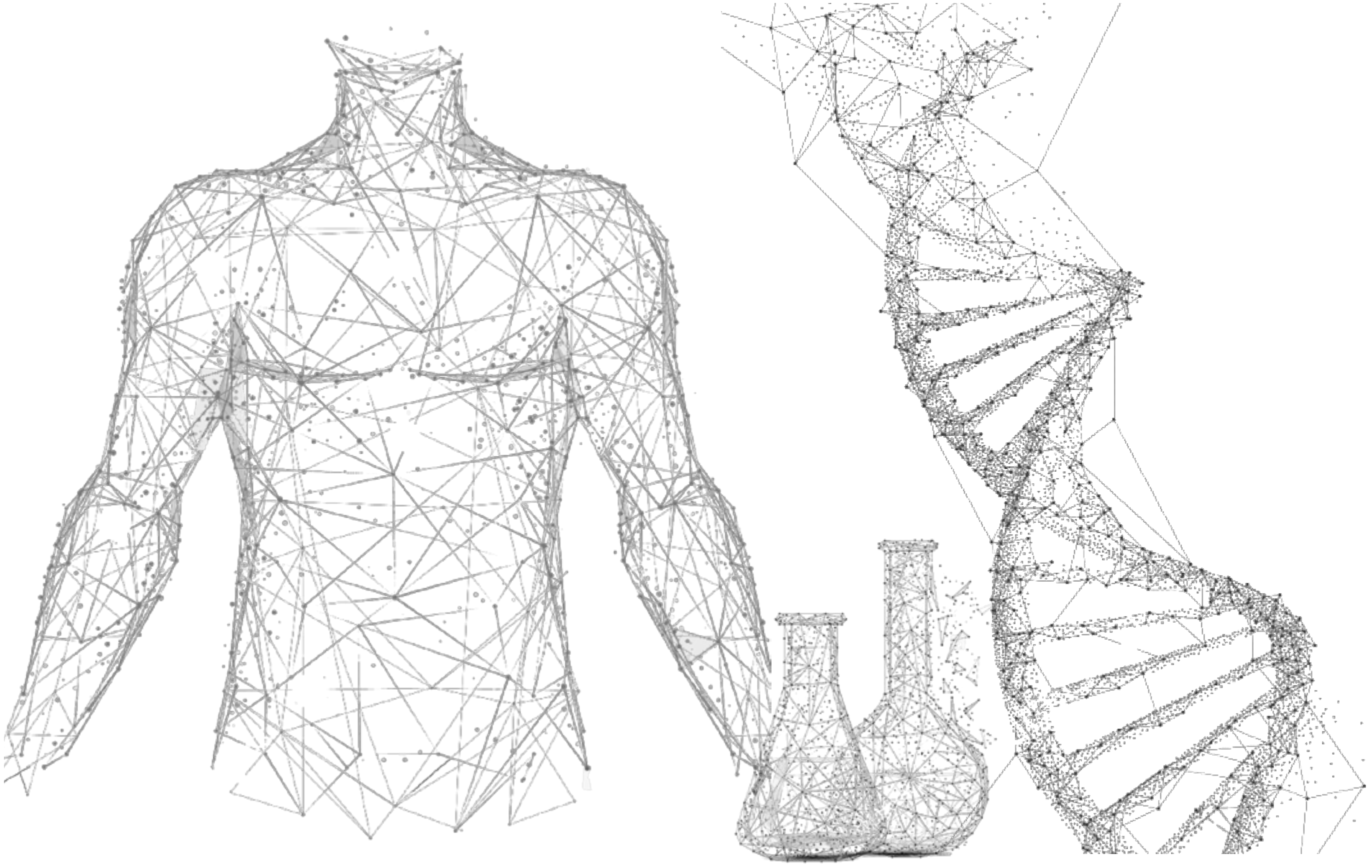


Transcription Factors

San Luis Potosi State University (UASLP) Mexico
Molecular Biology Course, Faculty of Medicine graduate program

Dr. Christian A. García-Sepúlveda
Viral & Human Genomics BSL-3 Laboratory
Last updated January 31, 2025 v3

Transcription factors

Proteins that regulate gene expression by binding to specific DNA sequences (enhancers, promoters, or silencers) to activate or repress transcription.

Also interact with RNA polymerase II, chromatin-modifying enzymes, and other transcription factors.

Key role in development, differentiation, cell cycle regulation, & response to environmental stimuli.

General TFs (also known as universal or basal TFs)
Required for the transcription of all genes (e.g., TFIID, TBP, TFIIB).

Sequence-Specific TFs
Bind to specific DNA motifs to regulate target genes (e.g., p53, NF-κB, MYC, HIF-1).

Pioneer Factors
Can bind closed chromatin and facilitate chromatin remodeling (e.g., FOXA1, SOX2).

General transcription factors

Essential proteins for RNA Pol II transcription initiation.

Form pre-initiation complex (PIC) at core promoter of protein-coding genes.

TFIID TATA-binding protein (TBP) and TBP-associated factors (TAFs) binds to the promoter.

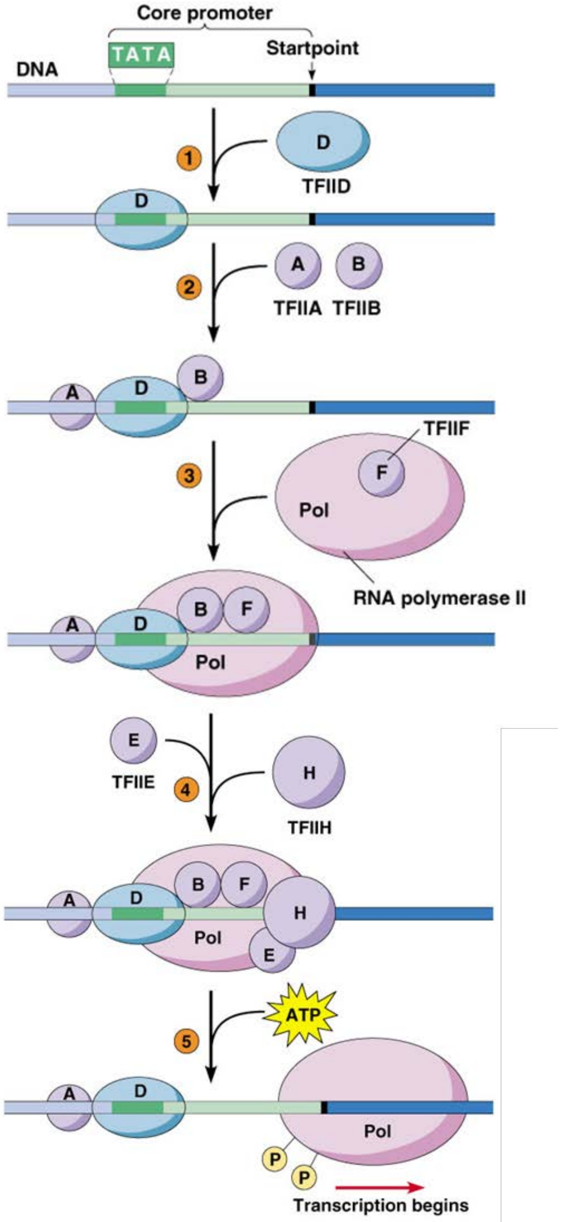
TFIIA Stabilizes TBP-DNA binding, counteracts repressors.

TFIIB Bridges TFIID and RNA polymerase II.

TFIIF Recruits RNA polymerase II to the promoter.

TFIIE Recruits and regulates TFIIH.

TFIIH Contains helicase (XPB, XPD) and kinase (CDK7) activities; unwinds DNA, phosphorylates the C-terminal domain (CTD) of RNA polymerase II, initiating transcription.



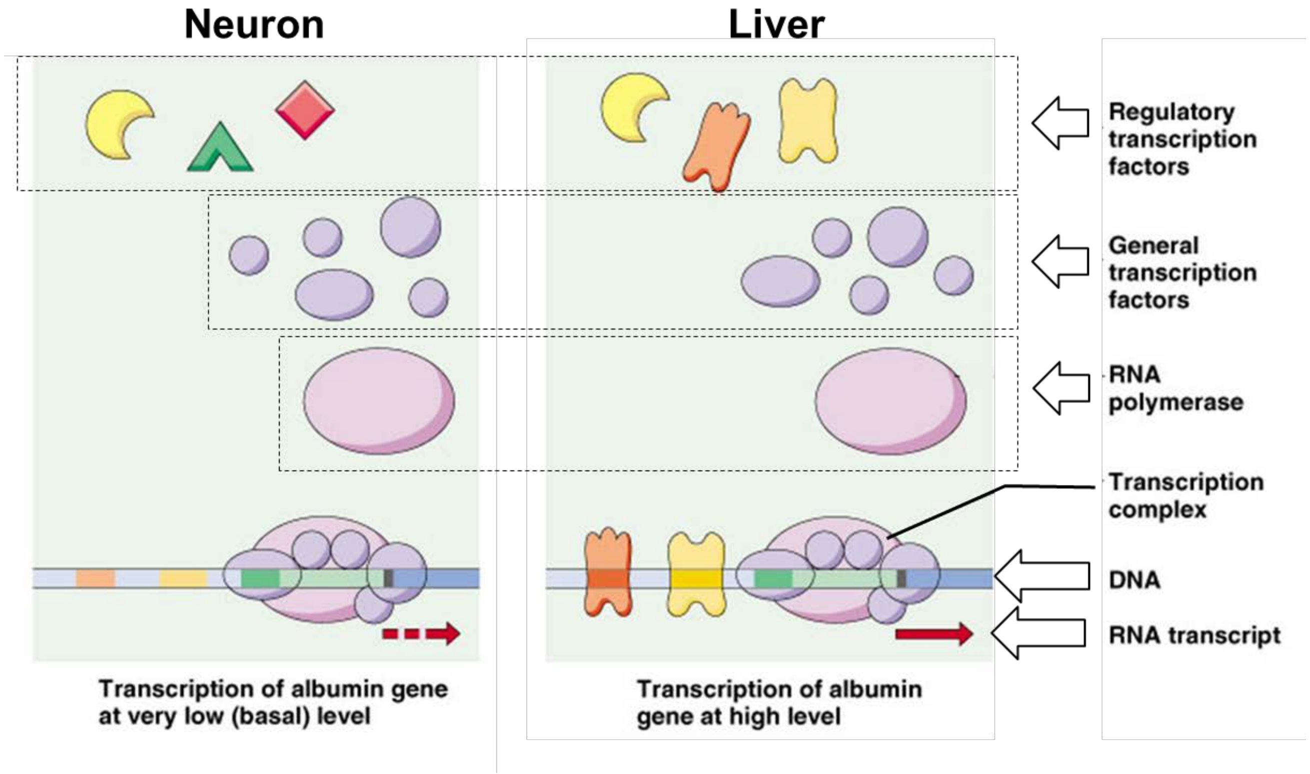
Copyright © 2003 Pearson Education, Inc., publishing as Benjamin Cummings.

Special transcription factors

Tissue and lineage specific accessory transcription factors.

Complement or further enhance GTF function.

Combinatorial hypothesis Explains tissue or lineage differences in gene expression



How do transcription factors access DNA

TF is synthesized when needed.

TF requires phosphorylation.

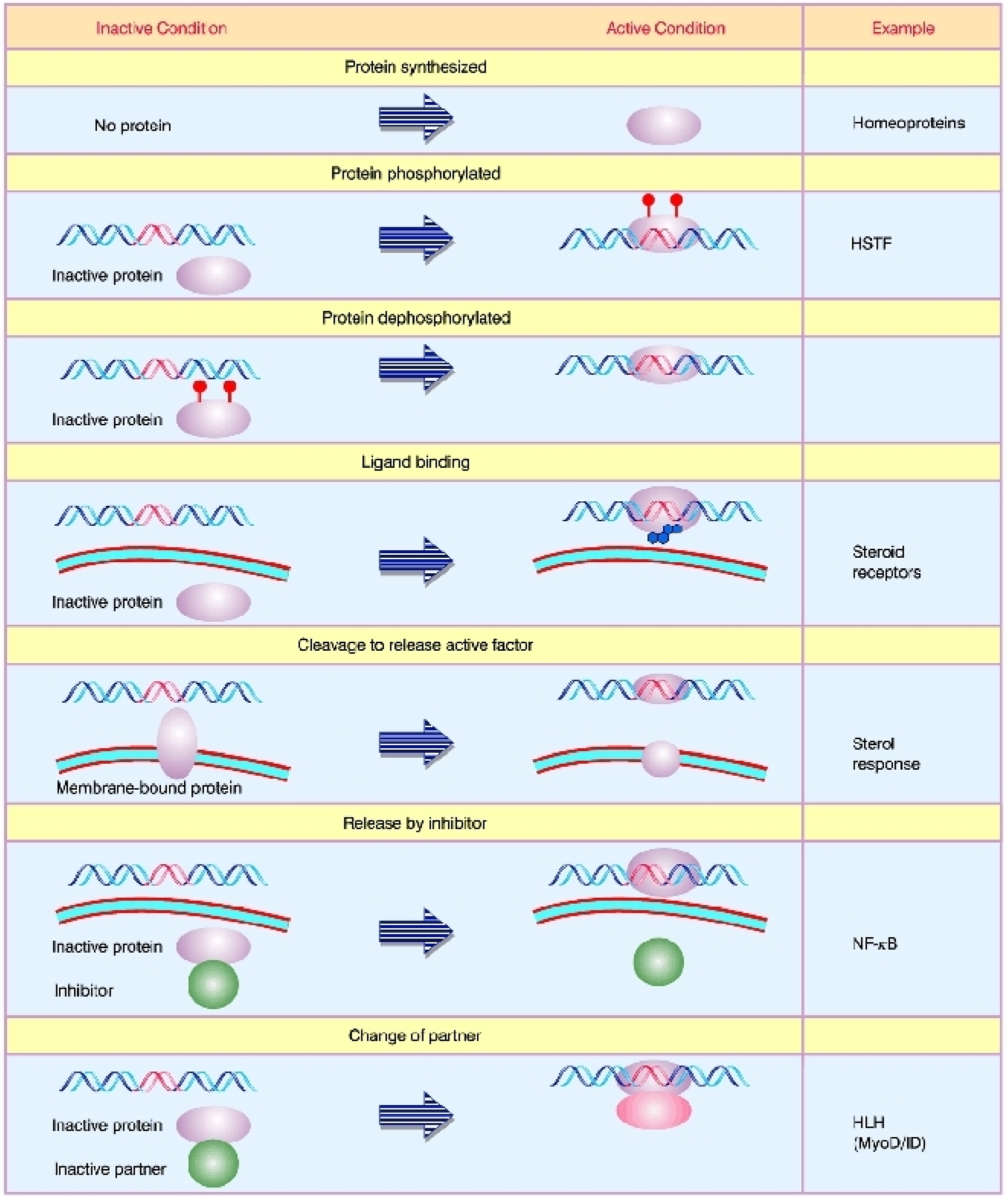
TF requires dephosphorylation.

TF need steroid for nuclear translocation.

TF has to be cleaved from sterol.

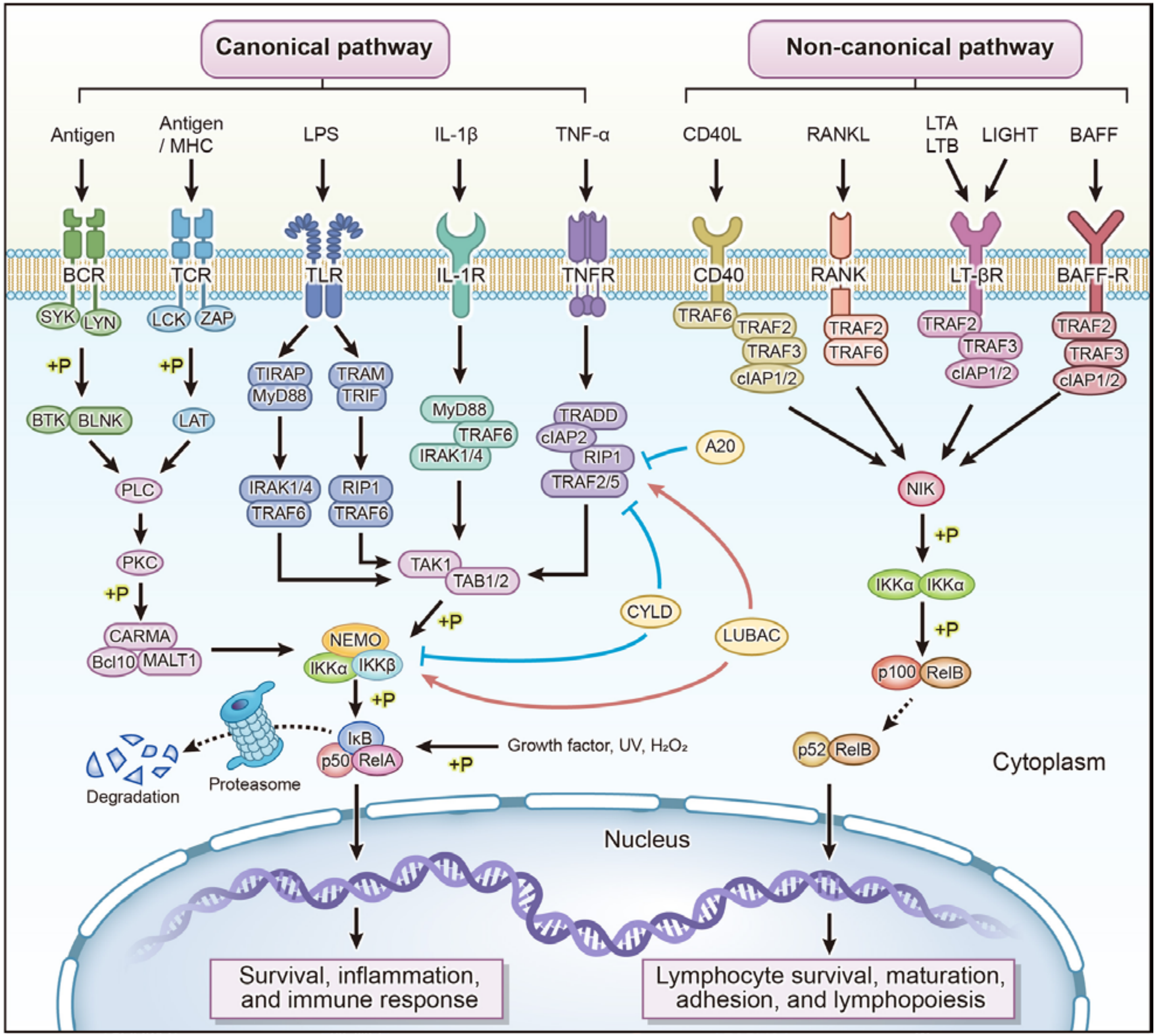
TF needs to loose inhibitor/repressor.

TF need top exhange partners.



Inactive Condition	Active Condition	Example
Protein synthesized		
No protein		Homeoproteins
Protein phosphorylated		
Inactive protein		HSTF
Protein dephosphorylated		
Inactive protein		
Ligand binding		
Inactive protein		Steroid receptors
Cleavage to release active factor		
Membrane-bound protein		Sterol response
Release by inhibitor		
Inactive protein Inhibitor		NF-κB
Change of partner		
Inactive protein Inactive partner		HLH (MyoD/ID)

Immune signaling cascades related to transcription factors



NF-kappaB in biology and targeted therapy. Guo Q, et al. Signal Transduct Target Ther. 2024 Mar 4;9(1):53.

Major families of sequence-specific TFs

Homeodomain TFs
Regulate embryonic development (HOX, PAX, SOX).

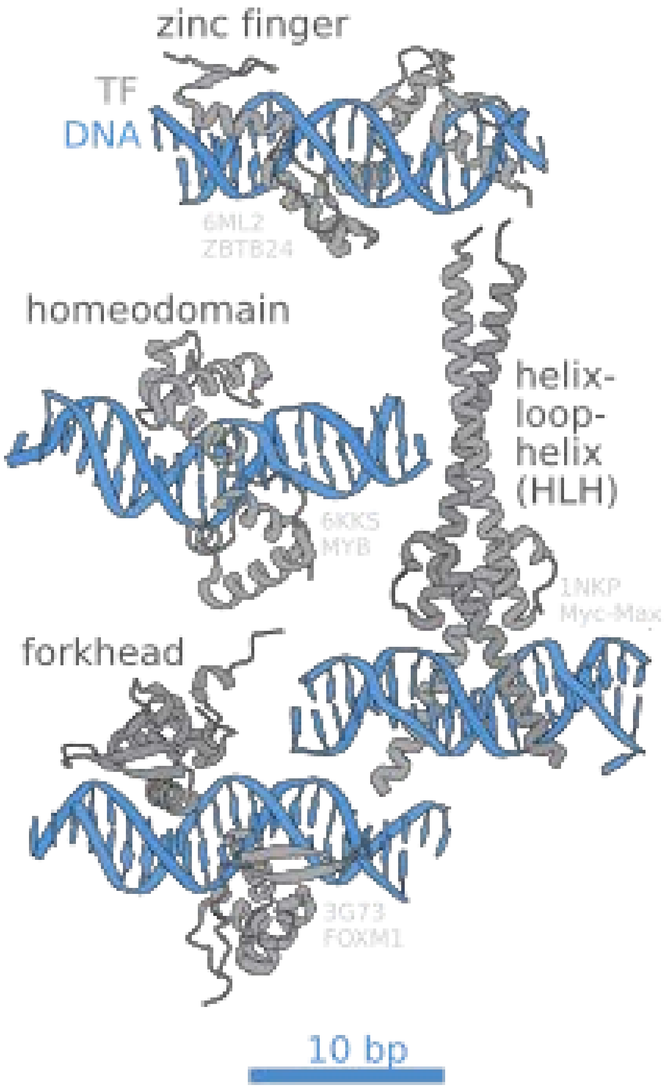
Zinc Finger TFs
Largest family, DNA repair & cell differentiation (SP1, WT1, KLF4).

Leucine Zipper TFs
Function as dimers, regulate proliferation & stress responses (AP-1, C/EBP, MYC).

Helix-Loop-Helix (bHLH) TFs
Control neurogenesis, myogenesis, and metabolism (MYOD, HIF-1, c-MYC).

Nuclear Receptors
Ligand-activated TFs regulating metabolism and immune responses (ER, GR, PPARs).

Forkhead TFs (FOXO, FOXP)
Involved in longevity, apoptosis, and immune function.



Homeodomain

Homeodomain is a ~60 amino acid DNA-binding motif that forms a helix-turn-helix (HTH).

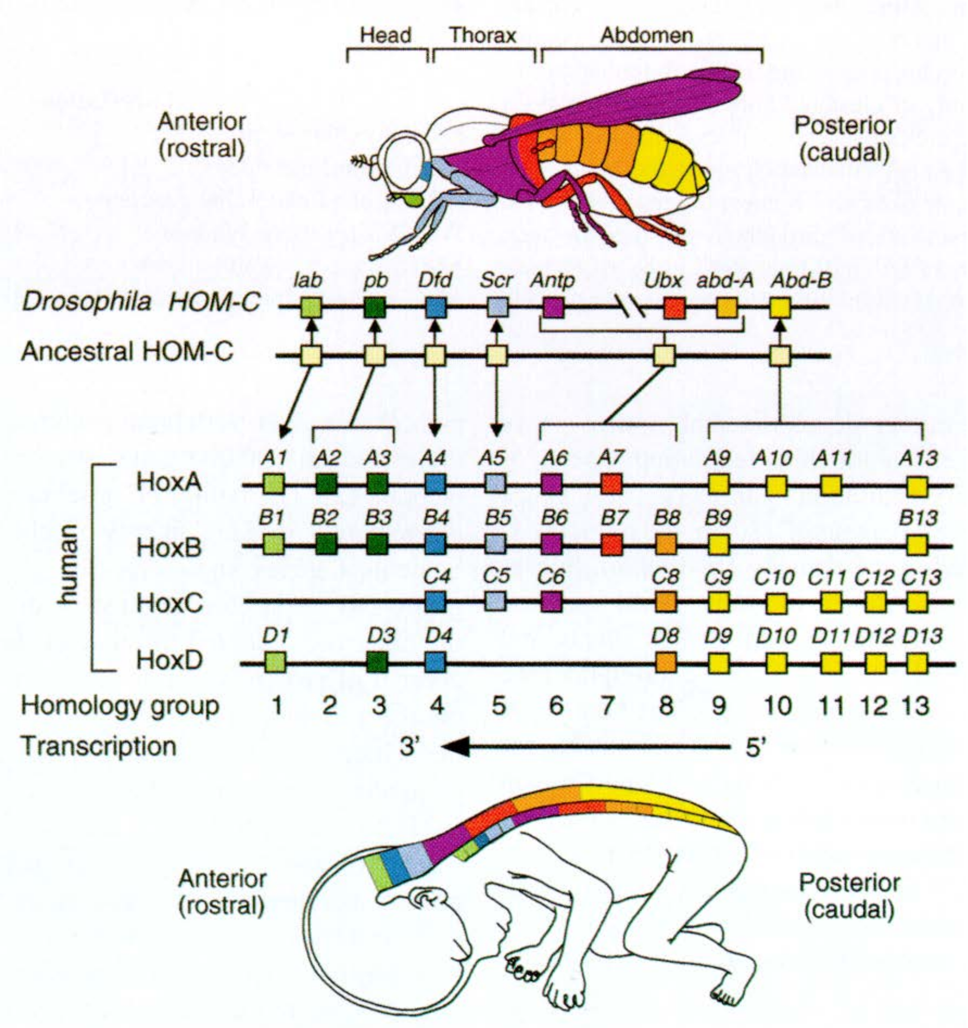
Bind to TAAT motifs in promoter or enhancer regions of target gene.

Regulate embryonic development, organogenesis, and cell differentiation.

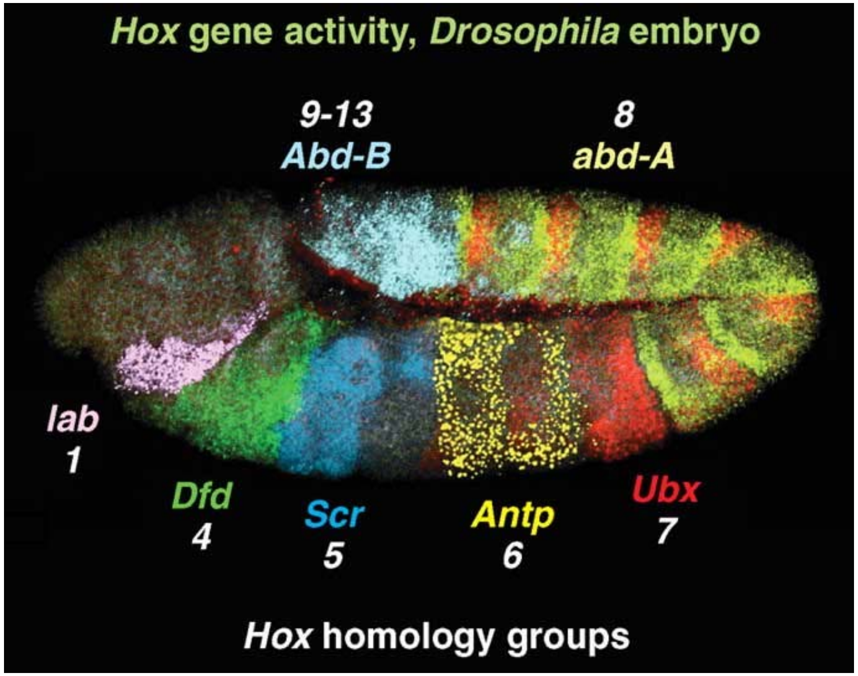
Play crucial roles in body patterning and segmental identity (e.g., Hox genes).

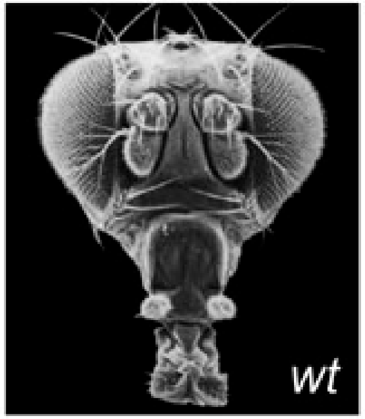
Control stem cell maintenance and lineage specification.

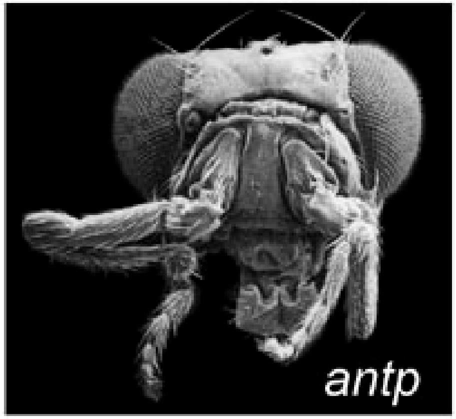
Involved in neurodevelopment, limb formation, and hematopoiesis.

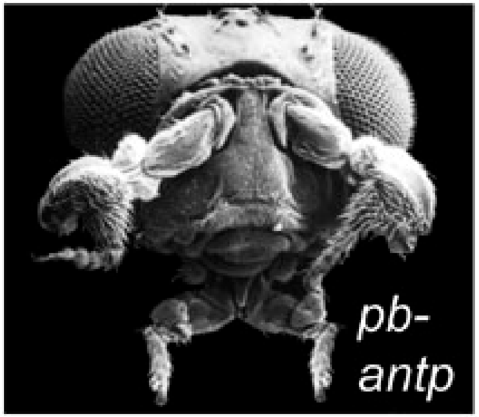


Homeodomain



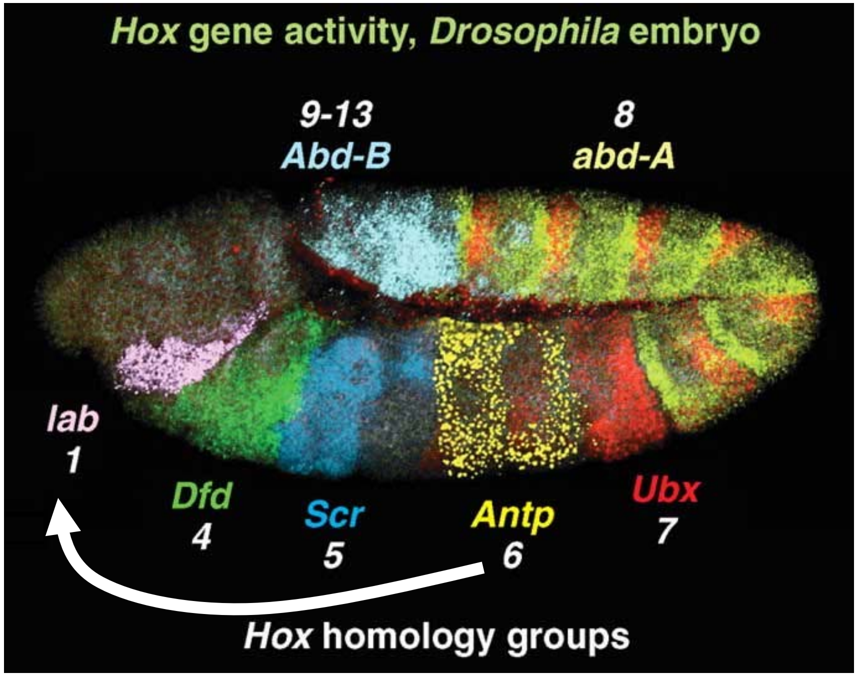


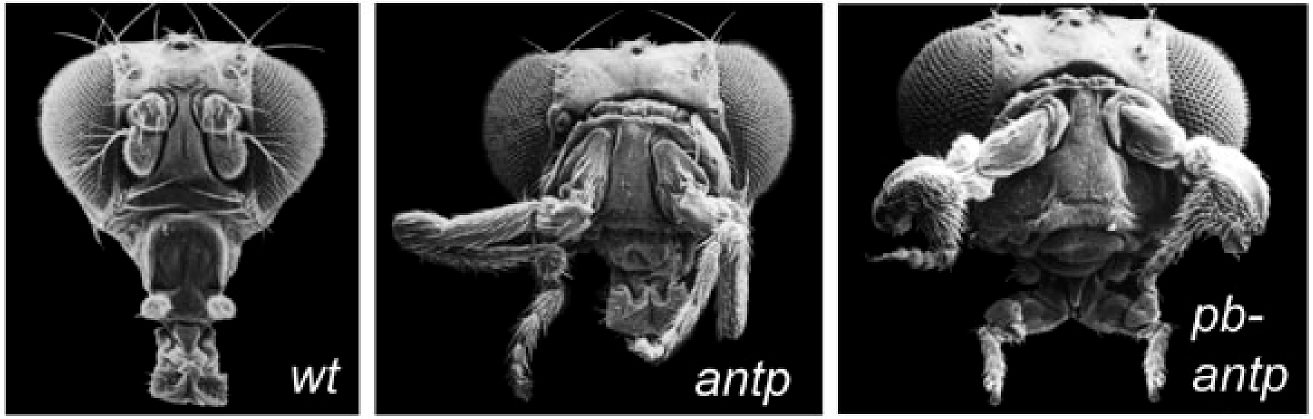




Ultrabithorax

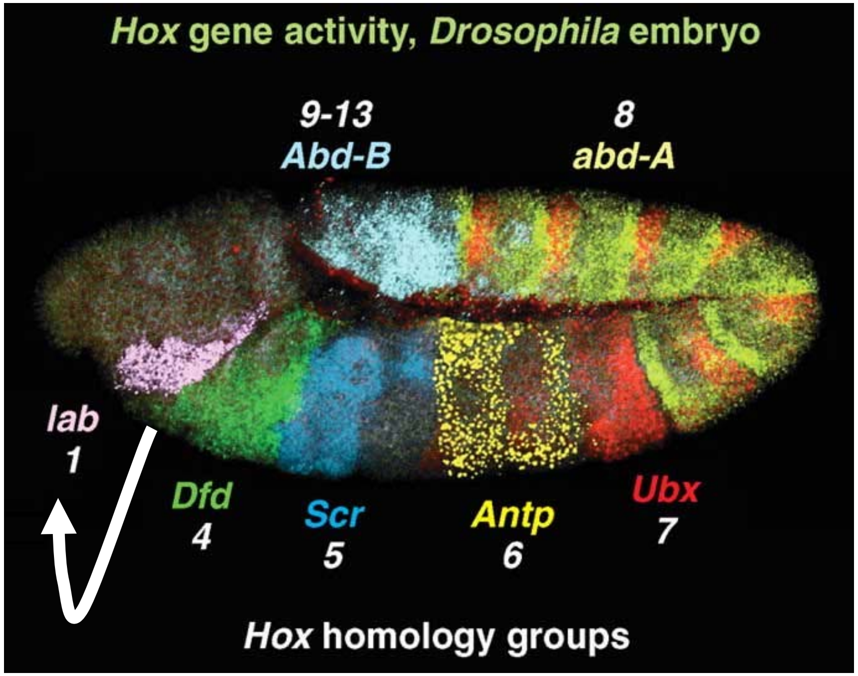
Homeodomain

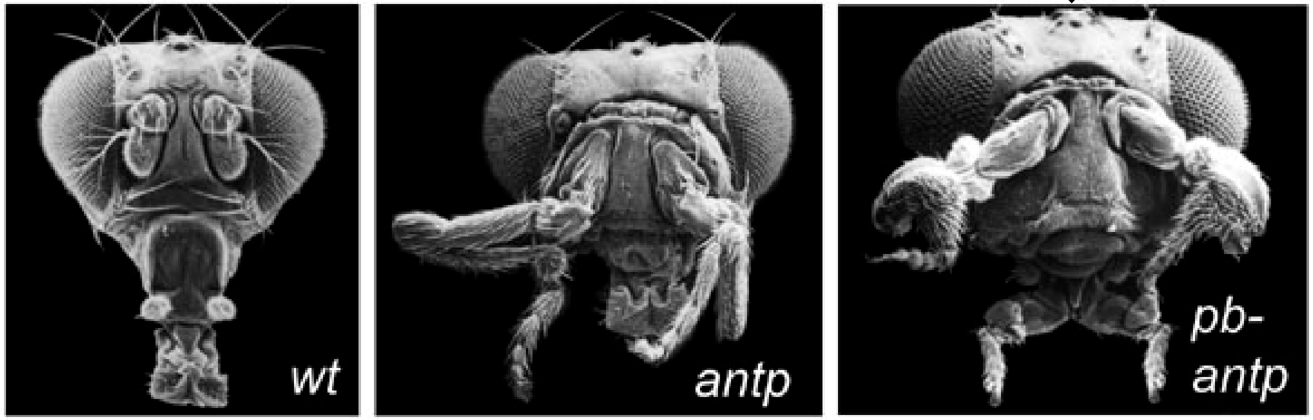




Ultrabithorax

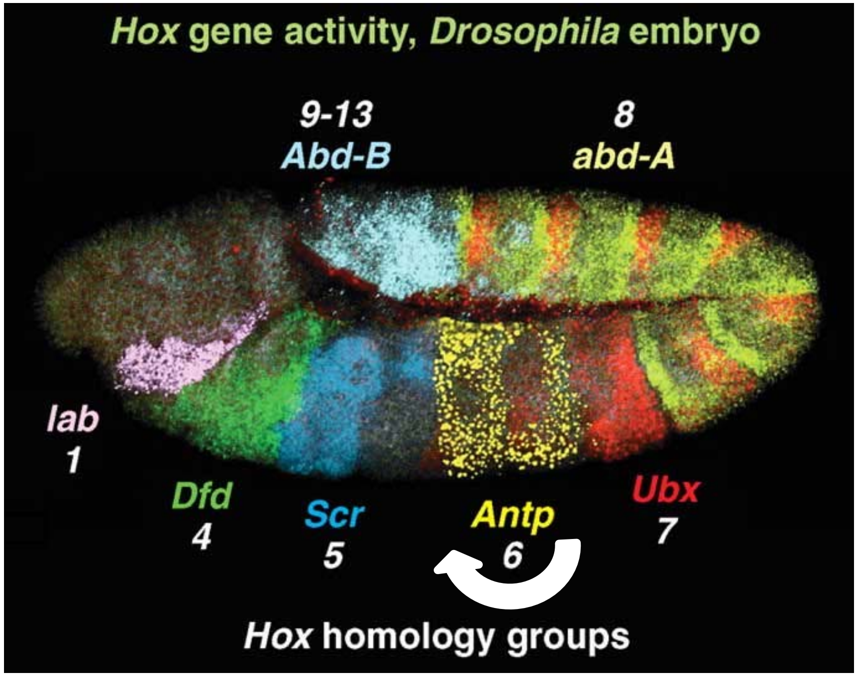
Homeodomain



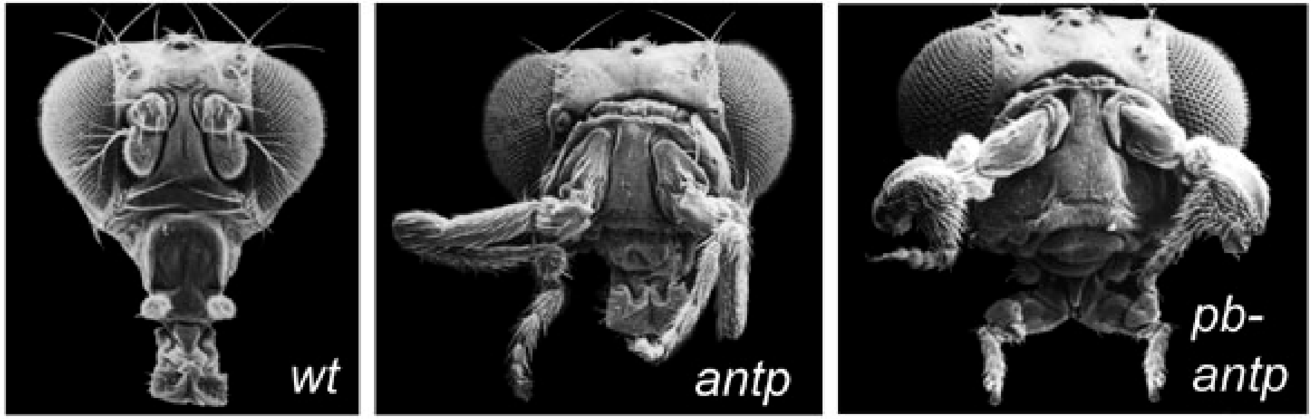


Ultrabithorax

Homeodomain



Ultrabithorax



Homeodomain

HOX Genes

Define anterior-posterior body axis (HOXA, HOXB, HOXC, HOXD clusters).

PAX (Paired Box)

Regulate organogenesis and neurogenesis (PAX6 in eye development).

SOX (SRY-related HMG box)

Involved in sex determination (SRY) and stem cell regulation (SOX2 in pluripotency).

NKX Family

Important in heart and lung development (NKX2-5 in cardiogenesis).

POU Domain Factors

Regulate neural and immune system development (POU5F1/OCT4 in stem cells).

LIM-Homeodomain TFs

Control neuronal and muscle development (LHX1, LHX2).

Homeodomain TF mutations and human disease

HOXA13 Mutation – Hand-Foot-Genital Sx (HFGS)
HOXD13 Mutation – Synpolydactyly (SPD)
HOXA11 Mutation – Radial Club Hand (RCH)
HOXA11 & HOXD11 Mutations – Müllerian and Renal Defects
HOXD10 Mutation – Congenital Vertical Talus (CVT)
HOXA10 Mutation – Uterine and Skeletal Defects
HOXA9 overexpression in acute myeloid leukemia (AML).
HOXB7 in Breast and Lung Cancer
HOXC6, HOXC8, and HOXD10 in Prostate and Colorectal Cancer
PAX6 Mutation – Aniridia
PAX3 Mutation – Waardenburg Sx Type 1 & 3
PAX2 Mutation – Renal-Coloboma Sx
SOX2 Mutation – Anophthalmia-Esophageal-Genital Sx
SOX9 Mutation – Campomelic Dysplasia
SRY Mutation – 46,XY Complete Gonadal Dysgenesis (Swyer Sx)
NKX2-5 Mutation – Congenital Heart Defects (CHDs)
NKX2-1 Mutation – Brain-Lung-Thyroid Syndrome
LHX3 and LHX4 Mutations – Combined Pituitary Hormone Deficiency (CPHD)
OTX2 Mutation – Microphthalmia and Pituitary Dysfunction
PAX3-FOXO1 Fusion – Rhabdomyosarcoma (RMS)
PITX1 Silencing – Colorectal Cancer
Etc.

Zinc fingers

Contain one or more finger-like domains having cysteine (Cys) and histidine (His) residues that bind Zn^{2+} .

C2H2 Zinc Finger (Classical Type)

Have two Cys

Found in KLF, SP1, WT1, and EGR1 transcription factors

C4 Zinc Finger (Nuclear Receptor Type)

Four Cys

Found in estrogen receptor, glucocorticoid receptor, etc.

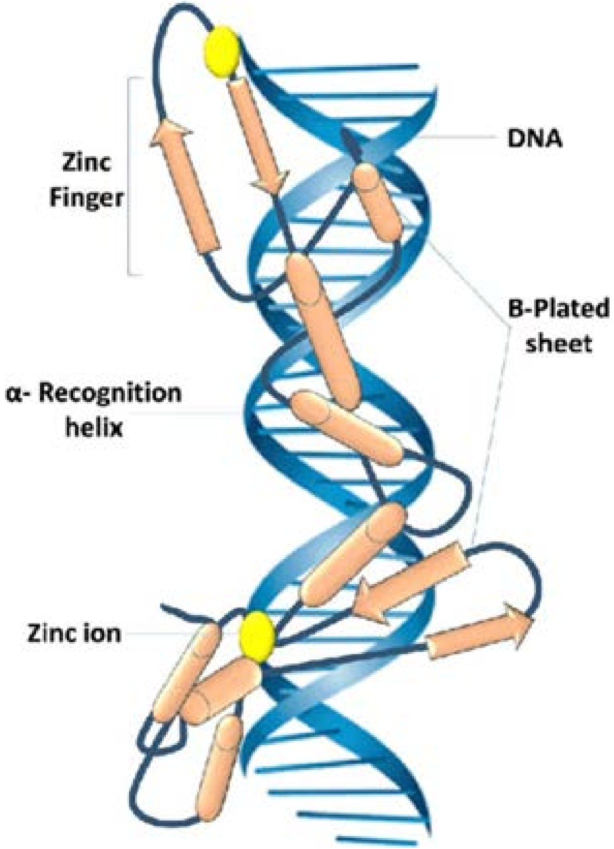
C6 Zinc Finger (Yeast Gal4-Type)

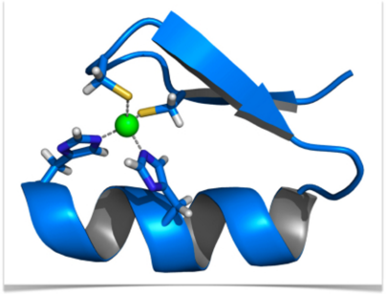
Fungal transcription factors that regulate metabolism

Ring Finger & PHD Finger Domains

Protein-protein interactions and chromatin remodeling

Found in BRCA1 and RING1.





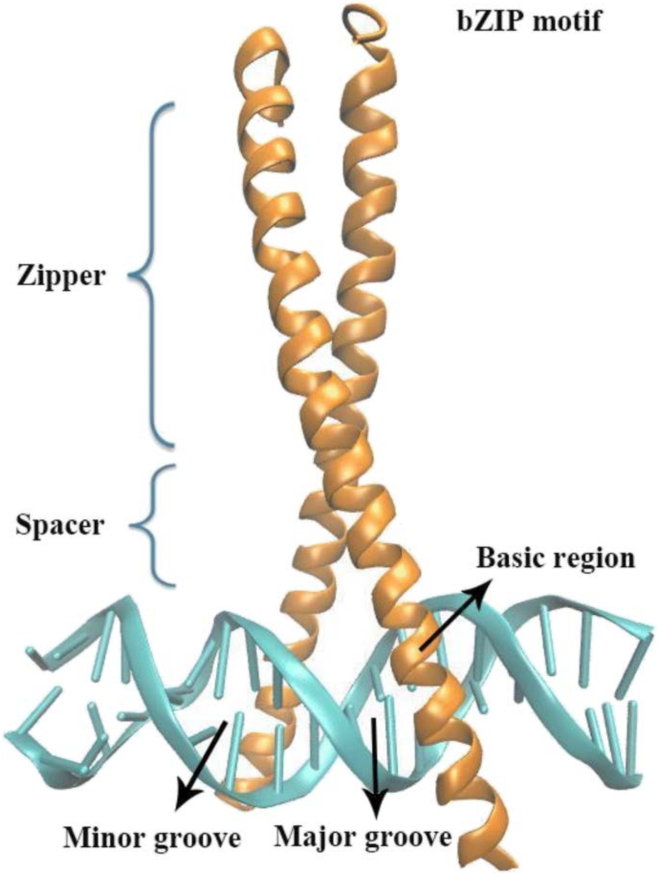
Leucine zipper (bZIP) TFs

Consist of Leu residues repeats every 7 amino acids in an α-helix, facilitating dimerization.

Regulate gene expression in cell growth, differentiation, apoptosis, & stress responses.

Mediate cellular signaling pathways such as MAPK, Wnt, and cAMP signaling.

Act as homodimers or heterodimers.



Major families of leucine zipper TFs

AP-1 (Activator Protein-1) Family

- Includes c-FOS, FOSB, c-JUN, JUNB, JUND, ATF, & MAF proteins.
- Regulates inflammation, stress response, and oncogenesis.
- c-JUN and c-FOS heterodimers control proliferation.

CREB/ATF (cAMP Response Element-Binding Protein) Family

- Includes CREB, ATF1, ATF2.
- Bind to cAMP response elements (CRE sites, 5'-TGACGTCA-3').
- Regulate memory formation, stress response, and metabolism.

C/EBP (CCAAT-Enhancer Binding Proteins) Family

- Includes C/EBPα, C/EBPβ, C/EBPγ, C/EBPδ.
- Control adipogenesis, immune response, and liver function.
- C/EBPα mutation is linked to acute myeloid leukemia (AML).

MAF (Musculoaponeurotic Fibrosarcoma Oncogene) Family

- Includes MAFA, MAFB, c-MAF, NRL.
- Involved in lens development, immune response, and β-cell function in diabetes.

XBP1 (X-Box Binding Protein 1)

- Regulate unfolded protein response (UPR) during endoplasmic reticulum stress.
- Implicated in cancer survival, inflammation, and neurodegenerative diseases.

Helix-Loop-Helix (bHLH) TFs

Basic Helix-Loop-Helix (bHLH) regulate cell differentiation, neurogenesis, myogenesis, hematopoiesis, and metabolism.

- Two α-helices connected by a loop region, enabling dimerization.
- A basic region adjacent to the HLH domain, responsible for DNA binding.
- Bind to E-box sequences (5'-CANNTG-3') in promoters or enhancers.

Class I (E-proteins E2A, TCF3, E47, E12, HEB, and E2-2).
Ubiquitously expressed and required for heterodimerization with other bHLH factors.

Class II (Tissue-Specific MYOD, MYF5, MASH1, HAND1, HAND2, ATOH1).
Regulate cell fate determination in muscle, neurons, and cardiac tissues.
Require Class I partners for activation.

Class III (bHLH-PAS Factors HIF-1α, ARNT, CLOCK, BMAL1).
Regulate hypoxia response, circadian rhythm, and environmental adaptation.

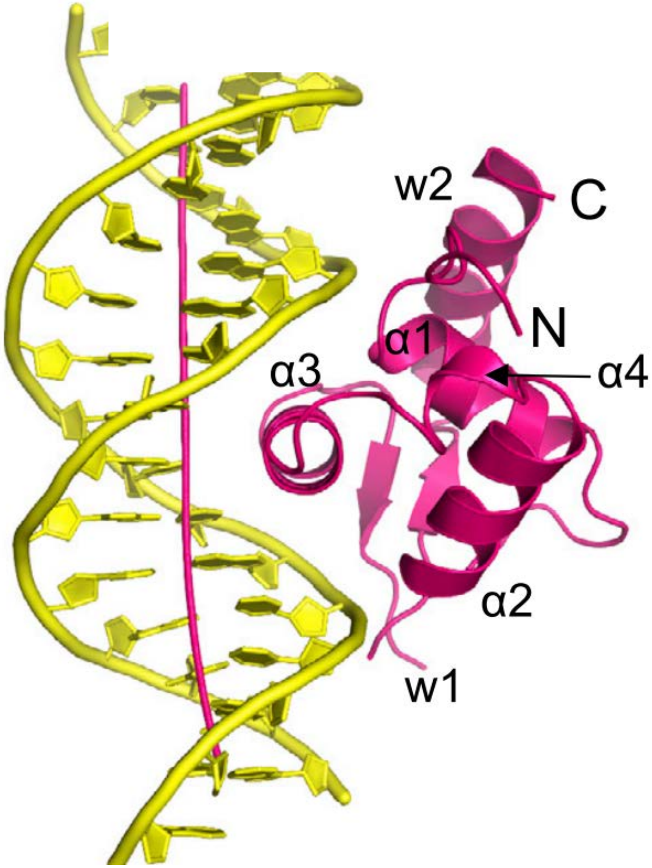
Class IV (Id Proteins)
Inhibitors of Differentiation.

Forkhead TFs (FOXO, FOXP)

Large family of DNA-binding proteins that regulate development, metabolism, longevity, immune function, and cancer progression.

Highly conserved Forkhead DNA-binding domain, a winged-helix structure that recognizes specific 5'-TTGTTTAC-3’ motifs.

Classified into 19 subfamilies (FOXA-FOXS).



Forkhead Transcription Factor FoxN3. Rogers JM, et al. Mol Cell. 2019 Apr 18;74(2):245-253

Forkhead transcription factor classification

Classified into 19 subfamilies (FOXA-FOXS), the most important being:

FOXO (Forkhead Box O) FOXO1, FOXO3, FOXO4 & FOXO6
Regulates stress response, longevity, apoptosis, autophagy, and metabolism.
FOXO3 mutations are linked to human longevity.

FOXP (Forkhead Box P) FOXP1, FOXP2, FOXP3 & FOXP4
Regulates neurodevelopment, immunity.
FOXP2 "language gene" involved in speech and neurodevelopment, FOXP3 master Treg regulator.

FOXA (Forkhead Box A) FOXA1, FOXA2, FOXA3
Regulates liver, pancreas, lung & endodermal organogenesis, opens chromatin for transcription.

FOXC (Forkhead Box C) FOXC1, FOXC2
Cardiovascular & neural crest development
FOXC1 mutations cause Axenfeld-Rieger syndrome (glaucoma, craniofacial defects).
FOXC2 regulates lymphatic vessel development.

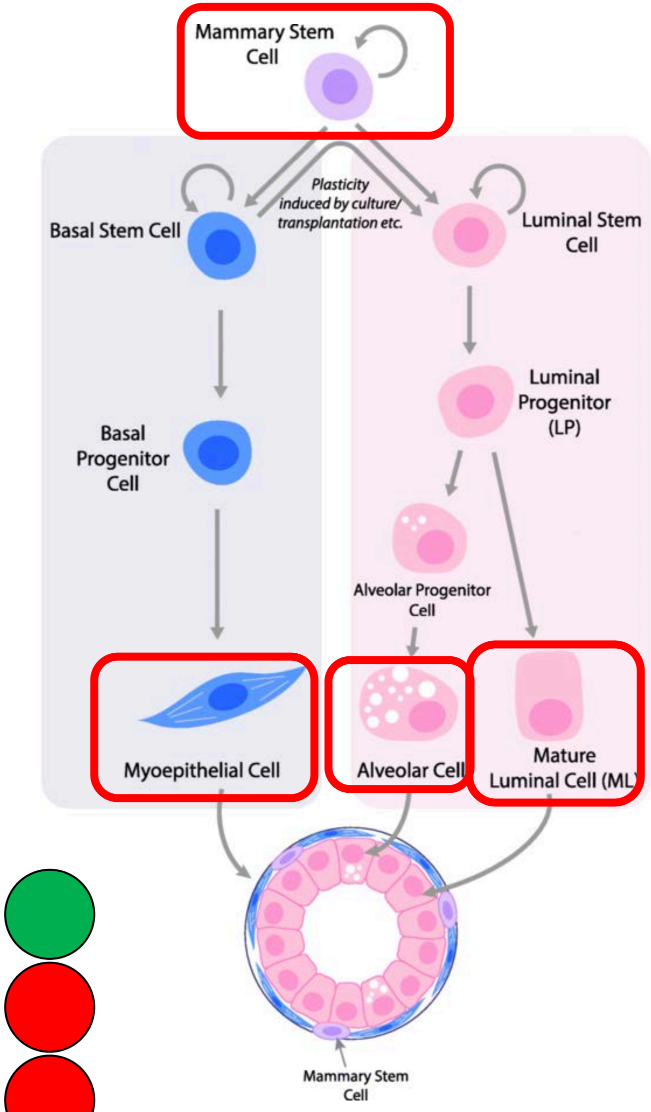
FOXD (Forkhead Box D) FOXD1, FOXD2, FOXD3
Neurogenesis & differentiation, FOXD3 is crucial for neural crest differentiation.

FOXE (Forkhead Box E) FOXE1, FOXE3
Metabolism & Endocrine Regulation, FOXE1 mutations cause hypothyroidism and cleft palate.

Lineage commitment within the mammary gland

Model for the regulation of lineage commitment within the mammary gland epithelium.

Mammary stem cells (MaSC) can give rise to both myoepithelial and luminal/alveolar lineages during embryonic development and when subjected to dissociation and transplantation or culture.



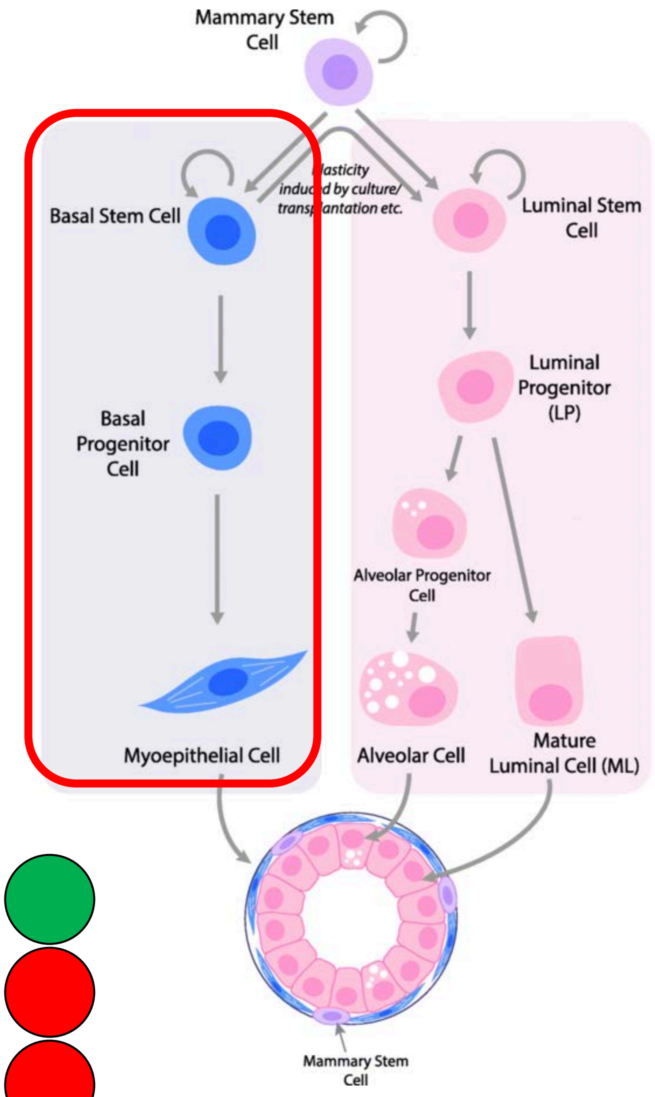
Basal identity genes 🟢

LP identity genes 🔴

ML identity genes 🔴

Lineage commitment within the mammary gland

Cells within the basal compartment (stem, progenitor, and differentiated myoepithelial) have an epigenetic landscape that allows basal identity genes to be turned on and luminal identity genes to be turned off.



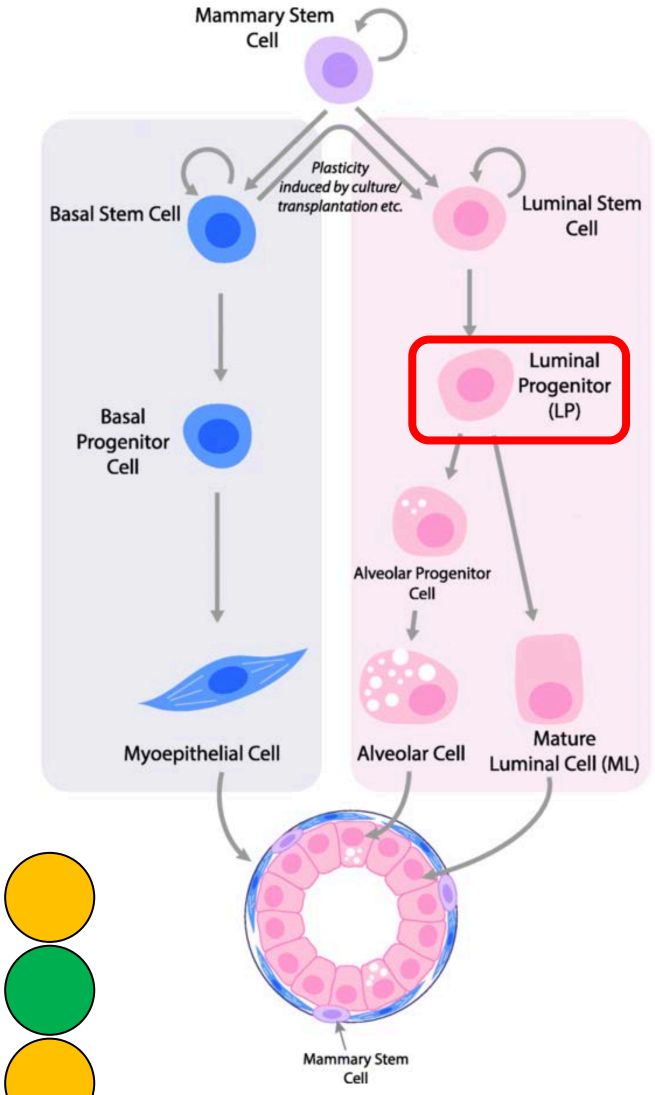
Basal identity genes 🟢

LP identity genes 🔴

ML identity genes 🔴

Lineage commitment within the mammary gland

Luminal progenitor cells have intermediate epigenetic features between basal and mature luminal cells.



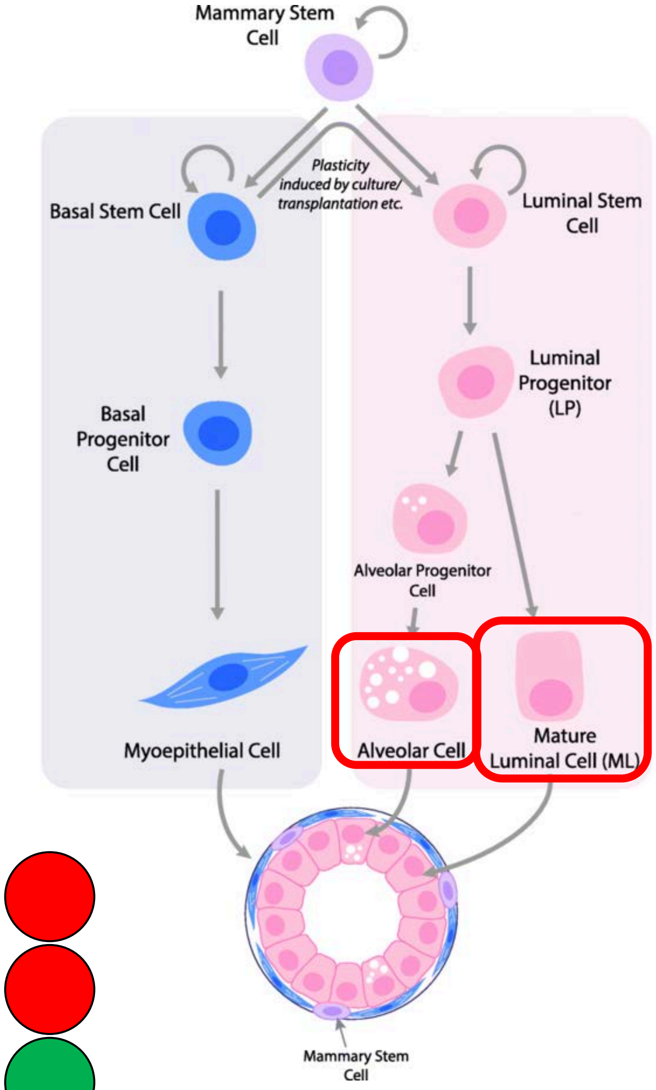
Basal identity genes 🟡

LP identity genes 🟢

ML identity genes 🟡

Lineage commitment within the mammary gland

Mature luminal cells have repressive epigenetic features in basal and luminal progenitor identity genes and active epigenetic features in mature luminal identity genes.



Basal identity genes 🔴

LP identity genes 🔴

ML identity genes 🟢

Laboratorio de Genómica Viral y Humana

Instalaciones de Alta Contención Biológica Nivel de Bioseguridad 3 (BSL-3) CDC-certificadas

Facultad de Medicina UASLP
San Luis Potosí, México

www.genomica.uaslp.mx GenomicaUASLP YouTube GenomicaUASLP

Content copyright and license

The Viral and Human Genomics Laboratory is committed to promoting the human rights of free access to knowledge and to receiving the benefits of scientific progress and its applications by providing universal access to all the resources and publications it produces. This is in agreement with article 15 of the United Nations International Covenant on Economic, Social and Cultural Rights published on April 30, 2020.

All information included in this document is in the public domain, was compiled by the licensor and is distributed under a Creative Commons Attribution 4.0 International (CC BY 4.0 DEED) license which grants the licensee (you) the right to copy, remix, transform, develop and redistribute the material in any medium or format for any purpose, including commercial purposes provided that

1) Corresponding credit is given to the licensor as “CA García-Sepúlveda, Laboratory of Viral and Human Genomics UASLP”,

2) Any changes to the original document are indicated and,

3) In no way suggest that the licensor endorses the derivative work.

All rights reserved © 2024 CA García-Sepúlveda, Laboratory of Viral and Human Genomics UASLP

(Last updated August 23, © 2024.)