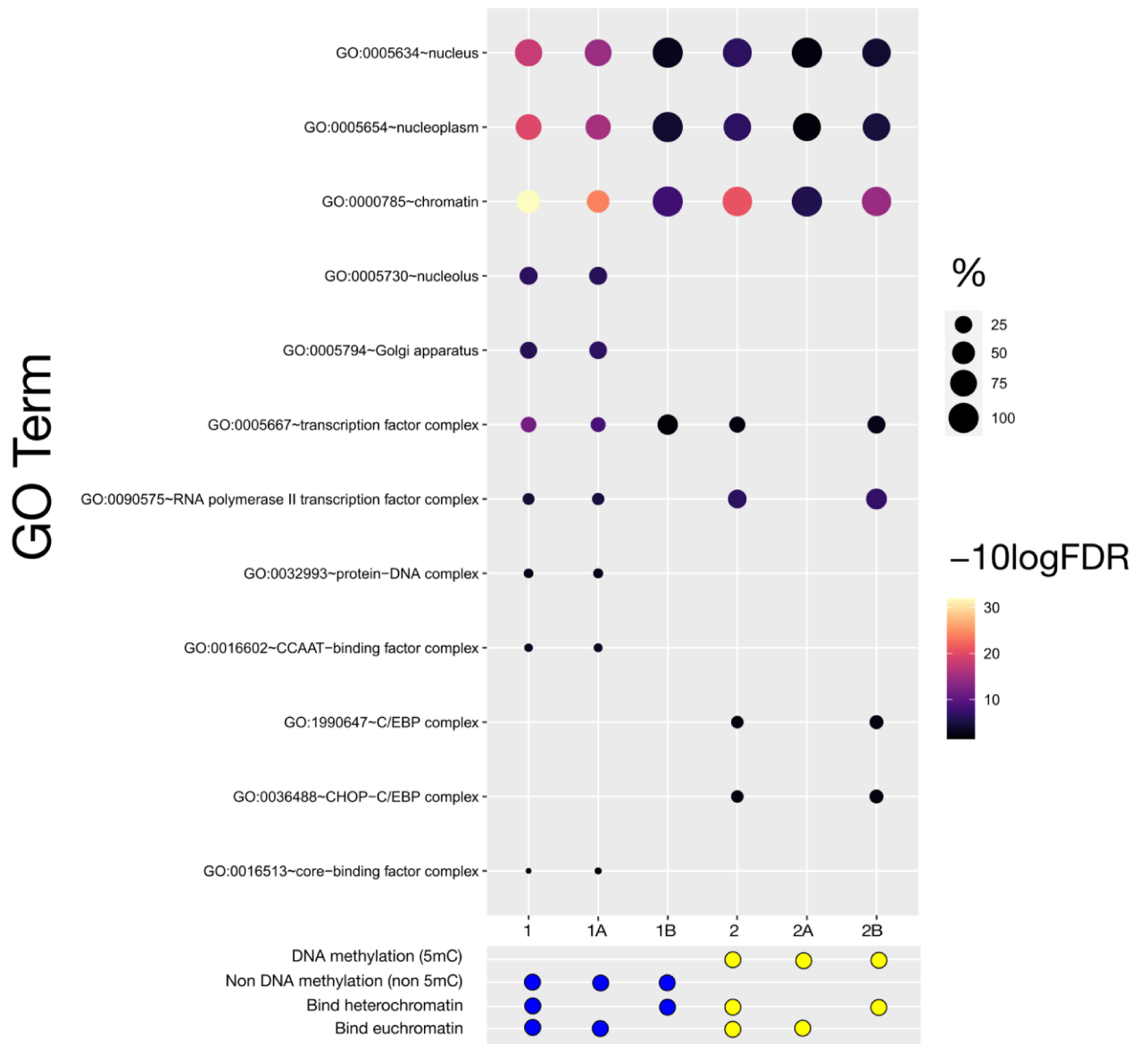
Figure 6. GO analysis of Cellular Component

The presented figure delineates the outcomes of the DAVID Gene Ontology analysis, specifically focusing on cellular processes within each category and subcategory. The figure utilizes two visual representations. Firstly, the dot size corresponds to the percentage of DAVID genes associated with a specific annotation term in each category. Secondly, a color gradient ranging from dark blue to light yellow signifies the -log10 false discovery rate (FDR).

Group	TF	Group	TF
1A	NF.Y.A.B.C	1A	Dux4
1A	Etv2	1A	FoxA2
1A	NEUROD1	1A	FoxA3
1A	HOXB13	1B	NANOG
1A	TCF3	1B	Myc
1A	SPI1	1B	MEF2
1A	Sox9	1B	FoxC1
1A	Sox6	2A	FLI1
1A	Sox5	2A	EBF1
1A	Sox2	2A	TFAP2.A.B.C
1A	RUNX3	2A	Klf1
1A	RUNX1	2B	MAFK
1A	PBX3	2B	Flf13
1A	PBX2	2B	Klf10
1A	PBX1	2B	Klf9
1A	Pax7	2B	KLF5
1A	Pax5	2B	Klf4
1A	Oct4	2B	GRHL2
1A	GATA6	2B	GATA4
1A	GATA5	2B	FoxK2
1A	GATA3	2B	FoxH1
1A	GATA2	2B	CEBPA
1A	GATA1	2B	CEBPB
1A	FOXQ1	2B	CEBPD
1A	FOXK1	2B	CEBPG
1A	FOXP1	1A	FoxL1
1A	FoxO3	1A	FoxD3

| 1A | FoxO1 | 1A | FoxA1 |
| 1A | FoxM1 | 1A | Ascl1 |

Table 10. Transcription factors groups classification

3.2.2 Analysis of Biological Process (BP)

The analysis successfully identified anticipated common terms related to transcriptional regulation across all recently established groups of pTFs (Figure 7).

In particular, Group 1, pTFs lacking the ability to bind 5-mC, showcases a diverse involvement in chondrocyte and erythrocyte differentiation. Conversely, Group 2, comprising TFs that bind 5-mC, exhibits a unique involvement in adipocyte differentiation and negative regulation of transcription. Group 1 was also found to be involved in various chromatin related processes, including histone PTM.
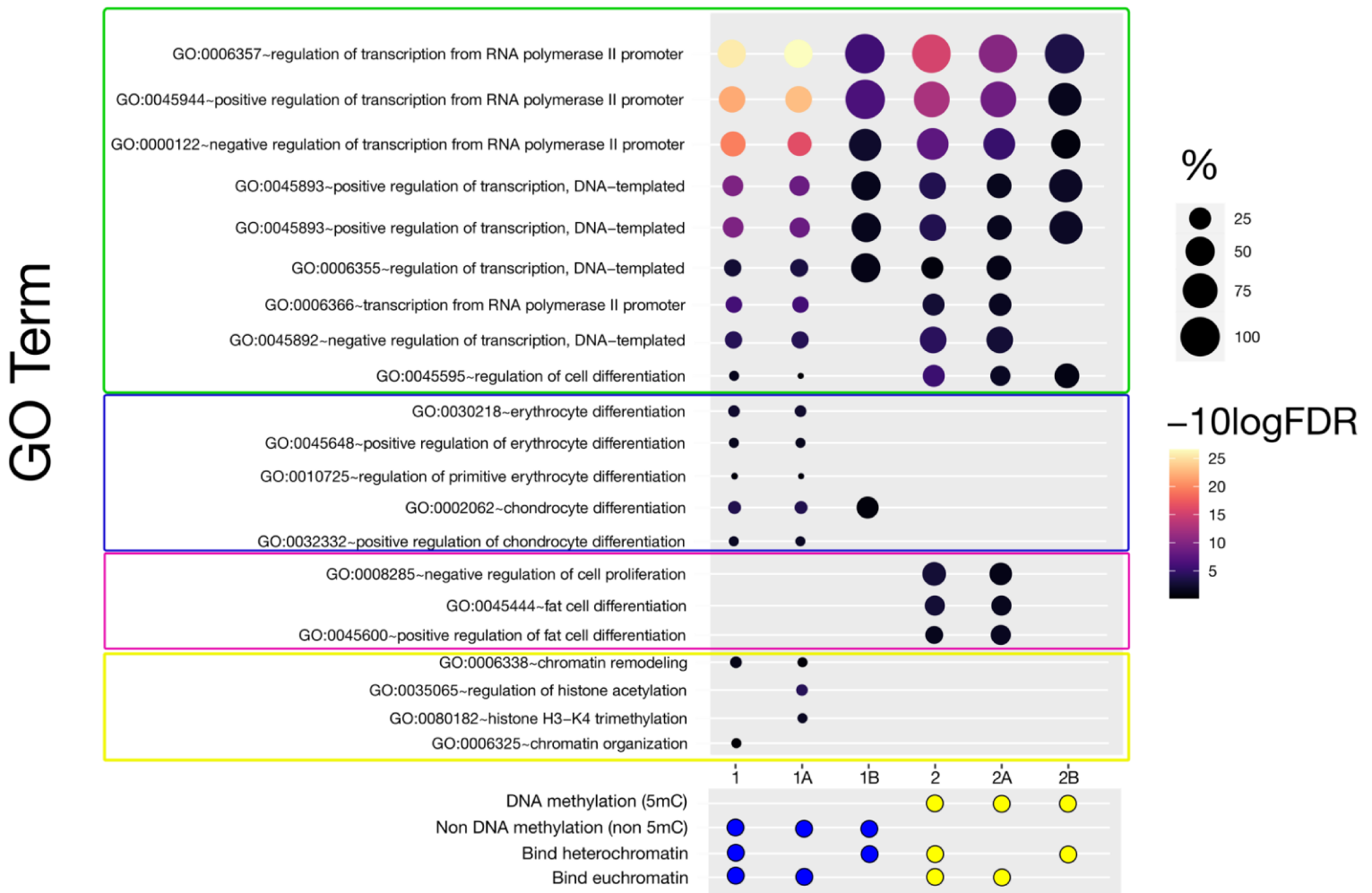
Figure 7. GO Analysis of Biological Process

The figure presented illustrates the results of the DAVID Gene Ontology analysis, specifically focusing on biological processes, for each category and subcategory. The figure employs two visual representations of the results. Firstly, the percentage of DAVID genes in the list associated with a particular annotation term associated with each category is represented by the size of the dots. Secondly, the color gradient, ranging from dark blue to light yellow, represents the -log10 false discovery rate (FDR). The shared Gene Ontology (GO) categories between Group 1 and Group 2 are highlighted in green, indicating the common biological processes. Additionally, unique GO categories for Group 1, Group 1A, and Group 2 are highlighted in blue, yellow, and pink, respectively.

3.2.2 Analysis of Molecular Functions (MF)

In the context of Molecular Function, terms such as 'sequence-specific double-stranded DNA binding' and various descriptors related to transcription regulation emerged, aligning with the known TF nature of our genes of interest (Figure 8). Noteworthy terms include 'protein dimerization activity' and 'histone deacetylase binding,' specific to Group 1, and 'co-SMAD binding,' exclusive to Group 2.
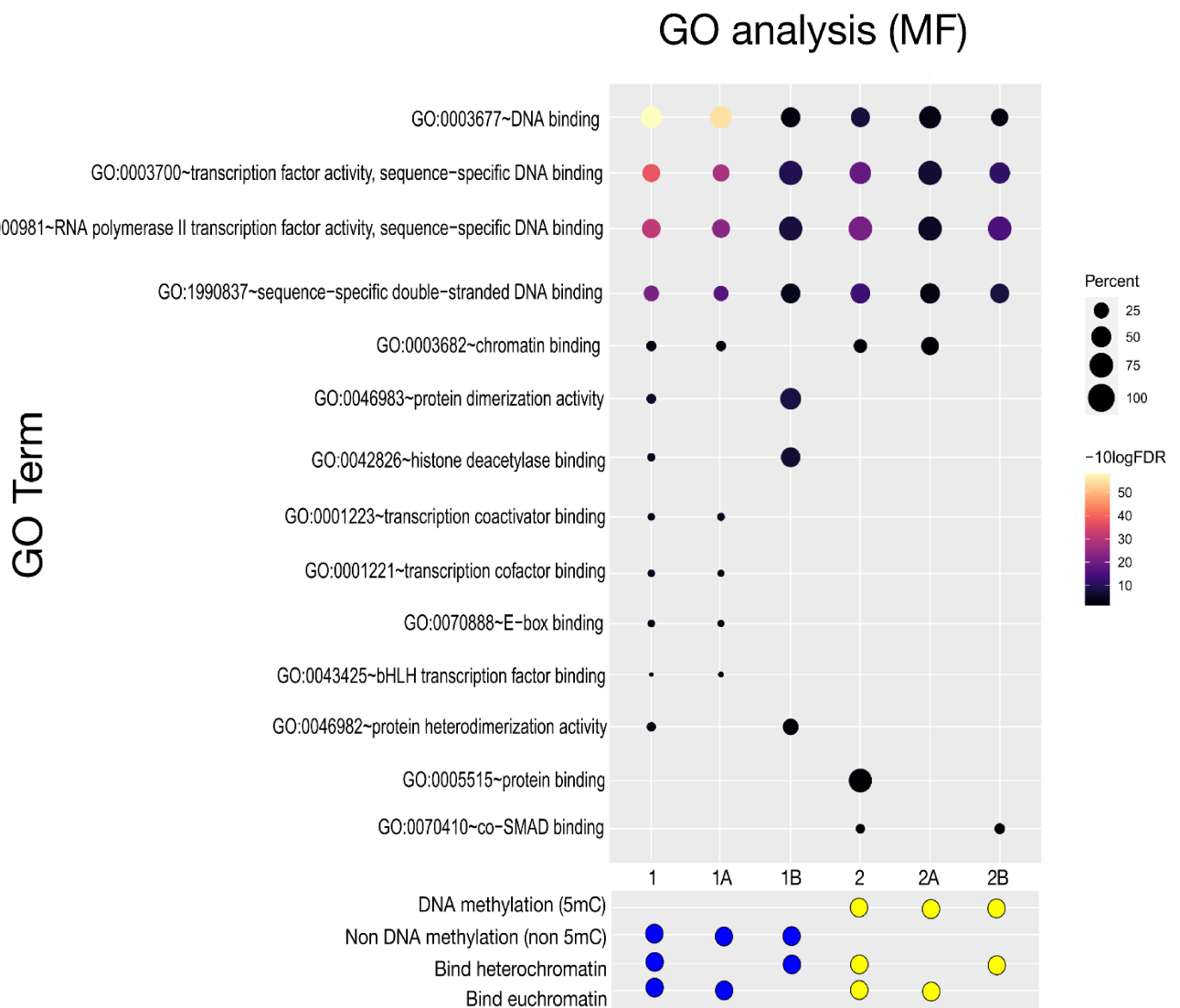
Figure 8. GO Analysis of Molecular Functions

The figure provides a visual representation of the outcomes from the DAVID Gene Ontology analysis, with a specific emphasis on molecular functions within each category and subcategory. The size of the dots corresponds to the percentage of DAVID genes in the list

associated with a specific annotation term for each category. The color gradient, transitioning from dark blue to light yellow, is utilized to portray the -log10 false discovery rate (FDR).

3.3. Pathway Analysis

Pathway analysis, utilizing the Reactome database, was executed to visualize the top 25 pathways considered highly significant for each group and subgroup (Figure 9). Several pathways were identified as common across all groups, primarily associated with developmental biology processes such as gastrulation, cell fate determination, and transcription regulation—attributes inherent to TFs. Group 1, consisting of TFs unable to bind methylated DNA was found to take part in the transcriptional regulation of pluripotent stem cells and involvement in gene expression in endocrine cell specificity and immune cell development. Group 2, known for its role in adipose tissue regulation, corroborated similar findings in the Gene Ontology analysis. Furthermore Group 2 has a connection to the reaction to environmental stimuli. Additionally, Group 2 demonstrated an association with the sumoylation of TFs and protein metabolism.
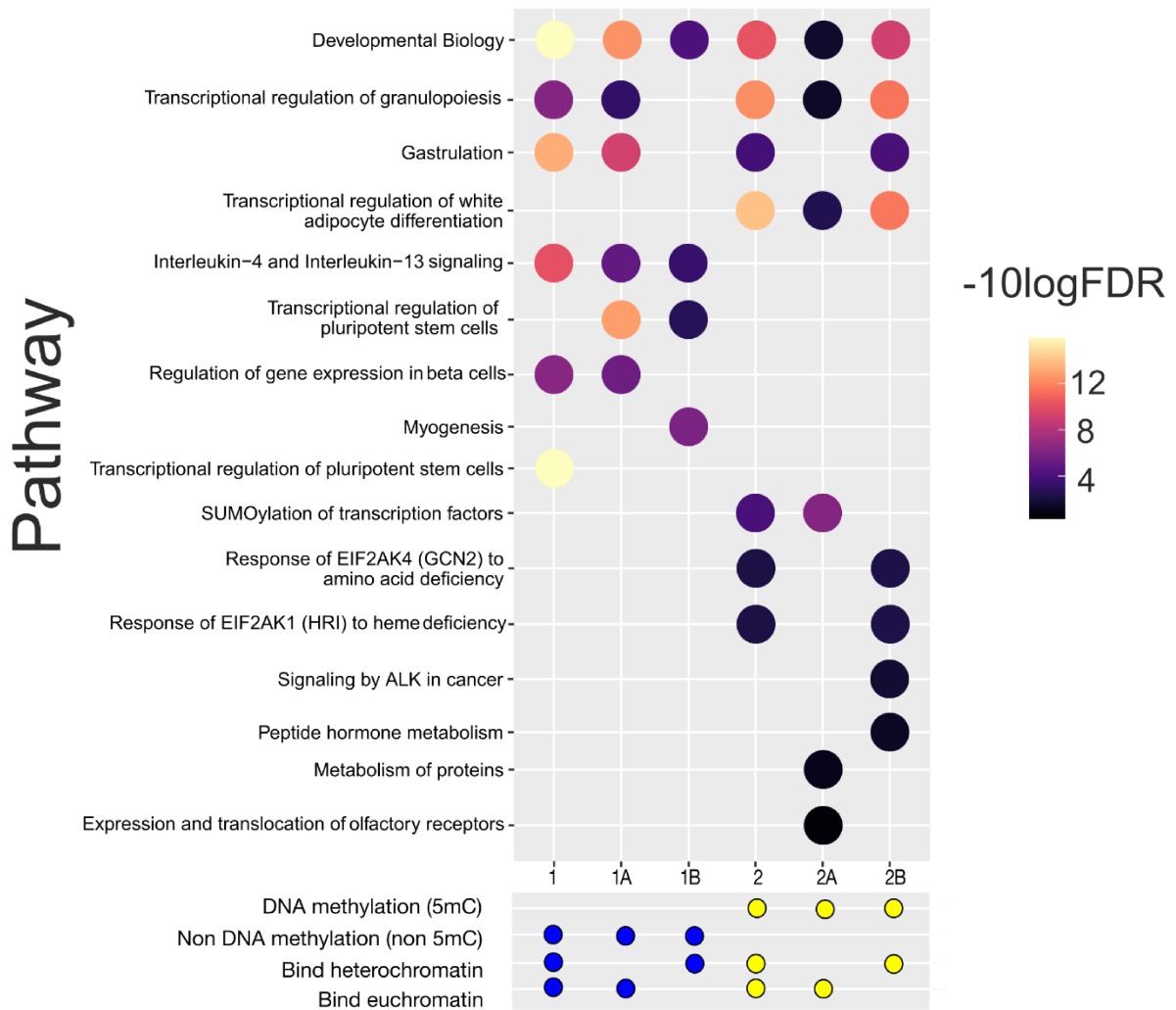
Figure 9. Pathway analysis

The figure provides a visual representation of results obtained from Reactome pathway analysis. On the left side, Reactome pathways are categorized and listed. The gradient of dots for each pathway represents the -log10 transformation of the False Discovery Rate (FDR), transitioning from dark blue to light yellow. This color spectrum indicates the significance of each result, where blue corresponds to the least significant values.

3.4. Gene Expression Analysis

3.4.1. Baseline Gene Expression

The baseline gene expression analysis of the entire set of transcription factors of interest identified three primary baseline gene expression groups, each providing unique insights into the transcriptional landscape (Figure 10).

The first group comprises TFs that exhibit high expression levels across numerous tissues, suggesting a ubiquitous and widespread nature. This finding implies a potential role for these TFs in fundamental and essential cellular processes that transcend tissue-specific boundaries.

Conversely, the second group showcases an opposing trend, featuring TFs with consistently low overall expression levels across all studied tissues. This pattern suggests a distinctive class of TFs that maintain a basal expression state, potentially highlighting their regulatory relevance under specific conditions or in response to particular stimuli.

The third group introduces a fascinating dimension, encompassing TFs characterized by high expression levels in specific tissues while concurrently maintaining low expression levels in others. This distinctive specialization implies a tissue-specific regulatory role for these TFs, indicating their potential involvement in the modulation of cellular processes exclusive to certain tissues.

A crucial observation stemming from the analysis is the apparent divergence within a single family of TFs. Notably, discrepancies in baseline expression patterns are evident within the CEBP family. Specifically, CEBPB and CEBPD exhibit widespread expression across multiple tissues, while CEBPA demonstrates high expression exclusively in breast and adipose tissues, as illustrated in Figure 10.

Illumina Body Map Basic Expression Human Adult
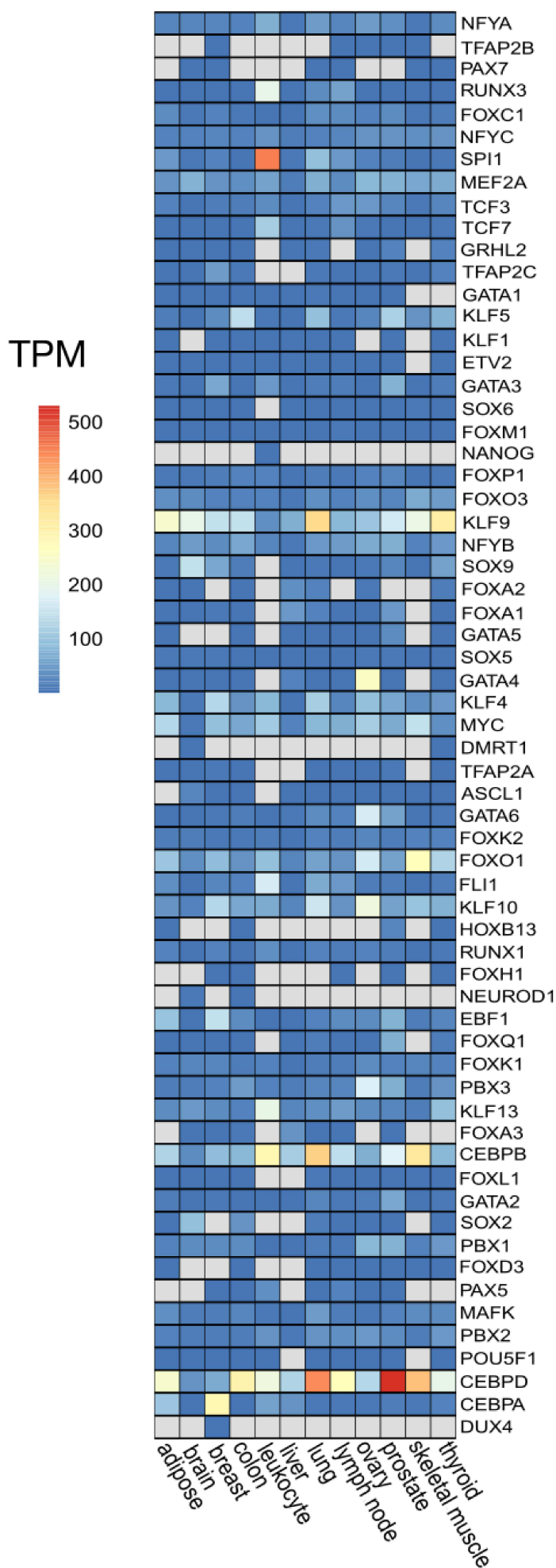
Figure 10. Baseline Gene Expression - Illumina Body Map

The figure depicts a matrix with tissue names along the bottom axis and the names of TFs of interest on the right side. The expression levels, denoted by Transcripts Per Million (TPM) values, are presented numerically in the table. The matrix is visually enhanced with color fillings corresponding to the TPM values. A gradient of colors ranges from blue, representing low values, to red, indicative of high values, with white blanks representing NA.

GTEx Basic Expression Human Adult
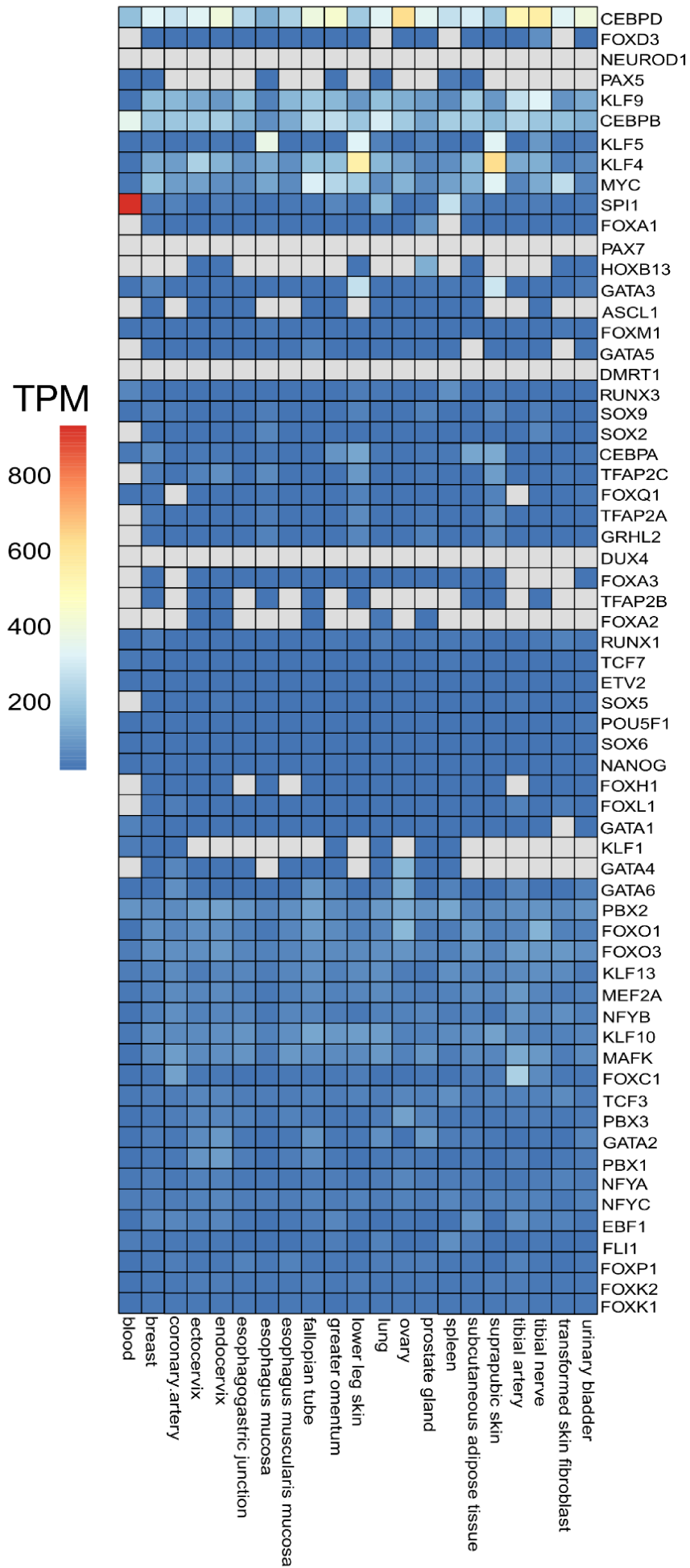
Figure 11. Baseline Gene Expression - Genotype-Tissue Expression (GTEx) Project

The illustration showcases a matrix where tissue names align along the horizontal axis, and the names of the TFs of interest are positioned on the right side. The numerical presentation of expression levels, indicated by Transcripts Per Million (TPM) values, is provided in the table. The matrix is visually enriched by employing color fills that align with the TPM values. The color spectrum transitions from blue, denoting lower values, to red, signifying higher values, with white blanks representing NA.

FANTOM5 Baseline Expression
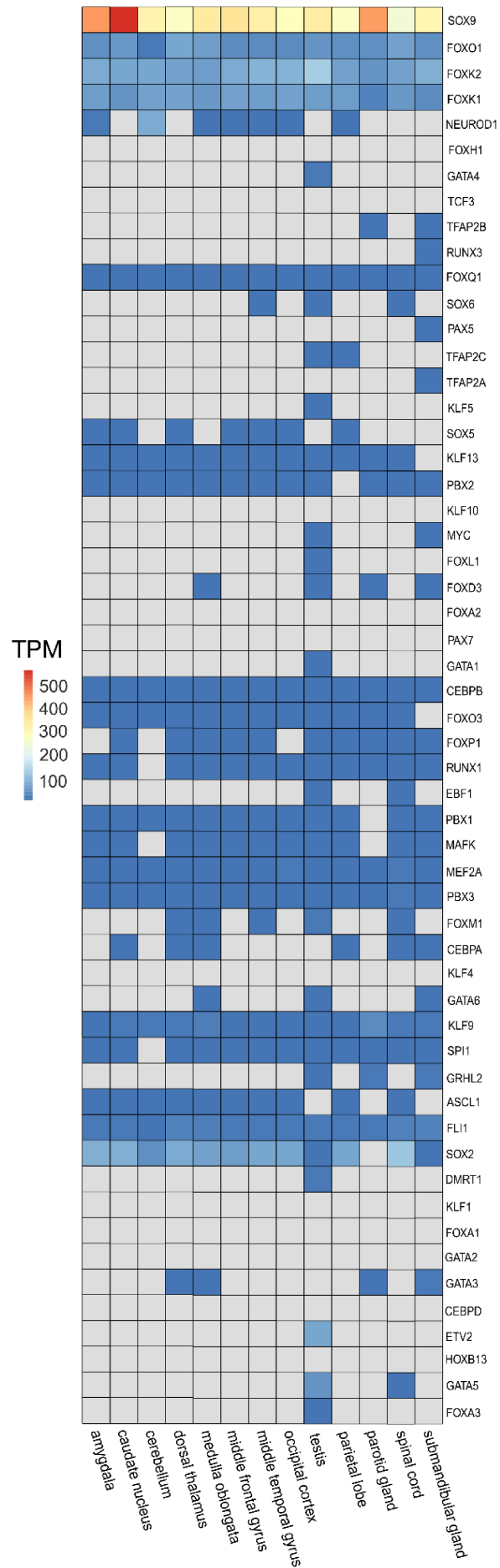
Figure 12. Baseline Gene Expression - FANTOM5 project

This figure features a matrix with tissue names arranged along the horizontal axis and the names of the specified TFs along the vertical axis. The table provides a numerical display of expression levels, denoted by Transcripts Per Million (TPM) values. The matrix incorporates color fills corresponding to the TPM values. The color spectrum transitions from blue, indicating lower values, to red, signifying higher values.

3.4.2. Differential Gene Expression

In our exploration of distinct patterns of differential expression among our genes of interest, we utilized the 'differential expression' function within the Expression Atlas. This function facilitated pairwise comparisons, pinpointing genes with notable differential expression (adjusted p-value < 0.05 and $\log_2$ fold-change > 1). The ensuing data is visually depicted in Figure 13 for upregulated genes and Figure 14 for downregulated genes.

Particularly noteworthy are certain processes in the analysis that showcase a plethora of differentially regulated genes. For instance, the comparison between 'induced pluripotent stem cell' and 'fibroblast of dermis' in both Klinefelter's syndrome and the normal condition reveals over 10 genes that are either upregulated or downregulated. This richness in differential regulation suggests that certain TFs may participate in a myriad of processes and pathologies, distinguishing them from TFs with more limited functional roles.

Differential gene expression - Upregulation
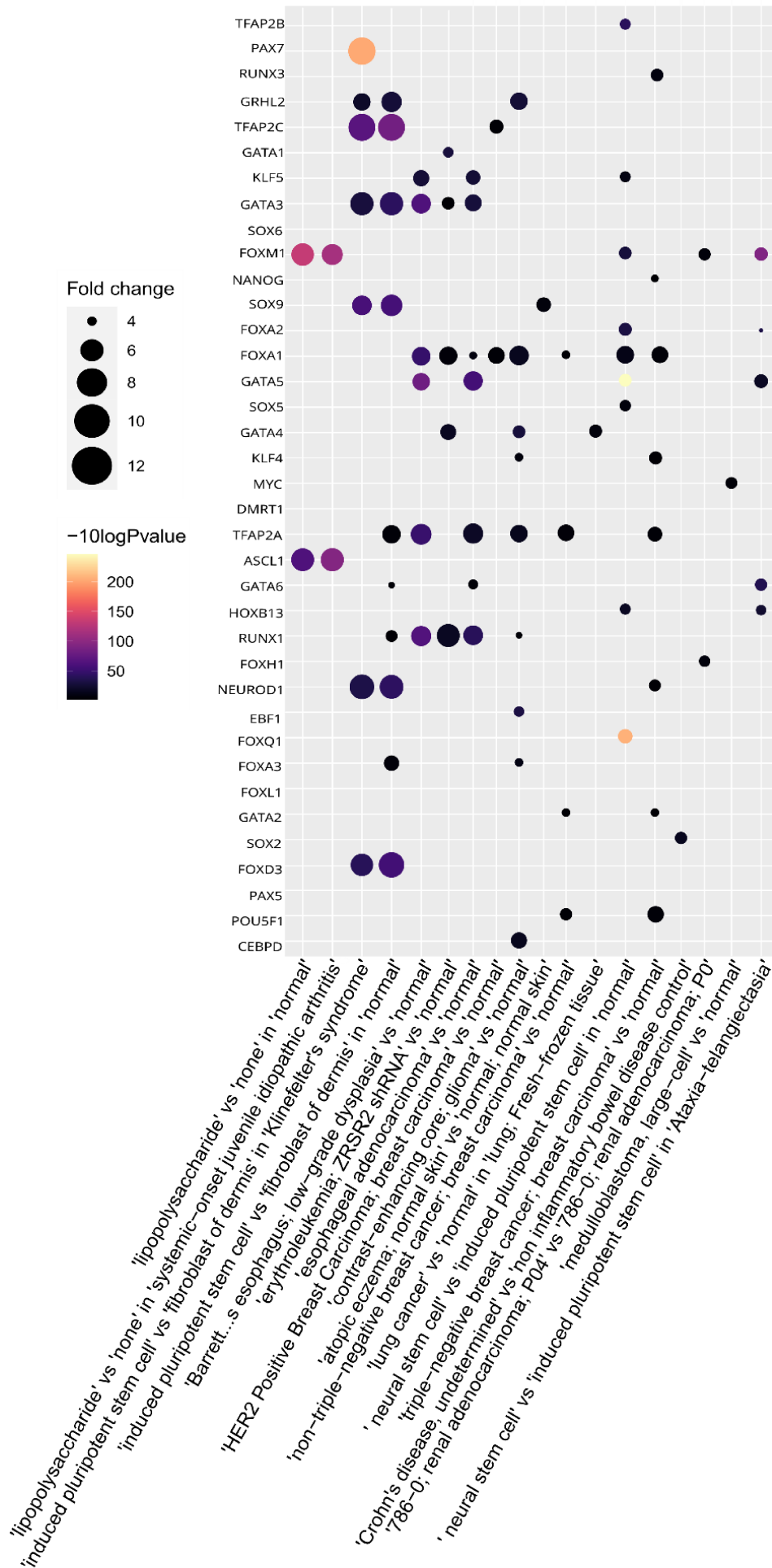
Figure 13. Differential Gene Expression – Upregulation

The figure showcases a list of upregulated genes on the left side, with the names of comparison pairs listed along the bottom axis. The visualization utilizes a color gradient transitioning from dark blue to yellow, effectively portraying the increasing -log10 p-value. Additionally, the size of the data points corresponds to the fold change.
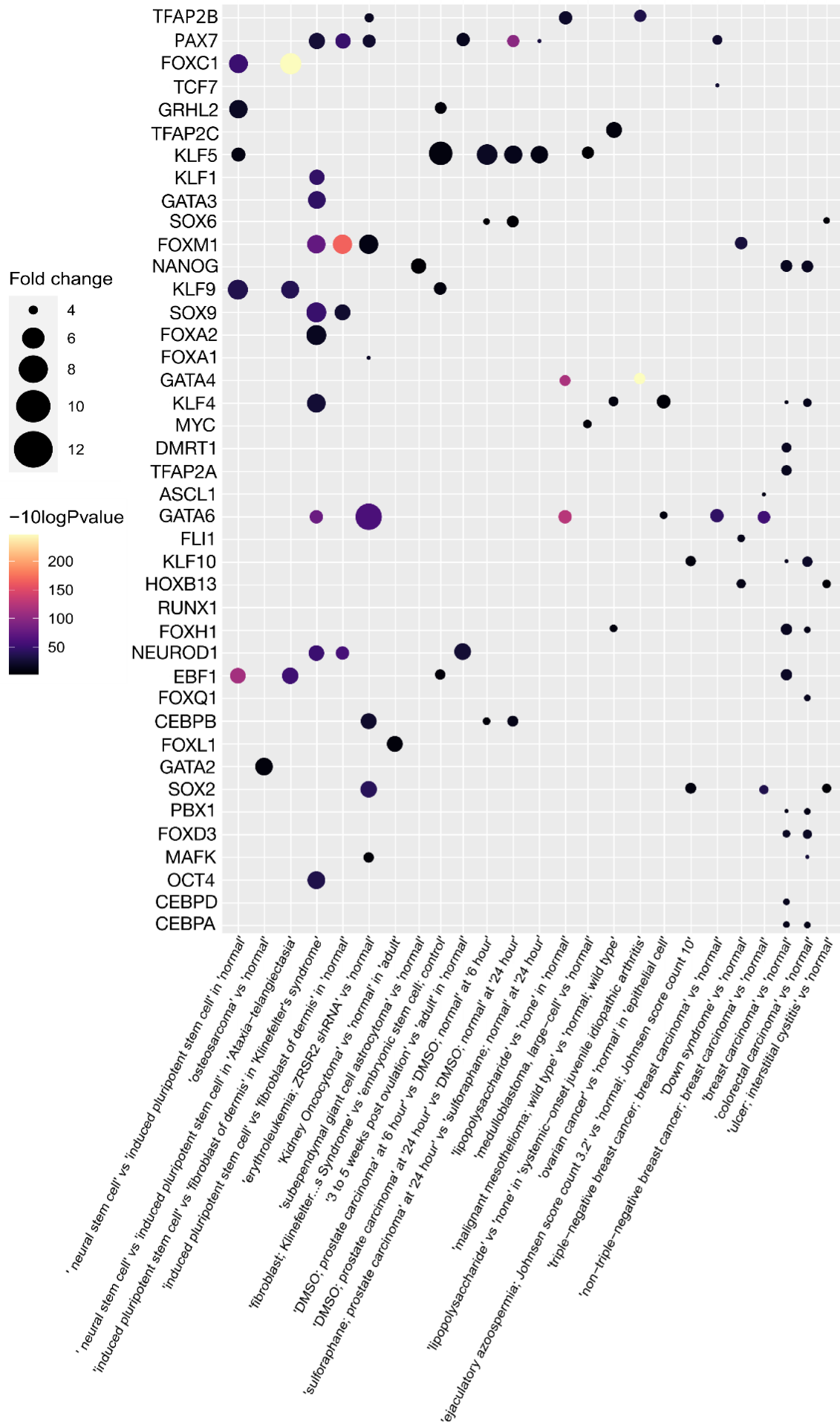
Differential gene expression - Downregulation

Figure 14. Differential Gene Expression – Downregulation

The figure presents a list of downregulated genes on the left side, accompanied by the names of comparison pairs listed along the bottom axis. The visualization employs a color gradient that smoothly transitions from dark blue to yellow, effectively illustrating the increasing log10 p-value. Simultaneously, the size of the data points aligns with the fold change, providing a comprehensive visual representation of both the statistical significance (indicated by color) and the magnitude of fold change for each gene across the diverse comparison pairs.

IV. Discussion

This study reveals pioneer TFs and their roles in cellular processes. It identifies 61 TFs data collection, categorizing them based on criteria including heterochromatin binding, interaction with methylated DNA, DNA accessibility, cell fate impact, and cancer roles. Heatmap analysis demonstrates distinct TF groups based on DNA methylation and heterochromatin binding patterns.

Group 1 encompassed TFs not binding to 5-methylcytosine (5-mC), further divided into subgroups 1A and 1B based on their interactions with closed chromatin. Group 2 consisted of TFs binding to 5-mC, with subgroups 2A and 2B distinguishing between those binding only to 5-mC and those binding to both 5-mC and heterochromatin.

The members of Group 1A, exemplified by OCT4, and SOX2, have been identified as pTFs with a notable capacity to interact with heterochromatin, distinct from their known interactions with 5-mC. This interaction implies a regulatory role in gene expression within these genomic regions. Recent studies corroborate this notion, particularly observing OCT4 and SOX2's ability to selectively perturb nucleosomal DNA (Michael et al. 2020), aligning with their classification as Yamanaka pluripotency factors (Xiao et al. 2016). Their unique capability to bind heterochromatin underscores their pivotal role in modulating chromatin architecture, crucial for maintaining pluripotency in stem cells and orchestrating cell fate determination during developmental processes (Rizzino and Wuebben 2016). Analyzing the binding preferences of TFs sheds light on cellular identity maintenance and differentiation mechanisms. TFs unable to bind to 5-mC-marked regions may lead to gene silencing (Vanzan et al. 2021), while those binding to heterochromatin can regulate gene expression through chromatin remodeling. This insight holds therapeutic promise (Stolzenburg et al. 2012), as evidenced by studies showing efficacy in treating drug-resistant breast cancer using specific cytotoxic T lymphocytes targeting OCT4 and SOX2 combined with a PD-1 inhibitor like nivolumab (Peng et al. 2022).

The group 1B examples, including NANOG, MYC, MEF2, and FoxC1, are notable for their unique characteristic of not binding to either 5-mC or heterochromatin. Despite this, they possess chromatin-modulating capabilities (Choi et al. 2022), raising intriguing questions about their mechanism of action (González-Rico et al. 2020). It's noteworthy that they may exert their effects through interaction with other TFs. A prominent illustration of this is the association between NANOG, MYC, and the subgroup 1A factors OCT4 and SOX2 (Amaya and Bryan 2015). These partnerships are pivotal for maintaining and inducing pluripotency (Kashyap et al. 2009). This suggests that direct binding to 5-mC and heterochromatin may not be essential for certain activities such as nucleosome disintegration, as some pTFs can compensate by collaborating synergistically with others possessing these abilities (Nepon-Sixt, Bryant, and Alexandrow 2019).

The group 2A includes pTFs that bind 5-mC but do not bind the heterochromatin. The examples are FLI1, EBF1, KLF1. Their ability to bind 5-mC but not heterochromatin is interesting as it's commonly supposed that heterochromatin region is usually tightly methylated which means that this region is repressed and not active. This unique pattern is consistent with the event of chromatin priming. Chromatin priming refers to the process by which specific genomic regions are marked or prepared for subsequent transcriptional activation or repression. This preparation involves the recruitment of various factors, including TFs, chromatin modifiers, and chromatin remodeling complexes, to modify the chromatin structure and establish a conducive environment for gene expression or repression. This suggests that group 2A might contain and therefore differentiate the pTFs which act specifically as a priming agent. TFs that interact with 5-mC contribute to the priming of chromatin by integrating epigenetic signals, modifying chromatin structure, and preparing genomic regions for subsequent transcriptional activation or repression. An illustrative example is EBF1 – pTF responsible for initiation of B cell differentiation (Strid et al. 2021). This factor stands out as the primary driver of cell-specific

gene expression initiation. Remarkably, EBF1 exhibits the unique capability to bind to chromatin independently of other factors (Boller et al. 2016). Consequently, EBF1 earns recognition for its role in priming and inducing a permissive chromatin state, essential for subsequent hematopoiesis processes.

Group 2B comprises factors that exhibit a unique dual affinity for both 5-methylcytosine (5-mC) and heterochromatin. Notable examples within this group include CEBPB, GATA4, and MAFK. Their remarkable binding versatility implies a crucial role in gene expression regulation, as they can autonomously bind without reliance on other factors. Consequently, these factors possess the direct ability to activate target genes (He et al. 2014). Moreover, their distinctive binding properties suggest involvement not only in chromatin landscape modulation but also in mediating repressed chromatin assembly. For instance, MAFK has been observed binding to intronic regions of IRF8, serving as a chromatin repressive element and influencing histone post-translational modifications (PTMs) (Fourier et al. 2020). This highlights a suggestive characteristic of this group: the maintenance of a repressive state in regions abundant in binding sites.

The classification rationale outlined above underscores the significance of diverse chromatin interaction patterns as a crucial criterion for categorizing these proteins. It highlights that binding preferences and the influence on chromatin dynamics are far more significant differentiating factors than any attributes related to the molecular structure alone. This approach recognizes that the functional role of these proteins in gene regulation is intricately tied to their ability to interact with chromatin in various ways, reflecting a nuanced understanding of their biological activity.

The Gene Ontology (GO) enrichment analysis conducted in Section 3.2 examined the functional implications of the TF groups. For cellular component analysis, common terms across all groups such as 'nucleus,' 'chromatin,' and 'TF complex' were identified as shared

cellular components. Additionally, group-specific terms like 'C/EBP complex' for Group 2 and 'Golgi apparatus' and 'nucleolus' for Group 1 were highlighted. This unexpected finding may be elucidated by the presence of E-twenty six (ETS) family TFs within Group 1. Notably, the transcriptional activity of ETS family members in Group 1 is associated with the regulation of spliceosome function and the induction of cell death, both of which are linked to Golgi stress (Baumann et al. 2018).

Regarding the biological procceses, Group 1demonstrates a diverse involvement in chondrocyte and erythrocyte differentiation. Notably, the regulation of erythrocyte differentiation, primarily governed by the GATA family of TFs, aligns well with the composition of Group 1 pTFs (Welch et al. 2004). Furthermore, the presence of RUX and SOX families in Group 1 corresponds with their documented roles in chondrocyte differentiation (Komori 2018; C.-F. Liu et al. 2017). Conversely, Group 2, was found in adipocyte differentiation, which is consistent with the presence of the CEBP TF family. Additionally, this group plays a role in the negative regulation of transcription—an intriguing finding that supports our classification process. This is because the ability to bind 5-mC regions is associated with inactive transcription (Miller and Grant 2013).

In the context of Molecular Function, noteworthy terms include 'protein dimerization activity' and 'histone deacetylase binding,' specific to Group 1, and 'co-SMAD binding,' exclusive to Group 2. The significance of protein dimerization for transcription activation (Ortega et al. 2018; Peterson et al. 2012) with our earlier findings for Group 1. Moreover, the identification of 'histone deacetylase binding' aligns coherently with our prior results, where Group 1 demonstrated a substantial impact on histone acetylation regulation and therefore transcription activation (Mizzen and Allis 1998; Shvedunova and Akhtar 2022; Kim et al. 2023). The specific term 'co-SMAD binding' in Group 2 gains significance as Common SMAD (co-SMAD) is a subtype of the main signal transducers for receptors of the transforming growth factor-beta

(TGF-B) (Moustakas, Souchelnytskyi, and Heldin 2001). TFs in Group 2 actively recruit SMADs, forming a complex of coregulators that modify chromatin structure (Hill 2016). This observation aligns with the fact that Group 2 includes TFs known as pioneer factors, demonstrating an affinity for binding methylated DNA. In summary, the comprehensive GO analysis not only underscores the shared functional aspects between groups but also highlights specific molecular functions that substantiate and enrich the rationale behind our classification. Pathway analysis revealed intriguing insights into the functional roles of Group 1 and Group 2 TFs (TFs). Group 1, known for its pioneer status, was found to exert influence over the fate of pluripotent stem cells, indicating its pivotal role in early developmental processes (Ng et al. 2021). Additionally, Group 1 TFs were implicated in gene expression related to endocrine cell specificity (Drouin 2014) and immune cell development (Yuan et al. 2022; Recaldin and Fear 2016), highlighting their multifaceted involvement in diverse cellular processes. On the other hand, Group 2, recognized for its significance in adipose tissue regulation, exhibited similar patterns in the Gene Ontology analysis, further affirming its role in metabolic pathways and adipogenesis. Notably, novel findings emerged for Group 2, indicating its association with the response to environmental stimuli, such as amino acid and heme deficiency. CEBPB, group 2, was found to be upregulated in the response to amino acid deficiency detected by EIF2AK4 (GCN2) (Thiaville et al. 2008; Averous et al. 2003). Moreover, Group 2 TFs demonstrated an intriguing association with the sumoylation of TFs. SUMOylation, a crucial post-translational modification mediated by small ubiquitin-like modifiers, has been linked to the regulation of gene expression, often resulting in transcriptional repression (Du, McConnell, and Yang 2010). This connection suggests a potential mechanism by which Group 2 TFs, with their ability to bind methylated DNA, may exert regulatory control over target genes by modulating the activity of other transcriptional regulators through SUMOylation. Overall, these comprehensive

analyses deepen our understanding of the functional roles of Group 1 and Group 2 TFs, shedding light on their involvement in diverse biological processes and regulatory mechanisms.

The Gene Expression Analysis conducted in Section 3.4 introduced additional complexity to the study by categorizing TFs into distinct groups based on their baseline expression patterns. The first pattern comprises TFs with consistently low expression levels across all examined tissues. Examples include FOXM1, FOXK2, and FOXC1. This suggests that TFs in this category may need to be present in high quantities to support normal cellular processes, and elevated expression levels could indicate malignancy. Notably, FOXM1 has been observed to be overexpressed in cancer (Barger et al. 2019; Rachmadi et al. 2022).

The second pattern consists of TFs that exhibit high expression in three or more tissues. Examples include CEBPB, CEBPD, and KLF9. This indicates that these TFs are abundant across multiple tissues but are maintained at higher expression levels compared to those in the first group. For instance, CEBPD has been identified as critical for various tissues such as prostate, lung, adipose, and colon. This finding aligns with existing literature, which identifies CEBPD as a key regulator of adipose tissue differentiation (Kusuyama et al. 2017). Moreover, CEBPD is activated through direct binding by the androgen receptor, contributing to normal apoptosis levels in prostate tissue (Yang et al. 2001).

The third pattern comprises TFs with high expression restricted to a specific tissue. Examples include CEBPA, SPI1, and GATA4. Despite belonging to the same family as CEBPB and CEBPD, CEBPA is predominantly expressed in breast tissue. Research indicates that CEBPA binds to the progesterone receptor and serves as a regulator of cell growth in hormone-dependent breast cancer (Nacht et al. 2019).

The utility of the results obtained from the Genotype-Tissue Expression (GTEx) and FANTOM5 projects was not as pronounced compared to the insights derived from the Illumina Body Map. Several factors could contribute to the perceived lower utility of GTEx and

FANTOM5 results such as tissue coverage, methodological variances, sample size and resolution. While the results from GTEx and FANTOM5 may not have been as immediately useful in the current context, it's essential to consider their strengths and limitations. The comparative analysis of multiple datasets can enhance the robustness of conclusions and provide a more comprehensive understanding of TF expression across diverse tissues. This emphasizes the pivotal significance of categorizing TFs not only according to their functional roles but also based on their expression levels within specific processes. The plethora of differentially regulated genes identified in these comparisons strongly suggests the direct participation of TFs in orchestrating these intricate biological processes. Furthermore, the adoption of this classification approach emerges as a valuable instrument for refining and predicting the enigmatic properties of pTFs, leveraging their observed behavior.

The analysis of differential gene expression revealed distinct patterns akin to those observed in baseline expression, resulting in the categorization of two main groups. The first group comprises factors exhibiting differential expression, either positively or negatively, across three or more processes. For instance, RUNX1 demonstrated overexpression in processes such as esophageal cancers and erythroleukemia. Notably, RUNX1 is among the most frequently mutated genes in various hematological malignancies and has been implicated in cancer metastasis modulation (Liau, Ngoc, and Sanda 2017; K. Liu et al. 2021). This suggests the potential utility of RUNX1 as both a cancer biomarker and a target for pharmaceutical intervention. On the other hand, the second group encompasses factors with highly specific functions. For example, CEBPD was found to be overexpressed in gliomas. It is well-established that CEBPD plays a critical role in the synthesis of nerve growth factor, underscoring its specific involvement in neural-related processes, which supports our findings (McCauslin et al. 2006).

The urgency in devising an innovative classification method arises from the recognition that TFs within the same family may possess distinct functional peculiarities, as evidenced by their disparate baseline expression profiles. A more nuanced classification approach, potentially integrating expression data and functional characteristics, becomes essential for accurately delineating the roles and regulatory impact of pTFs. This shift towards a more functionally informed classification methodology is pivotal for advancing our understanding of transcriptional regulation and its intricate complexities.

While this study provides valuable insights, it is essential to acknowledge and address potential limitations to enhance the robustness and applicability of the findings.

Firstly, the incorporation of datasets introduces the concept of Dataset Selection Bias. The investigation heavily depends on specific datasets, namely Illumina Body Map, GTEx, and FANTOM5, each potentially carrying inherent biases related to tissue representation and experimental conditions. Maintaining data consistency across diverse databases is another significant challenge. Updates or modifications made in one database may not propagate correctly to others, leading to inconsistencies unless robust synchronization mechanisms are in place. Performance variability is also a concern. Each database architecture performs differently based on factors like query complexity and data volume. Achieving consistent performance across different systems requires careful optimization tailored to each architecture. To overcome this challenge, additional datasets should be incorporated to validate and cross-reference findings.

Another critical aspect to consider is the criteria employed to categorize TFs as pioneer factors, which are susceptible to interpretation and may not completely encapsulate the nature of their functions. Continuous refinement of classification criteria by staying attuned to emerging literature and experimental evidence is advisable. Fostering collaborative efforts with domain

experts ensures that the classification criteria remain dynamic and reflective of the evolving understanding in the field.

The outcomes of Gene Ontology (GO) analysis can be impacted by the thoroughness and precision of gene annotation databases, introducing the possibility of biased functional interpretations. To mitigate this challenge, a proactive strategy involves the routine updating of gene annotation databases. Additionally, exploring alternative functional annotation methods enhances the robustness of the analysis.

Furthermore, conducting supplementary experimental validations becomes essential to validate and affirm the predicted functions of TFs across diverse cellular contexts. This multifaceted approach ensures a more reliable and comprehensive understanding of functional implications derived from GO analysis.

The integration of diverse datasets and the subsequent interpretation of results can introduce complexities and potential errors. To address these challenges, a recommended strategy involves the utilization of advanced bioinformatics tools and methods specifically designed for data integration. Collaborating with experts in data integration and statistics is crucial to ensuring a robust analysis and accurate interpretation of the integrated data. This collaborative approach not only enhances the reliability of the results but also leverages specialized knowledge to navigate the intricacies associated with diverse datasets, ultimately contributing to a more comprehensive and accurate understanding of the integrated data.

The findings from the current research on pTFs open up avenues for several follow-up studies to deepen our understanding of the molecular mechanisms and functional implications of these TFs. For example, cross-species comparative studies could be conducted to assess the conservation and divergence of pTFs across different organisms. Investigating whether the identified pTFs in human cells have homologs or functional equivalents in model organisms would provide valuable insights into the evolutionary aspects of these TFs. In summary, this

research not only contributes significantly to our current understanding of pTFs but also sets the stage for further studies that can deepen our knowledge of their molecular mechanisms and functional implications across diverse organisms and cellular contexts.

In conclusion, this comprehensive research significantly contributes to the understanding of pTFs, unraveling their diverse roles, regulatory mechanisms, and implications in various cellular processes. The findings presented in each section offer a multi-faceted view of TFs, enriching our knowledge in the field of transcriptional regulation.

References

Adams, Grace E., Aditya Chandru, and Shaun M. Cowley. 2018. "Co-Repressor, Co-Activator and General Transcription Factor: The Many Faces of the Sin3 Histone Deacetylase (HDAC) Complex." *Biochemical Journal* 475 (24): 3921–32. https://doi.org/10.1042/BCJ20170314.

Adhikari, Nilanjan, Sandip Kumar Baidya, and Tarun Jha. 2020. "Effective Anti-Aromatase Therapy to Battle against Estrogen-Mediated Breast Cancer: Comparative SAR/QSAR Assessment on Steroidal Aromatase Inhibitors." *European Journal of Medicinal Chemistry* 208 (December): 112845. https://doi.org/10.1016/j.ejmech.2020.112845.

Aguilar, Rodrigo, Kerrie B. Spencer, Barry Kesner, Noreen F. Rizvi, Maulik D. Badmalia, Tyler Mrozowich, Jonathan D. Mortison, et al. 2022. "Targeting Xist with Compounds That Disrupt RNA Structure and X Inactivation." *Nature* 604 (7904): 160–66. https://doi.org/10.1038/s41586-022-04537-z.

Akalin, Altuna, Matthias Kormaksson, Sheng Li, Francine E. Garrett-Bakelman, Maria E. Figueroa, Ari Melnick, and Christopher E. Mason. 2012. "MethylKit: A Comprehensive R Package for the Analysis of Genome-Wide DNA Methylation Profiles." *Genome Biology* 13 (10): R87. https://doi.org/10.1186/gb-2012-13-10-r87.

Akazawa, Hiroshi, and Issei Komuro. 2005. "Cardiac Transcription Factor Csx/Nkx2-5: Its Role in Cardiac Development and Diseases." *Pharmacology & Therapeutics* 107 (2): 252–68. https://doi.org/10.1016/j.pharmthera.2005.03.005.

Akey, Joshua M., Shameek Biswas, Jeffrey T. Leek, and John D. Storey. 2007. "On the Design and Analysis of Gene Expression Studies in Human Populations." *Nature Genetics*. nature.com. https://doi.org/10.1038/ng0707-807.

Allen, R. D., 3rd, T. P. Bender, and G. Siu. 1999. "C-Myb Is Essential for Early T Cell Development." *Genes & Development* 13 (9): 1073–78. https://doi.org/10.1101/gad.13.9.1073.

Almalki, Sami G., and Devendra K. Agrawal. 2016. "Key Transcription Factors in the Differentiation of Mesenchymal Stem Cells." *Differentiation; Research in Biological Diversity* 92 (1–2): 41–51. https://doi.org/10.1016/j.diff.2016.02.005.

Andrews, Simon, Christel Krueger, Maravillas Mellado-Lopez, Myriam Hemberger, Wendy Dean, Vicente Perez-Garcia, and Courtney W. Hanna. 2023. "Mechanisms and Function of de Novo DNA Methylation in Placental Development Reveals an Essential Role for DNMT3B." *Nature Communications* 14 (1): 371. https://doi.org/10.1038/s41467-023-36019-9.

Andrews, Simon, and Others. 2010. "FastQC: A Quality Control Tool for High Throughput Sequence Data." Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

Apicella, Clara, Camino S. M. Ruano, Céline Méhats, Francisco Miralles, and Daniel Vaiman. 2019. "The Role of Epigenetics in Placental Development and the Etiology of Preeclampsia." *International Journal of Molecular Sciences* 20 (11). https://doi.org/10.3390/ijms20112837.

Aranda, A., and A. Pascual. 2001. "Nuclear Hormone Receptors and Gene Expression." *Physiological Reviews* 81 (3): 1269–1304. https://doi.org/10.1152/physrev.2001.81.3.1269.

Aref, Rasha, Marwa N. M. E. Sanad, and Hans-Joachim Schüller. 2021. "Forkhead Transcription Factor Fkh1: Insights into Functional Regulatory Domains Crucial for Recruitment of Sin3 Histone Deacetylase Complex." *Current Genetics* 67 (3): 487–99. https://doi.org/10.1007/s00294-021-01158-3.

Ashery-Padan, R., and P. Gruss. 2001. "Pax6 Lights-up the Way for Eye Development."

Current Opinion in Cell Biology 13 (6): 706–14. https://doi.org/10.1016/s0955-0674(00)00274-x.

Augello, Michael A., Theresa E. Hickey, and Karen E. Knudsen. 2011. "FOXA1: Master of Steroid Receptor Function in Cancer." *The EMBO Journal* 30 (19): 3885–94. https://doi.org/10.1038/emboj.2011.340.

Aydin, Begüm, Michael Sierk, Mireia Moreno-Estelles, Link Tejavibulya, Nikathan Kumar, Nuria Flames, Shaun Mahony, and Esteban O. Mazzoni. 2022. "Foxa2 and Pet1 Direct and Indirect Synergy Drive Serotonergic Neuronal Differentiation." *Frontiers in Neuroscience* 16 (June): 903881. https://doi.org/10.3389/fnins.2022.903881.

"Babraham Bioinformatics - Trim Galore!" n.d. Accessed October 15, 2023. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

Bahadoram, Sara, Mohammad Davoodi, Shakiba Hassanzadeh, Mohammad Bahadoram, Maedeh Barahman, and Ladan Mafakher. 2022. "Renal Cell Carcinoma: An Overview of the Epidemiology, Diagnosis, and Treatment." *Giornale Italiano Di Nefrologia: Organo Ufficiale Della Societa Italiana Di Nefrologia* 39 (3). https://www.ncbi.nlm.nih.gov/pubmed/35819037.

Baker, Luke R., Bonnie M. Weasner, Athena Nagel, Sarah D. Neuman, Arash Bashirullah, and Justin P. Kumar. 2018. "Eyeless/Pax6 Initiates Eye Formation Non-Autonomously from the Peripodial Epithelium." *Development* 145 (15). https://doi.org/10.1242/dev.163329.

Bartlett, Anna, Ronan C. O'Malley, Shao-Shan Carol Huang, Mary Galli, Joseph R. Nery, Andrea Gallavotti, and Joseph R. Ecker. 2017. "Mapping Genome-Wide Transcription-Factor Binding Sites Using DAP-Seq." *Nature Protocols* 12 (8): 1659–72. https://doi.org/10.1038/nprot.2017.055.

Bartolomei, Marisa S., Rebecca J. Oakey, and Anton Wutz. 2020. "Genomic Imprinting: An Epigenetic Regulatory System." *PLoS Genetics* 16 (8): e1008970. https://doi.org/10.1371/journal.pgen.1008970.

Battistelli, Cecilia, Sabrina Garbo, and Rossella Maione. 2022. "MyoD-Induced Trans-Differentiation: A Paradigm for Dissecting the Molecular Mechanisms of Cell Commitment, Differentiation and Reprogramming." *Cells* 11 (21). https://doi.org/10.3390/cells11213435.

Baumann, Jan, Tatiana I. Ignashkova, Sridhar R. Chirasani, Silvia Ramírez-Peinado, Hamed Alborzinia, Mathieu Gendarme, Kyra Kuhnigk, Valentin Kramer, Ralph K. Lindemann, and Jan H. Reiling. 2018. "Golgi Stress-Induced Transcriptional Changes Mediated by MAPK Signaling and Three ETS Transcription Factors Regulate MCL1 Splicing." *Molecular Biology of the Cell* 29 (1): 42–52. https://doi.org/10.1091/mbc.E17-06-0418.

Beato, M., and J. Klug. 2000. "Steroid Hormone Receptors: An Update." *Human Reproduction Update* 6 (3): 225–36. https://doi.org/10.1093/humupd/6.3.225.

Belo, Jamie, Mansa Krishnamurthy, Amanda Oakie, and Rennian Wang. 2013. "The Role of SOX9 Transcription Factor in Pancreatic and Duodenal Development." *Stem Cells and Development* 22 (22): 2935–43. https://doi.org/10.1089/scd.2013.0106.

Bensimhon, M., J. Gabarro-Arpa, R. Ehrlich, and C. Reiss. 1983. "Physical Characteristics in Eucaryotic Promoters." *Nucleic Acids Research* 11 (13): 4521–40. https://doi.org/10.1093/nar/11.13.4521.

Bianco-Miotto, Tina, Benjamin T. Mayne, Sam Buckberry, James Breen, Carlos M. Rodriguez Lopez, and Claire T. Roberts. 2016. "Recent Progress towards Understanding the Role of DNA Methylation in Human Placental Development." *Reproduction* 152 (1): R23-30. https://doi.org/10.1530/REP-16-0014.

Bird, A. P. 1986. "CpG-Rich Islands and the Function of DNA Methylation." *Nature* 321 (6067): 209–13. https://doi.org/10.1038/321209a0.

Bird, Adrian. 2002. "DNA Methylation Patterns and Epigenetic Memory." *Genes & Development* 16 (1): 6–21. https://doi.org/10.1101/gad.947102.

Blij, Stephanie, Anthony Parenti, Neeloufar Tabatabai-Yazdi, and Amy Ralston. 2015. "Cdx2 Efficiently Induces Trophoblast Stem-like Cells in Naïve, but Not Primed, Pluripotent Stem Cells." *Stem Cells and Development* 24 (11): 1352–65. https://doi.org/10.1089/scd.2014.0395.

Blumberg, H., A. Eisen, A. Sledziewski, D. Bader, and E. T. Young. 1987. "Two Zinc Fingers of a Yeast Regulatory Protein Shown by Genetic Evidence to Be Essential for Its Function." *Nature* 328 (6129): 443–45. https://doi.org/10.1038/328443a0.

Bock, Christoph, Eleni M. Tomazou, Arie B. Brinkman, Fabian Müller, Femke Simmer, Hongcang Gu, Natalie Jäger, Andreas Gnirke, Hendrik G. Stunnenberg, and Alexander Meissner. 2010. "Quantitative Comparison of Genome-Wide DNA Methylation Mapping Technologies." *Nature Biotechnology* 28 (10): 1106–14. https://doi.org/10.1038/nbt.1681.

Bohmann, D. 1990. "Transcription Factor Phosphorylation: A Link between Signal Transduction and the Regulation of Gene Expression." *Cancer Cells* 2 (11): 337–44. https://www.ncbi.nlm.nih.gov/pubmed/2149275.

Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30 (15): 2114–20. https://doi.org/10.1093/bioinformatics/btu170.

Bommi-Reddy, Archana, Sungmi Park-Chouinard, David N. Mayhew, Esteban Terzo, Aparna Hingway, Michael J. Steinbaugh, Jonathan E. Wilson, Robert J. Sims 3rd, and Andrew R. Conery. 2022. "CREBBP/EP300 Acetyltransferase Inhibition Disrupts FOXA1-Bound Enhancers to Inhibit the Proliferation of ER+ Breast Cancer Cells." *PloS One* 17 (3): e0262378. https://doi.org/10.1371/journal.pone.0262378.

Bouhlel, Mohamed Amine, Melanie Lambert, and Marie-Helene David-Cordonnier. 2015. "Targeting Transcription Factor Binding to DNA by Competing with DNA Binders as an Approach for Controlling Gene Expression." *Current Topics in Medicinal Chemistry* 15 (14): 1323–58. https://doi.org/10.2174/1568026615666150413154713.

Breit, Andreas, Valeria Besik, Hans Jürgen Solinski, Susanne Muehlich, Evi Glas, Stephen J. Yarwood, and Thomas Gudermann. 2015. "Serine-727 Phosphorylation Activates Hypothalamic STAT-3 Independently from Tyrosine-705 Phosphorylation." *Molecular Endocrinology* 29 (3): 445–59. https://doi.org/10.1210/me.2014-1300.

Brenner, Carmen, Rachel Deplus, Céline Didelot, Axelle Loriot, Emmanuelle Viré, Charles De Smet, Arantxa Gutierrez, et al. 2005. "Myc Represses Transcription through Recruitment of DNA Methyltransferase Corepressor." *The EMBO Journal* 24 (2): 336–46. https://doi.org/10.1038/sj.emboj.7600509.

Brindle, P. K., and M. R. Montminy. 1992. "The CREB Family of Transcription Activators." *Current Opinion in Genetics & Development* 2 (2): 199–204. https://doi.org/10.1016/s0959-437x(05)80274-6.

Brinkman, Arie B., Femke Simmer, Kelong Ma, Anita Kaan, Jingde Zhu, and Hendrik G. Stunnenberg. 2010. "Whole-Genome DNA Methylation Profiling Using MethylCap-Seq." *Methods* 52 (3): 232–36. https://doi.org/10.1016/j.ymeth.2010.06.012.

Brockdorff, Neil. 2019. "Localized Accumulation of Xist RNA in X Chromosome Inactivation." *Open Biology* 9 (12): 190213. https://doi.org/10.1098/rsob.190213.

Bundo, Miki, Fumiko Sunaga, Junko Ueda, Kiyoto Kasai, Tadafumi Kato, and Kazuya Iwamoto. 2012. "A Systematic Evaluation of Whole Genome Amplification of Bisulfite-Modified DNA." *Clinical Epigenetics* 4 (1): 22. https://doi.org/10.1186/1868-7083-4-22.

Buratowski, S., and H. Zhou. 1993. "Functional Domains of Transcription Factor TFIIB."

Proceedings of the National Academy of Sciences of the United States of America 90 (12): 5633–37. https://doi.org/10.1073/pnas.90.12.5633.

Burda, P., P. Laslo, and T. Stopka. 2010. "The Role of PU.1 and GATA-1 Transcription Factors during Normal and Leukemogenic Hematopoiesis." *Leukemia* 24 (7): 1249–57. https://doi.org/10.1038/leu.2010.104.

Caglar, Hasan Onur, and Zekeriya Duzgun. 2023. "Identification of Upregulated Genes in Glioblastoma and Glioblastoma Cancer Stem Cells Using Bioinformatics Analysis." *Gene* 848 (January): 146895. https://doi.org/10.1016/j.gene.2022.146895.

Cai, Changmeng, Hongyun Wang, Youyuan Xu, Shaoyong Chen, and Steven P. Balk. 2009. "Reactivation of Androgen Receptor-Regulated TMPRSS2:ERG Gene Expression in Castration-Resistant Prostate Cancer." *Cancer Research* 69 (15): 6027–32. https://doi.org/10.1158/0008-5472.CAN-09-0395.

Campbell, Stephanie A., and Brad G. Hoffman. 2016. "Chromatin Regulators in Pancreas Development and Diabetes." *Trends in Endocrinology and Metabolism: TEM* 27 (3): 142–52. https://doi.org/10.1016/j.tem.2015.12.005.

Canse, Ceren, Ecem Yildirim, and Aylin Yaba. 2023. "Overview of Junctional Complexes during Mammalian Early Embryonic Development." *Frontiers in Endocrinology* 14 (April): 1150017. https://doi.org/10.3389/fendo.2023.1150017.

Cardoso, W. V. 1995. "Transcription Factors and Pattern Formation in the Developing Lung." *The American Journal of Physiology* 269 (4 Pt 1): L429-42. https://doi.org/10.1152/ajplung.1995.269.4.L429.

Castro-Mondragon, Jaime A., Rafael Riudavets-Puig, Ieva Rauluseviciute, Roza Berhanu Lemma, Laura Turchi, Romain Blanc-Mathieu, Jeremy Lucas, et al. 2022. "JASPAR 2022: The 9th Release of the Open-Access Database of Transcription Factor Binding Profiles." *Nucleic Acids Research* 50 (D1): D165–73. https://doi.org/10.1093/nar/gkab1113.

Cervantes-Camacho, Carla, Alicia Beltrán-Langarica, Ann-Karoll Ochoa-Uribe, Meytha Marsch-Moreno, Jorge-Tonatiuh Ayala-Sumuano, Cristina Velez-delValle, and Walid Kuri-Harcuch. 2015. "The Transient Expression of Klf4 and Klf5 during Adipogenesis Depends on GSK3β Activity." *Adipocyte* 4 (4): 248–55. https://doi.org/10.1080/21623945.2015.1007823.

Cheedipudi, Sirisha, Oriana Genolet, and Gergana Dobreva. 2014. "Epigenetic Inheritance of Cell Fates during Embryonic Development." *Frontiers in Genetics* 5 (February): 19. https://doi.org/10.3389/fgene.2014.00019.

Cheloufi, Sihem, and Konrad Hochedlinger. 2017. "Emerging Roles of the Histone Chaperone CAF-1 in Cellular Plasticity." *Current Opinion in Genetics & Development* 46 (October): 83–94. https://doi.org/10.1016/j.gde.2017.06.004.

Chinetti, G., J. C. Fruchart, and B. Staels. 2000. "Peroxisome Proliferator-Activated Receptors (PPARs): Nuclear Receptors at the Crossroads between Lipid Metabolism and Inflammation." *Inflammation Research: Official Journal of the European Histamine Research Society ... [et Al.]* 49 (10): 497–505. https://doi.org/10.1007/s000110050622.

Cirillo, Lisa Ann, Frank Robert Lin, Isabel Cuesta, Dara Friedman, Michal Jarnik, and Kenneth S. Zaret. 2002. "Opening of Compacted Chromatin by Early Developmental Transcription Factors HNF3 (FoxA) and GATA-4." *Molecular Cell* 9 (2): 279–89. https://doi.org/10.1016/s1097-2765(02)00459-8.

Coleman, R. A., A. K. Taggart, S. Burma, J. J. Chicca 2nd, and B. F. Pugh. 1999. "TFIIA Regulates TBP and TFIID Dimers." *Molecular Cell* 4 (3): 451–57. https://doi.org/10.1016/s1097-2765(00)80453-0.

Collings, Clayton K., and John N. Anderson. 2017. "Links between DNA Methylation and Nucleosome Occupancy in the Human Genome." *Epigenetics & Chromatin* 10 (April):

18. https://doi.org/10.1186/s13072-017-0125-5.

Compe, Emmanuel, Carlos M. Genes, Cathy Braun, Frederic Coin, and Jean-Marc Egly. 2019. "TFIIE Orchestrates the Recruitment of the TFIIH Kinase Module at Promoter before Release during Transcription." *Nature Communications* 10 (1): 2084. https://doi.org/10.1038/s41467-019-10131-1.

Cotton, Allison M., Lucia Lam, Joslynn G. Affleck, Ian M. Wilson, Maria S. Peñaherrera, Deborah E. McFadden, Michael S. Kobor, Wan L. Lam, Wendy P. Robinson, and Carolyn J. Brown. 2011. "Chromosome-Wide DNA Methylation Analysis Predicts Human Tissue-Specific X Inactivation." *Human Genetics* 130 (2): 187–201. https://doi.org/10.1007/s00439-011-1007-8.

Cui, Jiujie, Min Shi, Ming Quan, and Keping Xie. 2013. "Regulation of EMT by KLF4 in Gastrointestinal Cancer." *Current Cancer Drug Targets* 13 (9): 986–95. https://doi.org/10.2174/15680096113136660104.

Culig, Zoran, and Frédéric R. Santer. 2014. "Androgen Receptor Signaling in Prostate Cancer." *Cancer Metastasis Reviews* 33 (2–3): 413–27. https://doi.org/10.1007/s10555-013-9474-0.

Cvekl, Ales, Ying Yang, Bharesh K. Chauhan, and Kveta Cveklova. 2004. "Regulation of Gene Expression by Pax6 in Ocular Cells: A Case of Tissue-Preferred Expression of Crystallins in Lens." *The International Journal of Developmental Biology* 48 (8–9): 829–44. https://doi.org/10.1387/ijdb.041866ac.

Dady, Alwyn, Emmanuelle Havis, Virginie Escriou, Martin Catala, and Jean-Loup Duband. 2014. "Junctional Neurulation: A Unique Developmental Program Shaping a Discrete Region of the Spinal Cord Highly Susceptible to Neural Tube Defects." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 34 (39): 13208–21. https://doi.org/10.1523/JNEUROSCI.1850-14.2014.

Dahm, Ralf. 2005. "Friedrich Miescher and the Discovery of DNA." *Developmental Biology* 278 (2): 274–88. https://doi.org/10.1016/j.ydbio.2004.11.028.

Damante, G., D. Fabbro, L. Pellizzari, D. Civitareale, S. Guazzi, M. Polycarpou-Schwartz, S. Cauci, F. Quadrifoglio, S. Formisano, and R. Di Lauro. 1994. "Sequence-Specific DNA Recognition by the Thyroid Transcription Factor-1 Homeodomain." *Nucleic Acids Research* 22 (15): 3075–83. https://doi.org/10.1093/nar/22.15.3075.

Davidson, Mercy M., Claudia Nesti, Lluis Palenzuela, Winsome F. Walker, Evelyn Hernandez, Lev Protas, Michio Hirano, and Nithila D. Isaac. 2005. "Novel Cell Lines Derived from Adult Human Ventricular Cardiomyocytes." *Journal of Molecular and Cellular Cardiology* 39 (1): 133–47. https://doi.org/10.1016/j.yjmcc.2005.03.003.

Deaton, Aimée M., and Adrian Bird. 2011. "CpG Islands and the Regulation of Transcription." *Genes & Development* 25 (10): 1010–22. https://doi.org/10.1101/gad.2037511.

Desai, Shruti, Sophia C. E. Bolick, Michelle Maurin, and Kenneth L. Wright. 2009. "PU.1 Regulates Positive Regulatory Domain I-Binding Factor 1/Blimp-1 Transcription in Lymphoma Cells." *Journal of Immunology* 183 (9): 5778–87. https://doi.org/10.4049/jimmunol.0901120.

Diehl, Adam G., and Alan P. Boyle. 2018. "Conserved and Species-Specific Transcription Factor Co-Binding Patterns Drive Divergent Gene Regulation in Human and Mouse." *Nucleic Acids Research* 46 (4): 1878–94. https://doi.org/10.1093/nar/gky018.

Dobrinić, Paula, Aleksander T. Szczurek, and Robert J. Klose. 2021. "PRC1 Drives Polycomb-Mediated Gene Repression by Controlling Transcription Initiation and Burst Frequency." *Nature Structural & Molecular Biology* 28 (10): 811–24. https://doi.org/10.1038/s41594-021-00661-y.

Donovan, Benjamin T., Hengye Chen, Priit Eek, Zhiyuan Meng, Caroline Jipa, Song Tan, Lu Bai, and Michael G. Poirier. 2023. "Basic Helix-Loop-Helix Pioneer Factors Interact

with the Histone Octamer to Invade Nucleosomes and Generate Nucleosome-Depleted Regions." *Molecular Cell* 83 (8): 1251-1263.e6. https://doi.org/10.1016/j.molcel.2023.03.006.

Dror, Iris, Tamar Golan, Carmit Levy, Remo Rohs, and Yael Mandel-Gutfreund. 2015. "A Widespread Role of the Motif Environment in Transcription Factor Binding across Diverse Protein Families." *Genome Research* 25 (9): 1268–80. https://doi.org/10.1101/gr.184671.114.

Drouin, Jacques. 2014. "Minireview: Pioneer Transcription Factors in Cell Fate Specification." *Molecular Endocrinology* 28 (7): 989–98. https://doi.org/10.1210/me.2014-1084.

Du, James X., Beth B. McConnell, and Vincent W. Yang. 2010. "A Small Ubiquitin-Related Modifier-Interacting Motif Functions as the Transcriptional Activation Domain of Krüppel-like Factor 4." *The Journal of Biological Chemistry* 285 (36): 28298–308. https://doi.org/10.1074/jbc.M110.101717.

Dyer, Matthew, Quy Xiao Xuan Lin, Sofiia Shapoval, Denis Thieffry, and Touati Benoukraf. 2022. "Methmotif.org 2023: The Transcription Factor Binding Motifs Database That Integrates Cell-Specific Functions, Co-Factors and DNA Methylation Profiles." *Research Square*. https://doi.org/10.21203/rs.3.rs-2253164/v1.

Echigoya, Kenta, Masako Koyama, Lumi Negishi, Yoshimasa Takizawa, Yuka Mizukami, Hideki Shimabayashi, Akari Kuroda, and Hitoshi Kurumizaka. 2020. "Nucleosome Binding by the Pioneer Transcription Factor OCT4." *Scientific Reports* 10 (1): 11832. https://doi.org/10.1038/s41598-020-68850-1.

Edlund, H. 2001. "Factors Controlling Pancreatic Cell Differentiation and Function." *Diabetologia* 44 (9): 1071–79. https://doi.org/10.1007/s001250100623.

Egozcue, S., J. Blanco, F. Vidal, and J. Egozcue. 2002. "Diploid Sperm and the Origin of Triploidy." *Human Reproduction* 17 (1): 5–7. https://doi.org/10.1093/humrep/17.1.5.

Elstrodt, Fons, Antoinette Hollestelle, Jord H. A. Nagel, Michael Gorin, Marijke Wasielewski, Ans van den Ouweland, Sofia D. Merajver, Stephen P. Ethier, and Mieke Schutte. 2006. "BRCA1 Mutation Analysis of 41 Human Breast Cancer Cell Lines Reveals Three New Deleterious Mutants." *Cancer Research* 66 (1): 41–45. https://doi.org/10.1158/0008-5472.CAN-05-2853.

Elvenes, Julianne, Eva Sjøttem, Turid Holm, Geir Bjørkøy, and Terje Johansen. 2010. "Pax6 Localizes to Chromatin-Rich Territories and Displays a Slow Nuclear Mobility Altered by Disease Mutations." *Cellular and Molecular Life Sciences: CMLS* 67 (23): 4079–94. https://doi.org/10.1007/s00018-010-0429-0.

Engel, J. D., H. Beug, J. H. LaVail, M. W. Zenke, K. Mayo, M. W. Leonard, K. P. Foley, Z. Yang, J. M. Kornhauser, and L. J. Ko. 1992. "Cis and Trans Regulation of Tissue-Specific Transcription." *Journal of Cell Science. Supplement* 16: 21–31. https://doi.org/10.1242/jcs.1992.supplement_16.4.

Fabre, P. J., M. Leleu, B. Mascrez, Q. Lo Giudice, J. Cobb, and D. Duboule. 2018. "Heterogeneous Combinatorial Expression of Hoxd Genes in Single Cells during Limb Development." *BMC Biology* 16 (1): 101. https://doi.org/10.1186/s12915-018-0570-z.

Falender, Allison E., Richard N. Freiman, Kenneth G. Geles, Kirk C. Lo, Keumsil Hwang, Dolores J. Lamb, Patricia L. Morris, Robert Tjian, and Joanne S. Richards. 2005. "Maintenance of Spermatogenesis Requires TAF4b, a Gonad-Specific Subunit of TFIID." *Genes & Development* 19 (7): 794–803. https://doi.org/10.1101/gad.1290105.

Fang, Dong, Yoshiaki Tsuji, and Vijayasaradhi Setaluri. 2002. "Selective Down-regulation of Tyrosinase Family Gene TYRP1 by Inhibition of the Activity of Melanocyte Transcription Factor, MITF." *Nucleic Acids Research* 30 (14): 3096–3106. https://doi.org/10.1093/nar/gkf424.

Fang, He, Ana R. Tronco, Giancarlo Bonora, Truong Nguyen, Jitendra Thakur, Joel B. Berletch,

Galina N. Filippova, et al. 2023. "CTCF-Mediated Insulation and Chromatin Environment Modulate Car5b Escape from X Inactivation." *BioRxiv : The Preprint Server for Biology*, May. https://doi.org/10.1101/2023.05.04.539469.

Felipe, Cayke, Jaeoh Shin, and Anatoly B. Kolomeisky. 2022. "How Pioneer Transcription Factors Search for Target Sites on Nucleosomal DNA." *The Journal of Physical Chemistry. B* 126 (22): 4061–68. https://doi.org/10.1021/acs.jpcb.2c01931.

Fernandez Garcia, Meilin, Cedric D. Moore, Katharine N. Schulz, Oscar Alberto, Greg Donague, Melissa M. Harrison, Heng Zhu, and Kenneth S. Zaret. 2019. "Structural Features of Transcription Factors Associating with Nucleosome Binding." *Molecular Cell* 75 (5): 921-932.e6. https://doi.org/10.1016/j.molcel.2019.06.009.

Francois, Mathias, Prue Donovan, and Frank Fontaine. 2020. "Modulating Transcription Factor Activity: Interfering with Protein-Protein Interaction Networks." *Seminars in Cell & Developmental Biology* 99 (March): 12–19. https://doi.org/10.1016/j.semcdb.2018.07.019.

Frejtag, W., Y. Zhang, R. Dai, M. G. Anderson, and N. F. Mivechi. 2001. "Heat Shock Factor-4 (HSF-4a) Represses Basal Transcription through Interaction with TFIIF." *The Journal of Biological Chemistry* 276 (18): 14685–94. https://doi.org/10.1074/jbc.M009224200.

Fuglerud, Bettina M., Marit Ledsaak, Marie Rogne, Ragnhild Eskeland, and Odd S. Gabrielsen. 2018. "The Pioneer Factor Activity of C-Myb Involves Recruitment of P300 and Induction of Histone Acetylation Followed by Acetylation-Induced Chromatin Dissociation." *Epigenetics & Chromatin* 11 (1): 35. https://doi.org/10.1186/s13072-018-0208-y.

Fujimoto, Mitsuaki, Eiichi Takaki, Ryosuke Takii, Ke Tan, Ramachandran Prakasam, Naoki Hayashida, Shun-Ichiro Iemura, Tohru Natsume, and Akira Nakai. 2012. "RPA Assists HSF1 Access to Nucleosomal DNA by Recruiting Histone Chaperone FACT." *Molecular Cell* 48 (2): 182–94. https://doi.org/10.1016/j.molcel.2012.07.026.

Fukaya, Takashi. 2023. "Enhancer Dynamics: Unraveling the Mechanism of Transcriptional Bursting." *Science Advances* 9 (31): eadj3366. https://doi.org/10.1126/sciadv.adj3366.

Fukushima, A., A. Okuda, M. Nishimoto, N. Seki, T. A. Hori, and M. Muramatsu. 1998. "Characterization of Functional Domains of an Embryonic Stem Cell Coactivator UTF1 Which Are Conserved and Essential for Potentiation of ATF-2 Activity." *The Journal of Biological Chemistry* 273 (40): 25840–49. https://doi.org/10.1074/jbc.273.40.25840.

Galupa, Rafael, and Edith Heard. 2018. "X-Chromosome Inactivation: A Crossroads Between Chromosome Architecture and Gene Regulation." *Annual Review of Genetics* 52 (November): 535–66. https://doi.org/10.1146/annurev-genet-120116-024611.

Garin-Shkolnik, Tali, Assaf Rudich, Gökhan S. Hotamisligil, and Menachem Rubinstein. 2014. "FABP4 Attenuates PPARγ and Adipogenesis and Is Inversely Correlated with PPARγ in Adipose Tissues." *Diabetes* 63 (3): 900–911. https://doi.org/10.2337/db13-0436.

Garros-Regulez, Laura, Idoia Garcia, Estefania Carrasco-Garcia, Aquilino Lantero, Paula Aldaz, Leire Moreno-Cugnon, Olatz Arrizabalaga, et al. 2016. "Targeting SOX2 as a Therapeutic Strategy in Glioblastoma." *Frontiers in Oncology* 6 (October): 222. https://doi.org/10.3389/fonc.2016.00222.

Gerace, L. 1995. "Nuclear Export Signals and the Fast Track to the Cytoplasm." *Cell* 82 (3): 341–44. https://doi.org/10.1016/0092-8674(95)90420-4.

Gheorghe, Marius, Geir Kjetil Sandve, Aziz Khan, Jeanne Chèneby, Benoit Ballester, and Anthony Mathelier. 2019. "A Map of Direct TF-DNA Interactions in the Human Genome." *Nucleic Acids Research* 47 (4): e21. https://doi.org/10.1093/nar/gky1210.

Gillespie, Marc, Bijay Jassal, Ralf Stephan, Marija Milacic, Karen Rothfels, Andrea Senff-Ribeiro, Johannes Griss, et al. 2022. "The Reactome Pathway Knowledgebase 2022." *Nucleic Acids Research* 50 (D1): D687–92. https://doi.org/10.1093/nar/gkab1028.

Goldstein, Leonard D., Yi Cao, Gregoire Pau, Michael Lawrence, Thomas D. Wu, Somasekar Seshagiri, and Robert Gentleman. 2016. "Prediction and Quantification of Splice Events from RNA-Seq Data." *PloS One* 11 (5): e0156132. https://doi.org/10.1371/journal.pone.0156132.

Gong, Ting, Heather Borgard, Zao Zhang, Shaoqiu Chen, Zitong Gao, and Youping Deng. 2022. "Analysis and Performance Assessment of the Whole Genome Bisulfite Sequencing Data Workflow: Currently Available Tools and a Practical Guide to Advance DNA Methylation Studies." *Small Methods* 6 (3): e2101251. https://doi.org/10.1002/smtd.202101251.

Greig, Kylie T., Sebastian Carotta, and Stephen L. Nutt. 2008. "Critical Roles for C-Myb in Hematopoietic Progenitor Cells." *Seminars in Immunology* 20 (4): 247–56. https://doi.org/10.1016/j.smim.2008.05.003.

Grünewald, Thomas G. P., Florencia Cidre-Aranaz, Didier Surdez, Eleni M. Tomazou, Enrique de Álava, Heinrich Kovar, Poul H. Sorensen, Olivier Delattre, and Uta Dirksen. 2018. "Ewing Sarcoma." *Nature Reviews. Disease Primers* 4 (1): 5. https://doi.org/10.1038/s41572-018-0003-x.

GTEx Consortium. 2015. "Human Genomics. The Genotype-Tissue Expression (GTEx) Pilot Analysis: Multitissue Gene Regulation in Humans." *Science* 348 (6235): 648–60. https://doi.org/10.1126/science.1262110.

Gulshan, Kailash, Bernice Thommandru, and W. Scott Moye-Rowley. 2012. "Proteolytic Degradation of the Yap1 Transcription Factor Is Regulated by Subcellular Localization and the E3 Ubiquitin Ligase Not4." *The Journal of Biological Chemistry* 287 (32): 26796–805. https://doi.org/10.1074/jbc.M112.384719.

Guo, Charles C., and Bogdan Czerniak. 2019. "Bladder Cancer in the Genomic Era." *Archives of Pathology & Laboratory Medicine* 143 (6): 695–704. https://doi.org/10.5858/arpa.2018-0329-RA.

Haberle, Vanja, and Alexander Stark. 2018. "Eukaryotic Core Promoters and the Functional Basis of Transcription Initiation." *Nature Reviews. Molecular Cell Biology* 19 (10): 621–37. https://doi.org/10.1038/s41580-018-0028-8.

Hahn, Yoonsoo, Seung Kyoung Yang, and Jae Hoon Chung. 2005. "Structure and Expression of the Zebrafish Mest Gene, an Ortholog of Mammalian Imprinted Gene PEG1/MEST." *Biochimica et Biophysica Acta* 1731 (2): 125–32. https://doi.org/10.1016/j.bbaexp.2005.09.004.

Hai, T., and T. Curran. 1991. "Cross-Family Dimerization of Transcription Factors Fos/Jun and ATF/CREB Alters DNA Binding Specificity." *Proceedings of the National Academy of Sciences of the United States of America* 88 (9): 3720–24. https://doi.org/10.1073/pnas.88.9.3720.

Haller, D., M. Mackiewicz, S. Gerber, D. Beyer, B. Kullmann, I. Schneider, J. S. Ahmed, and U. Seitzer. 2010. "Cytoplasmic Sequestration of P53 Promotes Survival in Leukocytes Transformed by Theileria." *Oncogene* 29 (21): 3079–86. https://doi.org/10.1038/onc.2010.61.

Hankey, William, Zhong Chen, and Qianben Wang. 2020. "Shaping Chromatin States in Prostate Cancer by Pioneer Transcription Factors." *Cancer Research* 80 (12): 2427–36. https://doi.org/10.1158/0008-5472.CAN-19-3447.

Hansen, Kasper D., Benjamin Langmead, and Rafael A. Irizarry. 2012. "BSmooth: From Whole Genome Bisulfite Sequencing Reads to Differentially Methylated Regions." *Genome Biology* 13 (10): R83. https://doi.org/10.1186/gb-2012-13-10-r83.

Hartman, Mariusz L., and Malgorzata Czyz. 2015. "MITF in Melanoma: Mechanisms behind Its Expression and Activity." *Cellular and Molecular Life Sciences: CMLS* 72 (7): 1249–60. https://doi.org/10.1007/s00018-014-1791-0.

He, Shengbo, and Xiaoqi Feng. 2022. "DNA Methylation Dynamics during Germline Development." *Journal of Integrative Plant Biology* 64 (12): 2240–51. https://doi.org/10.1111/jipb.13422.

Héberlé, Éléa, and Anaïs Flore Bardet. 2019. "Sensitivity of Transcription Factors to DNA Methylation." *Essays in Biochemistry* 63 (6): 727–41. https://doi.org/10.1042/EBC20190033.

Heinlein, Cynthia A., and Chawnshang Chang. 2004. "Androgen Receptor in Prostate Cancer." *Endocrine Reviews* 25 (2): 276–308. https://doi.org/10.1210/er.2002-0032.

Helman, Lee J., and Paul Meltzer. 2003. "Mechanisms of Sarcoma Development." *Nature Reviews. Cancer* 3 (9): 685–94. https://doi.org/10.1038/nrc1168.

Hemberg, Martin, and Gabriel Kreiman. 2011. "Conservation of Transcription Factor Binding Events Predicts Gene Expression across Species." *Nucleic Acids Research* 39 (16): 7092–7102. https://doi.org/10.1093/nar/gkr404.

Henry, N. L., A. M. Campbell, W. J. Feaver, D. Poon, P. A. Weil, and R. D. Kornberg. 1994. "TFIIF-TAF-RNA Polymerase II Connection." *Genes & Development* 8 (23): 2868–78. https://doi.org/10.1101/gad.8.23.2868.

Hermann, Andrea, Rachna Goyal, and Albert Jeltsch. 2004. "The Dnmt1 DNA-(Cytosine-C5)-Methyltransferase Methylates DNA Processively with High Preference for Hemimethylated Target Sites." *The Journal of Biological Chemistry* 279 (46): 48350–59. https://doi.org/10.1074/jbc.M403427200.

Hervouet, Eric, François M. Vallette, and Pierre-François Cartron. 2010. "Dnmt1/Transcription Factor Interactions: An Alternative Mechanism of DNA Methylation Inheritance." *Genes & Cancer* 1 (5): 434–43. https://doi.org/10.1177/1947601910373794.

Hesse, Eric, Hiroaki Saito, Riku Kiviranta, Diego Correa, Kei Yamana, Lynn Neff, Daniel Toben, et al. 2010. "Zfp521 Controls Bone Mass by HDAC3-Dependent Attenuation of Runx2 Activity." *The Journal of Cell Biology* 191 (7): 1271–83. https://doi.org/10.1083/jcb.201009107.

Hill, Caroline S. 2016. "Transcriptional Control by the SMADs." *Cold Spring Harbor Perspectives in Biology* 8 (10). https://doi.org/10.1101/cshperspect.a022079.

Hitzler, Johann K., Joseph Cheung, Yue Li, Stephen W. Scherer, and Alvin Zipursky. 2003. "GATA1 Mutations in Transient Leukemia and Acute Megakaryoblastic Leukemia of Down Syndrome." *Blood* 101 (11): 4301–4. https://doi.org/10.1182/blood-2003-01-0013.

Ho, Y., M. Lewis, and M. Rosenberg. 1982. "Purification and Properties of a Transcriptional Activator. The CII Protein of Phage Lambda." *The Journal of Biological Chemistry* 257 (15): 9128–34. https://www.ncbi.nlm.nih.gov/pubmed/6212585.

Hochberg, Yosef. 1988. "A Sharper Bonferroni Procedure for Multiple Tests of Significance." *Biometrika* 75 (4): 800–802. https://doi.org/10.1093/biomet/75.4.800.

Hoefflin, Rouven, Sabine Harlander, Silvia Schäfer, Patrick Metzger, Fengshen Kuo, Désirée Schönenberger, Mojca Adlesic, et al. 2020. "HIF-1α and HIF-2α Differently Regulate Tumour Development and Inflammation of Clear Cell Renal Cell Carcinoma in Mice." *Nature Communications* 11 (1): 4111. https://doi.org/10.1038/s41467-020-17873-3.

Hoeijmakers, J. H., J. M. Egly, and W. Vermeulen. 1996. "TFIIH: A Key Component in Multiple DNA Transactions." *Current Opinion in Genetics & Development* 6 (1): 26–33. https://doi.org/10.1016/s0959-437x(96)90006-4.

Hoekman, Marco F. M., Frank M. J. Jacobs, Marten P. Smidt, and J. Peter H. Burbach. 2006. "Spatial and Temporal Expression of FoxO Transcription Factors in the Developing and Adult Murine Brain." *Gene Expression Patterns: GEP* 6 (2): 134–40. https://doi.org/10.1016/j.modgep.2005.07.003.

Hogeveen, Kevin N., and Paolo Sassone-Corsi. 2006. "Regulation of Gene Expression in Post-

Meiotic Male Germ Cells: CREM-Signalling Pathways and Male Fertility." *Human Fertility* 9 (2): 73–79. https://doi.org/10.1080/14647270500463400.

Horisawa, Kenichi, and Atsushi Suzuki. 2023. "The Role of Pioneer Transcription Factors in the Induction of Direct Cellular Reprogramming." *Regenerative Therapy* 24 (December): 112–16. https://doi.org/10.1016/j.reth.2023.06.002.

Hsieh, C. L. 1994. "Dependence of Transcriptional Repression on CpG Methylation Density." *Molecular and Cellular Biology* 14 (8): 5487–94. https://doi.org/10.1128/mcb.14.8.5487-5494.1994.

Hu, Jianzhong, Jing Cai, Su Jung Park, Kwanuk Lee, Yuxia Li, Yao Chen, Jae-Young Yun, Tao Xu, and Hunseung Kang. 2021. "N6 -Methyladenosine MRNA Methylation Is Important for Salt Stress Tolerance in Arabidopsis." *The Plant Journal: For Cell and Molecular Biology* 106 (6): 1759–75. https://doi.org/10.1111/tpj.15270.

Hu, Shaohui, Jun Wan, Yijing Su, Qifeng Song, Yaxue Zeng, Ha Nam Nguyen, Jaehoon Shin, et al. 2013. "DNA Methylation Presents Distinct Binding Sites for Human Transcription Factors." *ELife* 2 (September): e00726. https://doi.org/10.7554/eLife.00726.

Huang, Tinghua, Hong Xiao, Qi Tian, Zhen He, Cheng Yuan, Zezhao Lin, Xuejun Gao, and Min Yao. 2022. "Identification of Upstream Transcription Factor Binding Sites in Orthologous Genes Using Mixed Student's t-Test Statistics." *PLoS Computational Biology* 18 (6): e1009773. https://doi.org/10.1371/journal.pcbi.1009773.

Huang, Xinru, Weiwei Zhang, Yongling Liao, Jiabao Ye, and Feng Xu. 2023. "Contemporary Understanding of Transcription Factor Regulation of Terpenoid Biosynthesis in Plants." *Planta* 259 (1): 2. https://doi.org/10.1007/s00425-023-04268-z.

Hulme, Amy J., Jeffrey R. McArthur, Simon Maksour, Sara Miellet, Lezanne Ooi, David J. Adams, Rocio K. Finol-Urdaneta, and Mirella Dottori. 2020. "Molecular and Functional Characterization of Neurogenin-2 Induced Human Sensory Neurons." *Frontiers in Cellular Neuroscience* 14 (December): 600895. https://doi.org/10.3389/fncel.2020.600895.

Hurwitz, Jerard. 2005. "The Discovery of RNA Polymerase." *The Journal of Biological Chemistry* 280 (52): 42477–85. https://doi.org/10.1074/jbc.X500006200.

Huskova, Hana, Katarina Korecka, Josef Karban, Jarmila Vargova, Karina Vargova, Nina Dusilkova, Marek Trneny, and Tomas Stopka. 2015. "Oncogenic MicroRNA-155 and Its Target PU.1: An Integrative Gene Expression Study in Six of the Most Prevalent Lymphomas." *International Journal of Hematology* 102 (4): 441–50. https://doi.org/10.1007/s12185-015-1847-4.

Imanishi, Miki, Chika Imamura, Chika Higashi, Wei Yan, Shigeru Negi, Shiroh Futaki, and Yukio Sugiura. 2010. "Zinc Finger-Zinc Finger Interaction between the Transcription Factors, GATA-1 and Sp1." *Biochemical and Biophysical Research Communications* 400 (4): 625–30. https://doi.org/10.1016/j.bbrc.2010.08.116.

Imhof, A., X. J. Yang, V. V. Ogryzko, Y. Nakatani, A. P. Wolffe, and H. Ge. 1997. "Acetylation of General Transcription Factors by Histone Acetyltransferases." *Current Biology: CB* 7 (9): 689–92. https://doi.org/10.1016/s0960-9822(06)00296-x.

Inoue, Satoshi, Taichi Mizushima, Kazutoshi Fujita, Abdelrazak Meliti, Hiroki Ide, Seiji Yamaguchi, Hiroaki Fushimi, George J. Netto, Norio Nonomura, and Hiroshi Miyamoto. 2017. "GATA3 Immunohistochemistry in Urothelial Carcinoma of the Upper Urinary Tract as a Urothelial Marker and a Prognosticator." *Human Pathology* 64 (June): 83–90. https://doi.org/10.1016/j.humpath.2017.04.003.

Inukai, Sachi, Kian Hong Kock, and Martha L. Bulyk. 2017. "Transcription Factor-DNA Binding: Beyond Binding Site Motifs." *Current Opinion in Genetics & Development* 43 (April): 110–19. https://doi.org/10.1016/j.gde.2017.02.007.

Ito, Masaki. 2005. "Conservation and Diversification of Three-Repeat Myb Transcription

Factors in Plants." *Journal of Plant Research* 118 (1): 61–69. https://doi.org/10.1007/s10265-005-0192-8.

Ivascu, Claudia, Reinhold Wasserkort, Ralf Lesche, Jun Dong, Harald Stein, Andreas Thiel, and Florian Eckhardt. 2007. "DNA Methylation Profiling of Transcription Factor Genes in Normal Lymphocyte Development and Lymphomas." *The International Journal of Biochemistry & Cell Biology* 39 (7–8): 1523–38. https://doi.org/10.1016/j.biocel.2007.02.006.

Jackson, Jacob T., Stephen L. Nutt, and Matthew P. McCormack. 2023. "The Haematopoietically-Expressed Homeobox Transcription Factor: Roles in Development, Physiology and Disease." *Frontiers in Immunology* 14 (June): 1197490. https://doi.org/10.3389/fimmu.2023.1197490.

Jagarlamudi, Krishna R., and Aleksandar Rajkovic. 2011. "Lhx8 in Oocytes Is Required for Survival of Primordial Follicle Pool and Follicular Development." *Biology of Reproduction* 85 (Suppl_1): 31–31. https://doi.org/10.1093/biolreprod/85.s1.31.

Jenkins, Lisa M. Miller, Stewart R. Durell, Sharlyn J. Mazur, and Ettore Appella. 2012. "P53 N-Terminal Phosphorylation: A Defining Layer of Complex Regulation." *Carcinogenesis* 33 (8): 1441–49. https://doi.org/10.1093/carcin/bgs145.

Jiang, Shuai, Tian Li, Zhi Yang, Wei Hu, and Yang Yang. 2018. "Deciphering the Roles of FOXO1 in Human Neoplasms." *International Journal of Cancer. Journal International Du Cancer* 143 (7): 1560–68. https://doi.org/10.1002/ijc.31338.

Jiao, Zhongxian, Ramin Mollaaghababa, William J. Pavan, Anthony Antonellis, Eric D. Green, and Thomas J. Hornyak. 2004. "Direct Interaction of Sox10 with the Promoter of Murine Dopachrome Tautomerase (Dct) and Synergistic Activation of Dct Expression with Mitf." *Pigment Cell Research / Sponsored by the European Society for Pigment Cell Research and the International Pigment Cell Society* 17 (4): 352–62. https://doi.org/10.1111/j.1600-0749.2004.00154.x.

Jin, H., and C. Martin. 1999. "Multifunctionality and Diversity within the Plant MYB-Gene Family." *Plant Molecular Biology* 41 (5): 577–85. https://doi.org/10.1023/a:1006319732410.

Johnsen, O., N. Skammelsrud, L. Luna, M. Nishizawa, H. Prydz, and A. B. Kolstø. 1996. "Small Maf Proteins Interact with the Human Transcription Factor TCF11/Nrf1/LCR-F1." *Nucleic Acids Research* 24 (21): 4289–97. https://doi.org/10.1093/nar/24.21.4289.

Johnson, P. F., and S. L. McKnight. 1989. "Eukaryotic Transcriptional Regulatory Proteins." *Annual Review of Biochemistry* 58: 799–839. https://doi.org/10.1146/annurev.bi.58.070189.004055.

Johnston, S. A., and J. E. Hopper. 1982. "Isolation of the Yeast Regulatory Gene GAL4 and Analysis of Its Dosage Effects on the Galactose/Melibiose Regulon." *Proceedings of the National Academy of Sciences of the United States of America* 79 (22): 6971–75. https://doi.org/10.1073/pnas.79.22.6971.

Jolma, Arttu, Jian Yan, Thomas Whitington, Jarkko Toivonen, Kazuhiro R. Nitta, Pasi Rastas, Ekaterina Morgunova, et al. 2013. "DNA-Binding Specificities of Human Transcription Factors." *Cell* 152 (1–2): 327–39. https://doi.org/10.1016/j.cell.2012.12.009.

Karadkhelkar, Nishant M., Mingliang Lin, Lisa M. Eubanks, and Kim D. Janda. 2023. "Demystifying the Druggability of the MYC Family of Oncogenes." *Journal of the American Chemical Society* 145 (6): 3259–69. https://doi.org/10.1021/jacs.2c12732.

Kazemian, Majid, Hannah Pham, Scot A. Wolfe, Michael H. Brodsky, and Saurabh Sinha. 2013. "Widespread Evidence of Cooperative DNA Binding by Transcription Factors in Drosophila Development." *Nucleic Acids Research* 41 (17): 8237–52. https://doi.org/10.1093/nar/gkt598.

Kele, Julianna, Nicolas Simplicio, Anna L. M. Ferri, Helena Mira, François Guillemot, Ernest

Arenas, and Siew-Lan Ang. 2006. "Neurogenin 2 Is Required for the Development of Ventral Midbrain Dopaminergic Neurons." *Development* 133 (3): 495–505. https://doi.org/10.1242/dev.02223.

Khodadadi, Ehsan, Leila Fahmideh, Ehsaneh Khodadadi, Sounkalo Dao, Mehdi Yousefi, Sepehr Taghizadeh, Mohammad Asgharzadeh, Bahman Yousefi, and Hossein Samadi Kafil. 2021. "Current Advances in DNA Methylation Analysis Methods." *BioMed Research International* 2021 (March): 8827516. https://doi.org/10.1155/2021/8827516.

Kim, Bohyun, Insoon Jang, Kwangsoo Kim, Minsun Jung, Cheol Lee, Jeong Hwan Park, Young A. Kim, and Kyung Chul Moon. 2021. "Comprehensive Gene Expression Analyses of Immunohistochemically Defined Subgroups of Muscle-Invasive Urinary Bladder Urothelial Carcinoma." *International Journal of Molecular Sciences* 22 (2). https://doi.org/10.3390/ijms22020628.

Kim, Heesun, Mikang Sim, Nayoung Park, Kisang Kwon, Junyoung Kim, and Jaebum Kim. 2022. "MsPIPE: A Pipeline for the Analysis and Visualization of Whole-Genome Bisulfite Sequencing Data." *BMC Bioinformatics* 23 (1): 383. https://doi.org/10.1186/s12859-022-04925-2.

Kim, Joshua, and Gerhard A. Coetzee. 2004. "Prostate Specific Antigen Gene Regulation by Androgen Receptor." *Journal of Cellular Biochemistry* 93 (2): 233–41. https://doi.org/10.1002/jcb.20228.

Kim, Tae Hun, Michael L. Nosella, Nicolas Bolik-Coulon, Robert W. Harkness, Shuya Kate Huang, and Lewis E. Kay. 2023. "Correlating Histone Acetylation with Nucleosome Core Particle Dynamics and Function." *Proceedings of the National Academy of Sciences of the United States of America* 120 (15): e2301063120. https://doi.org/10.1073/pnas.2301063120.

Kim, Yea Woon, Yujin Kang, Jin Kang, and Aeri Kim. 2020. "GATA-1-Dependent Histone H3K27 Acetylation Mediates Erythroid Cell-Specific Chromatin Interaction between CTCF Sites." *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 34 (11): 14736–49. https://doi.org/10.1096/fj.202001526R.

Klinge, C. M. 2001. "Estrogen Receptor Interaction with Estrogen Response Elements." *Nucleic Acids Research* 29 (14): 2905–19. https://doi.org/10.1093/nar/29.14.2905.

Koch, Frederic, Romain Fenouil, Marta Gut, Pierre Cauchy, Thomas K. Albert, Joaquin Zacarias-Cabeza, Salvatore Spicuglia, et al. 2011. "Transcription Initiation Platforms and GTF Recruitment at Tissue-Specific Enhancers and Promoters." *Nature Structural & Molecular Biology* 18 (8): 956–63. https://doi.org/10.1038/nsmb.2085.

Kohler, Sarah, and Lisa Ann Cirillo. 2010. "Stable Chromatin Binding Prevents FoxA Acetylation, Preserving FoxA Chromatin Remodeling." *The Journal of Biological Chemistry* 285 (1): 464–72. https://doi.org/10.1074/jbc.M109.063149.

Kolde, Raivo. 2019. "Pheatmap: Pretty Heatmaps. R Package Version 1.0. 12."

Komori, Toshihisa. 2019. "Regulation of Proliferation, Differentiation and Functions of Osteoblasts by Runx2." *International Journal of Molecular Sciences* 20 (7). https://doi.org/10.3390/ijms20071694.

Kosugi, Shunichi, Masako Hasebe, Nobutaka Matsumura, Hideaki Takashima, Etsuko Miyamoto-Sato, Masaru Tomita, and Hiroshi Yanagawa. 2009. "Six Classes of Nuclear Localization Signals Specific to Different Binding Grooves of Importin Alpha." *The Journal of Biological Chemistry* 284 (1): 478–85. https://doi.org/10.1074/jbc.M807017200.

Kouzarides, T., and E. Ziff. 1989. "Leucine Zippers of Fos, Jun and GCN4 Dictate Dimerization Specificity and Thereby Control DNA Binding." *Nature* 340 (6234): 568–71. https://doi.org/10.1038/340568a0.

Kribelbauer, Judith F., Ryan E. Loker, Siqian Feng, Chaitanya Rastogi, Namiko Abe, H. Tomas Rube, Harmen J. Bussemaker, and Richard S. Mann. 2020. "Context-Dependent Gene Regulation by Homeodomain Transcription Factor Complexes Revealed by Shape-Readout Deficient Proteins." *Molecular Cell* 78 (1): 152-167.e11. https://doi.org/10.1016/j.molcel.2020.01.027.

Krueger, Felix, and Simon R. Andrews. 2011. "Bismark: A Flexible Aligner and Methylation Caller for Bisulfite-Seq Applications." *Bioinformatics* 27 (11): 1571–72. https://doi.org/10.1093/bioinformatics/btr167.

Kubik, Slawomir, Maria Jessica Bruzzone, and David Shore. 2017. "Establishing Nucleosome Architecture and Stability at Promoters: Roles of Pioneer Transcription Factors and the RSC Chromatin Remodeler." *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 39 (5). https://doi.org/10.1002/bies.201600237.

Kueh, Hao Yuan, Ameya Champhekar, Stephen L. Nutt, Michael B. Elowitz, and Ellen V. Rothenberg. 2013. "Positive Feedback between PU.1 and the Cell Cycle Controls Myeloid Differentiation." *Science* 341 (6146): 670–73. https://doi.org/10.1126/science.1240831.

Kung, Johnny T., Barry Kesner, Jee Young An, Janice Y. Ahn, Catherine Cifuentes-Rojas, David Colognori, Yesu Jeon, et al. 2015. "Locus-Specific Targeting to the X Chromosome Revealed by the RNA Interactome of CTCF." *Molecular Cell* 57 (2): 361–75. https://doi.org/10.1016/j.molcel.2014.12.006.

Labbé, David P., and Myles Brown. 2018. "Transcriptional Regulation in Prostate Cancer." *Cold Spring Harbor Perspectives in Medicine* 8 (11). https://doi.org/10.1101/cshperspect.a030437.

Laird, Peter W. 2010. "Principles and Challenges of Genomewide DNA Methylation Analysis." *Nature Reviews. Genetics* 11 (3): 191–203. https://doi.org/10.1038/nrg2732.

Lalmansingh, Avin S., Sudipan Karmakar, Yetao Jin, and Akhilesh K. Nagaich. 2012. "Multiple Modes of Chromatin Remodeling by Forkhead Box Proteins." *Biochimica et Biophysica Acta* 1819 (7): 707–15. https://doi.org/10.1016/j.bbagrm.2012.02.018.

Langmead, Ben, and Steven L. Salzberg. 2012. "Fast Gapped-Read Alignment with Bowtie 2." *Nature Methods* 9 (4): 357–59. https://doi.org/10.1038/nmeth.1923.

Law, Julie A., and Steven E. Jacobsen. 2010. "Establishing, Maintaining and Modifying DNA Methylation Patterns in Plants and Animals." *Nature Reviews. Genetics* 11 (3): 204–20. https://doi.org/10.1038/nrg2719.

Le, Nhat-Tu, James F. Martin, Keigi Fujiwara, and Jun-Ichi Abe. 2017. "Sub-Cellular Localization Specific SUMOylation in the Heart." *Biochimica et Biophysica Acta, Molecular Basis of Disease* 1863 (8): 2041–55. https://doi.org/10.1016/j.bbadis.2017.01.018.

Lee, Catherine S., Newman J. Sund, Rüdiger Behr, Pedro L. Herrera, and Klaus H. Kaestner. 2005. "Foxa2 Is Required for the Differentiation of Pancreatic α-Cells." *Developmental Biology* 278 (2): 484–95. https://doi.org/10.1016/j.ydbio.2004.10.012.

Lee, Ji Min, Henrik M. Hammarén, Mikhail M. Savitski, and Sung Hee Baek. 2023. "Control of Protein Stability by Post-Translational Modifications." *Nature Communications* 14 (1): 201. https://doi.org/10.1038/s41467-023-35795-8.

Lee, Kihyun, Hyunwoo Cho, Robert W. Rickert, Qing V. Li, Julian Pulecio, Christina S. Leslie, and Danwei Huangfu. 2019. "FOXA2 Is Required for Enhancer Priming during Pancreatic Differentiation." *Cell Reports* 28 (2): 382-393.e7. https://doi.org/10.1016/j.celrep.2019.06.034.

Lee, Seon-Hyeong, Joon Hee Kang, Ji Sun Ha, Jae-Seon Lee, Su-Jin Oh, Hyun-Jung Choi, Jaewhan Song, and Soo-Youl Kim. 2020. "Transglutaminase 2-Mediated P53 Depletion Promotes Angiogenesis by Increasing HIF-1α-P300 Binding in Renal Cell Carcinoma."

International Journal of Molecular Sciences 21 (14). https://doi.org/10.3390/ijms21145042.

Lemma, Roza B., Marit Ledsaak, Bettina M. Fuglerud, Geir Kjetil Sandve, Ragnhild Eskeland, and Odd S. Gabrielsen. 2021. "Chromatin Occupancy and Target Genes of the Haematopoietic Master Transcription Factor MYB." *Scientific Reports* 11 (1): 9008. https://doi.org/10.1038/s41598-021-88516-w.

Levy, Carmit, Mehdi Khaled, and David E. Fisher. 2006. "MITF: Master Regulator of Melanocyte Development and Melanoma Oncogene." *Trends in Molecular Medicine* 12 (9): 406–14. https://doi.org/10.1016/j.molmed.2006.07.008.

Li, Heng, and Nils Homer. 2010. "A Survey of Sequence Alignment Algorithms for Next-Generation Sequencing." *Briefings in Bioinformatics* 11 (5): 473–83. https://doi.org/10.1093/bib/bbq015.

Li, Yi, Hitoshi Ishiguro, Takashi Kawahara, Eiji Kashiwagi, Koji Izumi, and Hiroshi Miyamoto. 2014. "Loss of GATA3 in Bladder Cancer Promotes Cell Migration and Invasion." *Cancer Biology & Therapy* 15 (4): 428–35. https://doi.org/10.4161/cbt.27631.

Li, Yihao, Ke Jin, Gabi W. van Pelt, Hans van Dam, Xiao Yu, Wilma E. Mesker, Peter Ten Dijke, Fangfang Zhou, and Long Zhang. 2016. "C-Myb Enhances Breast Cancer Invasion and Metastasis through the Wnt/β-Catenin/Axin2 Pathway." *Cancer Research* 76 (11): 3364–75. https://doi.org/10.1158/0008-5472.CAN-15-2302.

Li, Ziming, Jia Huang, Shengping Shen, Zhenping Ding, Qingquan Luo, Zhiwei Chen, and Shun Lu. 2018. "SIRT6 Drives Epithelial-to-Mesenchymal Transition and Metastasis in Non-Small Cell Lung Cancer via Snail-Dependent Transrepression of KLF4." *Journal of Experimental & Clinical Cancer Research: CR* 37 (1): 323. https://doi.org/10.1186/s13046-018-0984-z.

Liang, G., and T. Hai. 1997. "Characterization of Human Activating Transcription Factor 4, a Transcriptional Activator That Interacts with Multiple Domains of CAMP-Responsive Element-Binding Protein (CREB)-Binding Protein." *The Journal of Biological Chemistry* 272 (38): 24088–95. https://doi.org/10.1074/jbc.272.38.24088.

Lin, Quy Xiao Xuan, Denis Thieffry, Sudhakar Jha, and Touati Benoukraf. 2020. "TFregulomeR Reveals Transcription Factors' Context-Specific Features and Functions." *Nucleic Acids Research* 48 (2): e10. https://doi.org/10.1093/nar/gkz1088.

Lindberg, Marie K., Sofia Movérare, Stanko Skrtic, Hui Gao, Karin Dahlman-Wright, Jan-Åke Gustafsson, and Claes Ohlsson. 2003. "Estrogen Receptor (ER)-β Reduces ERα-Regulated Gene Transcription, Supporting a 'Ying Yang' Relationship between ERα and ERβ in Mice." *Molecular Endocrinology* 17 (2): 203–8. https://doi.org/10.1210/me.2002-0206.

Liu, Chia-Feng, William E. Samsa, Guang Zhou, and Véronique Lefebvre. 2017. "Transcriptional Control of Chondrocyte Specification and Differentiation." *Seminars in Cell & Developmental Biology* 62 (February): 34–49. https://doi.org/10.1016/j.semcdb.2016.10.004.

Liu, L. Angela, and Joel S. Bader. 2009. "Structure-Based Ab Initio Prediction of Transcription Factor-Binding Sites." *Methods in Molecular Biology* 541: 23–41. https://doi.org/10.1007/978-1-59745-243-4_2.

Liu, Mengya, and Yun Xu. 2021. "An Effective Method to Resolve Ambiguous Bisulfite-Treated Reads." *BMC Bioinformatics* 22 (1): 283. https://doi.org/10.1186/s12859-021-04204-6.

Liu, Nan, Victoria V. Hargreaves, Qian Zhu, Jesse V. Kurland, Jiyoung Hong, Woojin Kim, Falak Sher, et al. 2018. "Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch." *Cell* 173 (2): 430-442.e17.

https://doi.org/10.1016/j.cell.2018.03.016.

Lizio, Marina, Imad Abugessaisa, Shuhei Noguchi, Atsushi Kondo, Akira Hasegawa, Chung Chau Hon, Michiel de Hoon, et al. 2019. "Update of the FANTOM Web Resource: Expansion to Provide Additional Transcriptome Atlases." *Nucleic Acids Research* 47 (D1): D752–58. https://doi.org/10.1093/nar/gky1099.

Love, J. J., X. Li, D. A. Case, K. Giese, R. Grosschedl, and P. E. Wright. 1995. "Structural Basis for DNA Bending by the Architectural Transcription Factor LEF-1." *Nature* 376 (6543): 791–95. https://doi.org/10.1038/376791a0.

Lu, Juane, Tao Wu, Biao Zhang, Suke Liu, Wenjun Song, Jianjun Qiao, and Haihua Ruan. 2021. "Types of Nuclear Localization Signals and Mechanisms of Protein Import into the Nucleus." *Cell Communication and Signaling: CCS* 19 (1): 60. https://doi.org/10.1186/s12964-021-00741-y.

Lupien, Mathieu, Jérôme Eeckhoute, Clifford A. Meyer, Qianben Wang, Yong Zhang, Wei Li, Jason S. Carroll, X. Shirley Liu, and Myles Brown. 2008. "FoxA1 Translates Epigenetic Signatures into Enhancer-Driven Lineage-Specific Transcription." *Cell* 132 (6): 958–70. https://doi.org/10.1016/j.cell.2008.01.018.

Luse, Donal S. 2012. "Rethinking the Role of TFIIF in Transcript Initiation by RNA Polymerase II." *Transcription* 3 (4): 156–59. https://doi.org/10.4161/trns.20725.

Lyko, Frank. 2018. "The DNA Methyltransferase Family: A Versatile Toolkit for Epigenetic Regulation." *Nature Reviews. Genetics* 19 (2): 81–92. https://doi.org/10.1038/nrg.2017.80.

Ma, J., and M. Ptashne. 1987. "A New Class of Yeast Transcriptional Activators." *Cell* 51 (1): 113–19. https://doi.org/10.1016/0092-8674(87)90015-8.

Ma, Meiji Kit-Wan, Carol Heath, Alan Hair, and Adam G. West. 2011. "Histone Crosstalk Directed by H2B Ubiquitination Is Required for Chromatin Boundary Integrity." *PLoS Genetics* 7 (7): e1002175. https://doi.org/10.1371/journal.pgen.1002175.

Ma, Q., C. Fode, F. Guillemot, and D. J. Anderson. 1999. "Neurogenin1 and Neurogenin2 Control Two Distinct Waves of Neurogenesis in Developing Dorsal Root Ganglia." *Genes & Development* 13 (13): 1717–28. https://doi.org/10.1101/gad.13.13.1717.

MacIsaac, Kenzie D., Kinyui A. Lo, William Gordon, Shmulik Motola, Tali Mazor, and Ernest Fraenkel. 2010. "A Quantitative Model of Transcriptional Regulation Reveals the Influence of Binding Location on Expression." *PLoS Computational Biology* 6 (4): e1000773. https://doi.org/10.1371/journal.pcbi.1000773.

Macrae, Trisha A., Julie Fothergill-Robinson, and Miguel Ramalho-Santos. 2023. "Regulation, Functions and Transmission of Bivalent Chromatin during Mammalian Development." *Nature Reviews. Molecular Cell Biology* 24 (1): 6–26. https://doi.org/10.1038/s41580-022-00518-2.

Marchal, Claire, Pierre-Antoine Defossez, and Benoit Miotto. 2022. "Context-Dependent CpG Methylation Directs Cell-Specific Binding of Transcription Factor ZBTB38." *Epigenetics: Official Journal of the DNA Methylation Society* 17 (13): 2122–43. https://doi.org/10.1080/15592294.2022.2111135.

Marinho, H. Susana, Carla Real, Luísa Cyrne, Helena Soares, and Fernando Antunes. 2014. "Hydrogen Peroxide Sensing, Signaling and Regulation of Transcription Factors." *Redox Biology* 2 (February): 535–62. https://doi.org/10.1016/j.redox.2014.02.006.

Maroui, Mohamed Ali, Ghizlane Maarifi, Francis P. McManus, Frédéric Lamoliatte, Pierre Thibault, and Mounira K. Chelbi-Alix. 2018. "Promyelocytic Leukemia Protein (PML) Requirement for Interferon-Induced Global Cellular SUMOylation." *Molecular & Cellular Proteomics: MCP* 17 (6): 1196–1208. https://doi.org/10.1074/mcp.RA117.000447.

Martianov, Igor, Mohamed-Amin Choukrallah, Arnaud Krebs, Tao Ye, Stephanie Legras,

Erikjan Rijkers, Wilfred Van Ijcken, Bernard Jost, Paolo Sassone-Corsi, and Irwin Davidson. 2010. "Cell-Specific Occupancy of an Extended Repertoire of CREM and CREB Binding Loci in Male Germ Cells." *BMC Genomics* 11 (September): 530. https://doi.org/10.1186/1471-2164-11-530.

Martinez, Ivan, and Daniel Dimaio. 2011. "B-Myb, Cancer, Senescence, and MicroRNAs." *Cancer Research* 71 (16): 5370–73. https://doi.org/10.1158/0008-5472.CAN-11-1044.

Matoba, Shogo, Yuting Liu, Falong Lu, Kumiko A. Iwabuchi, Li Shen, Azusa Inoue, and Yi Zhang. 2014. "Embryonic Development Following Somatic Cell Nuclear Transfer Impeded by Persisting Histone Methylation." *Cell* 159 (4): 884–95. https://doi.org/10.1016/j.cell.2014.09.055.

Matsuzaki, Hitomi, Eiichi Okamura, Daichi Kuramochi, Aki Ushiki, Katsuhiko Hirakawa, Akiyoshi Fukamizu, and Keiji Tanimoto. 2018. "Synthetic DNA Fragments Bearing ICR Cis Elements Become Differentially Methylated and Recapitulate Genomic Imprinting in Transgenic Mice." *Epigenetics & Chromatin* 11 (1): 36. https://doi.org/10.1186/s13072-018-0207-z.

Matys, V., E. Fricke, R. Geffers, E. Gößling, M. Haubrock, R. Hehl, K. Hornischer, et al. 2003. "TRANSFAC ® : Transcriptional Regulation, from Patterns to Profiles." *Nucleic Acids Research* 31 (1): 374–78. https://doi.org/10.1093/nar/gkg108.

Matys, V., O. V. Kel-Margoulis, E. Fricke, I. Liebich, S. Land, A. Barre-Dirrie, I. Reuter, et al. 2006. "TRANSFAC® and Its Module TRANSCompel®: Transcriptional Gene Regulation in Eukaryotes." *Nucleic Acids Research* 34 (suppl_1): D108–10. https://doi.org/10.1093/nar/gkj143.

Mayran, Alexandre, and Jacques Drouin. 2018. "Pioneer Transcription Factors Shape the Epigenetic Landscape." *The Journal of Biological Chemistry* 293 (36): 13795–804. https://doi.org/10.1074/jbc.R117.001232.

Mayran, Alexandre, Konstantin Khetchoumian, Fadi Hariri, Tomi Pastinen, Yves Gauthier, Aurelio Balsalobre, and Jacques Drouin. 2018. "Pioneer Factor Pax7 Deploys a Stable Enhancer Repertoire for Specification of Cell Fate." *Nature Genetics* 50 (2): 259–69. https://doi.org/10.1038/s41588-017-0035-2.

Meehan, R., J. Lewis, S. Cross, X. Nan, P. Jeppesen, and A. Bird. 1992. "Transcriptional Repression by Methylation of CpG." *Journal of Cell Science. Supplement* 16: 9–14. https://doi.org/10.1242/jcs.1992.supplement_16.2.

Meissner, Alexander, Tarjei S. Mikkelsen, Hongcang Gu, Marius Wernig, Jacob Hanna, Andrey Sivachenko, Xiaolan Zhang, et al. 2008. "Genome-Scale DNA Methylation Maps of Pluripotent and Differentiated Cells." *Nature* 454 (7205): 766–70. https://doi.org/10.1038/nature07107.

Meyer, T., J. A. Gustafsson, and J. Carlstedt-Duke. 1997. "Glucocorticoid-Dependent Transcriptional Repression of the Osteocalcin Gene by Competitive Binding at the TATA Box." *DNA and Cell Biology* 16 (8): 919–27. https://doi.org/10.1089/dna.1997.16.919.

Miller, Jaime L., and Patrick A. Grant. 2013. "The Role of DNA Methylation and Histone Modifications in Transcriptional Regulation in Humans." *Sub-Cellular Biochemistry* 61: 289–317. https://doi.org/10.1007/978-94-007-4525-4_13.

Mivelaz, Maxime Nicolas. 2021. "Dynamic Chromatin Invasion and Remodeling by the Transcription Factor Rap1." Lausanne, EPFL. https://doi.org/10.5075/EPFL-THESIS-8932.

Mizzen, C. A., and C. D. Allis. 1998. "Linking Histone Acetylation to Transcriptional Regulation." *Cellular and Molecular Life Sciences: CMLS* 54 (1): 6–20. https://doi.org/10.1007/s000180050121.

Monk, David. 2015. "Germline-Derived DNA Methylation and Early Embryo Epigenetic

Reprogramming: The Selected Survival of Imprints." *The International Journal of Biochemistry & Cell Biology* 67 (October): 128–38. https://doi.org/10.1016/j.biocel.2015.04.014.

Monk, M., M. Boubelik, and S. Lehnert. 1987. "Temporal and Regional Changes in DNA Methylation in the Embryonic, Extraembryonic and Germ Cell Lineages during Mouse Embryo Development." *Development* 99 (3): 371–82. https://doi.org/10.1242/dev.99.3.371.

Mony, Vinod K., Anna Drangowska-Way, Reka Albert, Emma Harrison, Abbas Ghaddar, Mary Kate Horak, Wenfan Ke, and Eyleen J. O'Rourke. 2021. "Context-Specific Regulation of Lysosomal Lipolysis through Network-Level Diverting of Transcription Factor Interactions." *Proceedings of the National Academy of Sciences of the United States of America* 118 (41). https://doi.org/10.1073/pnas.2104832118.

Moustakas, A., S. Souchelnytskyi, and C. H. Heldin. 2001. "Smad Regulation in TGF-Beta Signal Transduction." *Journal of Cell Science* 114 (Pt 24): 4359–69. https://doi.org/10.1242/jcs.114.24.4359.

Mueller, Elisabetta. 2014. "Understanding the Variegation of Fat: Novel Regulators of Adipocyte Differentiation and Fat Tissue Biology." *Biochimica et Biophysica Acta* 1842 (3): 352–57. https://doi.org/10.1016/j.bbadis.2013.05.031.

Mugnaini, Emiliano N., and Nilanjan Ghosh. 2016. "Lymphoma." *Primary Care* 43 (4): 661–75. https://doi.org/10.1016/j.pop.2016.07.012.

Murakami, Yota. 2019. "Phosphorylation of Repressive Histone Code Readers by Casein Kinase 2 Plays Diverse Roles in Heterochromatin Regulation." *Journal of Biochemistry* 166 (1): 3–6. https://doi.org/10.1093/jb/mvz045.

Murphy, L. C., S. L. Simon, A. Parkes, E. Leygue, H. Dotzlaw, L. Snell, S. Troup, A. Adeyinka, and P. H. Watson. 2000. "Altered Expression of Estrogen Receptor Coregulators during Human Breast Tumorigenesis." *Cancer Research* 60 (22): 6266–71. https://www.ncbi.nlm.nih.gov/pubmed/11103781.

Nagasaka, Mai, Chiharu Miyajima, Hiromasa Aoki, Mineyoshi Aoyama, Daisuke Morishita, Yasumichi Inoue, and Hidetoshi Hayashi. 2022. "Insights into Regulators of P53 Acetylation." *Cells* 11 (23). https://doi.org/10.3390/cells11233825.

Nair, Shalima S., Marcel W. Coolen, Clare Stirzaker, Jenny Z. Song, Aaron L. Statham, Dario Strbenac, Mark D. Robinson, and Susan J. Clark. 2011. "Comparison of Methyl-DNA Immunoprecipitation (MeDIP) and Methyl-CpG Binding Domain (MBD) Protein Capture for Genome-Wide DNA Methylation Analysis Reveal CpG Sequence Coverage Bias." *Epigenetics: Official Journal of the DNA Methylation Society* 6 (1): 34–44. https://doi.org/10.4161/epi.6.1.13313.

Nakamura, Manabu T., Barbara E. Yudell, and Juan J. Loor. 2014. "Regulation of Energy Metabolism by Long-Chain Fatty Acids." *Progress in Lipid Research* 53 (January): 124–44. https://doi.org/10.1016/j.plipres.2013.12.001.

Nebert, Daniel W. 2017. "Aryl Hydrocarbon Receptor (AHR): 'Pioneer Member' of the Basic-Helix/Loop/Helix per-Arnt-Sim (BHLH/PAS) Family of 'Sensors' of Foreign and Endogenous Signals." *Progress in Lipid Research* 67 (July): 38–57. https://doi.org/10.1016/j.plipres.2017.06.001.

Nerlov, C., and T. Graf. 1998. "PU.1 Induces Myeloid Lineage Commitment in Multipotent Hematopoietic Progenitors." *Genes & Development* 12 (15): 2403–12. https://doi.org/10.1101/gad.12.15.2403.

Ng, Alex H. M., Parastoo Khoshakhlagh, Jesus Eduardo Rojo Arias, Giovanni Pasquini, Kai Wang, Anka Swiersy, Seth L. Shipman, et al. 2021. "A Comprehensive Library of Human Transcription Factors for Cell Fate Engineering." *Nature Biotechnology* 39 (4): 510–19. https://doi.org/10.1038/s41587-020-0742-6.

Nioi, Paul, Truyen Nguyen, Philip J. Sherratt, and Cecil B. Pickett. 2005. "The Carboxy-Terminal Neh3 Domain of Nrf2 Is Required for Transcriptional Activation." *Molecular and Cellular Biology* 25 (24): 10895–906. https://doi.org/10.1128/MCB.25.24.10895-10906.2005.

Nitta, Kazuhiro R., Arttu Jolma, Yimeng Yin, Ekaterina Morgunova, Teemu Kivioja, Junaid Akhtar, Korneel Hens, et al. 2015. "Conservation of Transcription Factor Binding Specificities across 600 Million Years of Bilateria Evolution." *ELife* 4 (March). https://doi.org/10.7554/eLife.04837.

Niu, Chao, Li Yin, and Haji Akber Aisa. 2018. "Novel Furocoumarin Derivatives Stimulate Melanogenesis in B16 Melanoma Cells by Up-Regulation of MITF and TYR Family via Akt/GSK3β/β-Catenin Signaling Pathways." *International Journal of Molecular Sciences* 19 (3). https://doi.org/10.3390/ijms19030746.

Nosaka, T., T. Kawashima, K. Misawa, K. Ikuta, A. L. Mui, and T. Kitamura. 1999. "STAT5 as a Molecular Regulator of Proliferation, Differentiation and Apoptosis in Hematopoietic Cells." *The EMBO Journal* 18 (17): 4754–65. https://doi.org/10.1093/emboj/18.17.4754.

Obata, Yuuki, Álvaro Castaño, Stefan Boeing, Ana Carina Bon-Frauches, Candice Fung, Todd Fallesen, Mercedes Gomez de Agüero, et al. 2020. "Neuronal Programming by Microbiota Regulates Intestinal Physiology." *Nature* 578 (7794): 284–89. https://doi.org/10.1038/s41586-020-1975-8.

Ohkuma, Y. 1997. "Multiple Functions of General Transcription Factors TFIIE and TFIIH in Transcription: Possible Points of Regulation by Trans-Acting Factors." *Journal of Biochemistry* 122 (3): 481–89. https://doi.org/10.1093/oxfordjournals.jbchem.a021777.

Okano, M., S. Xie, and E. Li. 1998. "Cloning and Characterization of a Family of Novel Mammalian DNA (Cytosine-5) Methyltransferases." *Nature Genetics* 19 (3): 219–20. https://doi.org/10.1038/890.

Okuno, Yutaka, and Hiromichi Yuki. 2012. "PU.1 Is a Tumor Suppressor for B Cell Malignancies." *Oncotarget* 3 (12): 1495–96. https://doi.org/10.18632/oncotarget.800.

Olova, Nelly, Felix Krueger, Simon Andrews, David Oxley, Rebecca V. Berrens, Miguel R. Branco, and Wolf Reik. 2018. "Comparison of Whole-Genome Bisulfite Sequencing Library Preparation Strategies Identifies Sources of Biases Affecting DNA Methylation Data." *Genome Biology* 19 (1): 33. https://doi.org/10.1186/s13059-018-1408-2.

Ooga, Masatoshi, Masataka G. Suzuki, and Fugaku Aoki. 2015. "Involvement of Histone H2B Monoubiquitination in the Regulation of Mouse Preimplantation Development." *The Journal of Reproduction and Development* 61 (3): 179–84. https://doi.org/10.1262/jrd.2014-137.

Ortega, Esther, Srinivasan Rengachari, Ziad Ibrahim, Naghmeh Hoghoughi, Jonathan Gaucher, Alex S. Holehouse, Saadi Khochbin, and Daniel Panne. 2018. "Transcription Factor Dimerization Activates the P300 Acetyltransferase." *Nature* 562 (7728): 538–44. https://doi.org/10.1038/s41586-018-0621-1.

Ortega-Recalde, Oscar, Julian R. Peat, Donna M. Bond, and Timothy A. Hore. 2021. "Estimating Global Methylation and Erasure Using Low-Coverage Whole-Genome Bisulfite Sequencing (WGBS)." *Methods in Molecular Biology* 2272: 29–44. https://doi.org/10.1007/978-1-0716-1294-1_3.

Paige, Lisa A., Dale J. Christensen, Hanne Grøn, John D. Norris, Elizabeth B. Gottlin, Karen M. Padilla, Ching-Yi Chang, et al. 1999. "Estrogen Receptor (ER) Modulators Each Induce Distinct Conformational Changes in ER α and ER β." *Proceedings of the National Academy of Sciences* 96 (7): 3999–4004. https://doi.org/10.1073/pnas.96.7.3999.

Paillard, Guillaume, Cyril Deremble, and Richard Lavery. 2004. "Looking into DNA

Recognition: Zinc Finger Binding Specificity." *Nucleic Acids Research* 32 (22): 6673–82. https://doi.org/10.1093/nar/gkh1003.

Papatheodorou, Irene, Pablo Moreno, Jonathan Manning, Alfonso Muñoz-Pomer Fuentes, Nancy George, Silvie Fexova, Nuno A. Fonseca, et al. 2020. "Expression Atlas Update: From Tissues to Single Cells." *Nucleic Acids Research* 48 (D1): D77–83. https://doi.org/10.1093/nar/gkz947.

Park, Joo-Man, Seong-Ho Jo, Mi-Young Kim, Tae-Hyun Kim, and Yong-Ho Ahn. 2015. "Role of Transcription Factor Acetylation in the Regulation of Metabolic Homeostasis." *Protein & Cell* 6 (11): 804–13. https://doi.org/10.1007/s13238-015-0204-y.

Pataskar, Abhijeet, Johannes Jung, Pawel Smialowski, Florian Noack, Federico Calegari, Tobias Straub, and Vijay K. Tiwari. 2016. "NeuroD1 Reprograms Chromatin and Transcription Factor Landscapes to Induce the Neuronal Program." *The EMBO Journal* 35 (1): 24–45. https://doi.org/10.15252/embj.201591206.

Paul, Rachel, Rodrigue Peraldi, and Marie Kmita. 2024. "The Pioneering Function of the Hox Transcription Factors." *Seminars in Cell & Developmental Biology* 152–153 (January): 85–92. https://doi.org/10.1016/j.semcdb.2022.11.013.

Pavithran, Honey, and Ranjith Kumavath. 2021. "Emerging Role of Pioneer Transcription Factors in Targeted ERα Positive Breast Cancer." *Exploration of Targeted Anti-Tumor Therapy* 2 (1): 26–35. https://doi.org/10.37349/etat.2021.00031.

Payne, Victoria A., Wo-Shing Au, Christopher E. Lowe, Shaikh M. Rahman, Jacob E. Friedman, Stephen O'Rahilly, and Justin J. Rochford. 2009. "C/EBP Transcription Factors Regulate SREBP1c Gene Expression during Adipogenesis." *Biochemical Journal* 425 (1): 215–23. https://doi.org/10.1042/BJ20091112.

Peri, A., and M. Serio. 2000. "The CREM System in Human Spermatogenesis." *Journal of Endocrinological Investigation* 23 (9): 578–83. https://doi.org/10.1007/BF03343779.

Peterson, Tabitha A., Renee L. Benallie, Andrew M. Bradford, Sally C. Pias, Jaron Yazzie, Siamee N. Lor, Zachary M. Haulsee, et al. 2012. "Dimerization in the Grb7 Protein." *Journal of Molecular Recognition: JMR* 25 (8): 427–34. https://doi.org/10.1002/jmr.2205.

Plass, Christoph, and Paul D. Soloway. 2002. "DNA Methylation, Imprinting and Cancer." *European Journal of Human Genetics: EJHG* 10 (1): 6–16. https://doi.org/10.1038/sj.ejhg.5200768.

Ptashne, M. 1988. "How Eukaryotic Transcriptional Activators Work." *Nature* 335 (6192): 683–89. https://doi.org/10.1038/335683a0.

Pulimood, Nisha S., Minerva Contreras, Molly E. Pruitt, Agnieszka Tarasiewicz, and Alexandre E. Medina. 2021. "Phosphorylation of CREB at Serine 142 and 143 Is Essential for Visual Cortex Plasticity." *ENeuro* 8 (5). https://doi.org/10.1523/ENEURO.0217-21.2021.

Qian, Minyi, Jun Liu, Danyang Zhao, Pengpeng Cai, Chuyue Pan, Wenxin Jia, Yingsheng Gao, et al. 2022. "Aryl Hydrocarbon Receptor Deficiency in Intestinal Epithelial Cells Aggravates Alcohol-Related Liver Disease." *Cellular and Molecular Gastroenterology and Hepatology* 13 (1): 233–56. https://doi.org/10.1016/j.jcmgh.2021.08.014.

Qin, Hui, Yiwen Zhang, Yutao Lou, Zongfu Pan, Feifeng Song, Yujia Liu, Tong Xu, Xiaowei Zheng, Xiaoping Hu, and Ping Huang. 2022. "Overview of PROTACs Targeting the Estrogen Receptor: Achievements for Biological and Drug Discovery." *Current Medicinal Chemistry* 29 (22): 3922–44. https://doi.org/10.2174/0929867328666211110101018.

Ran, Hongbiao, Chunyan Li, Ming Zhang, Jincheng Zhong, and Hui Wang. 2023. "Neglected PTM in Animal Adipogenesis: E3-Mediated Ubiquitination." *Gene* 878 (August): 147574. https://doi.org/10.1016/j.gene.2023.147574.

Rashid, Mohsen, Leila Rostami Zadeh, Behzad Baradaran, Ommoleila Molavi, Zeinab Ghesmati, Mehdi Sabzichi, and Fatemeh Ramezani. 2021. "Up-down Regulation of HIF-1α in Cancer Progression." *Gene* 798 (September): 145796. https://doi.org/10.1016/j.gene.2021.145796.

Rauluseviciute, Ieva, Finn Drabløs, and Morten Beck Rye. 2019. "DNA Methylation Data by Sequencing: Experimental Approaches and Recommendations for Tools and Pipelines for Data Analysis." *Clinical Epigenetics* 11 (1): 193. https://doi.org/10.1186/s13148-019-0795-x.

Recaldin, T., and D. J. Fear. 2016. "Transcription Factors Regulating B Cell Fate in the Germinal Centre." *Clinical and Experimental Immunology* 183 (1): 65–75. https://doi.org/10.1111/cei.12702.

Reese, Rebecca M., Kyle T. Helzer, Kaelyn O. Allen, Christy Zheng, Natalia Solodin, and Elaine T. Alarid. 2022. "GRHL2 Enhances Phosphorylated Estrogen Receptor (ER) Chromatin Binding and Regulates ER-Mediated Transcriptional Activation and Repression." *Molecular and Cellular Biology* 42 (10): e0019122. https://doi.org/10.1128/mcb.00191-22.

Reiner, Anat, Daniel Yekutieli, and Yoav Benjamini. 2003. "Identifying Differentially Expressed Genes Using False Discovery Rate Controlling Procedures." *Bioinformatics* 19 (3): 368–75. https://doi.org/10.1093/bioinformatics/btf877.

Reizel, Yitzhak, Ashleigh Morgan, Long Gao, Jonathan Schug, Sarmistha Mukherjee, Meilín Fernández García, Greg Donahue, Joseph A. Baur, Kenneth S. Zaret, and Klaus H. Kaestner. 2021. "FoxA-Dependent Demethylation of DNA Initiates Epigenetic Memory of Cellular Identity." *Developmental Cell* 56 (5): 602-612.e4. https://doi.org/10.1016/j.devcel.2021.02.005.

Ren, Xingxing, Decai Wang, Guorong Zhang, Tingyue Zhou, Zheng Wei, Yi Yang, Yunjiang Zheng, et al. 2023. "Nucleic DHX9 Cooperates with STAT1 to Transcribe Interferon-Stimulated Genes." *Science Advances* 9 (5): eadd5005. https://doi.org/10.1126/sciadv.add5005.

Ren, Yu, Hitomi Suzuki, Krishna Jagarlamudi, Kayla Golnoski, Megan McGuire, Rita Lopes, Vassilis Pachnis, and Aleksandar Rajkovic. 2015. "Lhx8 Regulates Primordial Follicle Activation and Postnatal Folliculogenesis." *BMC Biology* 13 (June): 39. https://doi.org/10.1186/s12915-015-0151-3.

Rhee, Ho Sung, and B. Franklin Pugh. 2011. "Comprehensive Genome-Wide Protein-DNA Interactions Detected at Single-Nucleotide Resolution." *Cell* 147 (6): 1408–19. https://doi.org/10.1016/j.cell.2011.11.013.

Rimel, Jenna K., and Dylan J. Taatjes. 2018. "The Essential and Multifunctional TFIIH Complex." *Protein Science: A Publication of the Protein Society* 27 (6): 1018–37. https://doi.org/10.1002/pro.3424.

Rishi, Vikas, Paramita Bhattacharya, Raghunath Chatterjee, Julian Rozenberg, Jianfei Zhao, Kimberly Glass, Peter Fitzgerald, and Charles Vinson. 2010. "CpG Methylation of Half-CRE Sequences Creates C/EBPalpha Binding Sites That Activate Some Tissue-Specific Genes." *Proceedings of the National Academy of Sciences of the United States of America* 107 (47): 20311–16. https://doi.org/10.1073/pnas.1008688107.

Rizzino, Angie, and Erin L. Wuebben. 2016. "Sox2/Oct4: A Delicately Balanced Partnership in Pluripotent Stem Cells and Embryogenesis." *Biochimica et Biophysica Acta* 1859 (6): 780–91. https://doi.org/10.1016/j.bbagrm.2016.03.006.

Rocchi, Anna, Maria Cristina Manara, Marika Sciandra, Diana Zambelli, Filippo Nardi, Giordano Nicoletti, Cecilia Garofalo, et al. 2010. "CD99 Inhibits Neural Differentiation of Human Ewing Sarcoma Cells and Thereby Contributes to Oncogenesis." *The Journal of Clinical Investigation* 120 (3): 668–80. https://doi.org/10.1172/JCI36667.

Rodríguez, Jéssica, Jesús Mosquera, Jose R. Couceiro, M. Eugenio Vázquez, and José L. Mascareñas. 2015. "The AT-Hook Motif as a Versatile Minor Groove Anchor for Promoting DNA Binding of Transcription Factor Fragments." *Chemical Science* 6 (8): 4767–71. https://doi.org/10.1039/C5SC01415H.

Roeder, Robert G. 2005. "Transcriptional Regulation and the Role of Diverse Coactivators in Animal Cells." *FEBS Letters* 579 (4): 909–15. https://doi.org/10.1016/j.febslet.2004.12.007.

Rosenbauer, Frank, Bronwyn M. Owens, Li Yu, Joseph R. Tumang, Ulrich Steidl, Jeffery L. Kutok, Linda K. Clayton, et al. 2006. "Lymphoid Cell Growth and Transformation Are Suppressed by a Key Regulatory Element of the Gene Encoding PU.1." *Nature Genetics* 38 (1): 27–37. https://doi.org/10.1038/ng1679.

Ruvolo, Vivian, Lorena Navarro, Clare E. Sample, Michael David, Seung Sung, and Sankar Swaminathan. 2003. "The Epstein-Barr Virus SM Protein Induces STAT1 and Interferon-Stimulated Gene Expression." *Journal of Virology* 77 (6): 3690–3701. https://doi.org/10.1128/jvi.77.6.3690-3701.2003.

Ryan, D. 2017. "MethylDackel: A (Mostly) Universal Methylation Extractor for BS-Seq Experiments, 2017."."

Saldanha, Sabita N., and Trygve O. Tollefsbol. 2018. "Chapter 7 - Epigenetic Approaches to Cancer Therapy." In *Epigenetics in Human Disease (Second Edition)*, edited by Trygve O. Tollefsbol, 6:219–47. Academic Press. https://doi.org/10.1016/B978-0-12-812215-0.00007-8.

Satoh, Jun-Ichi, Shinya Obayashi, Hiroko Tabunoki, Taeko Wakana, and Seung U. Kim. 2010. "Stable Expression of Neurogenin 1 Induces LGR5, a Novel Stem Cell Marker, in an Immortalized Human Neural Stem Cell Line HB1.F3." *Cellular and Molecular Neurobiology* 30 (3): 415–26. https://doi.org/10.1007/s10571-009-9466-3.

Schaefer, Ulf, Sebastian Schmeier, and Vladimir B. Bajic. 2011. "TcoF-DB: Dragon Database for Human Transcription Co-Factors and Transcription Factor Interacting Proteins." *Nucleic Acids Research* 39 (Database issue): D106-10. https://doi.org/10.1093/nar/gkq945.

Schiessel, Helmut, and Ralf Blossey. 2020. "Pioneer Transcription Factors in Chromatin Remodeling: The Kinetic Proofreading View." *Physical Review. E* 101 (4–1): 040401. https://doi.org/10.1103/PhysRevE.101.040401.

Schmeier, Sebastian, Tanvir Alam, Magbubah Essack, and Vladimir B. Bajic. 2017. "TcoF-DB v2: Update of the Database of Human and Mouse Transcription Co-Factors and Transcription Factor Interactions." *Nucleic Acids Research* 45 (D1): D145–50. https://doi.org/10.1093/nar/gkw1007.

Schmitz, A., and D. J. Galas. 1980. "Sequence-Specific Interactions of the Tight-Binding I12-X86 Lac Repressor with Non-Operator DNA." *Nucleic Acids Research* 8 (3): 487–506. https://doi.org/10.1093/nar/8.3.487.

Schnepf, Max, Marc von Reutern, Claudia Ludwig, Christophe Jung, and Ulrike Gaul. 2020. "Transcription Factor Binding Affinities and DNA Shape Readout." *IScience* 23 (11): 101694. https://doi.org/10.1016/j.isci.2020.101694.

Schoch, Hannah, and Ted Abel. 2014. "Transcriptional Co-Repressors and Memory Storage." *Neuropharmacology* 80 (May): 53–60. https://doi.org/10.1016/j.neuropharm.2014.01.003.

Schödel, Johannes, Steffen Grampp, Eamonn R. Maher, Holger Moch, Peter J. Ratcliffe, Paul Russo, and David R. Mole. 2016. "Hypoxia, Hypoxia-Inducible Transcription Factors, and Renal Cancer." *European Urology* 69 (4): 646–57. https://doi.org/10.1016/j.eururo.2015.08.007.

Schuh, R., W. Aicher, U. Gaul, S. Côté, A. Preiss, D. Maier, E. Seifert, U. Nauber, C. Schröder,

and R. Kemler. 1986. "A Conserved Family of Nuclear Proteins Containing Structural Elements of the Finger Protein Encoded by Krüppel, a Drosophila Segmentation Gene." *Cell* 47 (6): 1025–32. https://doi.org/10.1016/0092-8674(86)90817-2.

Sela, Itamar, and David B. Lukatsky. 2011. "DNA Sequence Correlations Shape Nonspecific Transcription Factor-DNA Binding Affinity." *Biophysical Journal* 101 (1): 160–66. https://doi.org/10.1016/j.bpj.2011.04.037.

Shaham, Ohad, Yotam Menuchin, Chen Farhy, and Ruth Ashery-Padan. 2012. "Pax6: A Multi-Level Regulator of Ocular Development." *Progress in Retinal and Eye Research* 31 (5): 351–76. https://doi.org/10.1016/j.preteyeres.2012.04.002.

Shaid, S., C. H. Brandts, H. Serve, and I. Dikic. 2013. "Ubiquitination and Selective Autophagy." *Cell Death and Differentiation* 20 (1): 21–30. https://doi.org/10.1038/cdd.2012.72.

Shakya, Arvind, Catherine Callister, Alon Goren, Nir Yosef, Neha Garg, Vahid Khoddami, David Nix, Aviv Regev, and Dean Tantin. 2015. "Pluripotency Transcription Factor Oct4 Mediates Stepwise Nucleosome Demethylation and Depletion." *Molecular and Cellular Biology* 35 (6): 1014–25. https://doi.org/10.1128/MCB.01105-14.

Shang, Anquan, Xuan Wang, Chenzheng Gu, Wenfang Liu, Junjun Sun, Bingjie Zeng, Chen Chen, et al. 2020. "Exosomal MiR-183-5p Promotes Angiogenesis in Colorectal Cancer by Regulation of FOXO1." *Aging* 12 (9): 8352–71. https://doi.org/10.18632/aging.103145.

Shapiro, Michael J., and Virginia Smith Shapiro. 2011. "Transcriptional Repressors, Corepressors and Chromatin Modifying Enzymes in T Cell Development." *Cytokine* 53 (3): 271–81. https://doi.org/10.1016/j.cyto.2010.11.013.

Sherman, Brad T., Ming Hao, Ju Qiu, Xiaoli Jiao, Michael W. Baseler, H. Clifford Lane, Tomozumi Imamichi, and Weizhong Chang. 2022. "DAVID: A Web Server for Functional Enrichment Analysis and Functional Annotation of Gene Lists (2021 Update)." *Nucleic Acids Research* 50 (W1): W216–21. https://doi.org/10.1093/nar/gkac194.

Shi, Xianping, Yueyuan Zheng, Liling Jiang, Bo Zhou, Wei Yang, Liyan Li, Lingwen Ding, et al. 2020. "EWS-FLI1 Regulates and Cooperates with Core Regulatory Circuitry in Ewing Sarcoma." *Nucleic Acids Research* 48 (20): 11434–51. https://doi.org/10.1093/nar/gkaa901.

Shimizu, Ritsuko, Takashi Kuroha, Osamu Ohneda, Xiaoqing Pan, Kinuko Ohneda, Satoru Takahashi, Sjaak Philipsen, and Masayuki Yamamoto. 2004. "Leukemogenesis Caused by Incapacitated GATA-1 Function." *Molecular and Cellular Biology* 24 (24): 10814–25. https://doi.org/10.1128/MCB.24.24.10814-10825.2004.

Shvedunova, Maria, and Asifa Akhtar. 2022. "Modulation of Cellular Processes by Histone and Non-Histone Protein Acetylation." *Nature Reviews. Molecular Cell Biology* 23 (5): 329–49. https://doi.org/10.1038/s41580-021-00441-y.

Singh, Neha, Showket Hussain, Mausumi Bharadwaj, Nandita Kakkar, S. K. Singh, and Ranbir C. Sobti. 2012. "Overexpression of Signal Transducer and Activator of Transcription (STAT-3 and STAT-5) Transcription Factors and Alteration of Suppressor of Cytokine Signaling (SOCS-1) Protein in Prostate Cancer." *Journal of Receptor and Signal Transduction Research* 32 (6): 321–27. https://doi.org/10.3109/10799893.2012.733885.

Smith, Zachary D., and Alexander Meissner. 2013. "DNA Methylation: Roles in Mammalian Development." *Nature Reviews. Genetics* 14 (3): 204–20. https://doi.org/10.1038/nrg3354.

Song, Jikui, Olga Rechkoblit, Timothy H. Bestor, and Dinshaw J. Patel. 2011. "Structure of DNMT1-DNA Complex Reveals a Role for Autoinhibition in Maintenance DNA

Methylation." *Science* 331 (6020): 1036–40. https://doi.org/10.1126/science.1195380.

Soni, Shefali, Nikolay Pchelintsev, Peter D. Adams, and James J. Bieker. 2014. "Transcription Factor EKLF (KLF1) Recruitment of the Histone Chaperone HIRA Is Essential for β-Globin Gene Expression." *Proceedings of the National Academy of Sciences of the United States of America* 111 (37): 13337–42. https://doi.org/10.1073/pnas.1405422111.

Sönmezer, Can, Rozemarijn Kleinendorst, Dilek Imanci, Guido Barzaghi, Laura Villacorta, Dirk Schübeler, Vladimir Benes, Nacho Molina, and Arnaud Regis Krebs. 2021. "Molecular Co-Occupancy Identifies Transcription Factor Binding Cooperativity In Vivo." *Molecular Cell* 81 (2): 255-267.e6. https://doi.org/10.1016/j.molcel.2020.11.015.

Soufi, Abdenour, Meilin Fernandez Garcia, Artur Jaroszewicz, Nebiyu Osman, Matteo Pellegrini, and Kenneth S. Zaret. 2015. "Pioneer Transcription Factors Target Partial DNA Motifs on Nucleosomes to Initiate Reprogramming." *Cell* 161 (3): 555–68. https://doi.org/10.1016/j.cell.2015.03.017.

Stadhouders, Ralph, Guillaume J. Filion, and Thomas Graf. 2019. "Transcription Factors and 3D Genome Conformation in Cell-Fate Decisions." *Nature* 569 (7756): 345–54. https://doi.org/10.1038/s41586-019-1182-7.

Stadler, Michael B., Rabih Murr, Lukas Burger, Robert Ivanek, Florian Lienert, Anne Schöler, Erik van Nimwegen, et al. 2011. "DNA-Binding Factors Shape the Mouse Methylome at Distal Regulatory Regions." *Nature* 480 (7378): 490–95. https://doi.org/10.1038/nature10716.

Stegmaier, Philip, Alexander Kel, Edgar Wingender, and Jürgen Borlak. 2013. "A Discriminative Approach for Unsupervised Clustering of DNA Sequence Motifs." *PLoS Computational Biology* 9 (3): e1002958. https://doi.org/10.1371/journal.pcbi.1002958.

Stephanou, A., D. A. Isenberg, K. Nakajima, and D. S. Latchman. 1999. "Signal Transducer and Activator of Transcription-1 and Heat Shock Factor-1 Interact and Activate the Transcription of the Hsp-70 and Hsp-90beta Gene Promoters." *The Journal of Biological Chemistry* 274 (3): 1723–28. https://doi.org/10.1074/jbc.274.3.1723.

Stirzaker, Clare, Phillippa C. Taberlay, Aaron L. Statham, and Susan J. Clark. 2014. "Mining Cancer Methylomes: Prospects and Challenges." *Trends in Genetics: TIG* 30 (2): 75–84. https://doi.org/10.1016/j.tig.2013.11.004.

Stone, Nicole R., Casey A. Gifford, Reuben Thomas, Karishma J. B. Pratt, Kaitlen Samse-Knapp, Tamer M. A. Mohamed, Ethan M. Radzinsky, et al. 2019. "Context-Specific Transcription Factor Functions Regulate Epigenomic and Transcriptional Dynamics during Cardiac Reprogramming." *Cell Stem Cell* 25 (1): 87-102.e9. https://doi.org/10.1016/j.stem.2019.06.012.

Sun, Haibo, De-Chen Lin, Qi Cao, Brendan Pang, David D. Gae, Victor Kwan Min Lee, Huey Jin Lim, et al. 2017. "Identification of a Novel SYK/c-MYC/MALAT1 Signaling Pathway and Its Potential Therapeutic Value in Ewing Sarcoma." *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 23 (15): 4376–87. https://doi.org/10.1158/1078-0432.CCR-16-2185.

Sun, Linchong, Huafeng Zhang, and Ping Gao. 2022. "Metabolic Reprogramming and Epigenetic Modifications on the Path to Cancer." *Protein & Cell* 13 (12): 877–919. https://doi.org/10.1007/s13238-021-00846-7.

Sunkel, Benjamin, Meng Wang, Stephanie LaHaye, Benjamin J. Kelly, James R. Fitch, Frederic G. Barr, Peter White, and Benjamin Z. Stanton. 2021. "Pioneer Activity of an Oncogenic Fusion Transcription Factor at Inaccessible Chromatin." *BioRxiv*. https://doi.org/10.1101/2020.12.11.420232.

Swain, Niharika, Mansee Thakur, Jigna Pathak, and Biswaranjan Swain. 2020. "SOX2, OCT4

and NANOG: The Core Embryonic Stem Cell Pluripotency Regulators in Oral Carcinogenesis." *Journal of Oral and Maxillofacial Pathology: JOMFP* 24 (2): 368–73. https://doi.org/10.4103/jomfp.JOMFP_22_20.

Swinstead, Erin E., Ville Paakinaho, Diego M. Presman, and Gordon L. Hager. 2016. "Pioneer Factors and ATP-Dependent Chromatin Remodeling Factors Interact Dynamically: A New Perspective: Multiple Transcription Factors Can Effect Chromatin Pioneer Functions through Dynamic Interactions with ATP-Dependent Chromatin Remodeling Factors." *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 38 (11): 1150–57. https://doi.org/10.1002/bies.201600137.

Takayama, Yuki, and G. Marius Clore. 2012. "Interplay between Minor and Major Groove-Binding Transcription Factors Sox2 and Oct1 in Translocation on DNA Studied by Paramagnetic and Diamagnetic NMR *." *The Journal of Biological Chemistry* 287 (18): 14349–63. https://doi.org/10.1074/jbc.M112.352864.

Tan, M. H. Eileen, Jun Li, H. Eric Xu, Karsten Melcher, and Eu-Leong Yong. 2015. "Androgen Receptor: Structure, Role in Prostate Cancer and Drug Discovery." *Acta Pharmacologica Sinica* 36 (1): 3–23. https://doi.org/10.1038/aps.2014.18.

Tao, Lian Jing, Dong Eun Seo, Benjamin Jackson, Natalia B. Ivanova, and Fabio Rinaldo Santori. 2020. "Nuclear Hormone Receptors and Their Ligands: Metabolites in Control of Transcription." *Cells* 9 (12). https://doi.org/10.3390/cells9122606.

Tate, P. H., and A. P. Bird. 1993. "Effects of DNA Methylation on DNA-Binding Proteins and Gene Expression." *Current Opinion in Genetics & Development* 3 (2): 226–31. https://doi.org/10.1016/0959-437x(93)90027-m.

Taylor, Michael V., and Simon M. Hughes. 2017. "Mef2 and the Skeletal Muscle Differentiation Program." *Seminars in Cell & Developmental Biology* 72 (December): 33–44. https://doi.org/10.1016/j.semcdb.2017.11.020.

Timms, Kirsten M., Victor Abkevich, Elisha Hughes, Chris Neff, Julia Reid, Brian Morris, Saritha Kalva, et al. 2014. "Association of BRCA1/2 Defects with Genomic Scores Predictive of DNA Damage Repair Deficiency among Breast Cancer Subtypes." *Breast Cancer Research: BCR* 16 (6): 475. https://doi.org/10.1186/s13058-014-0475-x.

Tiwari, Anil, Chelsea L. Loughner, Sudha Swamynathan, and Shivalingappa K. Swamynathan. 2017. "KLF4 Plays an Essential Role in Corneal Epithelial Homeostasis by Promoting Epithelial Cell Fate and Suppressing Epithelial–Mesenchymal Transition." *Investigative Ophthalmology & Visual Science* 58 (5): 2785–95. https://doi.org/10.1167/iovs.17-21826.

Tiwari, Neha, Nathalie Meyer-Schaller, Phil Arnold, Helena Antoniadis, Mikhail Pachkov, Erik van Nimwegen, and Gerhard Christofori. 2013. "Klf4 Is a Transcriptional Regulator of Genes Critical for EMT, Including Jnk1 (Mapk8)." *PloS One* 8 (2): e57329. https://doi.org/10.1371/journal.pone.0057329.

Toder, R., S. A. Wilcox, M. Smithwick, and J. A. Graves. 1996. "The Human/Mouse Imprinted Genes IGF2, H19, SNRPN and ZNF127 Map to Two Conserved Autosomal Clusters in a Marsupial." *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology* 4 (4): 295–300. https://doi.org/10.1007/BF02263680.

Todeschini, Anne-Laure, Adrien Georges, and Reiner A. Veitia. 2014. "Transcription Factors: Specific DNA Binding and Specific Gene Regulation." *Trends in Genetics: TIG* 30 (6): 211–19. https://doi.org/10.1016/j.tig.2014.04.002.

Tognon, Manuel, Rosalba Giugno, and Luca Pinello. 2023. "A Survey on Algorithms to Characterize Transcription Factor Binding Sites." *Briefings in Bioinformatics* 24 (3). https://doi.org/10.1093/bib/bbad156.

Tonetti, D. A., and V. C. Jordan. 1997. "The Role of Estrogen Receptor Mutations in

Tamoxifen-Stimulated Breast Cancer." *The Journal of Steroid Biochemistry and Molecular Biology* 62 (2–3): 119–28. https://doi.org/10.1016/s0960-0760(97)00034-4.

Tong, Yi-Fan. 2016. "Mutations of NKX2.5 and GATA4 Genes in the Development of Congenital Heart Disease." *Gene* 588 (1): 86–94. https://doi.org/10.1016/j.gene.2016.04.061.

Ura, Hiroki, Kazuhiro Murakami, Tadayuki Akagi, Keita Kinoshita, Shukuro Yamaguchi, Shinji Masui, Hitoshi Niwa, Hiroshi Koide, and Takashi Yokota. 2011. "Eed/Sox2 Regulatory Loop Controls ES Cell Self-Renewal through Histone Methylation and Acetylation." *The EMBO Journal* 30 (11): 2190–2204. https://doi.org/10.1038/emboj.2011.126.

Viatour, Patrick, Marie-Paule Merville, Vincent Bours, and Alain Chariot. 2005. "Phosphorylation of NF-KappaB and IkappaB Proteins: Implications in Cancer and Inflammation." *Trends in Biochemical Sciences* 30 (1): 43–52. https://doi.org/10.1016/j.tibs.2004.11.009.

Vihervaara, Anniina, and Lea Sistonen. 2014. "HSF1 at a Glance." *Journal of Cell Science* 127 (Pt 2): 261–66. https://doi.org/10.1242/jcs.132605.

Vimalraj, S., B. Arumugam, P. J. Miranda, and N. Selvamurugan. 2015. "Runx2: Structure, Function, and Phosphorylation in Osteoblast Differentiation." *International Journal of Biological Macromolecules* 78 (April): 202–8. https://doi.org/10.1016/j.ijbiomac.2015.04.008.

Voorhees, J. J., E. A. Duell, D. A. Chambers, and C. L. Marcelo. 1976. "Regulation of Cell Cycles." *The Journal of Investigative Dermatology* 67 (1): 15–19. https://doi.org/10.1111/1523-1747.ep12512472.

Walsh, C. P., and T. H. Bestor. 1999. "Cytosine Methylation and Mammalian Development." *Genes & Development* 13 (1): 26–34. https://doi.org/10.1101/gad.13.1.26.

Walsh, C. P., J. R. Chaillet, and T. H. Bestor. 1998. "Transcription of IAP Endogenous Retroviruses Is Constrained by Cytosine Methylation." *Nature Genetics* 20 (2): 116–17. https://doi.org/10.1038/2413.

Wang, Dan, Li Yang, Weina Yu, Qian Wu, Jingyao Lian, Feng Li, Shasha Liu, et al. 2019. "Colorectal Cancer Cell-Derived CCL20 Recruits Regulatory T Cells to Promote Chemoresistance via FOXO1/CEBPB/NF-KB Signaling." *Journal for Immunotherapy of Cancer* 7 (1): 215. https://doi.org/10.1186/s40425-019-0701-2.

Wang, Hao, Matthew T. Maurano, Hongzhu Qu, Katherine E. Varley, Jason Gertz, Florencia Pauli, Kristen Lee, et al. 2012. "Widespread Plasticity in CTCF Occupancy Linked to DNA Methylation." *Genome Research* 22 (9): 1680–88. https://doi.org/10.1101/gr.136101.111.

Wang, Hua, Zheng Fan, Pavel V. Shliaha, Matthew Miele, Ronald C. Hendrickson, Xuejun Jiang, and Kristian Helin. 2023. "H3K4me3 Regulates RNA Polymerase II Promoter-Proximal Pause-Release." *Nature* 615 (7951): 339–48. https://doi.org/10.1038/s41586-023-05780-8.

Wang, Liangjun, J. Lesley Brown, Ru Cao, Yi Zhang, Judith A. Kassis, and Richard S. Jones. 2004. "Hierarchical Recruitment of Polycomb Group Silencing Complexes." *Molecular Cell* 14 (5): 637–46. https://doi.org/10.1016/j.molcel.2004.05.009.

Wang, Xiaopei, Yanli Niu, and Yuan Zheng. 2021. "Multiple Functions of MYB Transcription Factors in Abiotic Stress Responses." *International Journal of Molecular Sciences* 22 (11). https://doi.org/10.3390/ijms22116125.

Wang, Zhengpin, Chen-Yu Liu, Yangu Zhao, and Jurrien Dean. 2020. "FIGLA, LHX8 and SOHLH1 Transcription Factor Networks Regulate Mouse Oocyte Growth and Differentiation." *Nucleic Acids Research* 48 (7): 3525–41. https://doi.org/10.1093/nar/gkaa101.

Wang, Zi-Chuan, Jia-Qi Zhang, and Lei Li. 2022. "In Vitro Investigation of Mammalian Early Embryonic Development." *Yi Chuan = Hereditas / Zhongguo Yi Chuan Xue Hui Bian Ji* 44 (4): 269–74. https://doi.org/10.16288/j.yczz.22-034.

Wapenaar, Hannah, and Frank J. Dekker. 2016. "Histone Acetyltransferases: Challenges in Targeting Bi-Substrate Enzymes." *Clinical Epigenetics* 8 (May): 59. https://doi.org/10.1186/s13148-016-0225-2.

Watson, J. D., and F. H. Crick. 1953. "Molecular Structure of Nucleic Acids; a Structure for Deoxyribose Nucleic Acid." *Nature* 171 (4356): 737–38. https://doi.org/10.1038/171737a0.

Weber, Michael, Jonathan J. Davies, David Wittig, Edward J. Oakeley, Michael Haase, Wan L. Lam, and Dirk Schübeler. 2005. "Chromosome-Wide and Promoter-Specific Analyses Identify Sites of Differential DNA Methylation in Normal and Transformed Human Cells." *Nature Genetics* 37 (8): 853–62. https://doi.org/10.1038/ng1598.

Wee, Boyoung, Alexander Pietras, Tatsuya Ozawa, Elena Bazzoli, Ondrej Podlaha, Christophe Antczak, Bengt Westermark, et al. 2016. "ABCG2 Regulates Self-Renewal and Stem Cell Marker Expression but Not Tumorigenicity or Radiation Resistance of Glioma Cells." *Scientific Reports* 6 (July): 25956. https://doi.org/10.1038/srep25956.

Welch, John J., Jason A. Watts, Christopher R. Vakoc, Yu Yao, Hao Wang, Ross C. Hardison, Gerd A. Blobel, Lewis A. Chodosh, and Mitchell J. Weiss. 2004. "Global Regulation of Erythroid Gene Expression by Transcription Factor GATA-1." *Blood* 104 (10): 3136–47. https://doi.org/10.1182/blood-2004-04-1603.

Wendt, Kristy D., and Ali Shilatifard. 2006. "Packing for the Germy: The Role of Histone H4 Ser1 Phosphorylation in Chromatin Compaction and Germ Cell Development." *Genes & Development* 20 (18): 2487–91. https://doi.org/10.1101/gad.1477706.

Wessely, Anja, Theresa Steeb, Carola Berking, and Markus Vincent Heppt. 2021. "How Neural Crest Transcription Factors Contribute to Melanoma Heterogeneity, Cellular Plasticity, and Treatment Resistance." *International Journal of Molecular Sciences* 22 (11). https://doi.org/10.3390/ijms22115761.

Weth, Oliver, and Rainer Renkawitz. 2011. "CTCF Function Is Modulated by Neighboring DNA Binding Factors." *Biochemistry and Cell Biology = Biochimie et Biologie Cellulaire* 89 (5): 459–68. https://doi.org/10.1139/o11-033.

Whyte, Warren A., David A. Orlando, Denes Hnisz, Brian J. Abraham, Charles Y. Lin, Michael H. Kagey, Peter B. Rahl, Tong Ihn Lee, and Richard A. Young. 2013. "Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes." *Cell* 153 (2): 307–19. https://doi.org/10.1016/j.cell.2013.03.035.

Wickham, Hadley, Winston Chang, and Maintainer Hadley Wickham. 2016. "Package 'Ggplot2.'" *Create Elegant Data Visualisations Using the Grammar of Graphics. Version* 2 (1): 1–189. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=af53fd2f5b9e81b6edec0c13e1b3babd34bda399.

Williamson, Iain, Ragnhild Eskeland, Laura A. Lettice, Alison E. Hill, Shelagh Boyle, Graeme R. Grimes, Robert E. Hill, and Wendy A. Bickmore. 2012. "Anterior-Posterior Differences in HoxD Chromatin Topology in Limb Development." *Development* 139 (17): 3157–67. https://doi.org/10.1242/dev.081174.

Willy, P. J., R. Kobayashi, and J. T. Kadonaga. 2000. "A Basal Transcription Factor That Activates or Represses Transcription." *Science* 290 (5493): 982–85. https://doi.org/10.1126/science.290.5493.982.

Wingender, E. 1990. "Transcription Regulating Proteins and Their Recognition Sequences." *Critical Reviews in Eukaryotic Gene Expression* 1 (1): 11–48. https://www.ncbi.nlm.nih.gov/pubmed/1983713.

Wingett, Steven W., and Simon Andrews. 2018. "FastQ Screen: A Tool for Multi-Genome Mapping and Quality Control." *F1000Research* 7 (August): 1338. https://doi.org/10.12688/f1000research.15931.2.

Wolf, Bennett K., Yanding Zhao, Andrew McCray, William H. Hawk, Luke T. Deary, Nicholas W. Sugiarto, Ian S. LaCroix, Scott A. Gerber, Chao Cheng, and Xiaofeng Wang. 2023. "Cooperation of Chromatin Remodeling SWI/SNF Complex and Pioneer Factor AP-1 Shapes 3D Enhancer Landscapes." *Nature Structural & Molecular Biology* 30 (1): 10–21. https://doi.org/10.1038/s41594-022-00880-x.

Wu, Xian, Yichang Chen, Anthony Luz, Guang Hu, and Erik J. Tokar. 2022. "Cardiac Development in the Presence of Cadmium: An in Vitro Study Using Human Embryonic Stem Cells and Cardiac Organoids." *Environmental Health Perspectives* 130 (11): 117002. https://doi.org/10.1289/EHP11208.

Xiao, Xiong, Nan Li, Dapeng Zhang, Bo Yang, Hongmei Guo, and Yuemin Li. 2016. "Generation of Induced Pluripotent Stem Cells with Substitutes for Yamanaka's Four Transcription Factors." *Cellular Reprogramming* 18 (5): 281–97. https://doi.org/10.1089/cell.2016.0020.

Xu, Yinfeng, and Wei Wan. 2023. "Acetylation in the Regulation of Autophagy." *Autophagy* 19 (2): 379–87. https://doi.org/10.1080/15548627.2022.2062112.

Xuan Lin, Quy Xiao, Stephanie Sian, Omer An, Denis Thieffry, Sudhakar Jha, and Touati Benoukraf. 2019. "MethMotif: An Integrative Cell Specific Database of Transcription Factor Binding Motifs Coupled with DNA Methylation Profiles." *Nucleic Acids Research* 47 (D1): D145–54. https://doi.org/10.1093/nar/gky1005.

Yaffe, B. M., and H. H. Samuels. 1984. "Hormonal Regulation of the Growth Hormone Gene. Relationship of the Rate of Transcription to the Level of Nuclear Thyroid Hormone-Receptor Complexes." *The Journal of Biological Chemistry* 259 (10): 6284–91. https://www.ncbi.nlm.nih.gov/pubmed/6327674.

Yan, Bowen, Jennifer Yang, Min Young Kim, Huacheng Luo, Nicholas Cesari, Tao Yang, John Strouboulis, et al. 2021. "HDAC1 Is Required for GATA-1 Transcription Activity, Global Chromatin Occupancy and Hematopoiesis." *Nucleic Acids Research* 49 (17): 9783–98. https://doi.org/10.1093/nar/gkab737.

Yang, Yanfang, Yu He, Xixi Wang, Ziwei Liang, Gu He, Peng Zhang, Hongxia Zhu, Ningzhi Xu, and Shufang Liang. 2017. "Protein SUMOylation Modification and Its Associations with Disease." *Open Biology* 7 (10). https://doi.org/10.1098/rsob.170167.

Yaşar, Pelin, Gamze Ayaz, Sırma Damla User, Gizem Güpür, and Mesut Muyan. 2017. "Molecular Mechanism of Estrogen-Estrogen Receptor Signaling." *Reproductive Medicine and Biology* 16 (1): 4–20. https://doi.org/10.1002/rmb2.12006.

Yin, Xin, Xu Teng, Tianyu Ma, Tianshu Yang, Jingyao Zhang, Miaomiao Huo, Wei Liu, et al. 2022. "RUNX2 Recruits the NuRD(MTA1)/CRL4B Complex to Promote Breast Cancer Progression and Bone Metastasis." *Cell Death and Differentiation* 29 (11): 2203–17. https://doi.org/10.1038/s41418-022-01010-2.

Yin, Yimeng, Ekaterina Morgunova, Arttu Jolma, Eevi Kaasinen, Biswajyoti Sahu, Syed Khund-Sayeed, Pratyush K. Das, et al. 2017. "Impact of Cytosine Methylation on DNA Binding Specificities of Human Transcription Factors." *Science* 356 (6337). https://doi.org/10.1126/science.aaj2239.

Yoshida, Kiyotsugu, and Yoshio Miki. 2004. "Role of BRCA1 and BRCA2 as Regulators of DNA Repair, Transcription, and Cell Cycle in Response to DNA Damage." *Cancer Science* 95 (11): 866–71. https://doi.org/10.1111/j.1349-7006.2004.tb02195.x.

Yuan, Yuan, Guangjian Fan, Yuqi Liu, Lu Liu, Tong Zhang, Pengfei Liu, Qing Tu, et al. 2022. "The Transcription Factor KLF14 Regulates Macrophage Glycolysis and Immune Function by Inhibiting HK2 in Sepsis." *Cellular & Molecular Immunology* 19 (4): 504–

15. https://doi.org/10.1038/s41423-021-00806-5.

Zamanighomi, Mahdi, Zhixiang Lin, Yong Wang, Rui Jiang, and Wing Hung Wong. 2017. "Predicting Transcription Factor Binding Motifs from DNA-Binding Domains, Chromatin Accessibility and Gene Expression Data." *Nucleic Acids Research* 45 (10): 5666–77. https://doi.org/10.1093/nar/gkx358.

Zeng, Yang, and Taiping Chen. 2019. "DNA Methylation Reprogramming during Mammalian Development." *Genes* 10 (4). https://doi.org/10.3390/genes10040257.

Zhang, D. E., S. Hohaus, M. T. Voso, H. M. Chen, L. T. Smith, C. J. Hetherington, and D. G. Tenen. 1996. "Function of PU.1 (Spi-1), C/EBP, and AML1 in Early Myelopoiesis: Regulation of Multiple Myeloid CSF Receptor Promoters." *Current Topics in Microbiology and Immunology* 211: 137–47. https://doi.org/10.1007/978-3-642-85232-9_14.

Zhang, Huiyuan, and Vijay Kuchroo. 2019. "Epigenetic and Transcriptional Mechanisms for the Regulation of IL-10." *Seminars in Immunology* 44 (101324): 101324. https://doi.org/10.1016/j.smim.2019.101324.

Zhang, Li, Yu-Zhi Xu, Xiao-Feng Xiao, Jun Chen, Xue-Qin Zhou, Wen-Yuan Zhu, Zong Dai, and Xiao-Yong Zou. 2015. "Development of Techniques for DNA-Methylation Analysis." *Trends in Analytical Chemistry: TRAC* 72 (October): 114–22. https://doi.org/10.1016/j.trac.2015.03.025.

Zhang, Wen Jie, Yewei Zhou, Yi Zhang, Ying Hua Su, and Tongda Xu. 2023. "Protein Phosphorylation: A Molecular Switch in Plant Signaling." *Cell Reports* 42 (7): 112729. https://doi.org/10.1016/j.celrep.2023.112729.

Zhao, Yingdong, Ming-Chung Li, Mariam M. Konaté, Li Chen, Biswajit Das, Chris Karlovich, P. Mickey Williams, Yvonne A. Evrard, James H. Doroshow, and Lisa M. McShane. 2021. "TPM, FPKM, or Normalized Counts? A Comparative Study of Quantification Measures for the Analysis of RNA-Seq Data from the NCI Patient-Derived Models Repository." *Journal of Translational Medicine* 19 (1): 269. https://doi.org/10.1186/s12967-021-02936-w.

Zhao, Yongbing. 2023. "TFSyntax: A Database of Transcription Factors Binding Syntax in Mammalian Genomes." *Nucleic Acids Research* 51 (D1): D306–14. https://doi.org/10.1093/nar/gkac849.

Zhao, Yun, Xi Wang, and Kai Wang. 2023. "Transcription Factor-Mediated Programming of Stem Cell Fate." *Trends in Cell Biology* 33 (8): 621–24. https://doi.org/10.1016/j.tcb.2023.05.004.

Zhou, Pingzhu, Yan Zhang, Isha Sethi, Lincai Ye, Michael A. Trembley, Yangpo Cao, Brynn N. Akerberg, et al. 2022. "GATA4 Regulates Developing Endocardium Through Interaction With ETS1." *Circulation Research* 131 (11): e152–68. https://doi.org/10.1161/CIRCRESAHA.120.318102.

Zhou, Tao, Jun Yao, and Zhanjiang Liu. 2017. "Gene Ontology, Enrichment Analysis, and Pathway Analysis." In *Bioinformatics in Aquaculture*, 150–68. Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118782392.ch10.

Zhu, Heng, Guohua Wang, and Jiang Qian. 2016. "Transcription Factors as Readers and Effectors of DNA Methylation." *Nature Reviews. Genetics* 17 (9): 551–65. https://doi.org/10.1038/nrg.2016.83.

Zhu, Lili, Krishna Choudhary, Barbara Gonzalez-Teran, Yen-Sin Ang, Reuben Thomas, Nicole R. Stone, Lei Liu, et al. 2022. "Transcription Factor GATA4 Regulates Cell Type-Specific Splicing Through Direct Interaction With RNA in Human Induced Pluripotent Stem Cell-Derived Cardiac Progenitors." *Circulation* 146 (10): 770–87. https://doi.org/10.1161/CIRCULATIONAHA.121.057620.

Zhu, Ling, and Enamul Huq. 2011. "Mapping Functional Domains of Transcription Factors."

Methods in Molecular Biology 754: 167–84. https://doi.org/10.1007/978-1-61779-154-3_9.

Amaya, Clarissa N., and Brad A. Bryan. 2015. "Enrichment of the Embryonic Stem Cell Reprogramming Factors Oct4, Nanog, Myc, and Sox2 in Benign and Malignant Vascular Tumors." *BMC Clinical Pathology* 15 (September): 18.

Averous, Julien, Alain Bruhat, Sylvie Mordier, and Pierre Fafournoux. 2003. "Recent Advances in the Understanding of Amino Acid Regulation of Gene Expression." *The Journal of Nutrition* 133 (6 Suppl 1): 2040S-2045S.

Barger, Carter J., Connor Branick, Linda Chee, and Adam R. Karpf. 2019. "Pan-Cancer Analyses Reveal Genomic Features of FOXM1 Overexpression in Cancer." *Cancers* 11 (2). https://doi.org/10.3390/cancers11020251.

Baumann, Jan, Tatiana I. Ignashkova, Sridhar R. Chirasani, Silvia Ramírez-Peinado, Hamed Alborzinia, Mathieu Gendarme, Kyra Kuhnigk, Valentin Kramer, Ralph K. Lindemann, and Jan H. Reiling. 2018. "Golgi Stress-Induced Transcriptional Changes Mediated by MAPK Signaling and Three ETS Transcription Factors Regulate MCL1 Splicing." *Molecular Biology of the Cell* 29 (1): 42–52.

Boller, Sören, Senthilkumar Ramamoorthy, Duygu Akbas, Robert Nechanitzky, Lukas Burger, Rabih Murr, Dirk Schübeler, and Rudolf Grosschedl. 2016. "Pioneering Activity of the C-Terminal Domain of EBF1 Shapes the Chromatin Landscape for B Cell Programming." *Immunity* 44 (3): 527–41.

Choi, Kyoung-Jae, My Diem Quan, Chuangye Qi, Joo-Hyung Lee, Phoebe S. Tsoi, Mahla Zahabiyon, Aleksandar Bajic, et al. 2022. "NANOG Prion-like Assembly Mediates DNA Bridging to Facilitate Chromatin Reorganization and Activation of Pluripotency." *Nature Cell Biology* 24 (5): 737–47.

Drouin, Jacques. 2014. "Minireview: Pioneer Transcription Factors in Cell Fate Specification." *Molecular Endocrinology* 28 (7): 989–98.

Du, James X., Beth B. McConnell, and Vincent W. Yang. 2010. "A Small Ubiquitin-Related Modifier-Interacting Motif Functions as the Transcriptional Activation Domain of Krüppel-like Factor 4." *The Journal of Biological Chemistry* 285 (36): 28298–308.

Fourier, Nitsan, Maya Zolty, Aviva Azriel, Donato Tedesco, and Ben-Zion Levi. 2020. "MafK Mediates Chromatin Remodeling to Silence IRF8 Expression in Non-Immune Cells in a Cell Type-SpecificManner." *Journal of Molecular Biology* 432 (16): 4544–60.

González-Rico, Francisco J., Cristina Vicente-García, Almudena Fernández, Diego Muñoz-Santos, Lluís Montoliu, Antonio Morales-Hernández, Jaime M. Merino, Angel-Carlos Román, and Pedro M. Fernández-Salguero. 2020. "Alu Retrotransposons Modulate Nanog Expression through Dynamic Changes in Regional Chromatin Conformation via Aryl Hydrocarbon Receptor." *Epigenetics & Chromatin* 13 (1): 15.

He, Aibin, Fei Gu, Yong Hu, Qing Ma, Lillian Yi Ye, Jennifer A. Akiyama, Axel Visel, Len A. Pennacchio, and William T. Pu. 2014. "Dynamic GATA4 Enhancers Shape the Chromatin Landscape Central to Heart Development and Disease." *Nature Communications* 5 (September): 4907.

Hill, Caroline S. 2016. "Transcriptional Control by the SMADs." *Cold Spring Harbor Perspectives in Biology* 8 (10). https://doi.org/10.1101/cshperspect.a022079.

Kashyap, Vasundhra, Naira C. Rezende, Kymora B. Scotland, Sebastian M. Shaffer, Jenny Liao Persson, Lorraine J. Gudas, and Nigel P. Mongan. 2009. "Regulation of Stem Cell Pluripotency and Differentiation Involves a Mutual Regulatory Circuit of the NANOG, OCT4, and SOX2 Pluripotency Transcription Factors with Polycomb Repressive Complexes and Stem Cell MicroRNAs." *Stem Cells and Development* 18 (7): 1093–1108.

Kim, Tae Hun, Michael L. Nosella, Nicolas Bolik-Coulon, Robert W. Harkness, Shuya Kate Huang, and Lewis E. Kay. 2023. "Correlating Histone Acetylation with Nucleosome Core Particle Dynamics and Function." *Proceedings of the National Academy of Sciences of the United States of America* 120 (15): e2301063120.

Komori, Toshihisa. 2018. "Runx2, an Inducer of Osteoblast and Chondrocyte Differentiation." *Histochemistry and Cell Biology* 149 (4): 313–23.

Kusuyama, Joji, Tomokazu Ohnishi, Kenjiro Bandow, Muhammad Subhan Amir, Kaori Shima, Ichiro Semba, and Tetsuya Matsuguchi. 2017. "Constitutive Activation of P46JNK2 Is Indispensable for C/EBPδ Induction in the Initial Stage of Adipogenic Differentiation." *Biochemical Journal* 474 (20): 3421–37.

Liau, Wei-Siang, Phuong Cao Thi Ngoc, and Takaomi Sanda. 2017. "Roles of the RUNX1 Enhancer in Normal Hematopoiesis and Leukemogenesis." *Advances in Experimental Medicine and Biology* 962: 139–47.

Liu, Chia-Feng, William E. Samsa, Guang Zhou, and Véronique Lefebvre. 2017. "Transcriptional Control of Chondrocyte Specification and Differentiation." *Seminars in Cell & Developmental Biology* 62 (February): 34–49.

Liu, Kai, Huiying Hu, Huanyu Jiang, Haidong Zhang, Shanchun Gong, Dongmin Wei, and Zhenkun Yu. 2021. "RUNX1 Promotes MAPK Signaling to Increase Tumor Progression and Metastasis via OPN in Head and Neck Cancer." *Carcinogenesis* 42 (3): 414–22.

McCauslin, Christine Seitz, Victoria Heath, Anna Maria Colangelo, Radek Malik, Sook Lee, Alessandra Mallei, Italo Mocchetti, and Peter F. Johnson. 2006. "CAAT/Enhancer-Binding Protein Delta and CAMP-Response Element-Binding Protein Mediate Inducible Expression of the Nerve Growth Factor Gene in the Central Nervous System." *The Journal of Biological Chemistry* 281 (26): 17681–88.

Michael, Alicia K., Ralph S. Grand, Luke Isbel, Simone Cavadini, Zuzanna Kozicka, Georg Kempf, Richard D. Bunker, et al. 2020. "Mechanisms of OCT4-SOX2 Motif Readout on Nucleosomes." *Science* 368 (6498): 1460–65.

Miller, Jaime L., and Patrick A. Grant. 2013. "The Role of DNA Methylation and Histone Modifications in Transcriptional Regulation in Humans." *Sub-Cellular Biochemistry* 61: 289–317.

Mizzen, C. A., and C. D. Allis. 1998. "Linking Histone Acetylation to Transcriptional Regulation." *Cellular and Molecular Life Sciences: CMLS* 54 (1): 6–20.

Moustakas, A., S. Souchelnytskyi, and C. H. Heldin. 2001. "Smad Regulation in TGF-Beta Signal Transduction." *Journal of Cell Science* 114 (Pt 24): 4359–69.

Nacht, A. Silvina, Roberto Ferrari, Roser Zaurin, Valentina Scabia, José Carbonell-Caballero, Francois Le Dily, Javier Quilez, et al. 2019. "C/EBPα Mediates the Growth Inhibitory Effect of Progestins on Breast Cancer Cells." *The EMBO Journal* 38 (18): e101426.

Nepon-Sixt, Brook S., Victoria L. Bryant, and Mark G. Alexandrow. 2019. "Myc-Driven Chromatin Accessibility Regulates Cdc45 Assembly into CMG Helicases." *Communications Biology* 2 (March): 110.

Ng, Alex H. M., Parastoo Khoshakhlagh, Jesus Eduardo Rojo Arias, Giovanni Pasquini, Kai Wang, Anka Swiersy, Seth L. Shipman, et al. 2021. "A Comprehensive Library of Human Transcription Factors for Cell Fate Engineering." *Nature Biotechnology* 39 (4): 510–19.

Ortega, Esther, Srinivasan Rengachari, Ziad Ibrahim, Naghmeh Hoghoughi, Jonathan Gaucher, Alex S. Holehouse, Saadi Khochbin, and Daniel Panne. 2018. "Transcription Factor Dimerization Activates the P300 Acetyltransferase." *Nature* 562 (7728): 538–44.

Peng, Wei, Liang Chang, Wenqiang Li, Yanan Liu, and Min Zhang. 2022. "OCT4 and SOX2 Specific Cytotoxic T Cells Exhibit Not Only Good Efficiency but Also Synergize PD-

1 Inhibitor (Nivolumab) in Treating Breast Cancer Stem-Like Cells and Drug-Resistant Breast Cancer Mice." *Frontiers in Oncology* 12 (March): 781093.

Peterson, Tabitha A., Renee L. Benallie, Andrew M. Bradford, Sally C. Pias, Jaron Yazzie, Siamee N. Lor, Zachary M. Haulsee, et al. 2012. "Dimerization in the Grb7 Protein." *Journal of Molecular Recognition: JMR* 25 (8): 427–34.

Rachmadi, Lisnawati, Yayi Dwina Billianti Susanto, David Sitinjak, Amelia Fossetta Manatar, Meilania Saraswati, and Marlinda Adham. 2022. "HPV Infection Is Associated with FoxM1 Overexpression in Dysplastic Changes of Sinonasal Inverted Papilloma." *Asian Pacific Journal of Cancer Prevention: APJCP* 23 (12): 4293–98.

Recaldin, T., and D. J. Fear. 2016. "Transcription Factors Regulating B Cell Fate in the Germinal Centre." *Clinical and Experimental Immunology* 183 (1): 65–75.

Rizzino, Angie, and Erin L. Wuebben. 2016. "Sox2/Oct4: A Delicately Balanced Partnership in Pluripotent Stem Cells and Embryogenesis." *Biochimica et Biophysica Acta* 1859 (6): 780–91.

Shvedunova, Maria, and Asifa Akhtar. 2022. "Modulation of Cellular Processes by Histone and Non-Histone Protein Acetylation." *Nature Reviews. Molecular Cell Biology* 23 (5): 329–49.

Stolzenburg, Sabine, Marianne G. Rots, Adriana S. Beltran, Ashley G. Rivenbark, Xinni Yuan, Haili Qian, Brian D. Strahl, and Pilar Blancafort. 2012. "Targeted Silencing of the Oncogenic Transcription Factor SOX2 in Breast Cancer." *Nucleic Acids Research* 40 (14): 6725–40.

Strid, Tobias, Kazuki Okuyama, Johanna Tingvall-Gustafsson, Jacob Kuruvilla, Christina T. Jensen, Stefan Lang, Mahadesh Prasad, et al. 2021. "B Lymphocyte Specification Is Preceded by Extensive Epigenetic Priming in Multipotent Progenitors." *Journal of Immunology* 206 (11): 2700–2713.

Thiaville, Michelle M., Elizabeth E. Dudenhausen, Can Zhong, Yuan-Xiang Pan, and Michael S. Kilberg. 2008. "Deprivation of Protein or Amino Acid Induces C/EBPbeta Synthesis and Binding to Amino Acid Response Elements, but Its Action Is Not an Absolute Requirement for Enhanced Transcription." *Biochemical Journal* 410 (3): 473–84.

Vanzan, Ludovica, Hadrien Soldati, Victor Ythier, Santosh Anand, Simon M. G. Braun, Nicole Francis, and Rabih Murr. 2021. "High Throughput Screening Identifies SOX2 as a Super Pioneer Factor That Inhibits DNA Methylation Maintenance at Its Binding Sites." *Nature Communications* 12 (1): 3337.

Welch, John J., Jason A. Watts, Christopher R. Vakoc, Yu Yao, Hao Wang, Ross C. Hardison, Gerd A. Blobel, Lewis A. Chodosh, and Mitchell J. Weiss. 2004. "Global Regulation of Erythroid Gene Expression by Transcription Factor GATA-1." *Blood* 104 (10): 3136–47.

Xiao, Xiong, Nan Li, Dapeng Zhang, Bo Yang, Hongmei Guo, and Yuemin Li. 2016. "Generation of Induced Pluripotent Stem Cells with Substitutes for Yamanaka's Four Transcription Factors." *Cellular Reprogramming* 18 (5): 281–97.

Yang, G., C. W. Gregory, Q. Shang, D. A. O'Brien, and Y. L. Zhang. 2001. "Differential Expression of CCAAT/Enhancer-Binding Protein-Delta (c/EBPdelta) in Rat Androgen-Dependent Tissues and Human Prostate Cancer." *Journal of Andrology* 22 (3): 471–80.

Yuan, Yuan, Guangjian Fan, Yuqi Liu, Lu Liu, Tong Zhang, Pengfei Liu, Qing Tu, et al. 2022. "The Transcription Factor KLF14 Regulates Macrophage Glycolysis and Immune Function by Inhibiting HK2 in Sepsis." *Cellular & Molecular Immunology* 19 (4): 504–15.