

Circadian clocks in mammals and plants

Most organisms (animals, plants, fungi and cyanobacteria) enhance their fitness by coordinating their development with daily environmental changes through **molecular timekeepers (circadian clocks)**.

Mammals display circadian rhythms in behavioral and physiological processes, such as

- sleep
- feeding
- blood pressure and
- metabolism

Roles in **plants** e.g.:

- opening of flowers in the morning and their closure at night

Circadian rhythms are guided by **external light–dark signals** that are integrated through intrinsic central and peripheral molecular clocks

McClung Plant Cell 18, 792 (2006)

Circadian rhythms

(1) Circadian rhythms are the subset of biological rhythms with period of 24 h.
The term circadian combines the Latin words “circa” (about) and “dies” (day).

(2) Circadian rhythms are **endogenously generated** and **self-sustaining**.

They persist under constant environmental conditions, typically constant light (or dark) and constant temperature.

Under these controlled conditions, the free-running period of **24 h** is observed.

(3) For all circadian rhythms, the **period** remains relatively **constant** over a range of ambient temperatures.

This is thought to be one property of a general mechanism that buffers the clock against changes in cellular metabolism.

**In contrast, chemical reactions are usually faster at higher temperatures
(-> Arrhenius equation).**

$$k = A e^{\frac{-E_a}{RT}}$$

McClung Plant Cell 18, 792 (2006)

Essential elements of biological clocks

Our biological clocks contain 3 essential elements:

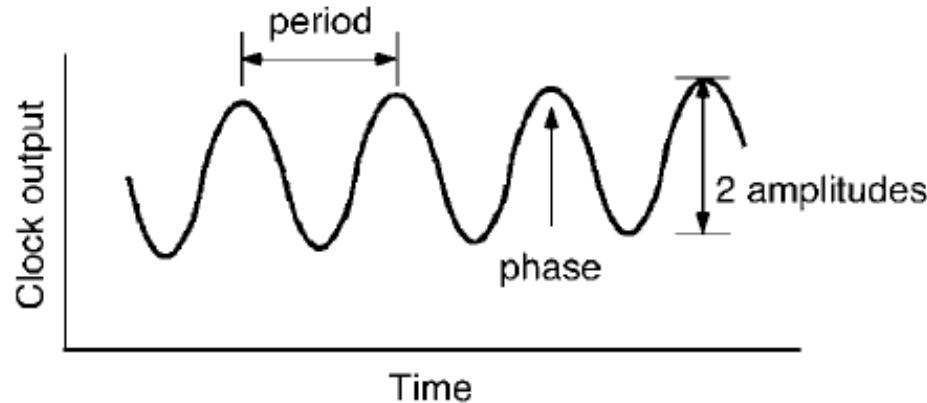
- (1) a **central oscillator** that keeps time;
- (2) the ability to **sense time cues** in the environment and to **reset the clock** as the seasons change; and
- (3) a series of outputs tied to distinct phases of the oscillator that regulate activity and physiology.

Gallego et al. Nat.Rev.Mol.Cell.Biol. 8, 140 (2007)

Parameters of Circadian clocks

Period : time to complete one cycle.

Amplitude of the rhythm :
one-half the peak-to-trough distance.



Phase : time of day for any given event.

E.g. if the peak in a rhythm occurred at dawn,
the phase of the peak would be defined as 0 h.

Phase is often defined in **zeitgeber time (ZT)**.

Zeitgeber is German for „time giver“, and any stimulus
that imparts time information to the clock is a zeitgeber.

The onset of light is a powerful zeitgeber, and dawn is defined as ZT0.

McClung Plant Cell 18, 792 (2006)

Suprachiasmatic nucleus (SCN)

In mammals, the central clock resides in the suprachiasmatic nucleus (SCN), a small region of the brain that contains ca. 20,000 neurons.

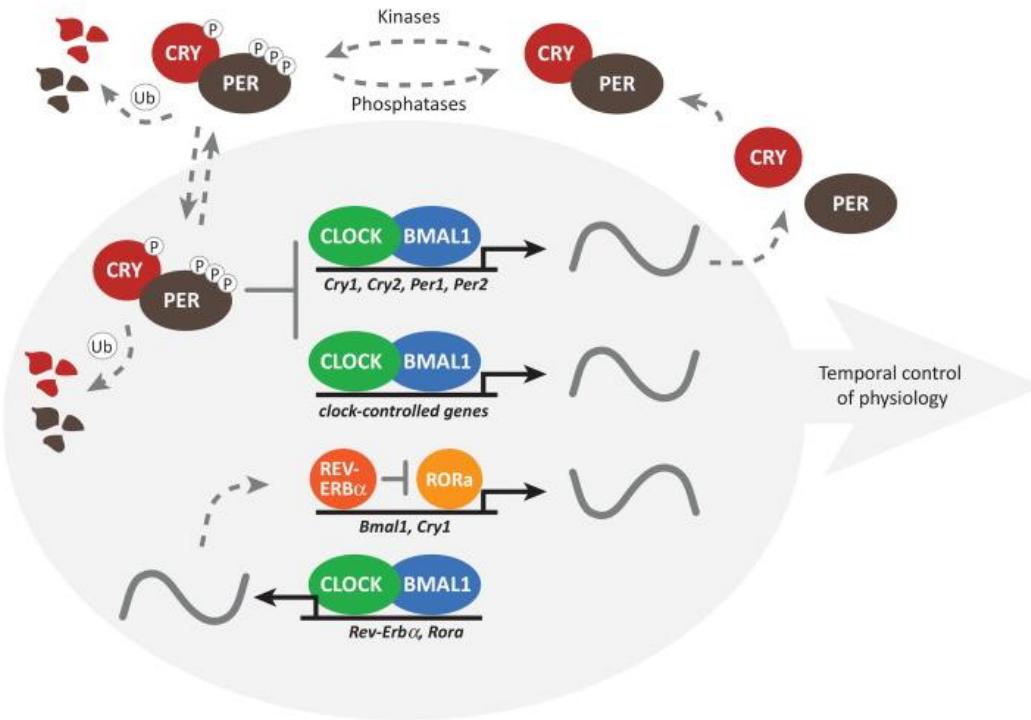
The SCN produces a **rhythmic output** that consists of a multitude of neural and hormonal signals that influence sleep and activity.

Most importantly, the SCN signals **set the peripheral clocks** present throughout the body.

The SCN clock is reset by **external light**, which is **sensed** by the ganglion cells of the **retina**.

Gallego et al. Nat.Rev.Mol.Cell.Biol. 8, 140 (2007)

The molecular circadian clock in mammals



BMAL1, brain and muscle ARNT-like 1
CLOCK, circadian locomotor output cycles kaput
CKI: casein kinases I CKI α , CKI δ , and CKI ϵ ;
CRY: cryptochromes
PER: period
PP: protein phosphatases PP1, PP5.

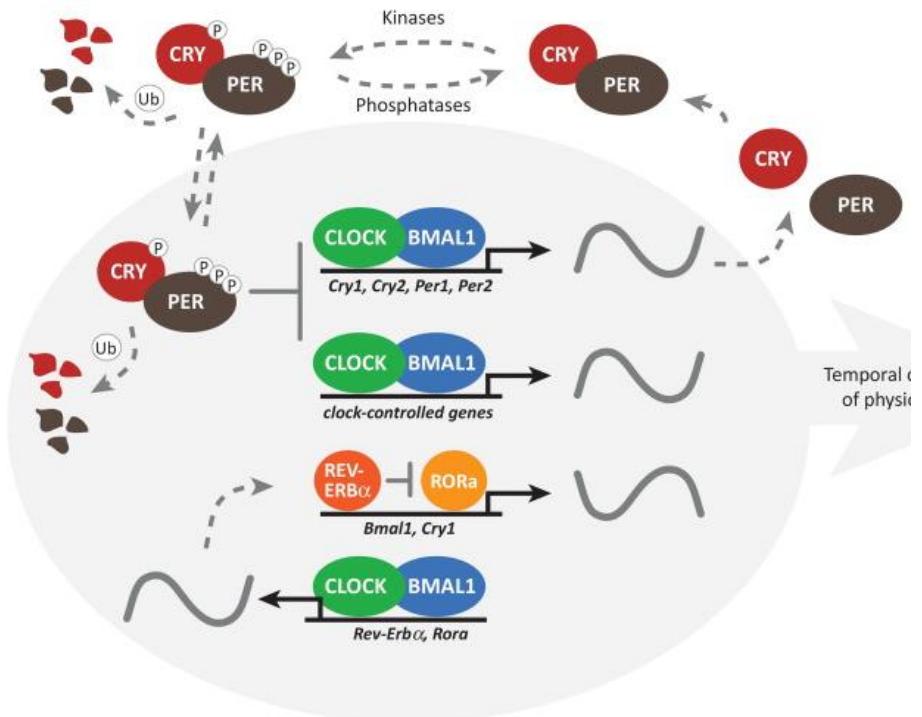
The cell-autonomous molecular clock in mammals is generated by 2 interlocking transcription/translation feedback loops (TTFL) that function together to produce robust 24 h rhythms of gene expression.

The core TTFL is driven by 4 integral clock proteins:

2 activators (CLOCK and BMAL1) and 2 repressors (PER and CRY), as well as by kinases and phosphatases that regulate the phosphorylation (P) and thereby localization and stability of these integral clock proteins.

Partch et al. Trends Cell Biol 24, 90 (2014)

The molecular circadian clock in mammals



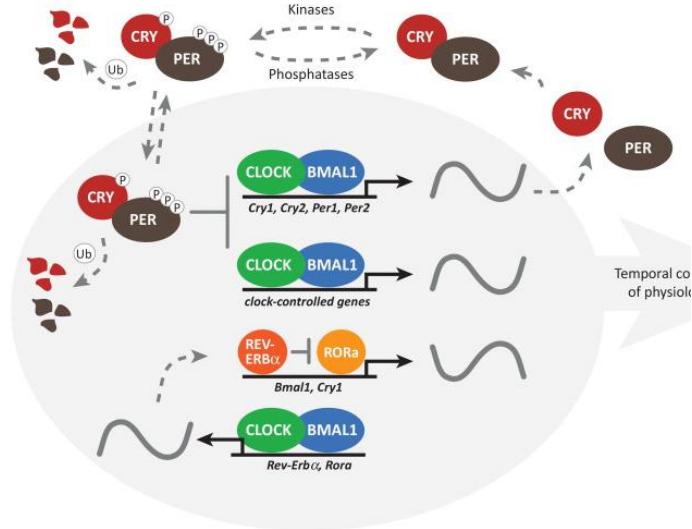
CLOCK and BMAL1 are subunits of the heterodimeric basic helix-loop-helix-PAS (PER-ARNT-SIM) transcription factor CLOCK:BMAL1, which activates transcription of the repressor *Per* and *Cry* genes, as well as other clock-controlled output genes.

PER and CRY proteins heterodimerize in the cytoplasm and translocate to the nucleus to interact with CLOCK:BMAL1, inhibiting further transcriptional activation.

As PER and CRY proteins are degraded through ubiquitin (Ub)-dependent pathways, repression on CLOCK:BMAL1 is relieved and the cycle begins again with ~24 h periodicity.

Partch et al. Trends Cell Biol 24, 90 (2014)

The molecular circadian clock in mammals



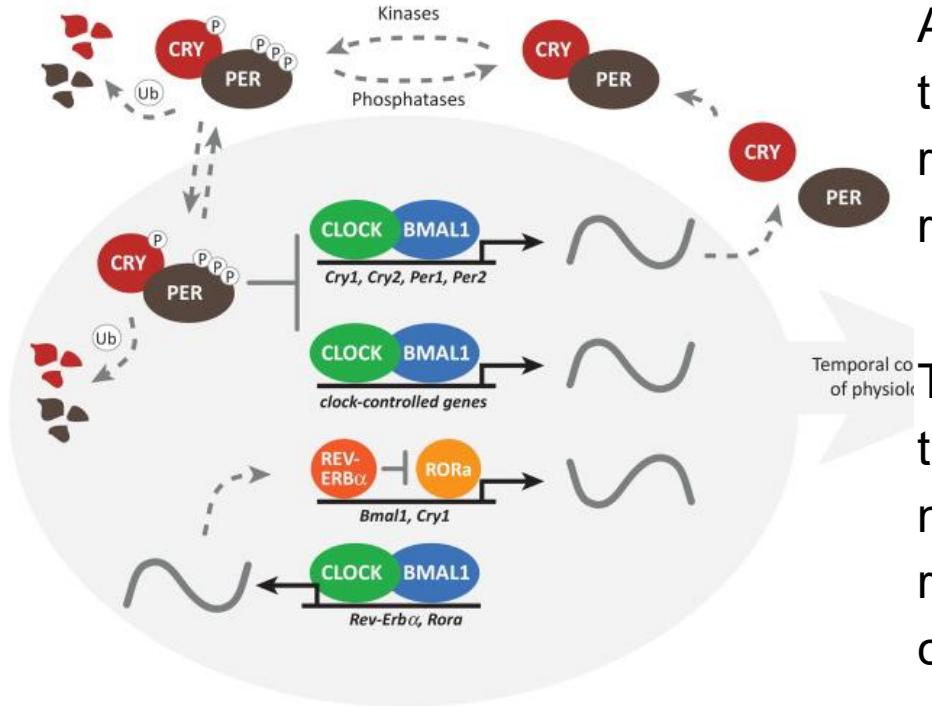
The casein kinases CKI δ and CKI ϵ play an important role in determining the intrinsic period of the clock by controlling the rate at which the PER:CRY complexes are either degraded or enter the nucleus, and their activity is either counteracted or regulated by the phosphatases PP1 and PP5, respectively.

Notably, **familial mutations** resulting in the loss of a single phospho-acceptor site on PER2 (S662G) or a loss-of-function mutation in CKI δ (T44A) shorten the intrinsic period of the clock in mice and give rise to sleep phase disorders in humans.

A key role for the casein kinases in establishing period length has also been demonstrated pharmacologically via modulation of the kinases with **small-molecule inhibitors**, which dramatically lengthen the period by modulating PER localization and stability.

Partch et al. Trends Cell Biol 24, 90 (2014)

The molecular circadian clock in mammals



A second TTFL is generated through transcriptional activation by the retinoid-related orphan receptors (ROR α , b, c) and repression by REV-ERB α /REV-ERB β .

This TTFL drives rhythmic changes in *Bmal1* transcription and introduces a **delay** in *Cry1* mRNA expression that offsets it from genes regulated strictly by CLOCK:BMAL1 and is crucial for proper circadian timing

The presence of cooperative, interlocking feedback loops provides **robustness** against noise and environmental perturbations to help maintain accurate circadian timing, and also helps to generate **phase delays** in circadian transcriptional output that optimally time gene expression for local physiology.

Partch et al. Trends Cell Biol 24, 90 (2014)

Detect unknown control mechanisms: Probe gene expression by microarrays

Already in the year 2000, Harmer *et al.* used oligonucleotide-based arrays to determine steady-state mRNA levels in *Arabidopsis* at 4-hour intervals during the subjective day and night.

→ They identified temporal patterns of gene expression in *Arabidopsis* plants under constant light conditions using GeneChip arrays representing about 8200 different genes.

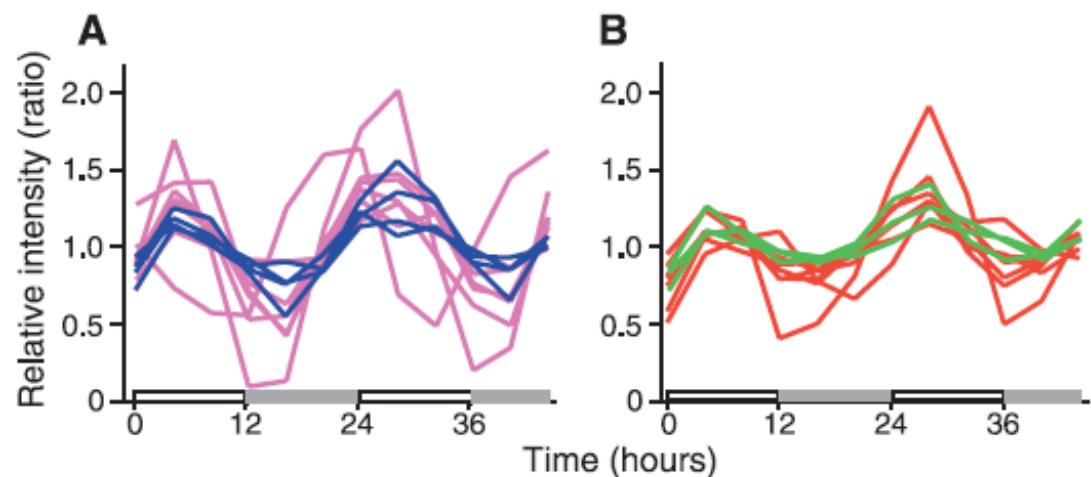
All genes were scored whether their expression is more correlated with a **cosine** test wave with a period between 20 and 28 hours than 95% of randomly permuted data sets.

→ consider those genes as circadian-regulated.

→ 453 genes (6% of the genes on the chip) were classified as **cycling**.

Harmer et al. Science 290, 2110 (2000)

Photosynthesis genes peak near the middle of the day

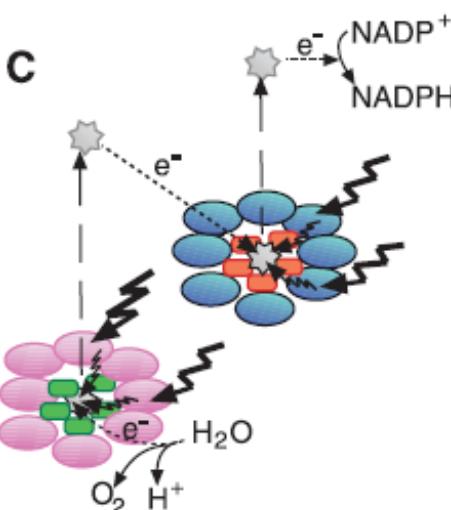


Results after normalization of peak maximum.

(A) *LHCA* genes are in blue; *LHCB* genes are in pink.

(B) Photosystem I genes are in red;. Photosystem II genes are in green;.

(C) Model for function of photosynthesis gene products in photosystems II (left) and I (right). Colors of proteins match colors of corresponding gene traces.



These results validate that the microarray captures the expected behavior. Light-harvesting complexes (A) and photosystems (B) show the expected oscillating expression pattern. They have peaks around lunch time. Time 0 corresponds to the time of sunrise.

Panel (C) is a cartoon of the photosynthesis machinery. Light-harvesting complexes (in the outer ring) collect photons (indicated by zig-zagged arrows) from the sun light and transmit the excitation energy to the green and red photosystem proteins (shown in the center of the rings). In the photosystems, excitation energy is used to for two rounds of charge separation, from the first ring to the second ring, and then onto the co-factor NADP+. As a net reaction, an electron loaded onto the co-factor.

Synchronized production of photoprotective pigments

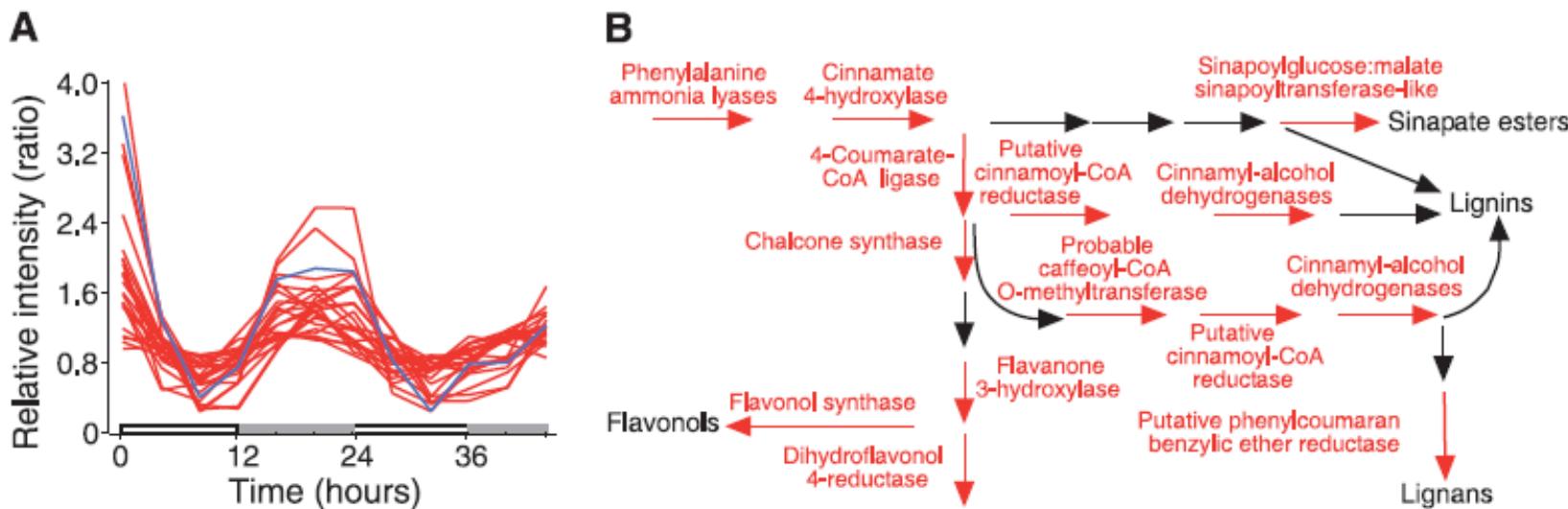


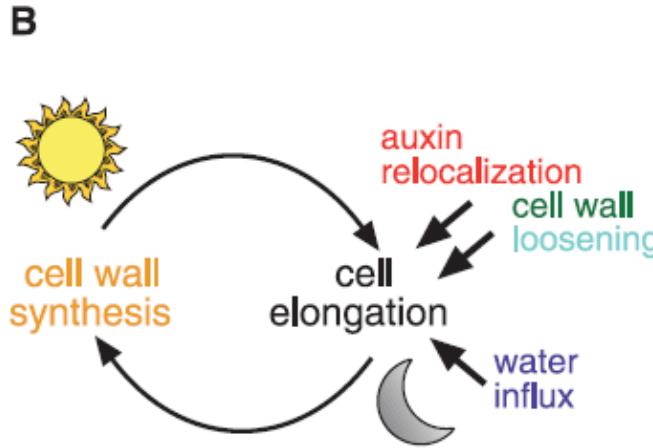
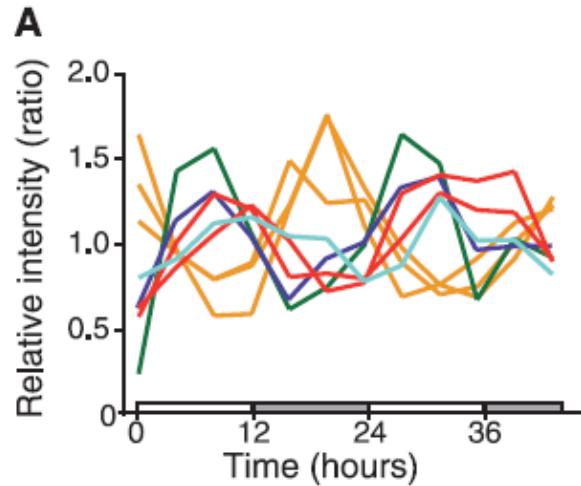
Fig. 2. Phenylpropanoid biosynthesis genes peak before subjective dawn. **(A)** The gene encoding the Myb transcription factor *PAP1* (accession number AAC83630) is in blue. The red traces represent phenylpropanoid biosynthesis genes. **(B)** Phenylpropanoid biosynthetic pathways. Genes encoding all enzymes indicated in red are clock-controlled. See Web table 2 (8) for gene names and accession numbers.

„**Phenolic sunscreen**“ is produced before sunrise.

Substances absorb light in the visible and UV range.

Harmer et al. Science 290, 2110 (2000)

Genes implicated in cell elongation are circadian-regulated



(B) Proposed mode of action of the products of these clock-controlled genes in cell wall remodeling.

The rigid plant cell wall normally prevents cell expansion, but a simultaneous loosening of cell wall components, uptake of water, and synthesis of cell wall components seems allowed.

(A) Genes encoding the auxin efflux carriers *PIN3* and *PIN7* (red), a putative expansin (green), a putative polygalacturonase (light blue), and aquaporin d-TIP (dark blue) all peak toward the end of the subjective day.

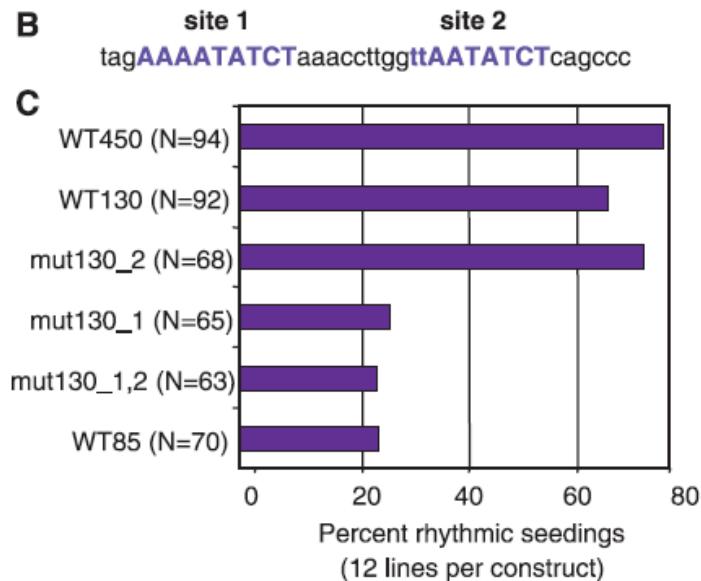
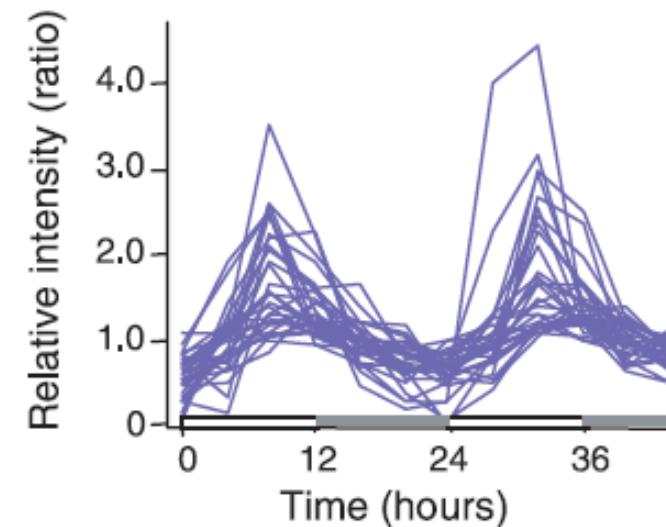
Auxins are phytohormones – they regulate cell extension.

3 enzymes implicated in cell wall synthesis (all in gold) peak toward the end of the subjective night.

Harmer et al. Science 290, 2110 (2000)

Master regulator sequence of circadian-regulated genes?

Check genomic DNA regions upstream of cycling genes for overrepresented promoter elements
→ absolutely conserved motif, AAAATATCT “**evening element**,” that occurs 46 times in the promoters of 31 cycling genes. All 31 genes demonstrated impressive coregulation. All but one peak toward the end of the subjective day.



Mutation of the conserved AAAATATCT motif, but not a closely related motif, greatly reduced the ability of a promoter to confer circadian rhythmicity on a luciferase reporter gene in plants.

Harmer et al. Science 290, 2110 (2000)

Summary

Most organisms enhance fitness by coordinating their development with daily environmental changes through molecular timekeepers known as circadian clocks.

Clocks are generated by a transcription-translation negative feedback loop with a crucial delay between stimulus and response.

This system of multiple connected loops increases the clock's robustness and provides numerous points of input and output to the clock.

Many metabolic pathways are regulated by circadian clocks in plants and animals.

Kay & Schroeder Science 318, 1730 (2007)

Effect of sleep duration on humans?

30% of civilian adults in the US sleep less than 6 hours per day ...

reasons: work, habits, studies ...

Importantly, **short sleep** duration (< 6 hours/day) has been associated with **negative health outcomes!**

Short sleep increases: overall mortality, obesity, diabetes, cardiovascular diseases ...

→ What happens on the molecular level?

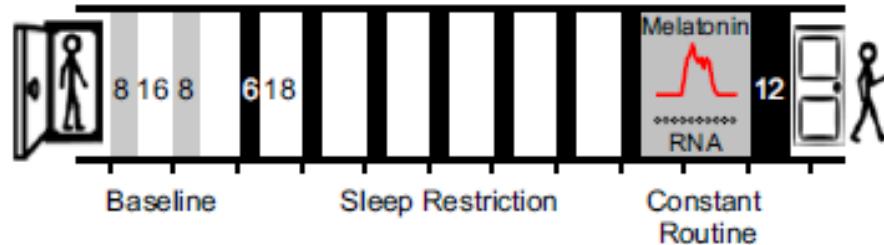
Effects of insufficient sleep on circadian rhythmicity and expression amplitude of the human blood transcriptome

PNAS (2013) 110, E1132-E1141

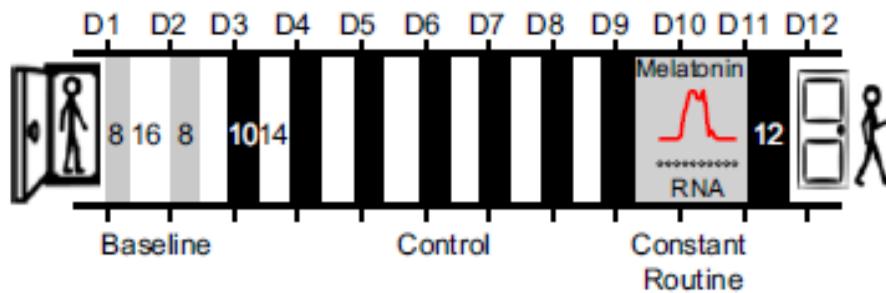
Carla S. Möller-Levet¹, Simon N. Archer¹, Giselda Bucca¹, Emma E. Laing, Ana Slak, Renata Kabiljo, June C. Y. Lo, Nayantara Santhi, Malcolm von Schantz, Colin P. Smith¹, and Derk-Jan Dijk^{1,2}

Cross-over design study

26 participants (volunteers) were first put into **sleep-restricted conditions** with only 6 hours of sleep opportunity per night (dark bars)



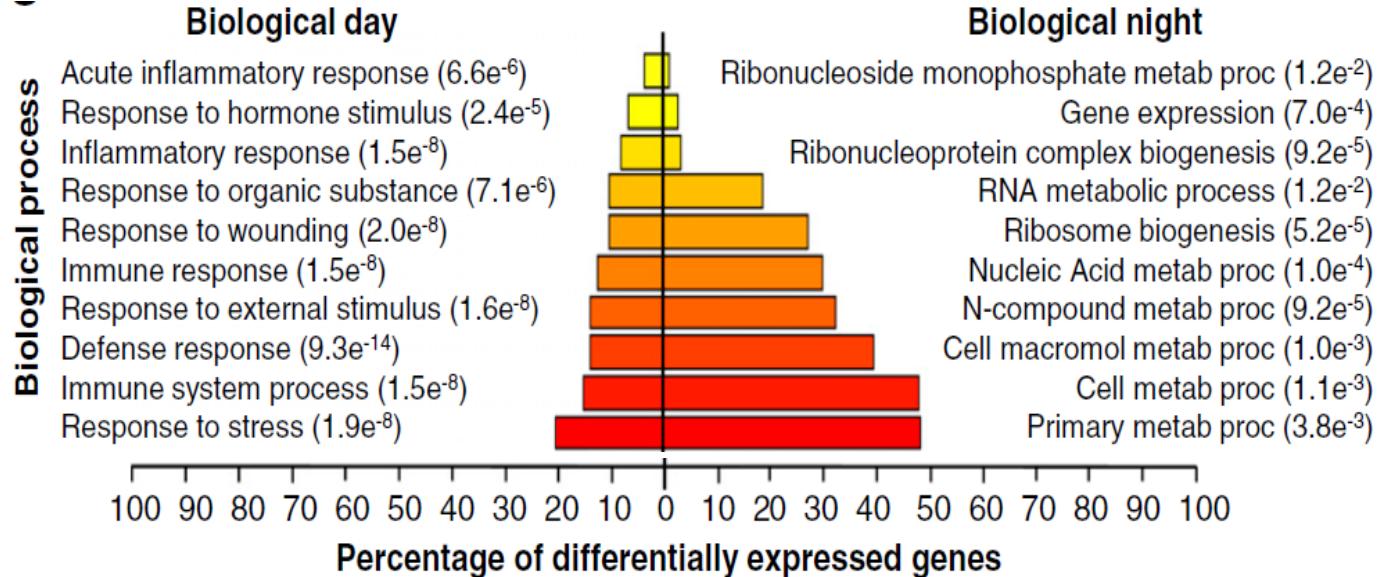
and then into conditions of **sufficient sleep** with 10 hours of sleep opportunity.
-> effects of genetic pre-disposition are minimized by using „matched samples“



D1 to D12: day 1 to day 12

PNAS 110, E1132 (2013)

Gene functions of „normal“ circadian genes



Top 10 enriched
GO BPs within the
circadian gene list
of the **control**
condition using
the human genome
as a background

Enrichment p-values are
given in brackets.

Immune, defense, stress and inflammatory responses, cytokine receptor activity, IL-1 receptor activity, NF- κ B signaling are more prominent during day time.

(Also found for rodents at night time when they are active).

Night time processes: “normal” maintenance + growth processes ...

Gene Ontology (GO)

Ontologies are structured vocabularies.

The **Gene Ontology** has 3 tracks:

- biological process (BP)
- molecular function (MF)
- cellular component (localisation).

Shown here is a part of the BP tree.

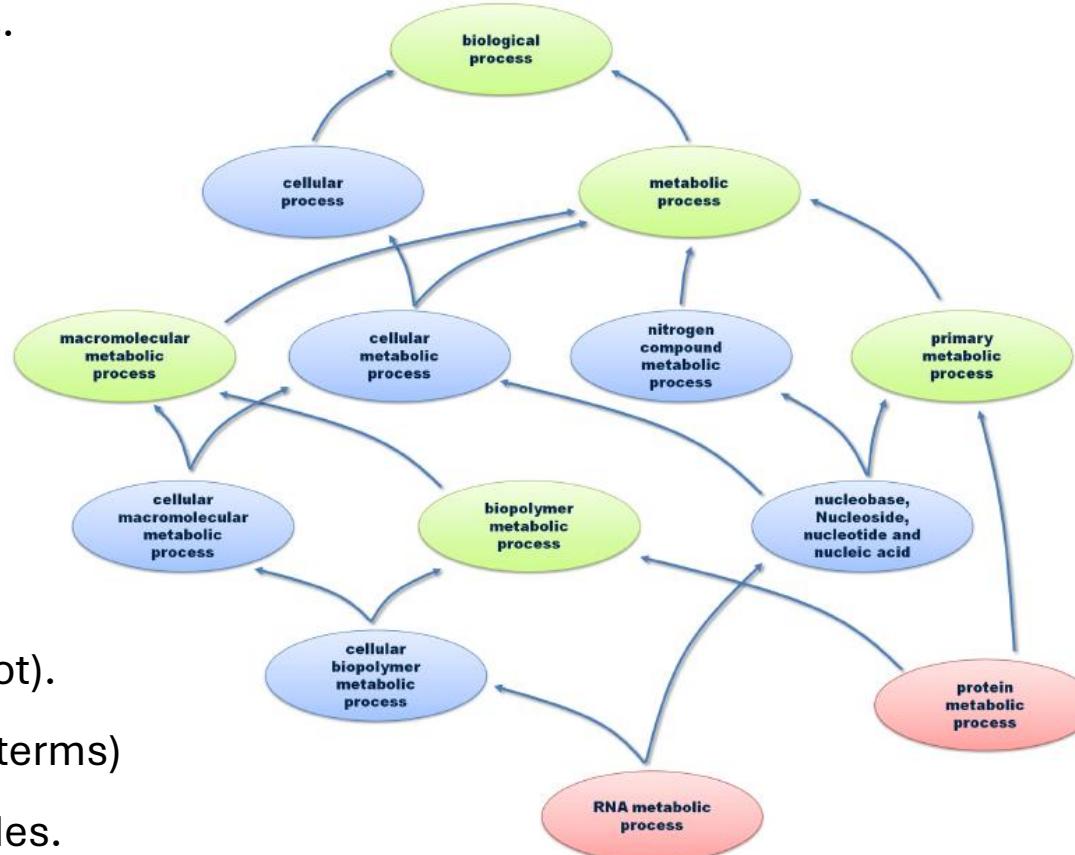
At the top: most general expression (root).

Red: leafs of the tree (very specific GO terms)

Green: common ancestors of 2 red nodes.

Blue: other nodes.

Lines: „Y is contained in X“- relationships



Dissertation Andreas Schlicker (UdS, 2010)

Over-representation analysis (WebGestalt)

Suppose that we have *n genes* in a “gene set of interest” (A) and *m genes* in the **reference gene set** (B).

Suppose further that there are *k genes* in A and *j genes* in B that belong to a particular functional category (C)
(e.g. a GO category, a KEGG pathway, a BioCarta pathway etc.).

Based on the reference gene set, the expected proportion k_{exp} would be

$$k_{\text{exp}} = (n/m) \times j$$

If *k* exceeds the above expected value,
category C is said to be **enriched**,
with a **ratio of enrichment** (*r*) given by $r = k/k_{\text{exp}}$.

Zhang, Kirov, Snoddy (2013)
Nucl Ac Res 33: W741-W748

Over-representation analysis (WebGestalt)

If B represents the population from which the genes in A are drawn,
WebGestalt uses the **hypergeometric test** to evaluate the significance
of enrichment for category C in gene set A,

$$P = \sum_{i=k}^n \frac{\binom{m-j}{n-i} \binom{j}{i}}{\binom{m}{n}}$$

Interpretation: draw $i = k$ genes for A that belong to category C from the j genes from B that belong to C.
→ The other $n - i$ genes in A do not belong to C. They are drawn from the $m - j$ genes in B that do not belong to C.
Normalization is done by the total number of possibilities to draw n genes from m genes.

If A and B are two independent gene sets,
WebGestalt uses **Fisher's exact test** instead,

$$P = \sum_{i=k}^n \frac{\binom{n}{i} \binom{m}{j+k-i}}{\binom{m+n}{j+k}}$$

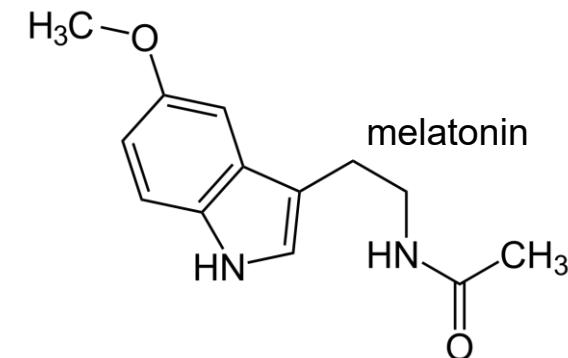
Zhang, Kirov, Snoddy (2013)
Nucl Ac Res 33: W741-W748

Effects of sleep deprivation on melatonin (SCN marker)

Melatonin is a hormone that regulates sleep-wake cycles.

On D10 + D11, melatonin **peaked significantly later** after sleep restriction:

04:15 am \pm 19 min →
Control
sleep restriction



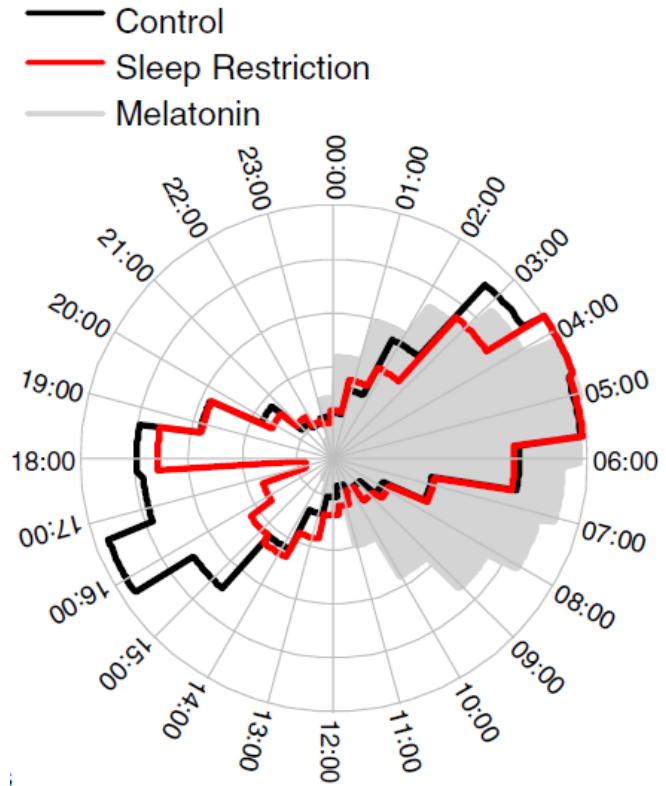
But the duration of melatonin secretion was only **insignificantly shortened**:

9 hours 53 min \pm 12 min →
sleep restriction

9 hours 35 min \pm 11 min

PNAS 110, E1132 (2013)
[wikipedia.de](https://en.wikipedia.org)

Peak times of expression



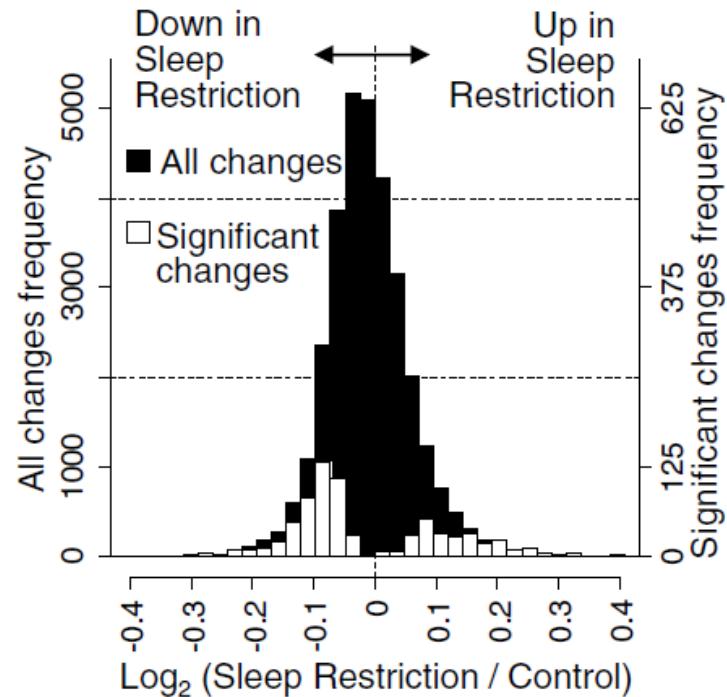
Shown are phase histogram of the peak times of prevalent circadian genes following sleep restriction or control.

The profiles of different individuals are aligned by their personal melatonin peaks.

Control: similar # of genes that peak during day / night, respectively.

Upon sleep restriction, clear reduction (> 50%) of the # of genes that peak during day time!

Global overview: changes open sleep deprivation



Frequency distribution of expression fold-changes after sleep restriction relative to control.

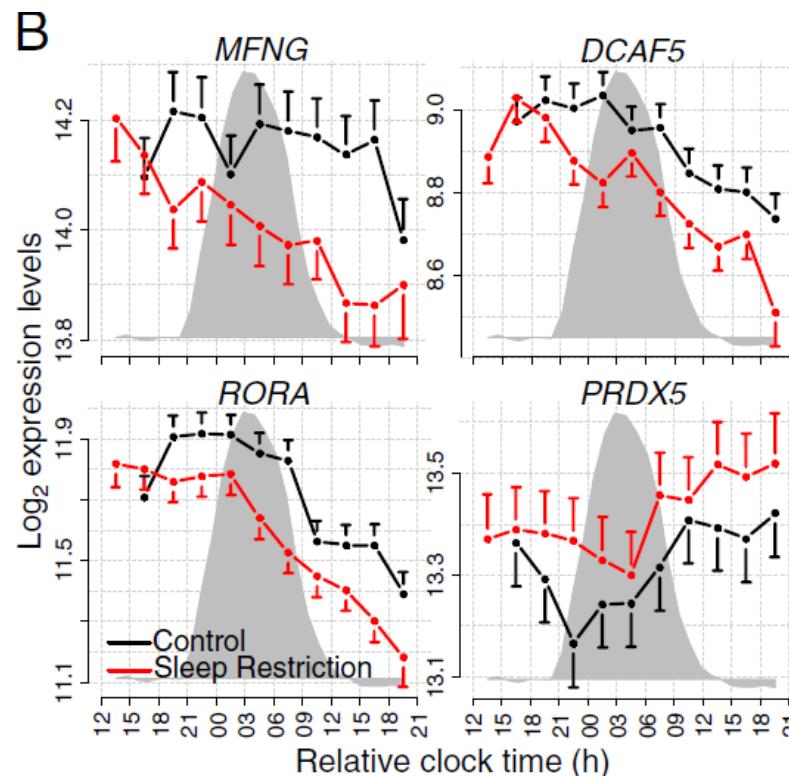
Filled area: Histogram of changes in all transcripts (31,685 probes that target 22,862 genes)

Open area: changes in transcripts identified as having a statistically significant (FDR-corrected p-value < 0.05) main effect of sleep condition (744 transcripts that target 711 genes).

444 genes are **down-regulated** upon sleep restriction (including the circadian rhythm-related genes RORA, IL6, PER2, PER3, TIMELESS, CAMK2D).

267 genes are **up-regulated** (including several circadian-rhythm related genes).

Examples of genes with significant effect of Sleep Condition



Most affected genes: $p < 10^{-6}$

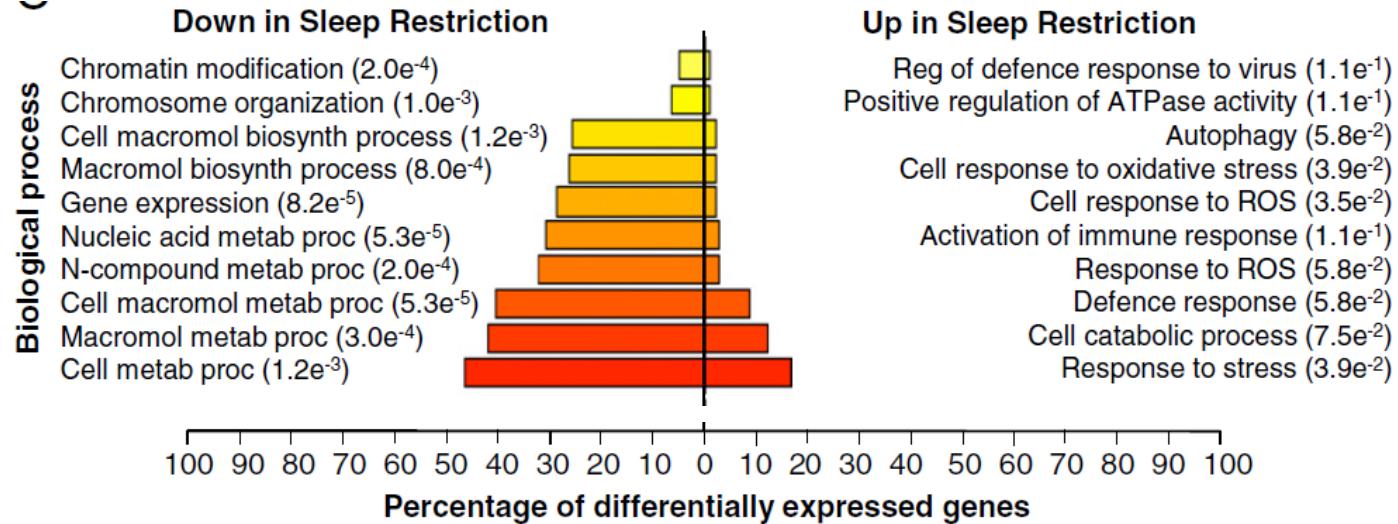
MFNG: O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase

DCAF5: is a protein-coding gene ...

RORA: retinoic acid receptor-related orphan receptor alpha is a nuclear hormone receptor – associated with circadian rhythms

PRDX5: peroxiredoxin 5

What sort of genes are differentially expressed upon sleep restriction?



Down-regulation: chromatin modification and organization, metabolism

Up-regulation: cellular response to oxidative stress and reactive oxygen

This does not sound healthy!

Small caveat: this is gene expression data, see paper #1 (protein expression may behave differently)

Top 10 enriched GO biological processes within the statistically significant differentially expressed gene list as identified by WebGestalt when using the human genome as background.

p-values are corrected by Benjamini-Hochberg method for multiple testing.

Cyclic cAMP levels in mouse brain

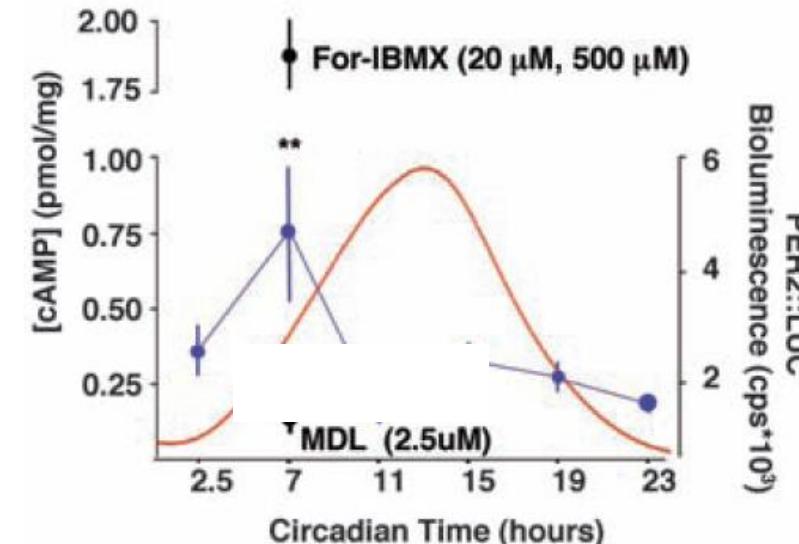
Track the molecular oscillations of the SCN as circadian emission of bioluminescence in slices from transgenic mouse brain.

As expected, a fusion protein of mPER2 and LUCIFERASE (mPER2::LUC) reported circadian protein synthesis rhythms (**red track** - right figure).

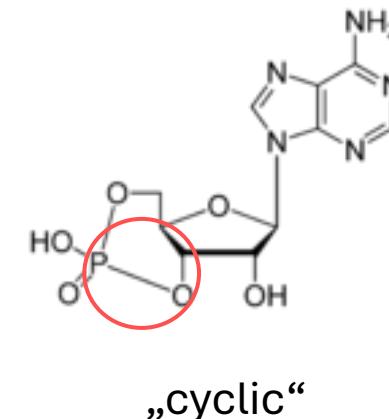
Novel finding: Under these conditions, the **cAMP** content of the SCN was also circadian (**blue track**).

O'Neill et al.
Science, 320, 949 (2008)

Cyclic adenosine monophosphate (cAMP) is a second messenger that is important in many biological processes.



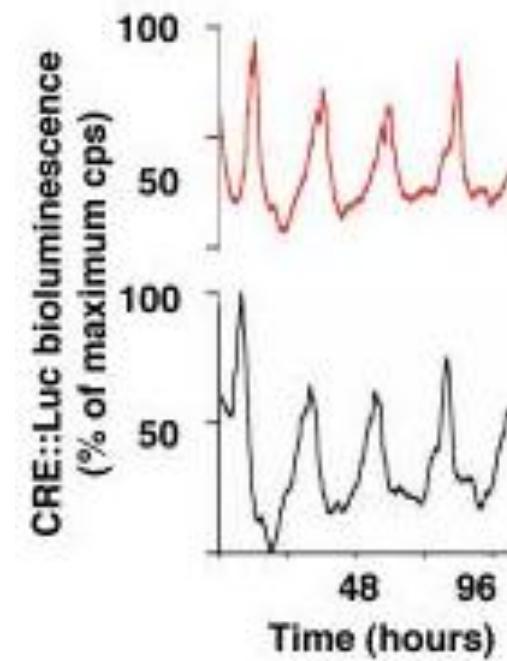
Circadian oscillation of cAMP concentration (blue) and PER2::LUC bioluminescence (red).



Cyclic cAMP levels in mouse brain

The circadian cAMP content of the SCN is accompanied by a circadian cycle in activity of cAMP response element sequences (CRE) reported by a *CRE::luciferase* adenovirus.

→ Oscillation is not specific only to *period* gene, but is also seen in CRE-response elements.



Circadian oscillation of CRE activity in two representative SCN slices (red and black) reported by *CRE:luciferase* adenovirus.

O'Neill et al.
Science, 320, 949 (2008)

Effect of MDL

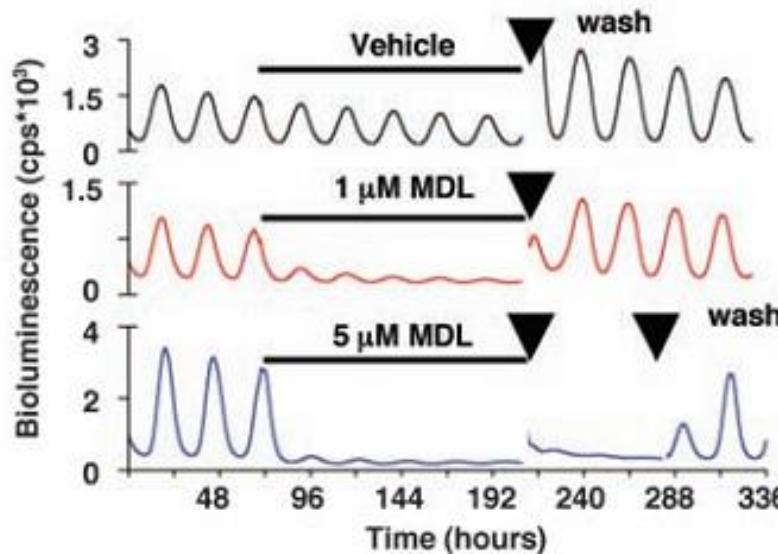
Idea: can one show that cAMP is the reason for the oscillations?

Realization: need to suppress cAMP-production in the cell.

Experiment: treat SCN slices with MDL, a potent, irreversible inhibitor of the enzyme adenylyl cyclase (that synthesizes cAMP) to reduce concentrations of cAMP to basal levels.

“Vehicle” is a control experiment.

O’Neill et al.
Science, 320, 949 (2008)



Interpretation: MDL rapidly suppressed circadian CRE:luciferase activity, presumably through loss of cAMP-dependent activation of CRE sequences.

This caused a dose-dependent **decrease** in the **amplitude** of cycles of circadian transcription and protein synthesis observed with mPer1::luciferase and mPER2::LUC.

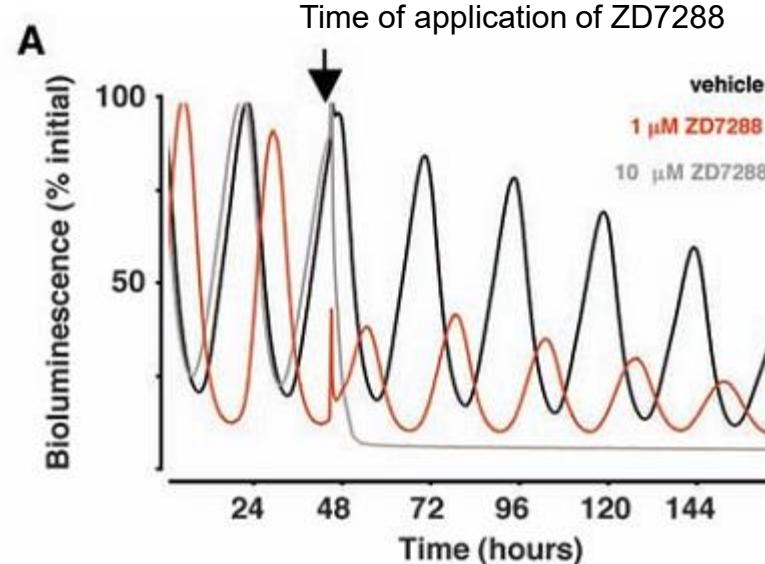
Can one block cAMP action?

Idea: If cAMP sustains the clock, interference with cAMP effectors should compromise pacemaking.

Plan A: treat brain slices with **inhibitors** of cAMP-dependent protein kinase. This had no effect, however, on circadian gene expression in the SCN.

Plan B: But cAMP also acts through hyperpolarizing cyclic nucleotide-gated ion (HCN) channels and through the guanine nucleotide-exchange factors Epac1 and Epac2 (Epac: exchange protein directly activated by cAMP).

O'Neill et al.
Science, 320, 949 (2008)



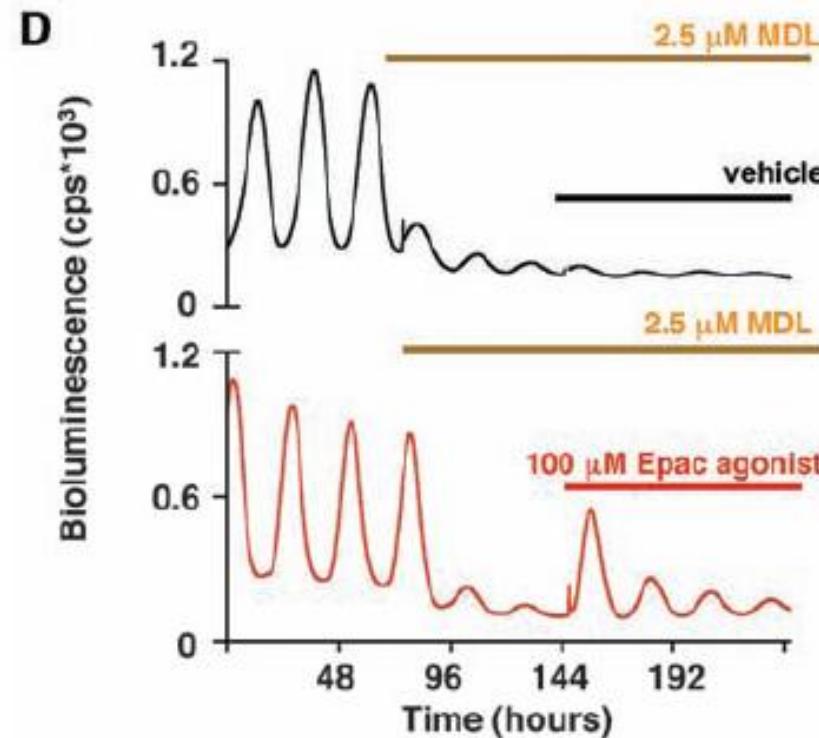
The irreversible HCN channel blocker ZD7288, which would be expected to hyperpolarize the neuronal membrane, dose-dependently damped circadian gene expression in the SCN. This is consistent with disruption of transcriptional feedback rhythms.

Can cAMP stimulation be recovered?

Experimentalists typically interrupt a cellular process and then restore it by a side-process.

Idea: **Direct activation** of the **effectors** might compensate for inactivation of adenylate cyclase by MDL.

Observation: A hydrolysis-resistant Epac agonist (bottom plot) transiently activated oscillations in transcriptional activity in SCN treated with MDL.



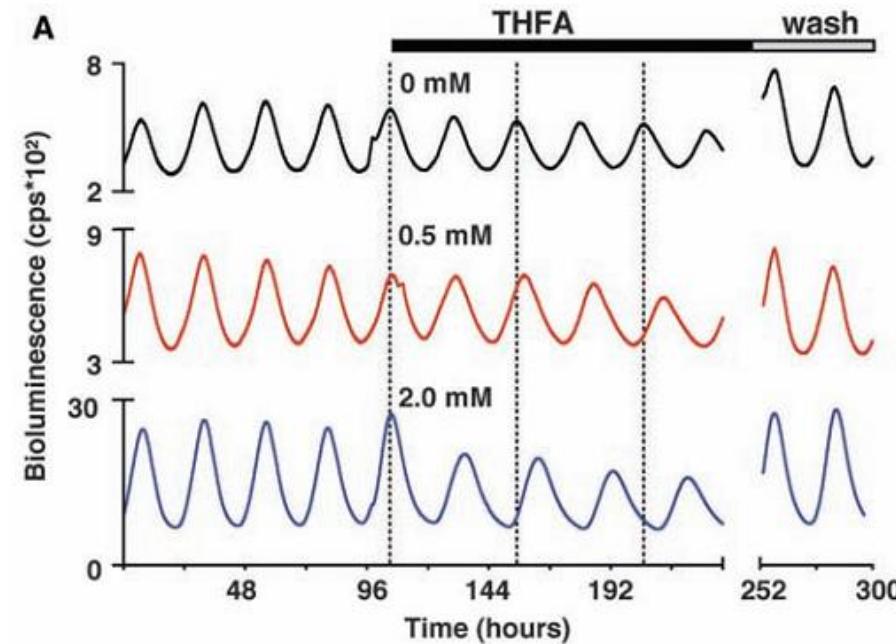
O'Neill et al.
Science, 320, 949 (2008)

slowing cAMP synthesis

Idea: if cAMP signaling is an integral component of the SCN pacemaker, altering the rate of cAMP synthesis should affect circadian period.

Experiment: 9-(tetrahydro-2-furyl)-adenine (THFA) is a noncompetitive inhibitor of adenylyl cyclase that slows the rate of G_s -stimulated cAMP synthesis, which attenuates peak concentrations.

O'Neill et al.
Science, 320, 949 (2008)



Interpretation: THFA dose-dependently increased the period of circadian pacemaking in the SCN, from 24 to 31 hours, with rapid reversal upon washout

Conclusions on cAMP-coupling

Circadian pacemaking in mammals is **sustained**.

Its canonical properties of **amplitude**, **phase**, and **period** are determined by a reciprocal interplay in which transcriptional and posttranslational feedback loops drive rhythms of cAMP signaling.

Dynamic changes in cAMP signaling, in turn, regulate transcriptional cycles.

Thus, output from the current cycle constitutes an input into subsequent cycles.

The interdependence between nuclear and cytoplasmic oscillator elements we describe for cAMP also occurs in the case of Ca^{2+} and cADPR.

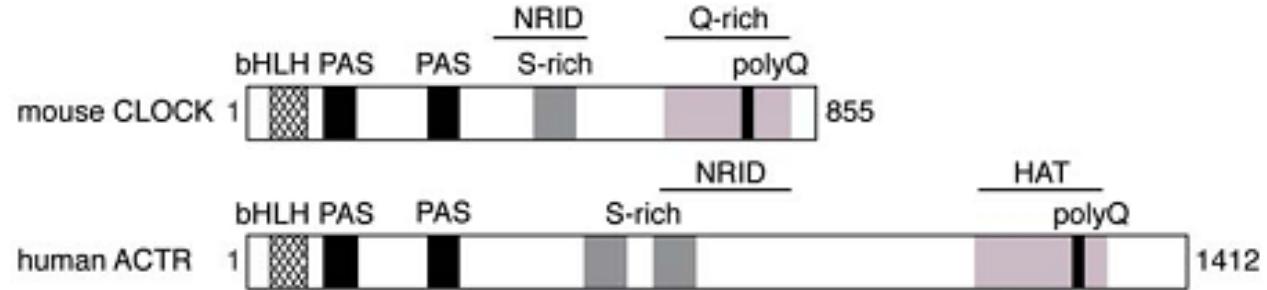
This highlights an important newly recognized common logic to circadian pacemaking in widely divergent taxa.

O'Neill et al.
Science, 320, 949 (2008)

(3) Circadian regulation of epigenetic chromatin

Circadian Regulator CLOCK Is a Histone Acetyltransferase

Masao Doi,^{1,3} Jun Hirayama,^{1,2} and Paolo Sassone-Corsi^{1,2,*}



Mouse CLOCK and human ACTR have very similar organization:
a basic helix-loop-helix (bHLH) motif (binds to DNA), Per-Arnt-Sim (PAS) domains,
serine-rich (S-rich) regions, a nuclear receptor interaction domain (NRID), and a
glutamine-rich (Q-rich) region containing a poly-glutamine (polyQ) stretch.

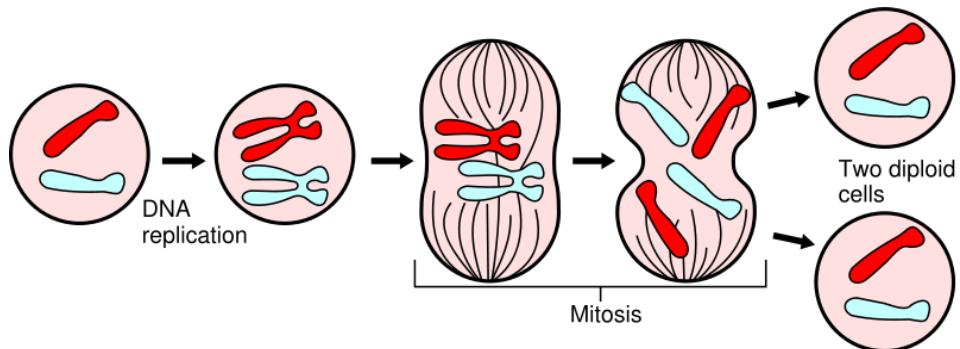
The polyQ region of hACTR is known to have **HAT activity**. Also that of CLOCK?

Histone acetyltransferases (HATs) are enzymes that acetylate conserved lysines on histone proteins by transferring an acetyl group from acetyl-CoA to form ϵ -N-acetyllysine.

V4 Cell Cycle

The cell cycle, or cell-division cycle, is the series of events that takes place in a cell leading to its division and duplication (replication).

In cells without a nucleus (prokaryotes), the cell cycle occurs via a process termed **binary fission**.



In cells with a nucleus (eukaryotes), the cell cycle can be divided in 2 brief periods:

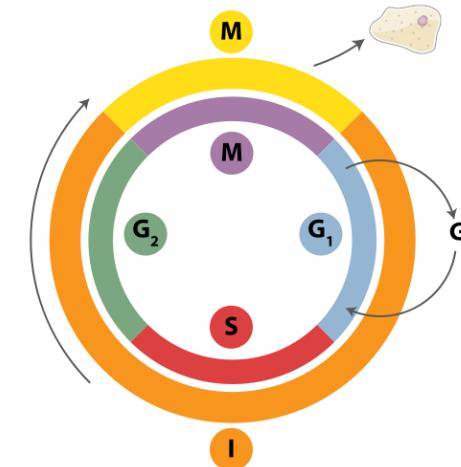
interphase—during which the cell grows, accumulating nutrients needed for mitosis and duplicating its DNA—and

the **mitosis (M) phase**, during which the cell splits itself into two distinct cells, often called "daughter cells".

Hint: read short introduction at:
<https://www.ncbi.nlm.nih.gov/books/NBK9876/>

Activity during 4 phases

State	Phase	Abbreviation	Description
quiescent/ senescent	Gap 0	G_0	A resting phase where the cell has left the cycle and has stopped dividing.
Interphase	Gap 1	G_1	Cells increase in size in Gap 1. The G_1 checkpoint control mechanism ensures that everything is ready for DNA synthesis.
	Synthesis	S	DNA replication occurs during this phase.
	Gap 2	G_2	During the gap between DNA synthesis and mitosis, the cell will continue to grow. The G_2 checkpoint control mechanism ensures that everything is ready to enter the M (mitosis) phase and divide.
Cell division	Mitosis	M	Cell growth stops at this stage and cellular energy is focused on the orderly division into two daughter cells. A checkpoint in the middle of mitosis (<i>Metaphase Checkpoint</i>) ensures that the cell is ready to complete cell division.

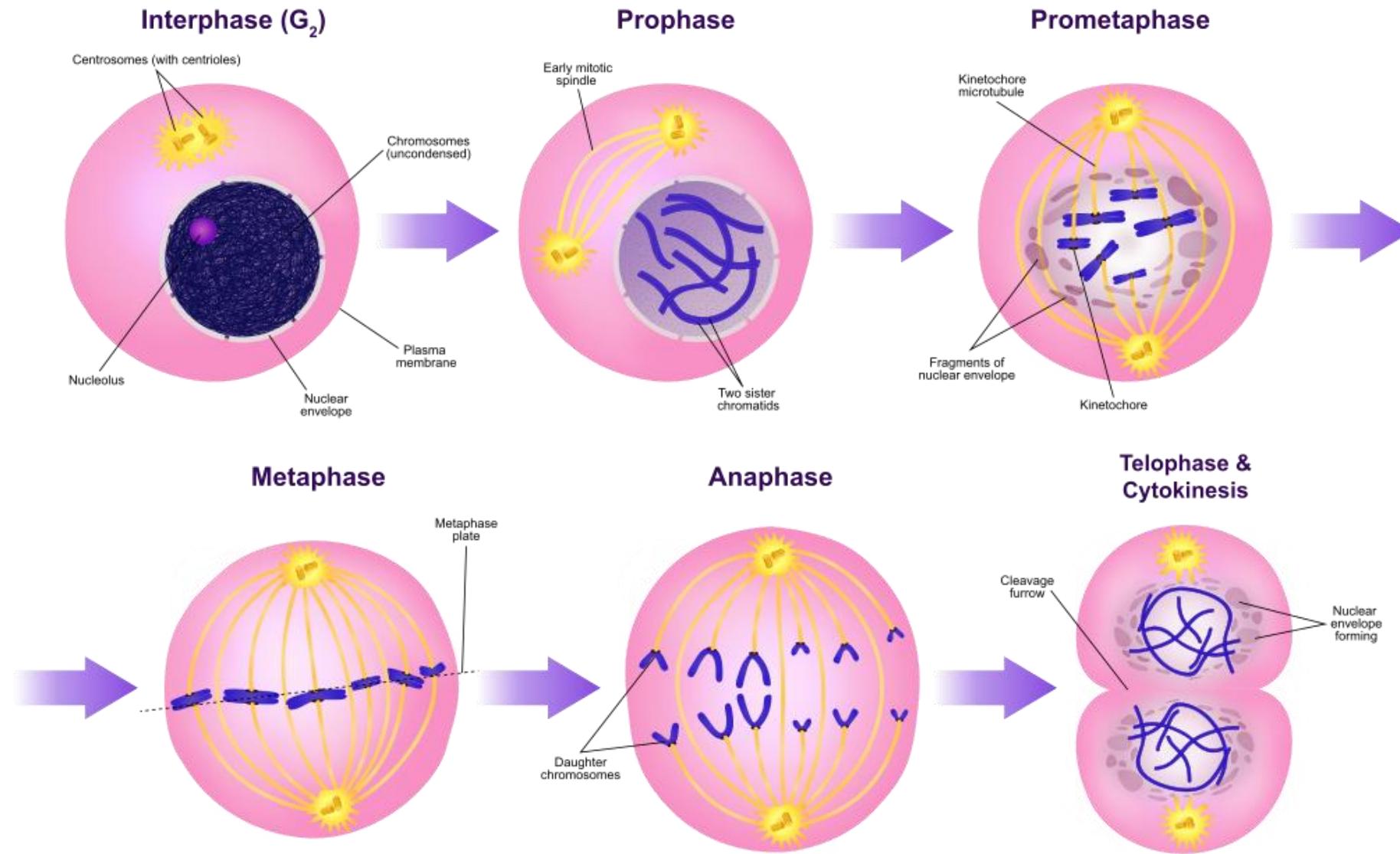


M phase itself is composed of 2 tightly coupled processes:

- **mitosis**, in which the cell's chromosomes are divided between the two daughter cells, and
- **cytokinesis**, in which the cell's cytoplasm divides in half forming distinct cells.

At the beginning of mitosis, the chromosomes condense and the **nuclear envelope breaks down**, resulting in the release of most of the contents of the nucleus into the cytoplasm. This is a very delicate stage for the life of the cell. Any mistake will lead to apoptosis.

Mitosis is complicated – involves > 1000 proteins



The yellow fibers are microtubules, formed from many copies of the protein tubulin. The **mitotic spindle** is composed of up to hundreds of thousands of microtubules (MTs) as the core constituent and roughly 1,000 additional proteins in unknown copy numbers. The exact molecular composition of the mechanism of spindle assembly over time is still unknown. Note the **centrioles** in the „centromeres“ in the upper left figure. Centrioles are figured prominently in paper #4.

Who regulates the cell cycle?

Genetic Control of the Cell-Division Cycle in Yeast, I. Detection of Mutants

Leland H. Hartwell,* Joseph Culotti, and Brian Reid†

DEPARTMENT OF GENETICS, UNIVERSITY OF WASHINGTON, SEATTLE

Later study:

L.H. Hartwell (professor), J. Culotti, J.R. Pringle, B.J. Reid (3 co-workers)

Genetic control of cell division cycle in yeast, Science, 183 (1974), pp. 46-51

In this later study, the authors characterized 150 **temperature-sensitive mutants** of the cell division cycle (cdc mutants) of *S. cerevisiae*.

These mutants are temperature-sensitive in the sense that they are unable to reproduce at 36° C (the “restrictive temperature”) but do grow normally at 23° C (the “permissive temperature”). The parent strain from which they were derived reproduces at both temperatures.

These mutations define **32 genes**, each of whose products plays an essential role in the successful completion of one event in the mitotic cycle, namely in initiation of DNA synthesis, bud emergence, DNA synthesis, medial nuclear division, late nuclear division, cytokinesis, and cell separation.

Temperature-sensitive mutations affecting the progression of the cell cycle were identified in 3 genes.

Accidental discovery of cyclins

Cell, Vol. 33, 389-396, June 1983, Copyright © 1983 by MIT

Cyclin: A Protein Specified by Maternal mRNA in Sea Urchin Eggs That Is Destroyed at Each Cleavage Division

Tom Evans,* Eric T. Rosenthal,†
Jim Youngblom,‡ Dan Distel,§ and
Tim Hunt†
Marine Biological Laboratory
Woods Hole, Massachusetts 02543

inhibition of protein synthesis by emetine or puromycin at (and Gross, 1971) blocks the stage, in which the nucleus is surrounded by stacked discs into lamellae (Messenger, 1971).

"It is difficult to believe that the behavior of the cyclins is not connected with processes involved in cell division, but at this stage we have no direct evidence that it is. . . ."

Unfortunately, we have no direct evidence as to the physiological role of cyclin, but one of its more plausible roles is promoting either directly or indirectly the breakdown of the nuclear envelope . . ."

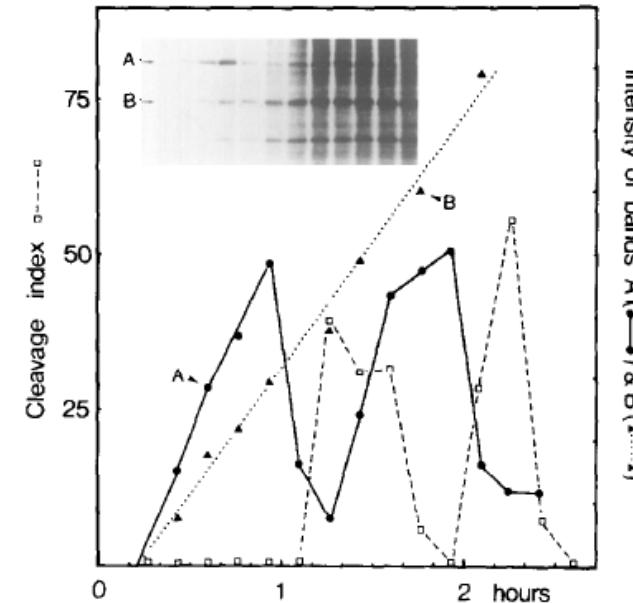


Figure 2. Correlation of the Level of Cyclin with the Cell Division Cycle
A suspension of eggs was fertilized, and after 6 min, ^{35}S -methionine was added to a final concentration of 25 $\mu\text{Ci}/\text{ml}$. Samples were taken for analysis on gels at 10 min intervals, starting at 16 min after fertilization. Samples were taken 20–30 sec later into 1% glutaraldehyde in calcium-free artificial seawater for later microscopic examination; the cleavage index is shown thus: $\square \cdots \square$. The autoradiograph shown as an inset was scanned to yield the data plotted thus: cyclin, ●—●; protein B, ▲—▲.

Who regulates the cell cycle?

The discovery of cyclin was one of 3 strands of work that came together to produce the first working model of the cell cycle oscillator.

Paul Nurse et al. identified a network of genes that controlled entry into mitosis. Its key component is the protein kinase **Cdk1**.

Masui and Smith identified **maturity-promoting factor** (MPF), a biochemical “activity” that induces meiosis and mitosis.

Lohka purified MPF. Its two subunits turned out to be Cdk1 and cyclin B.

Later work showed that

- different cyclin-Cdk complexes are activated at different points in the cell cycle,
- cyclins must be destroyed before cells can escape from mitosis

Murray AW, Cell 116, 221-234 (2004)

How are cyclins degraded?

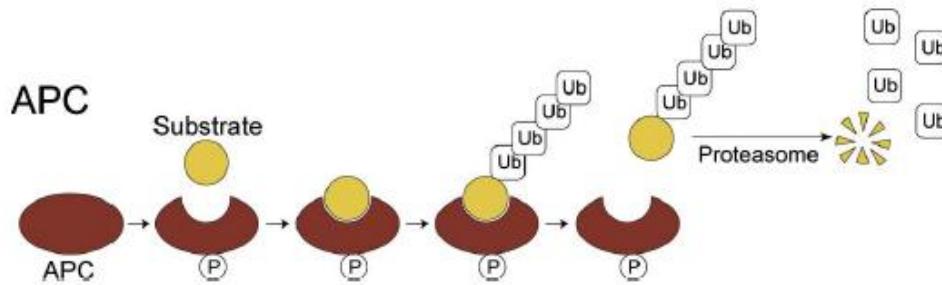
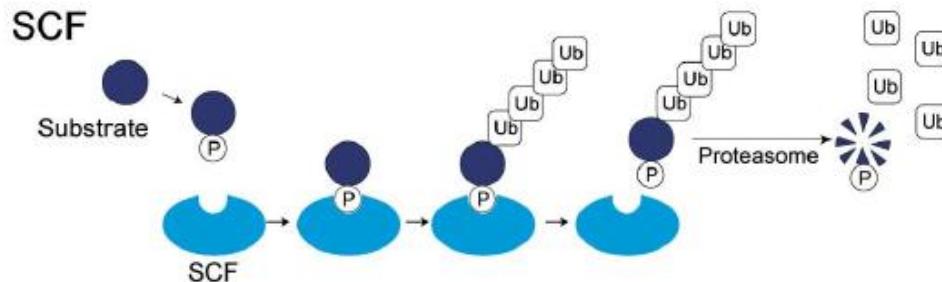
The obvious questions for cyclin were:

how is it degraded, by whom, and how is its degradation regulated?

It turned out that all known cyclins are targeted to the proteasome by the addition of a chain of ubiquitins.

Proteasome: big protein machinery that degrades proteins into 7 – 8 amino acid long peptides.

G1 cyclins are ubiquitinated by the **SCF complex**, whereas mitotic cyclins are ubiquitinated by the **anaphase-promoting complex (APC)**.



Murray AW, Cell 116, 221-234 (2004)

Cell cycle checkpoints

Cell cycle **checkpoints** are control mechanisms that ensure the fidelity of cell division in eukaryotic cells.

These checkpoints verify whether the processes at each phase of the cell cycle have been accurately completed before progression into the next phase.

An important function of many checkpoints is to **assess DNA damage**, which is detected by sensor mechanisms.

When damage is found, the checkpoint uses a signal mechanism either to stall the cell cycle until **repairs** are made or, if repairs cannot be made, to target the cell for destruction via **apoptosis** (effector mechanism).

All the checkpoints that assess DNA damage appear to utilize the same sensor-signal-effector mechanism.

Is the cyclin-CDK oscillator essential?

It has been shown that the cyclin–CDK oscillator governs the major events of the cell cycle.

In **embryonic systems** this oscillator functions in the absence of transcription, relying only on maternal stockpiles of messenger RNAs and proteins.

CDKs are also thought to act as the central oscillator in somatic cells and yeast.

Orlando et al. challenged this model and asked:

What happens in budding yeast mutant cells that do not express S-phase and mitotic cyclins?

...

Orlando et al., Nature 453, 944-947 (2008)

What happens in cyclin-mutant cells?

-> they investigated the dynamics of genome-wide transcription in budding yeast cells that are disrupted for all S-phase and mitotic cyclins (clb1,2,3,4,5,6).

These cyclin-mutant cells are unable to replicate DNA, to separate spindle pole bodies, to undergo isotropic bud growth or to complete nuclear division.

-> mutant cells are devoid of functional Clb–CDK complexes.

Expectations:

- $\Delta\text{clb1,2,3,4,5,6}$ cells should arrest at the G1/S border.
- if Clb–CDK activities are essential for triggering the transcriptional program, then periodic expression of S-phase-specific and G2/M-specific genes **should not be observed**.

Note that the yeast cells need to be **synchronized** in one particular state of the cell cycle, see e.g. <https://www.nature.com/articles/s41467-021-22689-w>

Orlando et al., Nature 453, 944-947 (2008)

Periodic transcripts in wt and cyclin-mutant cells

Aim: Identify periodically expressed genes.

For the expression x_i of each gene, i , a **Fourier score**, F_i , was computed as

$$F_i = \sqrt{\left(\sum_t \sin(\omega t) \cdot x_i(t) \right)^2 + \left(\sum_t \cos(\omega t) \cdot x_i(t) \right)^2}$$

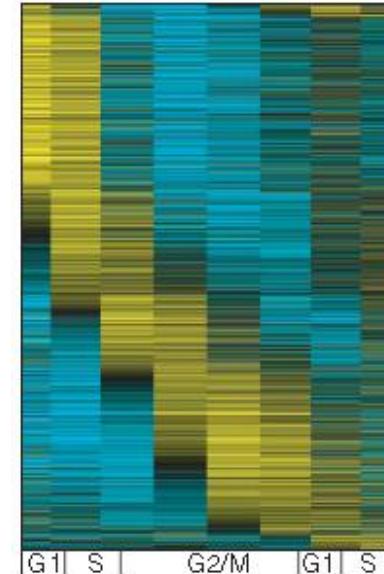
where $\omega = 2\pi/T$ and T is the interdivision time.

(If the expression x does not oscillate, projecting it on a sinus or cosinus function, will not give a large score.)

Similarly, scores were calculated for 1 000 000 artificial profiles constructed by **random shuffling** of the data points within the expression profile of the gene in question.

A p -value for significance of periodicity was calculated as the fraction of artificial profiles with Fourier scores equal to or larger than that observed for the real expression profile.

Orlando et al., Nature 453, 944-947 (2008)



Heat maps depicting mRNA levels of 1271 **periodic genes**

(having significant p -values) for wild-type cells.

Each row represents data for one gene.

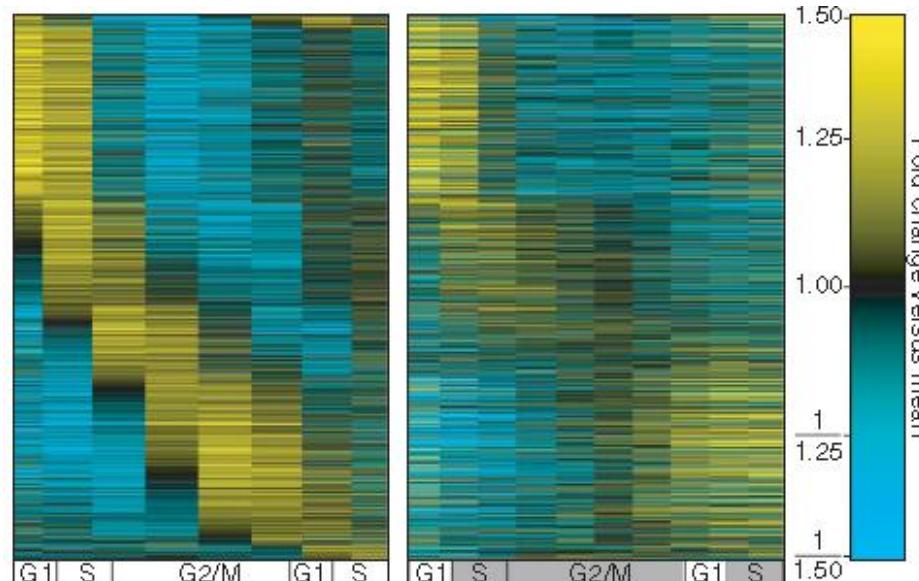
Periodic transcripts in wt and cyclin-mutant cells

mRNA levels of periodic genes for wild-type (a) and cyclin-mutant (b) cells.

Each row in a and b represents data for the same gene.

The S and G2/M phases of the cyclin-mutant timeline are shaded.

By conventional definitions, cyclin-mutant cells should arrest at the G1/S-phase border.



Observations

- (1) Expression of 883 genes is periodically altered in the mutant as well so that they are likely regulated by B-cyclin CDK,
- (2) Hence, although mutant cells are arrested at G1/S border, gene regulation program seems to continue ...

Orlando et al., Nature 453, 944-947 (2008)

Transcriptional dynamics of cyclin-CDK regulated genes

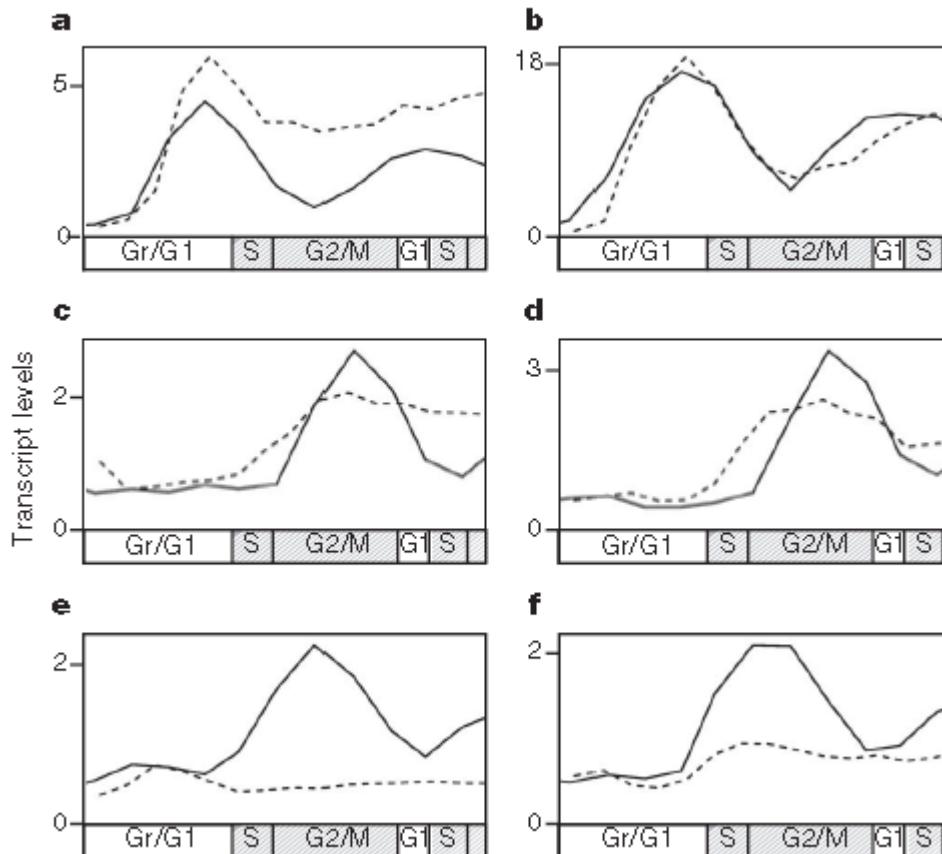
(a) Gene CLN2 (that is regulated by late G1-transcript SBF) is **not fully repressed** in mutant. Ok

(b) Gene RNR1 (that is regulated by MBF) is **not affected**.

Genes SIC1 (c) and NIS1 (d) are regulated by Ace2/Swi5. These TFs are usually excluded from the nucleus by CDK phosphorylation until late meiosis. In cyclin-mutant cells, nuclear exclusion of Swi5 and Ace2 is probably lost
-> **Early onset** observed in the mutant.

The Clb2-cluster genes CDC20 (e) and ACE2 (f) are **strongly down-regulated**.

Solid lines, wild-type cells;
dashed lines, cyclin-mutant cells.



Orlando et al., Nature 453, 944-947 (2008)

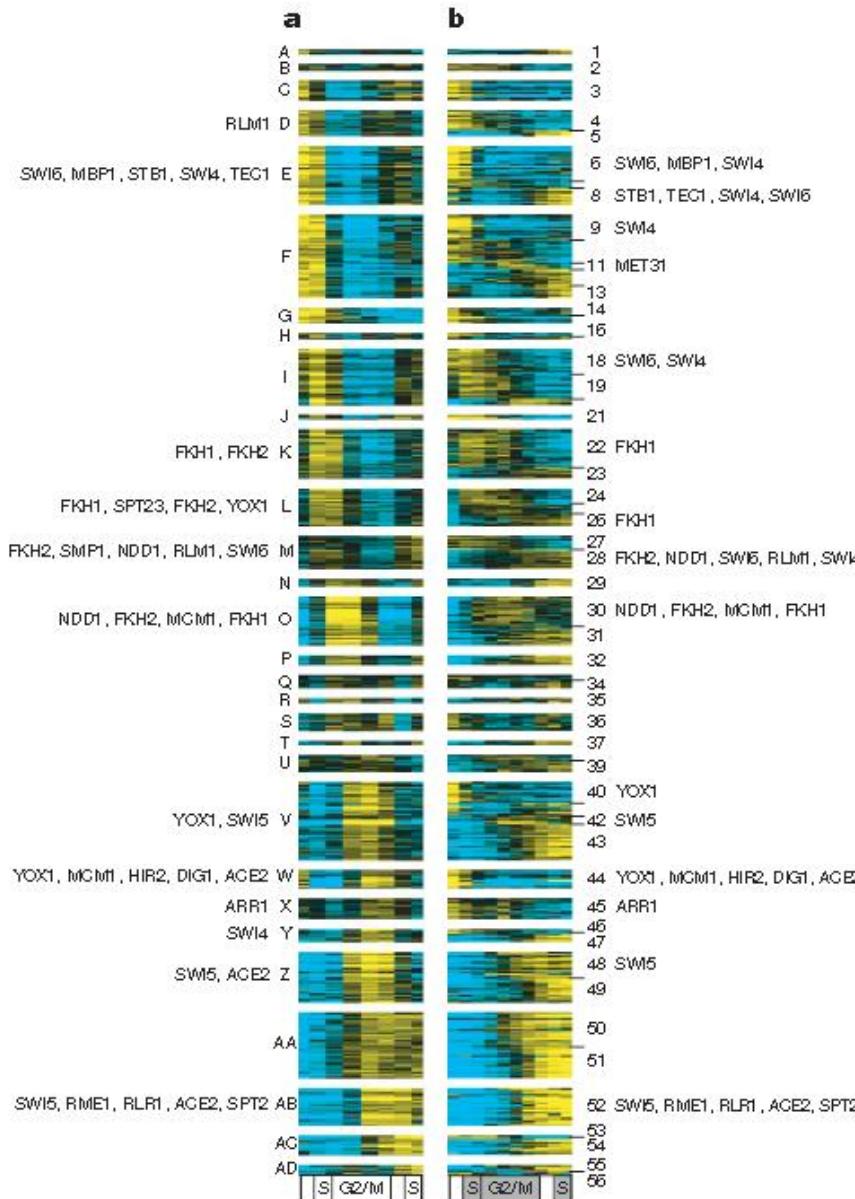
Clustering of genes

Cluster genes showing altered behaviors in cyclin-mutant cells.

a, Clusters of genes with similar expression patterns in **wild-type cells**.

b, Subclusters of genes with similarly altered expression patterns in **cyclin-mutant cells**.

Associate each cluster with up to 5 TFs that have over-represented binding motifs in the promoters of these genes (hypergeometric test).



Orlando et al., Nature 453, 944-947 (2008)

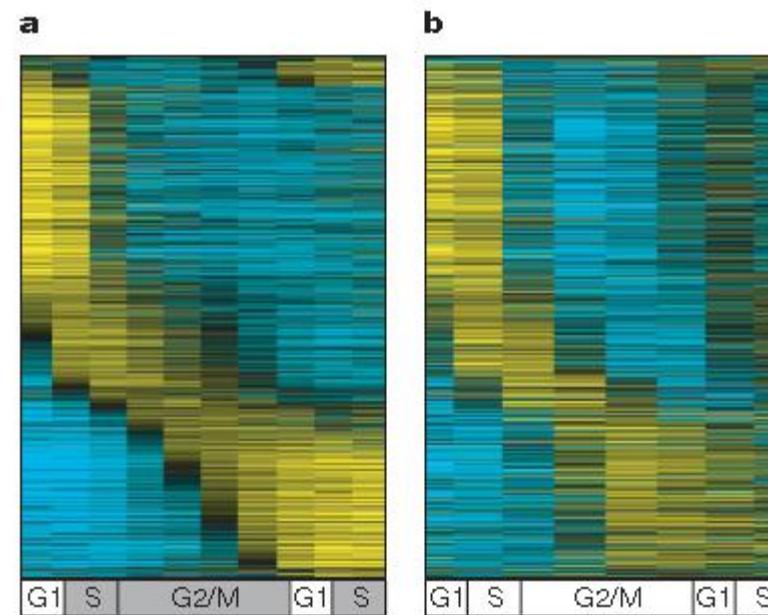
Independent transcriptional program

Unexpectedly, the periodic transcription program is largely intact in cyclin mutant cells that arrest at the G1/S border.

a, b, Genes maintaining periodic expression in cyclin-mutant cells (a) show similar dynamics in wildtype cells (b).

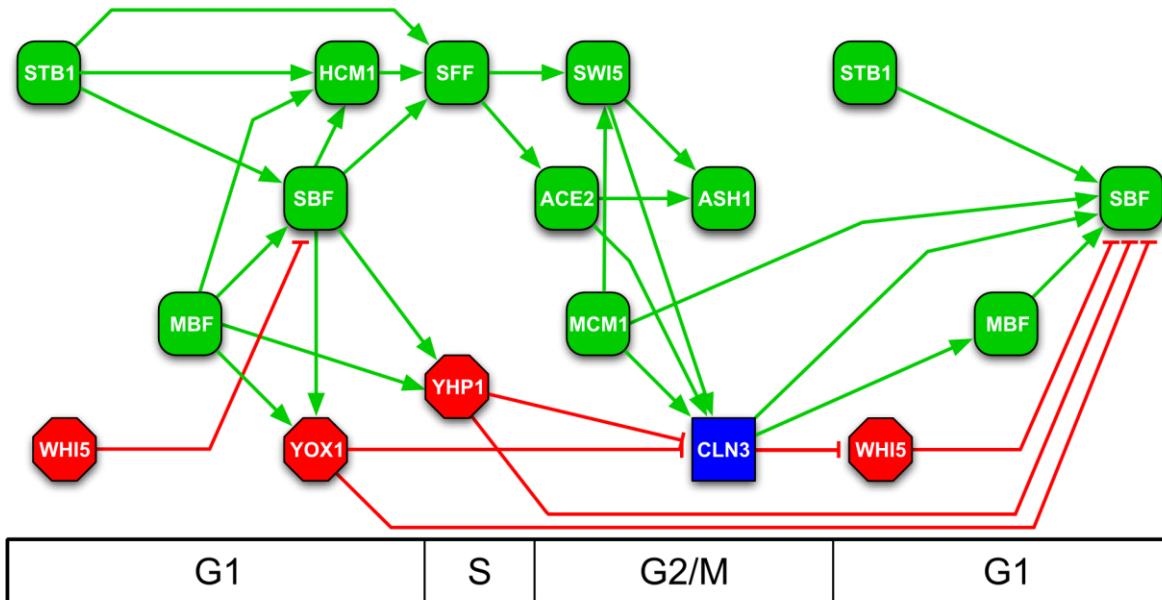
On the other hand, nearly 70% of the genes identified as periodic in wild-type cells are still expressed on schedule in cyclin-mutant cells.

This demonstrates the existence of a cyclin–CDK-independent mechanism that regulates temporal transcription dynamics during the cell cycle.



Orlando et al.,
Nature 453, 944-947 (2008)

Generate TF networks for wt and cyclin-mutant cells



Transcriptional activators are depicted in green, repressors in red, and the cyclin Cln3 in blue.

Periodically expressed TFs are placed on the cell-cycle timeline on the basis of the time of peak transcript levels.

Arrows indicate a documented interaction between a TF and promoter elements upstream of a gene encoding another TF.

Orlando et al., Nature 453, 944-947 (2008)

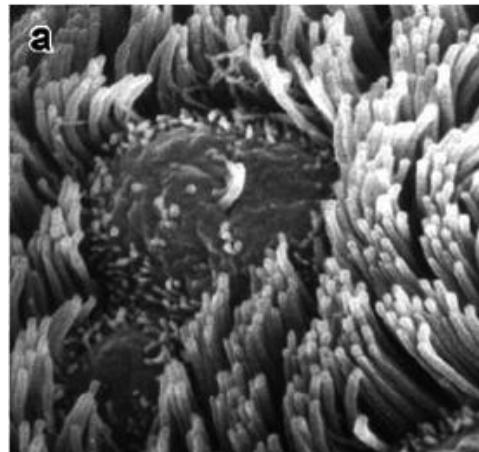
Summary

The cyclin–CDK oscillator governs the major events of the cell cycle.

Simple Boolean networks or ODE-models can generate oscillatory behavior.

There exists an independent TF network in yeast (in all higher eukaryotes?) that drives periodic expression of many genes throughout cell cycle.

Paper #4 (topic of assignment 2) describes how **multiciliated cells** develop. The machinery for this reuses some components of the cell cycle mentioned today.



Scanning electron microscopy (SEM)
image of oviduct cilia. The central cell
shows a primary cilium. Surrounding cells
are multiciliated with motile cilia

Satir, Christensen (2007) 69, 377, Ann Rev Physiol

V5: Protein phosphorylation during cell cycle

Protein **phosphorylation** and **dephosphorylation** are highly controlled biochemical processes that respond to various intracellular and extracellular stimuli. They belong to the class of post-translational modifications (PTMs).

Note: phosphorylation of histone tails also belongs to this class of PTMs.

Phosphorylation status modulates protein activity by

- influencing the tertiary and quaternary **structure** of a protein,
- controlling its **subcellular distribution** (e.g cytoplasm ⇔ nucleus for Per/Cry), and
- regulating its **interactions** with other proteins.

Regulatory protein phosphorylation is a **transient modification** that is often of low occupancy or “stoichiometry”.

Low occupancy means that only a fraction of the copies of a particular protein may be phosphorylated on a given site at any particular time.

Cell Cycle and the Phosphoproteome

CELL CYCLE

Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis

Jesper V. Olsen,^{1,2*} Michiel Vermeulen,^{1,3*} Anna Santamaria,^{4*} Chanchal Kumar,^{1,5*}
Martin L. Miller,^{2,6} Lars J. Jensen,² Florian Gnad,¹ Jürgen Cox,¹ Thomas S. Jensen,⁷
Erich A. Nigg,⁴ Søren Brunak,^{2,7} Matthias Mann^{1,2†}

(Published 12 January 2010; Volume 3 Issue 104 ra3)

www.SCIENCESIGNALING.org 12 January 2010 Vol 3 Issue 104 ra3

Aim: Analyze all proteins that are modified by phosphorylation during different stages of the cell cycle of human HeLa cells.

Ion-exchange chromatography + HPLC + MS + sequencing led to the identification of 6695 phosphorylated proteins („the phospho-proteome“). From this, 6027 quantitative cell cycle profiles of individual proteins were obtained.

A total of 24,714 phosphorylation events were identified.
20,443 of them were assigned to a specific residue with high confidence.

Finding: about 70% of all proteins get phosphorylated.

Protein quantification by SILAC

ARTICLE

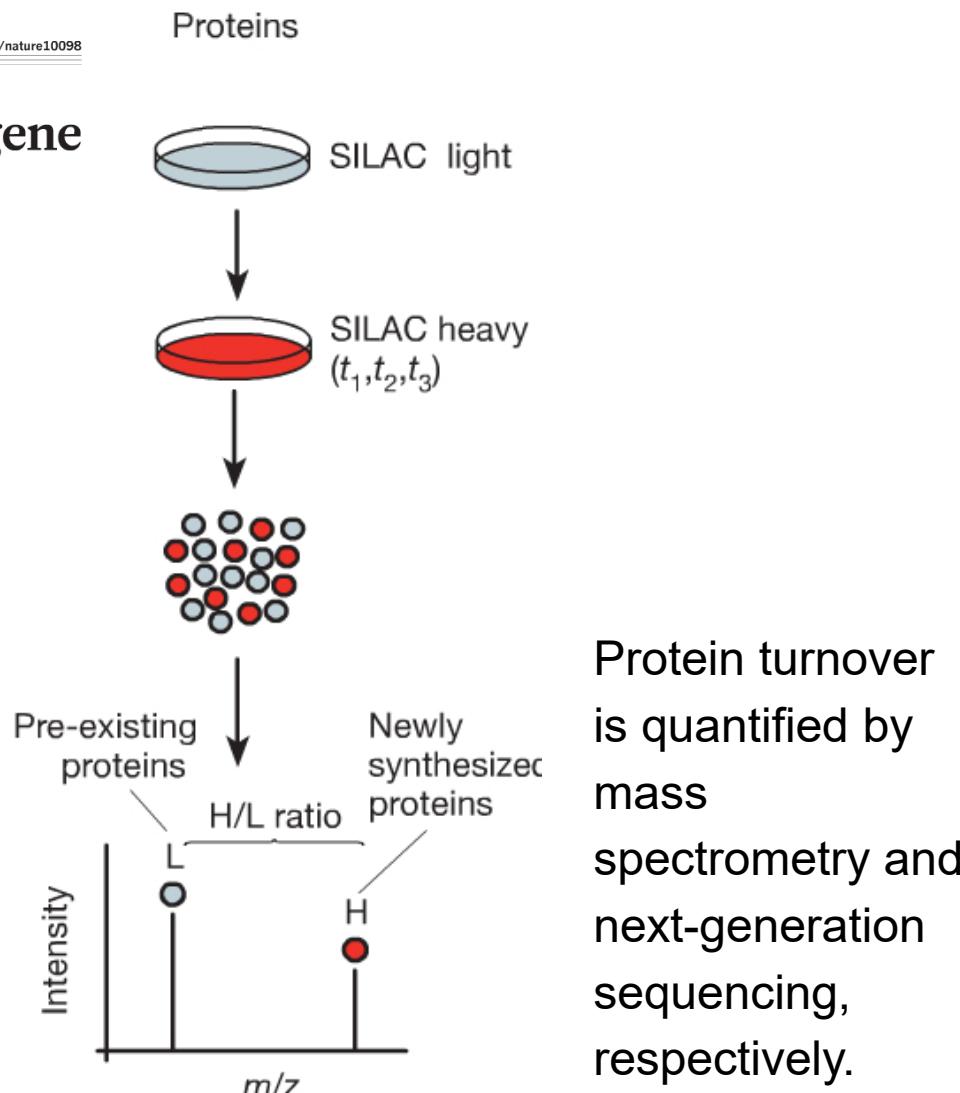
doi:10.1038/nature10098

Global quantification of mammalian gene expression control

Björn Schwanhäuser¹, Dorothea Busse¹, Na Li¹, Gunnar Dittmar¹, Johannes Schuchhardt², Jana Wolf¹, Wei Chen¹
& Matthias Selbach¹

SILAC: „stable isotope labelling by amino acids in cell culture“ means that cells are cultivated in a medium containing heavy stable-isotope versions of essential amino acids.

When non-labelled (i.e. light) cells are transferred to heavy SILAC growth medium, newly synthesized proteins incorporate the heavy label while pre-existing proteins remain in the light form.



Schwanhäuser et al. Nature 473, 337 (2011)

H/L ratios of individual proteins

Mass spectra of peptides for two proteins.

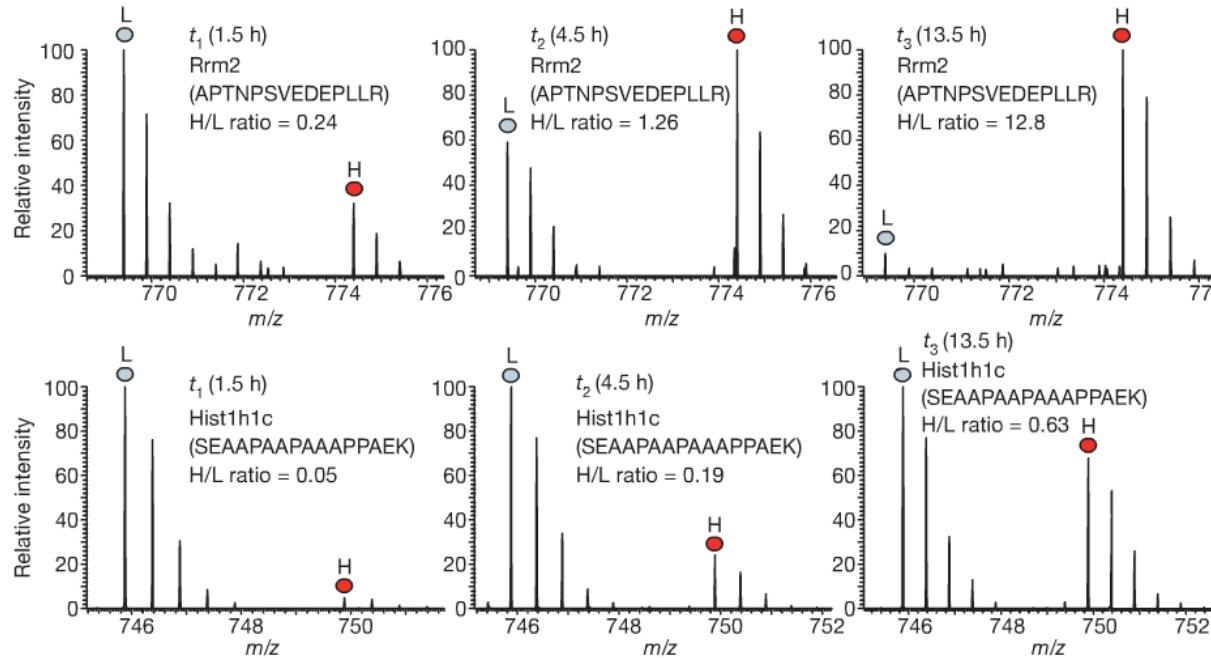
Top: **high-turnover protein**

Bottom: **low-turnover protein**.

Over time, the heavy to light (H/L) ratios increase.

H-concentration of high-turnover protein saturates.

That of low-turnover protein still increases.

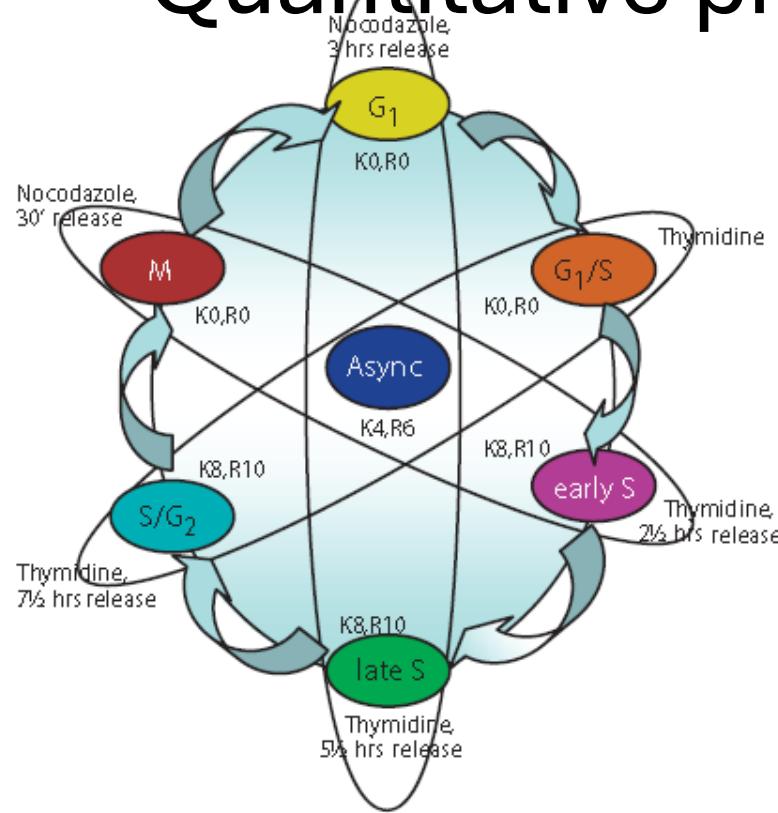


This example illustrates the principles of SILAC and mass spectroscopy signals (peaks).
 m/z : mass over charge ratio of a peptide fragment

In the Olson et al. study, the authors used H and L forms to label different stages of the cell cycle.

Schwanhäuser et al. Nature 473, 337 (2011)

Quantitative proteomic analysis



Center: asynchronously growing cell population as internal standard to allow normalization between experiments.

HeLa cells were SILAC-labeled with 3 different isotopic forms (light – medium – heavy) of arginine and lysine.

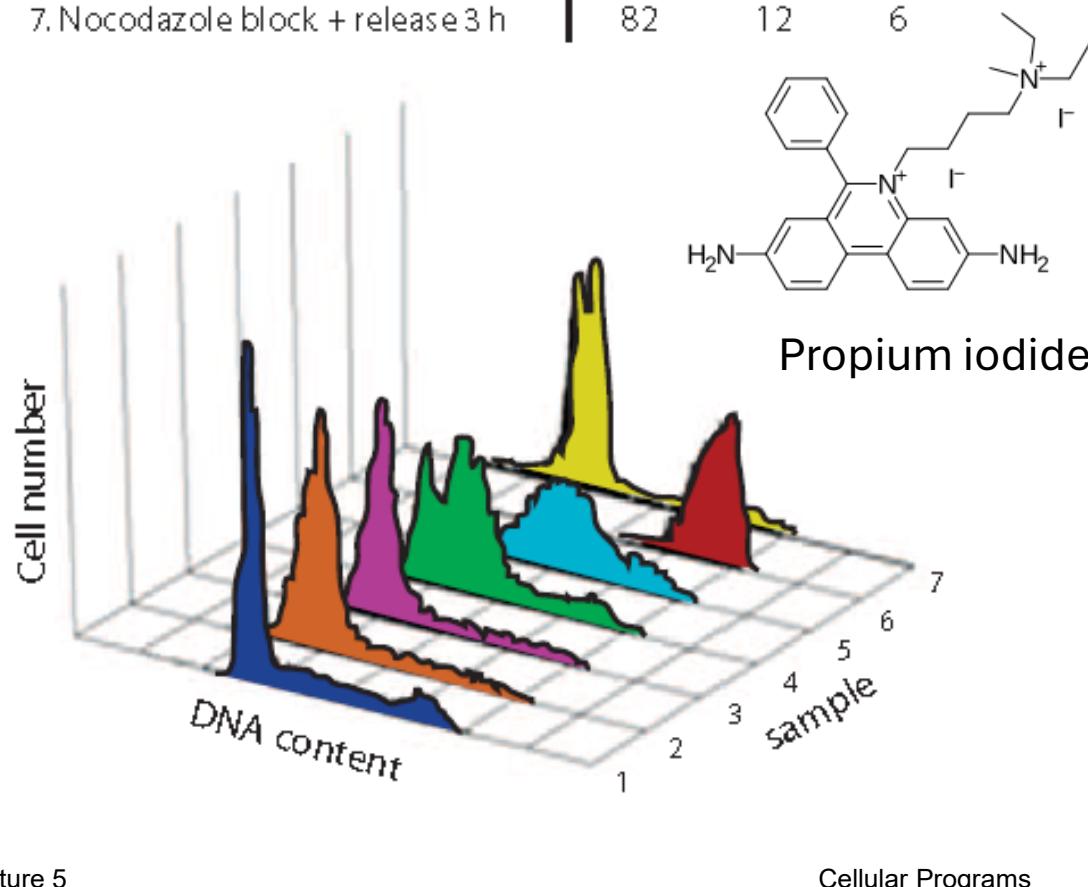
3 individual populations of heavy and light SILAC cells were synchronized with a **thymidine block** (analog of thymine, blocks entry into S phase).

Cells were then collected at 6 different time points across the cell cycle after release from the **thymidine arrest**.

Out of this, 2 samples were collected after a further **cell cycle arrest** with **nocodazole** and release. (Nocodazole interferes with polymerization of microtubules.)

FACS profiles of individual HeLa populations

	% Cells		
	G ₁	S	G ₂ /M
1. Asynchronous	64	27	9
2. Thymidine block	50	46	4
3. Thymidine block + release 2½ h	36	60	4
4. Thymidine block + release 5½ h	23	70	7
5. Thymidine block + release 7½ h	15	70	15
6. Nocodazole block + release ½ h	1	11	88
7. Nocodazole block + release 3 h	82	12	6



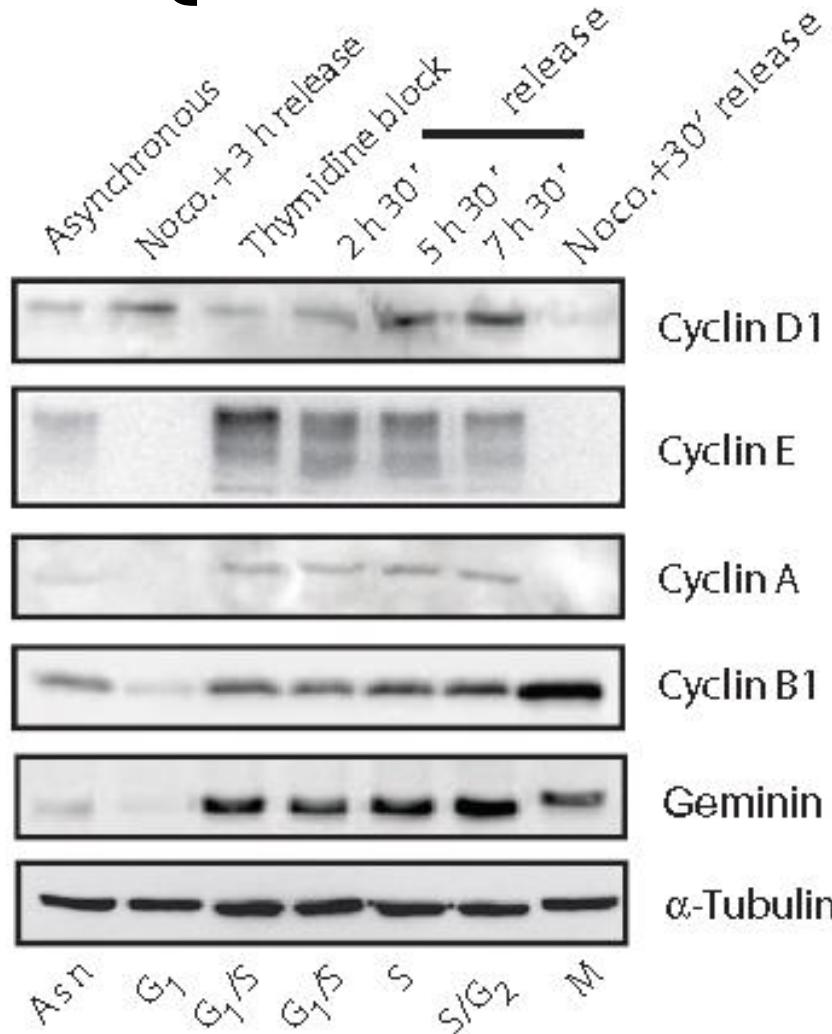
Cells were fixed and collected by centrifugation.

Then the **DNA content** of the cells was determined with propidium iodide.

The DNA content is the basis for classifying the **state** along the cell cycle.

→ Samples 1 – 5 are not pure states, but **mixtures**.
Nocodazole block is quite efficient in synchronizing cells (samples 6 and 7).

Quantification of cell cycle markers

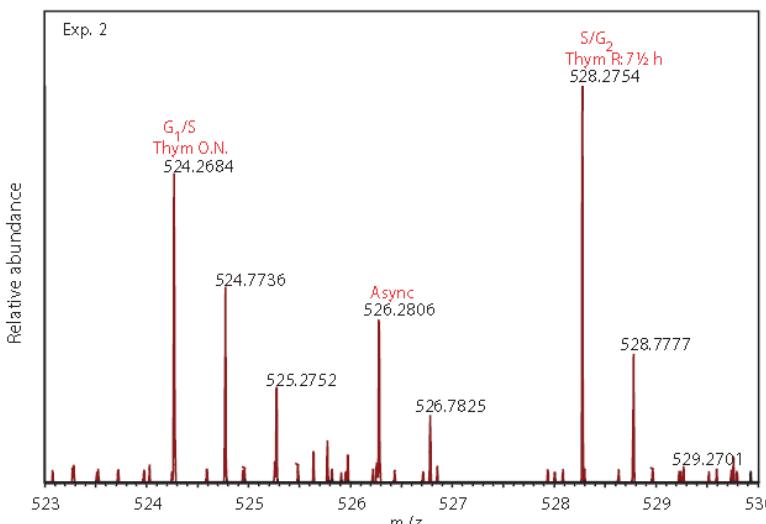
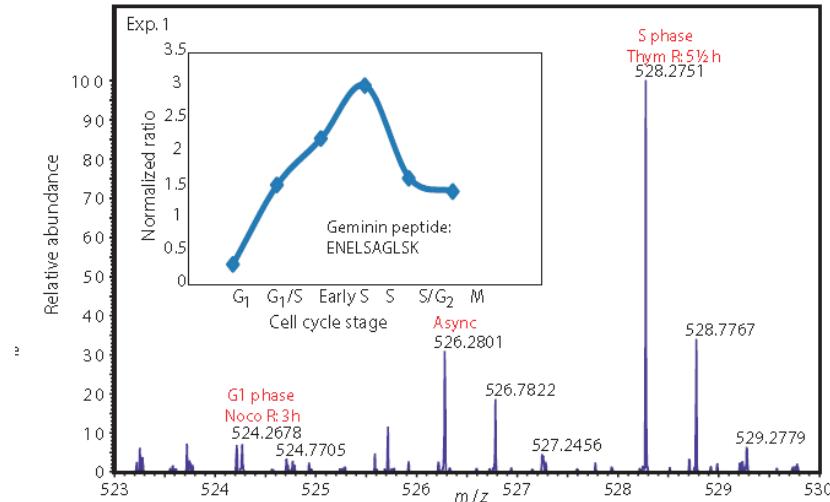
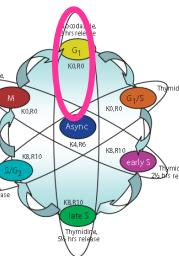


Immunoblot analysis of known cell cycle marker proteins in the different cell populations (α -tubulin is a control).

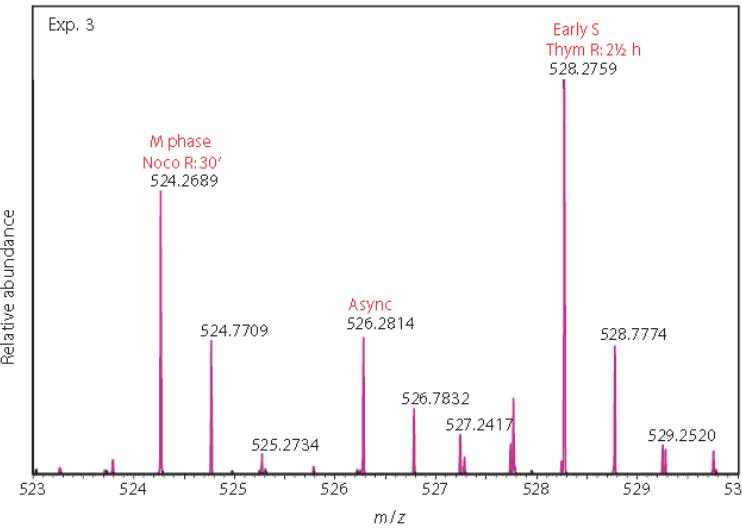
The abundance of 20% of the proteome changed by at least 4-fold throughout the cell cycle (difference between lowest and highest abundance).

Because a **fourfold change** best accounted for the dynamics of already described cell cycle components, this ratio was used as a threshold for subsequent analysis.

Experiment 1: mixture of
L = G1 phase
M = Async
H = S phase



Monitor protein abundance by MS



Representative MS data showing how the abundance of the proteins was monitored in 3 experiments to obtain information from the 6 stages of the cell cycle.

The data show the MS analysis of a tryptic SILAC peptide triplet derived from the cell cycle marker protein **Geminin**.

Relative peptide abundance changes were normalized to the medium SILAC peptide derived from the asynchronously grown cells in all three experiments.
The inset of Exp. 1 shows the combined six-time profile of Geminin over the cell cycle.

V5: Protein phosphorylation during cell cycle

Protein **phosphorylation** and **dephosphorylation** are highly controlled biochemical processes that respond to various intracellular and extracellular stimuli. They belong to the class of post-translational modifications (PTMs).

Note: phosphorylation of histone tails also belongs to this class of PTMs.

Phosphorylation status modulates protein activity by

- influencing the tertiary and quaternary **structure** of a protein,
- controlling its **subcellular distribution** (e.g cytoplasm ⇔ nucleus for Per/Cry), and
- regulating its **interactions** with other proteins.

Regulatory protein phosphorylation is a **transient modification** that is often of low occupancy or “stoichiometry”.

Low occupancy means that only a fraction of the copies of a particular protein may be phosphorylated on a given site at any particular time.

Cell Cycle and the Phosphoproteome

CELL CYCLE

Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis

Jesper V. Olsen,^{1,2*} Michiel Vermeulen,^{1,3*} Anna Santamaría,^{4*} Chanchal Kumar,^{1,5*}
Martin L. Miller,^{2,6} Lars J. Jensen,² Florian Gnad,¹ Jürgen Cox,¹ Thomas S. Jensen,⁷
Erich A. Nigg,⁴ Søren Brunak,^{2,7} Matthias Mann^{1,2†}

(Published 12 January 2010; Volume 3 Issue 104 ra3)

www.SCIENCESIGNALING.org 12 January 2010 Vol 3 Issue 104 ra3

Aim: Analyze all proteins that are modified by phosphorylation during different stages of the cell cycle of human HeLa cells.

Ion-exchange chromatography + HPLC + MS + sequencing led to the identification of 6695 phosphorylated proteins („the phospho-proteome“). From this, 6027 quantitative cell cycle profiles of individual proteins were obtained.

A total of 24,714 phosphorylation events were identified.
20,443 of them were assigned to a specific residue with high confidence.

Finding: about 70% of all proteins get phosphorylated.

Protein quantification by SILAC

ARTICLE

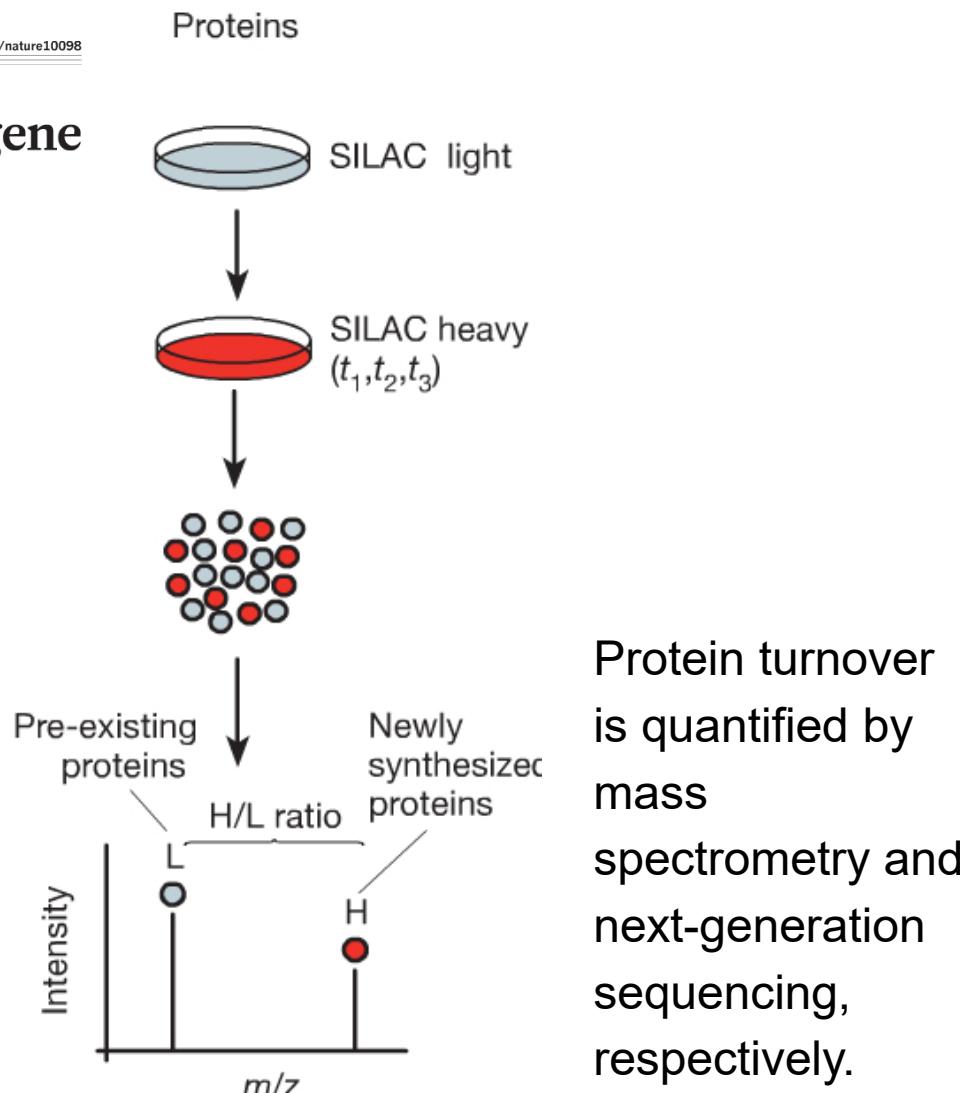
doi:10.1038/nature10098

Global quantification of mammalian gene expression control

Björn Schwanhäuser¹, Dorothea Busse¹, Na Li¹, Gunnar Dittmar¹, Johannes Schuchhardt², Jana Wolf¹, Wei Chen¹
& Matthias Selbach¹

SILAC: „stable isotope labelling by amino acids in cell culture“ means that cells are cultivated in a medium containing heavy stable-isotope versions of essential amino acids.

When non-labelled (i.e. light) cells are transferred to heavy SILAC growth medium, newly synthesized proteins incorporate the heavy label while pre-existing proteins remain in the light form.



Schwanhäuser et al. Nature 473, 337 (2011)

H/L ratios of individual proteins

Mass spectra of peptides for two proteins.

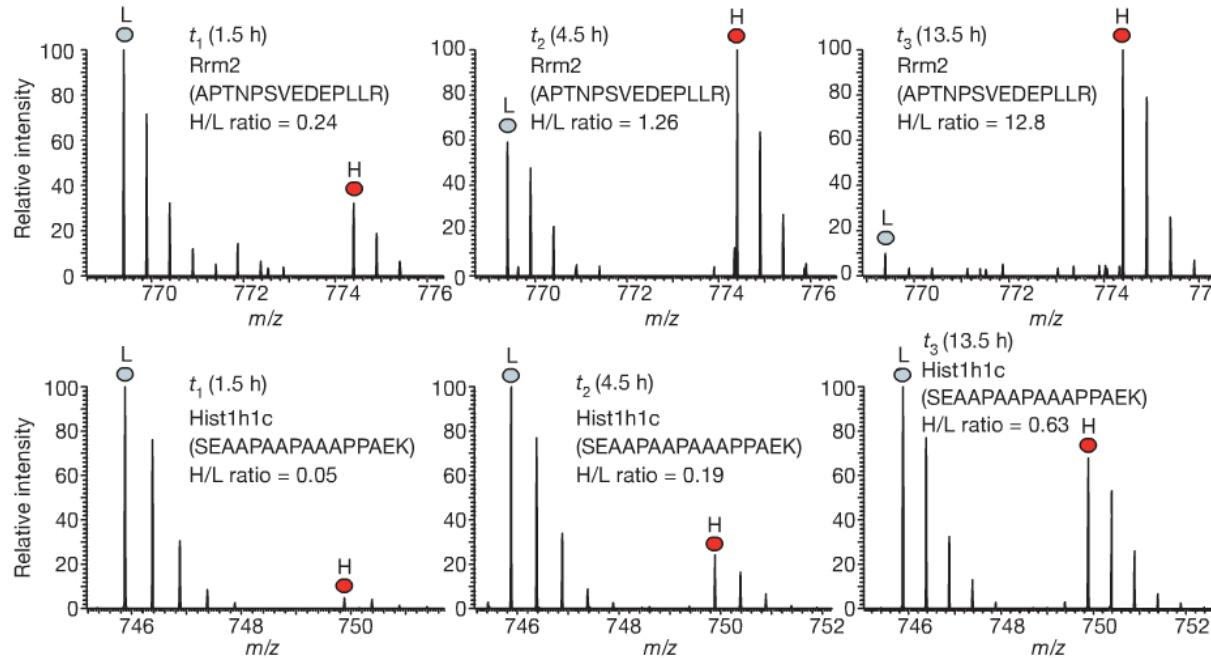
Top: **high-turnover protein**

Bottom: **low-turnover protein**.

Over time, the heavy to light (H/L) ratios increase.

H-concentration of high-turnover protein saturates.

That of low-turnover protein still increases.

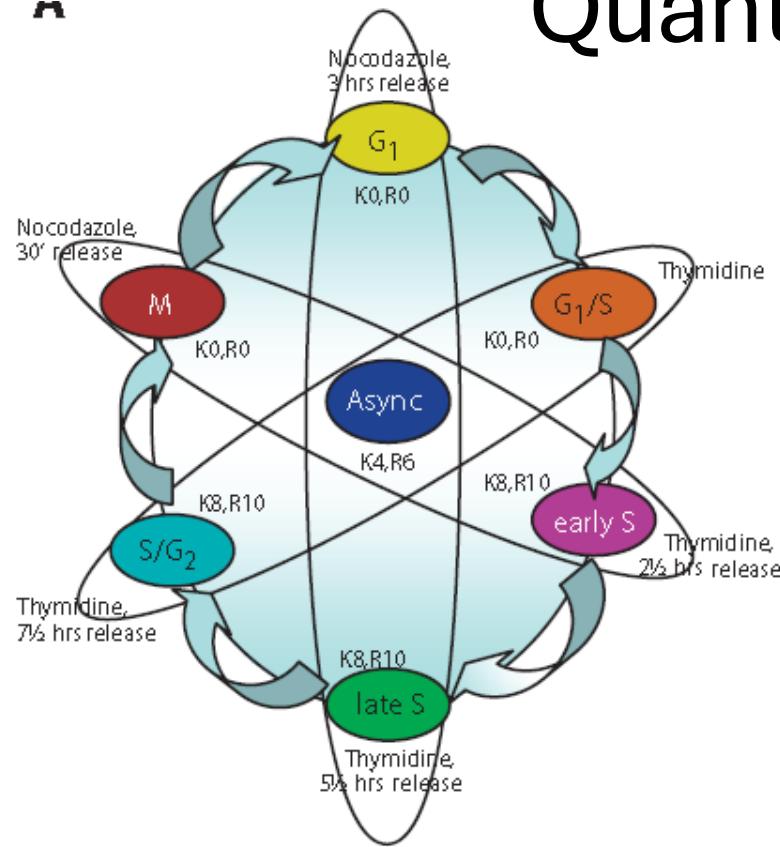


This example illustrates the principles of SILAC and mass spectroscopy signals (peaks).
 m/z : mass over charge ratio of a peptide fragment

In the Olson et al. study, the authors used H and L forms to label different stages of the cell cycle.

Schwanhäuser et al. Nature 473, 337 (2011)

Quantitative proteomic analysis



Center: asynchronously growing cell population as internal standard to allow normalization between experiments.

HeLa cells were SILAC-labeled with 3 different isotopic forms (light – medium – heavy) of arginine and lysine.

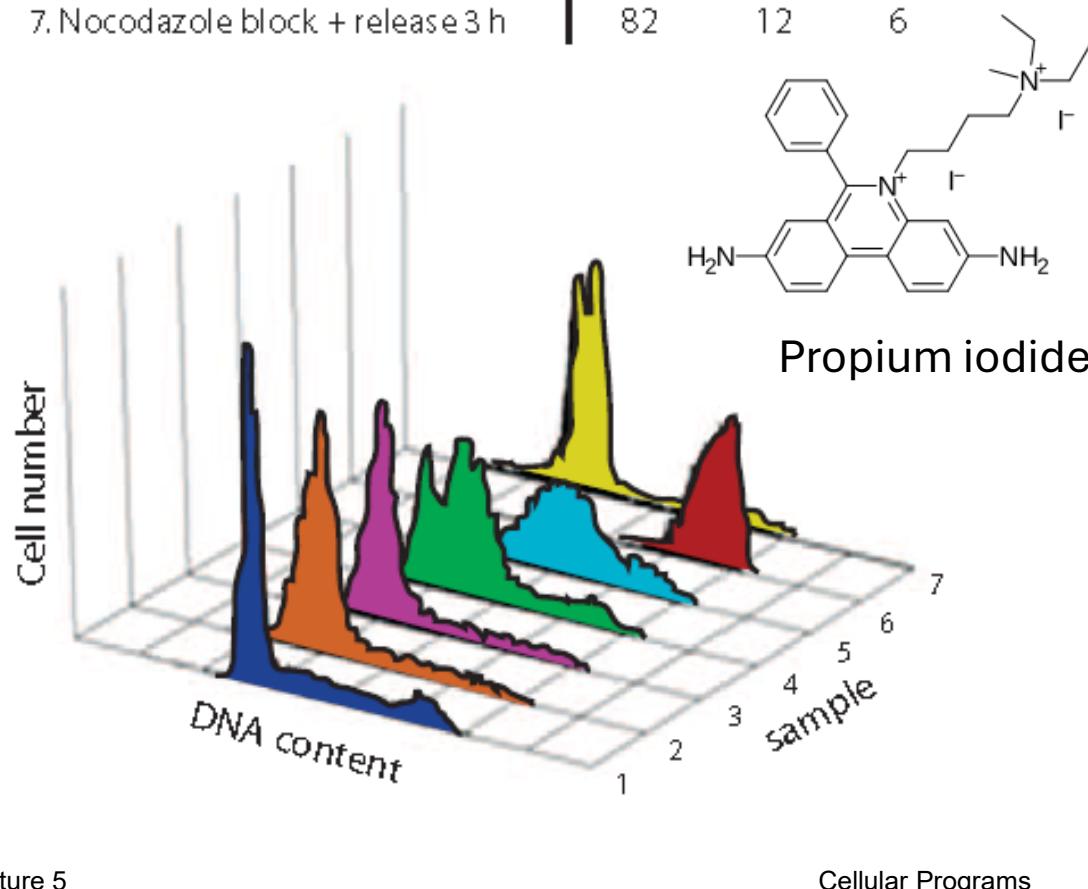
3 individual populations of heavy and light SILAC cells were synchronized with a **thymidine block** (analog of thymine, blocks entry into S phase).

Cells were then collected at 6 different time points across the cell cycle after release from the **thymidine arrest**.

Out of this, 2 samples were collected after a further **cell cycle arrest** with **nocodazole** and release. (Nocodazole interferes with polymerization of microtubules.)

FACS profiles of individual HeLa populations

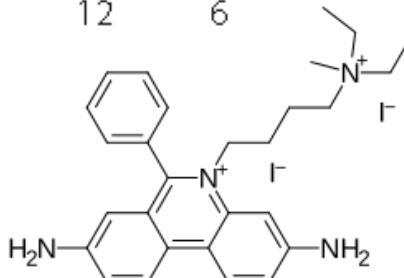
	% Cells		
	G ₁	S	G ₂ /M
1. Asynchronous	64	27	9
2. Thymidine block	50	46	4
3. Thymidine block + release 2½ h	36	60	4
4. Thymidine block + release 5½ h	23	70	7
5. Thymidine block + release 7½ h	15	70	15
6. Nocodazole block + release ½ h	1	11	88
7. Nocodazole block + release 3 h	82	12	6



Cells were fixed and collected by centrifugation.

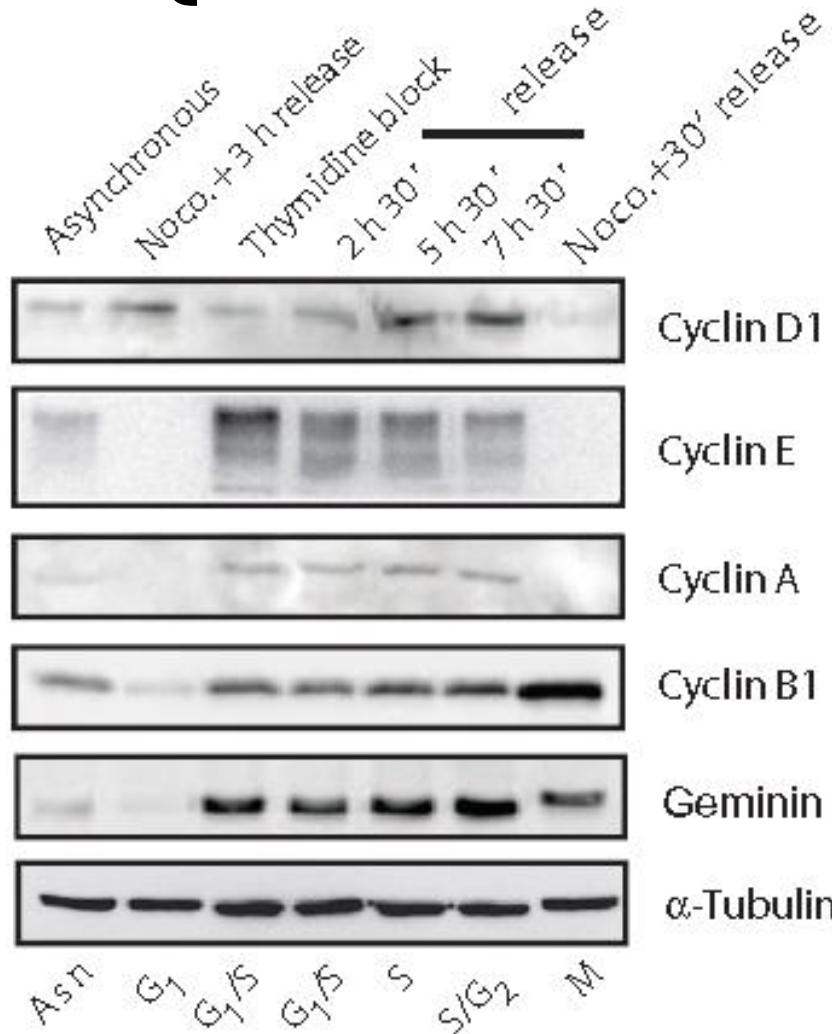
Then the **DNA content** of the cells was determined with propidium iodide.

The DNA content is the basis for classifying the **state** along the cell cycle.



→ Samples 1 – 5 are not pure states, but **mixtures**. Nocodazole block is quite efficient in synchronizing cells (samples 6 and 7).

Quantification of cell cycle markers

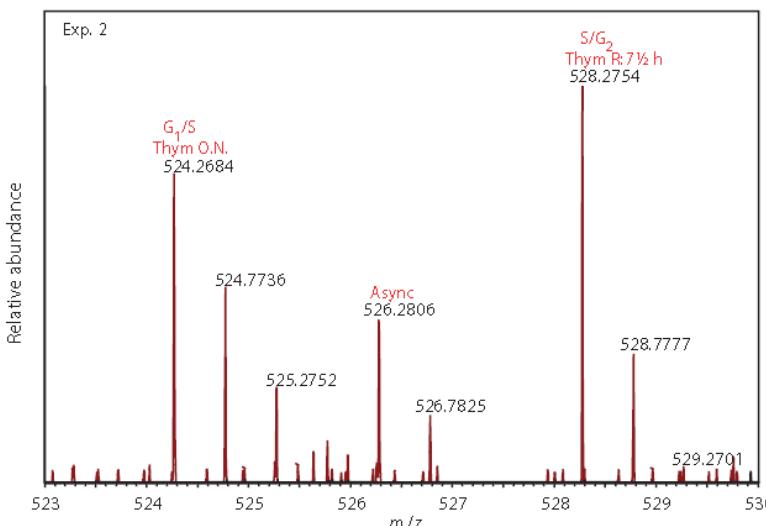
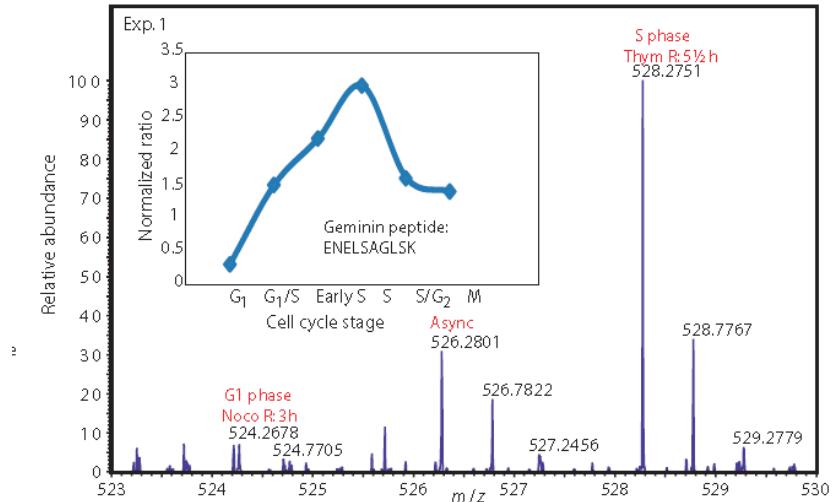
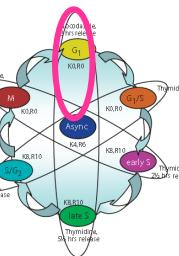


Immunoblot analysis of known cell cycle marker proteins in the different cell populations (α -tubulin is a control).

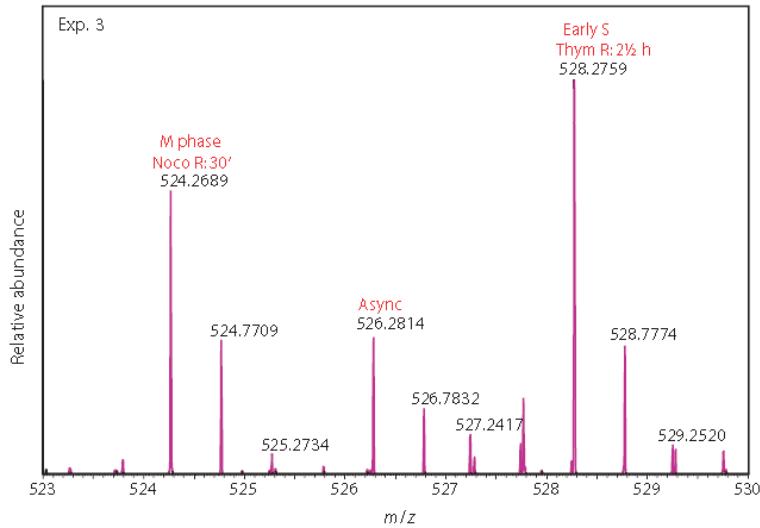
The abundance of 20% of the proteome changed by at least 4-fold throughout the cell cycle (difference between lowest and highest abundance).

Because a **fourfold change** best accounted for the dynamics of already described cell cycle components, this ratio was used as a threshold for subsequent analysis.

Experiment 1: mixture of
L = G1 phase
M = Async
H = S phase



Monitor protein abundance by MS

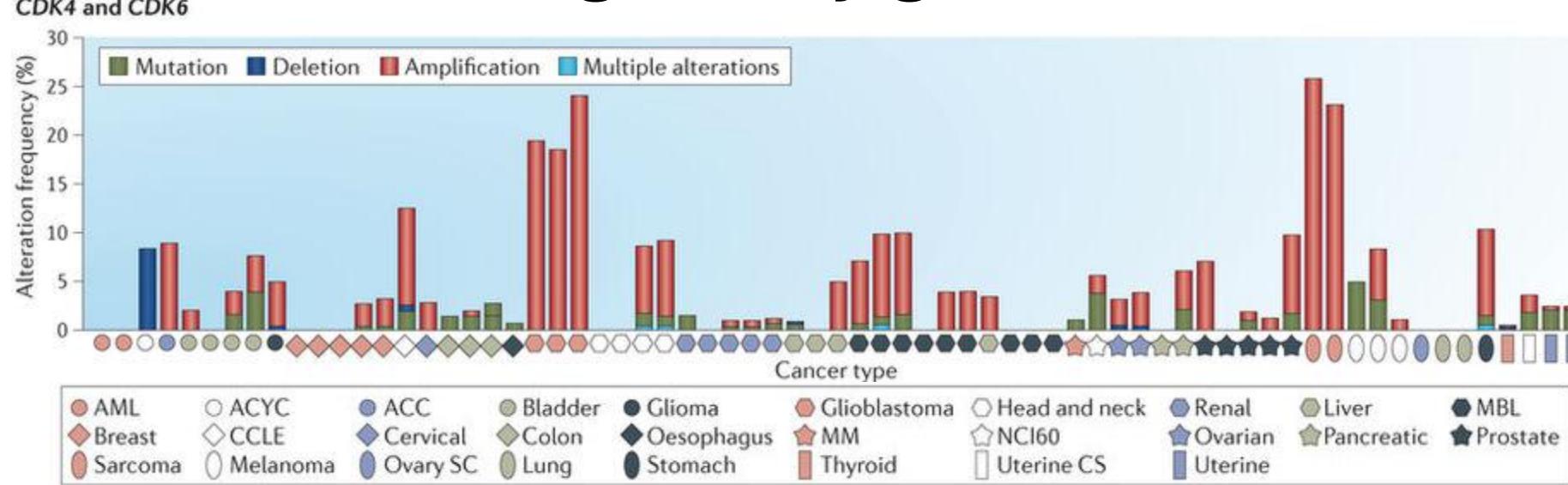


Representative MS data showing how the abundance of the proteins was monitored in 3 experiments to obtain information from the 6 stages of the cell cycle.

The data show the MS analysis of a tryptic SILAC peptide triplet derived from the cell cycle marker protein **Geminin**.

Relative peptide abundance changes were normalized to the medium SILAC peptide derived from the asynchronously grown cells in all three experiments.
The inset of Exp. 1 shows the combined six-time profile of Geminin over the cell cycle.

Deregulation of CDK regulatory genes in cancer.



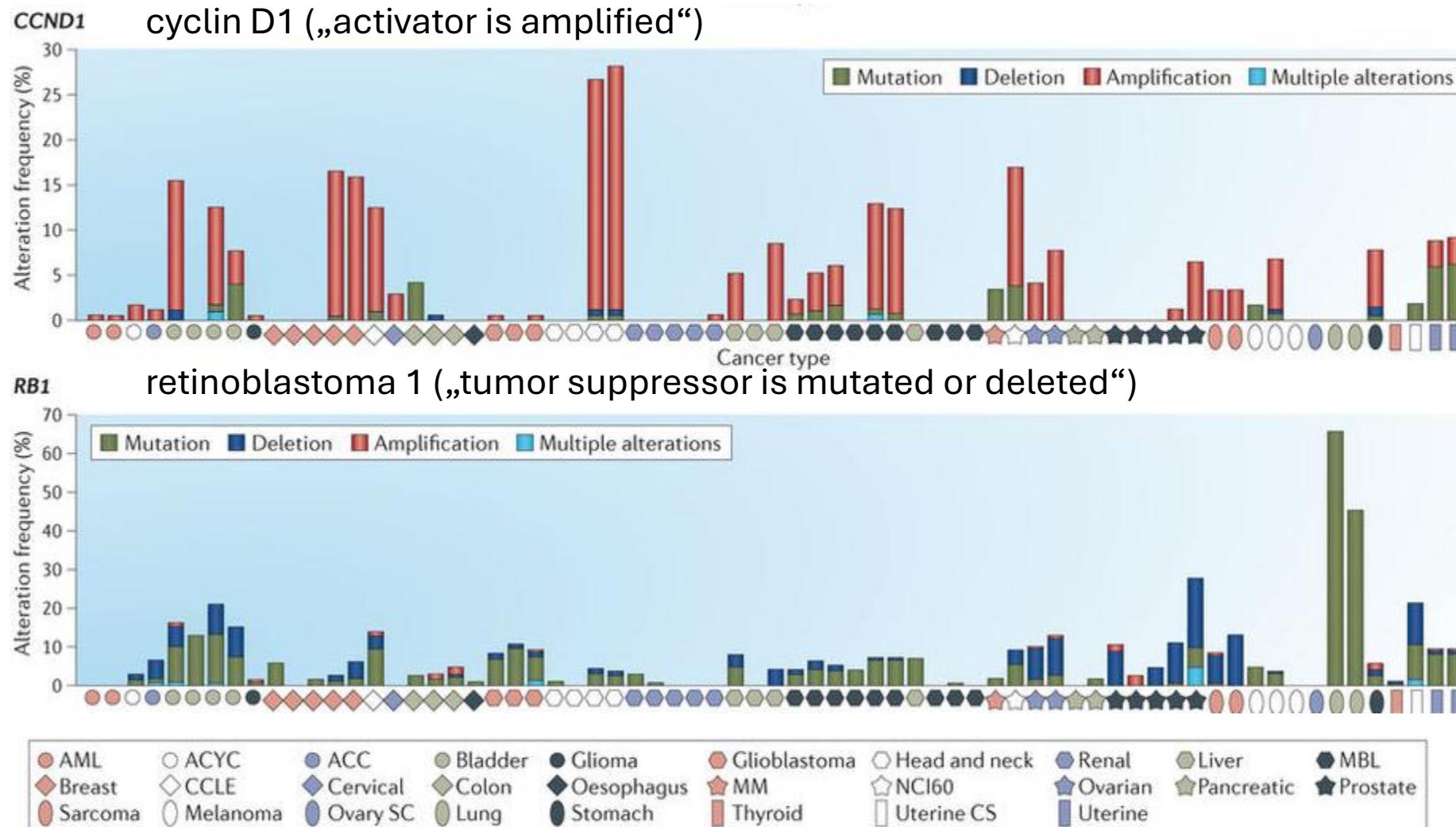
Frequencies of genetic amplification of *CDK4* and *CDK6* across multiple disease sites.

The frequencies of mutation (green), amplification (red) and homozygous deletion (dark blue) were determined using genetic data from >2,000 cancer cases.

Different types of cancer exhibit distinct predominant mechanisms of genetic alterations in cell cycle control.

<http://www.nature.com/articles/nrd4504>

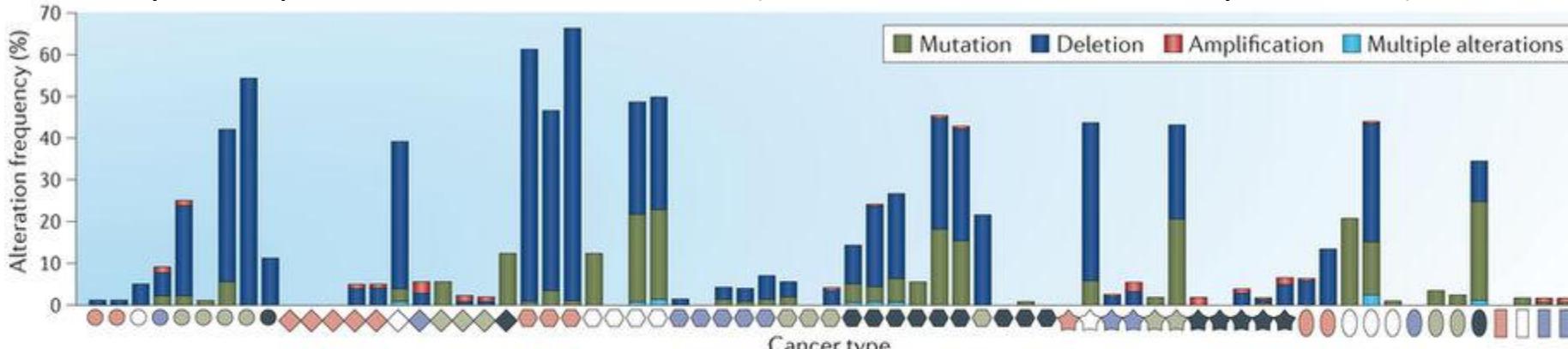
Deregulation of CDK regulatory genes in cancer.



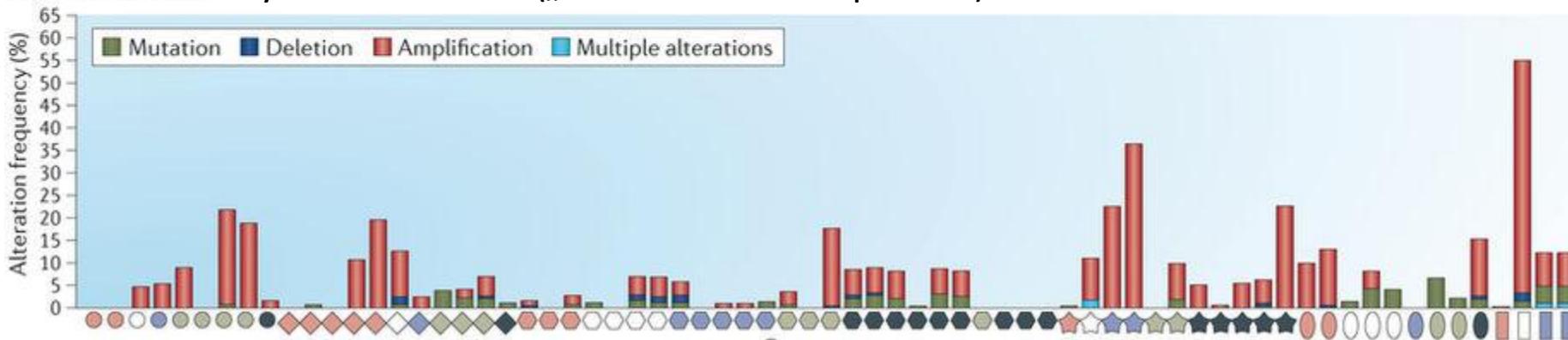
<http://www.nature.com/articles/nrd4504>

Deregulation of CDK regulatory genes in cancer.

CDKN2A cyclin-dependent kinase inhibitor 2A („inhibitors are shut down by deletion“)



CCNE1 and CCNE2 cyclins E1 and E2 („activators are amplified“)



AML	ACYC	ACC	Bladder	Glioma	Glioblastoma	Head and neck	Renal	Liver	MBL
Breast	CCLE	Cervical	Colon	Oesophagus	MM	NCI60	Ovarian	Pancreatic	Prostate
Sarcoma	Melanoma	Ovary SC	Lung	Stomach	Thyroid	Uterine CS	Uterine		

<http://www.nature.com/articles/nrd4504>

First generation of CDK inhibitors

Over the past 20 years, several small molecule inhibitors of CDKs have been developed as potential cancer therapeutics and tested in numerous trials and in several tumor types.

The first-generation CDK inhibitors developed were **relatively nonspecific** and may therefore be referred to as '**pan- CDK**' inhibitors.

Of these inhibitors, **flavopiridol** is the most extensively investigated CDK inhibitor so far, with **>60 clinical trials** carried out between 1998 and 2014.

Although flavopiridol can induce **cell cycle arrest** in G1 and G2 phases, in certain contexts it also induces a **cytotoxic response**.

Flavopiridol **did not meet the initial high expectations**. Low levels of clinical activity were seen in Phase II studies in several solid tumor types

Despite extensive investment, **no Phase III studies** have emerged and drug development of flavopiridol was discontinued in 2012.

<http://www.nature.com/articles/nrd4504>

Reasons for failure of broad-specificity CDK inhibitors

The general failure of non-selective CDK inhibitors in the clinic can be partly explained by at least 3 key underlying principles.

- (1) There was a lack of clear understanding of the **mechanism of action**. For many of the CDK inhibitors with low specificity, there remains a lack of clarity with regard to which CDKs are actually being inhibited *in vivo* and therefore the corresponding mechanism that could underlie the therapeutic effect.
- (2) There was a lack of **appropriate patient selection**. The vast majority of studies conducted with CDK inhibitors with low specificity were in unstratified patient cohorts. This is because there are essentially no **biomarkers** that may select for sensitive subpopulations for this class of inhibitors.
- (3) There is a lack of a **therapeutic window**. Many of these CDK inhibitors target several proteins that are critical to the **proliferation** (e.g. CDK1) and **survival** (e.g. CDK9) of **normal cells**. This limits the ability to achieve therapeutic levels of these drugs because of their intrinsic inability to discriminate between cancerous and healthy tissues.

<http://www.nature.com/articles/nrd4504>

V7: Cellular differentiation during development

In developmental biology, **cellular differentiation** is the process where a cell changes its **cell fate** from a **less differentiated** developmental stage to a **more specialized** cell stage.

Cell differentiation occurs numerous times during the development of a multicellular organism as it changes from a simple zygote to a complex system of tissues and cell types.

In nearly all multicellular organisms, differentiation continues in **adulthood** as **adult stem cells** divide and create fully differentiated daughter cells during tissue repair and during normal cell turnover.

Differentiation dramatically changes a cell's size, shape, membrane potential, metabolic activity, and responsiveness to signals.

These changes are largely due to highly controlled **modifications in gene expression** that are often controlled by **epigenetic** effects.

www.wikipedia.org

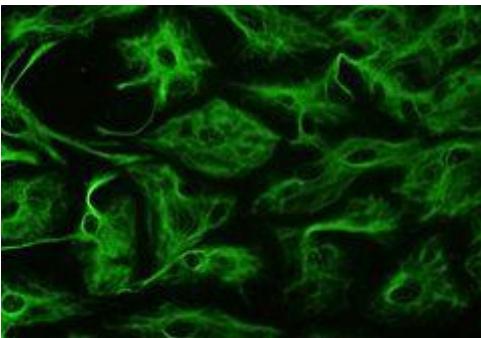
Different cell types

Complex genomes can generate a range of **different cell types** in a highly ordered and reproducible manner.

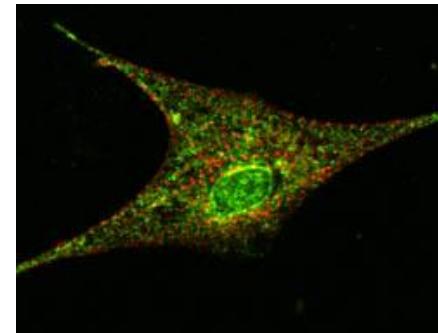
Transcriptional programs and **epigenetic modifications** are important for ‘programming’ lineage determination and cellular identity during development.



Fibroblast (connective tissue)
(wikipedia.org)



Astrocyte (nerve cell)
(wikipedia.org)



Cardiomyocyte (heart muscle)
(<http://www.kcl.ac.uk/content/1/c6/01/66/46/gautel3.jpeg>)

Cantone & Fisher,
Nature Struct Mol
Biol. 20, 292 (2013)

Model system *C. elegans*

The adult *Caenorhabditis elegans* hermaphrodite (dt. Zwitter) has 959 somatic cells (including 302 neurons); the male form has 1033 cells.



In 1963, **Sydney Brenner** proposed research into *C. elegans*, primarily in the area of neuronal development. He was a biologist from South Africa and won the 2002 noble prize in physiology or medicine for his work on *C. elegans*.

At the LMB / MRC Cambridge, Sydney Brenner also proposed the existence of messenger RNA (mRNA) and was involved in demonstrating the existence of the genetic code (nucleotide triplets are translated into amino acids)

An important aspect of *C. elegans* development is **apoptosis**, or programmed cell death, that leads to selective removal of certain cells. During the embryonic phase of worm development, 113 cells die as a result of apoptosis.



Sydney Brenner

www.wikipedia.org

Model system *C. elegans*

The *C. elegans* life cycle comprises of **four larval stages** — L1, L2, L3, L4 — which are followed by adulthood.

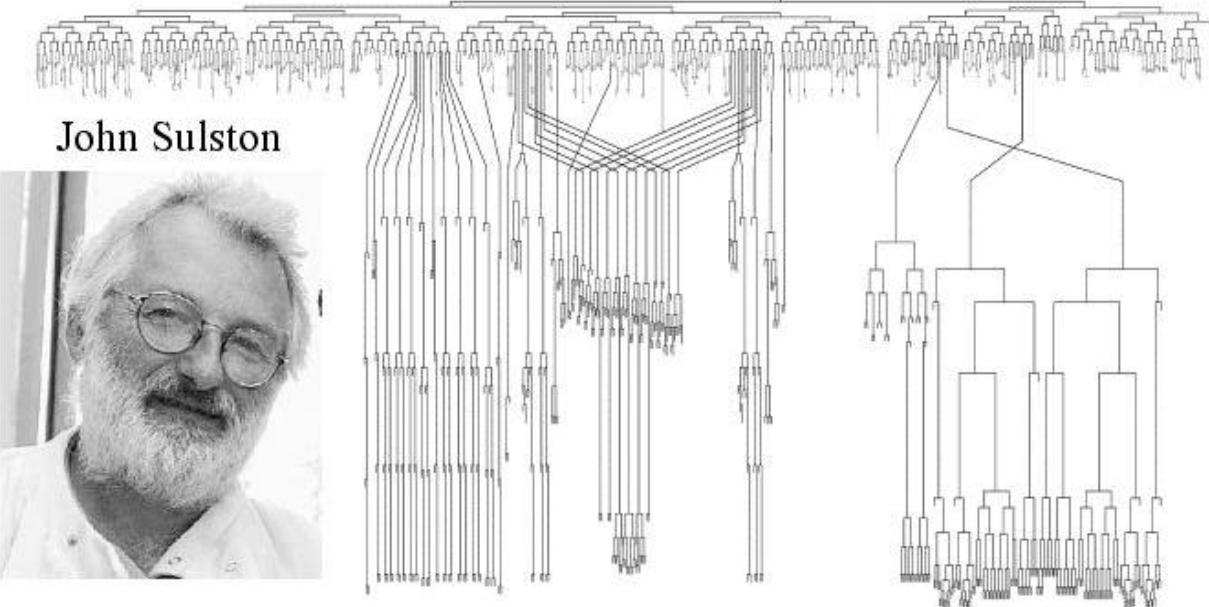
Under certain environmental conditions, such as scarcity of food, the late L1 or L2 larvae arrest and enter an alternative developmental program, called the **dauer stage**. The dauers can stay in this stage for many months, but upon availability of food they re-enter the normal.

The worms are **transparent**, allowing you to see the internal features of the worm.



<https://bastiani.biology.utah.edu/>

Cellular differentiation in *C. elegans*



Each worm consists of exactly 959 somatic cells.

John Sulston, the second of the 3 recipients of the Nobel Prize in Medicine in 2002 was able to identify the cell lineage of all 959 cells.

In the diagram, each vertical line represents a single cell.

Each horizontal line represents a cell division.

Any single cell will follow the same pattern of cell divisions in the worm, so the entire anatomy of the worm is known on a cellular level.

Because of this, circuit diagrams of the entire nervous system of the worm have been created in 2019.

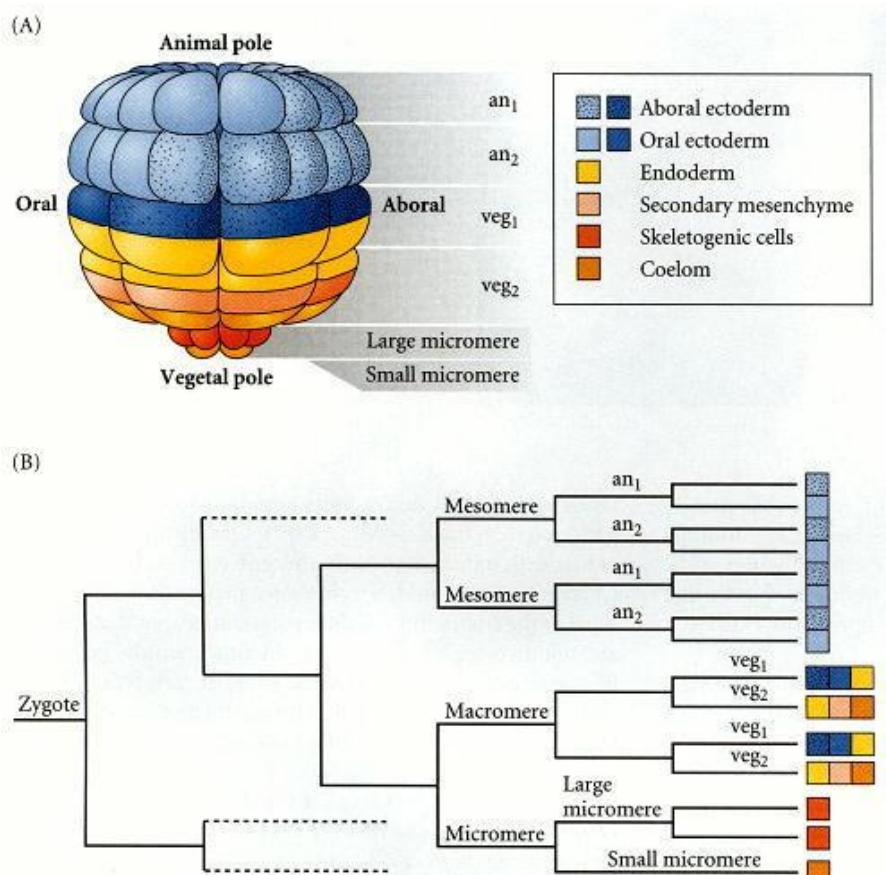
C. elegans was the first organism to have its connectome (neuronal "wiring diagram") completed.

<https://bastiani.biology.utah.edu/>



Model system Sea urchin

Fate map and cell lineage of the sea urchin *Strongylocentrotus purpuratus*, (dt. Seeigel) another model organism of developmental biology.



(A) 60-cell embryo. Blastomere fates are segregated along the animal-vegetal axis of the egg.

(B) Cell lineage map of the embryo. For simplicity, only one-quarter of the embryo is shown beyond second cleavage.

The veg₁ tier gives rise to both ectodermal and endodermal lineages, and the coelom comes from two sources: the second tier of micromeres, and some veg₂ cells. (After Wray 1999.)

Model system Sea urchin

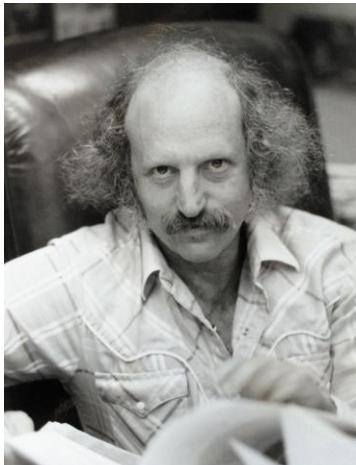
The initial discovery of three distinct eukaryotic DNA-dependent **RNA polymerases** was made using *S. purpuratus* as a model organism.

The full genome of the purple sea urchin was completely sequenced and annotated in 2006. One driver of this project was Eric H Davidson/Caltech (below).

The sea urchin genome is estimated to encode about 23,500 genes.

S. purpuratus has 353 protein kinases, containing members of 97% of human kinase subfamilies.

S. purpuratus's immune systems contains innate pathogen receptors like Toll-like receptors.

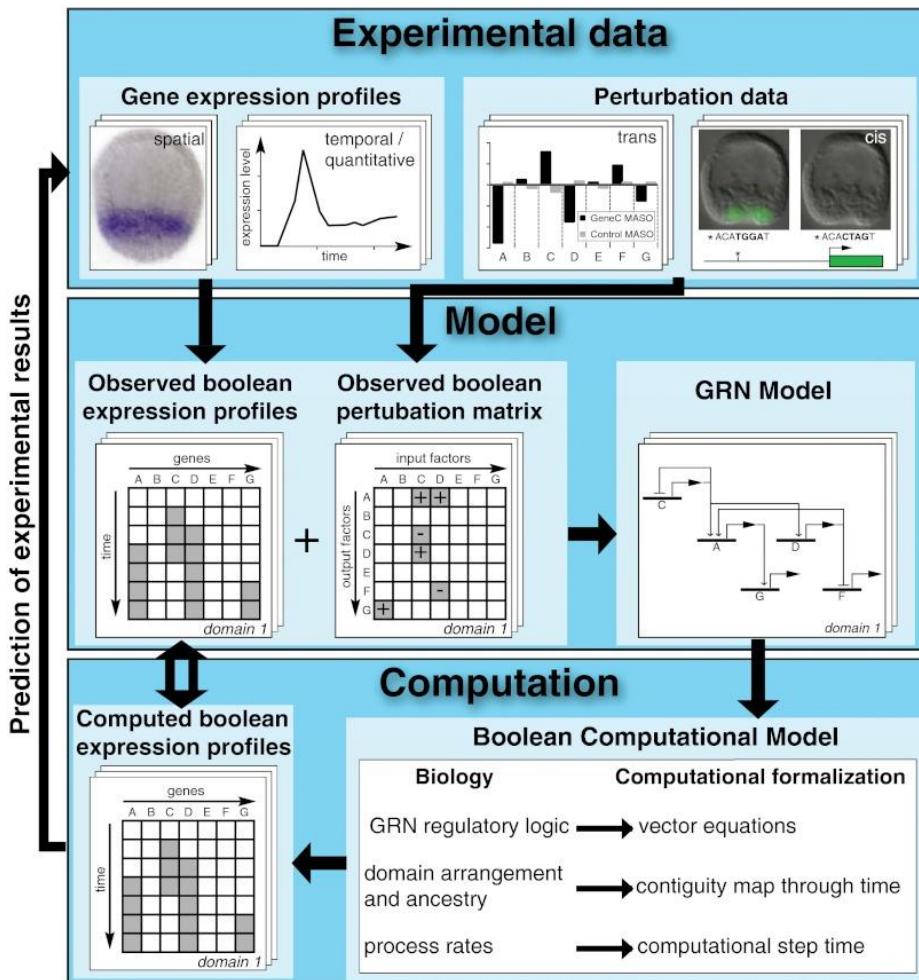


Eric H. Davidson
Caltech

www.wikipedia.org
www.nature.com

Gene regulatory networks (GRNs)

Eric H. Davidson was a pioneer in the establishment of GRNs.

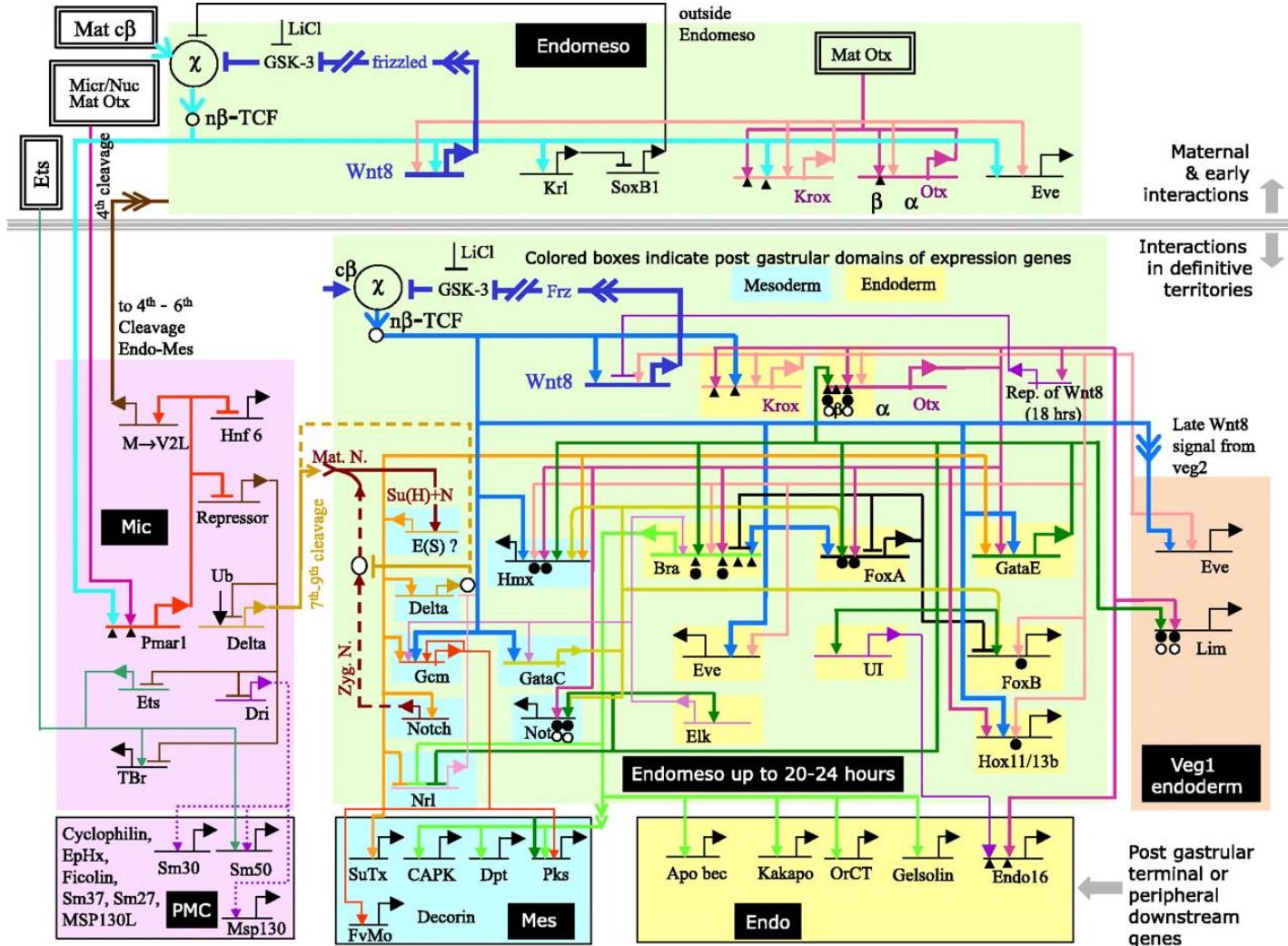


The figure illustrates the steps to derive a simple Boolean network (on/off) of gene activity.

The model incorporates essential aspects of the temporal and spatial biology of the embryo, and gives rise to a matrix of specific predictions of where, when, and for how long every gene is individually expressed over a 24-h period.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3478651/>

Gene regulatory networks (GRNs)



Endomesoderm
cis-regulatory
network for
development of
Strongylocentrotus
purpuratus.

Shown is which
transcription
factors activate
which target
genes.

[https://www.science.org/
doi/full/10.1126/science.
1069883](https://www.science.org/doi/full/10.1126/science.1069883)

Zygotes - fertilization

In living organisms that reproduce sexually, development starts from a single cell, the **zygote** (*dt: befruchtete Eizelle*).

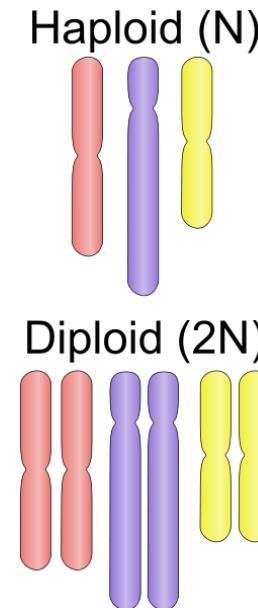
Zygotes are usually produced by a **fertilization** event between two **haploid** cells — an **ovum** from a female and a **sperm** cell from a male—which combine to form the single **diploid** cell.

Human sperm and egg (sex cells) have one complete set of chromosomes from the male or female parent.

Sex cells, also called **gametes**, combine to produce somatic cells. Somatic cells therefore have twice as many chromosomes.

In humans, gametes have 23 chromosomes.

Human **somatic cells** have 46 chromosomes.



some terms from developmental biology

somatic cells = cells forming the body of an organism

germ cells (dt. *Keimzelle, Ovulum*) are part of the germline.

germline (dt. *Keimbahn*) = line of germ cells that have genetic material that may be passed to a child/embryo. Germline cells are **immortal**.

Gametocyte = eukaryotic germ cell; includes spermatocytes (male) and oocytes (female)

primordial germ cells : predecessors of germ cells.

They migrate to the gonadal ridge (precursor of gonads).

They may be detected from expression of the marker gene/protein *Stella*.

gonad (dt. *Keimdrüse*)

www.wikipedia.org

How many eggs does a woman have?

Women are born with all the eggs they will ever produce.

During fetal development, a female embryo has about 6 million eggs.

At birth, there are approximately 1 million eggs left.

By the time girls reach puberty, only about 300,000 eggs remain.

The number of eggs women have continues to decline as they age and menstruate each cycle.

Fertility also declines with age due to the decreasing number and quality of your remaining eggs.

[https://my.clevelandclinic.org/health/
articles/9118-female-reproductive-system](https://my.clevelandclinic.org/health/articles/9118-female-reproductive-system)

Germline cells are produced by embryonic cleavage.

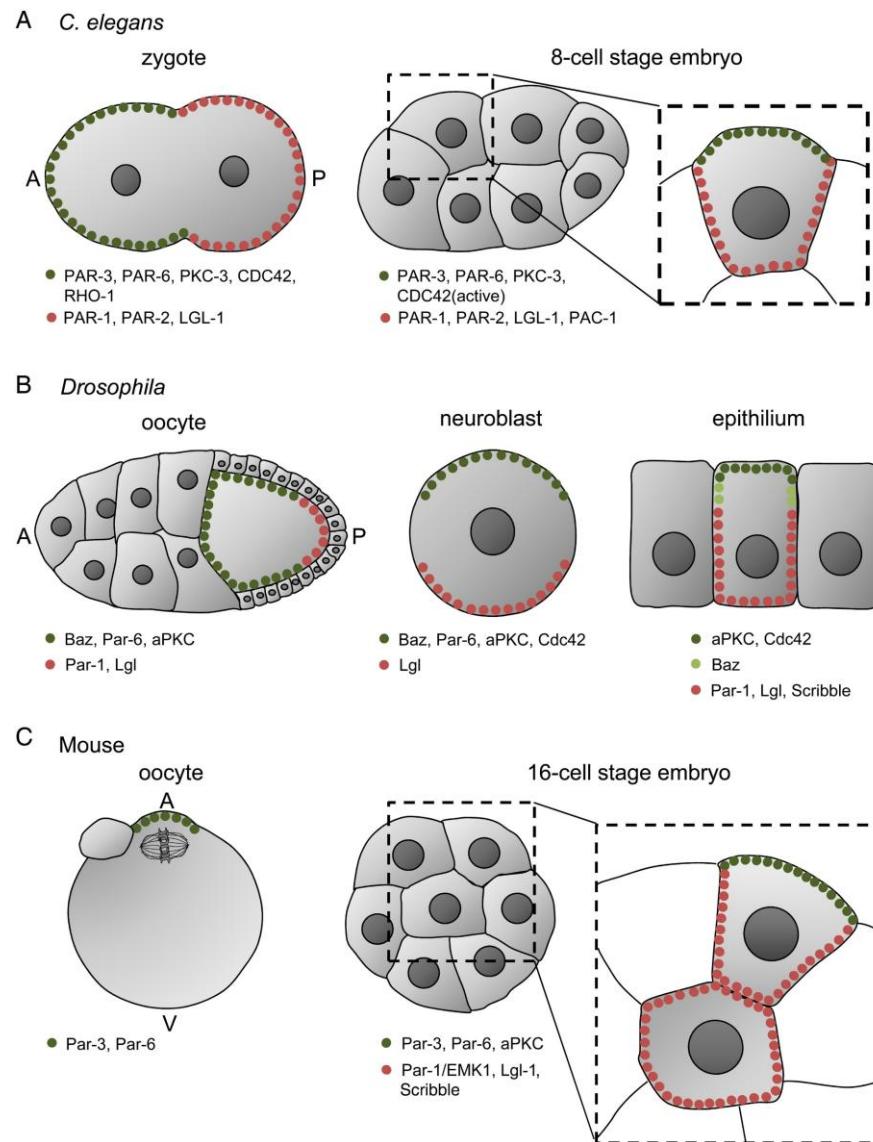
Cleavage: division of cells in the early embryo.

The **zygotes** of many species undergo rapid cell cycles with no significant growth.

Establishment of **cellular polarity** is one of the most important events during early embryonic divisions. In most species, including mammals, it enables cells to adopt distinct developmental fates.

Polarity in *C. elegans*, *Drosophila* and mouse oocytes and embryos. Polarized distribution of PAR proteins and accompanying factors in *C. elegans* zygote and 8-cell stage embryo (**A**), *Drosophila* oocyte, neuroblast and epithelium (**B**), and mouse oocyte and 16-cell stage embryo (**C**). In (A) and (B): A, anterior pole; P, posterior pole. In (C): A, animal pole; V, vegetal pole.

Germ line development



Ajduk et al. *Mol Hum Reprod.*,
22, 691–703 (2016)

Germ line development

The different cells derived from cleavage are called **blastomeres** and form a compact mass called the **morula** (because it resembles a mulberry/ *dt. Maulbeere*).

Cleavage ends with the formation of the **blastula**.

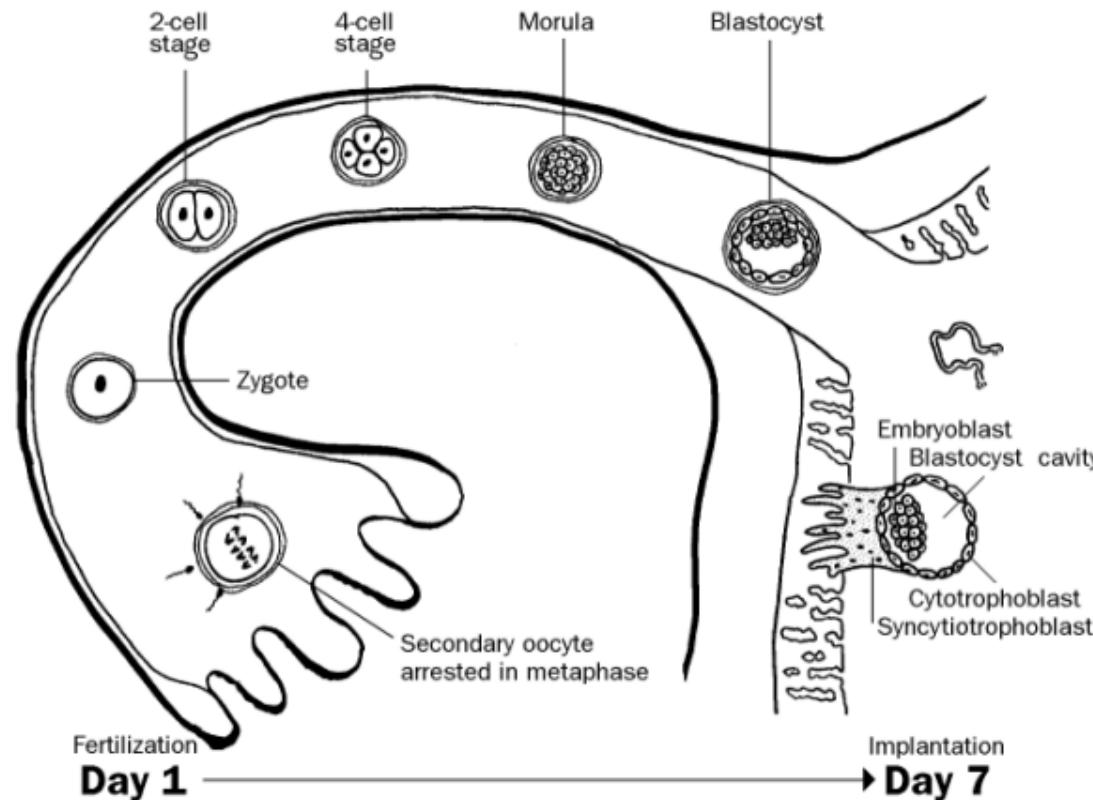
Cleavage in mammals is slow.

Cell division takes 12 – 24 hours and is asynchronous.

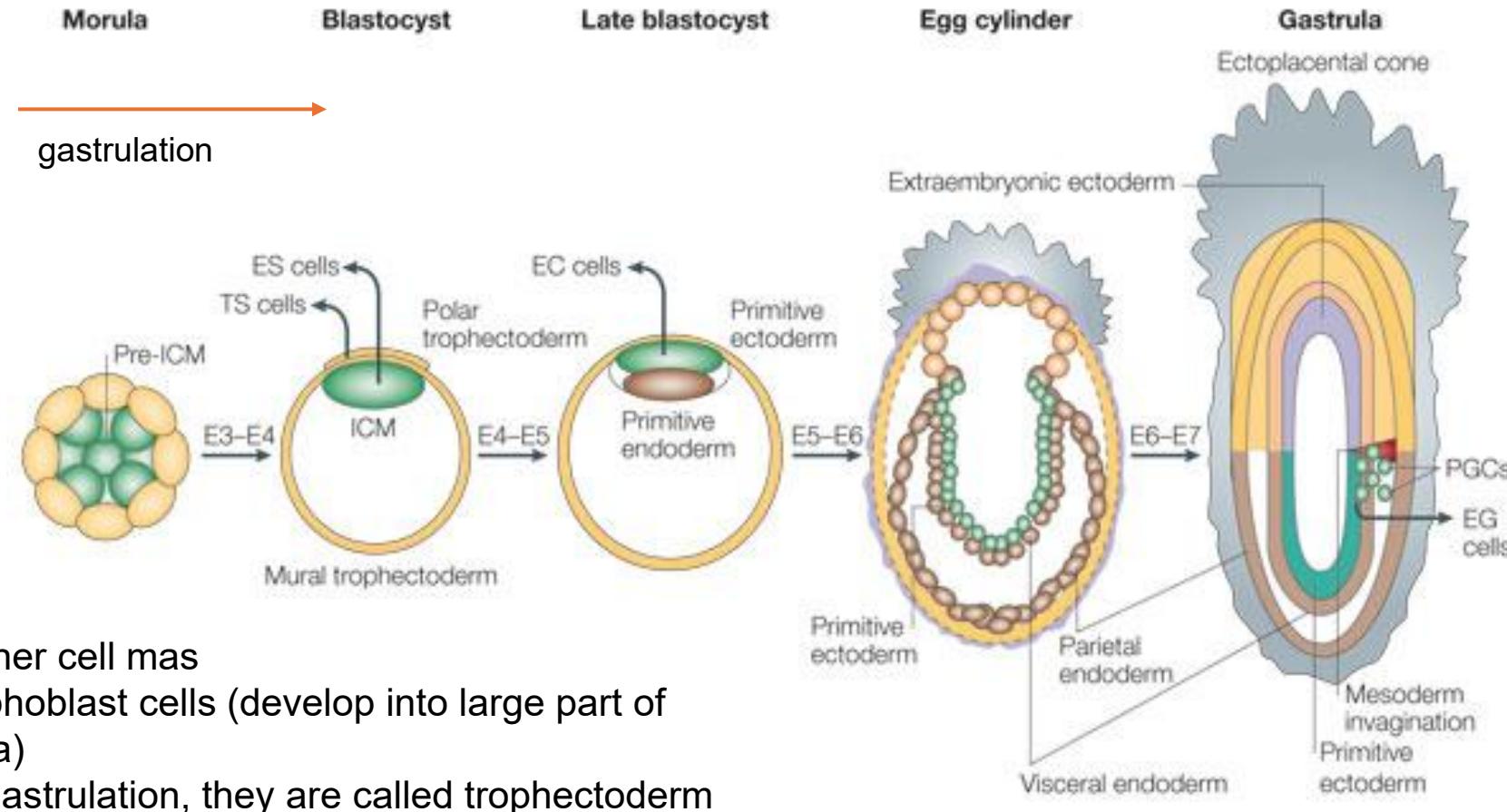
Figure shows human fertilization -> implantation. Implantation in the placenta is needed to ensure that the embryo can grow further.

www.wikipedia.org

<https://www.med.umich.edu/lrc/coursepages/m1/embryology/embryo/03firstweek.htm>



Embryonic development of mouse



ICM: Inner cell mass

TS: trophoblast cells (develop into large part of placenta)

- After gastrulation, they are called trophectoderm

PGCs: primordial germ cells (progenitors of germ cells)

E3: embryonic day 3

Copyright © 2005 Nature Publishing Group
Nature Reviews | Molecular Cell Biology

Boiani & Schöler, Nat Rev Mol Cell Biol 6, 872 (2005)

3 primary germ cell layers

The **ectoderm** is the **outer layer** of the early embryo.

It emerges first and forms from the outer layer of germ cells.

The ectoderm differentiates to form the nervous system (spine,

peripheral nerves and brain), tooth enamel and the epidermis.

It also forms the lining of mouth, anus, nostrils, sweat glands, hair and nails.

The **endoderm** develops at the **inner layer**.

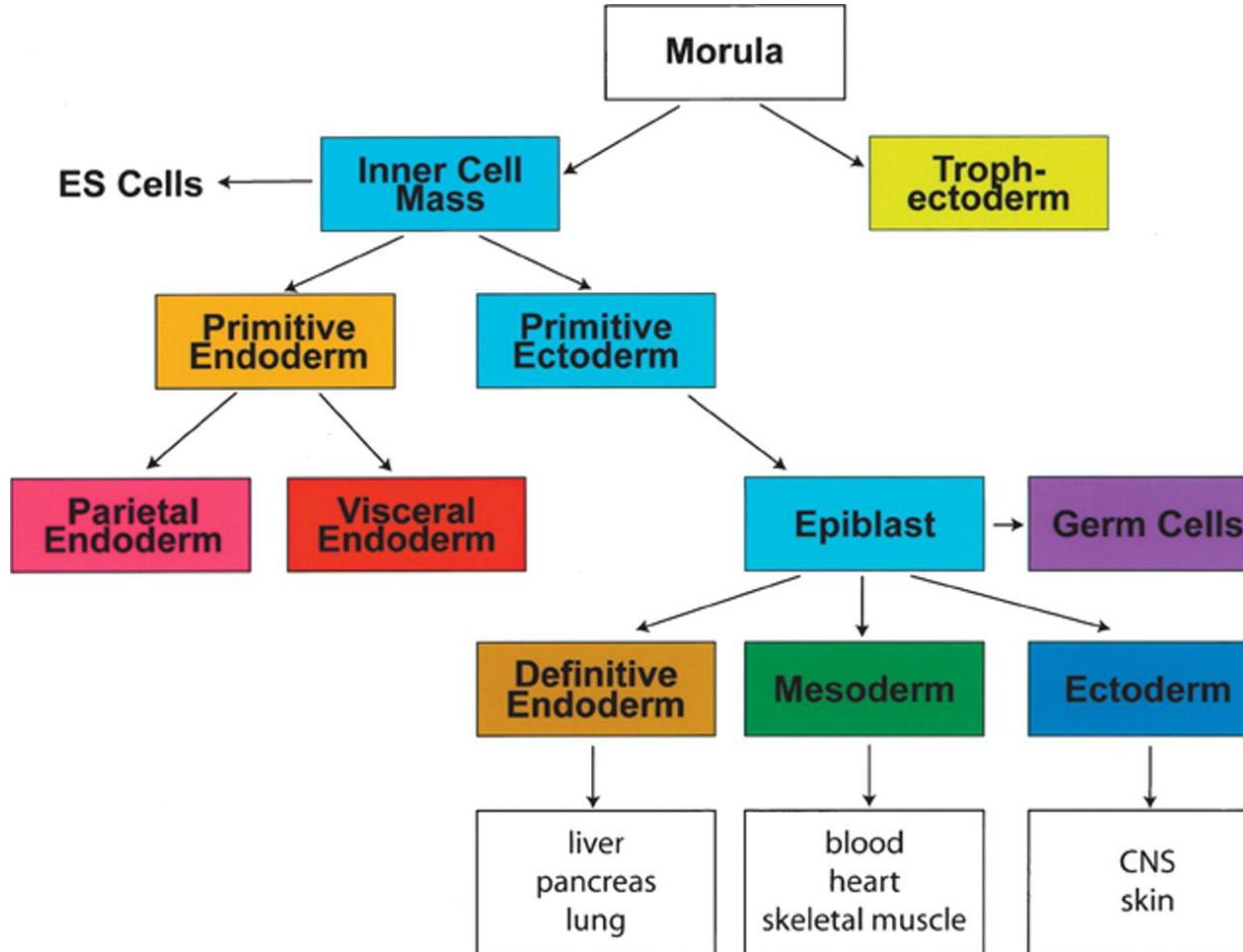
Its cells differentiate to form the gastrointestinal tract, the respiratory tract, endocrine glands and organs, auditory systems, and the urinary system.

The **mesoderm** is the **middle layer**.

It differentiates to give rise to a number of tissues and structures including bone, cartilage (*dt: Knorpel*), muscle, connective tissue (including that of the dermis), the middle layer of the skin, blood vascular, reproductive, excretory and urinogenital systems and contributes to some glands.

www.wikipedia.org

Cell populations in early mouse development



Scheme of **early mouse development** depicting the relationship of early cell populations to the primary germ layers

Keller, Genes & Dev.
(2005) 19: 1129-1155

Developmental Glossary (I)

Inner cell mass (ICM): Cells of the blastocyst embryo that appear transiently during development and give rise to the three germ layers of the developing embryo.

Embryonic stem (ES) cells:

Pluripotent cell line derived from the ICM upon explantation in culture.

In vitro, ES cells can differentiate into many different lineages and cell types.

Upon injection into blastocysts, ES cells can give rise to all tissues including the germline.

Primordial germ cells (PGCs):

In vivo, PGCs give rise to oocytes and sperm.

When explanted in vitro, PGCs give rise to embryonic germ (EG) cells.

Hochedlinger, Development 136, 509 (2009)

Adult stem cells

Embryonic stem cells only exist in the early embryo.

But we all possess **adult stem cells**, from which new specialized cells are formed throughout our life time.

Adult stem cells exist predominantly in **bone marrow** (*dt. Knochenmark*), but also in skin, fat tissue, umbilical cord (*dt. Nabelschnur*), brain, liver, and in pancreas (*dt. Bauchspeicheldrüse*).

Adult stem cells in cell culture have a much **reduced ability of self regeneration** and a reduced ability for differentiation compared to embryonic stem cells.

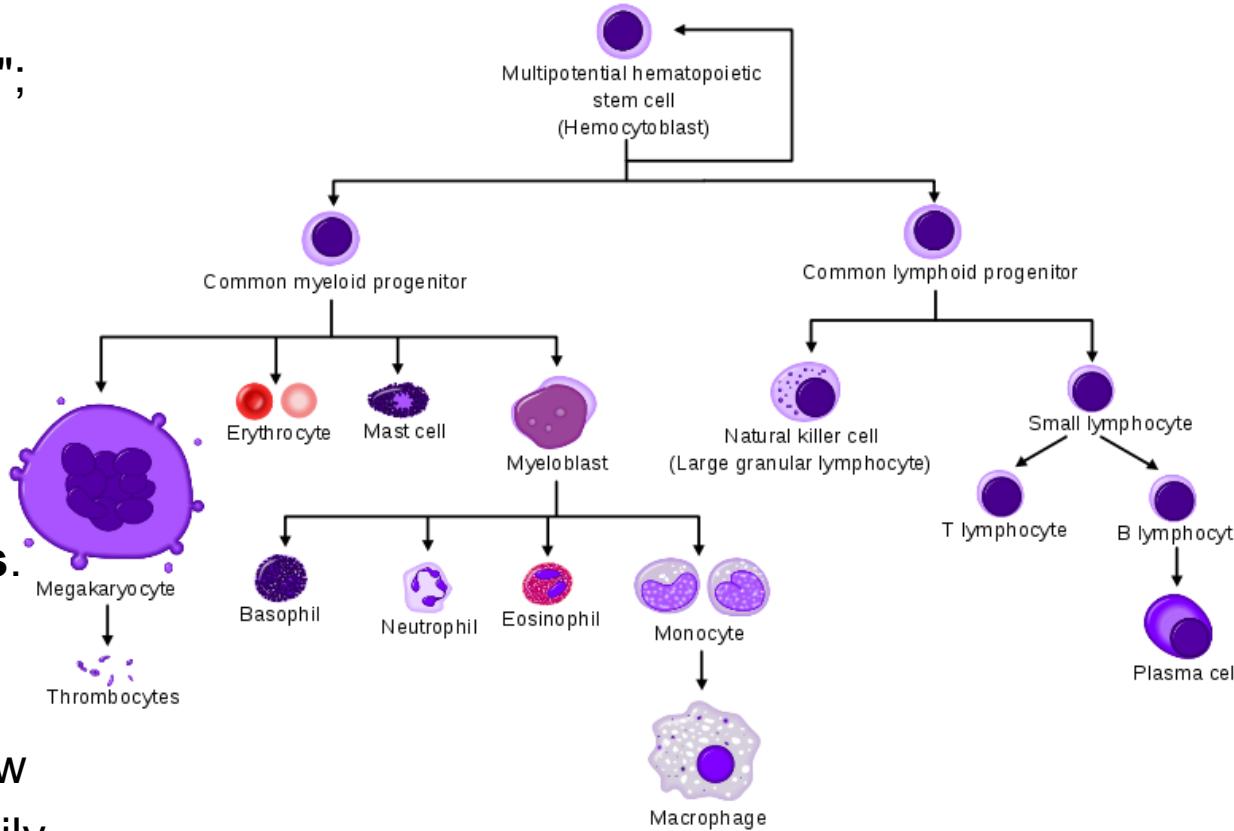
For example, neural stem cells can differentiate to all cell types of neural tissue (neurons, glia), but likely not into liver or muscle cells.

Haematopoiesis

Haematopoiesis (from Ancient Greek: αἷμα, "blood"; ποιεῖν "to make") is the formation of blood cellular components.

All cellular blood components are derived from **haematopoietic stem cells**.

In a healthy adult person, approximately 10^{11} – 10^{12} new blood cells are produced daily in order to maintain steady state levels in the peripheral circulation.



Development of different blood cells from haematopoietic stem cell to mature cells

Differentiation

(Review) A **zygote** is a eukaryotic cell formed by a fertilization event between two gametes.

Zygotes therefore contain DNA derived from both the mother and the father, and this provides all the genetic information necessary to form a new individual.

This property is named „**totipotency**“
(latin: totus – all, potentia – power/ability).

Continuous cell division produces daughter cells that start to specialize on individual functions.

This developmental process of cells and tissue from a less specialized to a more specialized state is called **differentiation** in developmental biology.

www.wikipedia.org

Glossary I

Totipotency Ability of a cell to give rise to all cells of an organism, including embryonic and extraembryonic tissues. Zygotes are totipotent.

Pluripotency Ability of a cell to give rise to all cells of the embryo. Cells of the inner cell mass (ICM) and its derivative, embryonic stem (ES) cells, are pluripotent.

Multipotency Ability of a cell to give rise to different cell types of a given cell lineage. These cells include most adult stem cells, such as gut stem cells, skin stem cells, hematopoietic stem cells and neural stem cells.

Unipotency Capacity of a cell to sustain only one cell type or cell lineage. Examples are terminally differentiated cells, certain adult stem cells (testis stem cells) and committed progenitors (erythroblasts).

Hochedlinger, Development 136, 509 (2009)

Types of body cells

3 basic categories of cells make up the mammalian body:

germ cells (oocytes and sperm cells)

somatic cells, and

stem cells.

Each of the approximately 100 trillion (10^{14}) cells in an adult human has its own copy or copies of the genome except certain cell types, such as red blood cells, that lack nuclei in their fully differentiated state.

Most cells are **diploid**; they have two copies of each chromosome.

Cells differentiate to specialize for different functions.

Somatic cells make up most of the human body, such as skin and muscle cells.

www.wikipedia.org

Development controlled by transcriptional programs

Embryonic development is a complex process that remains to be fully understood despite knowledge of the complete genome sequences of many species and rapid advances in genomic technologies.

Complication: experiments on human embryos are unethical.

Mouse development is similar to human in many aspects, but not identical.

A fundamental question is how the unique gene expression pattern in each cell type is established and maintained during embryogenesis.

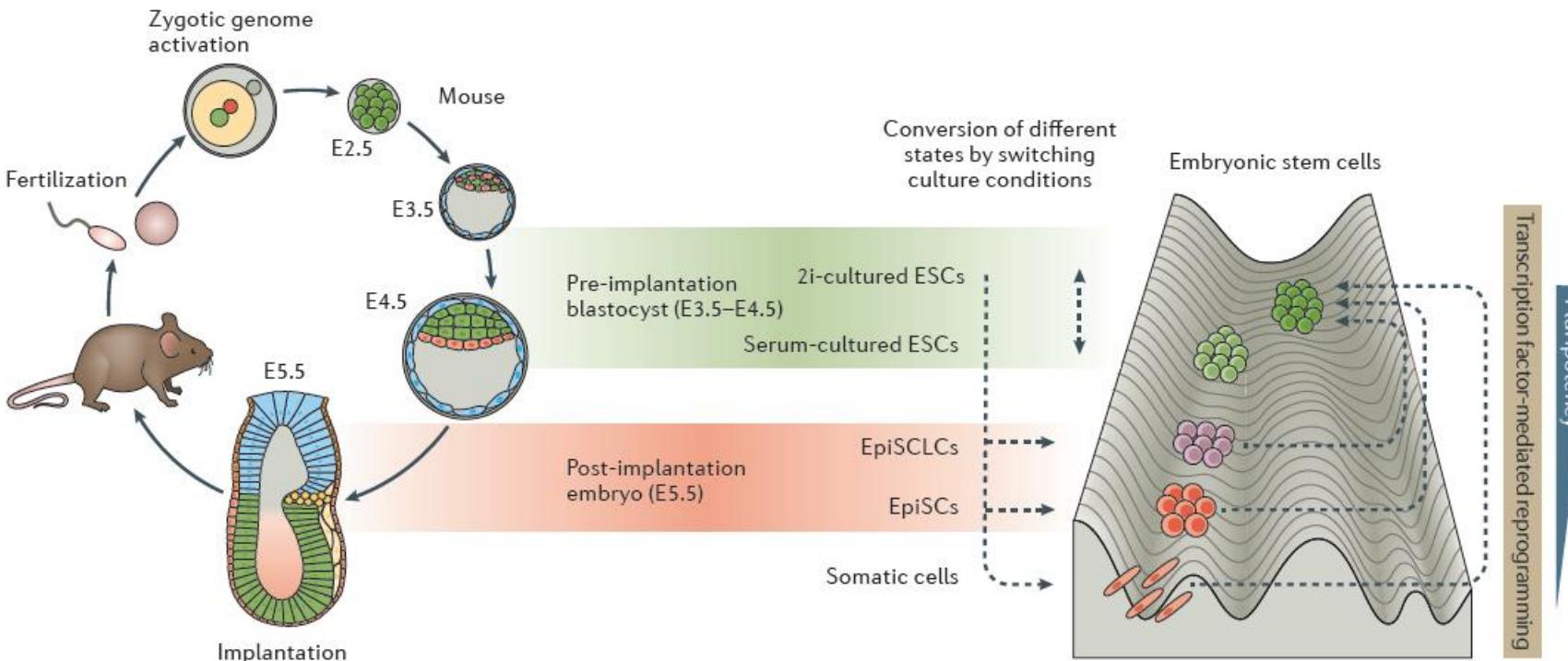
It is well accepted that the gene expression program encoded in the genome is executed by **transcription factors** that bind to cis-regulatory sequences and modulate gene expression in response to **environmental cues**.

V8: Cellular differentiation - Epigenetics

E4.5 epiblast cells: represent ground-state pluripotency

Implantation: stage of pregnancy at which the blastocyst adheres to the wall of the **uterus**.

After implantation (E5.5): **epiblast cells** undergo a strong wave of epigenetic reprogramming. They are now „primed“.



Atlasi & Stunnenberg, *Nature Rev Genet* **18**, 643–658 (2017)

Epigenetic mechanisms

Epigenetics refers to **alternate phenotypic states** that are **not based on differences in genotype**, and are potentially reversible, but are generally stably maintained during cell division.

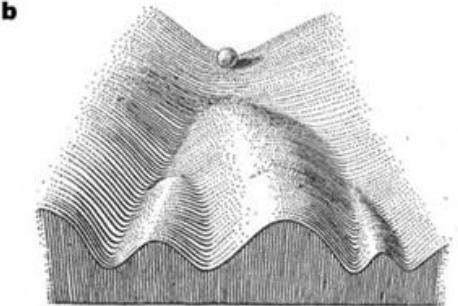
Examples: imprinting, twins, cancer vs. normal cells, differentiation, ...

Multiple mechanisms interact to collectively establish

- alternate states of chromatin structure (open – packed/condensed),
- **histone modifications**,
- composition of associated proteins (e.g. histones),
- transcriptional activity,
- activity of microRNAs, and
- in mammals, **cytosine-5 DNA methylation** at CpG dinucleotides.

Laird, Hum Mol Gen 14, R65 (2005)

Waddington's epigenetic landscape for embryology



Slack, Nature Rev Genet 3, 889-895 (2002)

Waddington worked in **embryology**

a) is a painting by John Piper that was used as the frontispiece for Waddington's book *Organisers and Genes*. It represents an **epigenetic landscape**.

Developmental pathways that could be taken by each cell of the embryo are metaphorically represented by the path taken by water as it flows down the valleys.

b) Later depiction of the epigenetic landscape. The ball represents a cell, and the bifurcating system of valleys represents bundles of trajectories in state space.

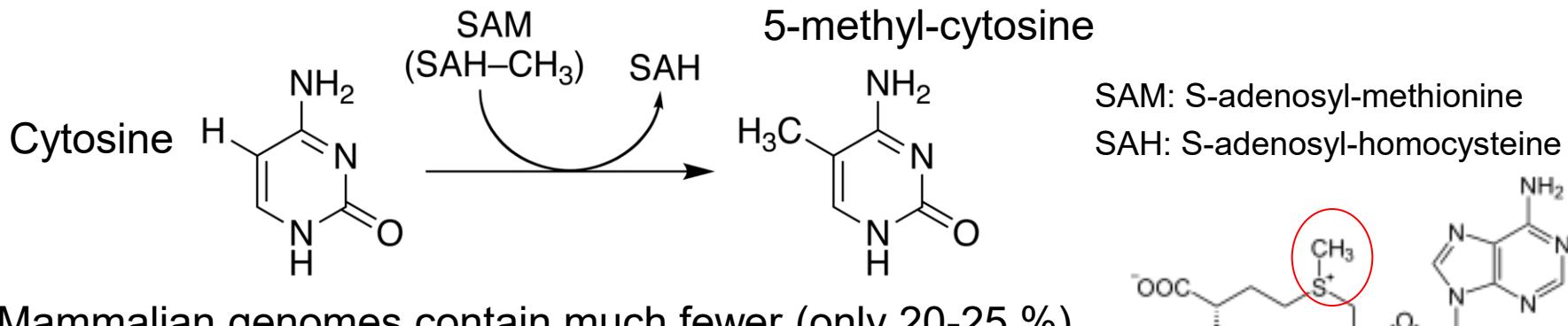


Conrad Hal Waddington
(1905 – 1975)
pictures.royalsociety.org

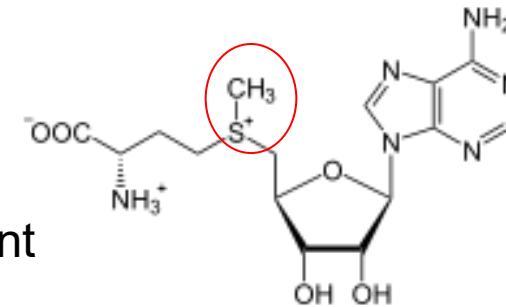
Cytosine methylation

Observation: 3-6 % of all cytosines are methylated in human DNA.

This methylation occurs (almost) exclusively when cytosine is followed by a guanine base -> **CpG dinucleotide**.



Mammalian genomes contain much fewer (only 20-25 %) of the CpG dinucleotide than is expected by the G+C content (we expect $1/16 \approx 6\%$ for any random dinucleotide).



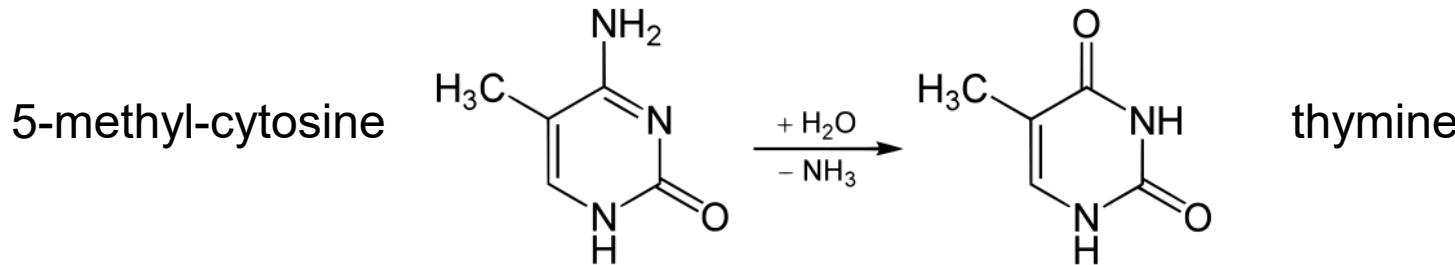
This is typically explained in the following way:

As most CpGs serve as targets of DNA methyltransferases, they are usually methylated (see following page)

Esteller, Nat. Rev. Gen. 8, 286 (2007)
www.wikipedia.org

Cytosine methylation

But 5-Methylcytosine can easily **deaminate to thymine**.



If this mutation is not repaired, the affected CpG is permanently converted to TpG (or CpA if the transition occurs on the reverse DNA strand).

Hence, methylCpGs represent **mutational hot spots** in the genome.

If such mutations occur in the germ line, they become heritable.

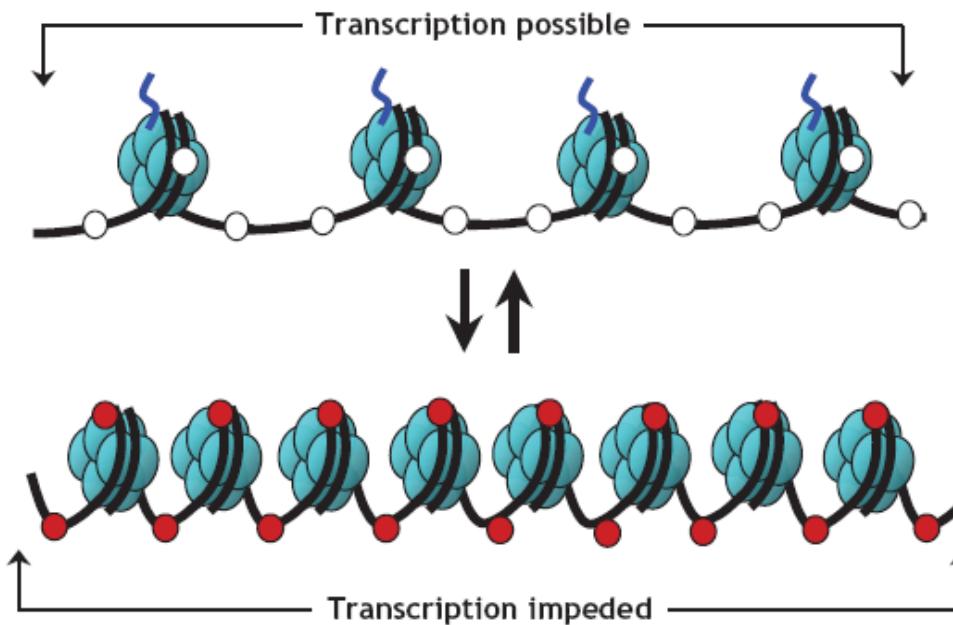
A constant loss of CpGs over thousands of generations
can explain the low frequency of this
special dinucleotide in the genomes of human and mouse.

Esteller, Nat. Rev. Gen. 8, 286 (2007)
www.wikipedia.org

chromatin organization affects gene expression

B

- Gene "switched on"
- Active (open) chromatin
 - Unmethylated cytosines (white circles)
 - Acetylated histones



- Gene "switched off"
- Silent (condensed) chromatin
 - Methylated cytosines (red circles)
 - Deacetylated histones

Schematic of the reversible changes in chromatin organization that influence gene expression:

genes are expressed (switched on) when the chromatin is **open** (active), and they are inactivated (switched off) when the chromatin is **condensed** (silent).

White circles = unmethylated cytosines;

red circles = methylated cytosines.

Rodenhiser, Mann, CMAJ 174, 341 (2006)

Altered DNA methylation upon cancerogenesis

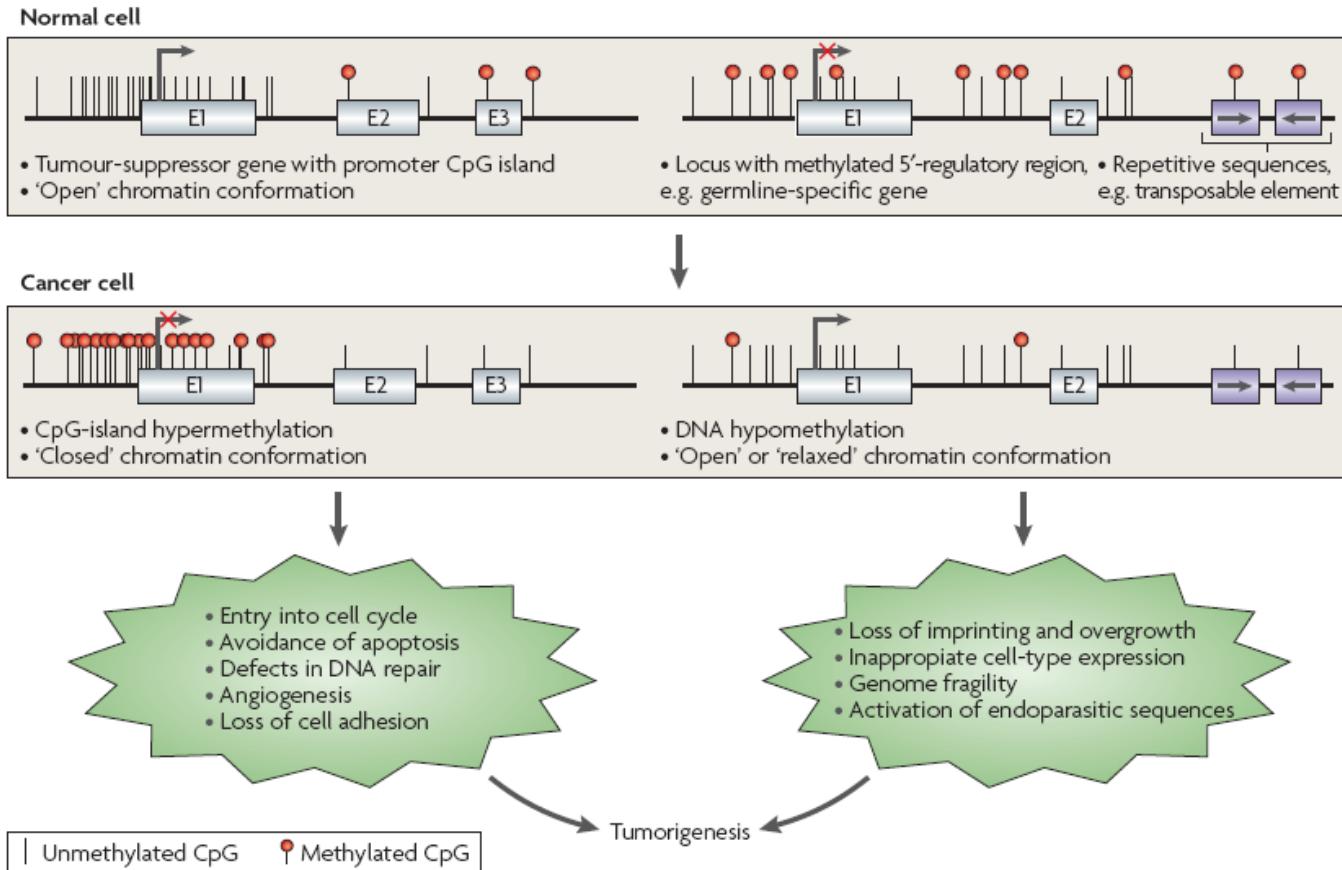


Figure 1 | Altered DNA-methylation patterns in tumorigenesis. The hypermethylation of CpG islands of tumour-suppressor genes is a common alteration in cancer cells, and leads to the transcriptional inactivation of these genes and the loss of their normal cellular functions. This contributes to many of the hallmarks of cancer cells. At the same time, the genome of the cancer cell undergoes global hypomethylation at repetitive sequences, and tissue-specific and imprinted genes can also show loss of DNA methylation. In some cases, this hypomethylation is known to contribute to cancer cell phenotypes, causing changes such as loss of imprinting, and might also contribute to the genomic instability that characterizes tumours. E, exon.

Esteller, Nat. Rev. Gen. 8, 286 (2007)

Genomic Imprinting:
Mono-allelic expression; one allele (either from the mother or the father) is silenced.

Typically, this is implemented by methylating the silenced allele.

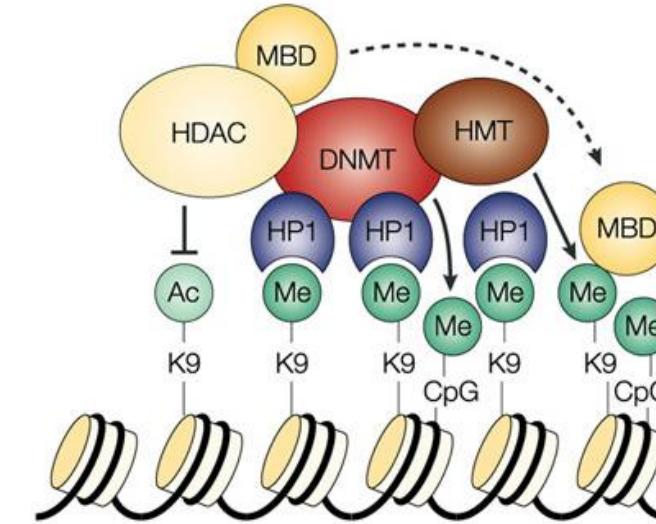
The human genome contains ca. 8% of **retroviral sequences**. Typically, these are also silenced by DNA methylation.

Enzymes that control DNA methylation and histone modifications

These dynamic chromatin states are controlled by reversible epigenetic patterns of **DNA methylation and histone modifications.**

Enzymes involved in this process include

- DNA methyltransferases (DNMTs),
- histone acetylases,
- histone deacetylases (HDACs),
- histone methyltransferases (HMT) and the
- methyl-binding domain protein MECP2 with its methyl-binding domain (MBD) that binds specifically to me-cytosine.



HP1: heterochromatin protein 1

Rodenhiser, Mann, CMAJ 174, 341 (2006)
Feinberg AP & Tycko P (2004) Nature Reviews: 143-153

DNA methylation

Typically, unmethylated clusters of CpG pairs are located in **tissue-specific genes** and in essential **housekeeping genes**.

(House-keeping genes are involved in routine maintenance roles and are expressed in most tissues.)

These clusters, or **CpG islands**, are targets for proteins that bind to unmethylated CpGs and initiate gene transcription.

In contrast, **methylated CpGs** are generally associated with silent DNA, can block methylation-sensitive proteins and can be easily mutated.

The **loss** of normal DNA methylation patterns is the best understood epigenetic cause of **disease**.

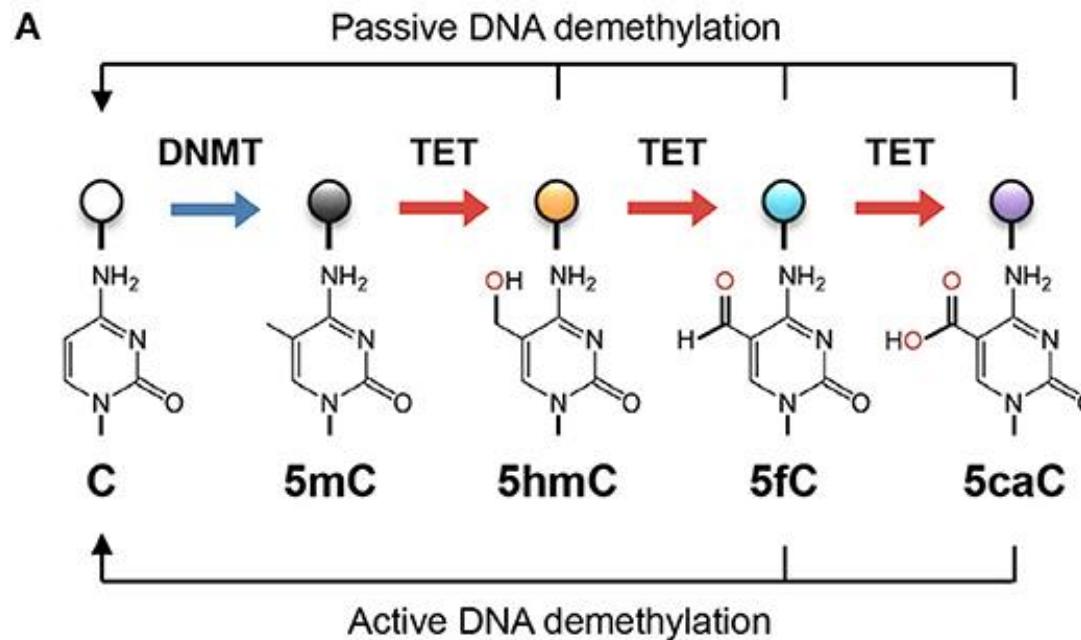
In animal experiments, the removal of genes that encode DNMTs is lethal; in humans, overexpression of these enzymes has been linked to a variety of cancers.

Rodenhiser, Mann, CMAJ 174, 341 (2006)

Higher forms of methylation – Tet enzymes

Unmodified cytosine (C) is methylated by DNA methyltransferases (DNMTs) at the 5 position to become 5-methylcytosine (5mC).

TET proteins oxidize 5mC into 5-hydroxymethylcytosine (5hmC), a stable epigenetic mark, and subsequently to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC).



TET can demethylate DNA via replication-dependent (passive) or replication-independent (active) mechanisms.

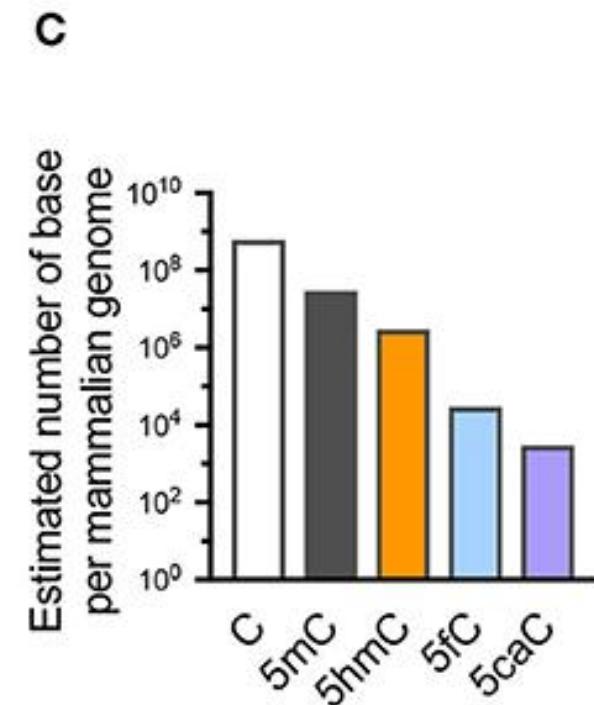
Lio & Rao, Front. Immunol. (2019)

Higher forms of methylation – abundance

The approximate abundance of unmodified and modified cytosines in the haploid human/mouse genome.

About 5% of cytosine is methylated (5mC); in most cells, the vast majority of 5mC is present at CG dinucleotides although it is low at CpG islands.

5hmC amounts to about 1-10% of 5mC (estimated at 10% in embryonic stem cells), while the levels of 5fC and 5caC are each about an order of magnitude lower than the previous oxidative modification.



Lio & Rao, Front. Immunol. (2019)

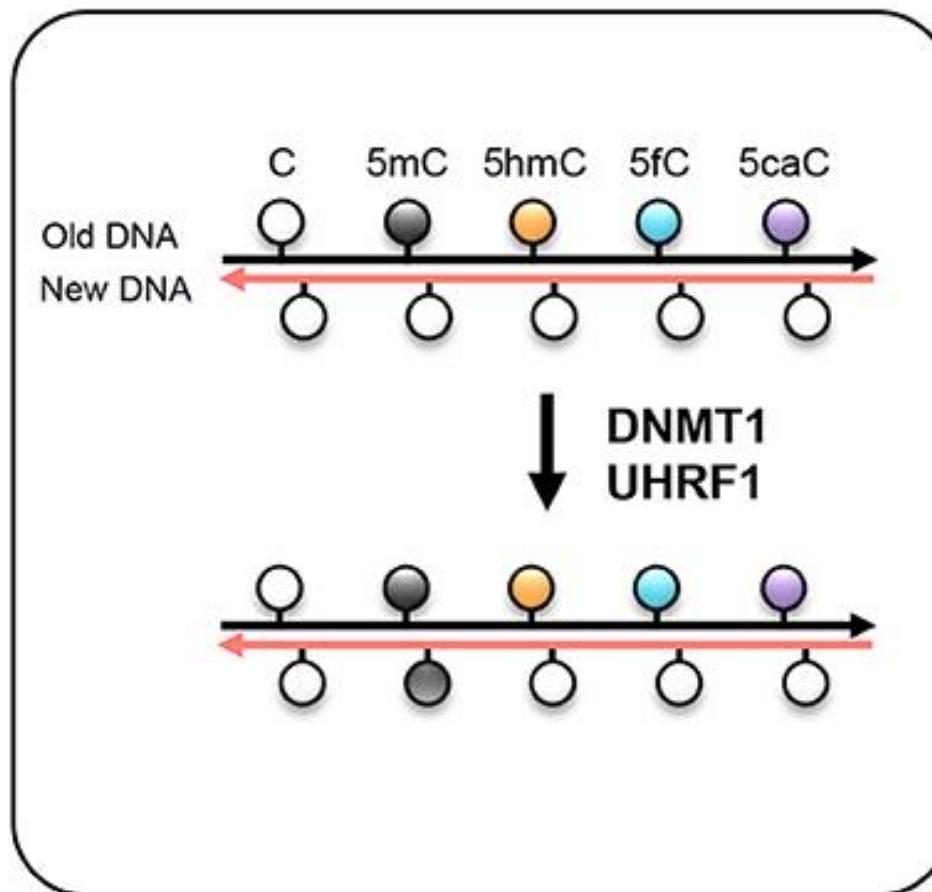
Passive DNA demethylation

The DNMT1/UHRF1 complex recognizes 5mC at the hemi-methylated CpG motif during DNA replication and methylates the unmodified cytosine on the newly synthesized DNA strand.

However, the oxidized methylcytosines 5hmC, 5fC, and 5caC are not recognized by DNMT1/UHRF1, resulting in unmodified cytosine on the new DNA strand.

Further DNA replication in the presence of continuing TET activity will result in progressive dilution of 5mC in the daughter cells.

Passive DNA demethylation

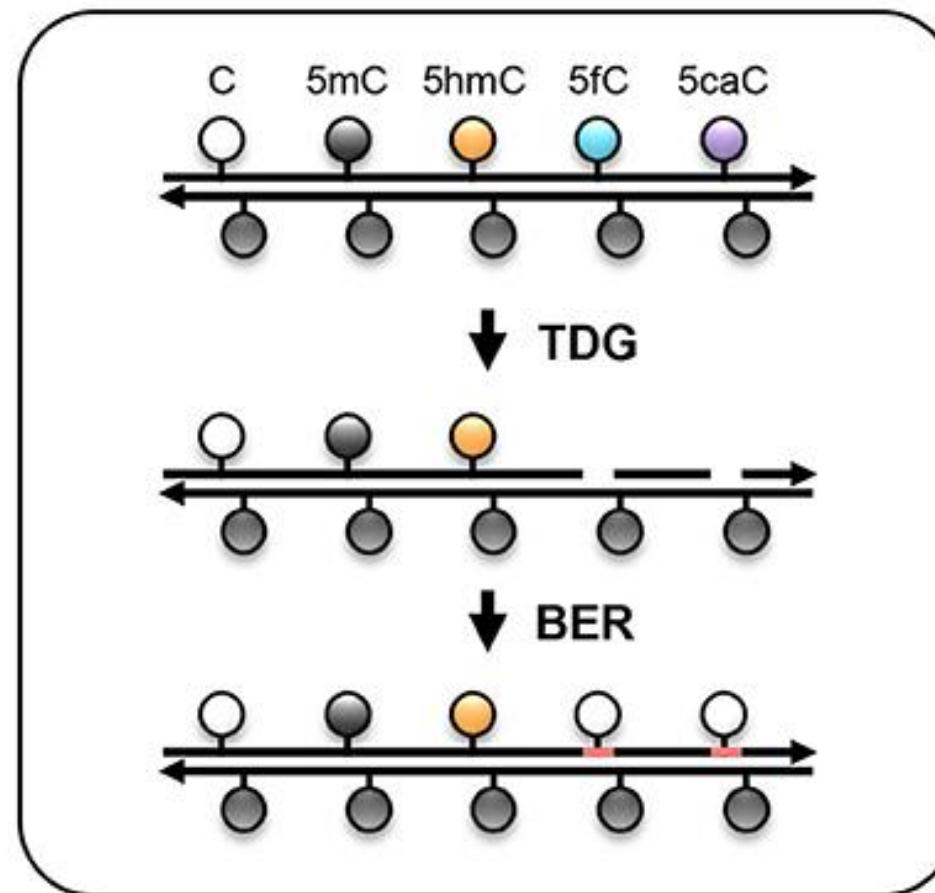


Lio & Rao, Front. Immunol. (2019)

Active DNA demethylation

While 5hmC is stable and persists in the genome, 5fC and 5caC can be recognized and **excised** by the enzyme thymine DNA glycosylase (TDG), and the resulting abasic sites are repaired as unmodified C by base excision repair (BER).

Active DNA demethylation



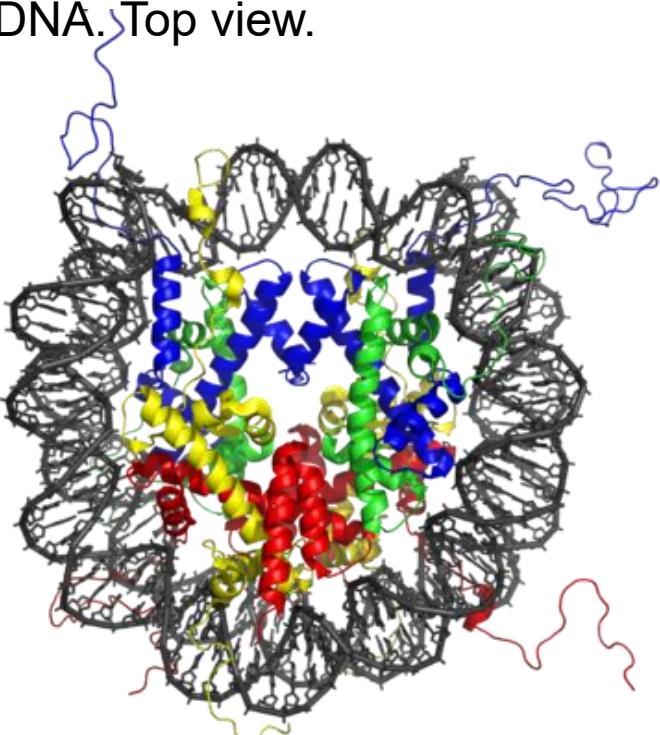
Lio & Rao, Front. Immunol. (2019)

(review V3) The histone code

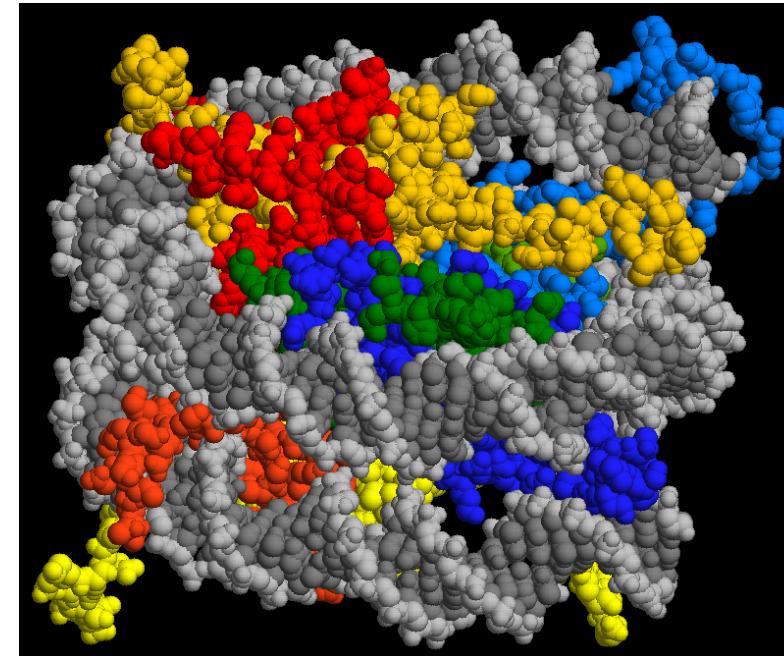
The DNA of eukaryotic organisms is packaged into chromatin, whose basic repeating unit is the **nucleosome**.

A nucleosome is formed by wrapping 147 base pairs of DNA twice around an octamer of four core histones, **H2A** , **H2B** , **H3** and **H4** (2 copies of each one).

X-ray structure of the nucleosome core particle consisting of core histones, and DNA. Top view.



Side view shows two windings of DNA and two histone layers



www.wikipedia.org

Post-translational modifications of histone tails

The disordered histone tails comprise 25-30% of the histone mass.

They extend from the compact histone multimer to provide a platform for various **post-translational modifications (PTMs)**.

These modifications affect the histones' ability to bind DNA and to other histones.

This, in turn, affects **gene expression**.

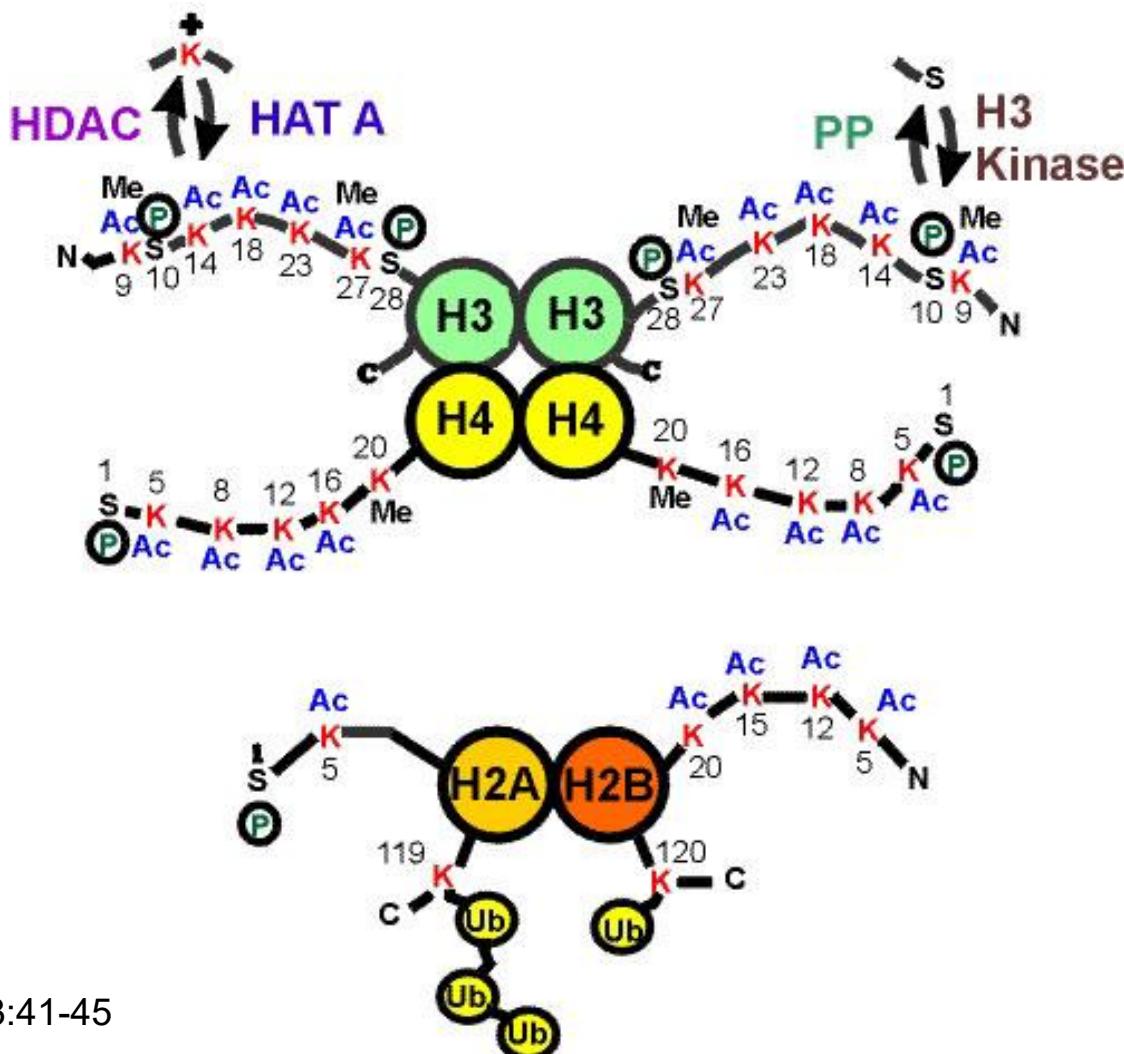
Strahl BD and Allis CD, 2000. Nature 403:41-45

ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS*

BY V. G. ALLFREY, R. FAULKNER, AND A. E. MIRSKY

THE ROCKEFELLER INSTITUTE

PNAS 1964;51:786
First report on PTMs of histones



Mode of action of histone PTMs

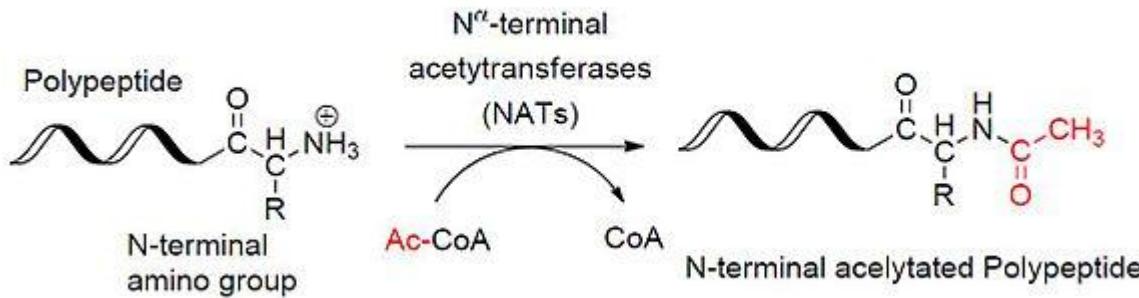
Histone PTMs exert their effects via two main mechanisms.

- (1) PTMs directly influence the overall structure of chromatin, either over short or long distances.
- (2) PTMs regulate (either positively or negatively) the binding of effector molecules.

Bannister, Kouzarides, Cell Res. (2011) 21: 381–395.

PTMs of histone tails

Histone **acetylation** and **phosphorylation** effectively reduce the positive charge of histones.



This potentially disrupts electrostatic interactions between histones and DNA.

This presumably leads to a less compact chromatin structure, thereby facilitating DNA access by protein machineries such as those involved in transcription.

Histone **methylation** mainly occurs on the side chains of lysines and arginines.

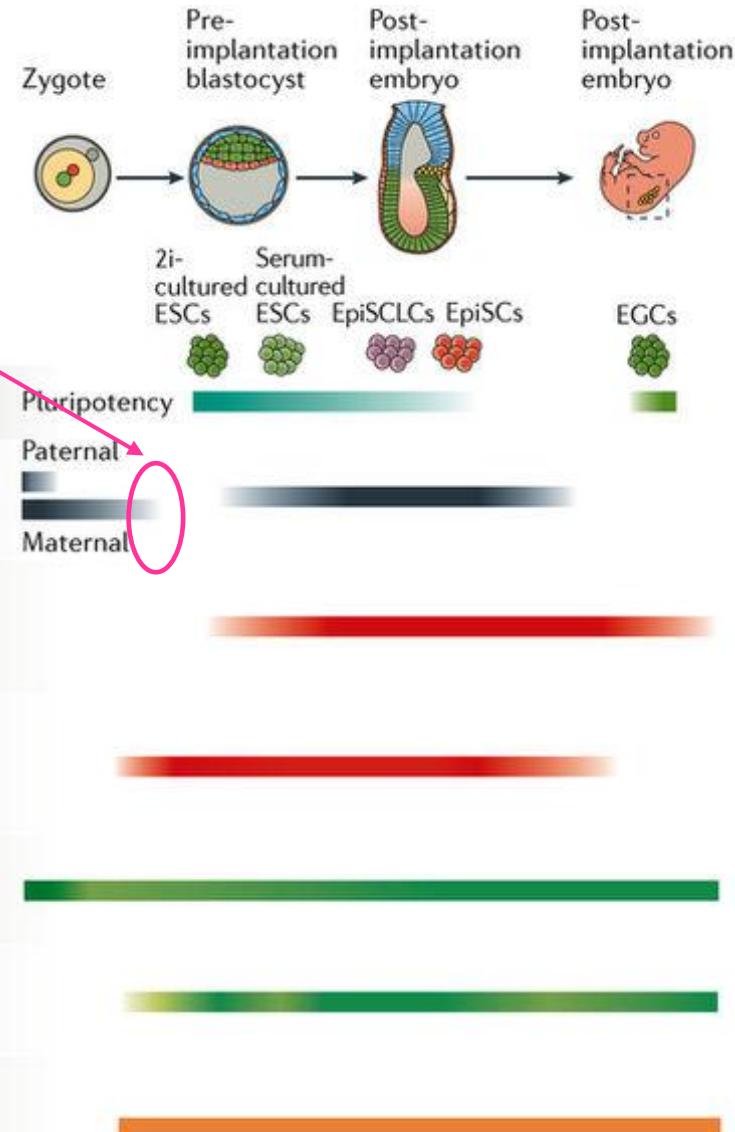
Unlike acetylation and phosphorylation, however, histone methylation does not alter the charge of the histone protein.

Bannister, Kouzarides, Cell Res. (2011) 21: 381–395.

By Ybs.Umich - Own work, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=31240656>

Dynamics of epigenetic modifications

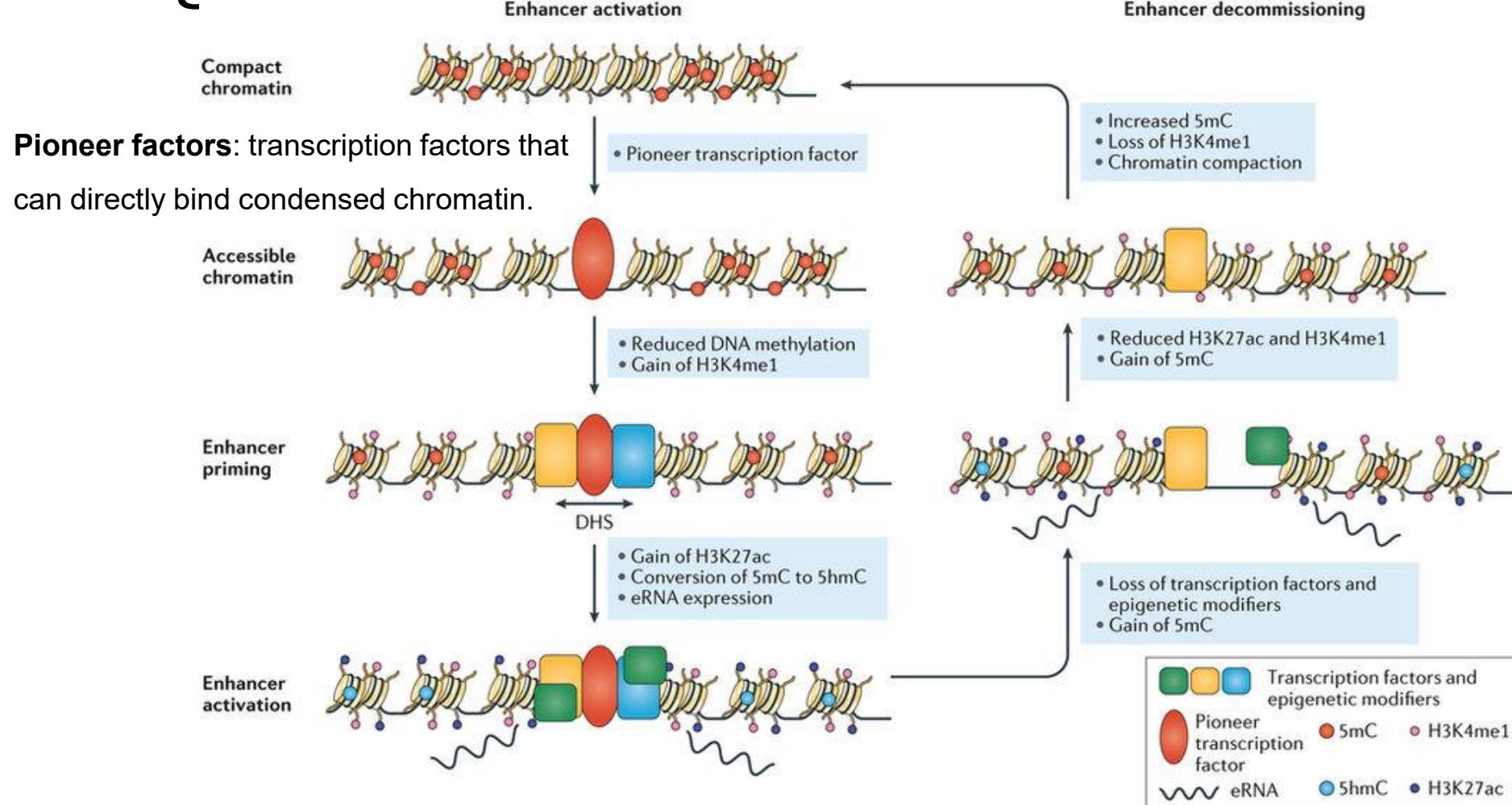
DNA methylation is erased in the paternal and maternal genomes after fertilization and is put back on at later developmental stages.



Chromatin modification	Writers	Erasers	Location	Function	Pluripotency
DNA methylation	DNMT1, DNMT3A and DNMT3B	TET1, TET2 and TET3	CpG dinucleotides	Silencing and others	Maternal
H3K27me3	PRC2	• UTX1 • JMJD3	CpG-rich promoters and intergenic regions	Silencing	Maternal
H3K9me2	G9A and GLP	• JMJD2A, JMJD2B, JMJD2C and JMJD2D • JMJD1A, JMJD1B and JMJD1C	Gene bodies, intergenic regions and enhancers	Silencing	Maternal
H3K4me3	COMPASS-like proteins (SET1, MLL1–MLL2)	• JARID1A, JARID1B, JARID1C and JARID1D • KDM2B	Mainly promoters	Possibly activating	Maternal
H3K27ac	HATs (including CBP/p300, GNATs and MYSTs)	HDACs and sirtuins	Promoters and enhancers	Activating	Maternal
H3K4me1	COMPASS-like proteins (MLL3–MLL4)	LSD1 and LSD2	Promoters, enhancers and intergenic regions	Priming and/or activating	Maternal

Atensi & Stunnenberg, *Nature Rev Genet* **18**, 643–658 (2017)

Events during enhancer activation / decommissioning



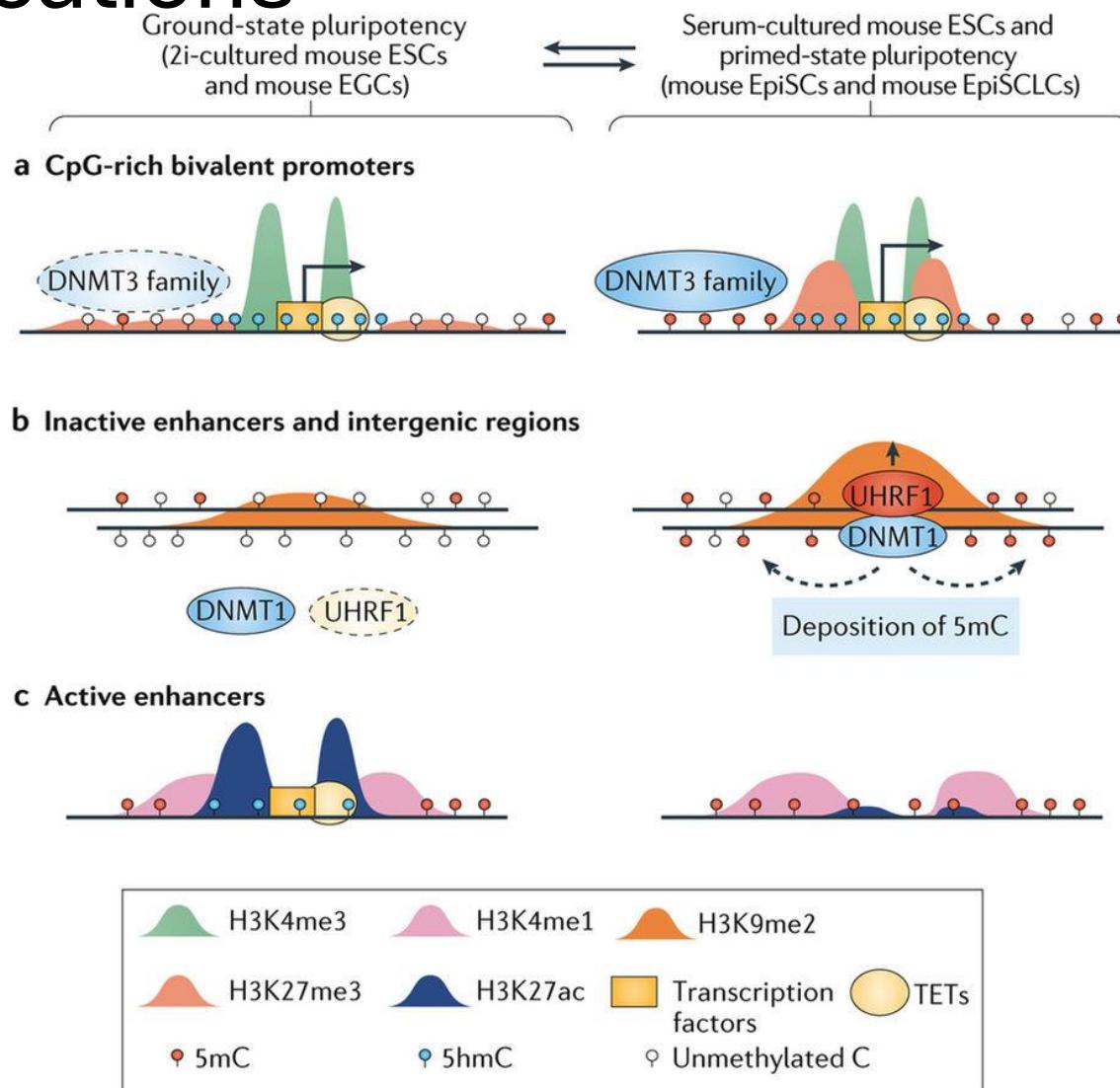
5mC: 5-methyl-cytosine

5hmC: 5-hydroxy-methyl-cytosine

Nature Reviews | Genetics

Atensi & Stunnenberg, *Nature Rev Genet* **18**, 643–658 (2017)

Interplay between DNA methylation and histone modifications



Bivalent chromatin are segments of DNA, bound to histone proteins, that have both repressing and activating epigenetic regulators in the same region. These regulators work to enhance or silence the expression of genes. Since these regulators work in opposition to each other, they normally interact with chromatin at different times. However, in bivalent chromatin, both types of regulators are interacting with the same domain at the same time. Bivalent chromatin domains are normally associated with promoters of transcription factor genes that are expressed at low levels. Bivalent domains have also been found to play a role in developmental regulation in pluripotent embryonic stem cells, as well as gene imprinting.

Atensi & Stunnenberg, *Nature Rev Genet* **18**, 643–658 (2017)
www.wikipedia.org

V9: Cancerogenesis

www.healthline.com specifies:

“A **neoplasm** is an abnormal growth of cells, also known as a **tumor**.

Neoplastic diseases are conditions that cause tumor growth — both benign and malignant.

Benign tumors are noncancerous growths. They usually grow slowly and can't spread to other tissues.

Malignant tumors are cancerous and can grow slowly or quickly.

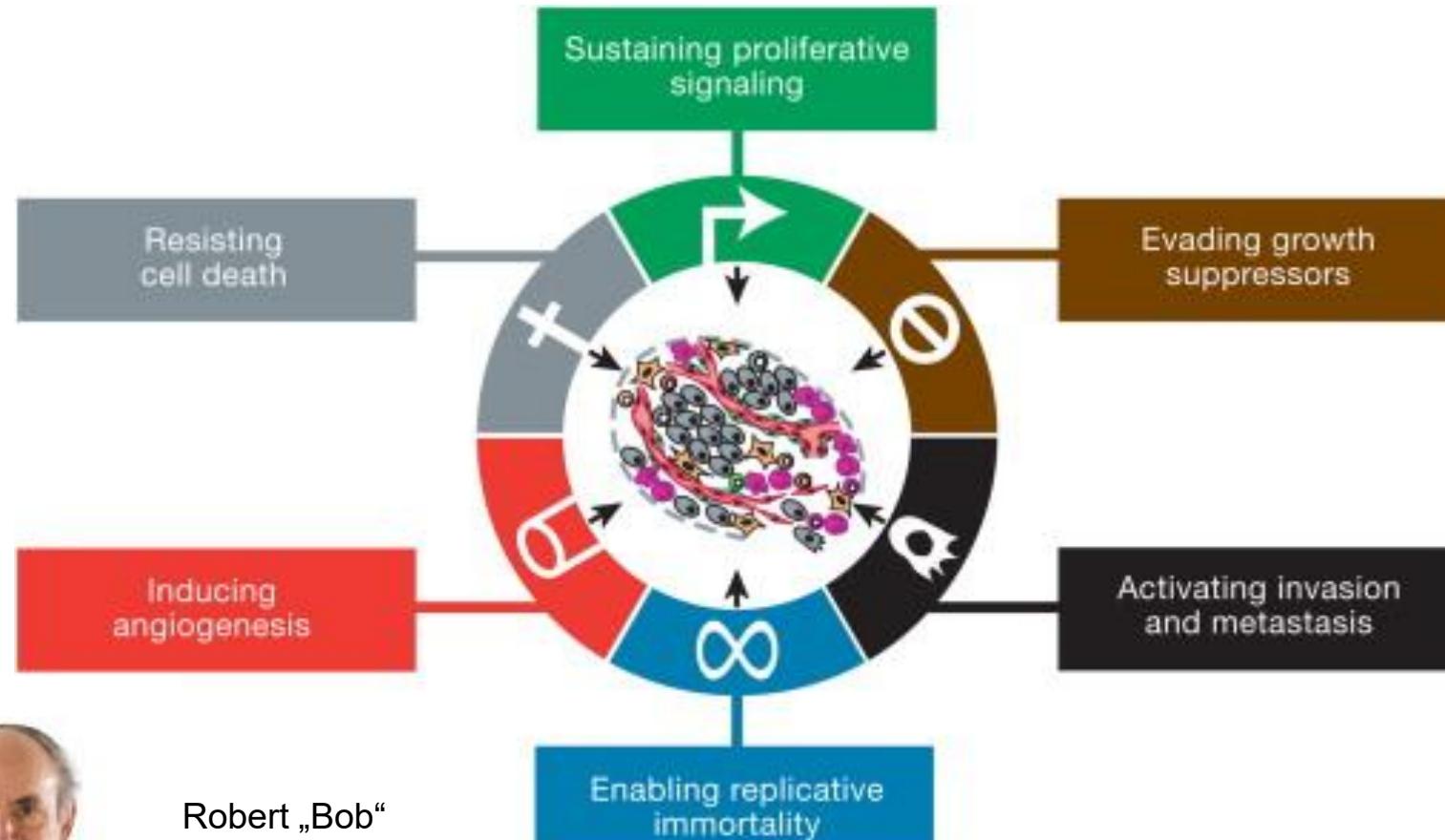
Malignant tumors carry the risk of **metastasis**, or spreading to multiple tissues and organs.”

<https://blog.dana-farber.org/> explains

“**Cancer** is a disease in which cells, almost anywhere in the body, begin to divide uncontrollably. A **tumor** is when this uncontrolled growth occurs in **solid tissue** such as an organ, muscle, or bone. Tumors may spread to surrounding tissues through the blood and lymph systems.”

Hallmarks of cancers

In 2000, Hanahan & Weinberg proposed that the six listed hallmarks of cancer are characteristic common properties of the diverse tumors (neoplastic diseases).



Robert „Bob“
A. Weinberg

Hanahan & Weinberg,
Cell 144, 646–674 (2011)

Mutations sustain proliferative signalling: Raf, ras, Pi3K

High-throughput DNA sequencing analyses of cancer cell genomes revealed somatic mutations in certain human tumors that predict constitutive activation of signaling circuits that are usually triggered by activated growth factor receptors.

~40% of human **melanomas** contain activating mutations affecting the structure of the **B-Raf protein**.

These lead to **constitutive signaling** through the Raf to mitogen-activated protein (MAP)-kinase pathway.

Similarly, mutations in the catalytic subunit of **phosphoinositide 3-kinase (PI3-kinase)** isoforms are detected in several tumor types, which serve to hyperactivate the PI3-kinase signaling circuitry, including its key **Akt/PKB** signal transducer.

Hanahan & Weinberg,
Cell **144**, 646–674 (2011)

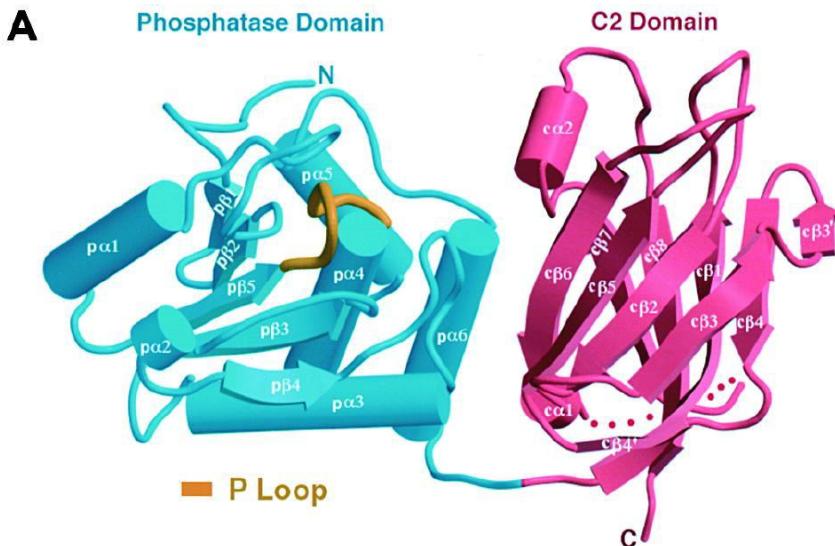
Attenuate proliferative signalling: PTEN

Negative-feedback mechanisms operate at multiple nodes within the proliferative signaling circuitry.

A prominent example involves the **PTEN phosphatase**, which counteracts PI3-kinase by degrading its product, phosphatidylinositol (3,4,5) trisphosphate (PIP_3).

Loss-of-function mutations in PTEN amplify **PI3K signaling** and promote tumorigenesis in a variety of experimental models of cancer.

In human tumors, PTEN expression is often lost by **promoter methylation**.



Hanahan & Weinberg,
Cell **144**, 646–674 (2011)

Lee, Nikola Pavletich,
Cell **99**, 323-334 (1999)

Attenuate proliferative signalling: PTEN

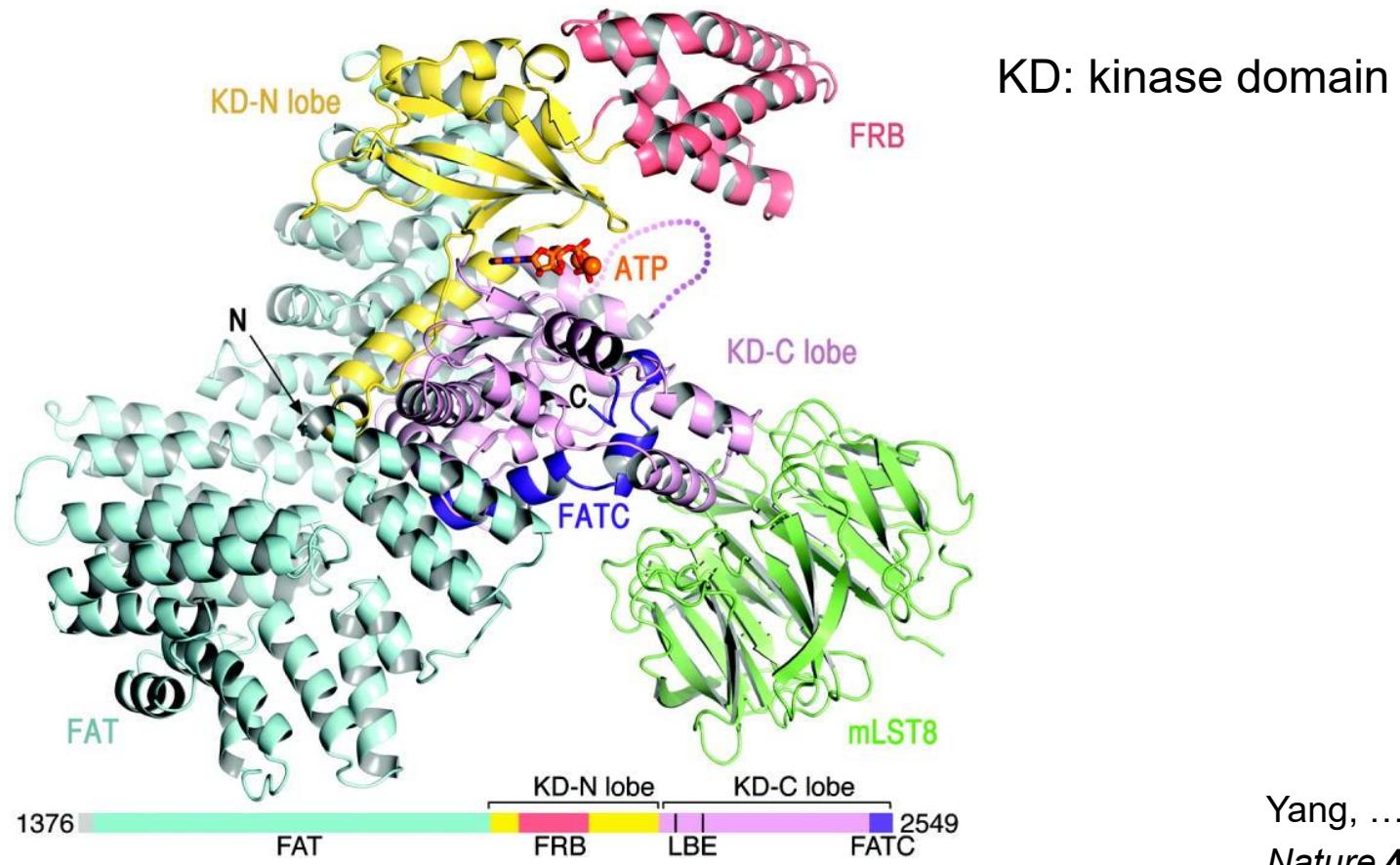


Conservation of PTEN and histogram of 93 **missense (red bars)** and **80 chain termination (green bars)** tumor-derived mutations known in 1999.

Attenuate proliferative signalling: mTOR

Another example involves the **mTOR kinase**. This enzyme coordinates cell growth and metabolism and is positioned both upstream and downstream of the PI3K pathway.

In the circuitry of some cancer cells, mTOR activation results, via negative feedback, in the inhibition of PI3K signaling.



Yang, ... , Nikola Pavletich,
Nature **497**, 217–223 (2013)

Corruption of the TGF-beta pathway

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4717672/> writes

„The TGF- β family is key to specifying the body plan during metazoan development. Members of this family, including nodal, activins, inhibins, bone morphogenetic proteins (BMPs) and growth differentiation factors (GDFs), specify the anterior/posterior and dorsal/ventral axes, endoderm, mesoderm and ectoderm, left-right asymmetry and details of individual organs.

TGF- β 1, TGF- β 2 and TGF- β 3 are important in development, wound healing, immune responses and tumour-cell growth and inhibition.”

TGF- β (**transforming growth factor beta**) has also anti-proliferative effects that need to be evaded by cancer cells.

In many late-stage tumors, TGF- β signaling is redirected away from suppressing cell proliferation. Instead, one finds that TGF- β activates a cellular program, termed the **epithelial-to-mesenchymal transition (EMT)**, that confers on cancer cells traits associated with high-grade malignancy.

Hanahan & Weinberg,
Cell **144**, 646–674 (2011)

Evading Growth Suppressors

Dozens of tumor suppressors that limit cell growth and proliferation in various ways have been discovered through their characteristic inactivation in one or another form of animal or human cancer.

Many of these genes have been validated as bona fide tumor suppressors through gain- or loss-of-function experiments in mice.

The two prototypical tumor suppressors encode the RB (retinoblastoma-associated) and TP53 proteins.

RB and TP53 operate as central control nodes within two key complementary cellular regulatory circuits that govern the **decisions of cells to proliferate** or, alternatively, **activate senescence and apoptotic programs**.

Hanahan & Weinberg,
Cell 144, 646–674 (2011)

Evading Growth Suppressors: RB

The RB protein integrates signals from diverse extracellular and intracellular sources and, in response, decides whether or not a cell should proceed through its growth-and-division cycle.

Cancer cells with defects in RB pathway function are thus missing the services of a critical gatekeeper of cell-cycle progression.

Absence of RB permits persistent cell proliferation.

Hanahan & Weinberg,
Cell **144**, 646–674 (2011)

“Guardian of the cell”: TP53

TP53 receives inputs from **stress** and **abnormality sensors** that function within the cell's intracellular operating systems:

if the degree of damage to the genome is excessive, or if the levels of nucleotide pools, growth-promoting signals, glucose, or oxygenation are suboptimal,

TP53 can call a **halt** to further **cell-cycle progression** until these conditions have been normalized.

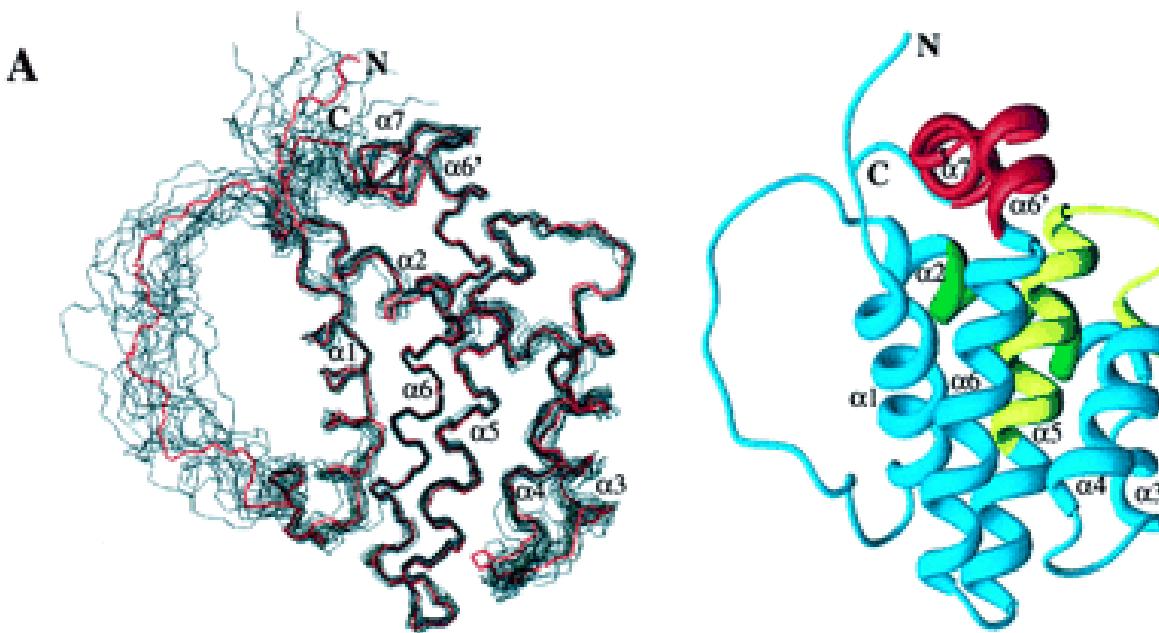
Alternatively, in the face of alarm signals indicating overwhelming or irreparable damage to such cellular subsystems, TP53 can trigger **apoptosis**.

The various effects of activated TP53 are complex and highly context dependent, varying by cell type as well as by the severity and persistence of conditions of cell stress and genomic damage.

Hanahan & Weinberg,
Cell 144, 646–674 (2011)

Resisting Cell Death: Bcl-2

The “apoptotic trigger” that conveys signals between the regulators and effectors is controlled by counterbalancing pro- and antiapoptotic members of the **Bcl-2** family of regulatory proteins.



Backbone (N, Ca, C') superposition of 15
low-energy NMR-derived structures and
Ribbons depiction of the average minimized
structure for Bcl-2.

Petros, ..., Stephen Fesik,
PNAS **98**, 3012 (2001)

Resisting Cell Death: Bcl-2

Bcl-2, along with its close relatives Bcl-x_L, Bcl-w, Mcl-1, A1 are **inhibitors of apoptosis**.

They act by binding to and thereby suppressing two proapoptotic triggering proteins (Bax and Bak) that are embedded in the mitochondrial outer membrane.

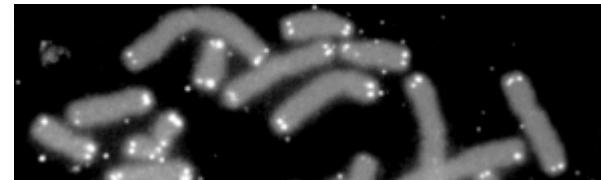
When relieved of inhibition by their antiapoptotic relatives, Bax and Bak disrupt the integrity of the outer mitochondrial membrane, causing the release of proapoptotic signaling proteins including cytochrome c.

The released cytochrome c activates, in turn, a cascade of **caspases** that act via their proteolytic activities to induce the multiple cellular changes associated with the apoptotic program.

Hanahan & Weinberg,
Cell **144**, 646–674 (2011)

Enabling Replicative Immortality: Telomeres

Telomeres protecting the ends of chromosomes are centrally involved in the capability for unlimited proliferation.



Chromosomes,
Telomeres marked
In white.

The telomeres, composed of multiple tandem hexanucleotide repeats, shorten progressively in nonimmortalized cells propagated in culture, eventually losing the ability to protect the ends of chromosomal DNAs from end-to-end fusions.

Such fusions generate unstable dicentric chromosomes whose resolution results in a scrambling of karyotype that threatens cell viability.

Accordingly, the length of telomeric DNA in a cell dictates how many successive cell generations its progeny can pass through before telomeres are largely eroded and have consequently lost their protective functions, triggering entrance into crisis.

Hanahan & Weinberg,
Cell **144**, 646–674 (2011)

Enabling Replicative Immortality: Telomerase

Telomerase, the specialized DNA polymerase that adds telomere repeat segments to the ends of telomeric DNA, is almost absent in nonimmortalized cells but expressed at functionally significant levels in the vast majority (~90%) of spontaneously immortalized cells, including human cancer cells.

By extending telomeric DNA, telomerase is able to counter the progressive telomere erosion that would otherwise occur in its absence.

The presence of telomerase activity, either in spontaneously immortalized cells or in the context of cells engineered to express the enzyme, is correlated with a resistance to induction of both senescence and crisis/apoptosis.

Conversely, suppression of telomerase activity leads to telomere shortening and to activation of one or the other of these proliferative barriers.

Hanahan & Weinberg,
Cell **144**, 646–674 (2011)

Inducing Angiogenesis: VEGF-A

The VEGF-A gene (Vascular Endothelial Growth Factor A) encodes ligands that are involved in **orchestrating new blood vessel growth** during embryonic and postnatal development, and then in homeostatic survival of endothelial cells, as well as in physiological and pathological situations in the adult.

VEGF signaling via three receptor tyrosine kinases (VEGFR-1–3) is regulated at multiple levels. VEGF gene expression can be upregulated both by hypoxia and by oncogene signaling.

Additionally, VEGF ligands can be sequestered in the extracellular matrix in latent forms that are subject to release and activation by extracellular matrix-degrading proteases (e.g., MMP-9).

In addition, other proangiogenic signals, such as members of the fibroblast growth factor (FGF) family, have been implicated in sustaining tumor angiogenesis when their expression is chronically upregulated.

Hanahan & Weinberg,
Cell **144**, 646–674 (2011)

Activating Invasion and metastasis: ECM, E-cadherin

When carcinomas arising from epithelial tissues progress to higher pathological grades of malignancy (reflected in local invasion and distant metastasis), the associated cancer cells typically develop alterations in their shape as well as in their **attachment** to other cells and to the **extracellular matrix (ECM)**.

In particular, **carcinoma cells loose E-cadherin**, a key cell-to-cell adhesion molecule. E-cadherin forms adhering junctions with adjacent epithelial cells. Thereby, E-cadherin helps to assemble epithelial cell sheets and maintain the quiescence of the cells within these sheets.

Increased expression of E-cadherin prevents invasion and metastasis, whereas reduced expression potentiates these phenotypes.

E-cadherin is frequently downregulated and occasionally inactivated by mutations in human carcinomas. This strongly supports its role as a key suppressor of this hallmark capability.

Hanahan & Weinberg,
Cell **144**, 646–674 (2011)

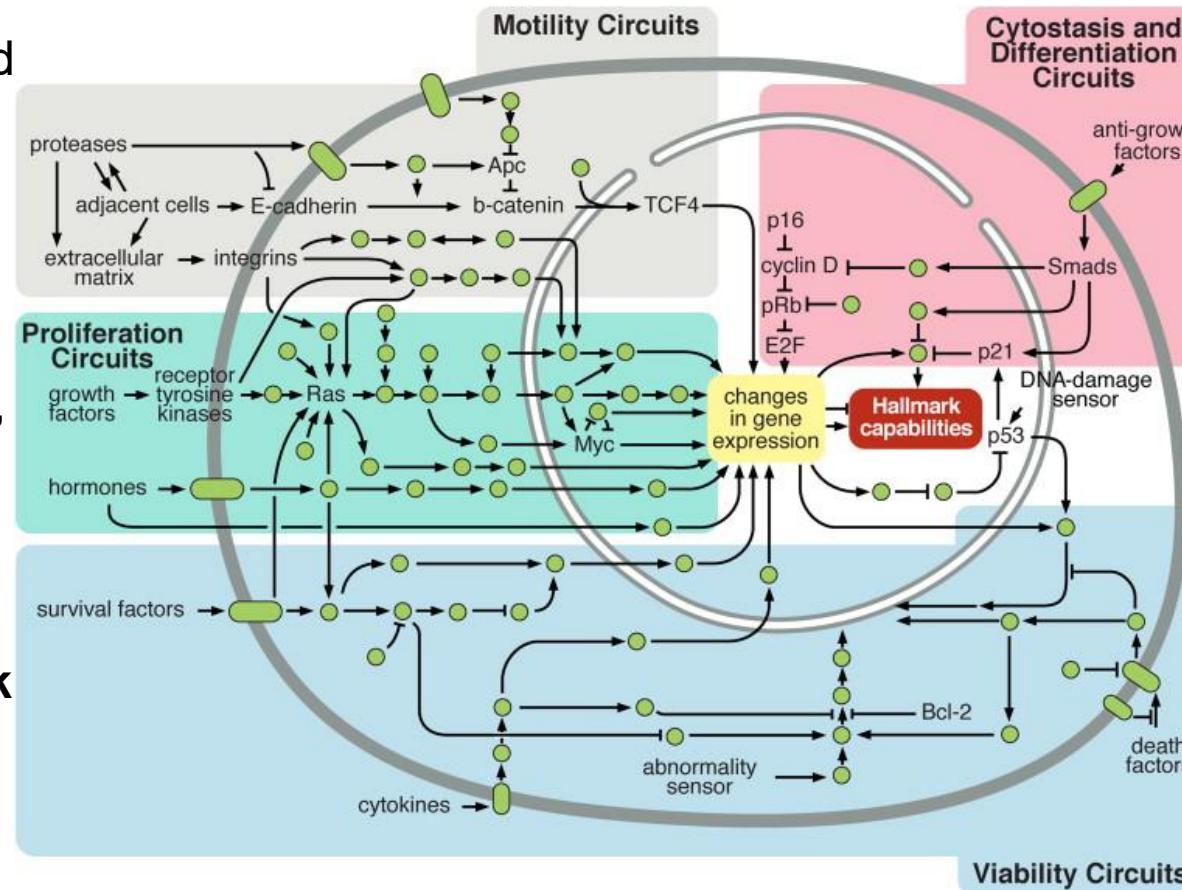
Intracellular Signaling Networks

An elaborate integrated circuit operates within normal cells and is reprogrammed to regulate hallmark capabilities within cancer cells.

Separate subcircuits, depicted here in differently colored fields, are specialized to orchestrate the various capabilities.

This depiction is simplistic. There is considerable **crosstalk** between such subcircuits.

Also, each of these subcircuits is connected with signals originating from other cells in the **tumor microenvironment**.



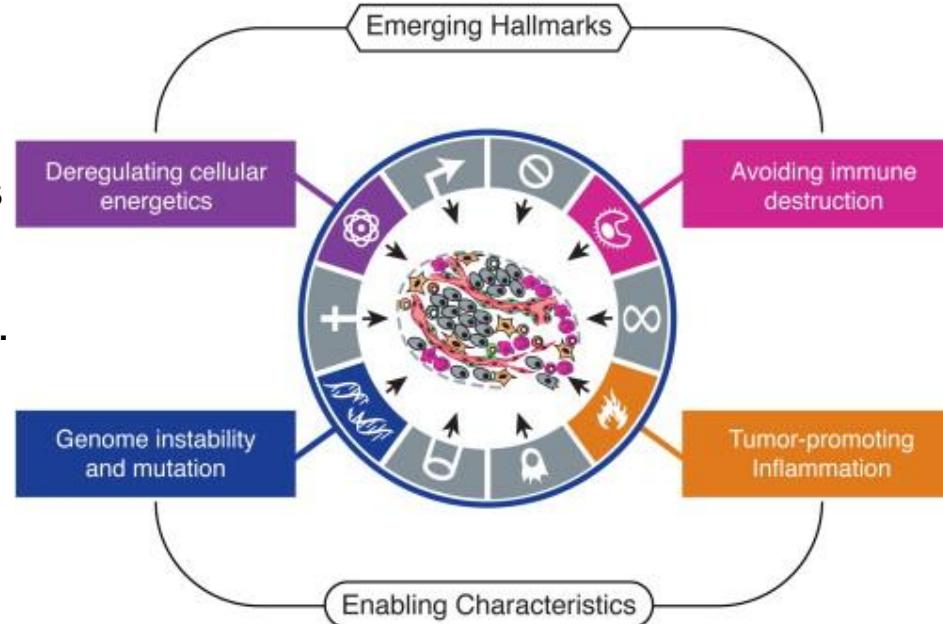
Hanahan & Weinberg,
Cell 144, 646–674 (2011)

Emerging Hallmarks

Additional hallmarks of cancers discussed in the 2011 paper.

(1) In tumors, the **cellular metabolism** is reprogrammed so that neoplastic proliferation is most effectively supported.

(2) Cancer cells **evade immunological destruction** by T and B lymphocytes, macrophages, and natural killer cells.



(3) **Genomic instability** and thus **mutability** endow cancer cells with genetic alterations that drive tumor progression.

(4) **Inflammation** by innate immune cells designed to fight infections and heal wounds can instead result in their inadvertent support of multiple hallmark capabilities. This manifests the tumor-promoting consequences of inflammatory responses.

Hanahan & Weinberg,
Cell 144, 646–674 (2011)

Driver and passenger mutations

Genomewide sequencing has shown that every tumor harbors thousands of genetic (and epigenetic) alterations that are not present in the patient's germline.

Only a very small fraction of these alterations are in “**driver genes**”.

When driver genes are mutated, this endows the tumor cell with a **growth advantage** over surrounding cells.

The remaining alterations are called “**passenger mutations**” .

They are found in tumor cells only because they occurred coincidentally during the long march toward tumorigenesis.

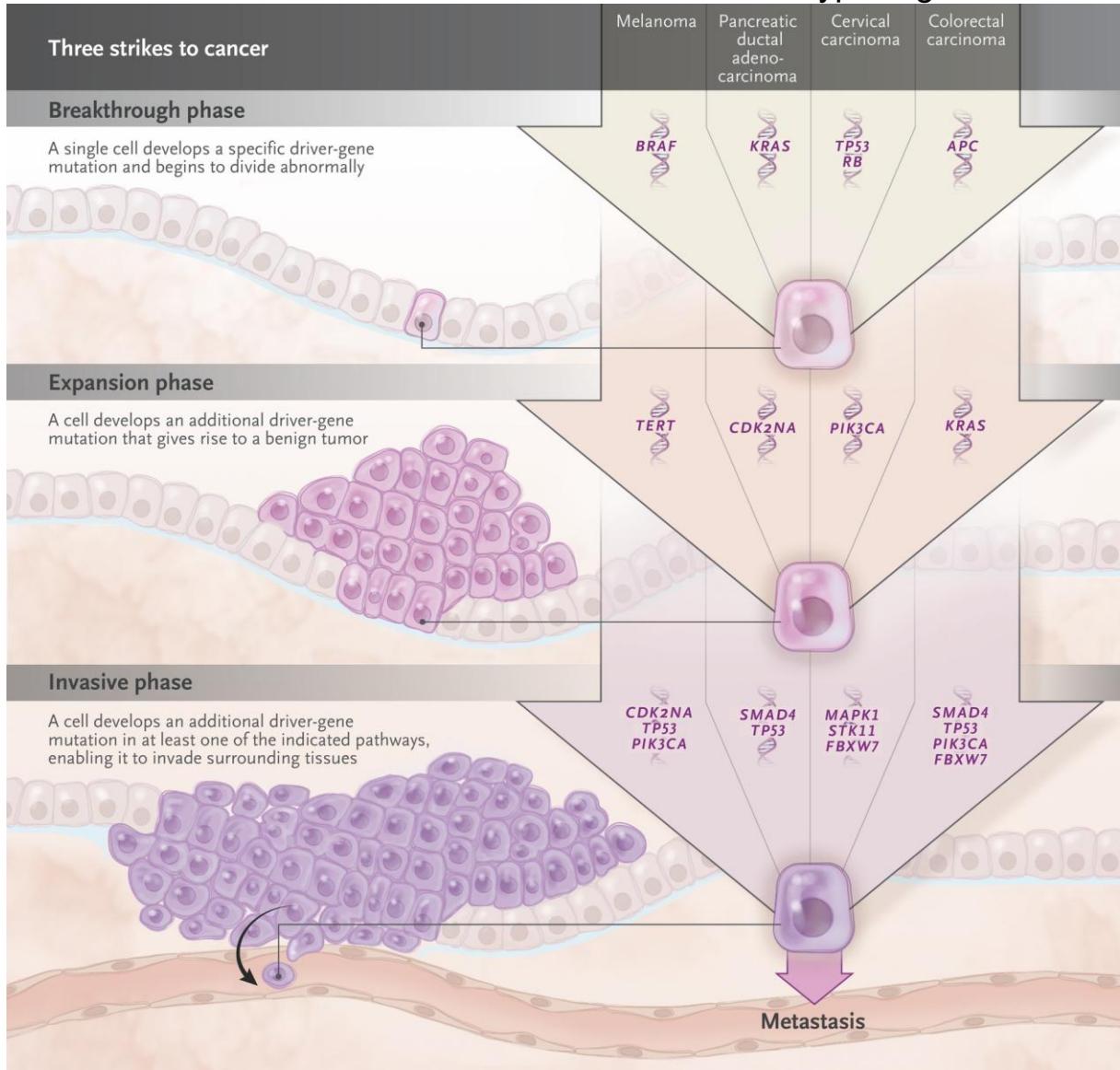
Only about 200 of the 20,000 genes in the human genome have been shown to act as driver genes for common cancers.

Driver genes appear to function through a limited number of pathways that regulate cells' growth and fate.

Vogelstein & Kinzler,
NEJM **373**, 1895-1898 (2015)

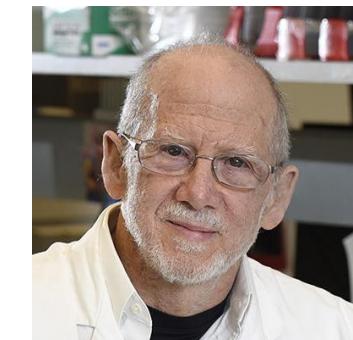
Three Strikes and You're Out

Mutations in typical genes.



Bert Vogelstein from Johns Hopkins is one of the most influential cancer researchers.

Tumors evolve in 3 phases shown in the figure.



Bert Vogelstein

Vogelstein & Kinzler,
NEJM 373, 1895-1898 (2015)
www.hhmi.org

Driver and passenger mutations

Driver-gene mutations should be viewed at a pathway level rather than at an individual-gene level. E.g., colorectal cancers are initiated by mutations in genes in the adenomatous polyposis coli (APC) pathway. This pathway includes the genes *APC*, *CTNNB1*, *SOX9*, *TCF7L1*, *TCF7L2*, and *AMER1*.

The order in which driver-gene mutations occur is also important. E.g., RAS-pathway mutations are the initiating events for pancreatic ductal adenocarcinomas and melanomas but occur later in colorectal tumorigenesis.

Genome-sequencing data **exclude the possibility of spontaneous tumors**: a normal adult cell cannot suddenly transform into a cancer cell.

Every time a cancer (or normal) cell divides, a few new mutations occur. Cancers continuously evolve, always generating more passenger mutations and occasionally another driver-gene mutation that increases growth.

Vogelstein & Kinzler,
NEJM **373**, 1895-1898 (2015)

The human cancer transcriptome

Analysis of transcriptomic data for ca 8000 patients suffering from 17 major cancer types showed that:

- a large fraction of cancer protein-coding genes are differentially expressed and, in many cases, have an impact on overall patient survival.
- shorter patient survival was generally associated with up-regulation of genes involved in mitosis and cell growth and down-regulation of genes involved in cellular differentiation.

Uhlen et al, *Science* **357**, eaan2507 (2017)

The human cancer transcriptome

On the one hand, 41% of the protein-coding genes were expressed in all analyzed cancers.

On the other hand, 46% of the protein-coding genes displayed tumor type-restricted expression.

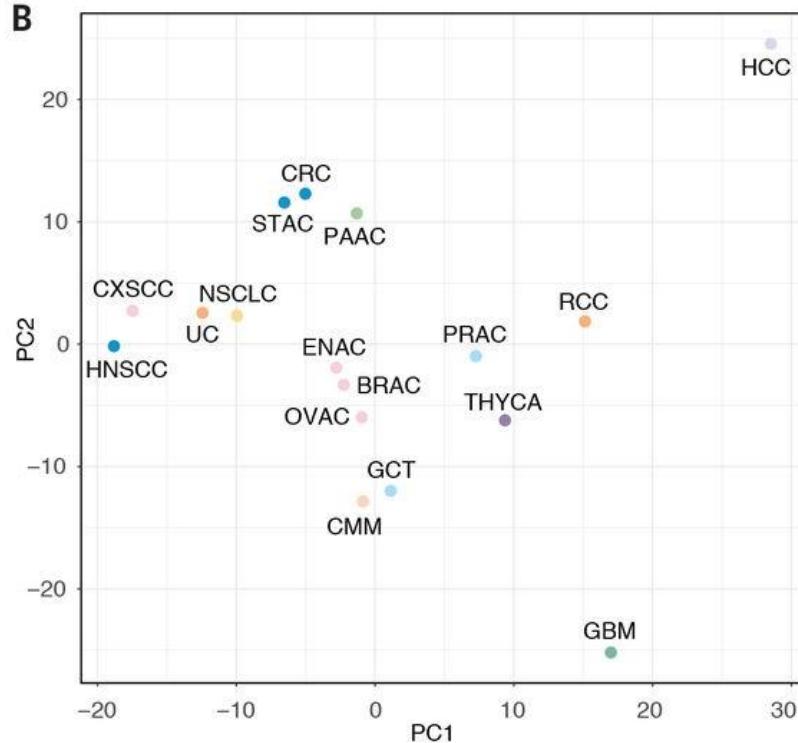
13% of the protein-coding genes were not detected in any tumor types investigated.

The majority of the genes ($n = 5772$) detected in all samples were shared between cancers and normal tissues, whereas 2401 additional genes were expressed in all cancers analyzed, but with more restricted expression in the normal tissues.

These “**housekeeping**” genes in tumors are enriched in biological functions related to DNA replication and the regulation of apoptosis and mitosis.

Uhlen et al, *Science* **357**, eaan2507 (2017)

PCA of the human cancer transcriptome



PCA analysis based on transcriptomic data:
Cancer types that share a similar tissue type
of origin or similar morphological features
and phenotypic expression patterns
are grouped together.

E.g., cancers with a dominating squamous cell carcinoma phenotype, such as cervical (CXSCC) or head and neck cancer (HNSCC), clustered together close to the related urothelial cell carcinoma and non–small cell lung cancer (NSCLC), which also contains a large fraction of squamous cell carcinoma.

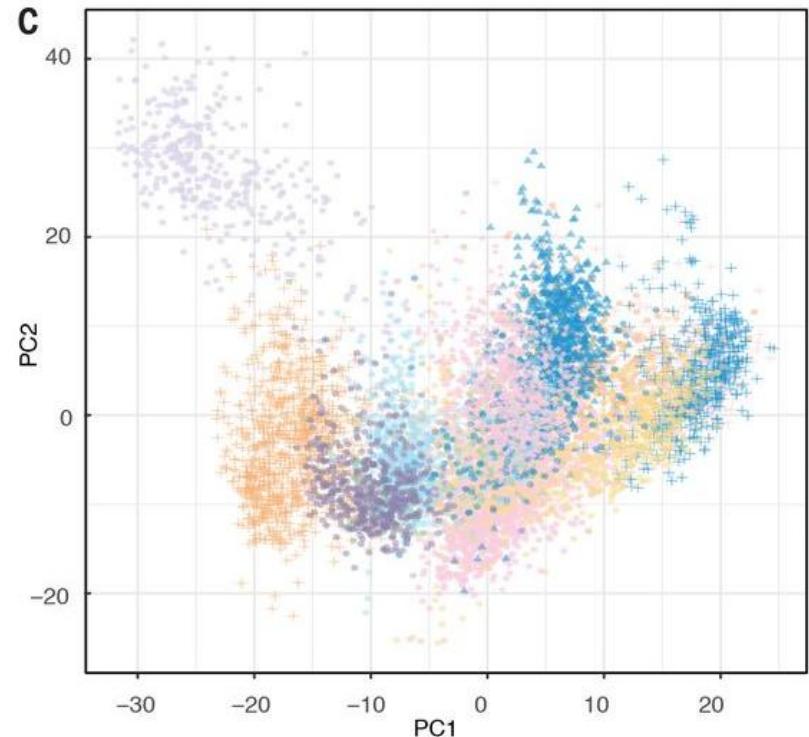
Adenocarcinomas that originate from the gastrointestinal tract, including pancreatic cancer (PAAC), cluster separately from the cluster containing the 3 adenocarcinomas representing female cancer (i.e., breast BRAC, endometrial ENAC, and ovarian cancer OVAC).

Glioma (brain GBM) and hepatocellular (liver HCC) carcinoma were the most divergent tumor types.

Uhlen et al, *Science* **357**, eaan2507 (2017)

The human cancer transcriptome

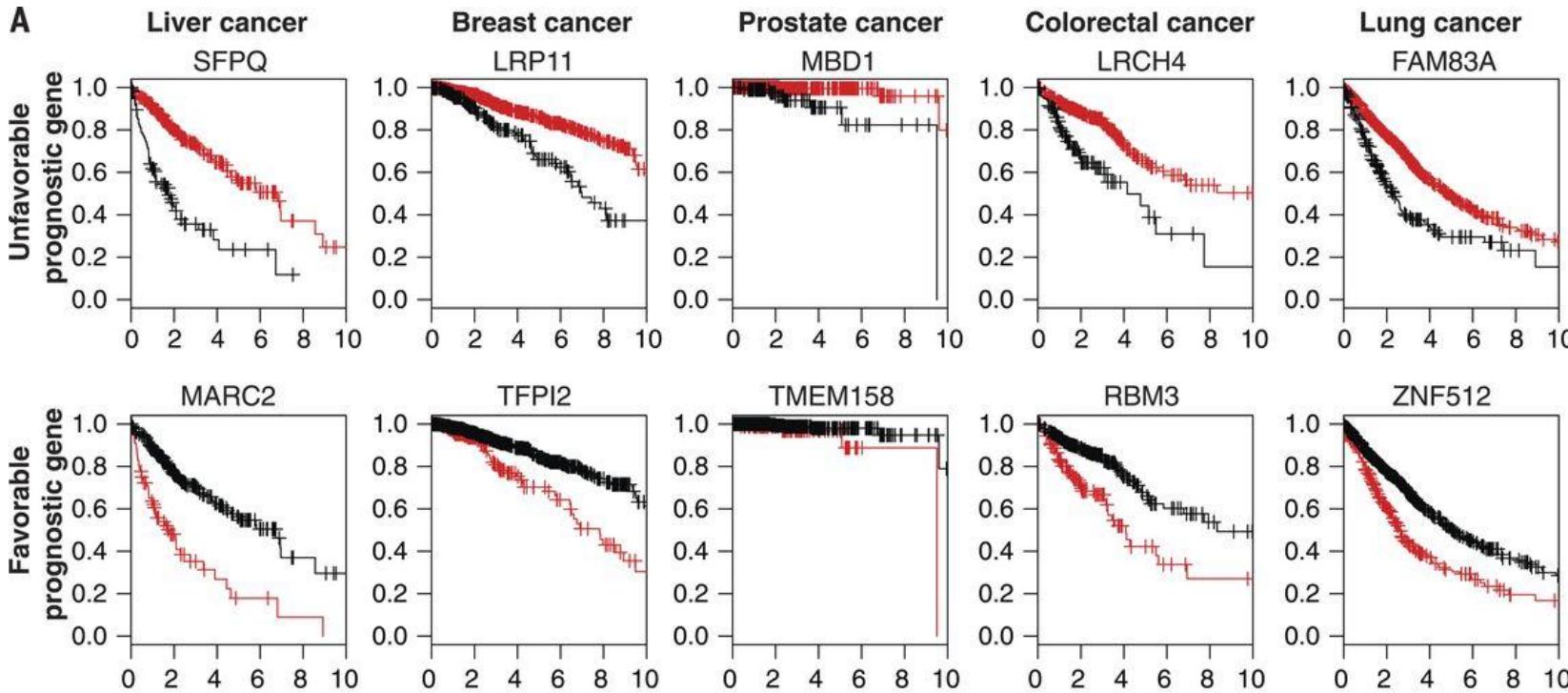
Same analysis for individual tumors



-> there is a large overlap in expression among individuals patients of different cancer types.

Uhlen et al, *Science* **357**, eaan2507 (2017)

Clinical outcome



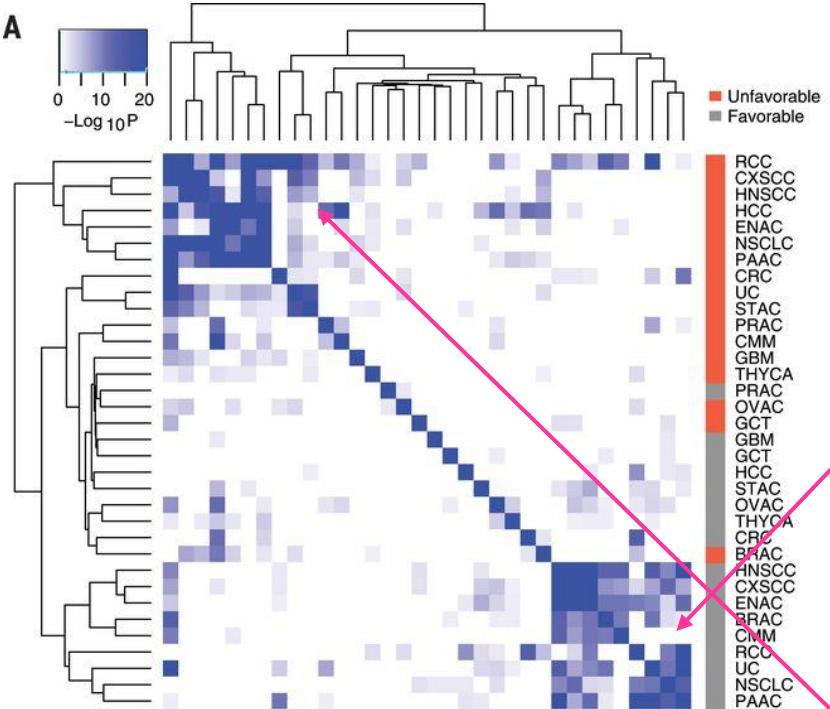
Black: high expression of this gene, red: low expression.

(Top) 5 genes with unfavorable effect on survival.

(Bottom) 5 genes with favorable effect on survival.

Uhlen et al, *Science* **357**, eaan2507 (2017)

Overlap of prognostic genes among cancer types



For most cancers, little correlation was observed, suggesting a relatively limited number of common prognostic genes.

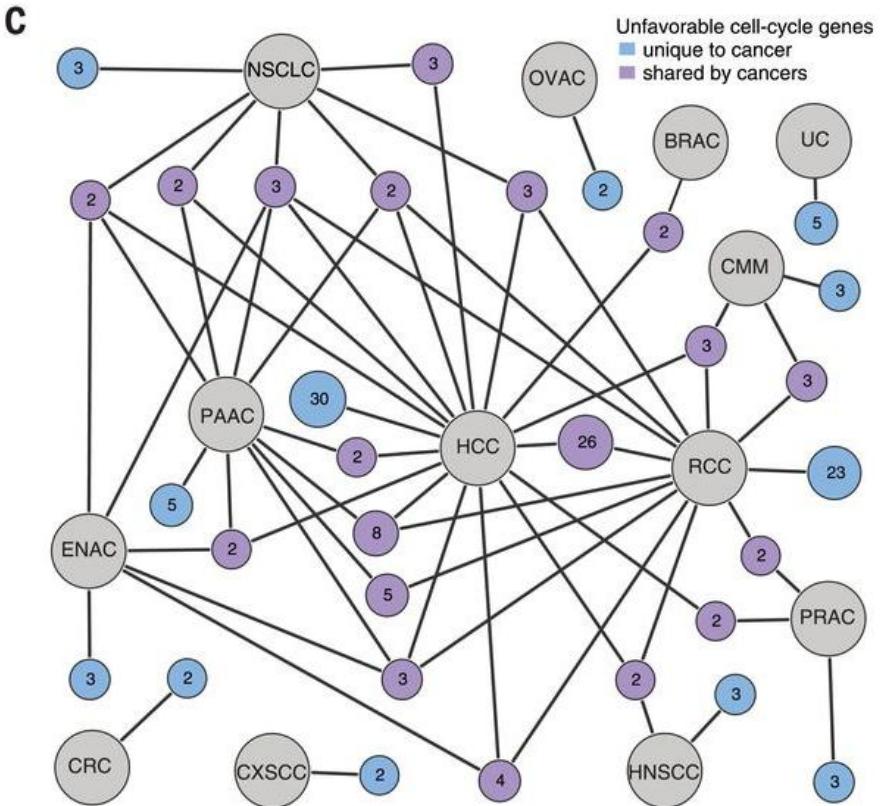
In contrast, a significant overlap of favorable prognostic genes was observed for other cancers (e.g., renal, breast, lung, and pancreatic cancers).

Similarly, unfavorable prognostic genes for some cancers, including renal, liver, lung, and pancreatic cancer, clustered together.

No prognostic genes were shared among more than 7 of the cancer types

Uhlen et al, *Science* **357**, eaan2507 (2017)

Functional annotation of prognostic genes



GO analysis of most significant prognostic genes:

many of the common unfavorable genes are related to cell proliferation, including mitosis, cell cycle regulation, and nucleic acid metabolism.

In contrast, few GO functions were significantly overrepresented by the common favorable genes; the most enriched GO functions were positive regulation of cell activation, regulation of immune cell activation, and cell-cell adhesion.

Uhlen et al, *Science* **357**, eaan2507 (2017)

V10: Cancerogenesis (II)

- **Oncogenic signaling pathways**

Sanchez-Vega et al, *Cell* **173**,
321-337.e10 (2018)

Genetic alterations in signaling pathways that control cell-cycle progression, apoptosis, and cell growth are common hallmarks of cancer.

- **Cancer driver genes**

Martinez-Jimenz et al, *Nature Rev Cancer* **20**, 555-572 (2020)

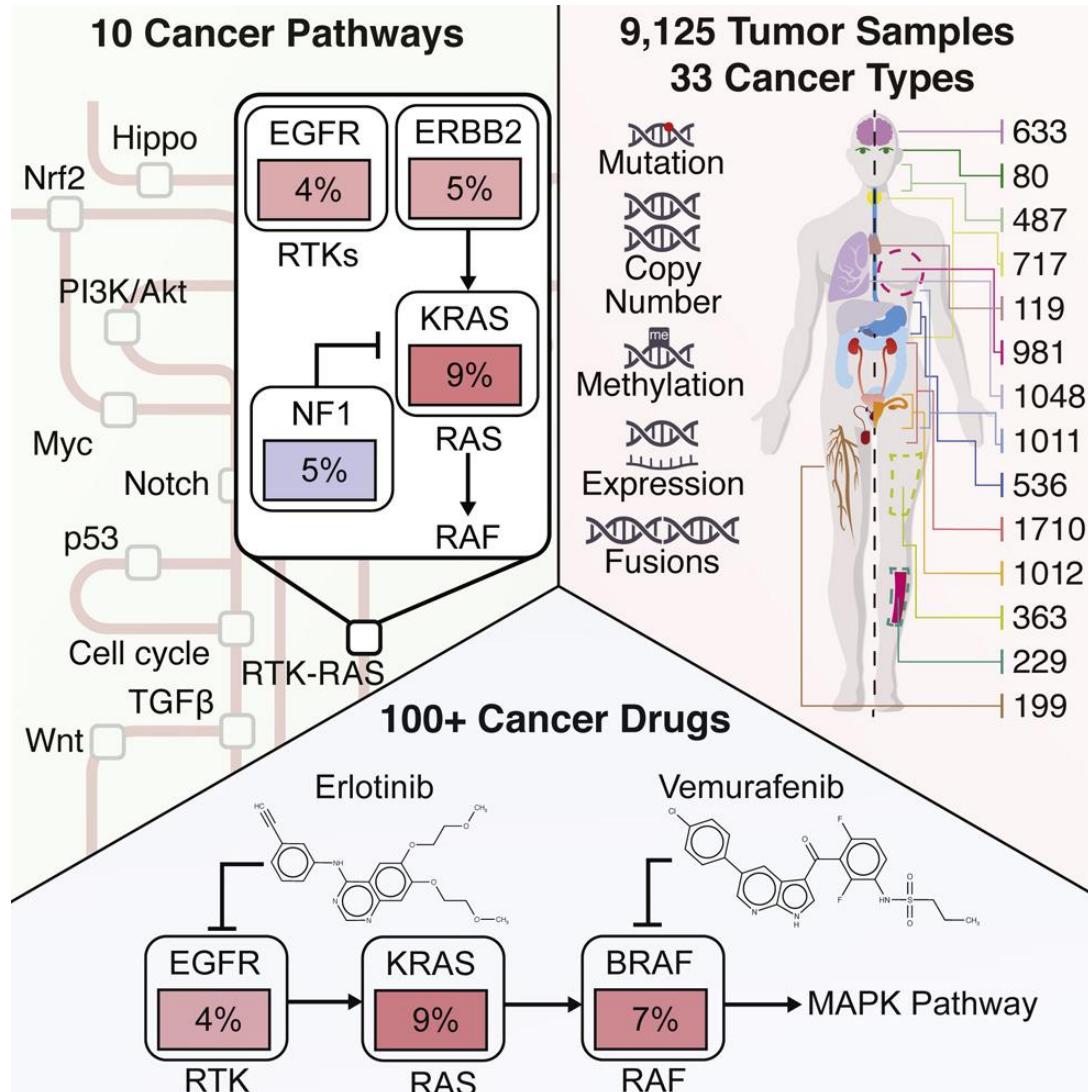
Cancers are diseases characterized by abnormal and uncontrolled cellular growth caused primarily by genetic mutations.

These mutations, called '**drivers**' due to their ability to drive tumorigenesis, confer on cells in a somatic tissue certain selective advantages with respect to neighboring cells.

These mutations occur in a set of genes called '**cancer driver genes**'.

Mutant forms of driver genes affect the homeostatic development of a set of key cellular functions.

Oncogenic Signaling Pathways in TCGA



Alteration map of 10 signaling pathways across 9,125 samples from 33 cancer types

57% of tumors have at least one potentially actionable alteration in these pathways.

The percentages reflect the mutation frequency of this gene in tumor genomes.

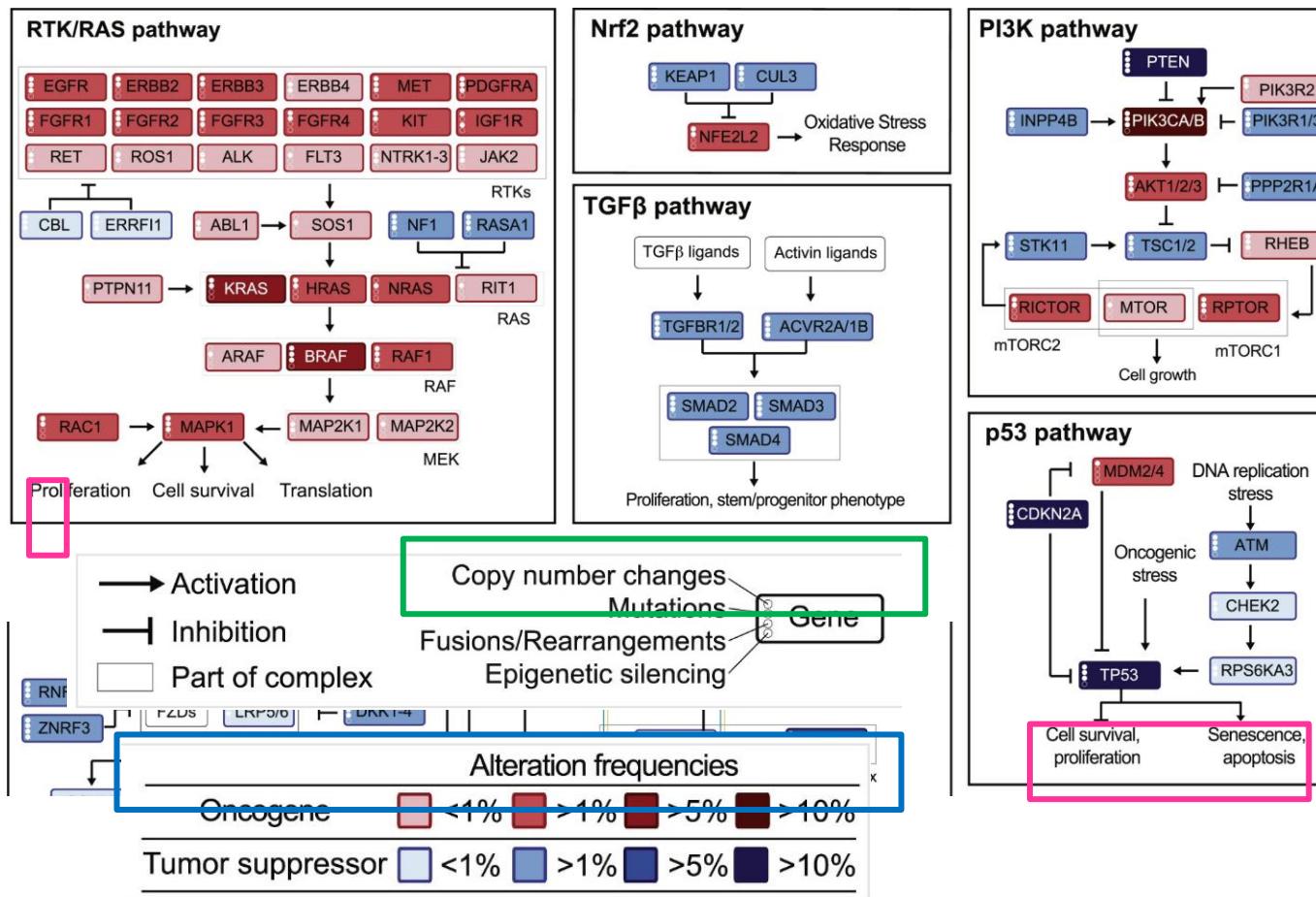
E.g. 9% of all tumor patients have mutations in KRAS.

Sanchez-Vega et al, *Cell* 173, 321-337.e10 (2018)

Members + interactions of major signalling pathways 1 - 5

Downstream effects of pathways are listed below each pathway.

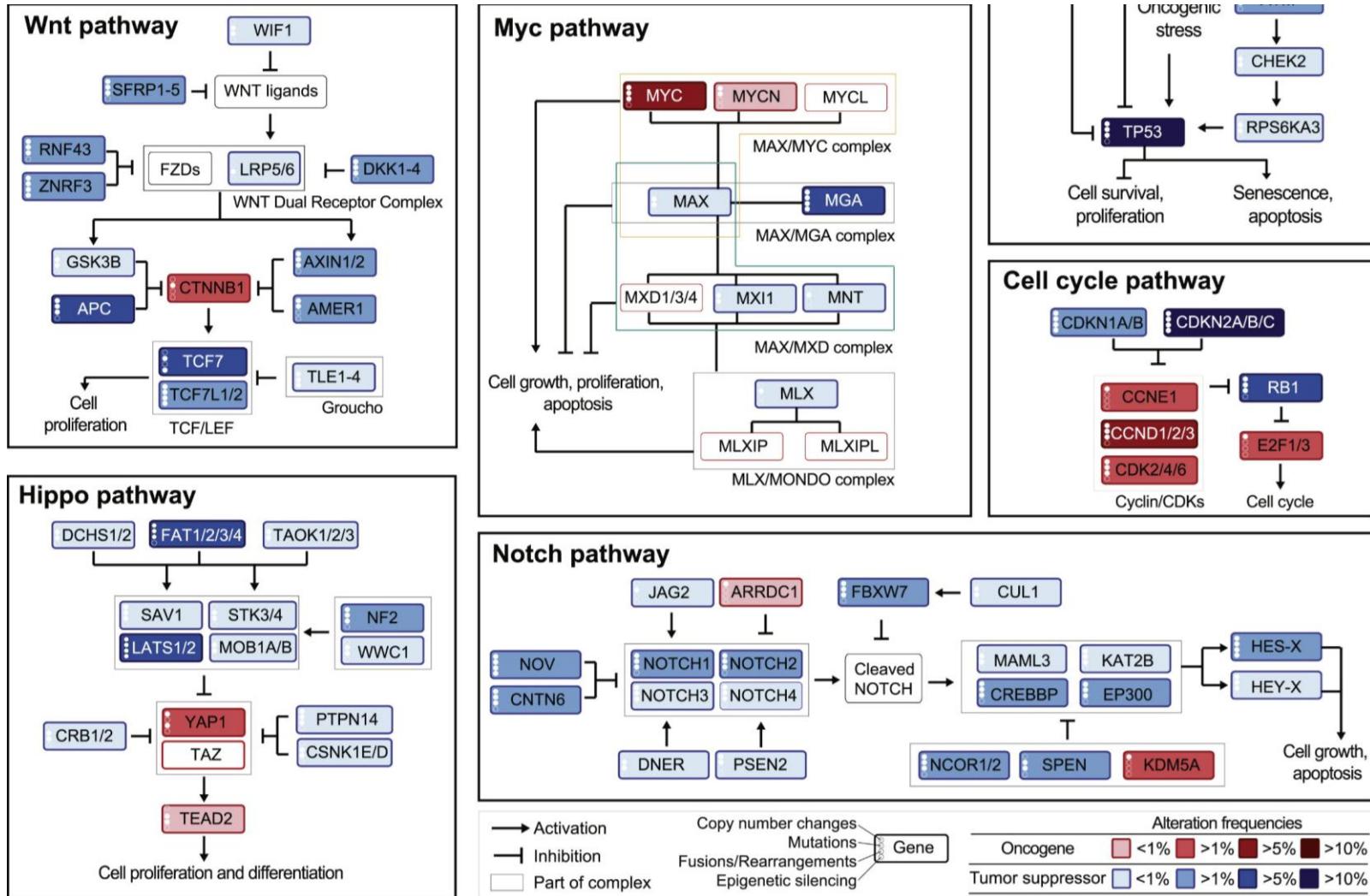
Genes are altered at different frequencies (see blue box at bottom) by oncogenic activations (red) or tumor suppressor inactivations (blue).



The types of somatic alteration considered for each gene (copy-number alterations, mutations, fusions or epigenetic silencing) are specified using a set of four vertical dots on the left of each gene symbol (green box).

Sanchez-Vega et al, *Cell* 173, 321-337.e10 (2018)

Major signalling pathways 6 - 10



Sanchez-Vega et al, *Cell* 173,
321-337.e10 (2018)

V11: Tumor microenvironment

A tumor is not simply a group of cancer cells, but rather a heterogeneous collection of cancer cells, infiltrating and resident host cells, secreted factors and extracellular matrix.

Tumor cells stimulate significant molecular, cellular and physical changes within their host tissues to support tumor growth and progression.

A tumor microenvironment is a complex and continuously evolving entity.

The composition of the tumor microenvironment varies between tumor types, but hallmark features include **immune cells**, **stromal cells**, **blood vessels**, and **extracellular matrix**.

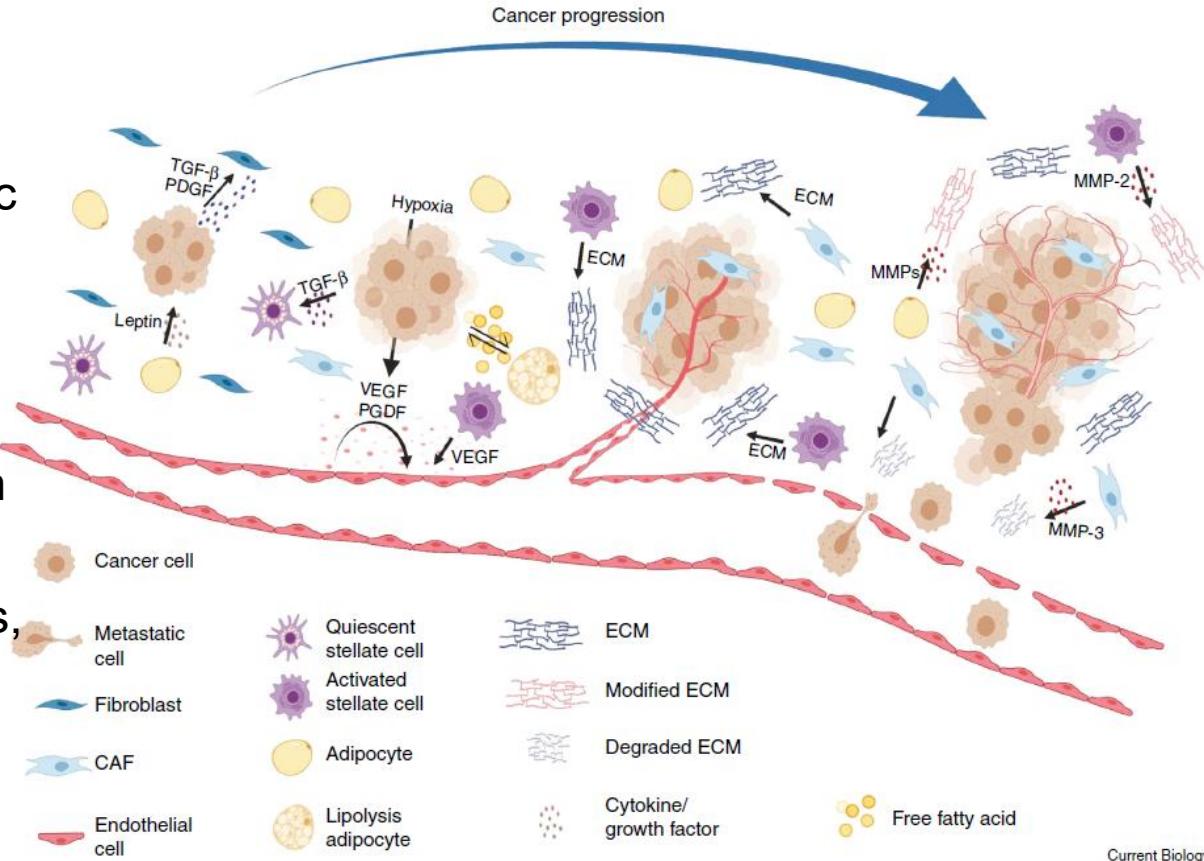
It is believed that the “tumor microenvironment is not just a silent bystander, but rather an active promoter of cancer progression”.

Anderson & Simon,
Current Biology 30, R921–R925, (2020)

Stromal cells of tumor microenvironment

In the tumor microenvironment, stromal cells from neighboring tissue and cancer cells are in a dynamic relationship promoting the cancer progression.

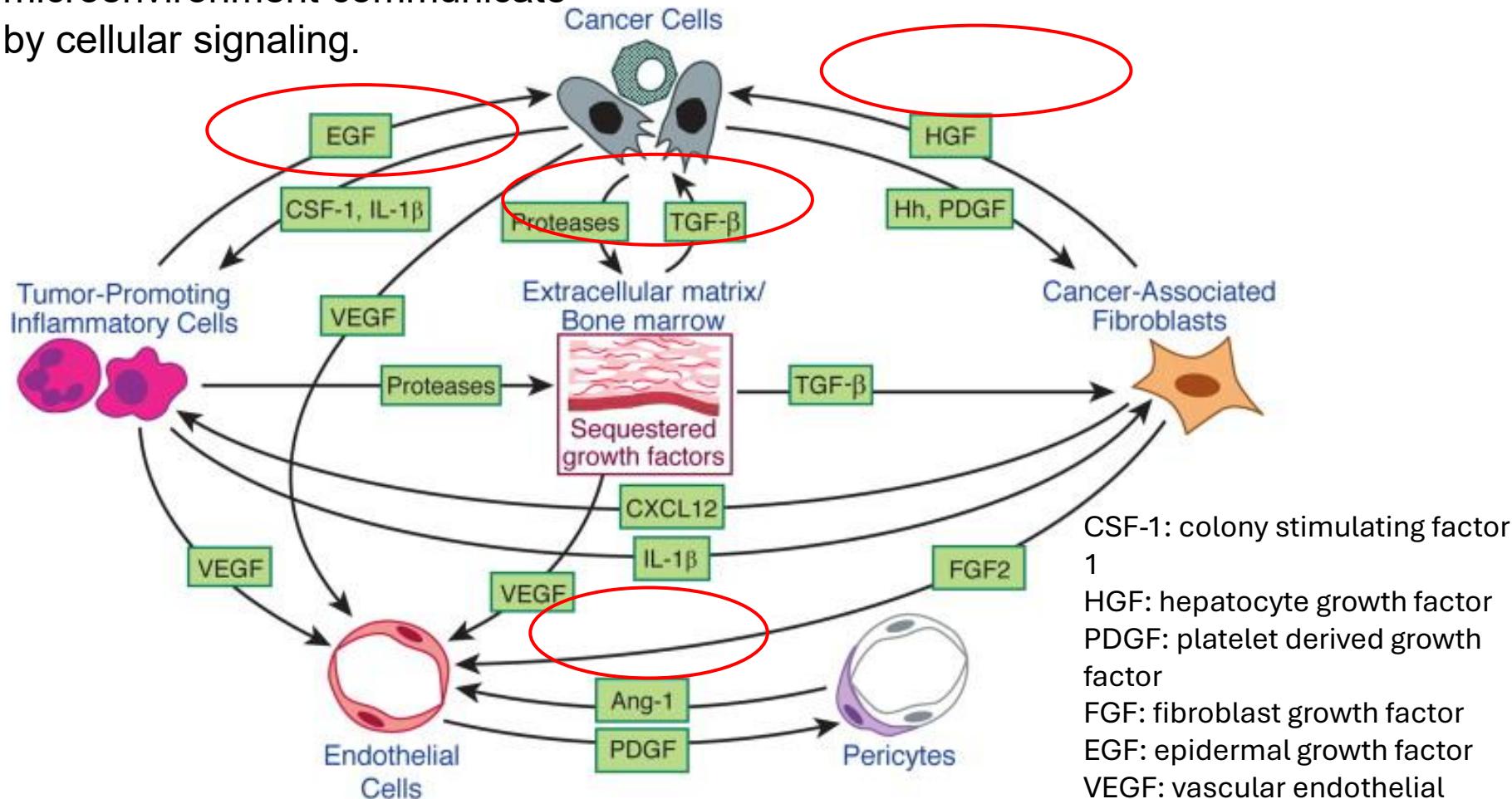
The stromal cell composition varies between tumor types but includes endothelial cells, fibroblasts, adipocytes and stellate cells.



The tumor microenvironment orchestrates angiogenesis, proliferation, invasion and metastasis through the secretion of **growth factors** and **cytokines**.

Communication by cellular signaling

Cell types in the tumor microenvironment communicate by cellular signaling.



CSF-1: colony stimulating factor 1
HGF: hepatocyte growth factor
PDGF: platelet derived growth factor
FGF: fibroblast growth factor
EGF: epidermal growth factor
VEGF: vascular endothelial growth factor
IL-1: interleukin 1

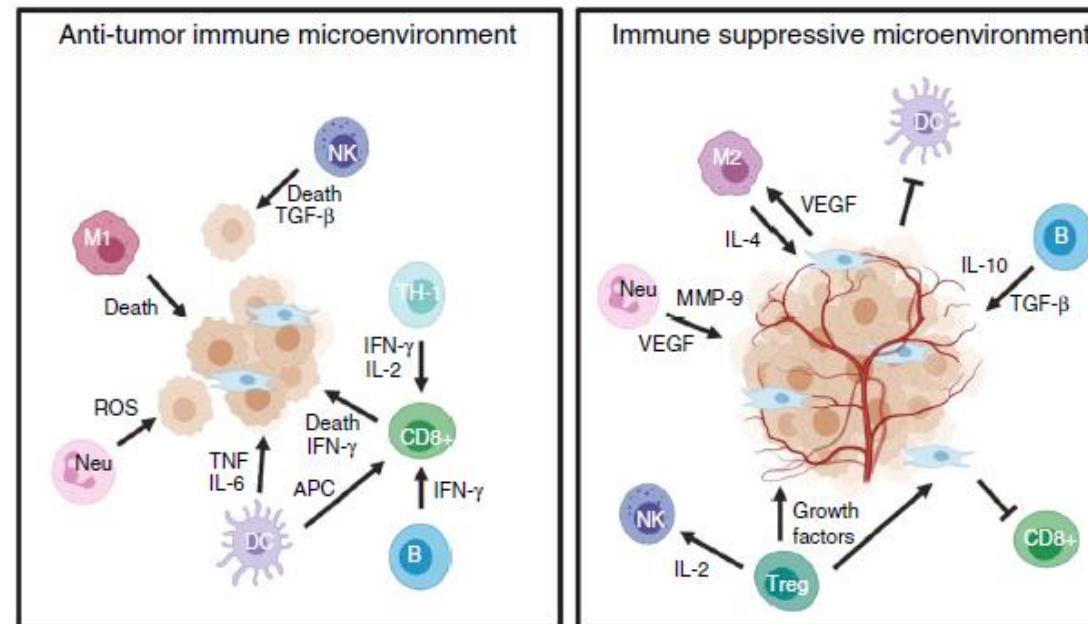
Hanahan & Weinberg,
Cell 144, 646–674 (2011)

Tumor microenvironment (II)

Early in tumor growth, a dynamic and reciprocal relationship develops between cancer cells and components of the tumor microenvironment that supports cancer cell survival, local invasion and metastatic dissemination.

To overcome a hypoxic (low oxygen supply) and acidic microenvironment, the tumor microenvironment coordinates a program that promotes **angiogenesis** to restore oxygen and nutrient supply and remove metabolic waste (**see slide 17**).

Tumors become infiltrated with diverse adaptive and innate immune cells that can perform both pro- and antitumorigenic functions.



Anderson & Simon,
Current Biology 30, R921–R925 (2020)
Cellular Programs

Immune cells

Immune cells are critical components of the tumor microenvironment.

Persistent **inflammation** due to chronic infection is a common mechanism underlying tumor formation in several types of cancer, including colorectal, hepatocellular and cervical cancer.

Broadly, immune cells fall into 2 categories: adaptive immune cells and innate immune cells. **Adaptive immunity** is activated by exposure to specific antigens and uses an **immunological memory** to ‘evaluate’ the threat and enhance immune responses.

T cells, B cells and natural killer (NK) cells belong to the adaptive immune response.

Innate immunity is a non-specific defense mechanism that comes into play within hours after a foreign antigen enters the body.

Cells that carry out an innate immune response include **macrophages, neutrophils and dendritic cells**.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)

Landscape of immune cells

The immune landscape within the tumor microenvironment falls into three main categories: immune infiltrated, immune excluded, and immune silent.

In an **immune infiltrated tumor**, immune cells (such as cytotoxic T cells) are homogeneously distributed throughout the tumor indicating an active immune response.

Alternatively, some tumors are classified as **immune excluded**.

In these cases T cells are only located at the periphery of the tumor and have not infiltrated the tumor microenvironment.

Finally, some tumors are categorized as '**immune silent**' and completely lack immune cell infiltrates, indicating no immune response to the tumor.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)

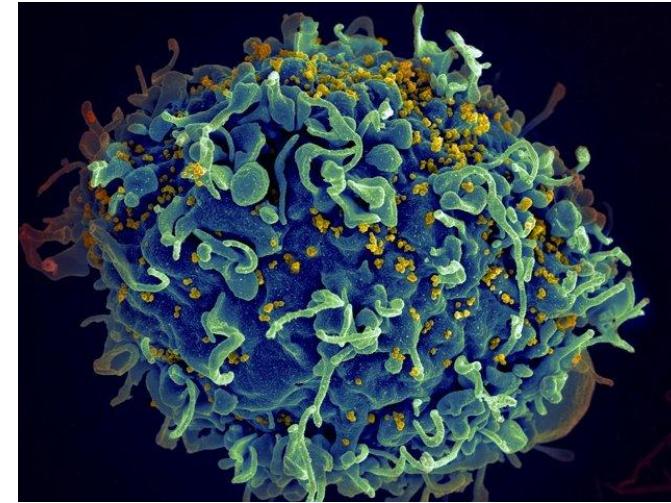
Immune cells: regulatory T-cells

Regulatory T cells (Tregs) are normally required to suppress inflammatory responses and control autoimmunity.

In the context of the tumor microenvironment, Tregs are ubiquitous and promote tumor development and progression by dampening anti-tumor immune responses..

E.g. Tregs secrete the cytokine interleukin 2 (IL-2), which modulates NK cell homeostasis and function.

Additionally, Tregs directly support the survival of cancer cells through the secretion of growth factors, and indirectly through interaction with stromal cells such as fibroblasts and endothelial cells.



Anderson & Simon,
Current Biology 30, R921–R925 (2020)
Picture: blog.hemacare.com

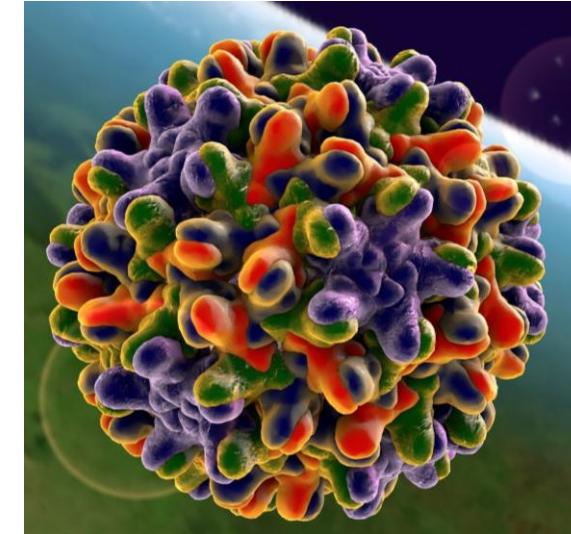
T cells (I)

Each T cell develops its own T-cell receptor (TCR) that recognizes a specific antigen. Within the tumor microenvironment, there are several distinct populations of T cells that influence tumorigenesis.

Cytotoxic T cells (CD8+) (abbreviated as CTLs) detect abnormal tumor antigens expressed on cancer cells and target them for destruction.

The presence of CTLs in the tumor microenvironment is often associated with a **positive prognosis** in cancer patients.

CTLs kill tumor cells and suppress angiogenesis through the secretion of interferon gamma (IFN- γ).



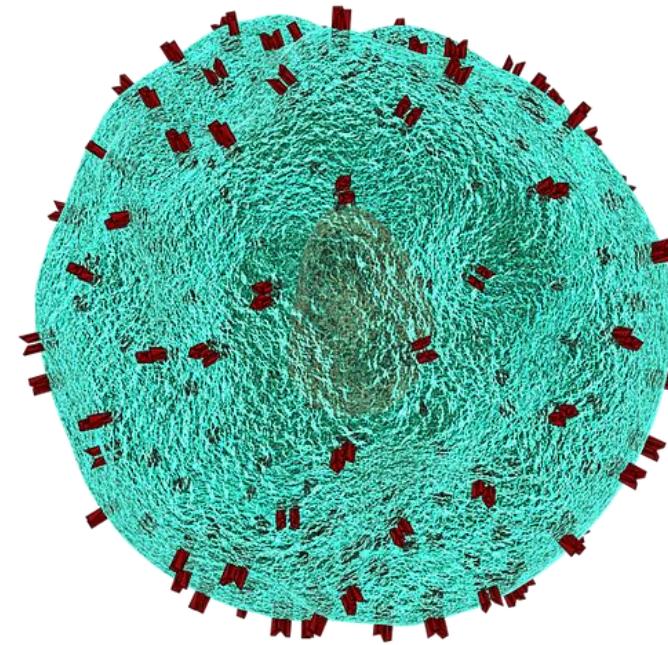
Anderson & Simon,
Current Biology 30, R921–R925 (2020)
Picture: blog.hemacare.com

T cells (II)

CD4+ T cells differentiate into a variety of subtypes and thus coordinate a wide range of immune responses within the context of the tumor microenvironment.

T helper 1 (Th-1) cells are proinflammatory CD4+ T cells that support CD8+ cells through the secretion of IL-2 and IFN- γ .

Increased levels of Th-1 cells within the tumor microenvironment are also associated with **positive outcomes** in many types of cancer.

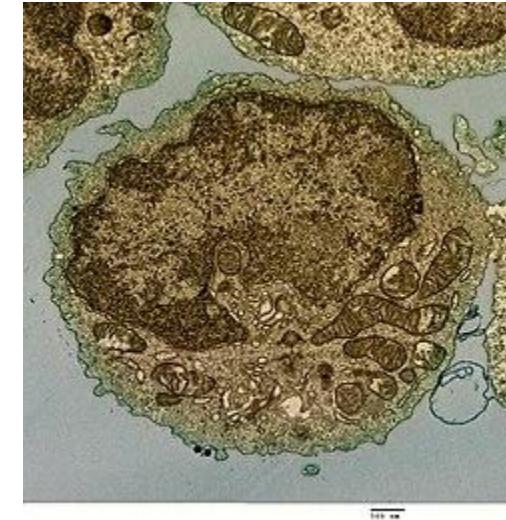


Anderson & Simon,
Current Biology 30, R921–R925 (2020)
Picture of T-helper cell: pixabay.com

B cells

B cells, also known as B lymphocytes, are a type of white blood cell of the lymphocyte subtype.

The **antitumorigenic** roles of B cells include antigen-presentation to T cells, anti-tumor antibody production and secretion of cytokines, like IFN- γ , that promote cytotoxic immune responses.



Alternatively, B cells can have **protumorigic** effects. Their presence in the tumor microenvironment can be predictive of **poor outcome** in bladder cancer, prostate cancer, and renal cell carcinoma.

Similar to Tregs, regulatory B cells promote tumor aggression through production of cytokines (including IL-10 and transforming growth factor beta (TGF- β)) that promote immune suppressive phenotypes in macrophages, neutrophils, and cytotoxic T cells.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)
Picture of B cell: Wikipedia.org

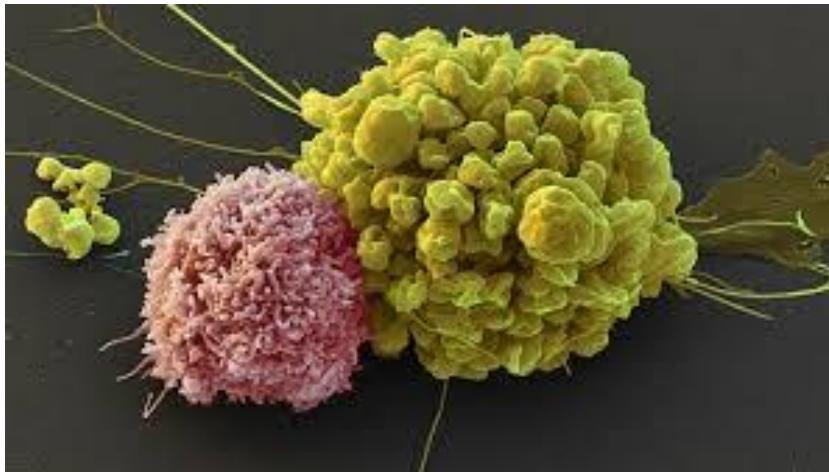
Natural killer cells

Natural killer cells typically patrol the bloodstream, seeking out virally infected host cells and tumor cells.

Functionally, natural killer cells can be broken down into two classes, those that directly participate in cell-mediated killing of tumor cells and those that secrete inflammatory cytokines.

Natural killer cells are highly efficient at killing tumor cells within the circulation and can participate in blocking metastasis, but they are less efficient at killing within the tumor microenvironment.

NK cell (left) attacks cancer cell (right)



Anderson & Simon,
Current Biology 30, R921–R925 (2020)
Picture: <https://nature.com>

Macrophages

Macrophages are critical components of the innate immune system that modulate immune responses through pathogen phagocytosis and antigen presentation.

In addition, macrophages are critical in wound healing and tissue repair.

Monocyte derived macrophages can be categorized as either

- **inflammatory M1 macrophages**, which phagocytize and kill cells, or
- **immune-suppressive M2 macrophages**, which participate in wound healing.

Both classes of macrophages can be found within a tumor. Yet, the tumor microenvironment promotes the M2 phenotype through hypoxia and the secretion of cytokines (such as IL-4) to support tumor growth and progression.

Certain tumor types can be heavily infiltrated with macrophages, which can comprise up to 50% of a tumor's mass.

Typically, high macrophage infiltration is associated with **poor patient prognosis** in many types of cancer, such as breast, lung, and gastric cancers.

Often, macrophages are found to surround blood vessels in the tumor microenvironment where they secrete vascular endothelial growth factor (VEGF) and induce new blood vessel formation.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)

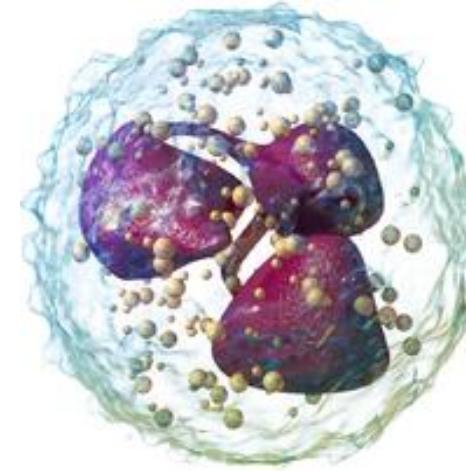
Neutrophils

Neutrophils make up up to 70% of circulating leukocytes and provide the first line of defense against many pathogens.

In the context of cancer, neutrophils can act to either suppress or promote tumor growth, depending on tumor type and stage of development.

As a tumor begins to grow, neutrophils are recruited to the tumor microenvironment and promote inflammation through release of cytokines and reactive oxygen species that promote tumor cell apoptosis.

Later in tumor development, neutrophils promote tumor growth through modification of the extracellular matrix, releasing VEGF and producing matrix metalloproteinase (MMP)-9 to stimulate angiogenesis and, ultimately, tumor progression and local invasion.



Neutrophils contain a nucleus divided into 2–5 lobes.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)
picture: Wikipedia.org

Dendritic cells

Dendritic cells play a critical role in the immune system: they recognize, capture and **present antigens** to T cells at secondary lymphoid organs (such as lymph nodes). Such antigens could e.g. result from pathogen infection.



The fate of dendritic cells in the tumor microenvironment is shaped by cues that promote either an anti-tumor immune response or tolerance.

Dendritic cells are inherently programmed to have an anti-tumorigenic function in the body, but the tumor microenvironment can co-opt dendritic cells to support tumor progression.

Specifically, cytokines secreted from the tumor microenvironment trigger dendritic cells to tolerate the presence of tumor cells and block the induction of an immune response.

Anderson & Simon, Current Biology 30, R921–R925 (2020) Picture: Sriram Subramaniam, National Cancer Institute (NCI) and Donny Bliss, National Library of Medicine (NLM).

Stromal cells

Cancer cells also recruit supporting cells from nearby endogenous tissue stroma to promote critical steps in tumor formation.

Stromal cell composition can vary significantly between tumor types and include vascular endothelial cells, fibroblasts, adipocytes and stellate cells.

Once recruited to the tumor microenvironment, stromal cells secrete many factors that influence angiogenesis, proliferation, invasion, and metastasis.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)

Endothelial cells (I)

Vascular endothelium is a thin monolayer of endothelial cells that help to orchestrate the formation of blood vessels.

Vascular endothelium separates circulating blood from tissues. Besides, it also delivers water and nutrients, maintains metabolic homeostasis, carries immune cells and participates in the formation of new blood vessels.

During the initial stages of tumor development, cancer cells rely on passive diffusion for gas exchange and the transport of nutrients.

Once tumors reach 1–2 mm³ in volume, insufficient oxygen and a build-up of metabolic waste results in the tumor microenvironment becoming hypoxic and acidic.

To overcome this, tumors must develop their own blood supply.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)

Endothelial cells (II)

Specifically, hypoxia-inducible factors initiate blood vessel sprouting by **instructing endothelial cells to secrete proangiogenic factors** such as platelet derived growth factor (PDGF), epidermal growth factor (EGF) and VEGF.

VEGF stimulates migration of endothelial cells to form new blood vessel lumens.

Endothelial cells are also critical in promoting cancer cell migration, invasion and metastasis. They are highly plastic in nature and can change cell fate.

During tumor progression, endothelial cells undergo what is called the '**endothelial–mesenchymal transition**' (EMT) to become **cancer-associated fibroblasts** (CAFs).

This transition is organized by TGF- β and bone morphogenetic protein (BMP), and leads to loss of cell-to cell connections, detachment and elongation, enhanced migration and loss of endothelial properties.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)

Endothelial cells (III)

Cancer associated fibroblasts are critical in stimulating migration and invasion of tumor cells.

Metastasis is a multistep process that involves translocation of cancer cells from the primary tumor microenvironment to distant locations.

Tumor cells must first escape the primary tumor site and enter the vasculature in a process known as **intravasation**.

During intravasation, tumor cells adhere to endothelial cells and this interaction changes the endothelial barrier, allowing tumor cells to migrate between two endothelial cells.

In addition, blood vessels formed in the tumor microenvironment are usually immature and lack proper cell-to-cell connections, enabling cancer cells to transverse the vasculature.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)

Cancer associated fibroblasts (I)

Cancer-associated fibroblasts (CAFs) are a major component of the tumor stroma and play a critical role in facilitating crosstalk between cancer cells and tumor microenvironment (see also <https://www.nature.com/articles/s41568-019-0238-1>).

CAFs are often (not always) derived from fibroblasts.

Upon injury, fibroblasts that normally reside within tissues can become reversibly induced to form **myofibroblasts**, which actively participate in wound healing.

Myofibroblasts are activated by TGF- β signaling and develop characteristics important in wound healing, such as proliferation, contractile properties, secretory phenotypes and ECM formation.

Tumors have been termed ‘wounds that never heal’.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)

Cancer associated fibroblasts (II)

In the tumor microenvironment, cancer and stromal cells secrete factors such as TGF- β , PGDF, and fibroblast growth factor 2 (FGF2) to convert fibroblasts into CAFs.

A build-up of CAFs within the tumor microenvironment is often associated with **poor prognosis** in many cancer types.

Despite this association, CAFs have been shown to both promote and restrain tumorigenesis.

Within the tumor microenvironment, CAFs produce the majority of extracellular components, including growth factors, cytokines and extracellular matrix components.

CAFs shape the tumor microenvironment in 4 main ways:

tumor proliferation and metastasis, neoangiogenesis, ECM remodeling and immunosuppression.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)

Cancer associated fibroblasts (III)

In tumors of epithelial origin (80-90%, see comments), the epithelial– mesenchymal transition (EMT) is a critical step in metastasis. During EMT, epithelial cells lose cell polarity and cell-to-cell adhesions and gain migratory and invasive phenotypes.

CAFs control metastasis by secreting TGF- β , which is required for the EMT and angiogenesis.

To facilitate migration of cancer cells through the tumor microenvironment, cancer-associated fibroblasts secrete the matrix metallo proteinase MMP-3, which degrades E-cadherin to promote cancer cell invasion → MMPs are candidate drug targets.

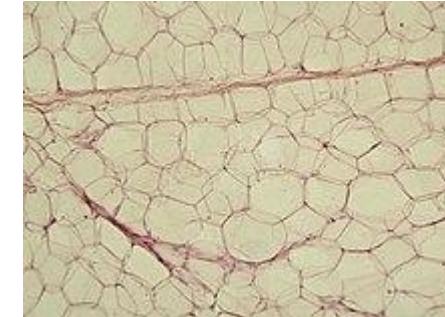
The ECM is also an important source of VEGF, which can be released by MMP-13 to promote angiogenesis.

In general, cancer-associated fibroblasts promote an immunosuppressive phenotype through the production of immune-modulatory chemokines and cytokines.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)

Adipocytes

Adipocytes are specialized cells within the body that regulate energy balance and are responsible for storing excess energy as fat.



Adipocytes exert their effects on the tumor microenvironment through secretion of metabolites, enzymes, hormones, growth factors and cytokines.

Within the context of the tumor microenvironment, adipocytes are in a dynamic and reciprocal relationship with tumor cells to **support tumor progression**.

Breast tissue is largely composed of white adipose tissue; therefore, adipocytes are a critical player in the breast cancer tumor microenvironment.

Anderson & Simon,
Current Biology 30, R921–R925 (2020)
Picture: Wikipedia.org

Non-cellular components: extracellular matrix (ECM)

The ECM is composed of collagen, fibronectin, elastin, and laminin. The ECM is an important molecular component of the tumor microenvironment because it provides a physical scaffold for cells and plays a key role in promoting tumor cell dissemination.

Solid tumors contain large ECM deposits that constitute up to 60% of tumor mass.

Large **collagen deposits**, together with a high percentage of fibroblast infiltration, result in desmoplasia, which is strongly linked to **poor patient prognosis**.

Many cells within the tumor microenvironment secrete components of the ECM, although cancer-associated fibroblasts are the predominant source.

The ECM is a depot for cytokines and growth factors, which are released by proteases like the MMPs. E.g., the ECM can be a deposit for proangiogenic factors, like VEGF, FGF, PDGF, TGF- β .

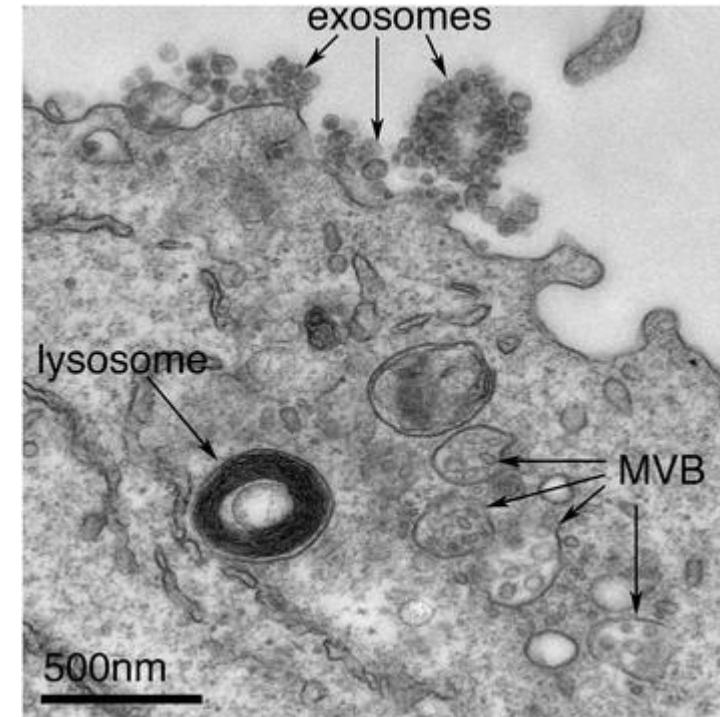
Extracellular components: exosomes

Exosomes are **microvesicles** that range in size from 30–200 nm.

Their contents reflect the cells from which they were derived, including protein, RNA, DNA and lipids.

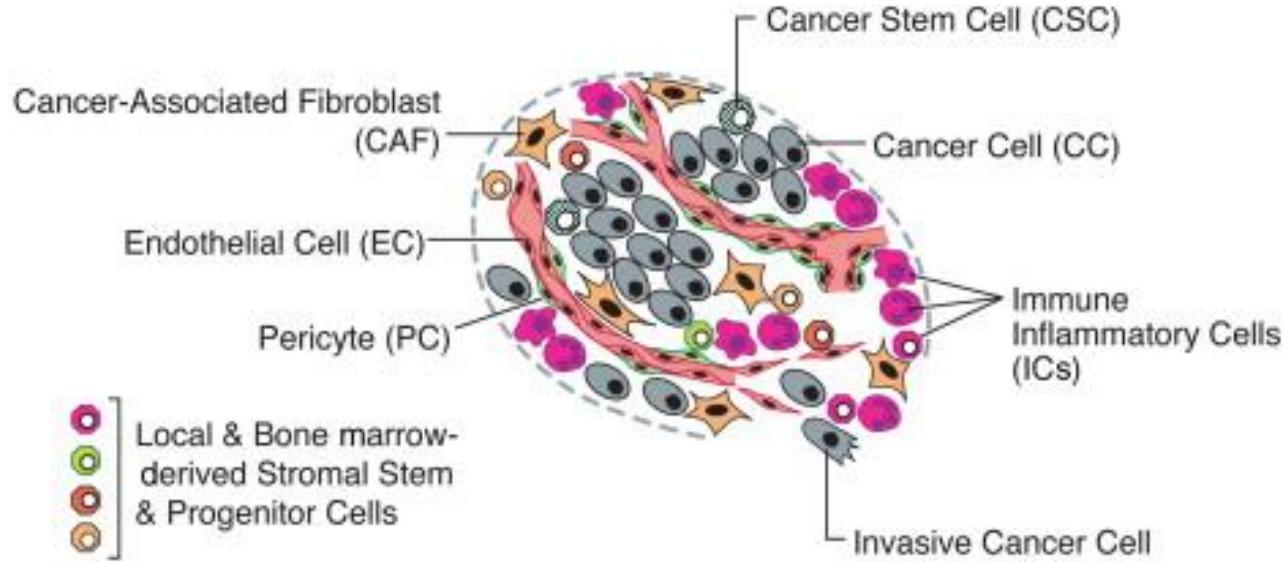
Within the tumor microenvironment, exosomes play a critical role in facilitating cross-talk between cancer cells and stromal cells.

Exosomes have been shown to promote inflammation, tumor progression, angiogenesis, and metastasis within the tumor microenvironment.



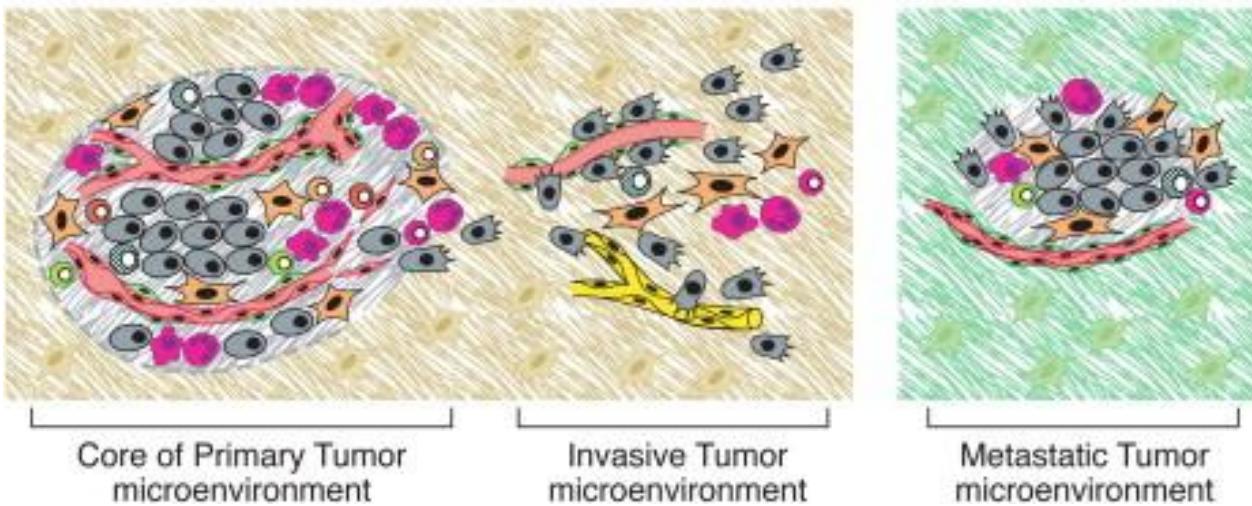
Anderson & Simon,
Current Biology 30, R921–R925 (2020)
Picture: <https://bmcbiol.biomedcentral.com/articles/10.1186/s12915-016-0268-z>

Tumor microenvironment



(TOP) Assemblage of different cell types that make up most solid tumors.

(BOTTOM) Different stages of the tumor face changing microenvironments.



Hanahan & Weinberg,
Cell 144, 646–674 (2011)

Summary

Cancer arises from mutations accruing within cancer cells, but both disease progression and responses to therapy are strongly modulated by non-mutant cells within the tumor microenvironment.

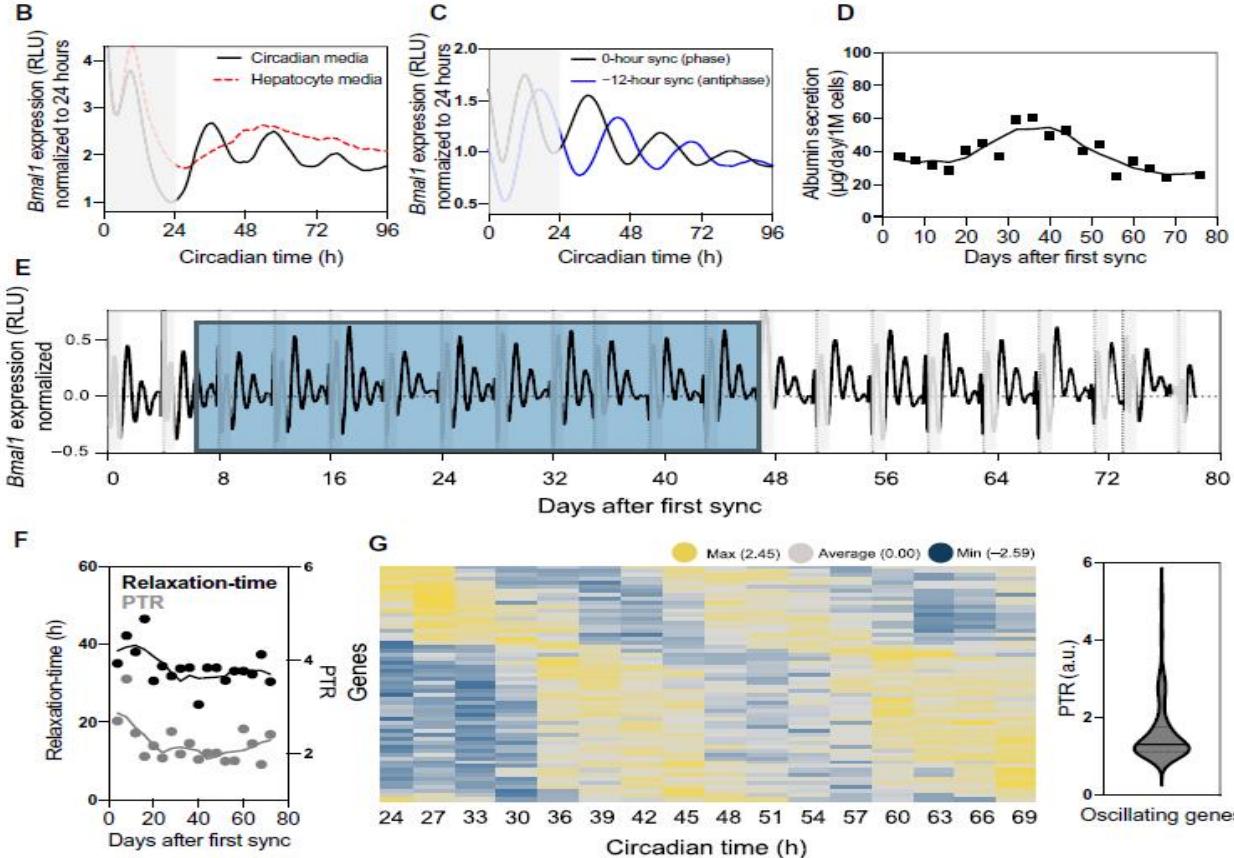
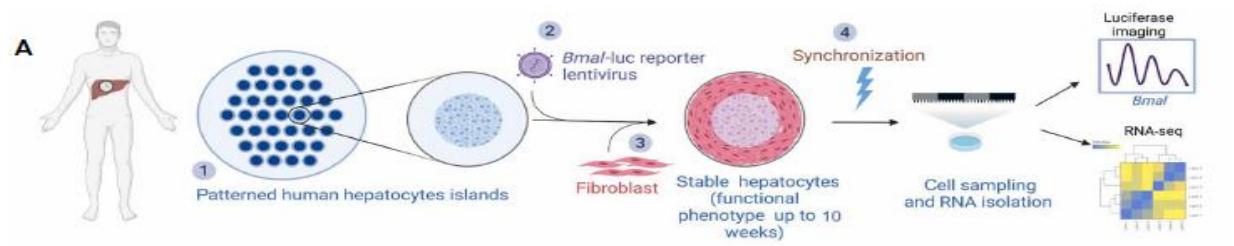
Cancer cells can hijack surrounding immune cells and epithelial cells for their own benefit, e.g. in securing oxygen supply and waste removal.

Remodelling of the ECM is important for EMT and metastasis.

CAFs are perhaps the most effective cell within the tumor microenvironment at depositing and remodeling the ECM.

Immunotherapy is a type of cancer treatment that helps our immune system in fighting cancer, see e.g.

<https://www.cancer.gov/about-cancer/treatment/types/immunotherapy>



Primary human hepatocytes display a circadian rhythm in culture.

(A) Experimental work-flow. PHHs are seeded on collagen-coated plates to create hepatocyte islands.
 (2) Transduction with *Bmal1*-luc reporter lentiviral particles.
 (3) Seeding of mouse fibroblasts, 24 hours later.
 (4) Synchronization and monitoring of luciferase-based *Bmal1* cyclic expression.

(B) Visualization of circadian rhythm in real time. Circadian rhythm of transduced PHH with *Bmal1*-luc reporter was monitored in real time at 20-s sampling resolution by light emission of luciferase over 4 days in free-running conditions. To synchronize the hepatocytes, the cultures were placed in specialized circadian (black) or hepatocyte media (red). For 96 hours, their circadian rhythm was monitored in real time. Circadian time is defined as hours after synchronization by media change.

(C) Anti-phasic *Bmal1* expression of PHH cultures. Synchronization of two sets of PHH cultures was performed with a circadian medium change 12 hours apart, allowing them to free-run under constant conditions for 96 hours. (D) Daily albumin secretion in PHH over 10-week period.

(E) Circadian rhythm of *Bmal1*-luc reporter-transduced PHH was observed over a period of 10 weeks. Between each 96-hour monitoring of the circadian rhythms, a new synchronization of the cultures was triggered by circadian media exchange.

(F) Hepatocytes show stable circadian relaxation-time and peak-to-trough ratio (PTR) of *Bmal1* gene expression over a 10-week period.

(G) Transcriptomic analysis of oscillating transcripts. Heatmap representation of oscillating transcripts ordered by the time of the oscillation (columns),

"Primary human hepatocytes display a circadian rhythm in culture," illustrates the experimental setup and key findings demonstrating autonomous circadian rhythms in primary human hepatocytes (PHHs) in an *in vitro* system.

Here's a summary of what each panel in Fig. 1 shows:

- **A: Experimental Workflow:** This panel outlines the steps involved in establishing and monitoring the circadian rhythm in PHH cultures.

1. PHHs are seeded on collagen-coated plates to form micropatterned hepatocyte islands.
2. They are then transduced with a **Bmal1-luc reporter lentiviral particle**, which expresses destabilized luciferase under the control of a Bmal1 promoter to monitor activity in real time.
3. Mouse 3T3-J2 fibroblasts are seeded between the hepatocyte islands to stabilize the functional phenotype of the PHHs.
4. The cultures are synchronized, typically by a medium exchange, and Bmal1 cyclic expression is monitored using luciferase activity.

- **B: Visualization of Circadian Rhythm in Real Time:** This graph compares the real-time circadian rhythm of Bmal1 expression in PHHs when cultured in standard hepatocyte medium (red line, not conducive to circadian expression) versus an optimized "circadian medium" (black line). The circadian medium successfully allows for the visualization of cyclic Bmal1 expression over 72 to 96 hours.

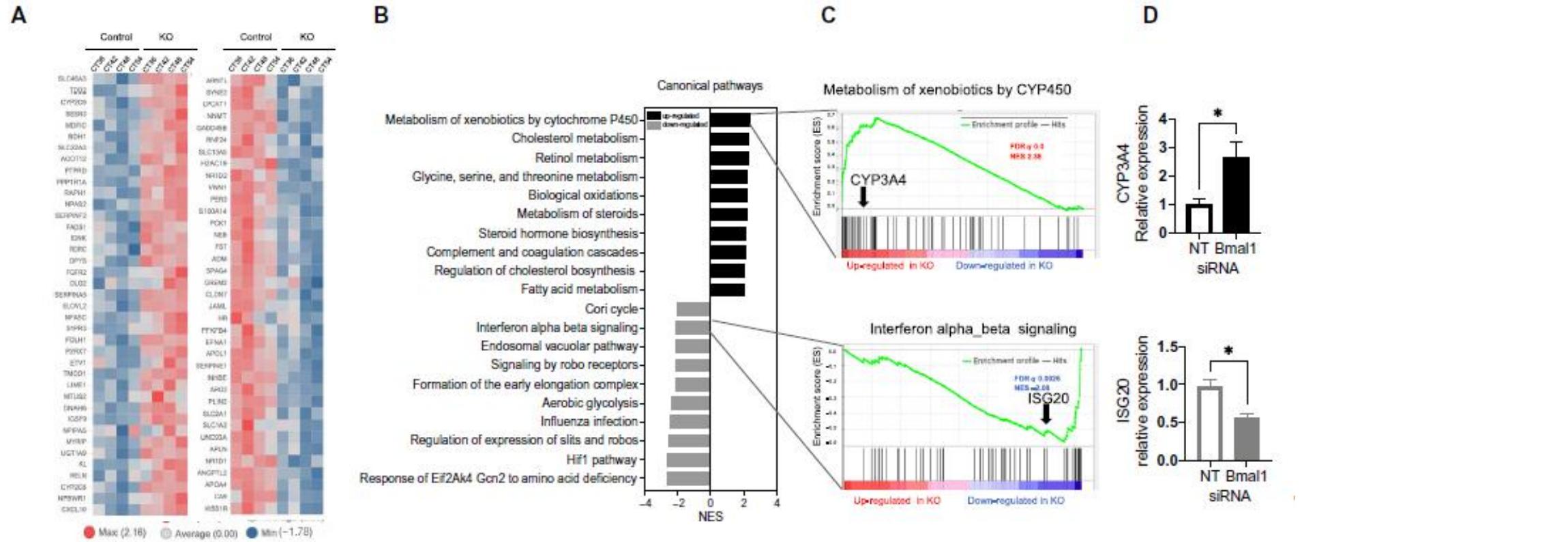
- **C: Anti-phasic Bmal1 Expression of PHH Cultures:** This panel shows that when two sets of PHH cultures are synchronized 12 hours apart, they exhibit oppositely phased rhythms (an "antiphase" relationship) for 96 hours, confirming a hallmark property of circadian rhythms.

- **D: Daily Albumin Secretion in PHH over 10-Week Period:** This graph demonstrates the sustained functional stability of the PHH cultures, indicated by consistent daily albumin secretion levels, for up to a 10-week period when circadian oscillations are present.

- **E: Circadian Rhythm of Bmal1-luc Reporter–Transduced PHH Observed over a Period of 10 Weeks:** This panel illustrates the long-term maintainability of circadian rhythms in the culture system. Although the amplitude of oscillations decreased over time within each cycle, the initial cycling pattern was restored for at least 96 hours after each new synchronization event triggered by circadian media exchange.

- **F: Hepatocytes Show Stable Circadian Relaxation-Time and Peak-to-Trough Ratio (PTR) of Bmal1 Gene Expression over a 10-Week Period:** This graph quantifies the stability of the oscillations over 80 days of culture, showing consistent relaxation-time values (how long it takes for oscillations to decay) and peak-to-trough ratios (fold difference between highest and lowest expression levels).

- **G: Transcriptomic Analysis of Oscillating Transcripts:** This heatmap and corresponding PTR graph visually represent the **59 oscillating genes** identified across two consecutive 24-hour cycles through transcriptomic analyses. The heatmap shows gene expression intensities (low in blue, high in yellow) ordered by their peak expression time, indicating a dynamic transcriptional state across 24 hours. These oscillating genes include core clock genes and those related to inflammation, drug metabolism, and energy homeostasis.



Loss of Bmal1 expression disrupts expression of genes involved in inflammatory signaling and drug metabolism in PHH.

(A) Unclustered heatmaps of genes differentially expressed (logFC of 1, adjusted P of 0.05 cutoffs) in Bmal1-siRNA– treated [knockout (KO)] versus those treated with a nontargeting siRNA construct (NT). CT, circadian time in hours.

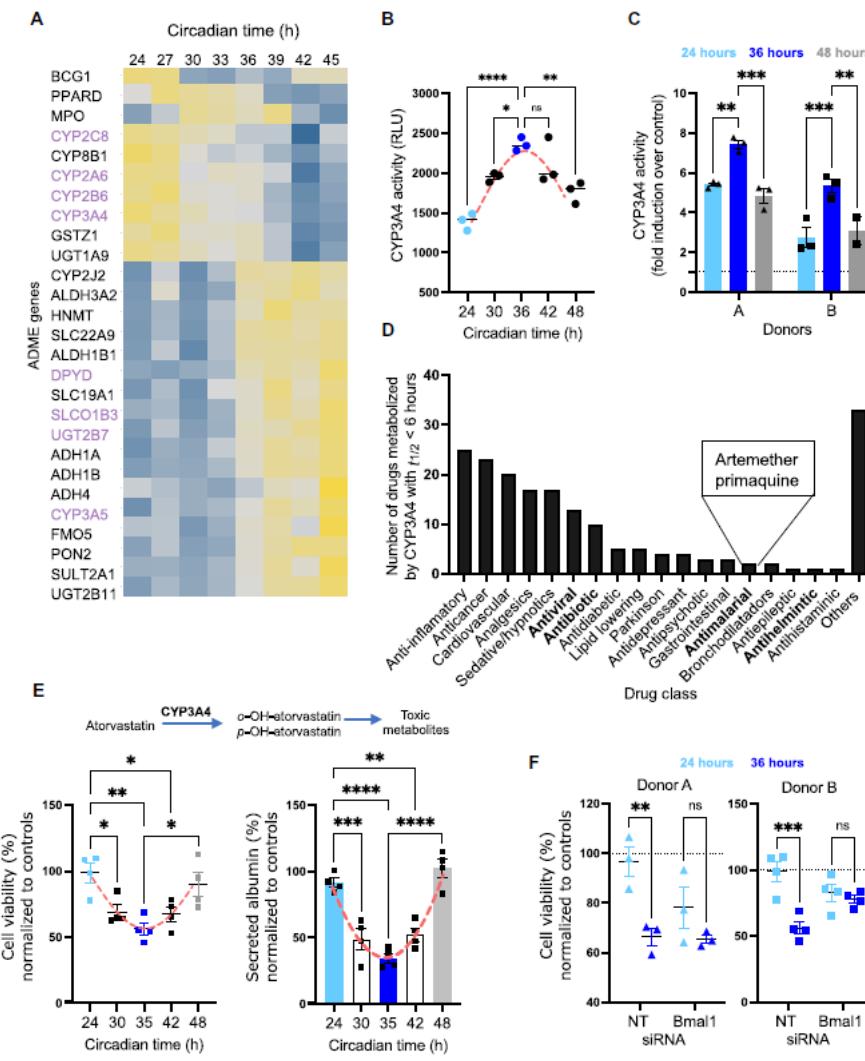
(B) Canonical pathways GSEA. Top 10 statistically differentially expressed gene sets that were up-(black) or down-regulated (gray) in *Bmal1* knockdown cultures from MSigDB. Analysis of the CT24 time point.

(C) Metabolism of xenobiotics by CYP450 (top) and IFN- α or IFN- β signaling (bottom) enrichment profiles. Arrows indicate the localization of *CYP3A4* and *ISG20* in the respective gene sets.

(D) qRT-PCR validation. Up-regulation of *CYP3A4* (top) (black) and down-regulation of *ISG20* (bottom) (gray) are observed in *Bmal1* knockdown cultures. Analysis performed at CT48. FDR, false discovery rate; NES, normalized enrichment score. * P < 0.05.

A: Unclustered heatmaps of genes differentially expressed: This panel visually represents the **genes that are significantly up- or down-regulated** when Bmal1 expression is silenced (knockout, KO) compared to control PHHs treated with a non-targeting siRNA construct (NT). The data are shown at various circadian time (CT) points (CT36, CT42, CT48, and CT54), with differential expression defined by a logFC of 1 and an adjusted P value of 0.05 cutoffs. This highlights global changes in gene expression due to Bmal1 knockdown.

- **B: Canonical pathways Gene Set Enrichment Analysis (GSEA):** This GSEA waterfall plot identifies the **top 10 statistically differentially expressed gene sets** that were either upregulated (black) or downregulated (gray) in Bmal1 knockdown cultures from the MSigDB database. A **substantial upregulation of pathways related to the metabolism of xenobiotics by CYP450 enzymes** is observed. Other dysregulated pathways detected include those involved in inflammatory responses and IFN signaling. This analysis helps to pinpoint the biological processes most affected by Bmal1 disruption.
- **C: Enrichment profiles for specific gene sets:** This panel provides a more detailed view of the enrichment for two key gene sets:
 - **Metabolism of xenobiotics by CYP450 (top):** It shows the enrichment profile, with arrows indicating the localization of **CYP3A4**, a critical enzyme in drug metabolism, within this gene set. Genes on the "leading edge" (far left) of this plot belong to the core drug absorption, distribution, metabolism, and excretion (ADME) gene set, which are directly involved in drug metabolism and clearance.
 - **IFN- α or IFN- β signaling (bottom):** This profile highlights the impact on inflammatory pathways, with arrows pointing to **ISG20** (an IFN-stimulated gene). This panel visually confirms that genes related to drug metabolism and inflammatory/IFN responses are significantly impacted by Bmal1 silencing.
- **D: qRT-PCR validation:** This panel presents **quantitative reverse transcription polymerase chain reaction (qRT-PCR)** validation of key findings from the transcriptomic analysis. It specifically shows:
 - **Up-regulation of CYP3A4 (top)** in Bmal1 knockdown cultures.
 - **Down-regulation of ISG20 (bottom)** in Bmal1 knockdown cultures. These results experimentally confirm that **Bmal1 controls the circadian expression of genes involved in inflammatory signaling and drug metabolism** in the cultured PHHs. The analysis was performed at CT48



PHH exhibit circadian-dependent patterns of drug metabolism.

(A) Transcriptomic analysis of oscillating transcripts involved in drug metabolism. Phase sorted heat map of cycling ADME genes over a 24-hour period. Each vertical column represents a time point (3-hourresolution). Each row is a cycling transcript, colored based on the expression intensities, from low (blue) to high (yellow) ($\text{BHQ} < 0.2$). Expression values are mean-normalized for each gene and are ordered by the peak of expression.

(B) CYP3A4 luminogenic activity assay. An activity-based luminogenic assay was used to determine CYP3A4 enzymatic activity in synchronized hepatocytes every 6 hours, over a 24-hour period.

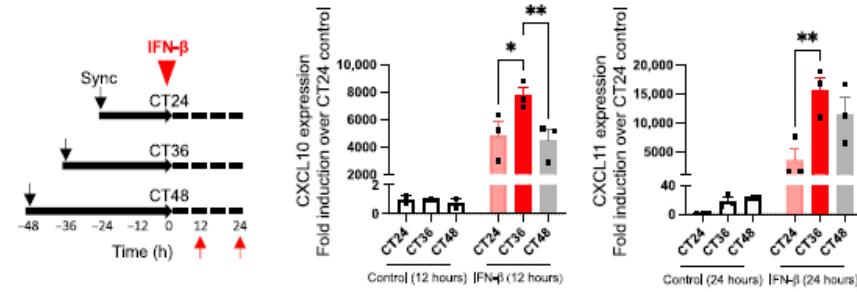
(C) Impact of rifampin induction on CYP3A4 activity. Synchronized cultures at CT24, CT36, and CT48 were dosed with rifampin (10 μM), and 24 hours later, CYP3A4 activity was measured via luminogenic assay.

(D) Putative candidates for chronotherapy. Drugs metabolized by CYP3A4 with $t_1 < 6$ hours.

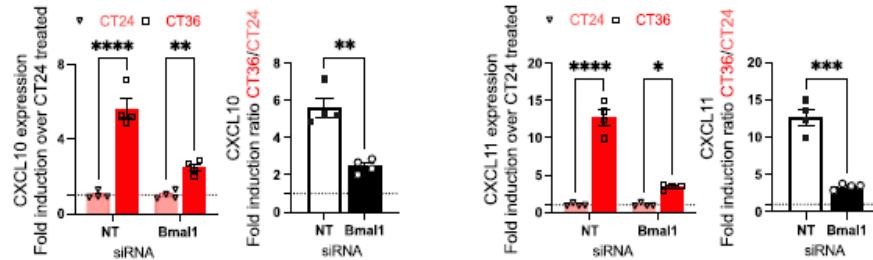
(E) Atorvastatin hepatotoxicity mediated by CYP3A4. Schematic showing the metabolism of atorvastatin by CYP3A4 and its by-products (top). Adenosine 5'-triphosphate (ATP)-based real-time viability assay to determine CYP3A4-metabolized toxicity of atorvastatin in PHH cultures synchronized 5 to 7 hours apart (CT24 versus CT30 versus CT35 versus CT42 versus CT48). Atorvastatin was added (200 μM) and viability was monitored for 2 hours (bottom left). Albumin levels measured 12 hours after dosing with atorvastatin (bottom right). CT24, CT30, CT35, CT42, and CT48 are defined, respectively as, 24, 30, 35, 42, and 48 hours after synchronization by media change.

(F) Circadian-dependent Atorvastatin hepatotoxicity in Bmal1-silenced hepatocytes. Atorvastatin was added at 200 μM , and real-time viability was measured for 2 hours in two independent

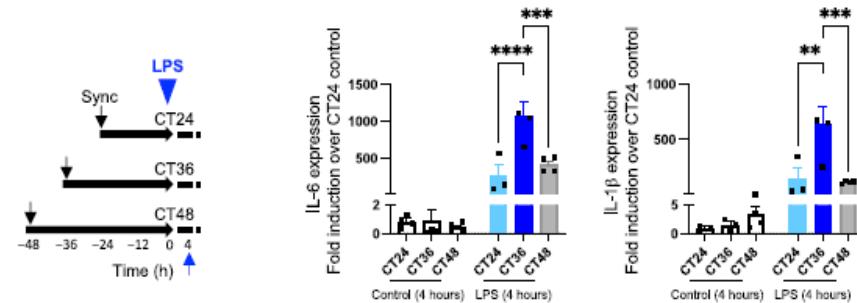
A



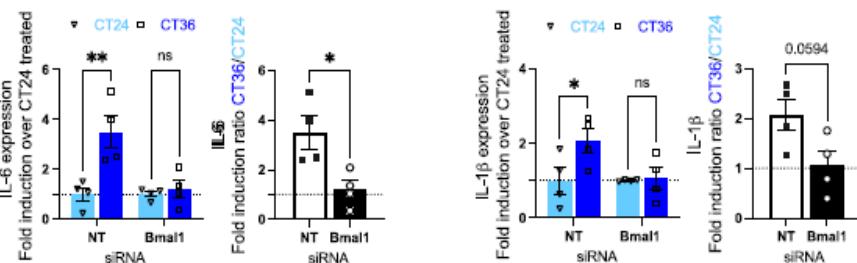
B



C



D



Circadian control of the hepatocyte inflammatory response.

(A) Circadian-dependent inflammatory response to an immune challenge mimicked by IFN- β exposure.

Schematic of induction via immune stimuli challenge (left). Expression of inflammatory cytokines was measured by qPCR. mRNA was isolated from synchronized hepatocyte cultures harvested at 12 and 24 hours after IFN- β treatment (1000 U/ml) at either CT24 (light red), CT36 (dark red), or CT48 (gray). Levels of cytokine mRNA were quantified [relative to glyceraldehyde-3-phosphate dehydrogenase (GAPDH)] and presented in relation to expression levels in hepatocytes harvested at CT24 from control groups [showing induction levels of cells harvested at 12 hours for CXCL10 (middle) or 24 hours for CXCL11 (right) after induction with IFN- β].

(B) Circadian-dependent inflammatory response to IFN- β in Bmal1-silenced hepatocytes.

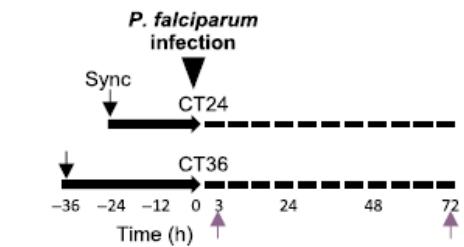
Expression of inflammatory cytokines was measured by qPCR. Levels of cytokine mRNA were quantified (relative to GAPDH) and presented in relation to expression levels in hepatocytes harvested at CT24 from nontargeting or Bmal1 siRNA IFN- β -treated groups.

(C) Inflammatory response to circadian-dependent LPS immune challenge. Schematic of induction with immune challenge (left). Expression of inflammatory cytokines was measured by qPCR.

mRNA was isolated from synchronized hepatocyte cultures harvested at 4 hours after LPS treatment at either CT24 (light blue), CT36 (dark blue), or CT48 (gray). Levels of inflammatory cytokine mRNA were quantified (relative to GAPDH) and are presented in relation to expression levels in hepatocytes harvested at CT24 from the control group.

(D) Circadian-dependent inflammatory response to LPS in Bmal1-silenced hepatocytes. Expression of inflammatory cytokines was measured by qPCR. Levels of cytokine mRNA were quantified (relative to GAPDH) and presented in relation to expression levels in hepatocytes harvested at CT24 from nontargeting or Bmal1 siRNA LPS-treated groups.

A



B

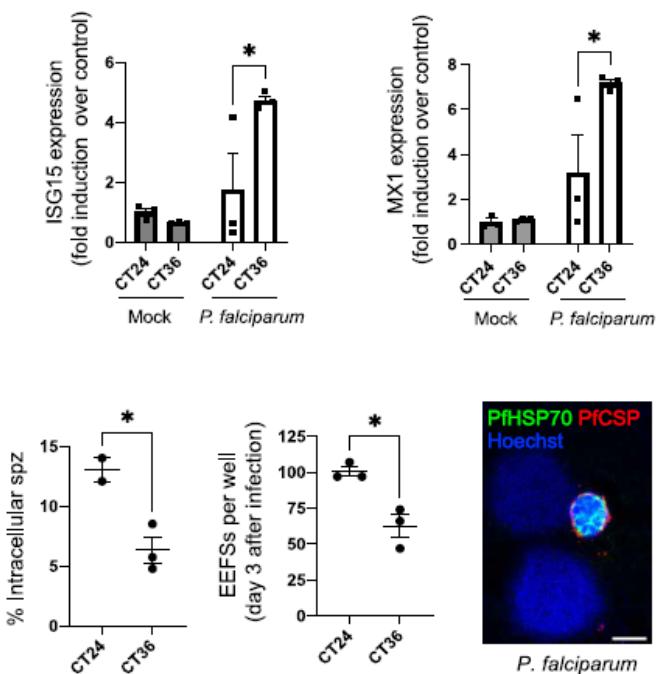
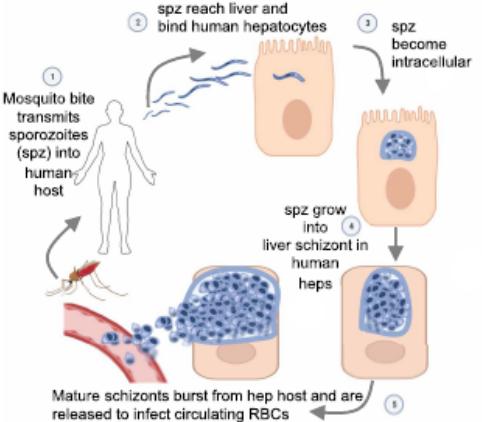
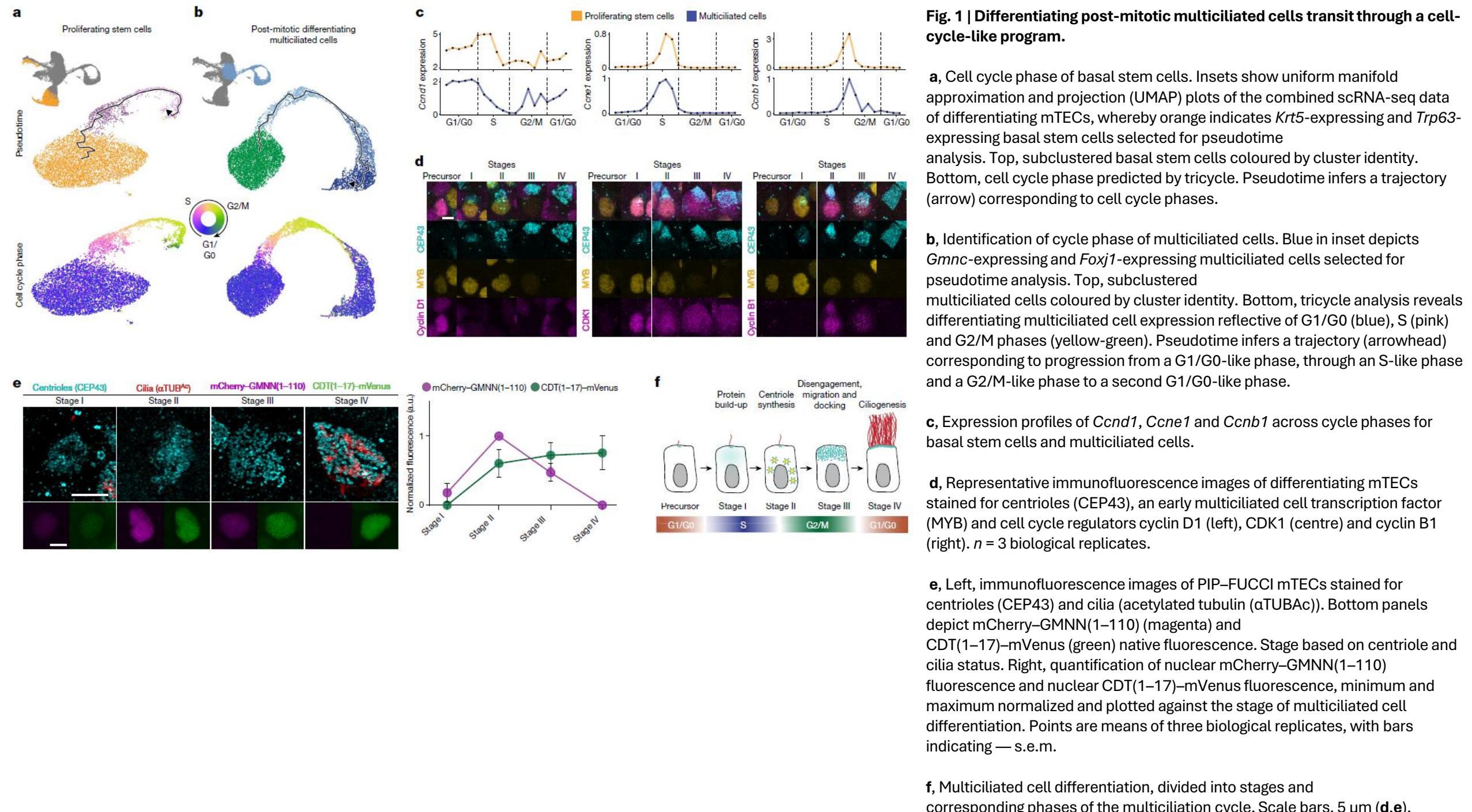
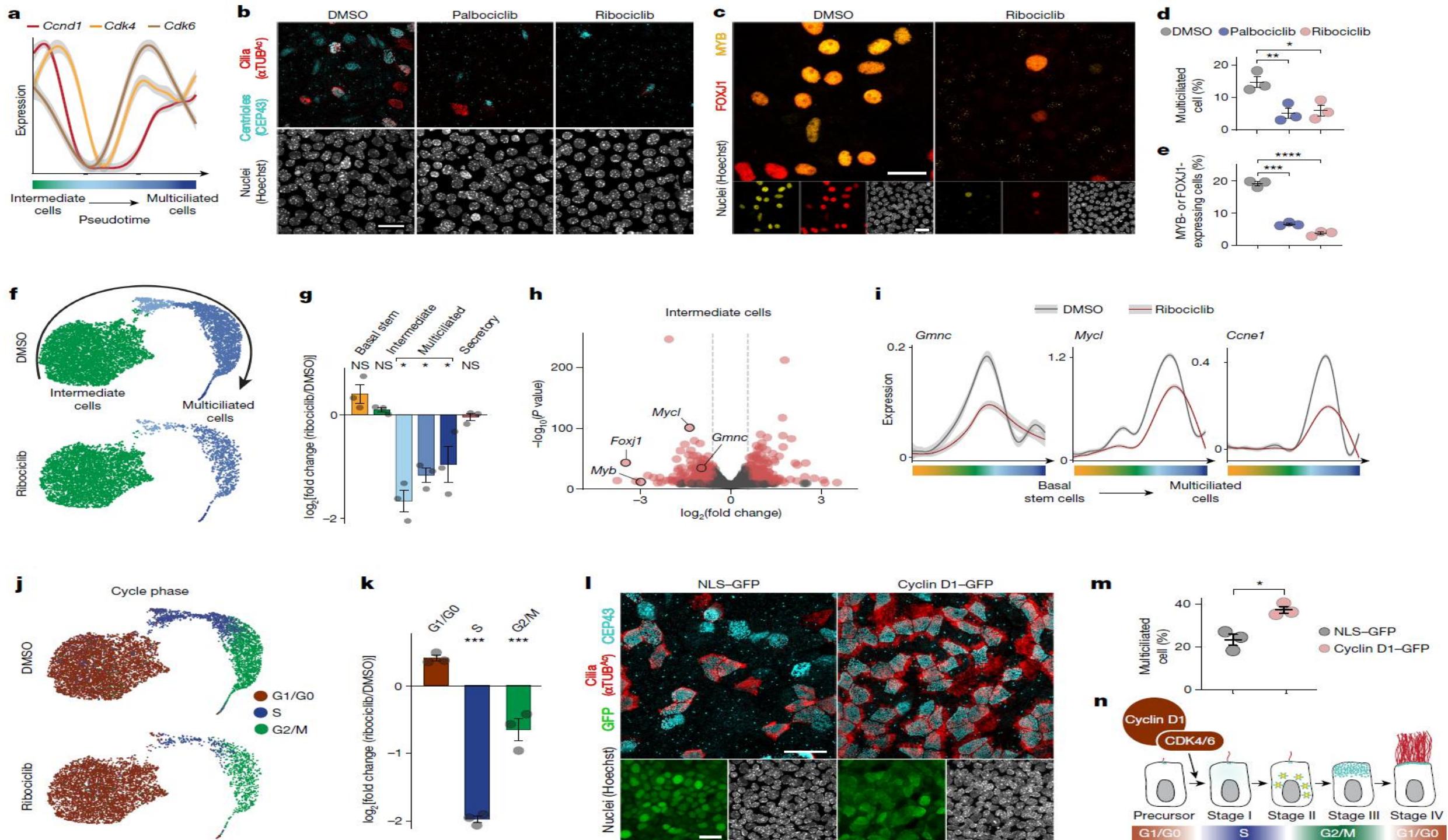


Fig. 5. Circadian control of hepatocyte infection.

(A) Infection of hepatocytes by malaria-causing *P. falciparum* sporozoites. Experimental workflow (left). mRNA was isolated from synchronized hepatocytes cultures harvested at 3 hours after infection at CT24 or CT36. Transcript levels of ISG15 and MX1 mRNA were first normalized relative to GAPDH and are presented in relation to expression levels in mock-infected hepatocytes harvested at CT24 (right). Data are mean \pm SEM. n = 3 independent wells. Statistical significance was determined using a two-way ANOVA. *P < 0.05.

(B) Malaria life cycle schematic depicting the obligate initial liver stage expansion before egress to the blood (left). Quantification of intracellular parasites at 3 hours (corresponding to the transition from stage 2 to 3 in the left hand schematic) and 3 days (corresponding to schematic stages 3 and 4) after infection. Representative image of *P. falciparum* infection (parasite EEFs are identified by anti-H SP70 (green) and anti-circumsporozoite protein (red) staining) in human MPCC s at day 3 (right). Scale bar, 5 μ m. Data are means \pm SEM. n = 3 independent wells. Statistical significance was determined using an unpaired t test. *P < 0.05. (left). RBCs, red blood cells. PfCSP, *P. falciparum* circumsporozoite protein.





Cyclin D1-CDK4/6 initiates multiciliated cell differentiation.

- a. Minimum and maximum normalized scRNA-seq expression across multiciliated cell differentiation pseudotime. Grey bars indicate 95% confidence intervals. Coloured x-axis indicates cluster identity.
- b. mTECs treated with DMSO, palbociclib or ribociclib and stained for centrioles (CEP43), cilia (α TUBAc) and nuclei.
- c. mTECs treated with DMSO or ribociclib and stained for MYB, FOXJ1 and nuclei.
- d. Percentage of multiciliated cells after DMSO, palbociclib or ribociclib treatment. Horizontal lines indicate means \pm s.e.m. of 3 biological replicates, ** $P = 0.0118$, * $P = 0.0183$ (one-way analysis of variance (ANOVA) with Dunnet's correction).
- e. Percentage of MYB-expressing or FOXJ1-expressing cells after DMSO, palbociclib or ribociclib treatment. Horizontal lines indicate means \pm s.e.m. of 3 biological replicates, *** $P = 0.000002$, **** $P = 0.0000007$ (one-way ANOVA with Dunnet's correction).
- f. scRNA-seq UMAP of mTEC intermediate and differentiating multiciliated cells after DMSO or ribociclib treatment. Arrow indicates differentiation trajectory.
- g. Change in cell cluster proportion after ribociclib treatment. Bars indicate means \pm s.e.m. of 3 biological replicates, *false discovery rate (FDR) < 0.02 (two-tailed Bayes quasi-likelihood F-test with Benjamini–Hochberg correction).
- h. Genes differentially expressed by ribociclib-treated intermediate cells compared with DMSO-treated cells. Red, fold-change > 1.5 and $P < 0.0005$ (two-tailed Wald test with Benjamini-Hochberg correction).
- i. scRNA-seq expression across multiciliated cell differentiation pseudotime (DMSO or ribociclib). Grey bars indicate 95% confidence intervals. Colours indicate cluster identity.
- j. scRNAseq UMAP of tricycle-based cycle phase after DMSO or ribociclib treatment.
- k. Change in cycle phase proportion after ribociclib treatment. Bars indicate means \pm s.e.m. of 3 biological replicates, log2(fold-change) < -0.5 and FDR < 0.0008 (two-tailed moderated t-test with Benjamini–Hochberg correction).
- l. mTECs expressing NLS-GFP or cyclin D1-GFP stained for centrioles (CEP43) and cilia (α TUBAc). Bottom panels show GFP and nuclei staining.
- m. Proportion of multiciliated cells after cyclin D1-GFP expression. Horizontal lines indicate means \pm s.e.m. of 3 biological replicates, * $P = 0.0102$ (unpaired two-tailed t-test).
- n. Cyclin D1-CDK4/6 promotes the transition of precursor cells from the G1/G0-like phase to S-like phase of the multiciliation cycle. Scale bars, 10 μ m (b,c,l).

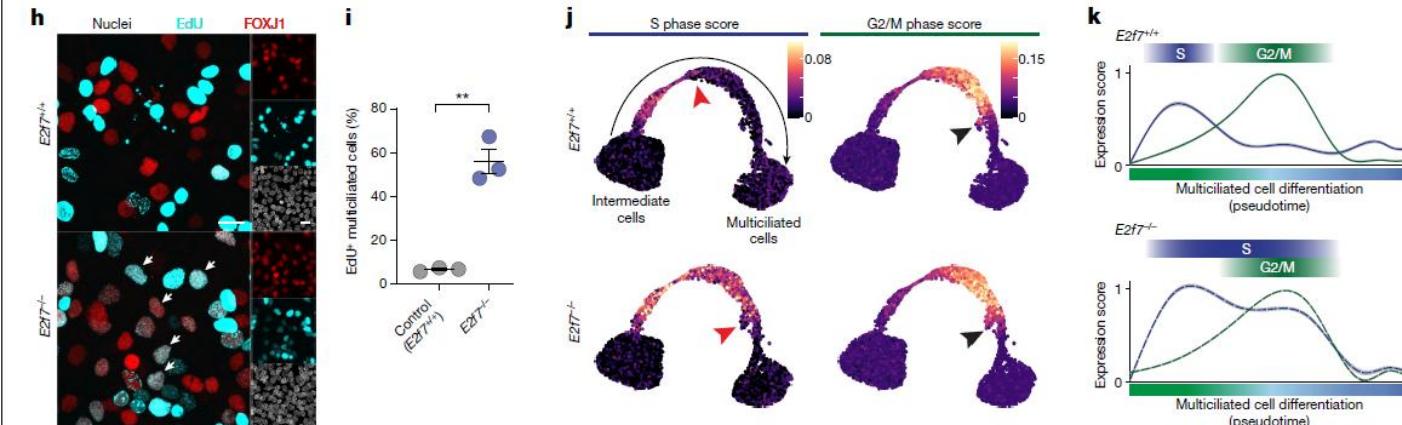
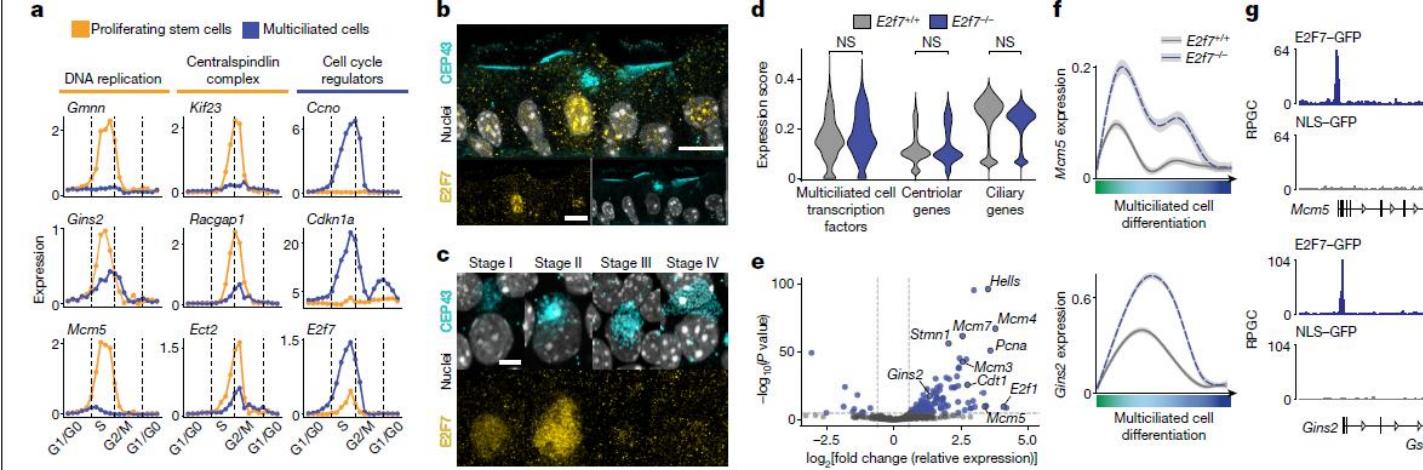
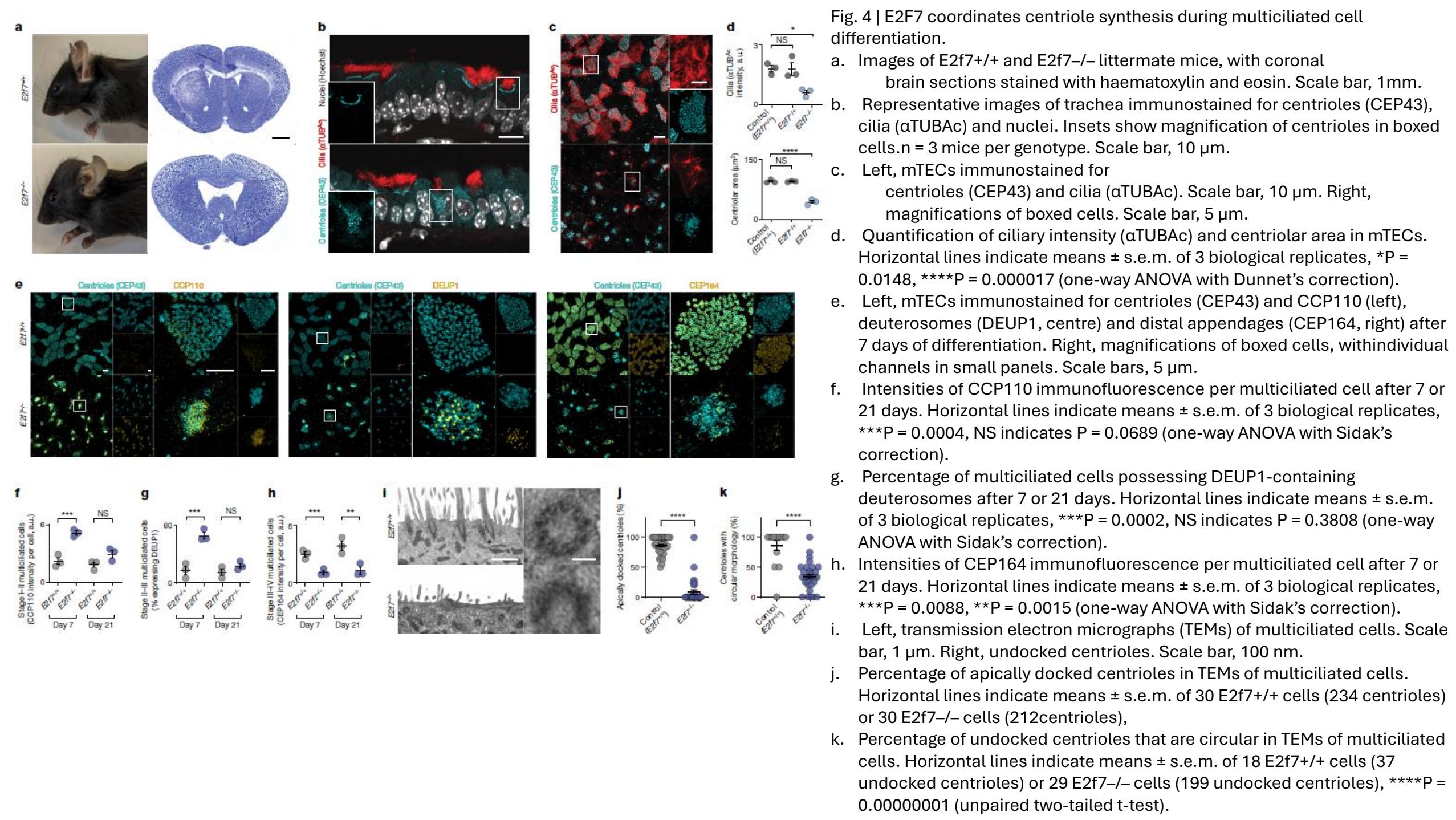


Fig. 3 | E2F7 blocks DNA replication and promotes progression through the multiciliation cycle.

- Average read counts across cycle phase of stem (orange) and multiciliated (blue) cells from the wild-type scRNA-seq dataset (Fig. 1).
- Representative image of adult mouse trachea immunostained for E2F7, centrioles (CEP43) and nuclei. n = 3 mice.
- mTECs immunostained for E2F7, centrioles (CEP43) and nuclei. Cells representative of each stage from three biological replicates are shown.
- Composite expression of multiciliated cell transcription factors, centriolar and ciliary genes in E2f7^{+/+} and E2f7^{-/-}-multiciliated cells. Scores are the normalized Mann–Whitney U-statistic of gene set expression. NS, not significant (multiple unpaired two-tailed t-tests with Holm–Sidak correction).
- Genes differentially expressed between E2f7^{-/-} and E2f7^{+/+}-differentiating multiciliated cells. Blue, fold-change > 1.5 and P < 0.00001 (two-tailed Wald test with Benjamini–Hochberg correction).
- scRNA-seq expression in E2f7^{+/+} and E2f7^{-/-}-multiciliated cells across pseudotime. Grey bars indicate 95% confidence intervals.
- E2F7-GFP or NLS-GFP CUT&RUN in mTECs, presented as reads per genomic content (RPGC).
- E2f7^{+/+} and E2f7^{-/-}-mTECs stained for EdU and FOXJ1. White arrows indicate E2f7^{-/-}-cells that express FOXJ1 and are EdU-positive.
- Percentage of E2f7^{+/+} and E2f7^{-/-}-mTECs expressing FOXJ1 that are EdU-positive. Horizontal lines indicate means ± s.e.m. of 3 biological replicates, **P = 0.001 (unpaired two-tailed t-test).
- S phase (left) and G2/M phase (right) gene signature scores derived from normalized sum rank of gene sets projected onto UMAPs of E2f7^{+/+} and E2f7^{-/-}-multiciliated cells. Colours indicate expression score of S or G2/M gene sets. Arrow indicates differentiation trajectory. Red and black arrowheads indicate the end of the S and G2/M phases, respectively, as defined by the half-maximal phase score.
- E2f7^{+/+} and E2f7^{-/-}-multiciliated cell minimum and maximum normalized S phase and G2/M phase gene signature scores across pseudotime. Grey bars indicate 95% confidence intervals Scale bars, 10 μm (b,h), 5 μm (c) or 1 kb (g).



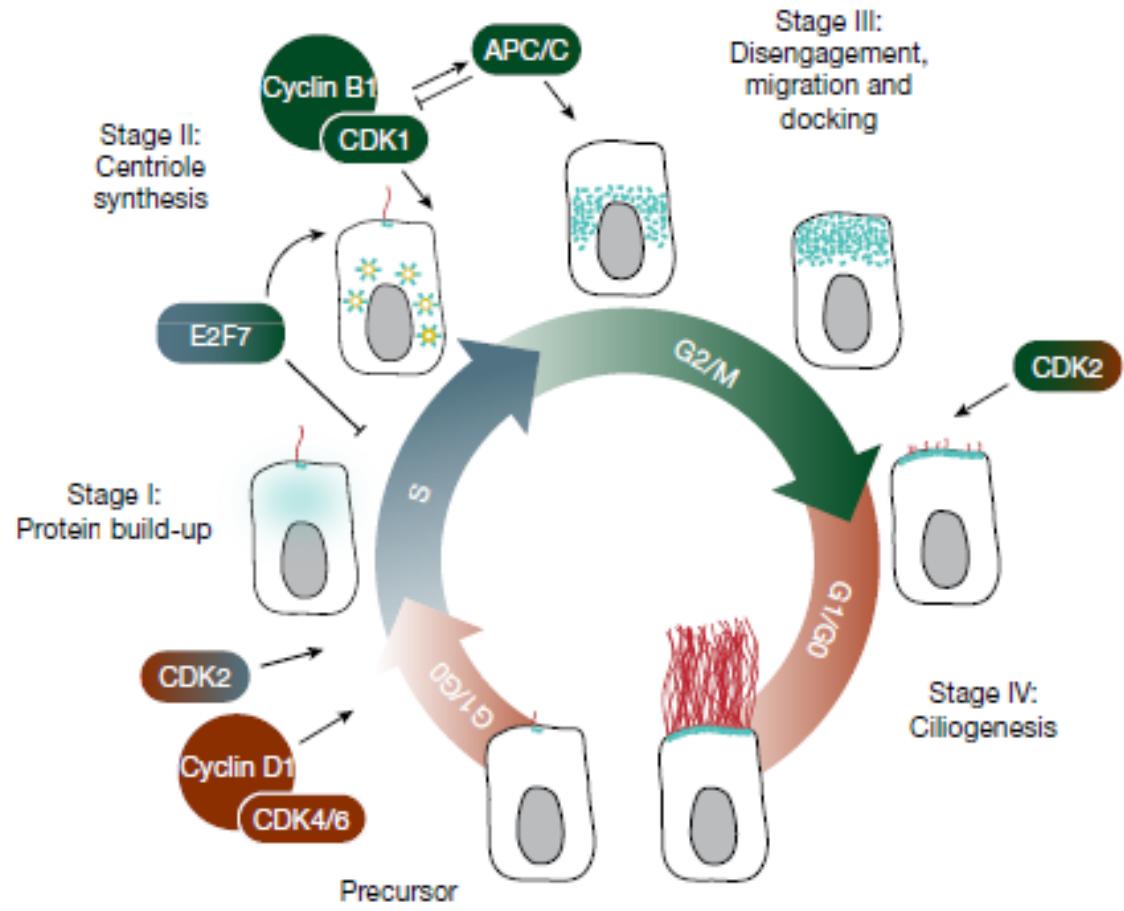
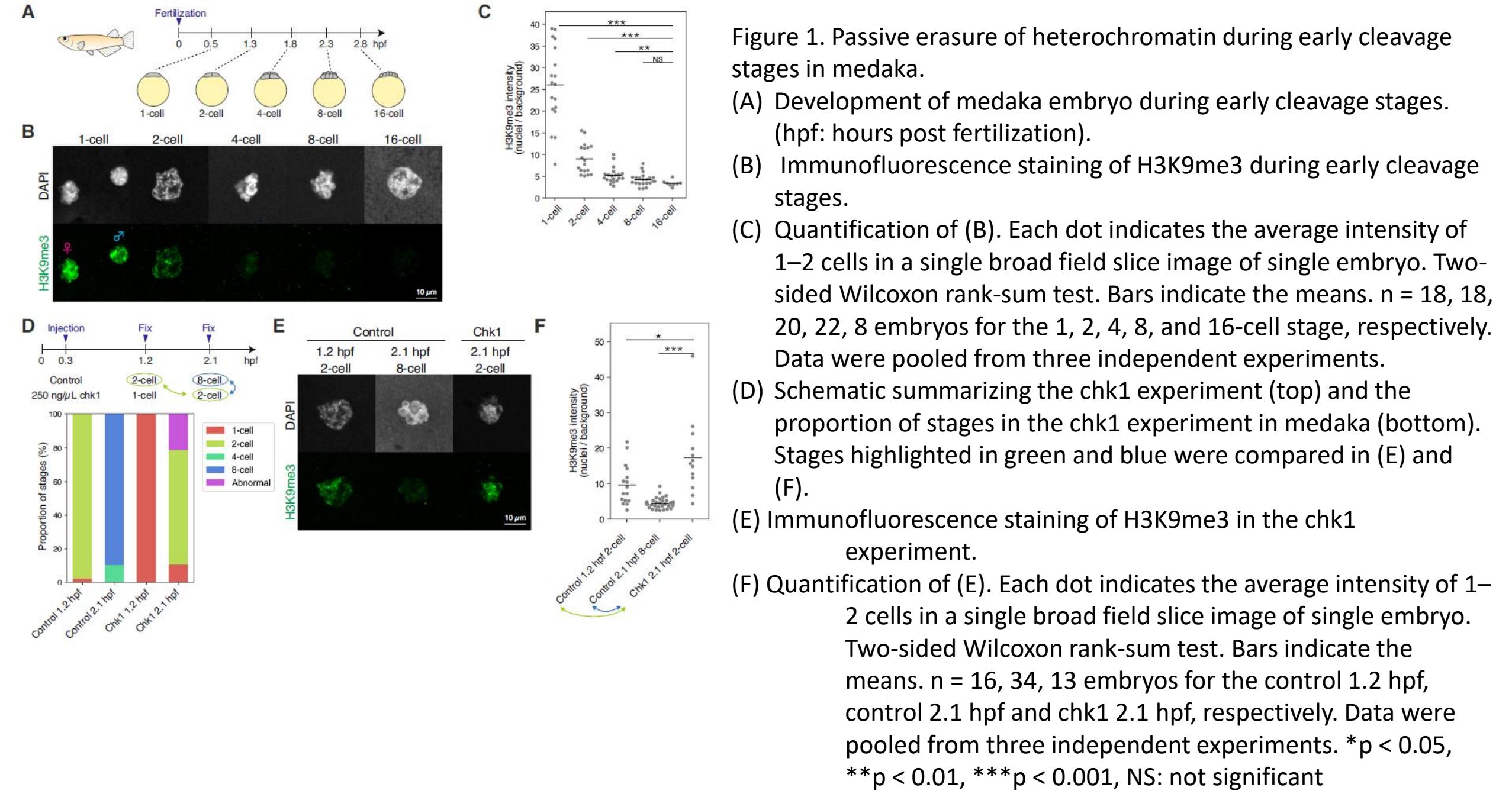


Fig. 5 | The multiciliation cycle is a cell cycle variant that coordinates differentiation. A model of how the multiciliation cycle coordinates multiciliated cell differentiation. Multiciliated cell precursors initiate differentiation in a G1/G0-like phase. Precursors progress into an S-like phase encompassing stage I and early stage II. Cyclin D1–CDK4/6 and CDK2 regulate entry into the S-like phase (this work and ref. 6). E2F7 suppresses DNA synthesis during the S-like phase and promotes the S-like to G2/M-like transition. During the G2/M-like phase, cyclin B1–CDK1 promotes the growth of newly forming centrioles and APC/C controls centriole number and progression to stage III of multiciliated cell differentiation, when centrioles dock to the membrane⁵. From the G2/M-like phase, differentiating multiciliated cells transition into the G1/G0-like phase corresponding to stage IV, ciliogenesis. CDK2 promotes this final stage of multiciliated cell differentiation.



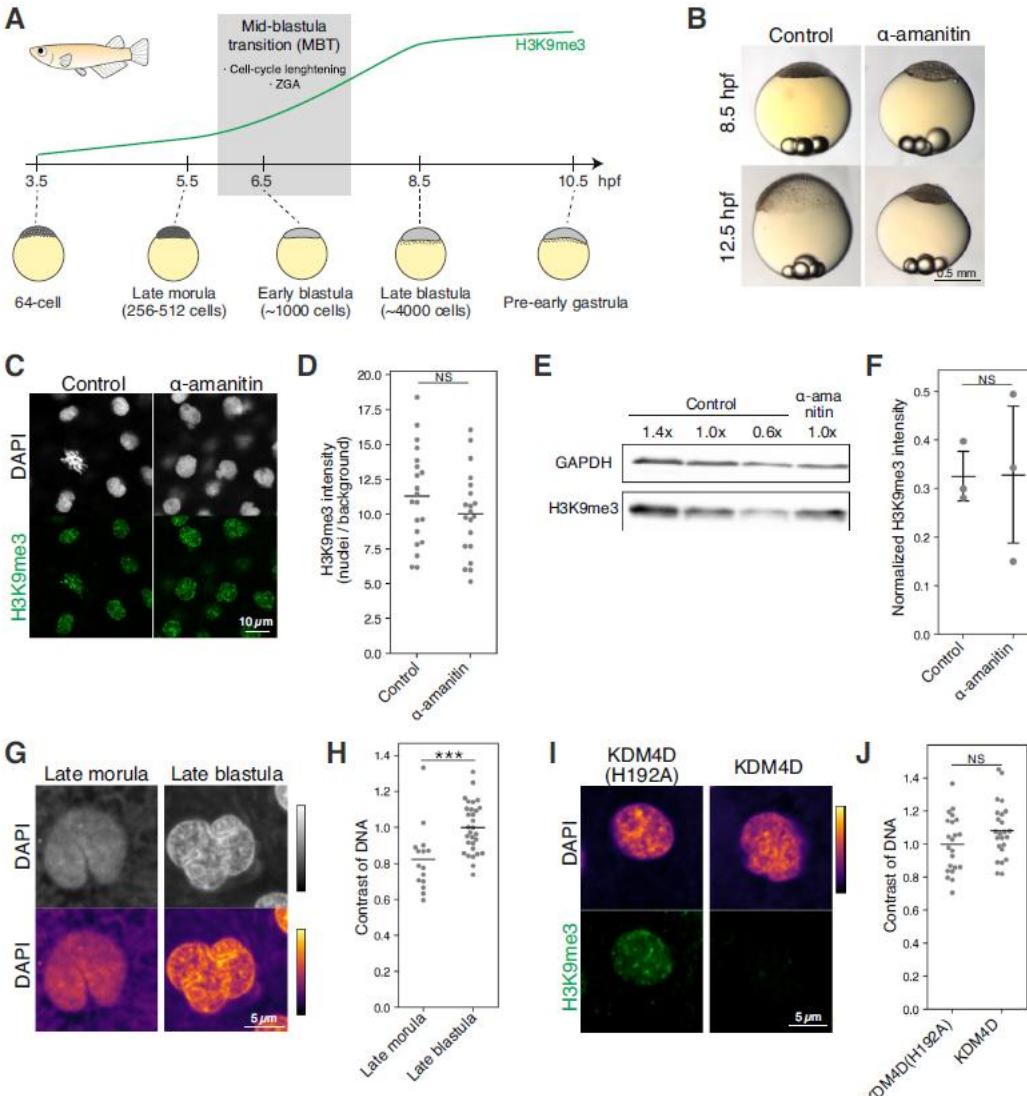


Figure 2. ZGA is dispensable for heterochromatin establishment during the MBT in medaka.

(A) Development of medaka embryos before and after the MBT.

(B) Phenotype of α -amanitin-injected medaka embryos.

(C) Immunofluorescence staining of H3K9me3 at the late blastula stage (8.5 hpf) in the α -amanitin injection experiment.

(D) Quantification of (C). Each dot indicates the average intensity of ~40 cells in a single broad field slice image of single embryo. Two-sided unpaired Student's t-test. Bars indicate the means. n = 20 embryos. Data were pooled from two independent experiments.

(E) Western blot of H3K9me3 and GAPDH using control and α -amanitin-injected embryos at the late blastula stage (8.5 hpf).

(F) Quantification of (E). H3K9me3 signal intensity was normalized by GAPDH signal intensity. Two-sided unpaired Student's t-test. Error bars indicate the mean \pm s.d. n = 3 biological replicates.

(G) DAPI-staining at the late morula and late blastula stages. Colormaps are shown at the bottom to better illustrate the appearance of DNA-dense regions at the late blastula stage.

(H) Quantification of DNA contrast in (G). Each dot indicates the DNA contrast of a single nucleus. ~6 embryos were analyzed. Two-sided Wilcoxon rank-sum test. Bars indicate the means. n = 15 and 30 nuclei for the late morula and late blastula, respectively. Data were pooled from two independent experiments.

(I) DAPI and immunofluorescence staining of embryos injected with human KDM4D, a demethylase of H3K9me3, or its catalytically inactive mutant KDM4D(H192A) at the late blastula stage. The pattern of DNA-dense domains was comparable irrespective of presence or absence of H3K9me3.

(J) Quantification of DNA contrast in (I). Each dot indicates the DNA contrast of a single nucleus. Five embryos were analyzed. Two-sided unpaired Student's t-test. Bars indicate the means. n = 22 and 24 nuclei for the KDM4D(H192A) and KDM4D, respectively. Data were pooled from two independent experiments. ***p < 0.001, NS: not significant.

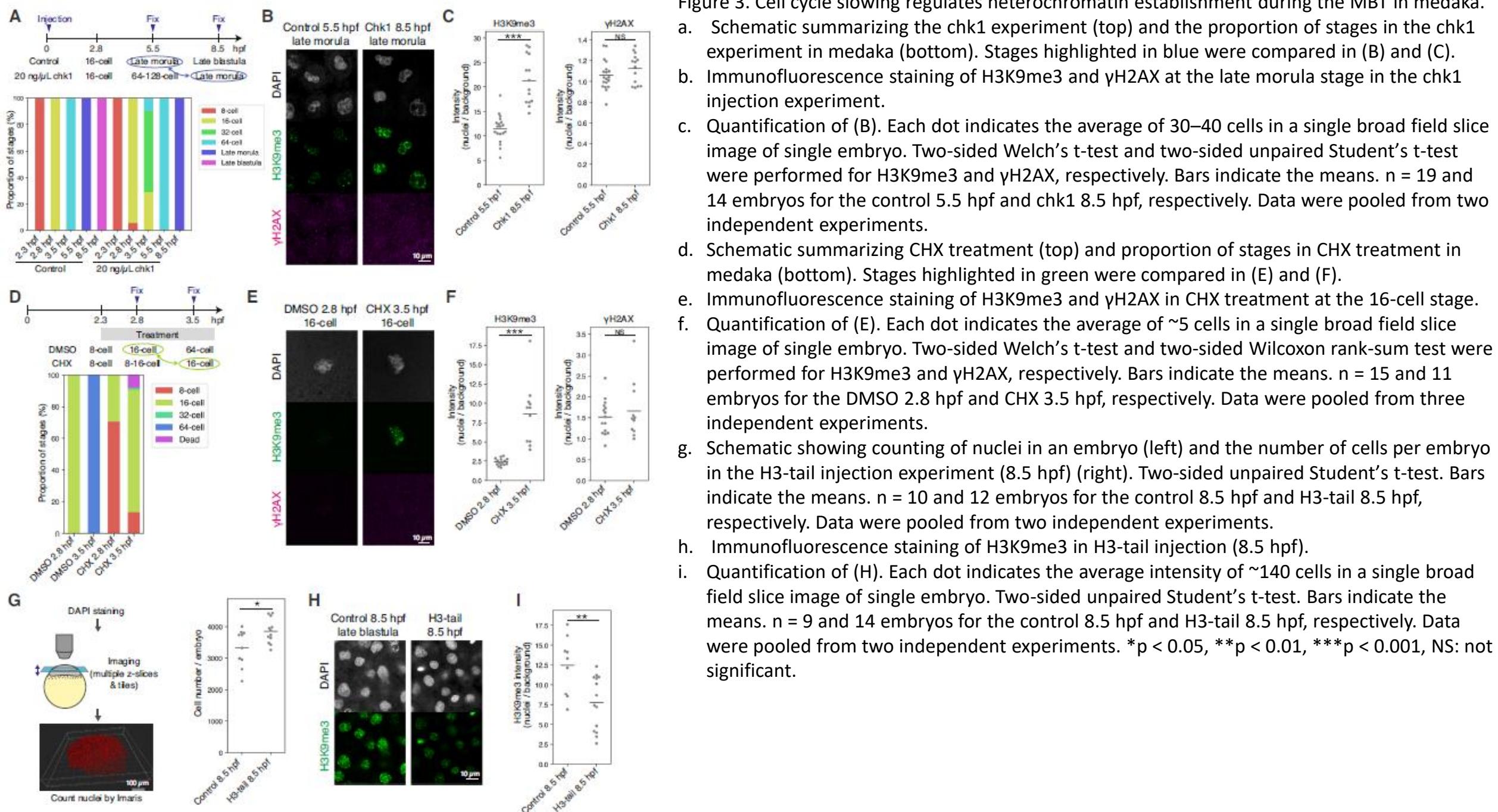
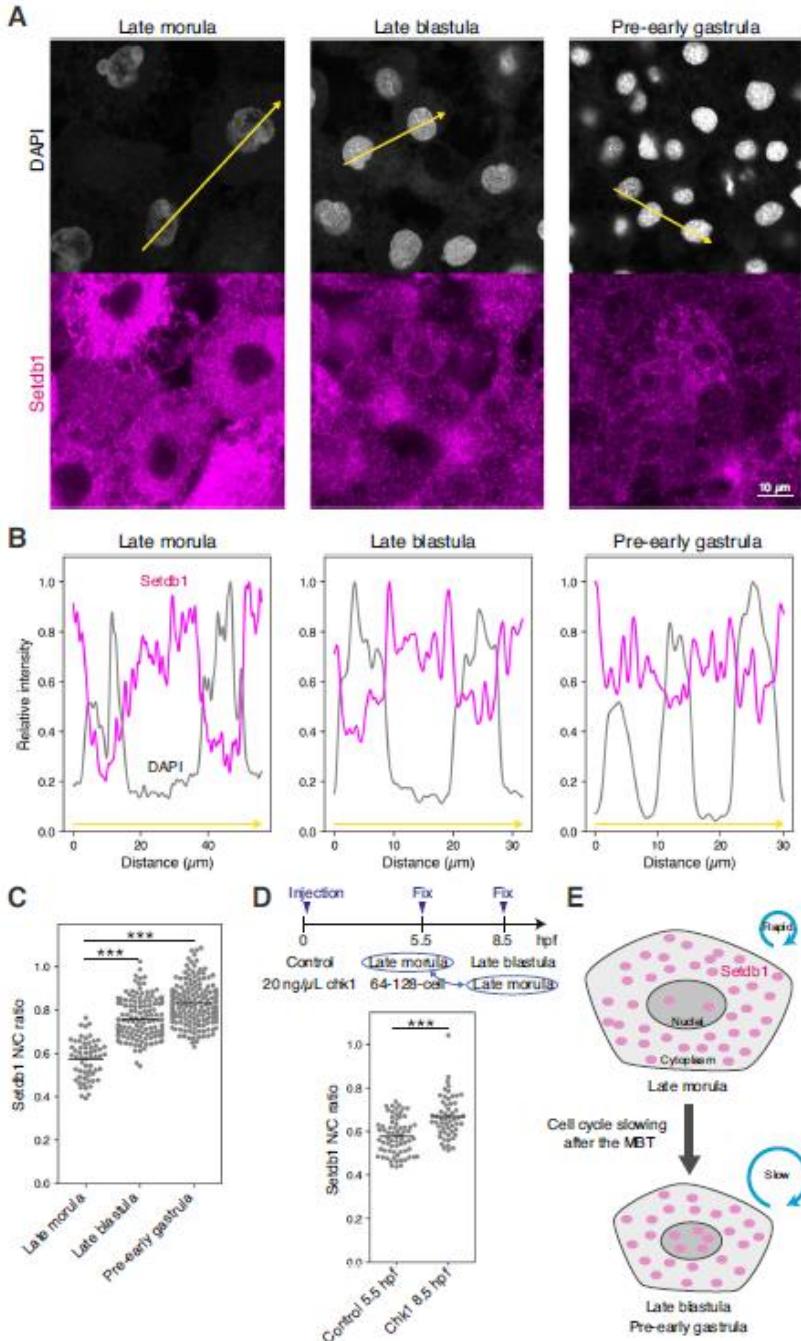


Figure 3. Cell cycle slowing regulates heterochromatin establishment during the MBT in medaka.

- Schematic summarizing the chk1 experiment (top) and the proportion of stages in the chk1 experiment in medaka (bottom). Stages highlighted in blue were compared in (B) and (C).
- Immunofluorescence staining of H3K9me3 and γH2AX at the late morula stage in the chk1 injection experiment.
- Quantification of (B). Each dot indicates the average of 30–40 cells in a single broad field slice image of single embryo. Two-sided Welch's t-test and two-sided unpaired Student's t-test were performed for H3K9me3 and γH2AX, respectively. Bars indicate the means. n = 19 and 14 embryos for the control 5.5 hpf and chk1 8.5 hpf, respectively. Data were pooled from two independent experiments.
- Schematic summarizing CHX treatment (top) and proportion of stages in CHX treatment in medaka (bottom). Stages highlighted in green were compared in (E) and (F).
- Immunofluorescence staining of H3K9me3 and γH2AX in CHX treatment at the 16-cell stage.
- Quantification of (E). Each dot indicates the average of ~5 cells in a single broad field slice image of single embryo. Two-sided Welch's t-test and two-sided Wilcoxon rank-sum test were performed for H3K9me3 and γH2AX, respectively. Bars indicate the means. n = 15 and 11 embryos for the DMSO 2.8 hpf and CHX 3.5 hpf, respectively. Data were pooled from three independent experiments.
- Schematic showing counting of nuclei in an embryo (left) and the number of cells per embryo in the H3-tail injection experiment (8.5 hpf) (right). Two-sided unpaired Student's t-test. Bars indicate the means. n = 10 and 12 embryos for the control 8.5 hpf and H3-tail 8.5 hpf, respectively. Data were pooled from two independent experiments.
- Immunofluorescence staining of H3K9me3 in H3-tail injection (8.5 hpf).
- Quantification of (H). Each dot indicates the average intensity of ~140 cells in a single broad field slice image of single embryo. Two-sided unpaired Student's t-test. Bars indicate the means. n = 9 and 14 embryos for the control 8.5 hpf and H3-tail 8.5 hpf, respectively. Data were pooled from two independent experiments. *p < 0.05, **p < 0.01, ***p < 0.001, NS: not significant.



- Figure 4. Setdb1 accumulates to nuclei upon the MBT in medaka.**
- Immunofluorescence staining of Setdb1 in medaka embryos before (late morula) and after the MBT (late blastula and pre-early gastrula). Signal intensities along the yellow arrows were quantified in (B).
 - Quantification of signal intensity of DAPI and Setdb1 along the yellow arrows in (A).
 - Quantification of nuclear/cytoplasmic ratio (N/C ratio) of Setdb1 in medaka embryos before (late morula) and after the MBT (late blastula and pre-early gastrula). Each dot indicates the N/C ratio of a single cell. 10, 8, and 10 embryos at the late morula, late blastula, and pre-early gastrula, respectively, were analyzed. Two-sided Wilcoxon rank-sum test. Bars indicate the means. n = 59, 142, and 193 cells for the late morula, late blastula, and the pre-early gastrula stage, respectively. Data were pooled from three independent experiments.
 - Schematic showing chk1 overexpression experiment (top) and quantification of N/C ratio of Setdb1 in control and chk1-injected embryos (bottom) at the late morula stage. Each dot indicates the N/C ratio of a single cell. Eleven embryos were analyzed for each condition. Two-sided Wilcoxon rank-sum test. Bars indicate the means. n = 69 and 55 cells for the control 5.5 hpf and chk1 8.5 hpf, respectively. Data were pooled from two independent experiments.
 - Schematic representation of the model of Setdb1 accumulation induced by cell cycle slowing during the MBT. ***p < 0.001. Source data are available online for this figure.