

# Genstruktur und Genexpression bei Eukaryoten

Kap. 21,  
23, 24, 26

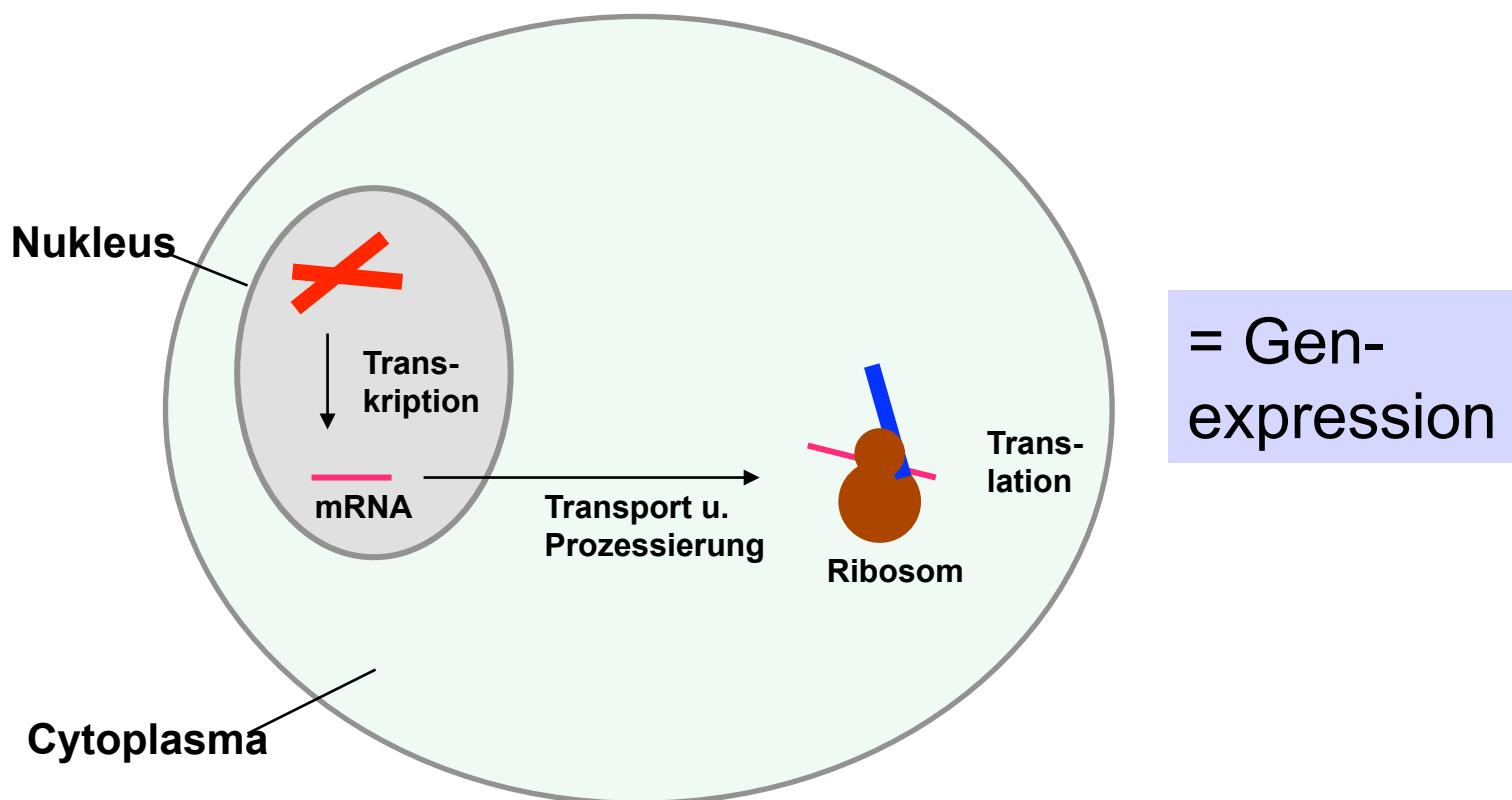


An animated primer on the basics of DNA, genes, and heredity.

Thomas Hankeln (iOME) & Julian König (IMB)

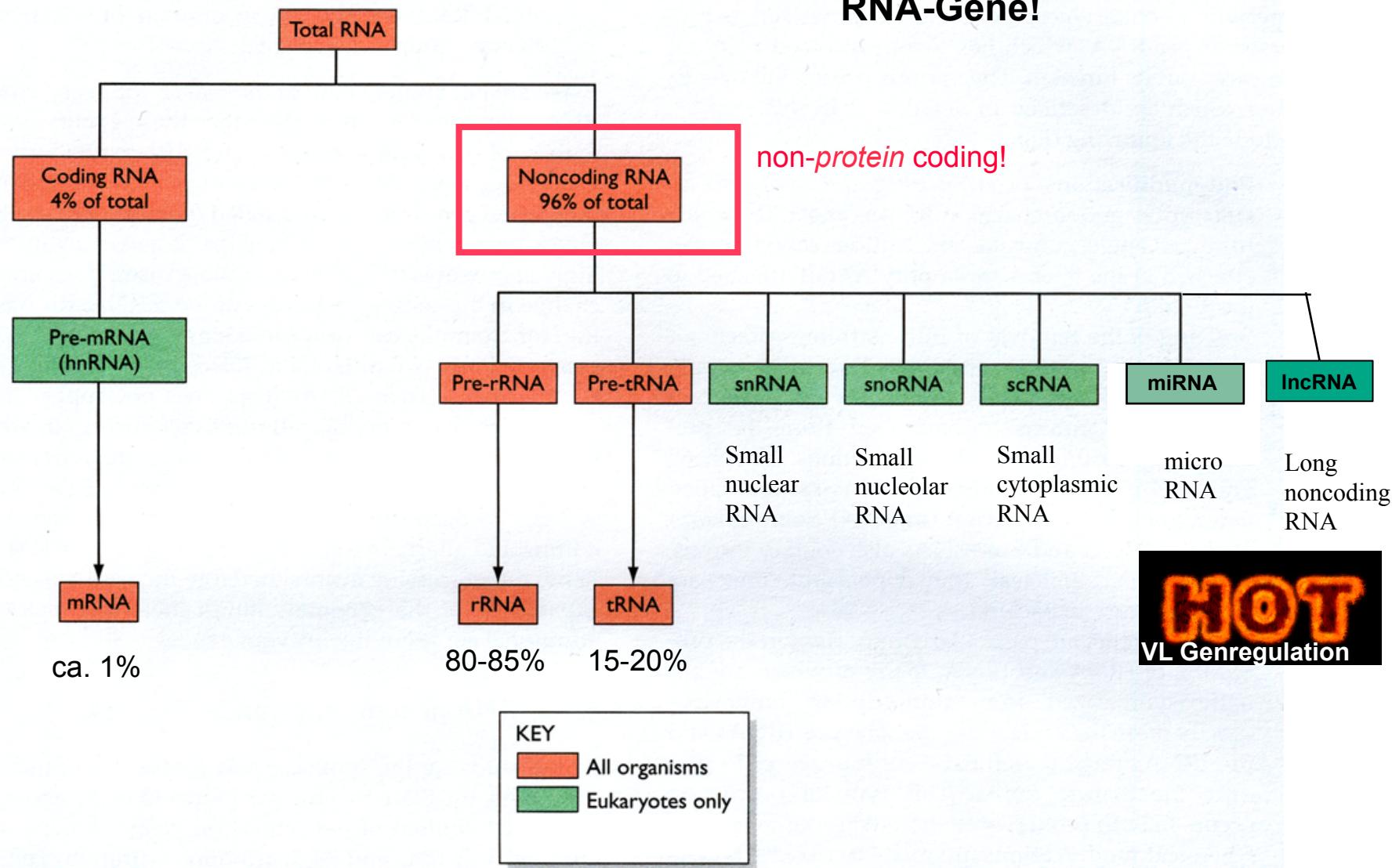
F. H. Crick 1957

# Das „Zentrale Dogma“ der Molekulargenetik



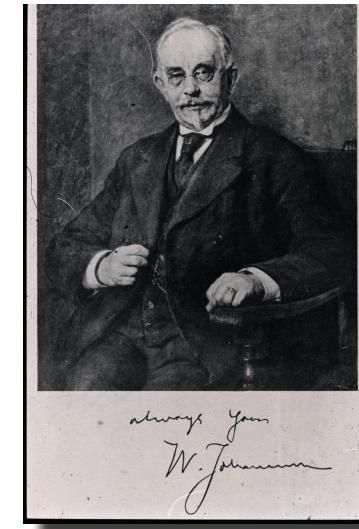
# Typen von RNA

Zusätzlich zu den ca. 20000 Protein-kodierenden Genen haben wir > 10.000 RNA-Gene!



# Genbegriff

- *Mendel*: „Faktor, der für Merkmal verantwortlich ist“
- *Bateson*: „Genetik“ (von *genetikos*, Hervorbringung)
- *Johannsen*: Begriff des „Gens“
- *Morgan*: „Gen ist chromosomaler Abschnitt“



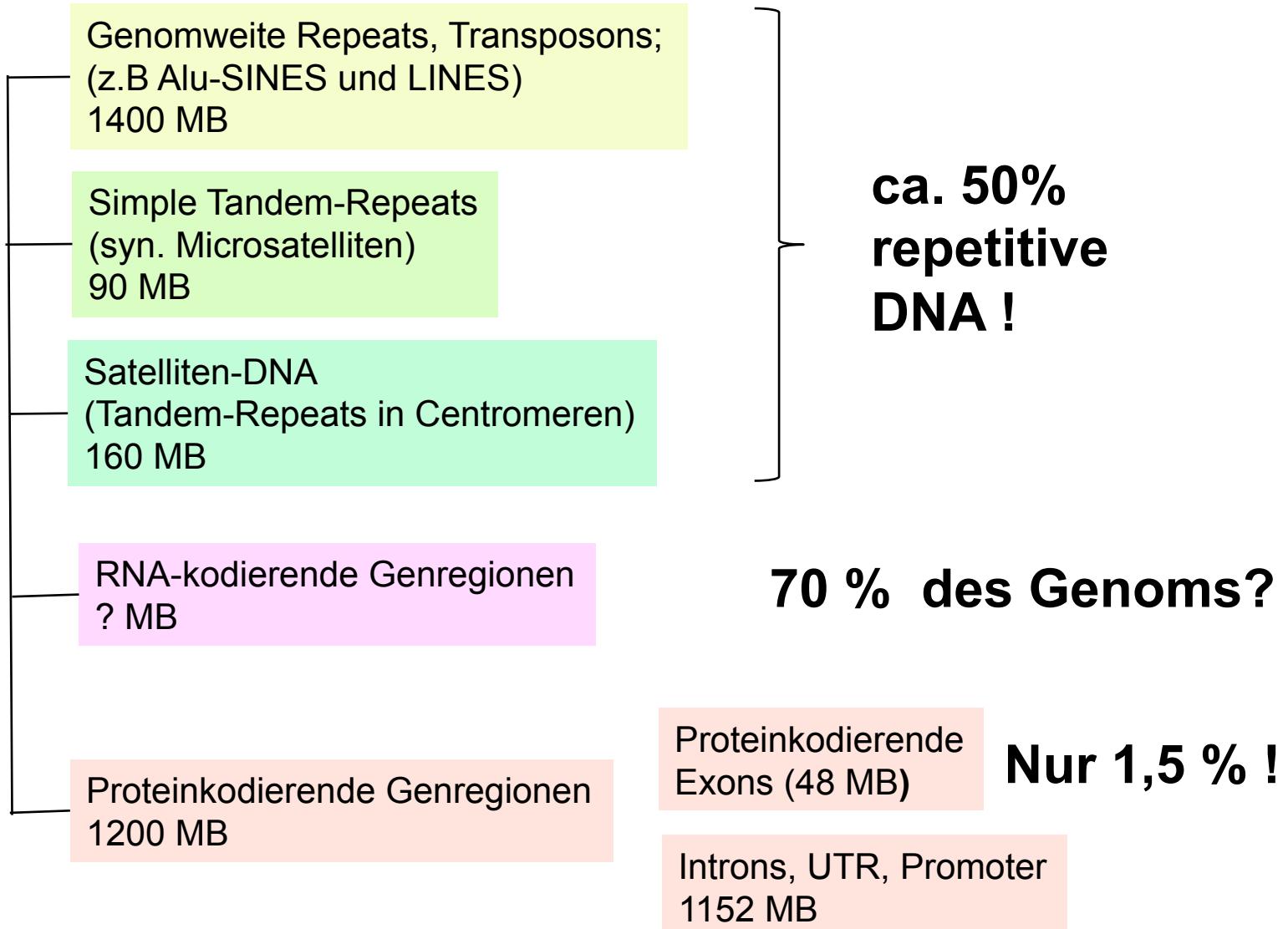
## Ein Gen...

- ist ein **funktioneller Abschnitt** der DNA, der entweder ein Protein oder eine RNA kodiert.
- umfasst auch die benachbarten regulatorischen Abschnitte (Promotor, Enhancer etc.)

# Genomkomponenten

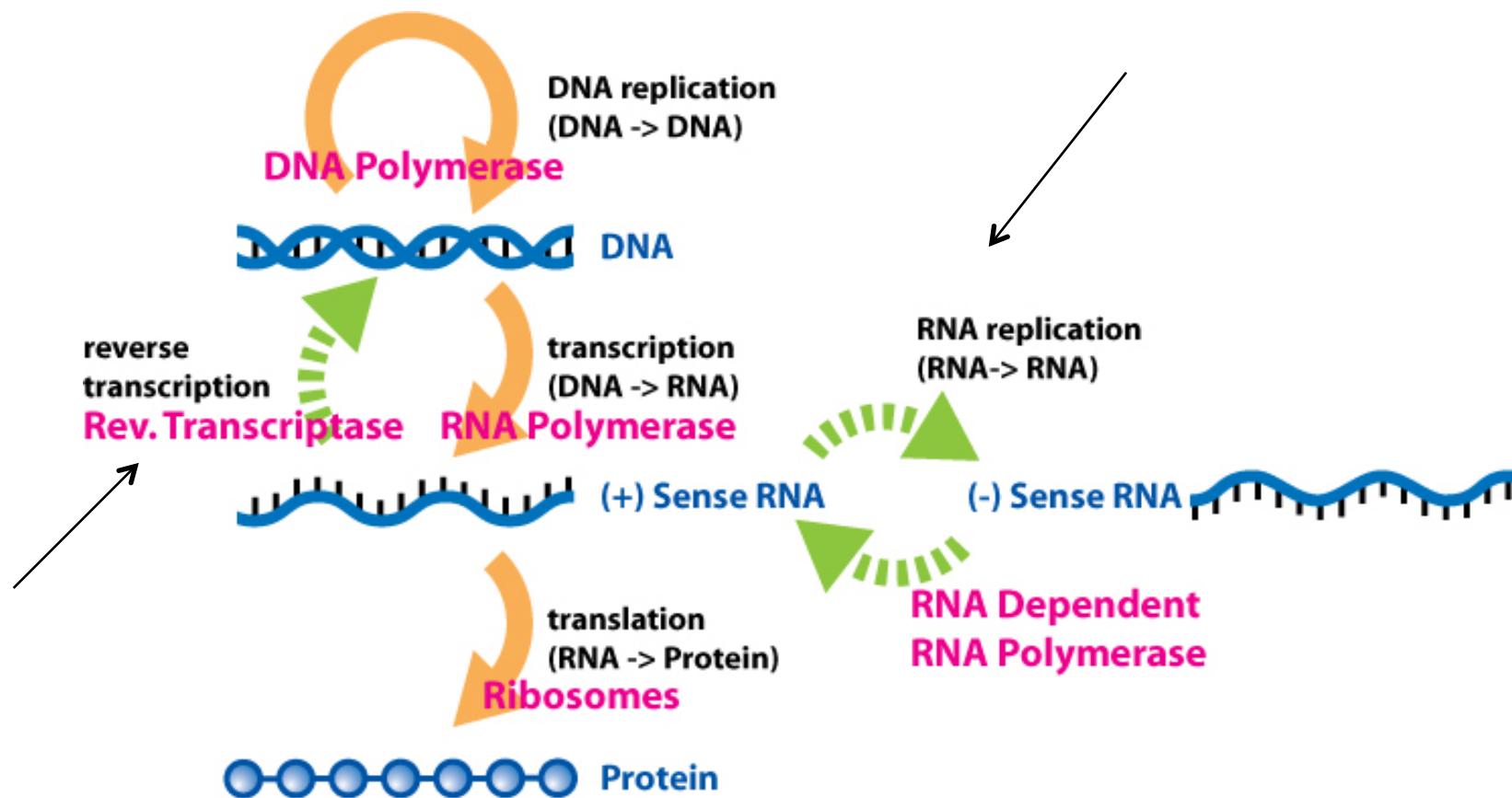


Humangenom  
3200 MB



MB = Mega-Basenpaare = 1 Mio Bp

# Moderne Erweiterungen des „Zentralen Dogmas“



# RNA: Vom Gen zur Funktion



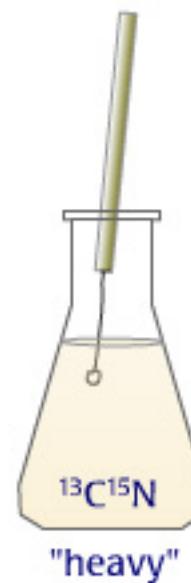
„...to solve the riddle of the RNA structure and to understand how it builds proteins!“

# 1960

The discovery of messenger RNA (mRNA) by Sydney Brenner (1927-), Francis Crick (1916-), François Jacob (1920-) and Jacques Monod (1910 - 1976)

Hello, I'm Sydney Brenner. With my colleagues, François Jacob and Matthew Meselson, I showed that rRNA was not the template for building proteins. There was a third type of RNA – an unstable intermediate – that carries the DNA message to the ribosome.

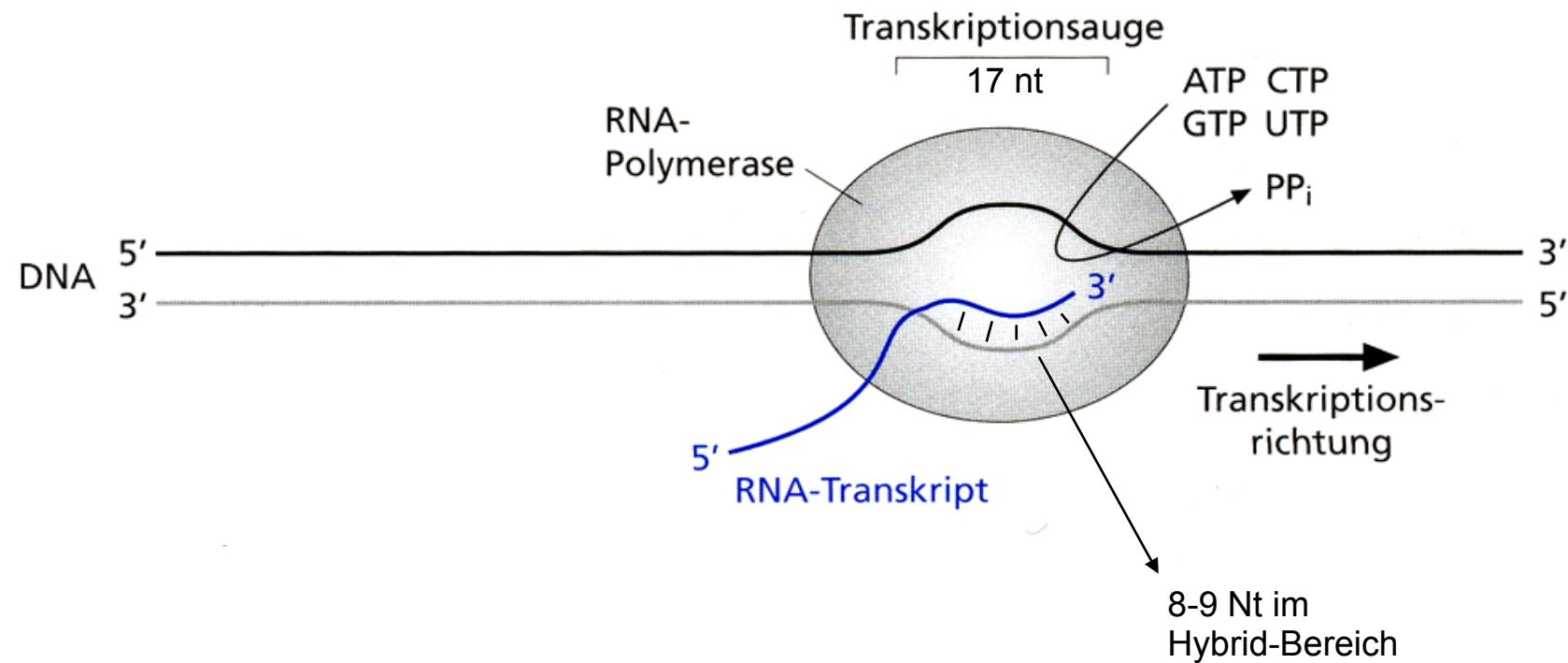
We did this by using phage-infected bacteria. We started by growing bacteria in "heavy" isotopes of carbon and nitrogen to radiolabel all of the bacterial RNA and proteins.



Also lautet ein  
Beschluß,  
daß der Mensch  
'was lesen muß!



# Transkription



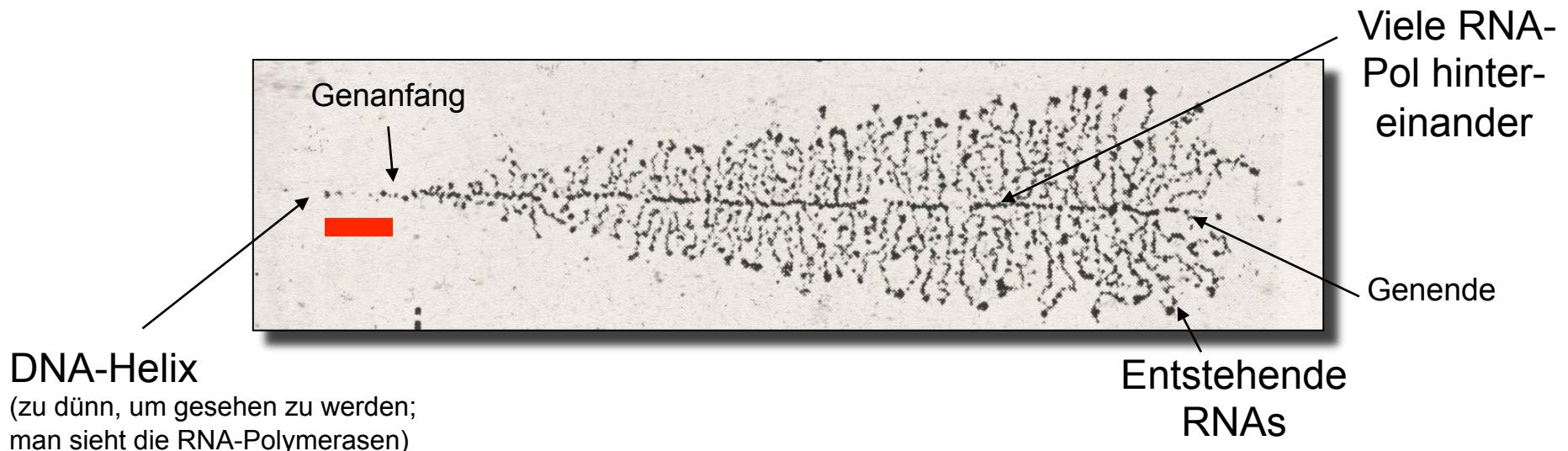
Die **RNA-Polymerase** addiert Ribonukleotide in **5' > 3'** –Richtung.

Sie braucht **keinen** Primer.

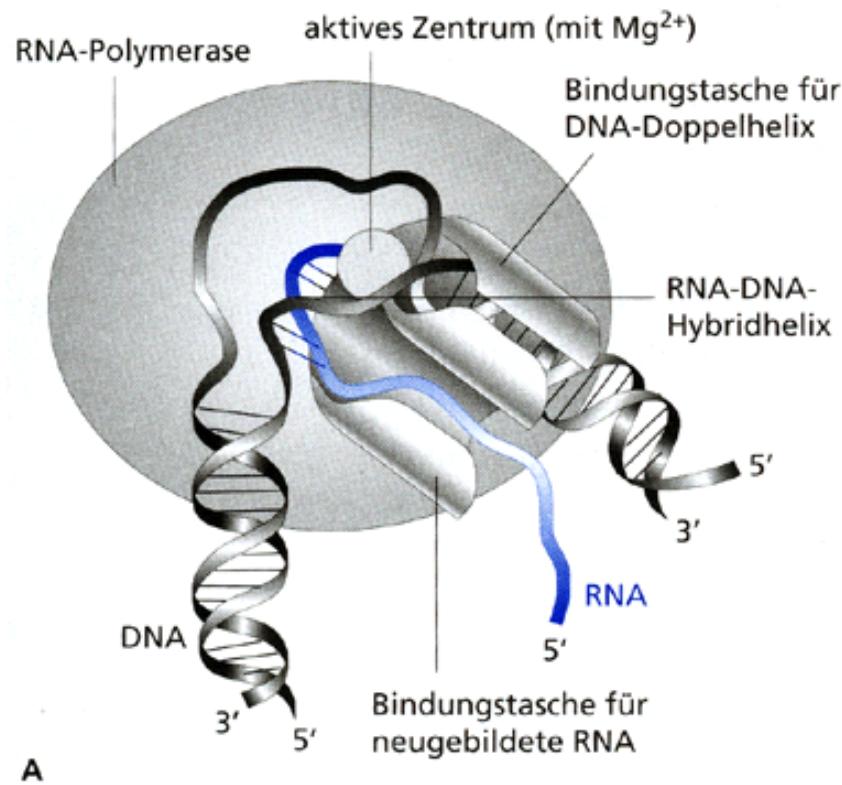


# Transkription

- Transkription startet dennoch nicht überall, sondern an definierten Stellen im Genom (**Promotor**).
- RNA-Pols können in kurzer Folge ein Gen transkribieren.

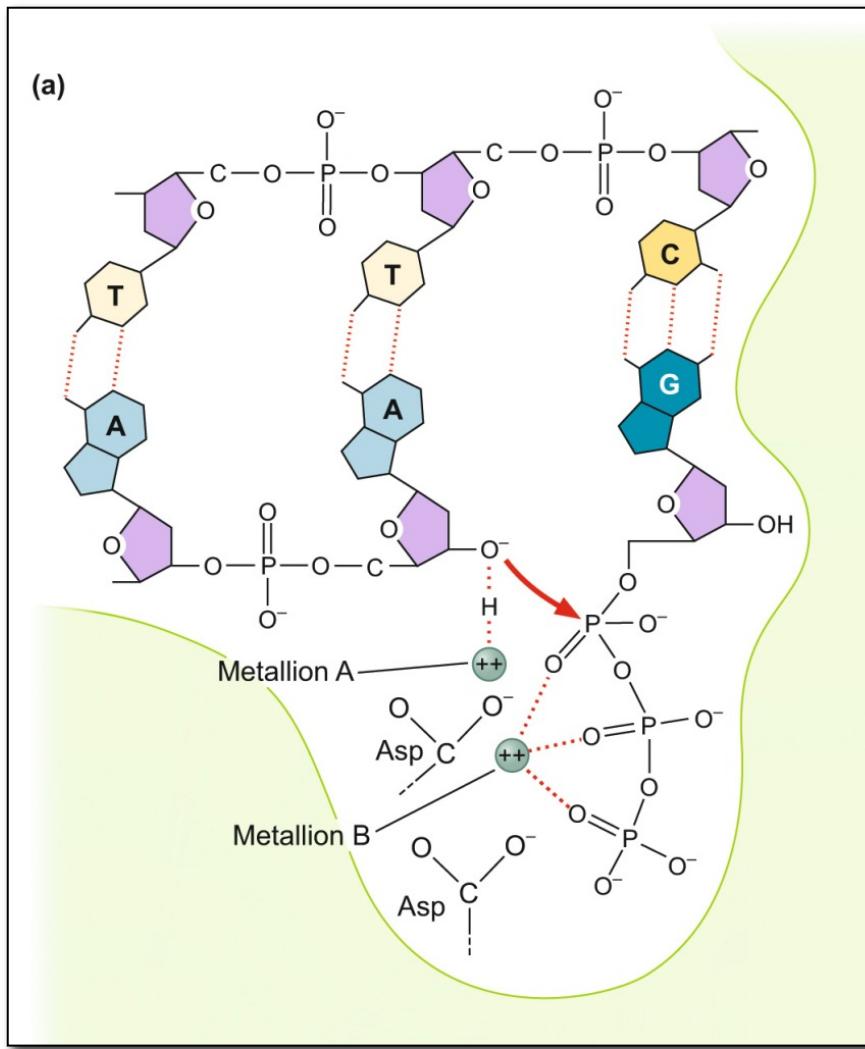


# RNA-Polymerase



- Enzym-Komplex aus mehreren Untereinheiten
- **Kein „proof reading“!**  
Macht alle  $10^4$ - $10^5$  Nt einen Fehleinbau. Egal...
- ca. 60-80 Nt pro sec  
(viel langsamer als Replikation)
- „eingebaute“ Helicase-Funktion
- Unterstützt Re-Hybridisierung der DNA-Stränge & löst RNA von der DNA-Matrize ab

# Aktives Zentrum

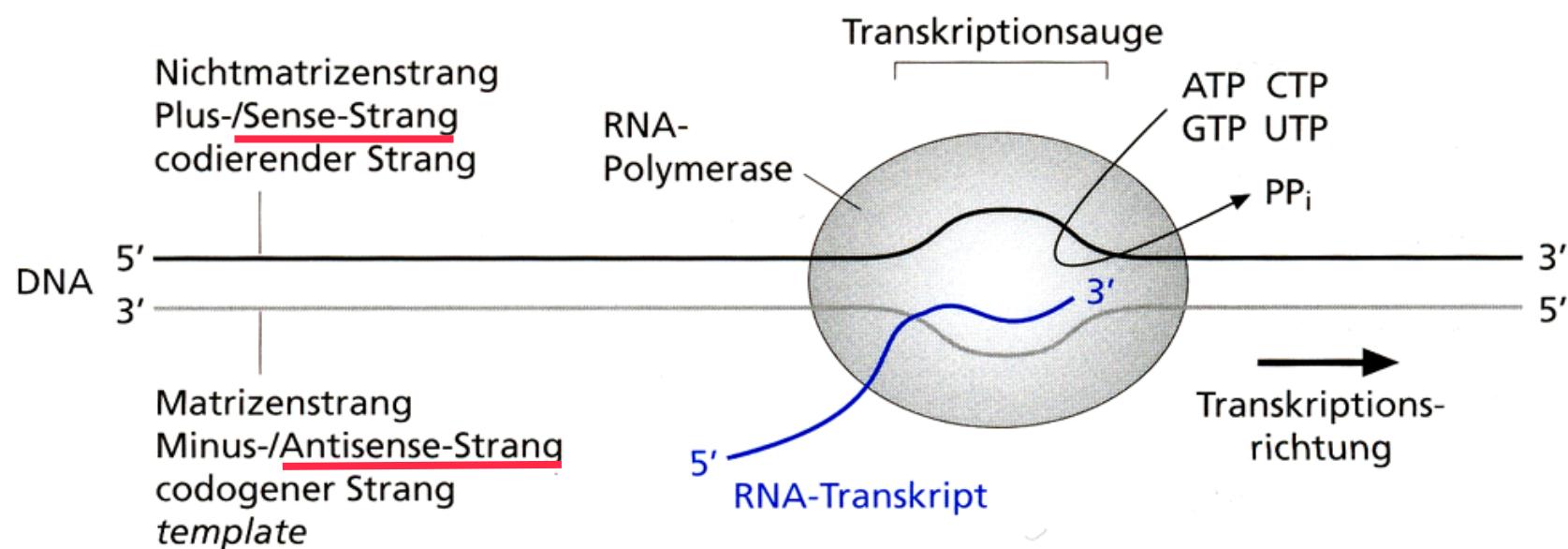


RNA Pol hat (ebenso wie die hier im Bild gezeigte DNA Pol)  
**Mg<sup>++</sup> Ionen im aktiven Zentrum.**

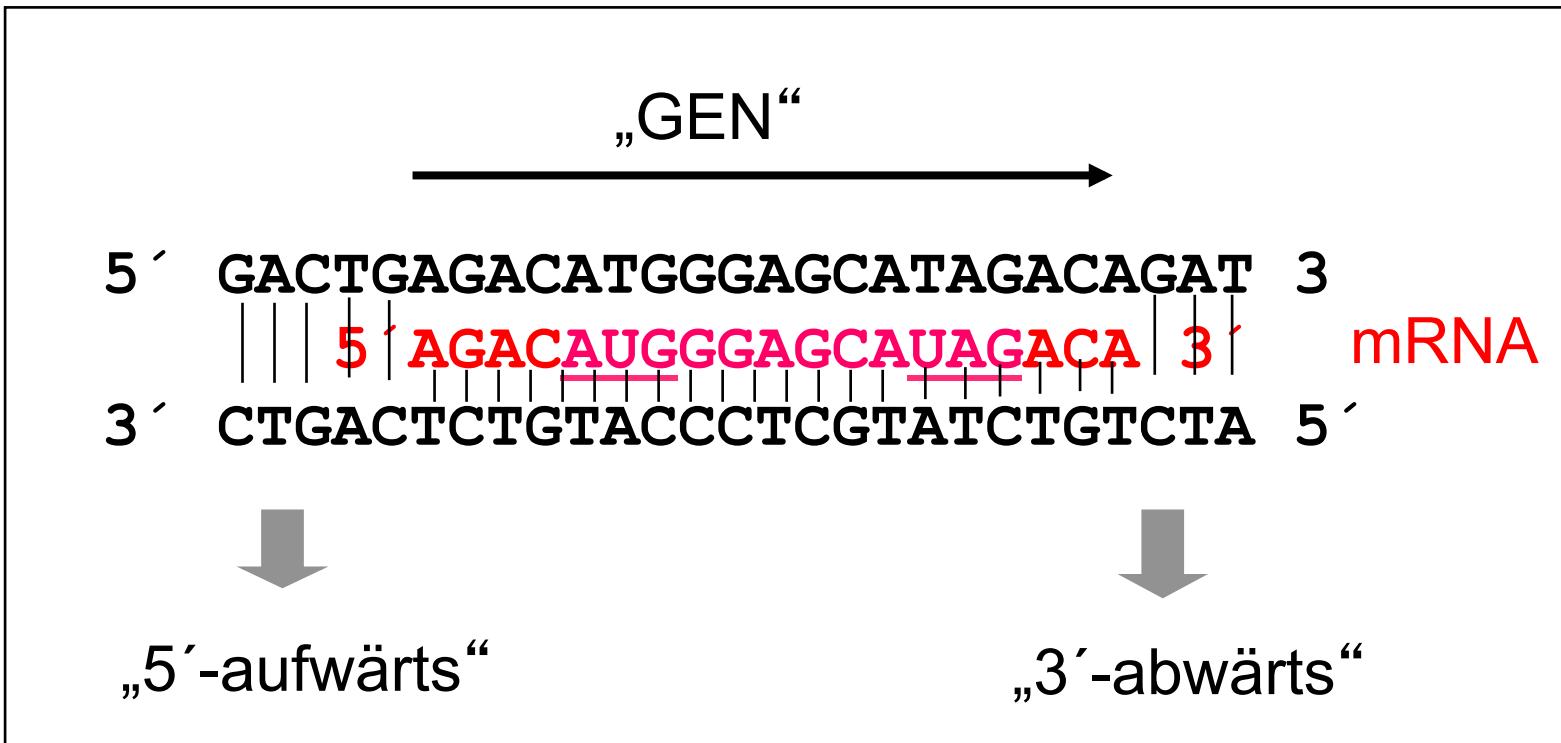
Ion A schwächt OH-Bindung und erhöht nukleophile Reaktivität des O-Atoms.

Ion B neutralisiert z. T. die negative Ladung des PPi.

# Transkription & Gen-Anordnung



# Transkription & Genanordnung

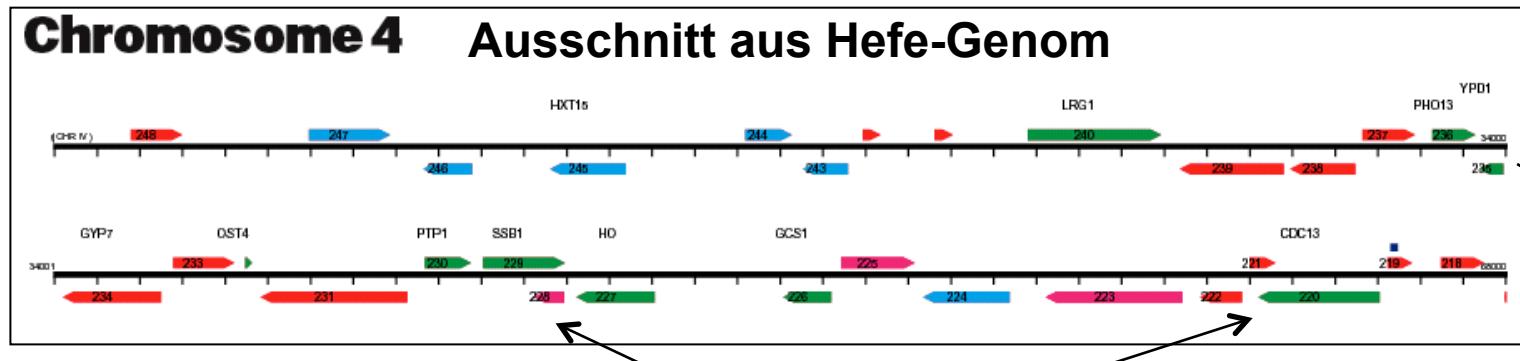


Konvention: Gen-Darstellung durch Schreiben derjenigen Sequenz,  
die der mRNA entspricht! („Sense-Strang“)

Pfeil gibt immer die transkriptionelle Orientierung an.

# Gene & Transkription

- Beide Stränge der DNA bei Pro-und Eukaryoten können transkribiert werden.

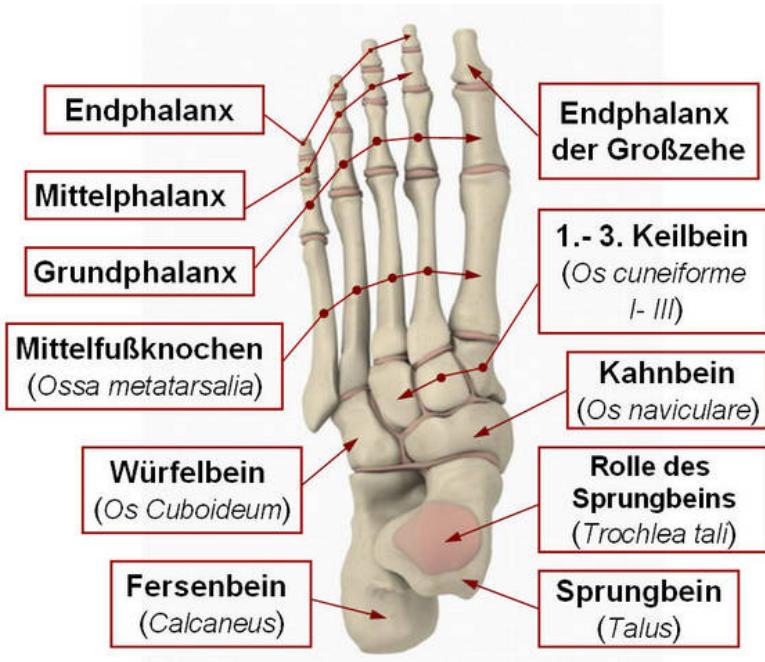


- Gene können **überlappen!**

- Transkribierter Genomanteil vermutlich **> 70%**!  
Protein-kodierender Anteil nur ca. 1,5 %!



14.6.2007



# > Anatomie der Gene

# 5 RNA-Polymerasen in Eukaryoten !

- RNA Pol I rRNA (18S - 5,8S - 28S rRNA)
  - RNA Pol II mRNA, snRNA, miRNA
  - RNA Pol III tRNA, 5S rRNA, 7SL RNA  
U6 sn RNA, snoRNA, miRNA

---

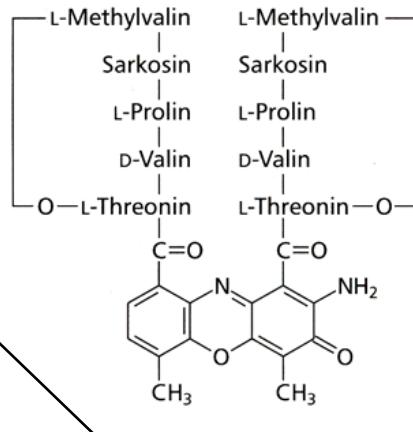
  - RNA Pol IV nur in Pflanzen (siRNA)

---

  - POLRMT RNA Pol für TK der mtDNA

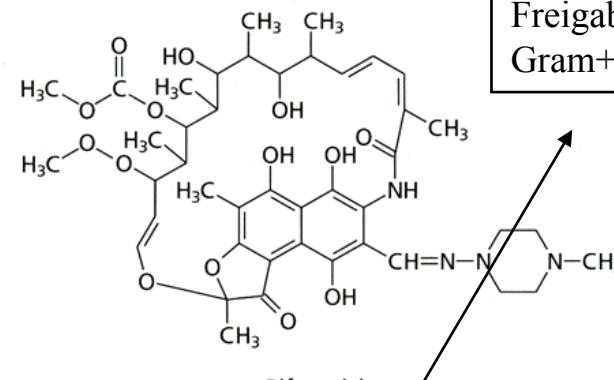
# RNA-Pol: Zielstruktur für Antibiotika!

Hemmt TK in Pro- u. Eukaryoten durch Interkalation

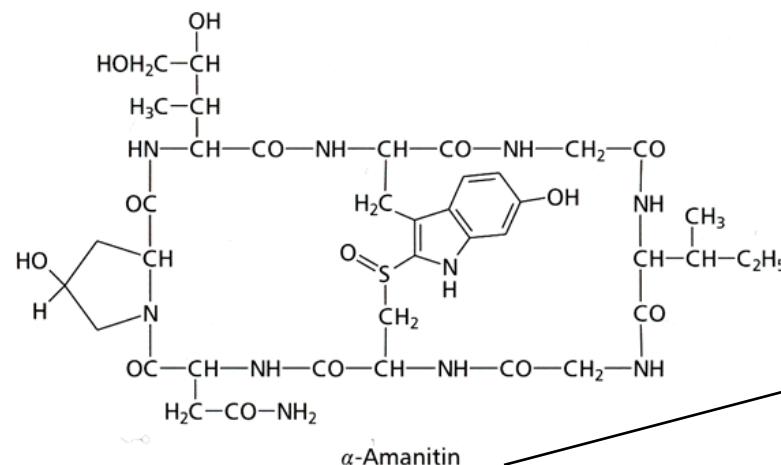


A

Bindet an  $\beta$ -UE der bakt. RNA Pol und verhindert Promoter-Freigabe. Gegen Gram+/- und Mykobakt.



B

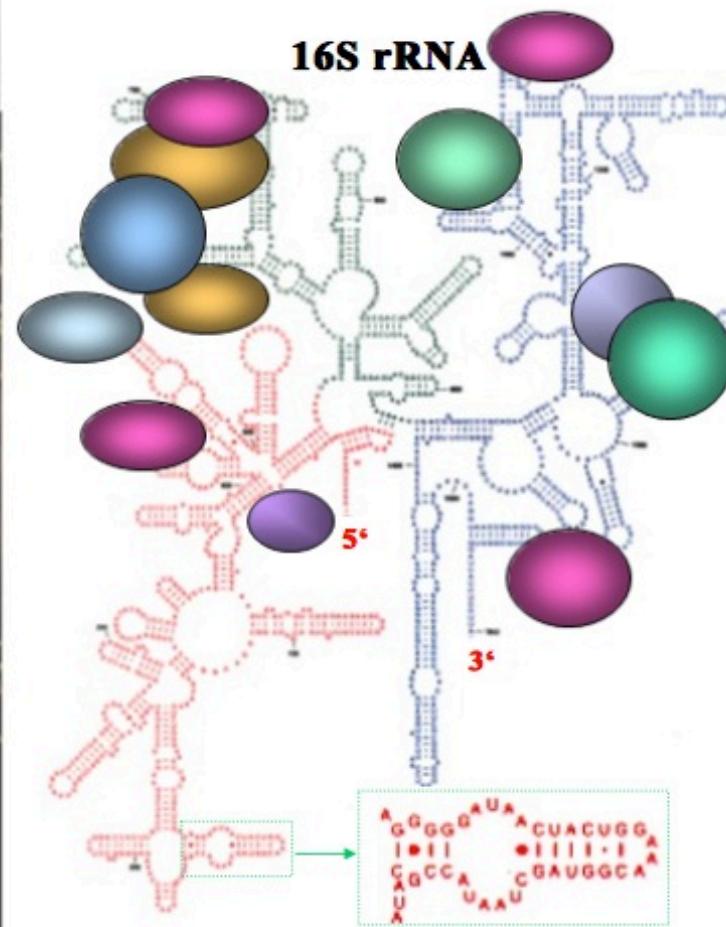
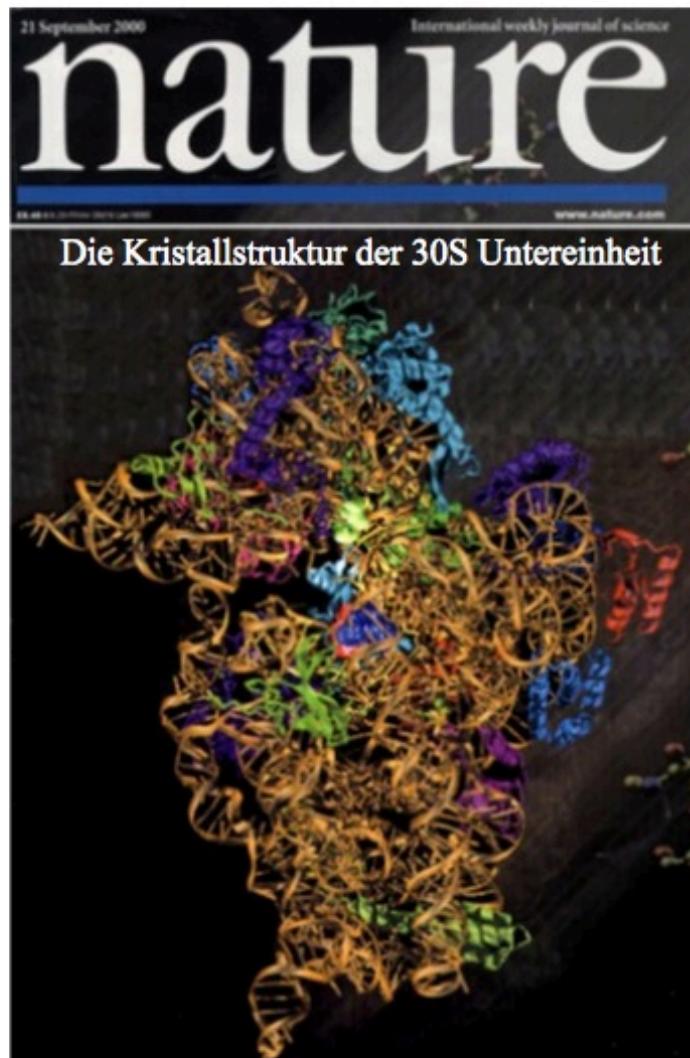


Hochtox. für RNA Pol II bei Eukaryoten. Wird in Leber transportiert und hemmt dort Proteinbiosynthese.

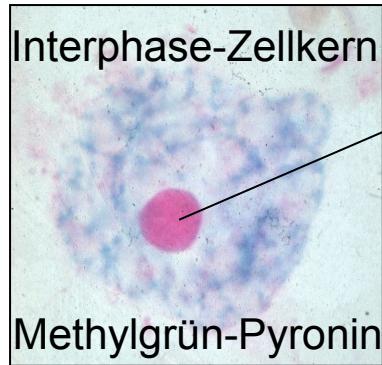


**RNA Pol I**  
**(rRNA-Gene)**

# Ribosomen = rRNA plus Proteine



# Organisation der rRNA-Gene („rDNA“)



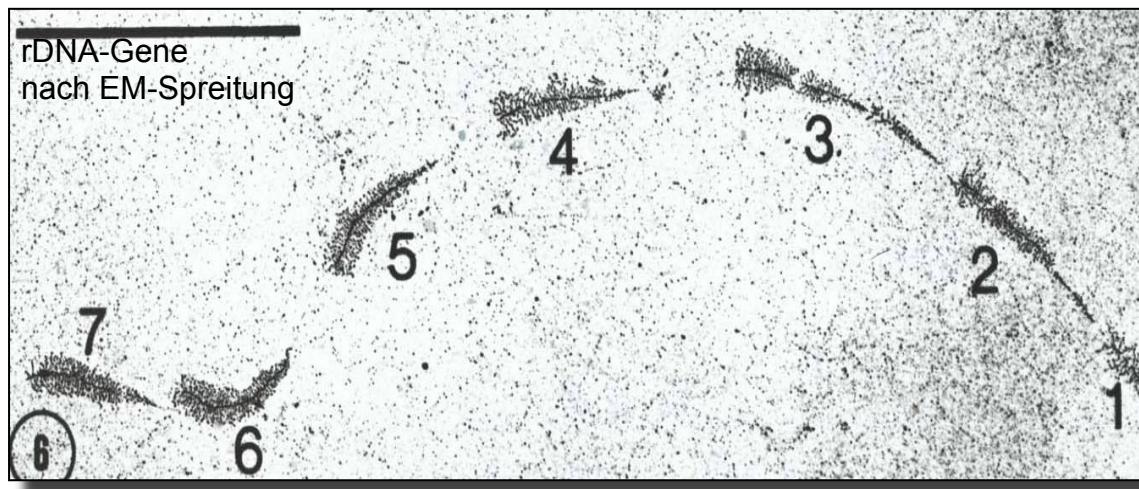
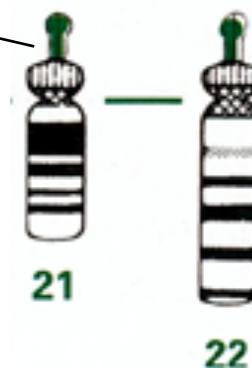
## In Interphase:

# Nukleolus

- Transkription der rDNA-Gene
  - Prozessierung der rRNA
  - Reifung der Ribosomen

## in Metaphase:

NOR-

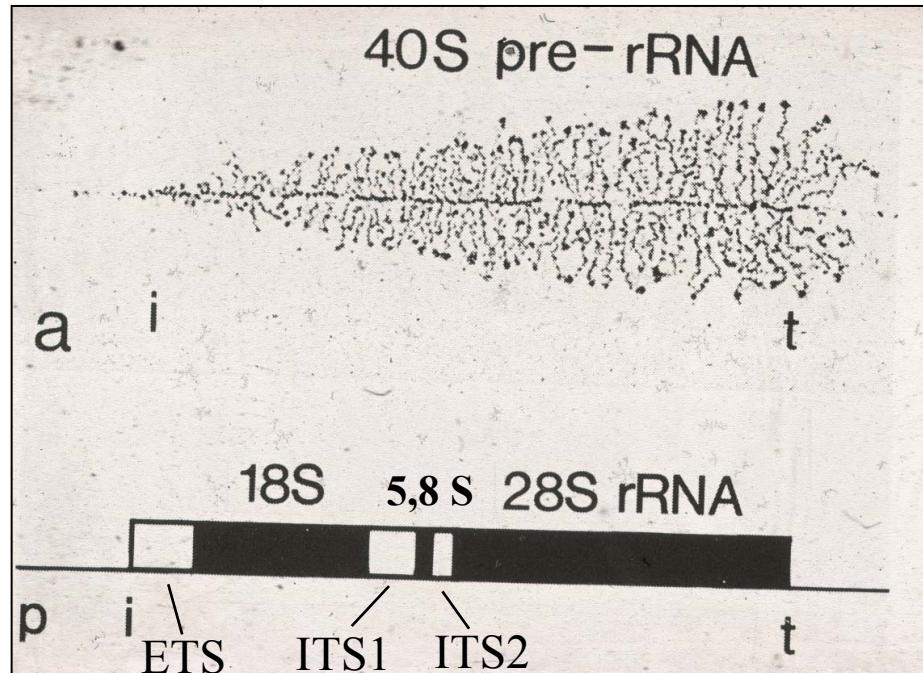


rDNA-Gene sind  
**tandem-repetitiv**  
organisiert

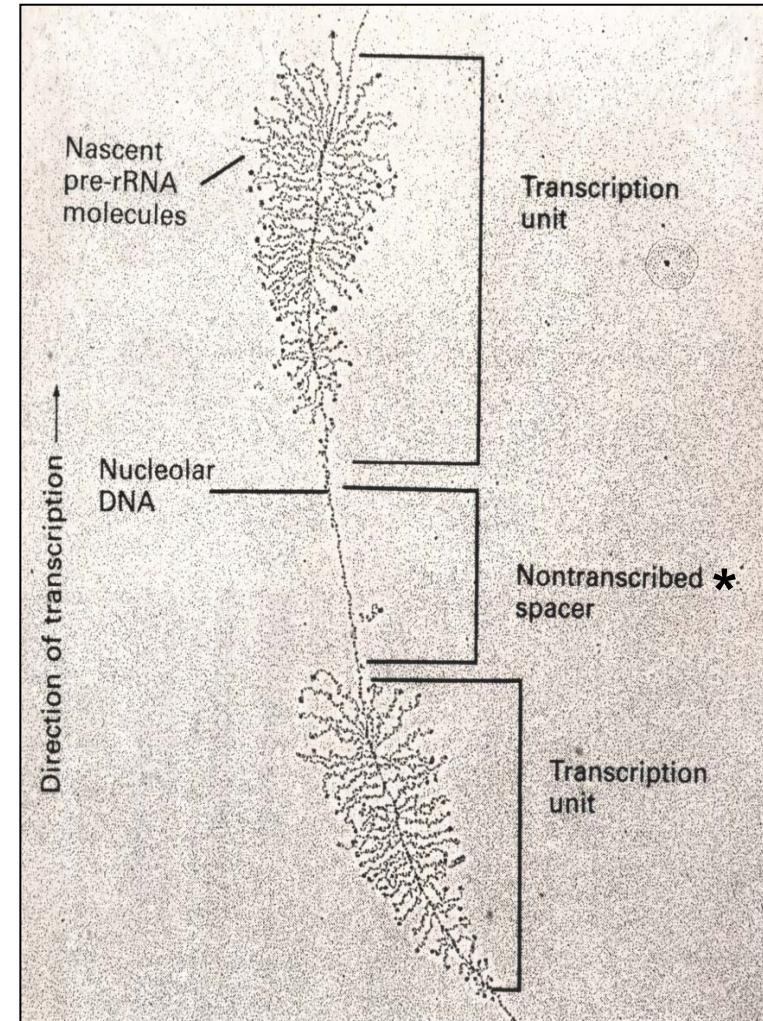
(ähneln also z. B.  
Satelliten-DNA im Zentromer)

# Organisation der rRNA-Gene

Spreitung > Elektronenmikroskop

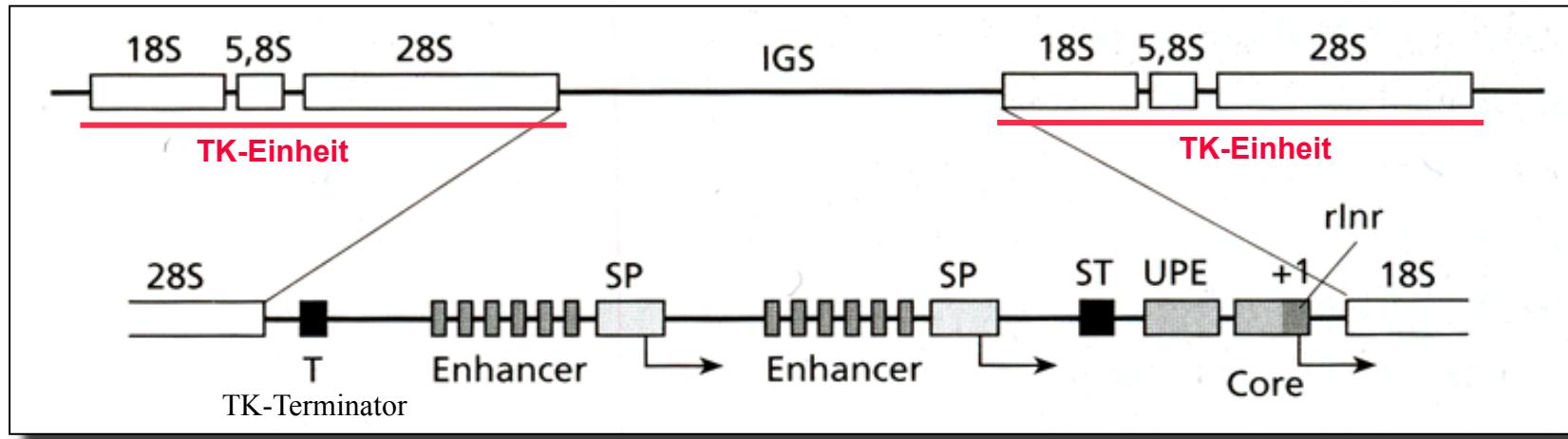


p Promoter  
i Initiation der TK  
t Termination der TK  
ITS internal transcribed spacer 1, 2  
ETS external transcribed spacer



\* = veraltet. Heute: Intergenic Spacer

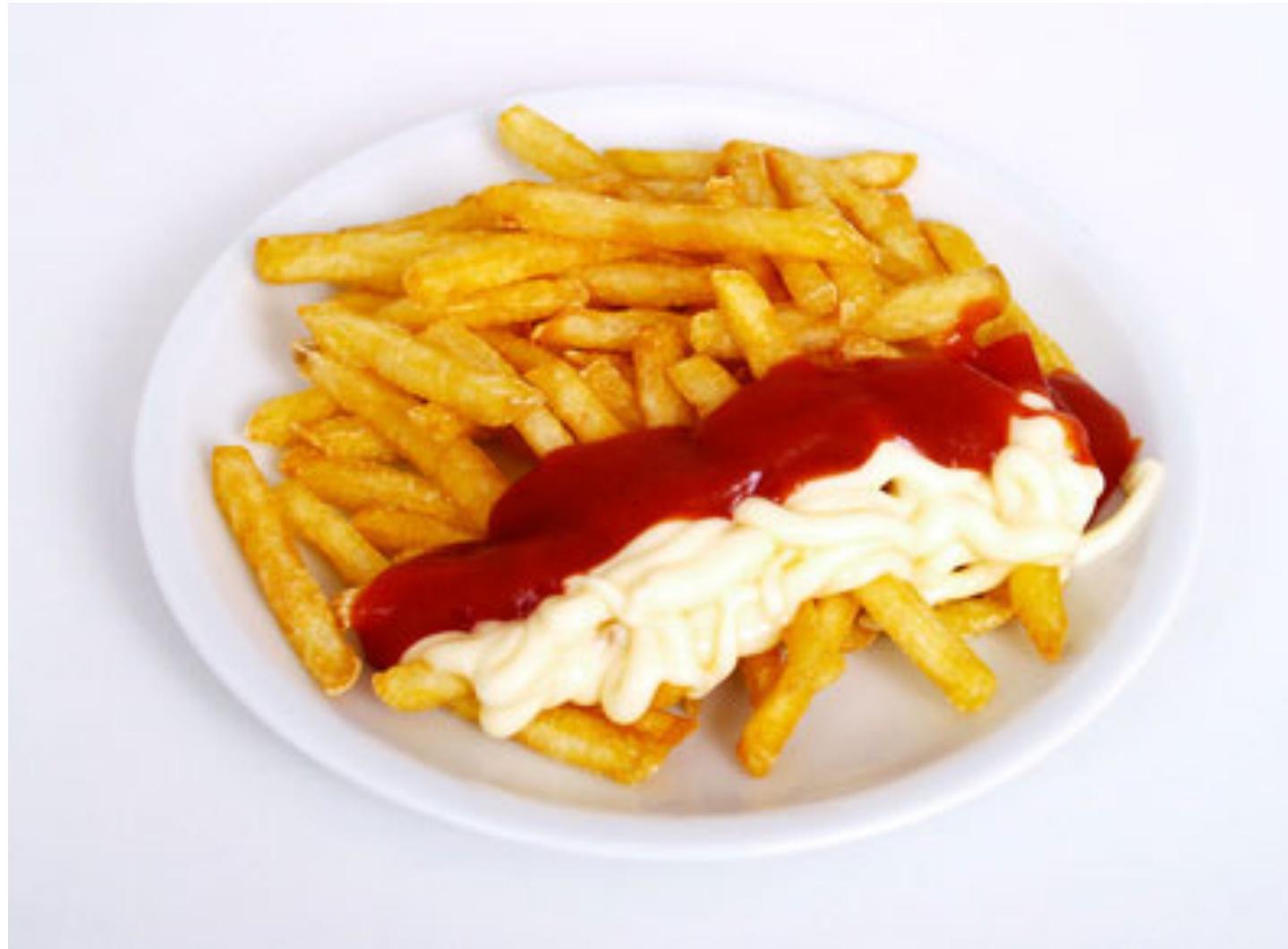
# RNA Pol I-Promoterstruktur der rDNA



- 18S-5,8S-28S Transkriptionseinheit ist **tandem-repetitiv**
- Geneinheiten sind durch *intergenic spacer* (IGS) getrennt
- Promoter-Elemente:
  - sp spacer promoter
  - st spacer terminator
  - upe upstream promoter element
  - rlnr ribosomales Initiator-Element



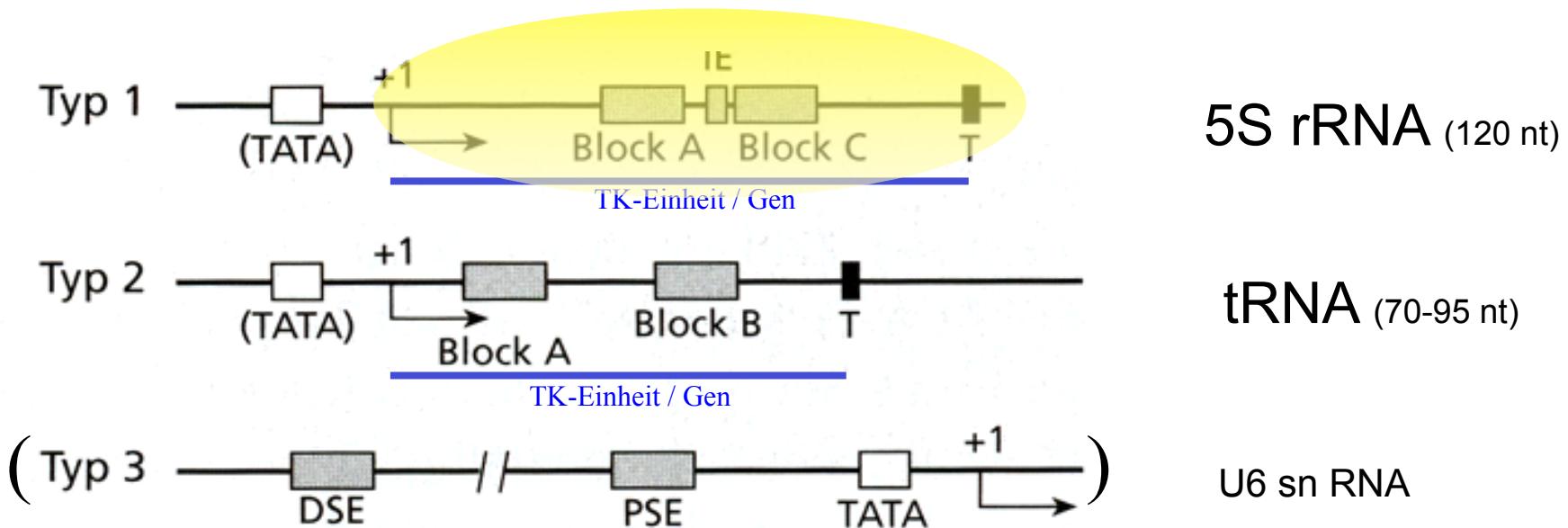
Ein Transkript für mehrere Genprodukte:  
**Stöchiometrische Verhältnisse...**



# **RNA Pol III**

## **(tRNA-Gene)**

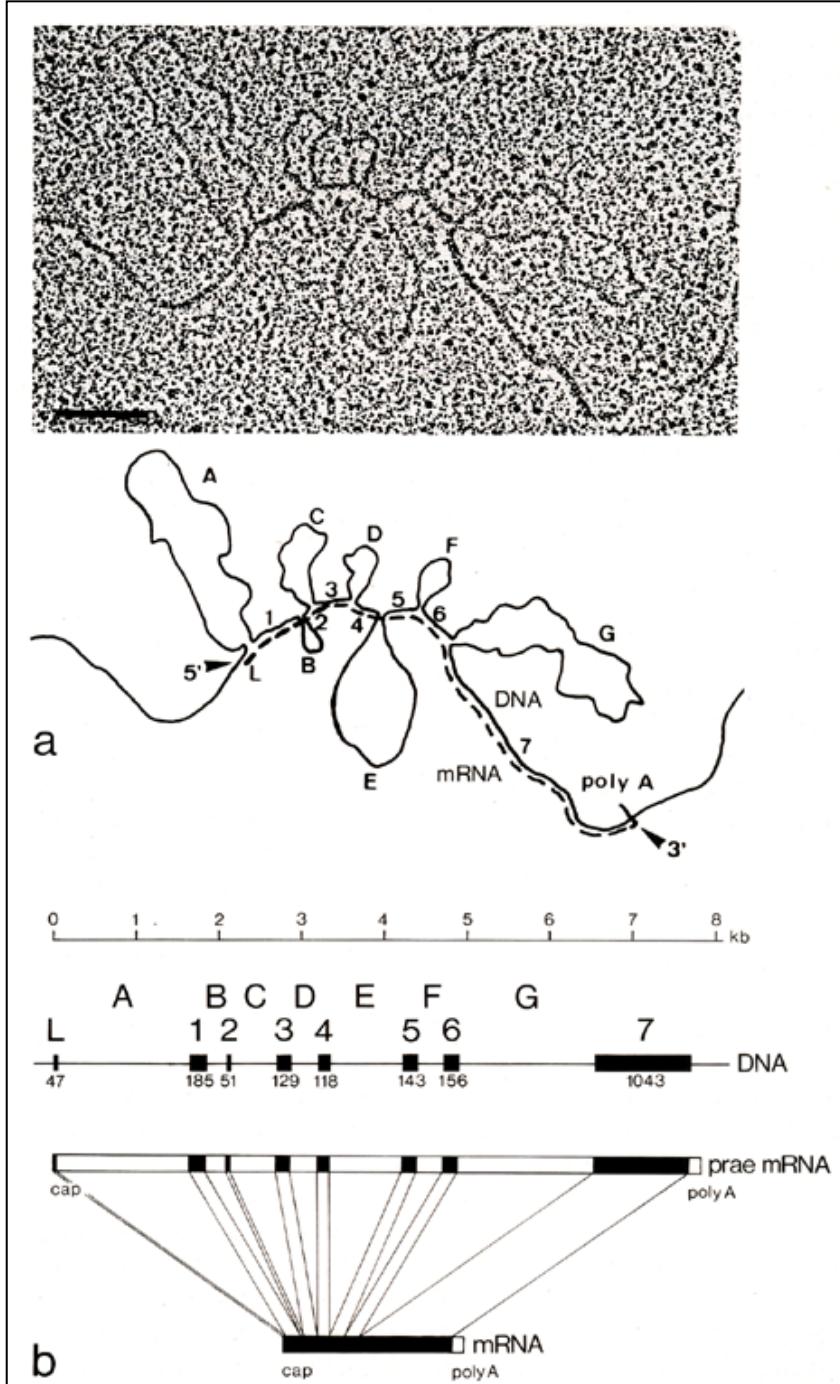
# Struktur von RNA Pol III-Genen



- Achtung: **intragenischer Promoter!**
- TATA-Box bei Typ 1,2 optional

# RNA Pol III

## (Protein- kodierende Gene)



# 1977: Mosaik-Struktur in eukaryotischen Gene

R-Loop Analyse:  
 ...DNA schmelzen  
 ...mit RNA hybridisieren  
 ...im EM anschauen



The Nobel Prize in Physiology or Medicine 1993  
 Richard J. Roberts, Phillip A. Sharp



The Nobel Prize in Physiology or Medicine 1993

Nobel Prize Award Ceremony

Richard J. Roberts

Philip A. Sharp



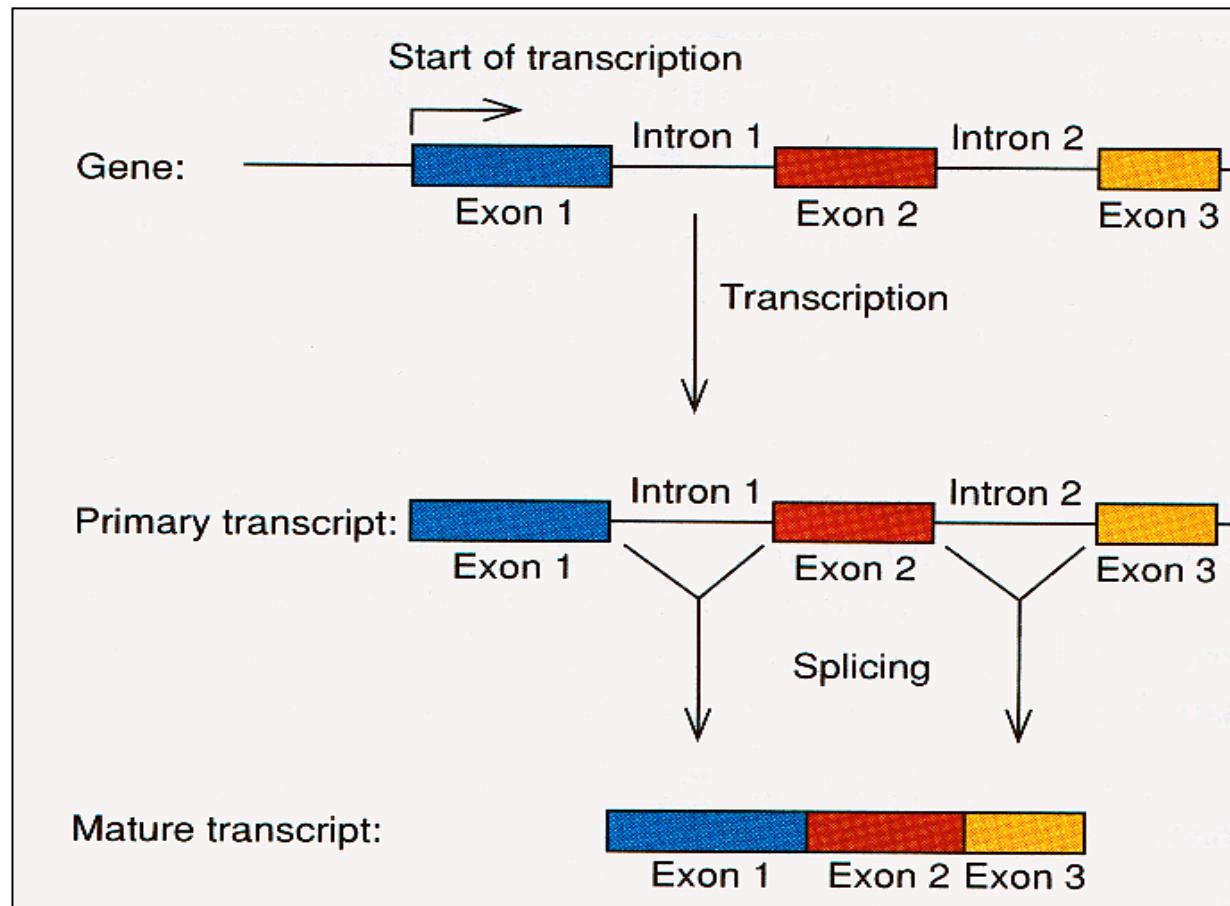
Richard J. Roberts



Phillip A. Sharp

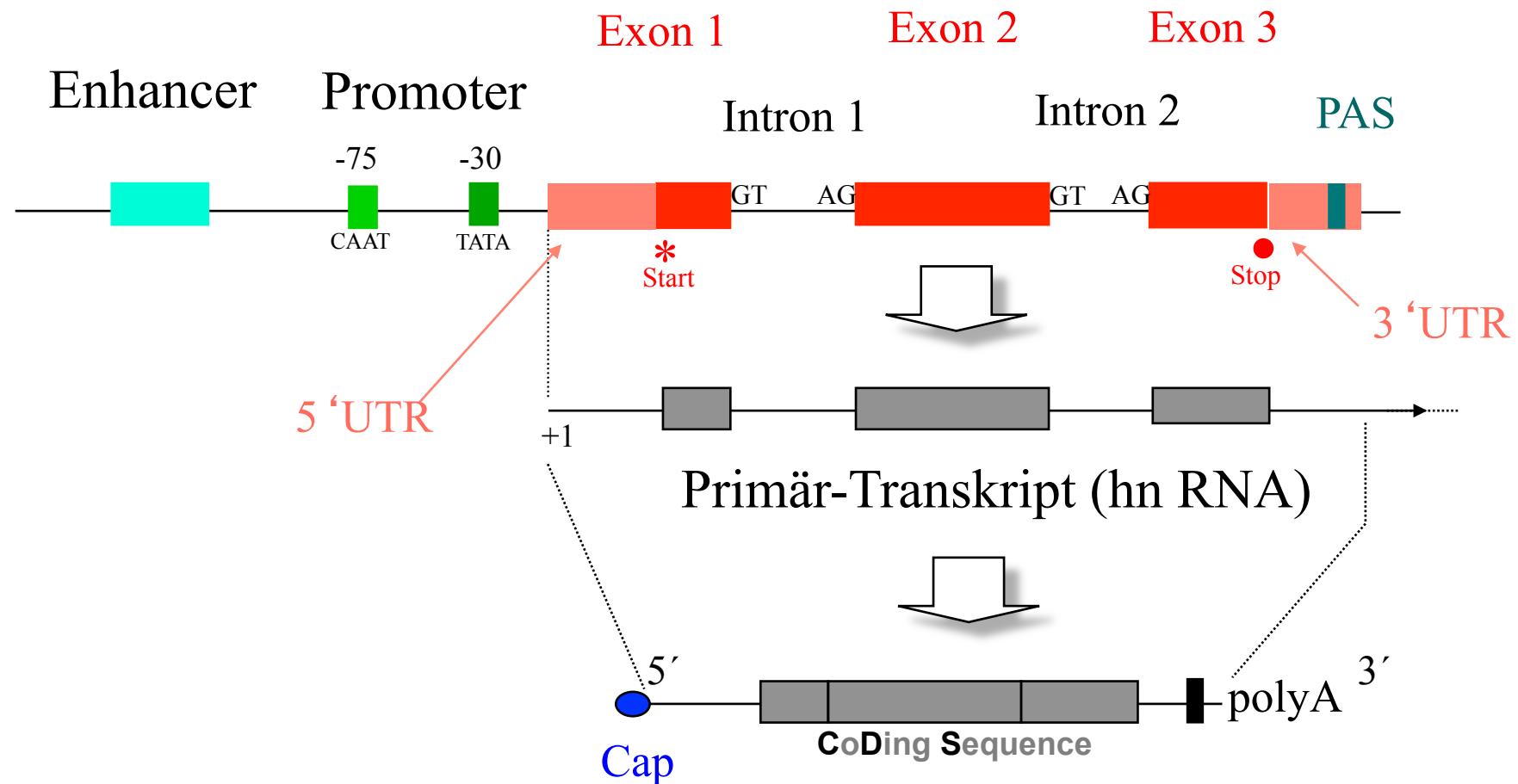
The Nobel Prize in Physiology or Medicine 1993 was awarded jointly to Richard J. Roberts and Phillip A. Sharp "for their discoveries of split genes"

# Introns werden aus mRNA herausgespleißt



**Def. „Exon“: Sequenz in reifer mRNA vorhanden!**

# Struktur von RNA Pol II-Genen

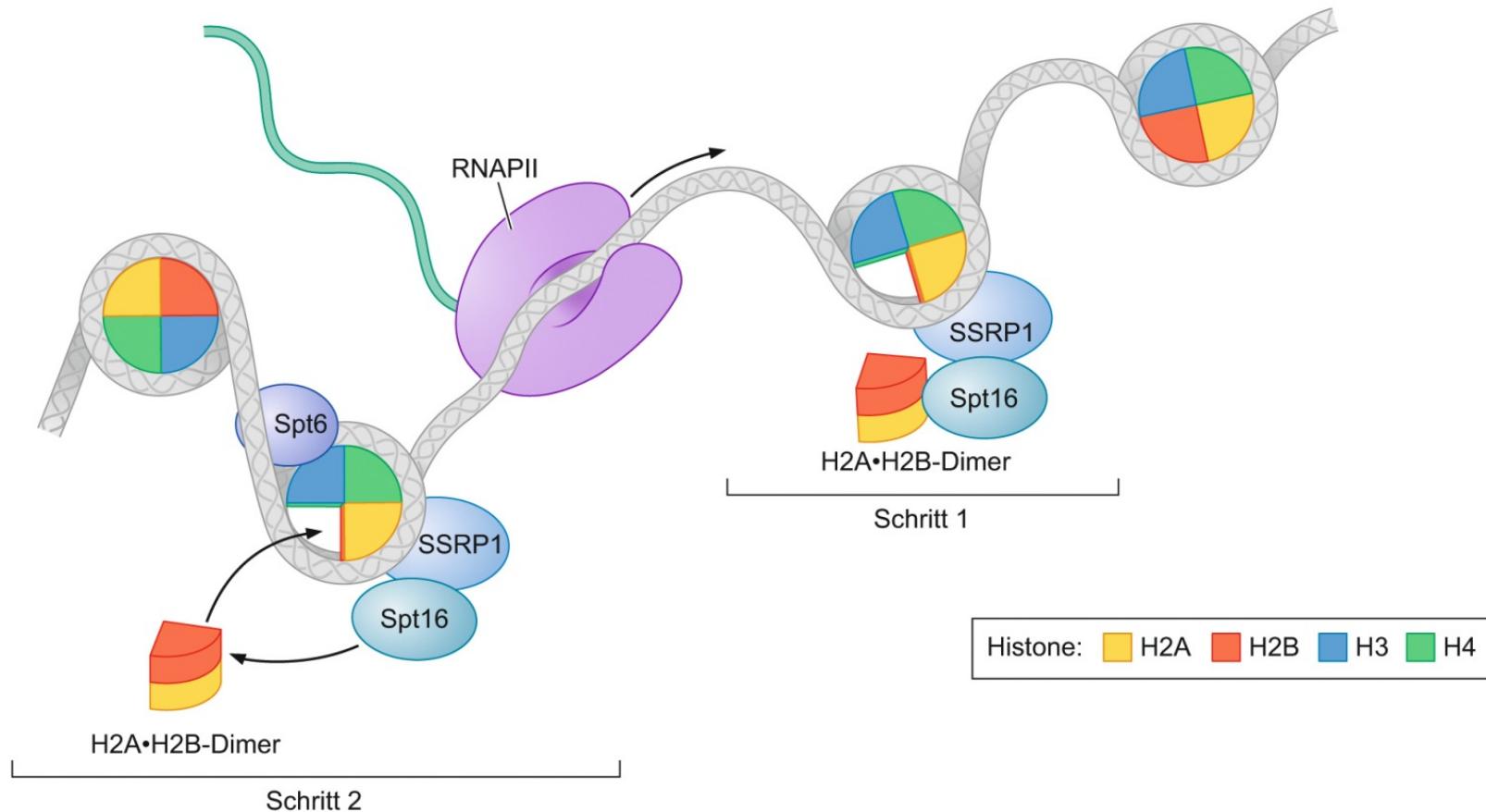


# Definitionen



- Exons sind Teil der reifen mRNA (E. sind nicht immer protein-kodierend!)
- Introns werden aus Primärtranskript (hnRNA) gespleißt.
- Introns beginnen mit GT und enden mit AG („GT-AG-Regel“)
- hnRNA = (längen)heterogene nukleäre RNA
- 5 '/3 ' UTR = 5 ' bzw. 3 ' liegende nicht-translatierte Regionen der reifen mRNA (können auch eigene Exons sein!)
- Enhancer bestimmen v.a. Spezifität der Transkription
- „+1“ = Transkriptionsstart (= erste transkribierte Nt-Position)
- PAS = Polyadenylierungssignal. Ca. 20 Nt abwärts des PAS wird die mRNA geschnitten und polyadenyliert.

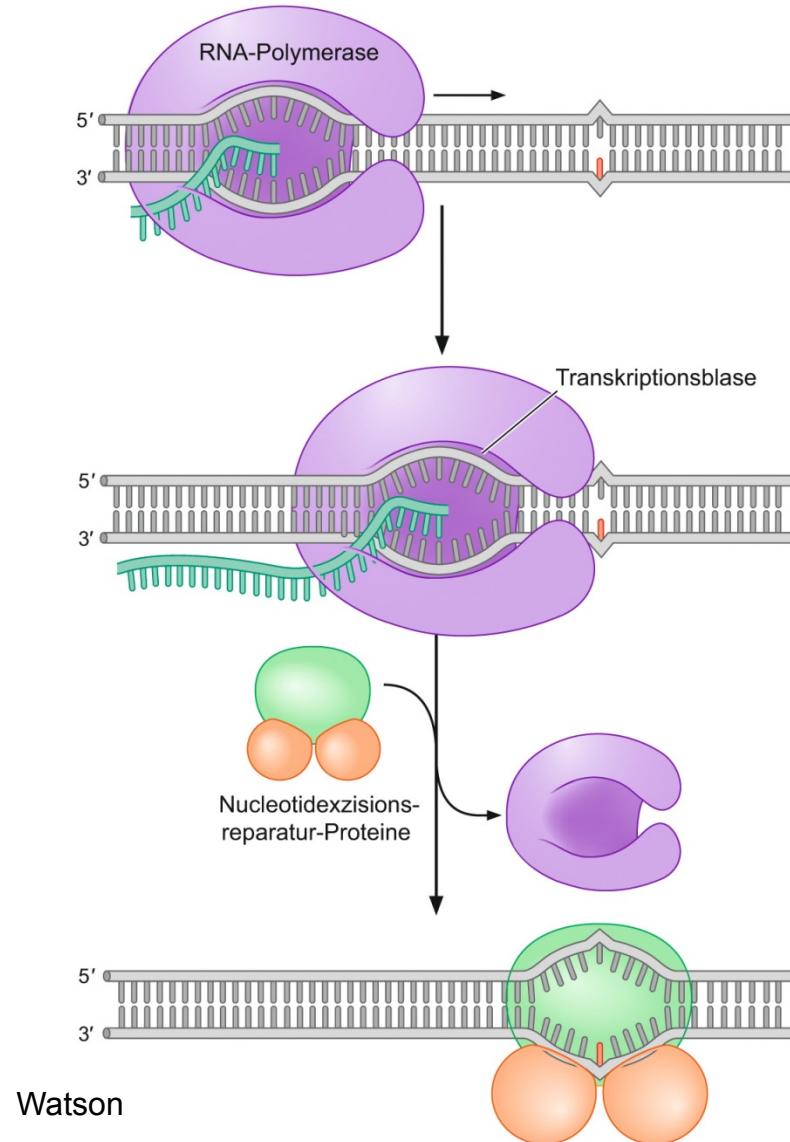
# Transkription & Chromatin



**Abbildung 12.22: Modell für die von FACT unterstützte Transkription nucleosomaler DNA.** Wie im Text beschrieben wird, kann der Faktor FACT, der hier als Heterodimer aus Spt16 und SSRP1 dargestellt ist, die Histon-Kerne von Nucleosomen vor der transkribierenden RNA-Polymerase II (teilweise) zerlegen (Schritt 1) und hinter der Polymerase (RNAPII) wieder zusammensetzen (Schritt 2). Dabei entfernt FACT vor der Polymerase jeweils ein H2A•H2B-Dimer und fügt hinter der Polymerase wieder ein H2A•H2B-Dimer hinzu. Ein weiterer Faktor namens Spt6 bindet an Histon H3 und unterstützt vermutlich die Regeneration des Histon-Kerns.

Watson

# TK-gekoppelte Reparatur der DNA!



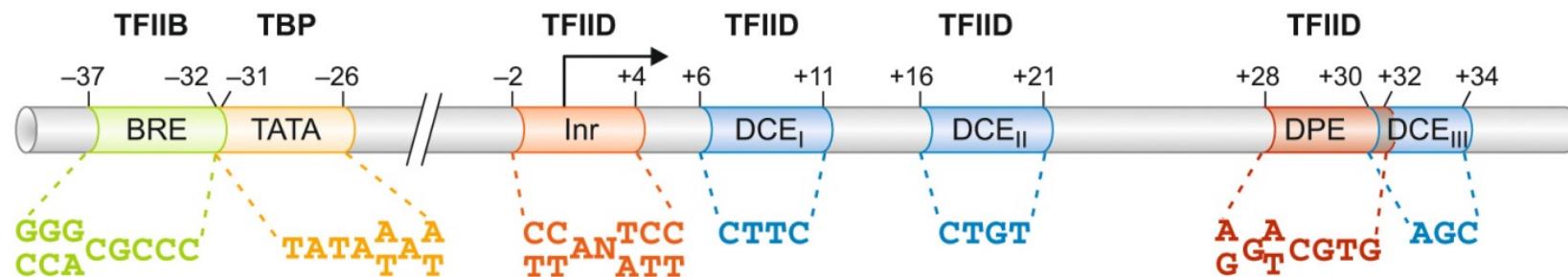
Gestoppte RNA Pol II  
detektiert DNA-Schaden  
und rekrutiert Proteine  
der Nukleotid-Exzisions-  
Reparatur.

## Cockayne-Syndrom

- autosom. rezess.
- selten
- Progerie
- Lebenserwartung 6-12 J.

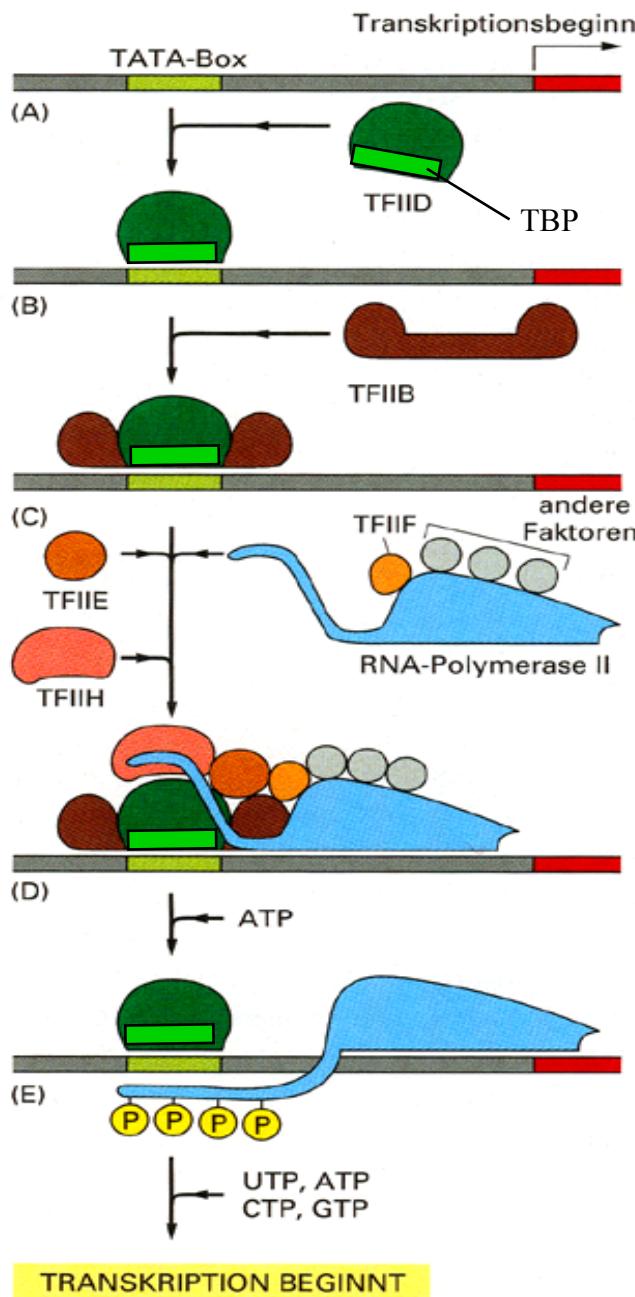


# Promoter eines RNA Pol II-Gens



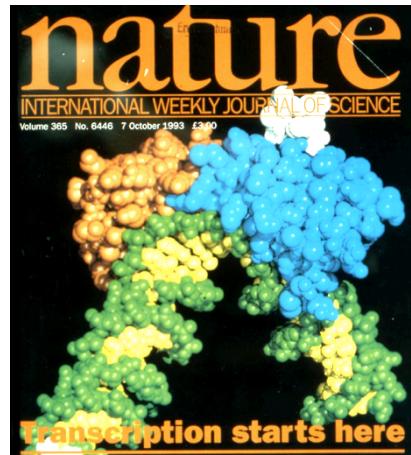
**Abbildung 12.14: Sequenzelemente des RNA-Polymerase-II-Kernpromotors.** Die Abbildung zeigt die Positionen der verschiedenen DNA-Sequenzelemente, die im Pol-II-Kernpromotor vorkommen, relativ zum Transkriptionsstartpunkt. Die Position des Startpunkts ist durch den Anfang des rechtwinkligen Pfeils über der DNA gekennzeichnet. Folgende Elemente, die im Text näher beschrieben werden, treten als Bestandteil des Kernpromotors auf: BRE (für engl. *B recognition element*, TFIIB-Erkennungselement); TATA (TATA-Box); Inr (Initiator); DPE (Downstream-Promotorelement); DCE (Downstream-Kernelement; engl. *downstream core element*). Ein weiteres Promotorelement namens MTE (für engl. *motif ten element*), das im Text beschrieben wird, in der Abbildung aber nicht dargestellt ist, liegt unmittelbar stromaufwärts vom DPE. Außerdem sind in der Abbildung die Konsensussequenzen der einzelnen Elemente angegeben, die auf die gleiche Weise bestimmt wurden, die für den Fall der bakteriellen Promotorelemente in Box 12.1 beschrieben wurde. Über den einzelnen Promotorelementen in der DNA sind darüber hinaus die allgemeinen Transkriptionsfaktoren genannt, durch welche die jeweiligen Elemente erkannt werden.

- Eukaryotische Promoter bestehen aus konservierten kurzen Sequenzabschnitten, sog. „Boxen“.
- An die „Boxen“ binden Proteine, die die Transkription regulieren („Transkriptionsfaktoren“).

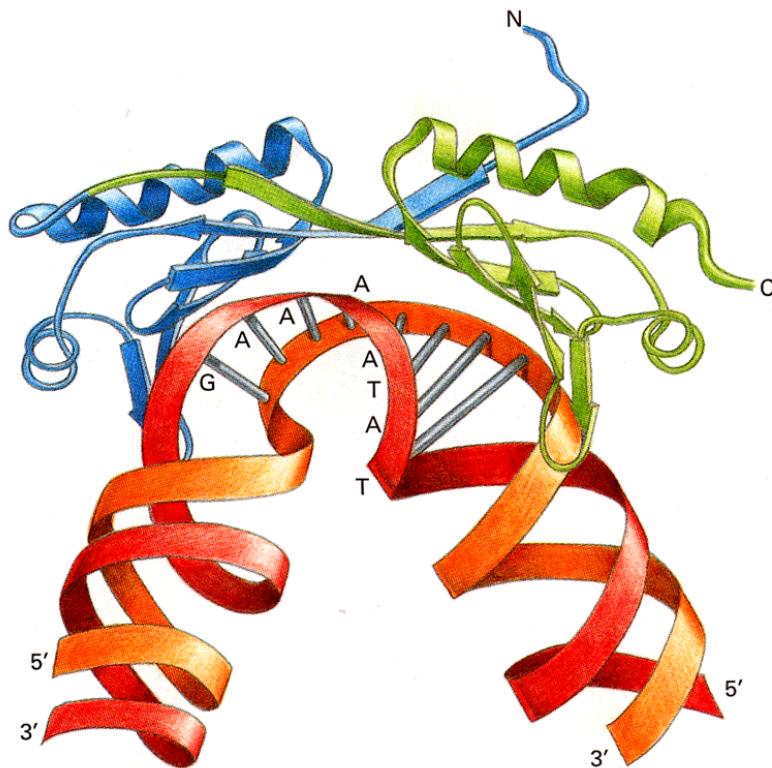


# RNA Pol II- Initiation durch „basale“ Transkriptionsfaktoren

- TBP/TFIID-Bindung an TATA-Box
- TFIIIF rekrutiert RNA Pol II
- TFIIH-Bindung:
  - wirkt als Helikase (ATP-Verbrauch!)
  - hängt als Kinase Phosphatgruppen an C-Terminale Domäne der RNA Pol II
- > RNA Pol II beginnt TK
- > die meisten TFs verschwinden
- > TBP initiert weitere TK-Runde

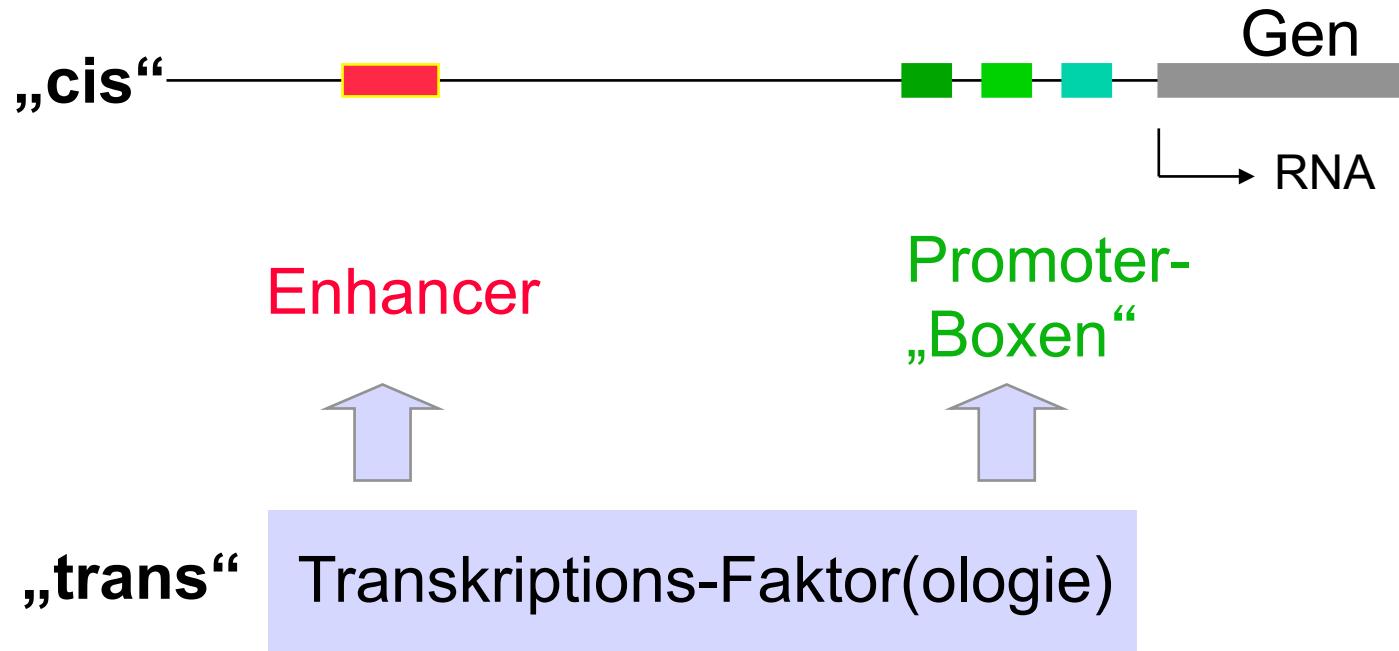


# Das TATA-Box-Bindeprotein TBP

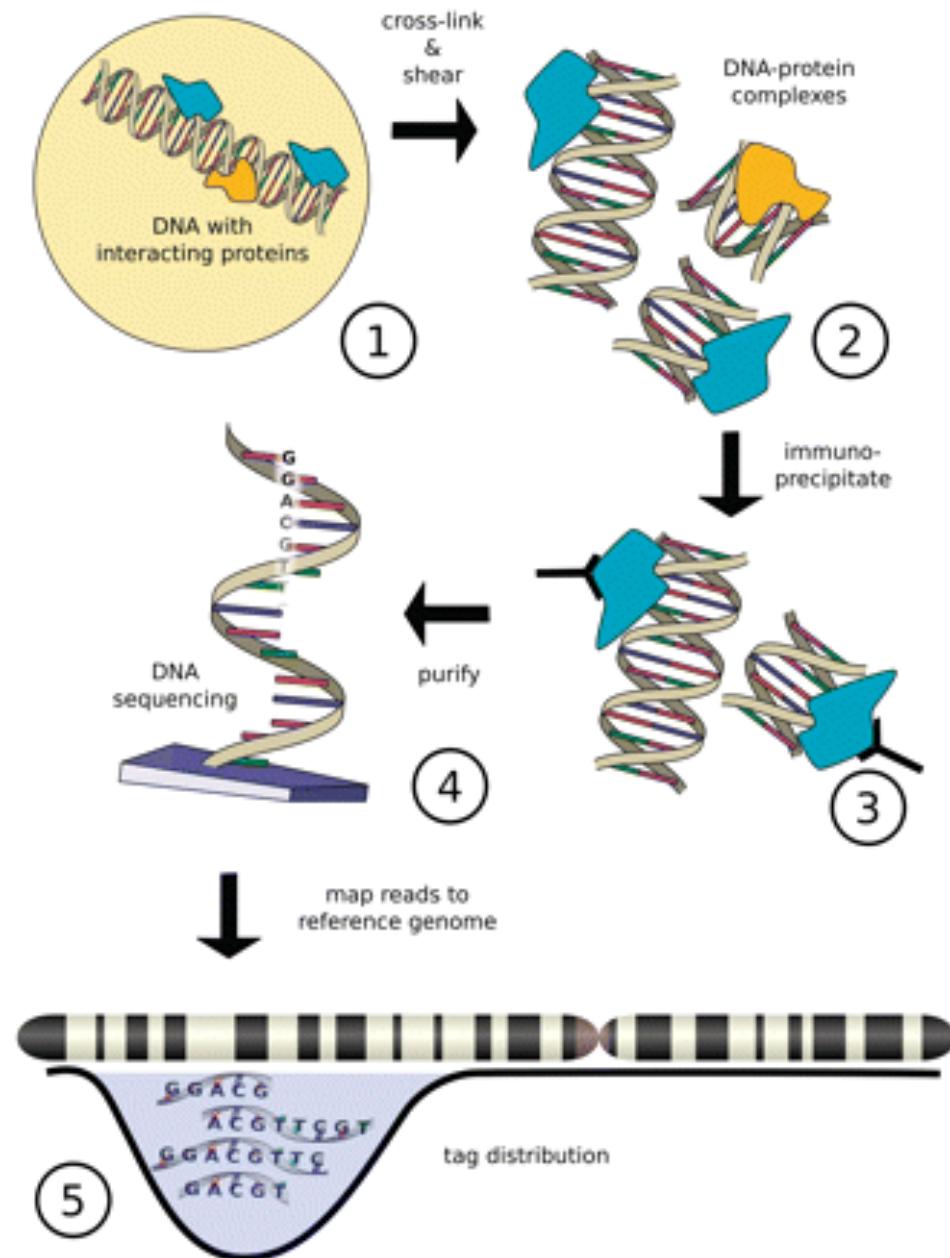


- TBP ist UE von **TFIID**
- TBP induziert Biegung der DNA  
  > partielle Entwindung
- A:T-Basenpaare erleichtern Biegung
- TBP bindet in der kleinen Grube sequenzspezifisch
- TBP wird von allen 3 RNA Pol 's benötigt!

# Genregulation in Eukaryoten: z.B. auf Transkriptionsebene



- „basale“ TK-Faktoren, in variabler Kombination in vielen Promotern
- Zell/Gewebe-spezifische TK-Faktoren, induzierbare Faktoren (siehe VL Genregulation)

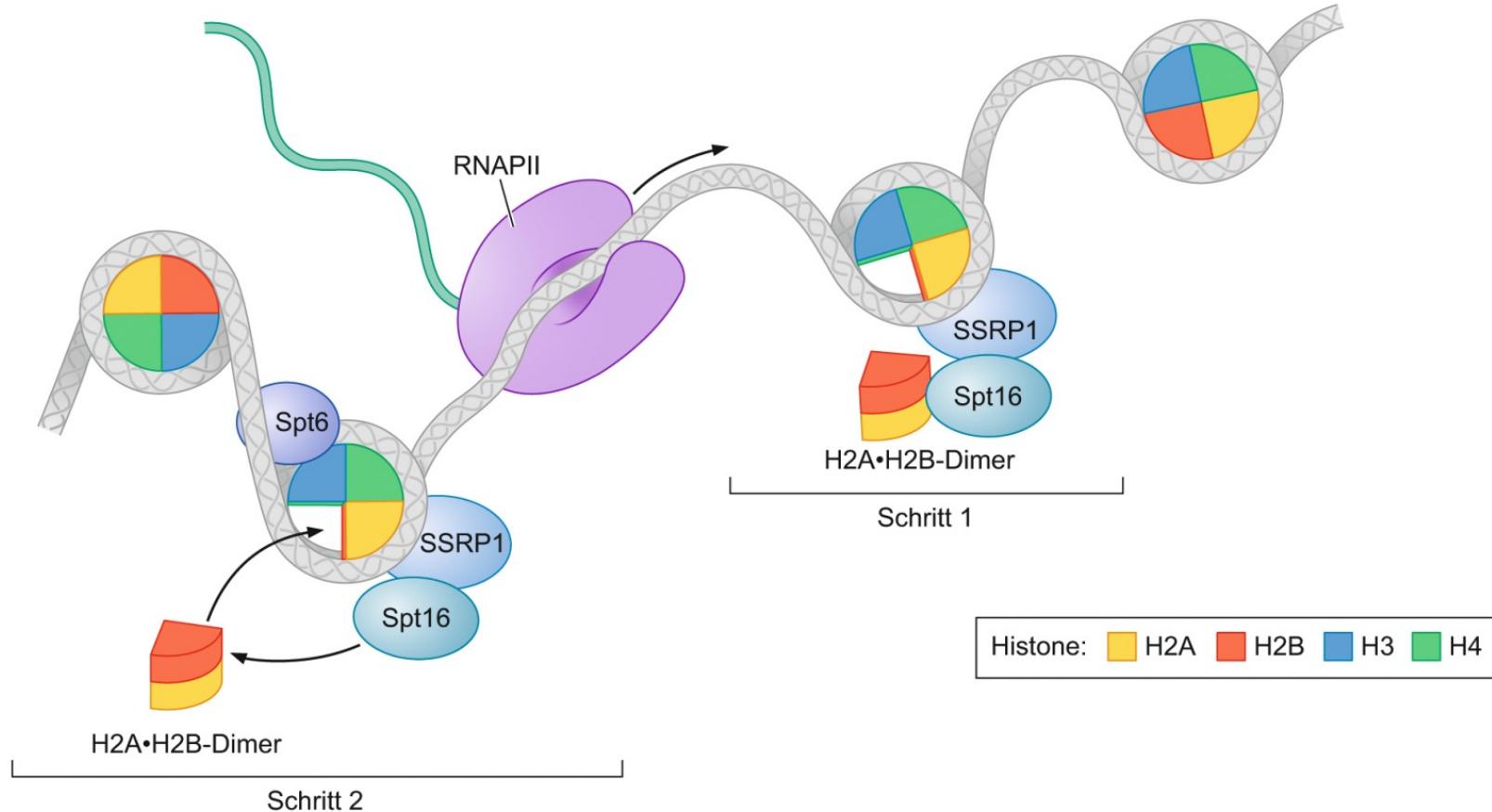


# Wo überall binden Regulatorproteine im ganzen Genom?

„ChIP-Seq“

**Chromatin-  
Immunopräzipitation &  
Sequenzierung**

# Was passiert mit dem Chromatin während der Transkription?

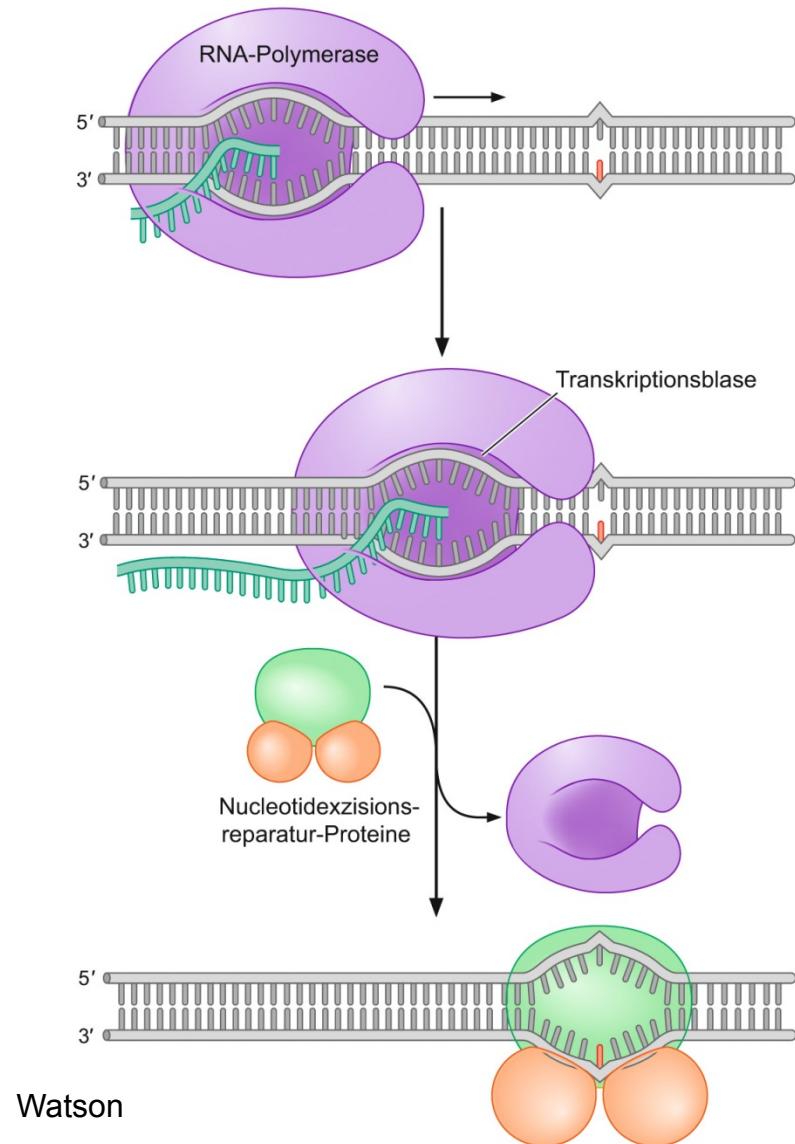


**Abbildung 12.22: Modell für die von FACT unterstützte Transkription nucleosomaler DNA.** Wie im Text beschrieben wird, kann der Faktor FACT, der hier als Heterodimer aus Spt16 und SSRP1 dargestellt ist, die Histon-Kerne von Nucleosomen vor der transkribierenden RNA-Polymerase II (teilweise) zerlegen (Schritt 1) und hinter der Polymerase (RNAPII) wieder zusammensetzen (Schritt 2). Dabei entfernt FACT vor der Polymerase jeweils ein H2A•H2B-Dimer und fügt hinter der Polymerase wieder ein H2A•H2B-Dimer hinzu. Ein weiterer Faktor namens Spt6 bindet an Histon H3 und unterstützt vermutlich die Regeneration des Histon-Kerns.

Watson

Alles hängt miteinander zusammen...

# TK-gekoppelte Reparatur der DNA!



Gestoppte RNA Pol II detektiert DNA-Schaden und rekrutiert Proteine der Nukleotid-Exzisions-Reparatur (VL dazu 2019).

## Cockayne-Syndrom

- autosom. rezess.
- selten
- Progerie
- Lebenserwartung 6-12 J.



# Prozessierung der mRNA in Eukaryoten

- Anhängen eines **CAP** am 5 'Ende
- Heraus-**Spleißen** der Introns
- Anhängen eines **poly A**-Schwanzes am 3 'Ende
- nachträgliche Korrektur der mRNA-Sequenz  
("RNA editing")

grobe  
zeitliche  
Abfolge

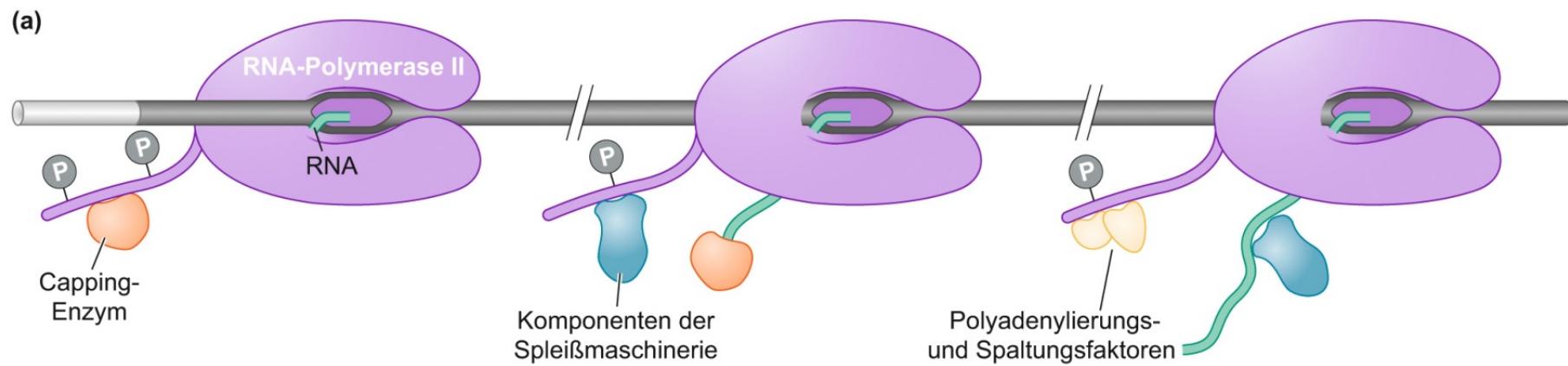
---

alles noch  
während  
Transkription  
im Zellkern

---

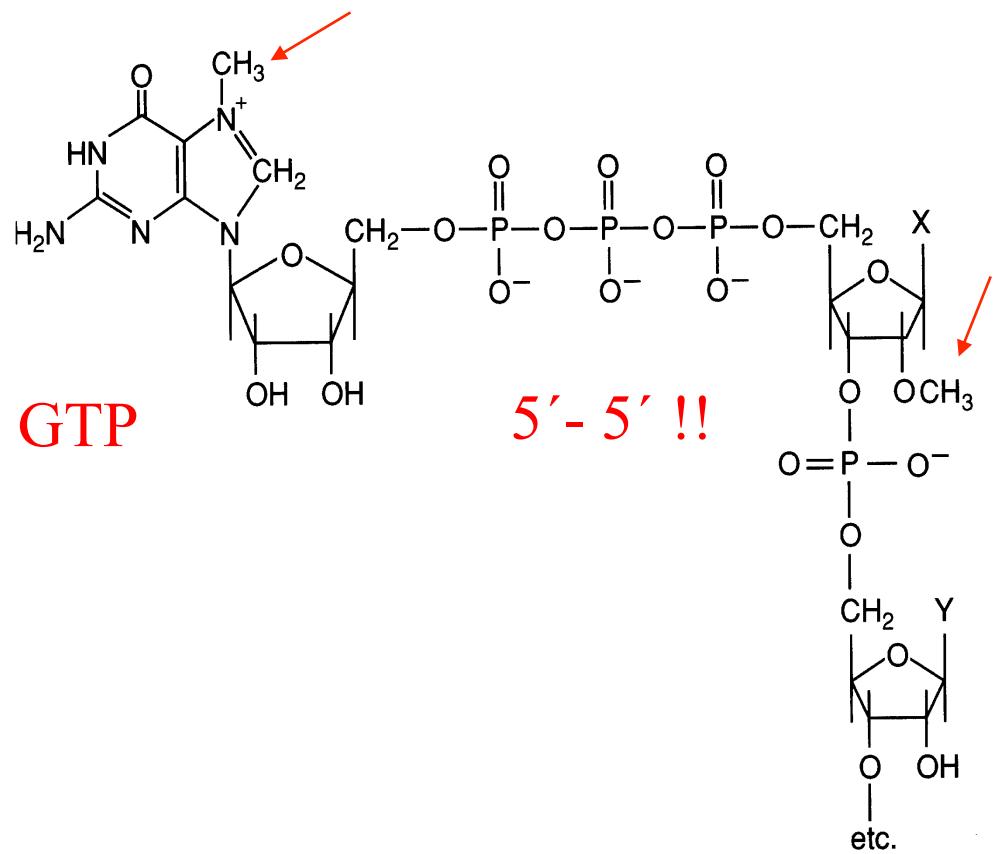
Bei tRNA und rRNA gibt es ebenfalls  
Prozessierung...

# Prozessierung der mRNA in Eukaryoten



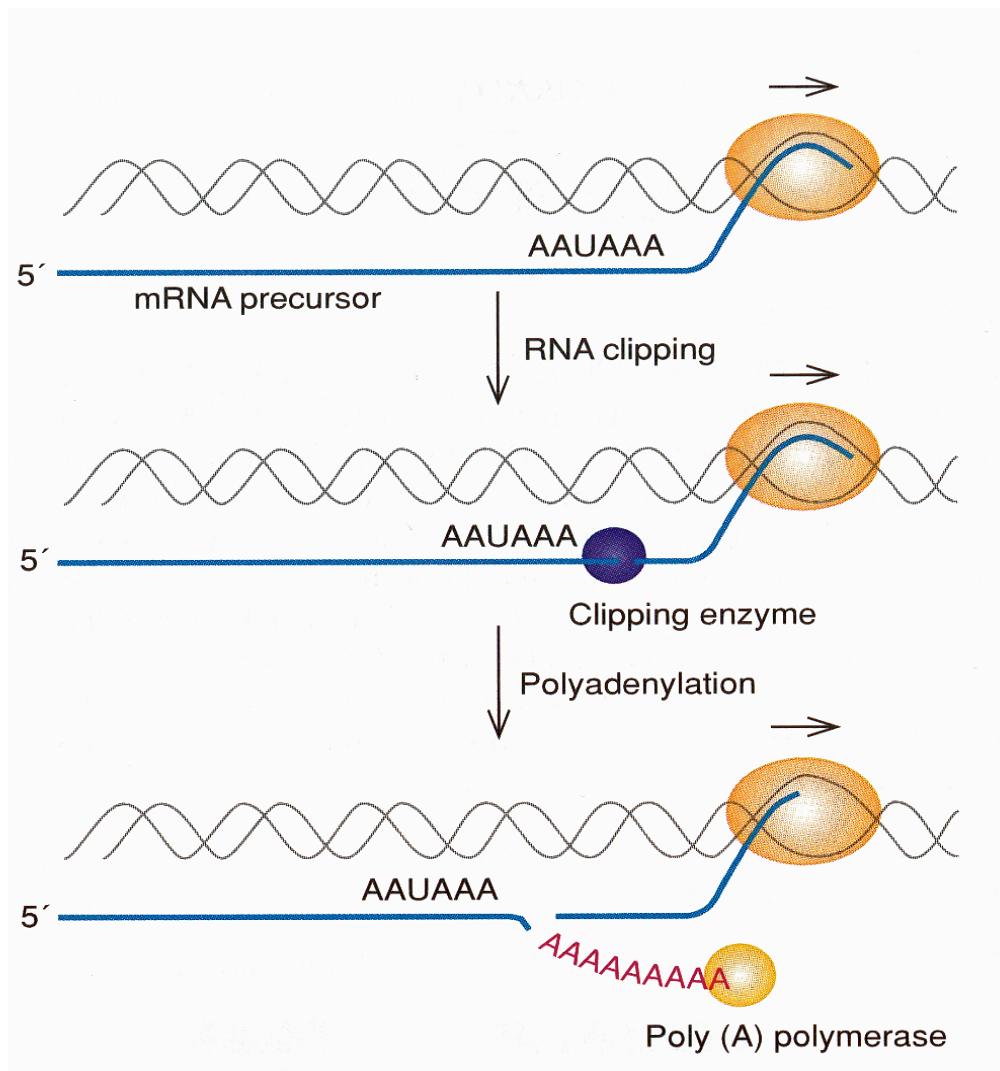
Die RNA Pol II koordiniert die Prozessierung der mRNA

# Cap am 5'Ende eukaryotischer mRNA



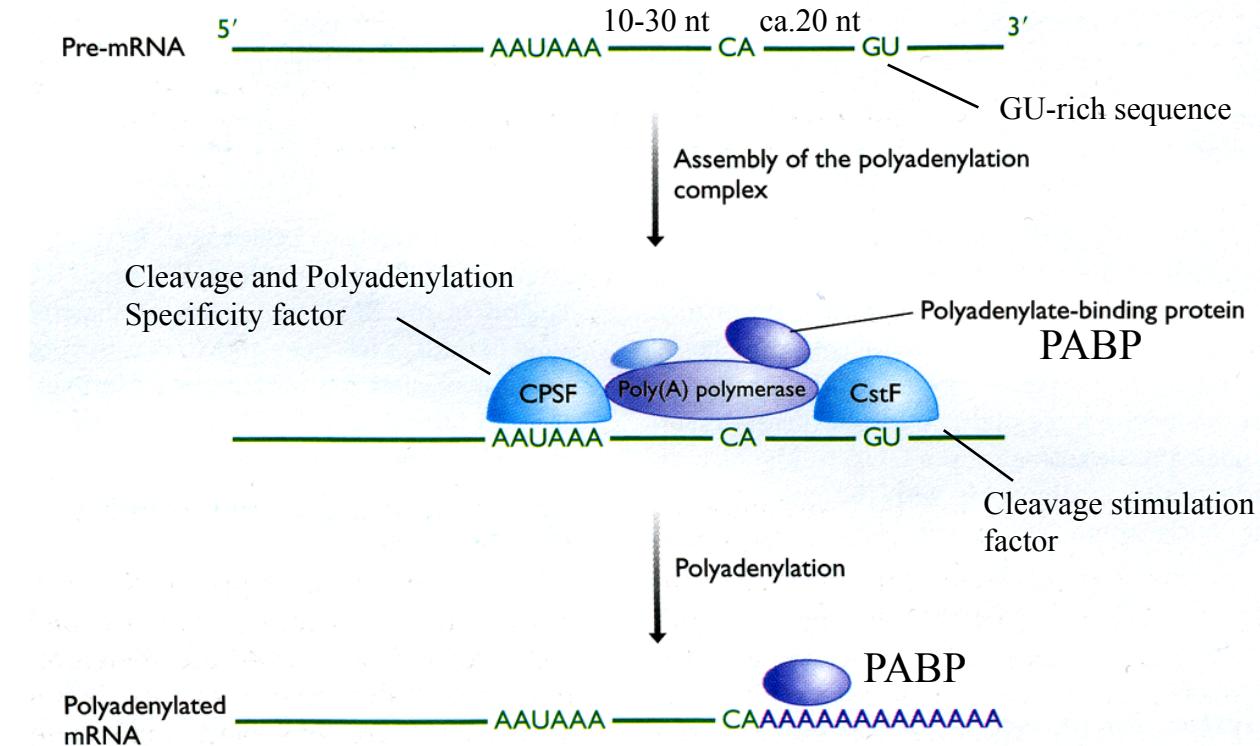
- Schutz der mRNA  
(>Genregulation)
- Bindung der mRNA  
an das Ribosom
- bereits im Nukleus  
angehängt durch  
Guanylyl-Transferase

# Termination der TK und Poly-Adenylierung am 3'-Ende eukaryotischer mRNA



- eigentliche Termination ist „inexakt“
- 150-200 A's noch im Kern angehängt
- Stabilität der mRNA (>Genregulation!)

# Poly-Adenylierung



- PABP hilft ' PolyA-Polymerase und stabilisiert ,tail ' (>Genregulation)
- Polyadenylierung ist Teil der TK!  
Die C-terminale Domäne der RNA Pol II wird für Polyadenylierung benötigt.

## REVIEW SUMMARY

### TRANSCRIPTION

# Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut

Nick J. Proudfoot\*

**BACKGROUND:** Genes correspond to single transcription units, starting from the promoter and ending at the terminator. Terminating gene transcription is directly coupled to mRNA processing, which occurs cotranscriptionally. When RNA polymerase II (Pol II) reaches the gene end, it first slows down over the terminator. This is partly because 3'-end cleavage and polyadenylation (CPA) complex is recruited onto Pol II when poly(A) signals appear in the nascent transcript. This nascent transcript will often invade the DNA duplex to form an R-loop structure, which induces further polymerase slowdown. During this time, CPA releases mRNA from chromatin into eventual cytoplasmic translation. Pol II continues to transcribe its DNA template after mRNA release. However, this is short-lived, as an exonuclease (Xrn2) degrades the transcript from its 5' end. When this molecular torpedo catches up with Pol II, then conformational shockwaves are transmitted into its active site, which releases Pol II from the DNA template.

#### ON OUR WEBSITE

Read the full article  
at <http://dx.doi.org/10.1126/science.aad9926>

Pol II is then free to restart transcription on another gene promoter.

**ADVANCES:** The above process of Pol II termination appears surprisingly complex, but provides unanticipated layers of gene regulation. First, most protein-coding genes generate multiple mRNAs of different lengths caused by the use of alternative poly(A) sites (APA), which in turn is dictated by gene termination. Alternative mRNAs with shorter or longer 3'-untranslated region (3'UTR) sequences possess different sequence codes for how long to survive or where in the cell to translate their proteins. Second, termination of transcription is employed as a quality-control mechanism. Transcription errors occur either because the DNA template is damaged or because the RNA is mis-synthesized and induce premature termination before reaching the gene end. These truncated transcripts are rapidly degraded and, if Pol II becomes arrested on the DNA template, then it

is degraded in situ by the proteasome to allow subsequent rounds of transcription.

Recent studies reveal that cellular stress such as osmotic or heat shock, as well as viral infection or cancer-inducing mutations, can all promote aberrant termination. Under these varied conditions, many genes fail to terminate transcription. The resulting extensive readthrough transcription can cause massive deregulation of downstream gene expression.

**OUTLOOK:** Many questions remain about the mechanism and regulation of transcriptional termination. Exactly how degradation of the transcript by Xrn2 together with CPA and various helicases promotes Pol II termination remains poorly understood. Structural changes occurring within Pol II to promote this effect are currently unknown. Their resolution will need new technology, such as cryo-electron microscopy.

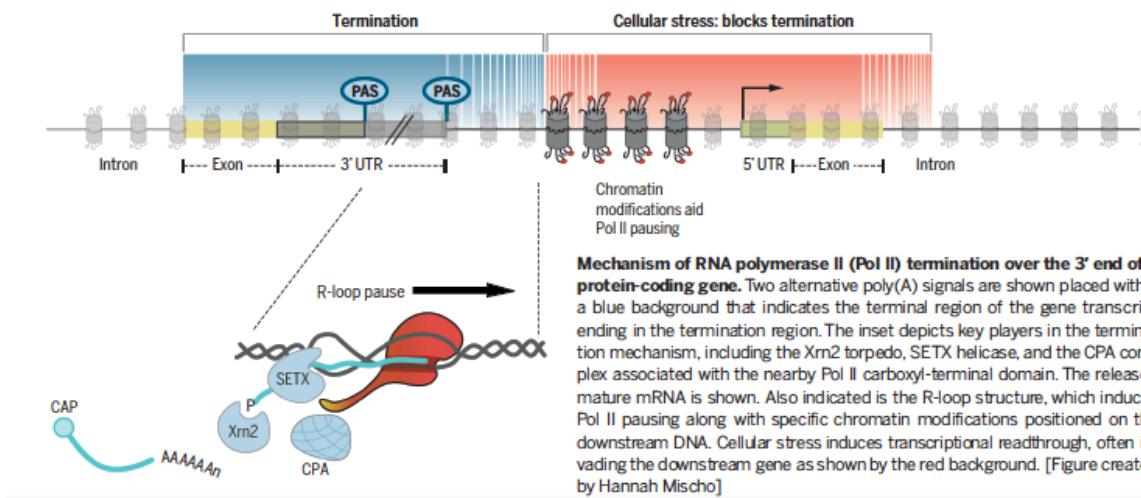
The regulation of APA to generate mRNA with different 3'UTRs is similarly poorly understood. Although changes in CPA factor levels or in the transcription process itself can affect APA, dominant factors and mechanisms used in biology to achieve this regulation remain enigmatic. The easy perturbation of termination resulting in readthrough transcripts appears to be at odds with the elaborate mechanisms in place to stop the Pol II juggernaut. It is evident that the field of Pol II termination has many surprises in store for future research into this fascinating process. ■



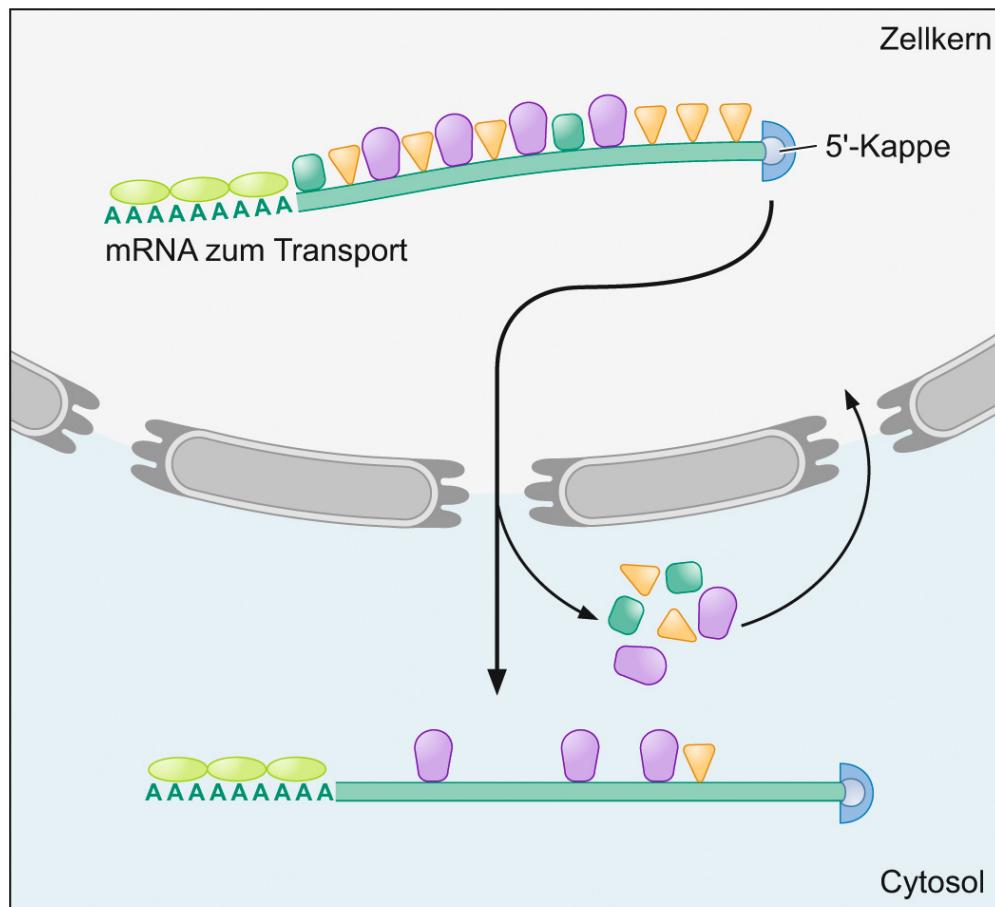
The list of author affiliations is available in the full article online.

\*Corresponding author. Email: nicholas.proudfoot@path.ox.ac.uk

Cite this article as N. J. Proudfoot, *Science* 352, aad9926 (2016). DOI: 10.1126/science.aad9926



# Export der mRNA ins Cytoplasma

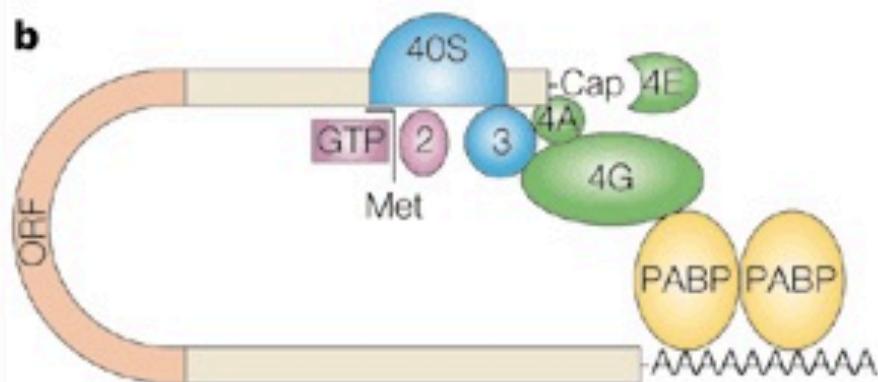
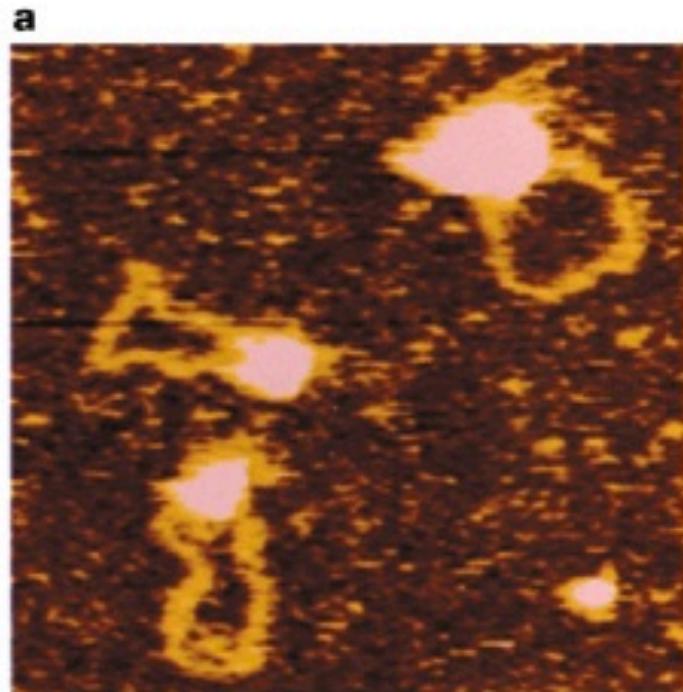


**Abbildung 13.32: Transport von mRNAs aus dem Zellkern heraus.** Der RNA-Export aus dem Zellkern ist ein aktiver Prozess und nur bestimmte (geeignete) RNAs werden für den Transport ausgewählt. Um dafür ausgewählt zu werden, muss die RNA die richtigen Proteine gebunden haben. Diese Proteine machen die RNA unterscheidbar von anderen, die im Zellkern zurückbleiben und gegebenenfalls dort abgebaut werden. Proteine, die etwa Exon:Exon-Grenzen erkennen, zeigen an, dass die mRNA richtig gespleißt ist. Proteine, die an Introns binden, zeigen dagegen an, dass eine RNA im Zellkern verbleiben soll. Einmal im Cytoplasma angelangt, dissoziieren einige Proteine ab, andere treten als Zeichen der Translationsfähigkeit hinzu.

Watson



# Vorbereitung der mRNA für die Translation

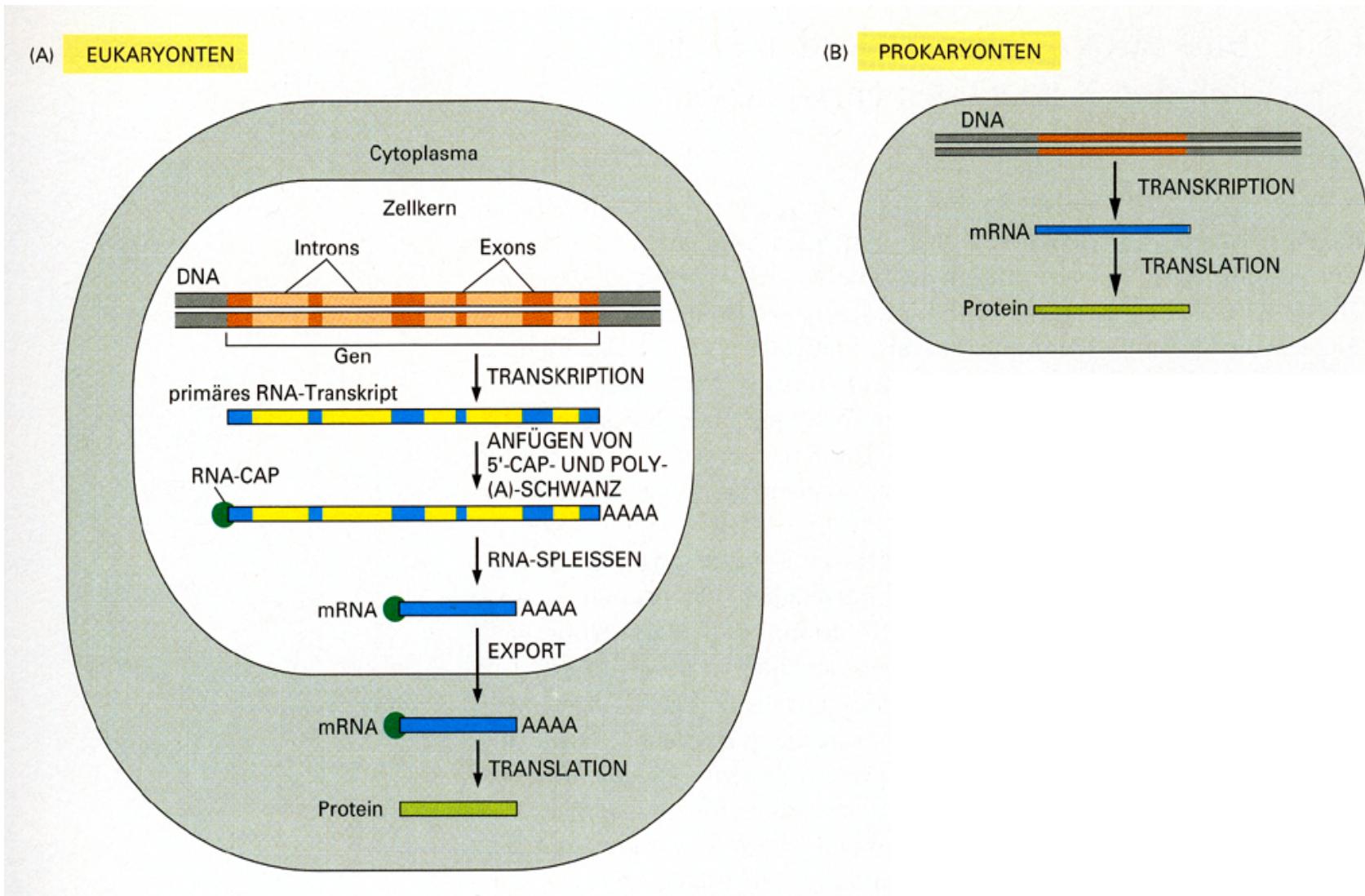


Über Proteine wird  
die mRNA zirkularisiert!

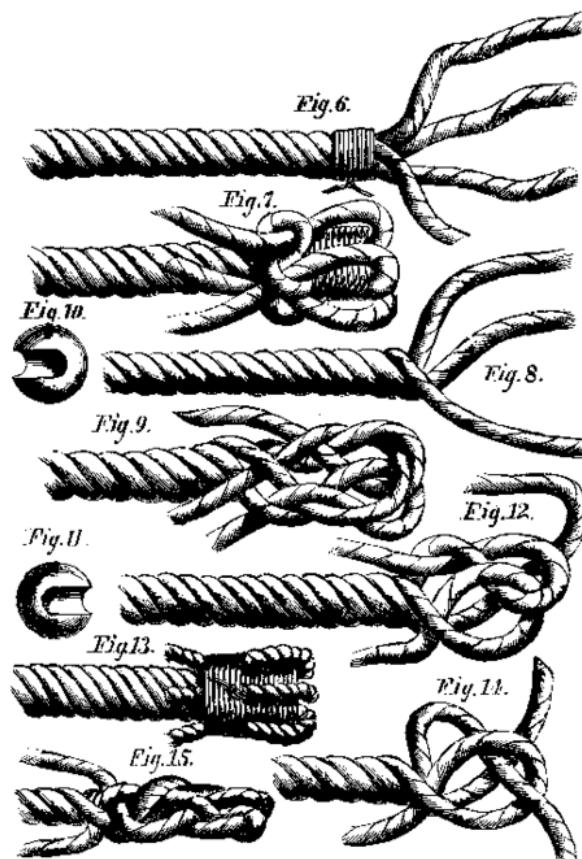
# Unterschiede? Gemeinsamkeiten?

|                 | <u>Prokaryoten</u> | <u>Eukaryoten</u> |
|-----------------|--------------------|-------------------|
| Operons         | ●                  | ○                 |
| Polycistr. mRNA | ●                  | ○                 |
| Introns/ Exons  | ○                  | ●                 |
| mRNA-cap        | -                  | ●                 |
| mRNA-polyA      | ○                  | ●                 |
| RNA-Polymerasen | 1                  | 3 (5)             |

# Überblick Genexpression

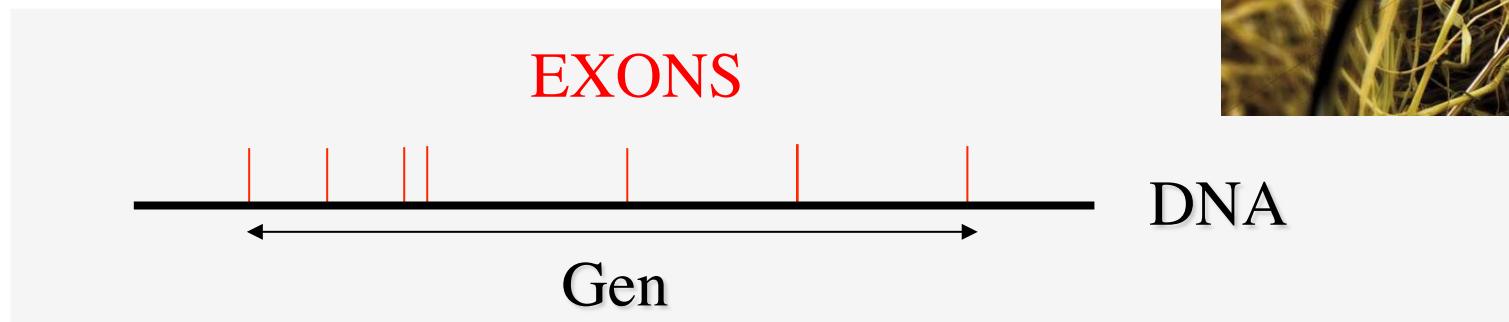


# Spezialität bei Eukaryoten: RNA Splicing & RNA Editing



# Gene von Eukaryoten bestehen aus Exons und Introns

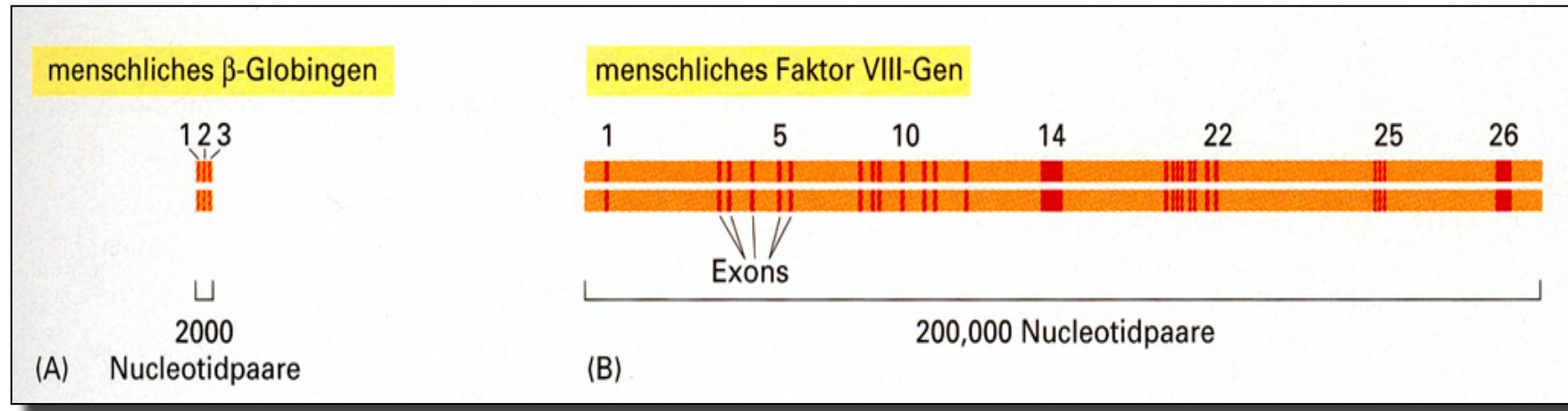
Ein Gen im Human-Genom...



Funktionelle Teile von proteinkodierenden Genen sind als **Exons** verteilt !

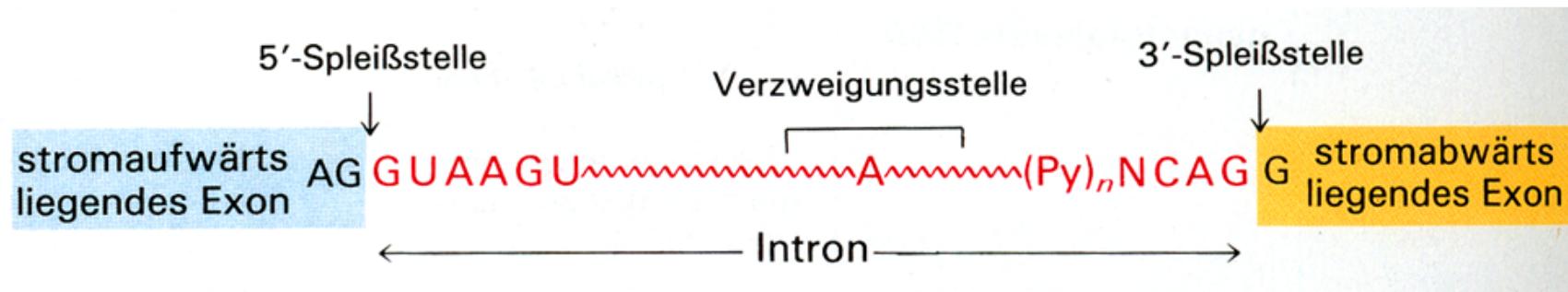
(Ø Länge 145 Bp; Ø 7 Exons pro Gen auf etwa 8 kb)

# Intron-Exon-Struktur



- Duchenne-Muskeldystrophie-Gen:
  - 78 Exons
  - verteilt über 2 400 kb
  - 99% des Genbereichs besteht aus Introns
  - Transkriptionsdauer ca. 20 Std. (**> Genregulation**)
- Introns haben 95%-Längenanteil an menschlichen Genen

# Erkennung von Spleißstellen



**Tabelle 29.3: Basensequenzen von Spleißpunkten bei Transkripten mit Introns**

| Genregion                            | Exon      | Intron          | Exon                  |
|--------------------------------------|-----------|-----------------|-----------------------|
| Ovalbumin, Intron 2                  | U A A G G | U G A G C ~~~~~ | U U A C A G G U U G   |
| Ovalbumin, Intron 3                  | U C A G G | U A C A G ~~~~~ | A U U C A G G U C U G |
| $\beta$ -Globin, Intron 1            | G C A G G | U U G G U ~~~~~ | C C U U A G G C U G   |
| $\beta$ -Globin, Intron 2            | C A G G G | U G A G U ~~~~~ | C C A C A G G U C U C |
| Immunglobulin $\lambda_1$ , Intron 1 | U C A G G | U C A G C ~~~~~ | U U G C A G G G G C   |
| SV40-Virus, frühes T-Antigen         | U A A G G | U A A A U ~~~~~ | U U U U A G A U U C   |

**„GU-AG-Regel“**

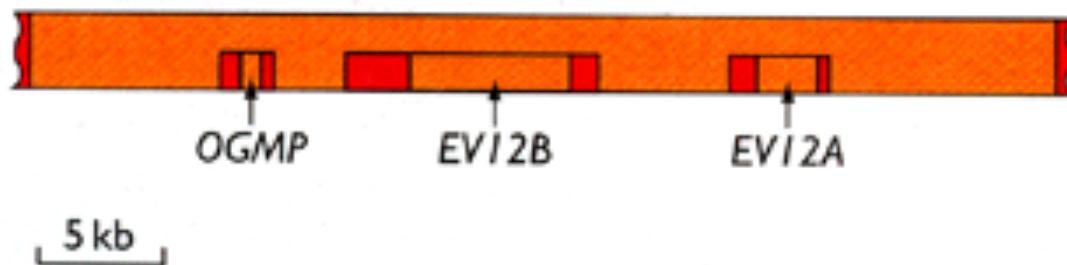
# „Intronphasen“

|         |     |        |
|---------|-----|--------|
| Phase 0 | AAG | CCA    |
|         | Lys | Pro    |
| Phase 1 | A   | AGCCA  |
|         | L   | ys Pro |
| Phase 2 | AA  | GCCA   |
|         | Ly  | s Pro  |

Introns können proteinkodierende Gene an allen Kodon-Positionen unterbrechen...

- **Genes-within-genes** are relatively common features of nuclear genomes, one gene being contained within an intron of a second gene. An example in the human genome is the neurofibromatosis type I gene, which has three short genes (called *OGMP*, *EVI2A* and *EVI2B*) within one of its introns. Each of these internal genes is also split into exons and introns.

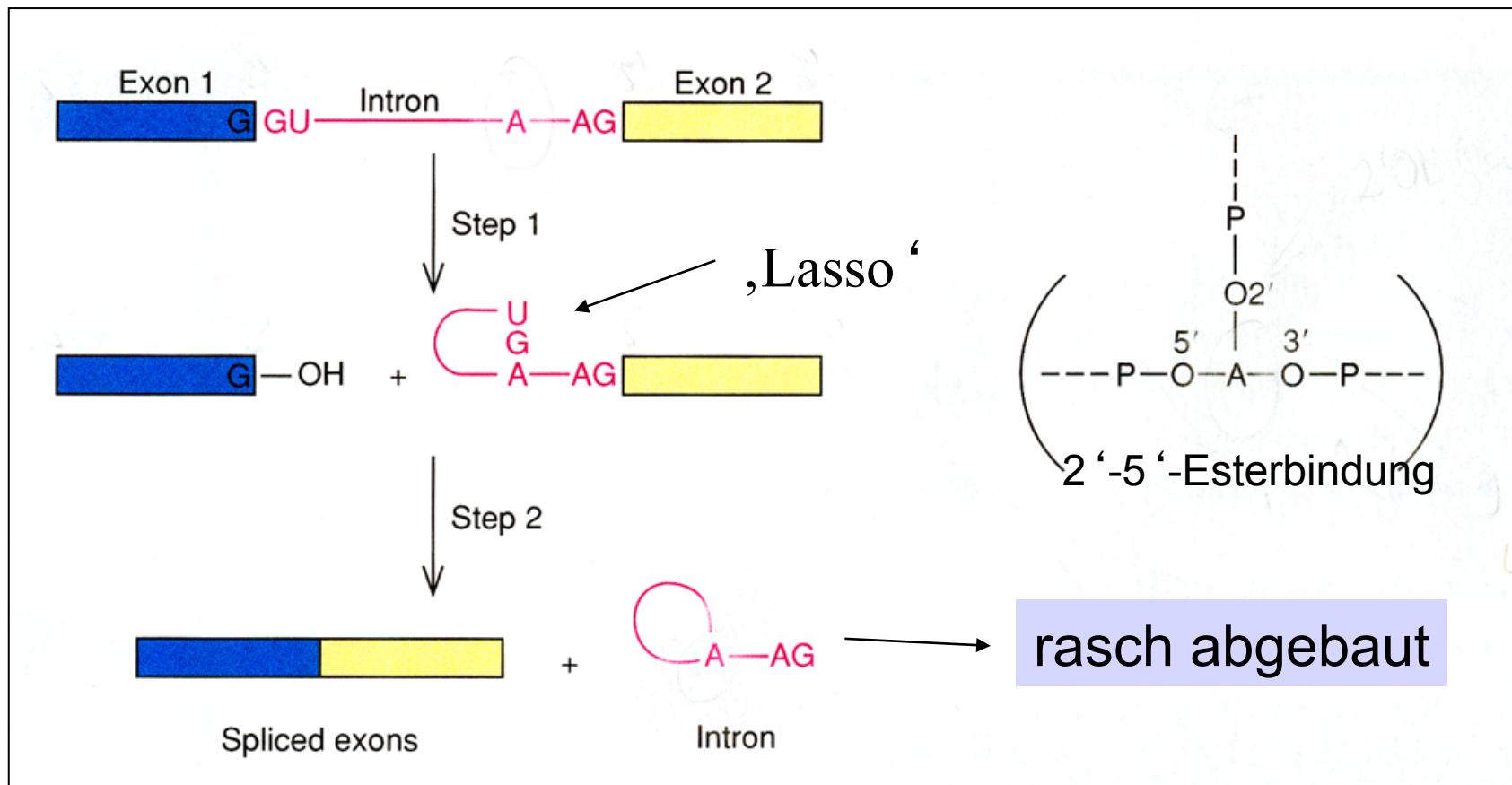
Neurofibromatosis type I gene  
Intron 27



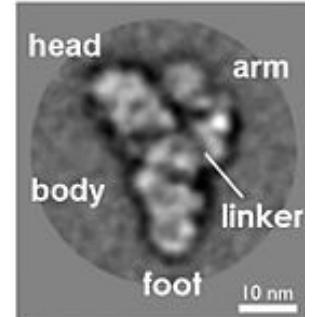
Recently, it has been discovered that many snoRNAs, which are involved in chemical modification of rRNAs (Section 9.4.1), are specified by genes within introns.

In Eukaryoten:  
„Nested Genes“

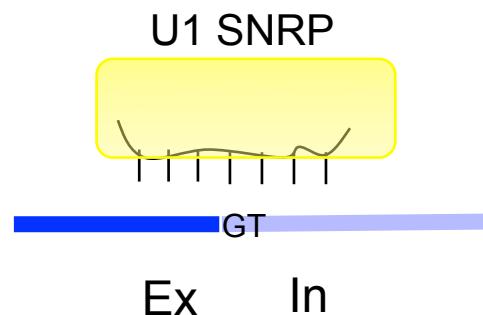
# Prinzip der Spleißereaktion



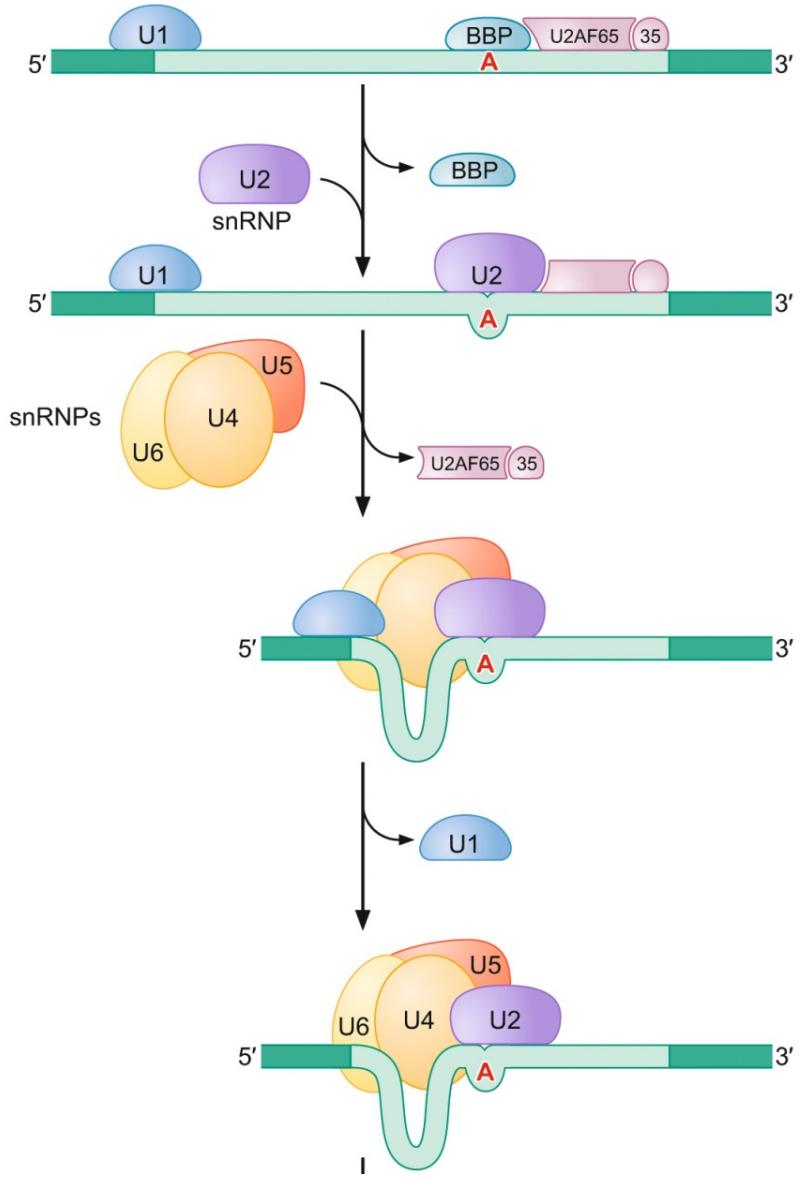
# Das Spleißosom...



- besteht aus 150 Proteinen und 5 'small nuclear ribo-nukleoproteins' (snRNPs)
- RNA-Anteile der 'Snurps' gehen **Basenpaarung mit den Exon/Introngrenzen ein**

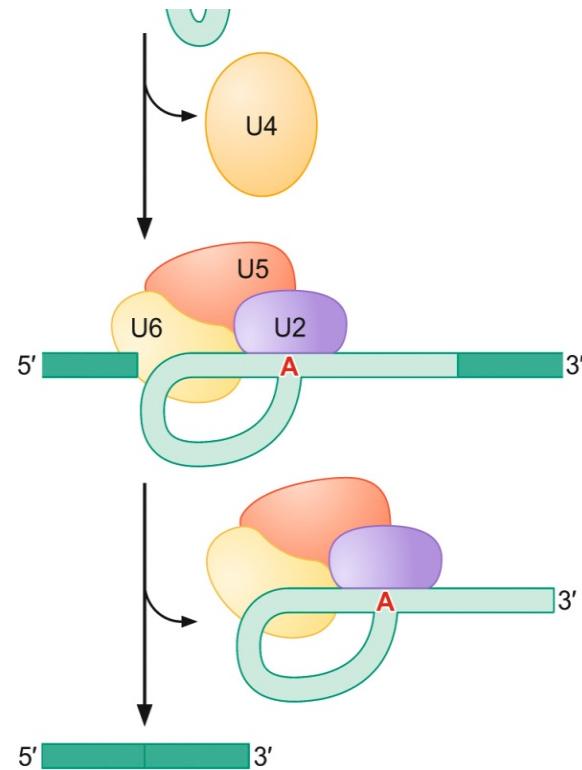


# Spleißosom: Abfolge der Ereignisse



**Abbildung 13.8: Schritte der spleißosom-vermittelten Spleißreaktion.** Die Zusammenlagerung und die Wirkungsweise des Spleißosoms sind dargestellt. Die Details der einzelnen Schritte finden Sie im Haupttext erläutert. Teile der Spleißmaschinerie treten bei jedem der Schritte hinzu oder verlassen die Szene. Diese Veränderungen bewirken strukturelle Umlagerungen des Spleißapparates, die für den Fortgang des Spleißprozesses notwendig sind. Die Bezeichnungen der einzelnen Komplexe sind rechts angegeben. Es gibt Hinweise darauf, dass einige der hier dargestellten Komponenten nicht genau in dem in dieser Abbildung angegebenen Schritt hinzutreten oder die Maschinerie verlassen. Sie können aber weiterhin am Ort vorhanden sein und zum Beispiel lediglich ihre Assoziation mit dem Komplex abschwächen, statt vollständig abzudissoziieren. Es herrscht gegenwärtig noch Unklarheit über die genaue Reihenfolge einiger der hier schematisch dargestellten Veränderungen. Dies gilt insbesondere für die beiden Schritte, die Änderungen der U6-Paarungen beinhalten: wenn diese Komponente an der 5'-Spleißstelle von U1 übernimmt, im Vergleich zur Übernahme der U2-Bindung von U4. Ungeachtet dieser verbleibenden Unsicherheiten, die durch weitere Forschungen ausgeräumt werden können, sind die entscheidenden Anteile verschiedener Komponenten der Maschinerie in den verschiedenen Stadien des Spleißvorgangs und das allgemeine dynamische Wesen des Spleißosoms hier dargestellt.

# Spleißosom: Abfolge der Ereignisse



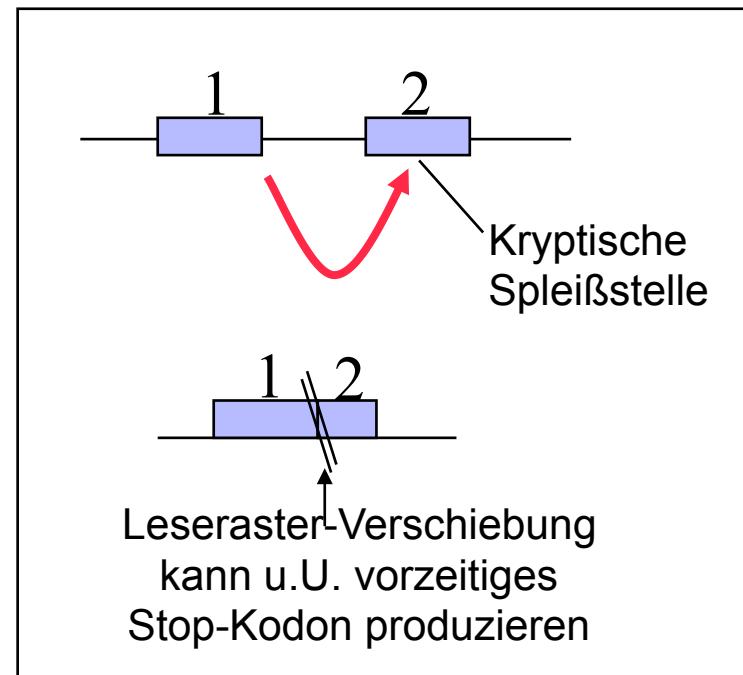
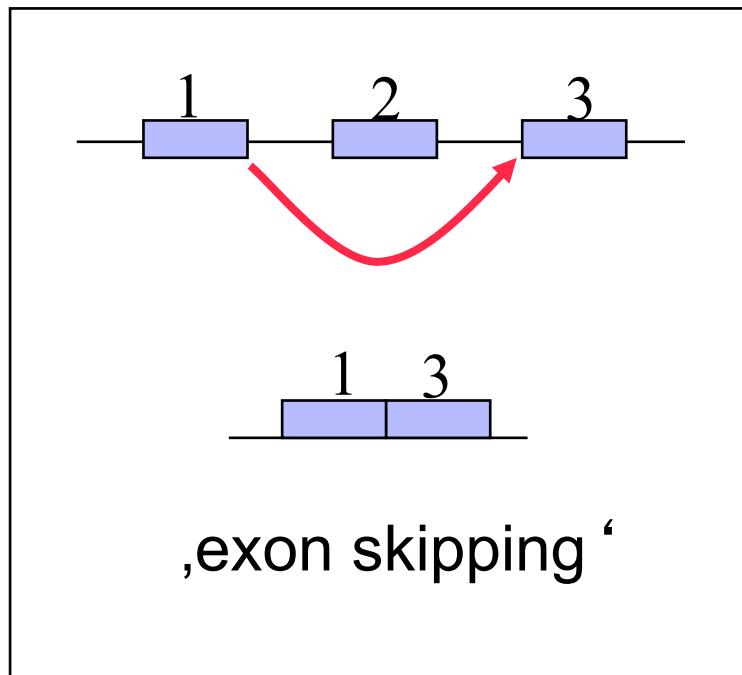
# ...eine ambitionierte Aufgabe!

- Introns sind lang (bis zu 800 kb) im Vergleich zu Exons (150 Bp im Durchschnitt)
- Sequenzen der Intron-Exon-Grenzen erscheinen kurz und daher nicht besonders spezifisch.



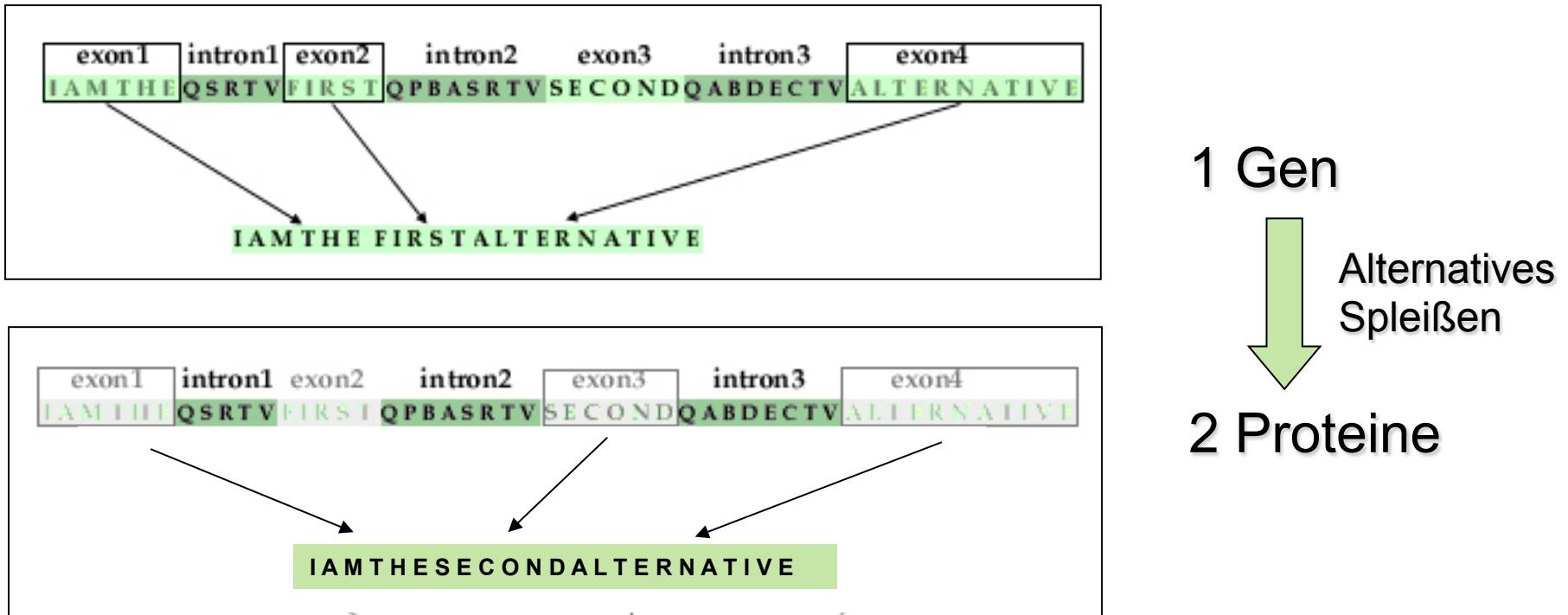
# Probleme & Möglichkeiten

- Distanz zwischen Spleiß-Stellen
- Auswahl der korrekten Spleißstelle



nicht nur Problem, sondern auch Möglichkeit:  
gezieltes „alternatives Spleißen“ schafft neue Proteine

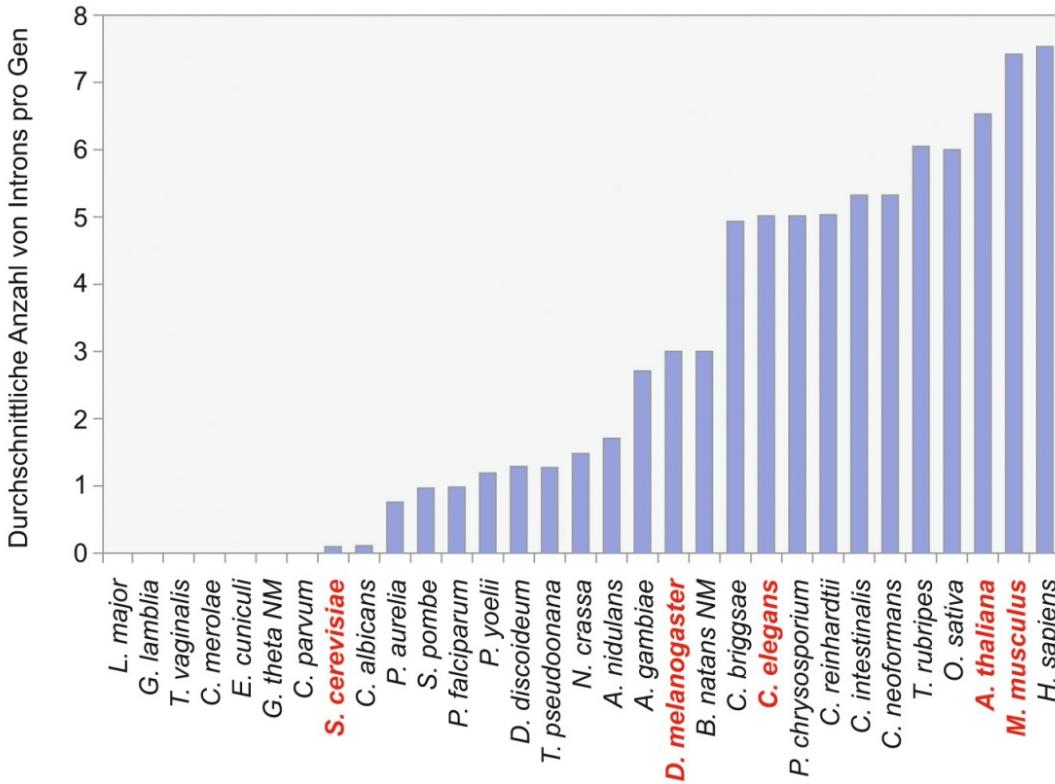
# Alternatives Spleißen



- > 75 % der menschlichen Gene werden alternativ gespleißt
  - oft > 5 Spleißformen/Gen
- (>Genregulation)



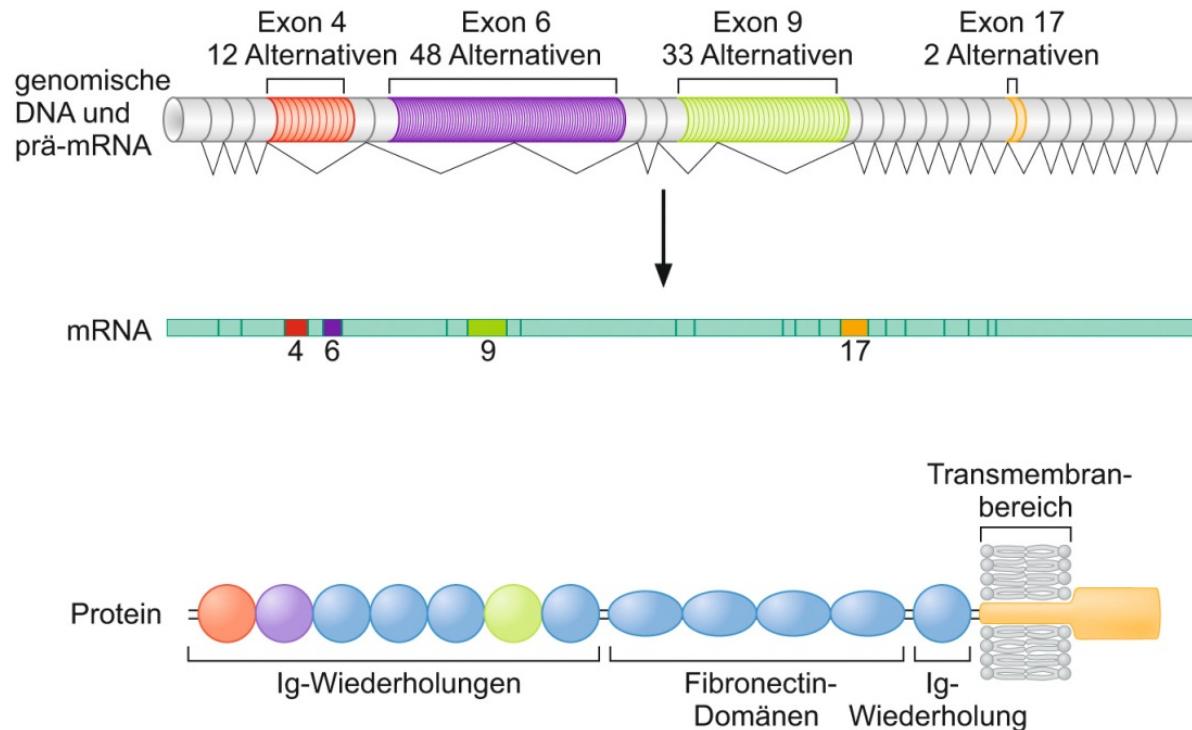
# Intron-Häufigkeit $\approx$ Komplexität



**Abbildung 13.2:** Anzahl der Introns pro Gen bei verschiedenen Eukaryontenarten. Die durchschnittliche Intronzahl pro Gen ist für eine Auswahl eukaryotischer Arten aufgezeigt. Die rot hervorgehobenen Artnamen zeigen gebräuchliche Modellorganismenarten an (siehe Kapitel 22): die Bäckerhefe (*Saccharomyces cerevisiae*), die Taufliege (*Drosophila melanogaster*), einen Fadenwurm (*Caenorhabditis elegans*), die Ackerschmalwand (*Arabidopsis thaliana*) sowie die Feldmaus (*Mus musculus*). Die übrigen Arten sind *Anopheles gambiae* (eine Steckmückenart), *Aspergillus nidulans* (ein Schimmelpilz), *Bigelovia natans* Nucleomorph, *Caenorhabditis briggsae*, *Candida albicans* (ein humanpathogener Hefepilz), *Chlamydomonas reinhardtii*, *Ciona intestinalis*, *Cryptococcus neoformans*, *Cryptosporidium parvum*, *Cyanidioschyzon merolae*, *Dictyostelium discoideum*, *Encephalitozoon cuniculi*, *Giardia lamblia*, *Guillarida theta* Nucleomorph, *Homo sapiens* (Mensch), *Leischmania major*, *Neurospora crassa* (eine Schimmelpilzart), *Oryza sativa* (Reis), *Paramecium aurelia* (Pantoffeltierchen), *Phaenochaete chrysosporium*, *Plasmodium falciparum*, *Plasmodium yoelii*, *Schizosaccharomyces pombe* (Spalthefe), *Takifugu rubripes* (ein Kugelfisch), *Thalassiosira pseudonana* und *Trichomonas vaginalis*.



# Weltrekord: *Drosophila DSCAM*



38 000 Splice-Möglichkeiten!!

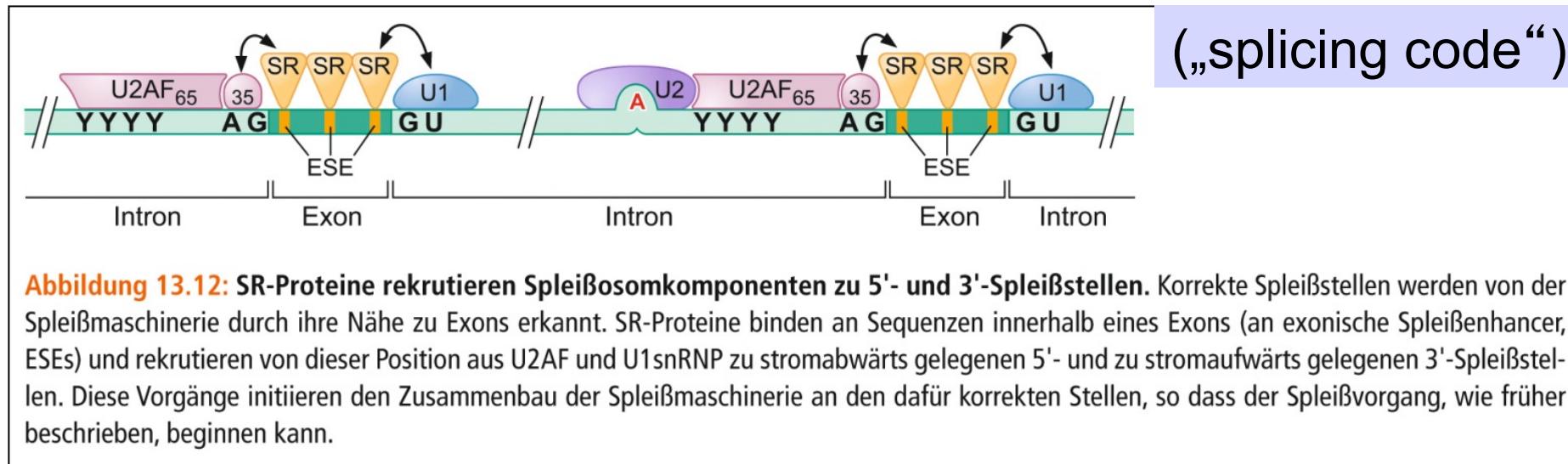
Zell-Zell-Interaktion zwischen Nervenzellen!

➤ Aufbau neuronaler Schaltkreise?

**Abbildung 13.19: Die multiplen Exons des *Drosophila*-Gens *Dscam*.** Das Gen *Dscam* (oben) ist 61,2 kb lang. Nach Transkription und Spleißen gehen aus ihm eine oder mehrere Varianten einer 7,8 kb langen, 24 Exons enthaltenden mRNA hervor (die Schemazeichnung zeigt eine allgemeine Form dieser mRNAs). Wie hier dargelegt, gibt es mehrere, sich wechselseitig ausschließende Alternativen für die Exons 4, 6, 9 und 17. Jede maturierte mRNA enthält eine von zwölf möglichen Alternativen des Exons 4 (rot), eine von 48 Alternativen des Exons 6 (lila), eine von 33 des Exons 9 (blau) und eine von zwei des Exons 17 (gelb). Die Exons 4, 6 und 9 codieren Teile der drei Ig-Domänen (Immunglobulindomänen), die in der Abbildung unten in den entsprechenden Farben dargestellt sind. Das Exon 17 codiert die Transmembrandomäne des Proteins. Falls alle zur Verfügung stehenden Kombinationen der existierenden Exons zum Zuge kommen, können aus dem *Dscam*-Gen 38.016 unterschiedliche mRNAs und Proteine hervorgehen.  
Watson



# SR-Proteine bewirken korrektes Spleißen durch Erkennen von Exons

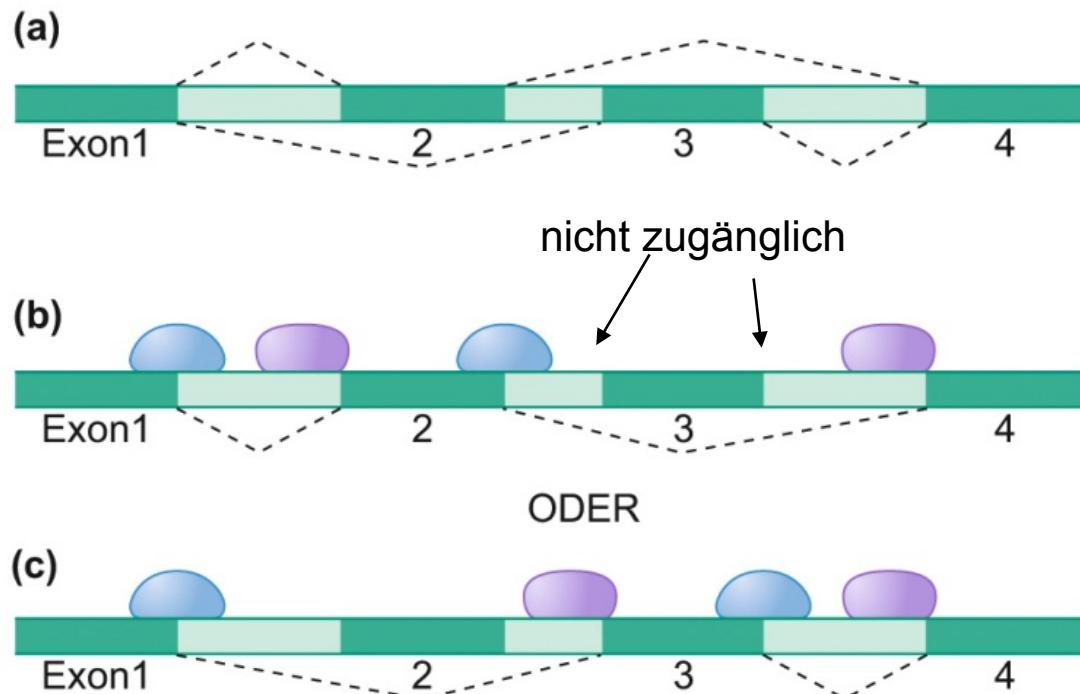


- SR = Serin und Arginin-reiche Proteine
- binden an purinreiche „exonic splicing enhancer“ (ESE)
- interagieren mit SNURPS und anderen nicht-Snurp-Proteinen

Watson



# Korrekte Spleißen durch sterische Hinderung



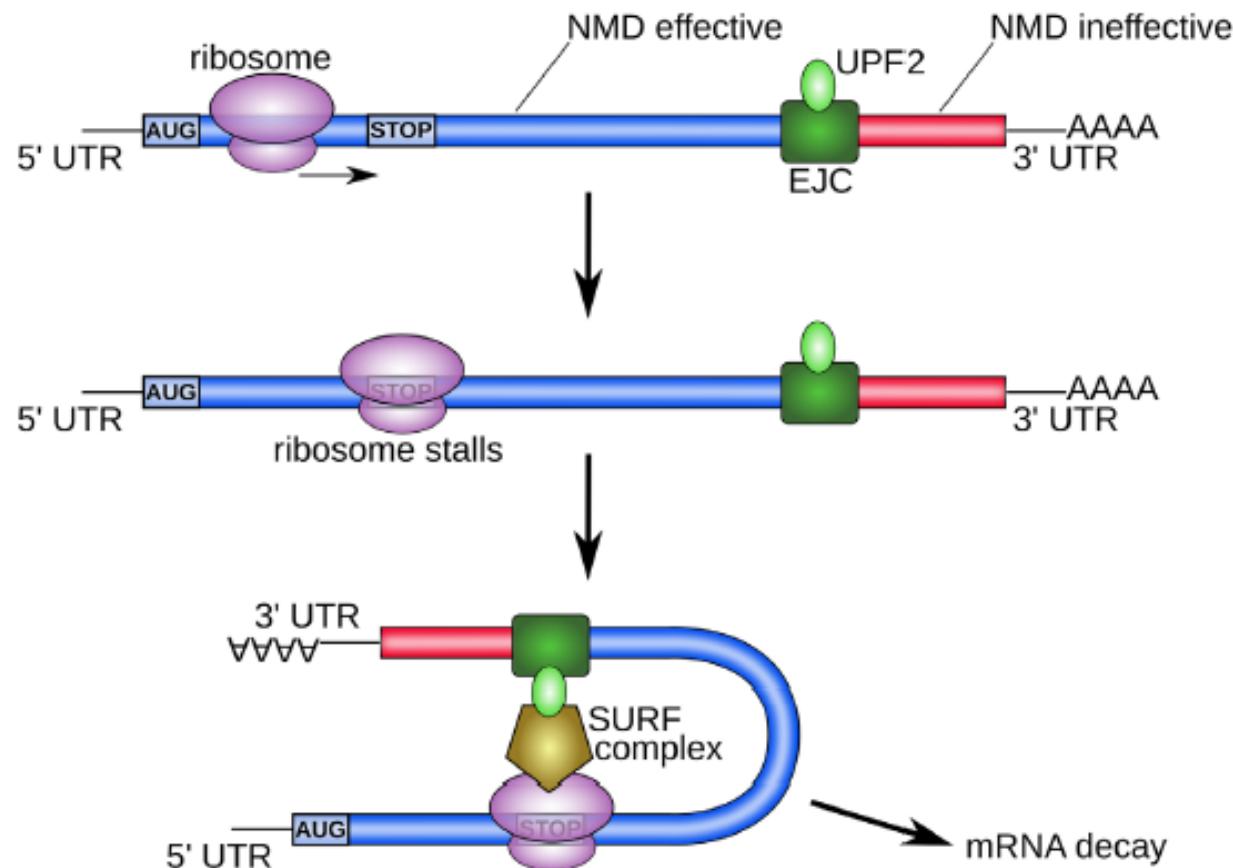
Watson

**Abbildung 13.17:** Sich wechselseitig ausschließende Spleißvorgänge durch sterische Hinderung. (a) Diese Ansicht zeigt die alternativen Spleißmöglichkeiten. (b) Die Bindung des U1-snRNPs an die 5'-Spleißstelle des zweiten Introns schließt eine Bindung von U2-snRNP an den Verzweigungspunkt desselben Introns aus (für die gleichzeitige Bindung beider Makromolekülkomplexe ist nicht genügend Platz vorhanden); die Bindung von U2 an das nächstgelegene Intron führt zum Ausschluss von Exon 3. (c) Hier verhindert die Bindung von U2-snRNP an den Verzweigungspunkt des zweiten Introns die Bindung von U1 an die 5'-Spleißstelle desselben Introns. In diesem Fall führt die Bindung von U1 an die 5'-Spleißstelle des ersten Introns zum Ausschluss von Exon 2.

Zur Sicherheit, falls Spleißen schief geht...



# Nonsense-mediated RNA decay

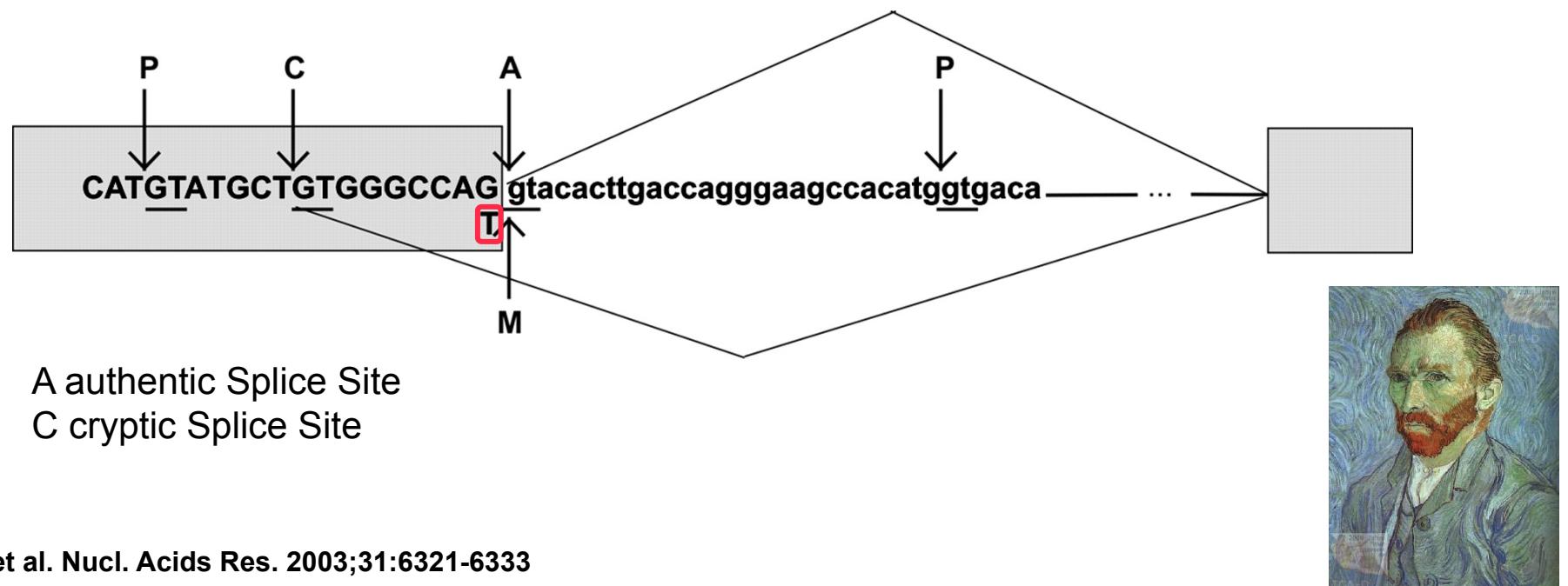


**Figure 1. EJC-dependent nonsense-mediated decay (NMD) [4].** After splicing, exonic-junction complexes (EJCs) remain 20–24 nucleotides upstream of every exon junction. These EJCs are then bound by UPF2, one of the core proteins of NMD. When the first ribosome translates the mRNA, it displaces the EJCs. However, if the ribosome encounters a premature stop codon and stalls, it forms a complex with a downstream EJC, mediated by UPF2 and another complex called SURF. (The SURF complex is named after its constituent proteins [5].) This complex then initiates mRNA decay. Because EJC-dependent NMD requires a downstream EJC, it is only effective in the coding regions upstream of the last EJC (indicated in blue). It cannot detect any premature stop codons downstream of the last EJC (indicated in red). Note that an alternative, less potent mode of NMD takes place in the absence of the EJC [6].

doi:10.1371/journal.pgen.1002296.g001

# Splicing & Erkrankung

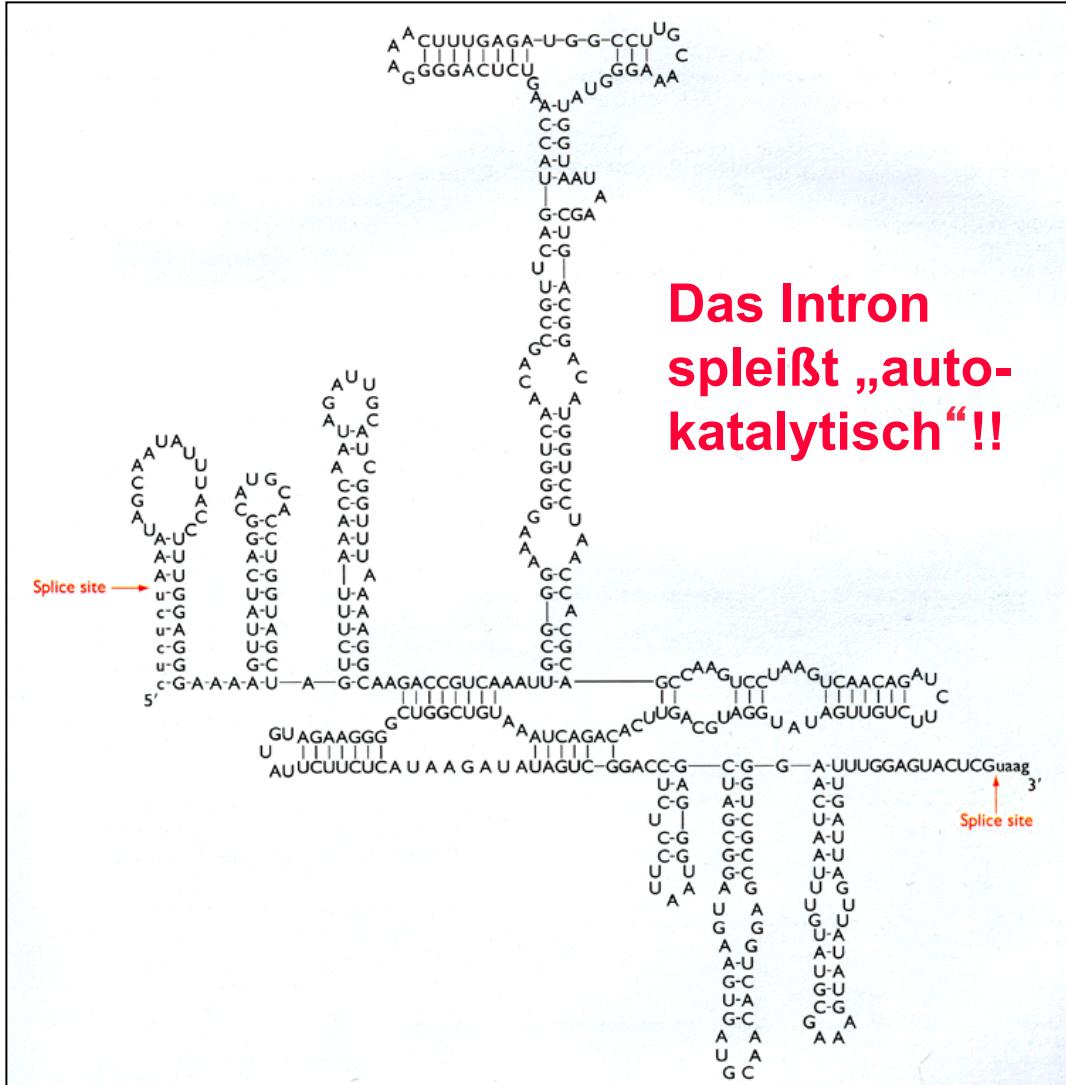
- ca. 50 % aller Mutationen in Exons verändern das Spleißen!
- Auch sog. „stille Mutationen“ können so drastische Auswirkungen haben.
- Bsp. porphobilinogen desaminase-Gen (PBGD) > Häm-Mangel



„...die Medizin ist der beste Lehrmeister der (Molekular-) Biologie.“

Lehrbuch: Watson et al. Molekularbiologie

# Aufklärung des Spleiß-Mechanismus



Struktur eines  
„Gruppe I“-Introns  
aus den rRNA-  
Genen des Einzellers  
Tetrahymena.



The Nobel Prize in Chemistry 1989  
Sidney Altman, Thomas R. Cech

## The Nobel Prize in Chemistry 1989

Nobel Prize Award Ceremony

Sidney Altman

Thomas R. Cech



Sidney Altman



Thomas R. Cech

The Nobel Prize in Chemistry 1989 was awarded jointly to Sidney Altman and Thomas R. Cech "for their discovery of catalytic properties of RNA"

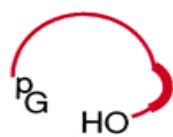
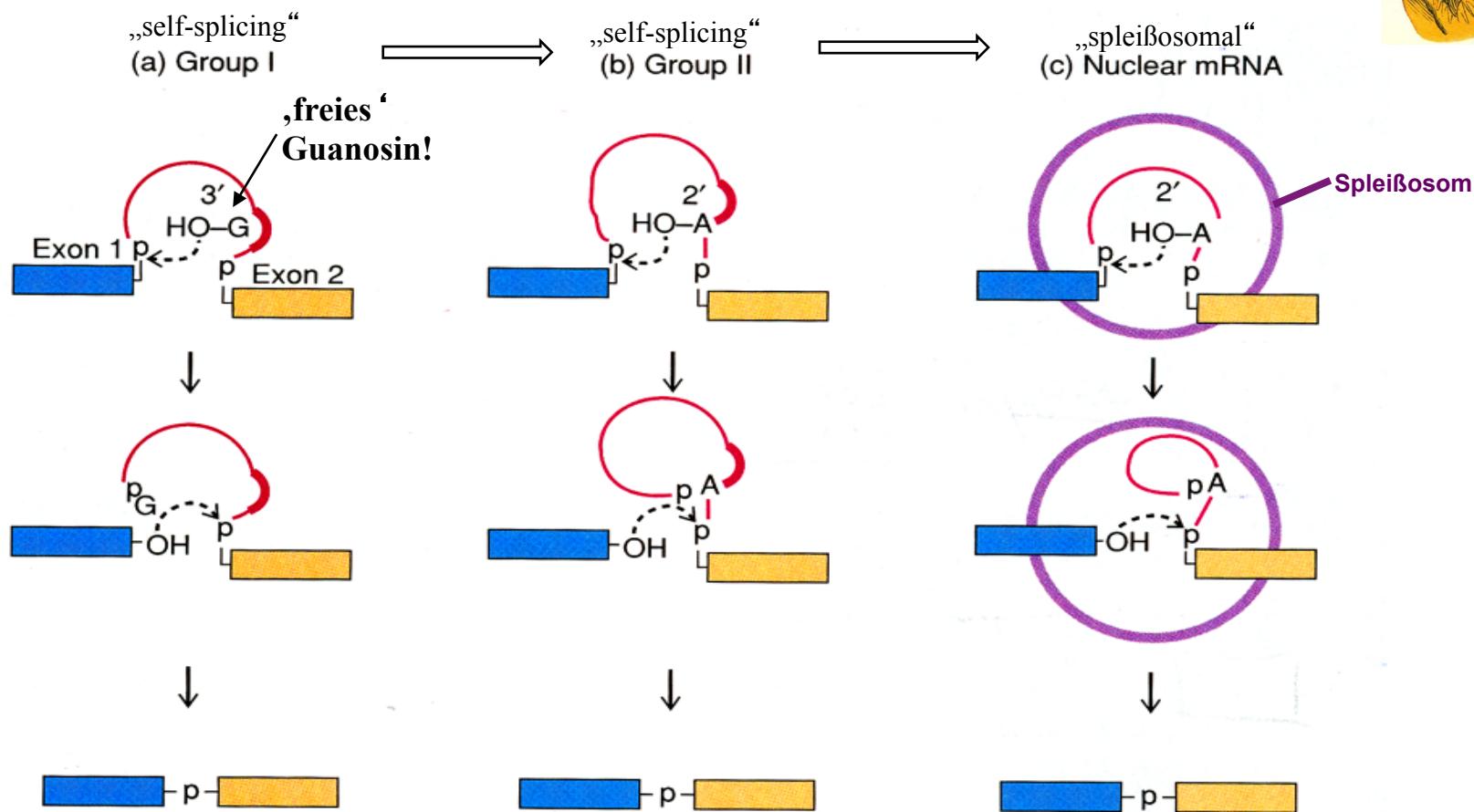
Spleißen ohne Spleißosom:

## „Selbstspleißende Introns“

- Gruppe I
  - prä-rRNA im Kern eukaryot. Einzeller
  - Organellen-mRNAs und -rRNAs
  - wenige bakt. RNAs (z. B. Phage T4!)
- Gruppe II
  - Organellen-RNAs in Pflanzen u. Pilzen
  - wenige bakt. RNAs
- Gruppe III
  - Organellen-RNAs
  - kürzer als Gruppe II, jedoch ähnlicher Mechanismus

- unterschieden nach typischer Faltungsstruktur sowie nach Spleiß-Mechanismus!

# Evolution der Spleiß-Mechanismen



# **Der „Kaiser“ zum Thema Evolution:**



„Sokrates, Platon und diese Leute haben sich schon vor 2000 Jahren Gedanken gemacht, da sind wir noch auf den Bäumen gehockt und haben uns vor den Wildschweinen gefürchtet.“

Seitdem haben sich nur ganz wenige weiterentwickelt.“

# Evolution der Exon-Intronstruktur eukaryotischer Gene

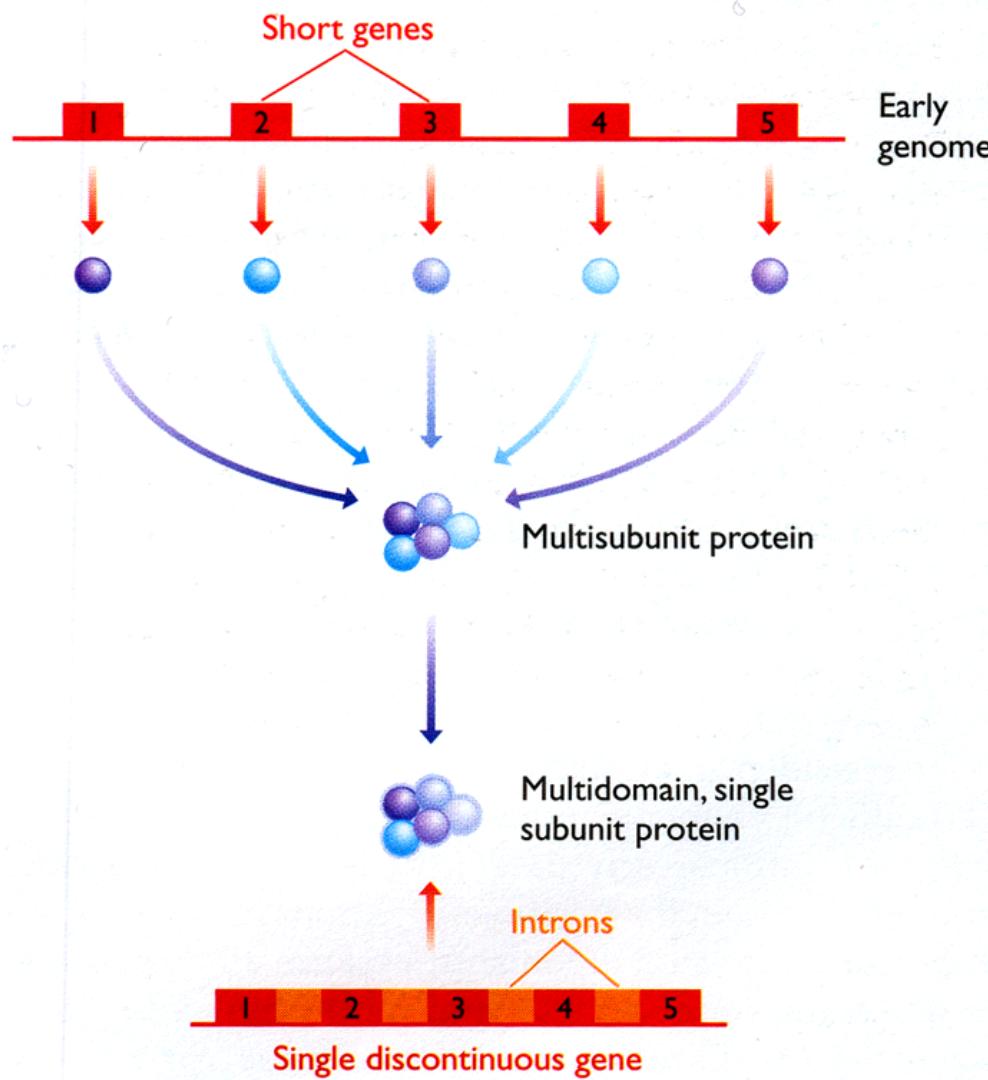
- „**Introns-early**“ (= „**exon theory of genes**“)
  - > Introns leiten sich aus Struktur des Ur-Genoms ab, Prokaryoten haben Introns weitgehend verloren

- „**Introns-late**“ („**insertional introns**“)
  - > Introns sind bewegliche Elemente und haben sich während der Eukaryotenevolution in zuvor intronlose Gene integriert

Woher kommt die Exon-Intron-Struktur evolutionär gesehen?

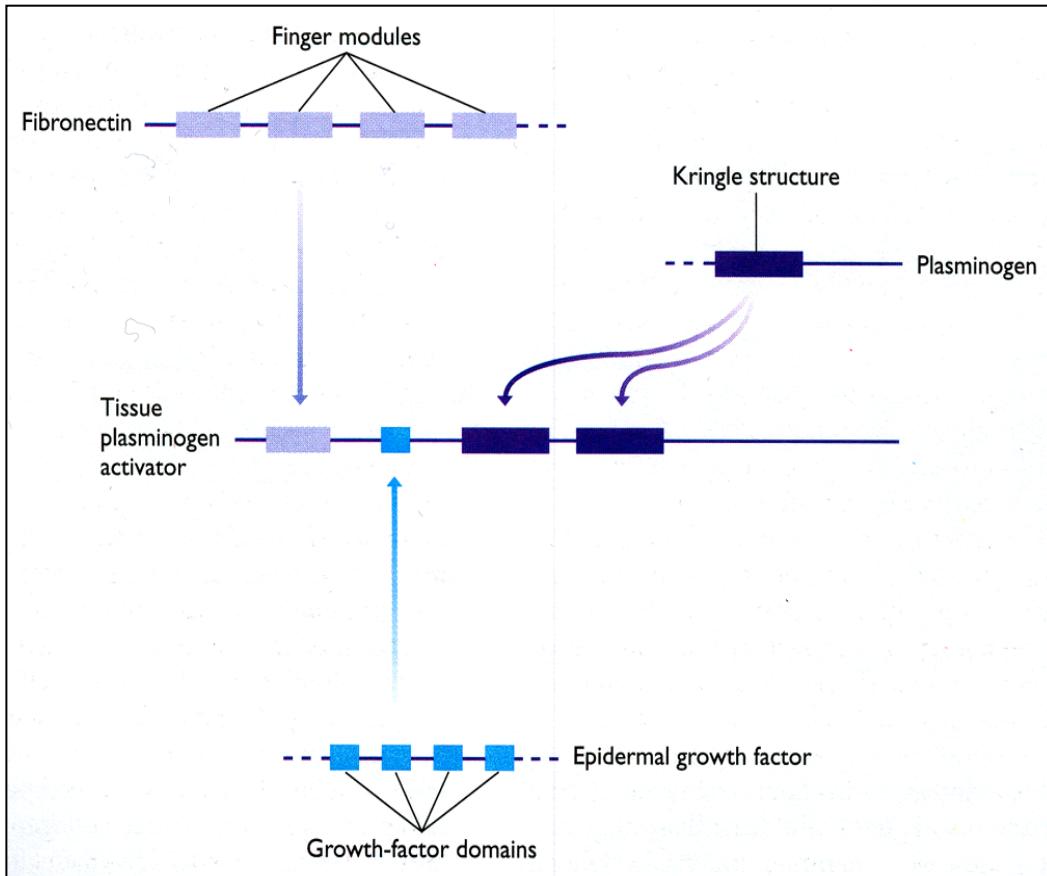
# „Exon theory of genes“

(Walter Gilbert)



- Exons = „alte Minigene“
- Introns = „alte Intergen-regionen“

# Vorteil der Exon-Intron-Struktur



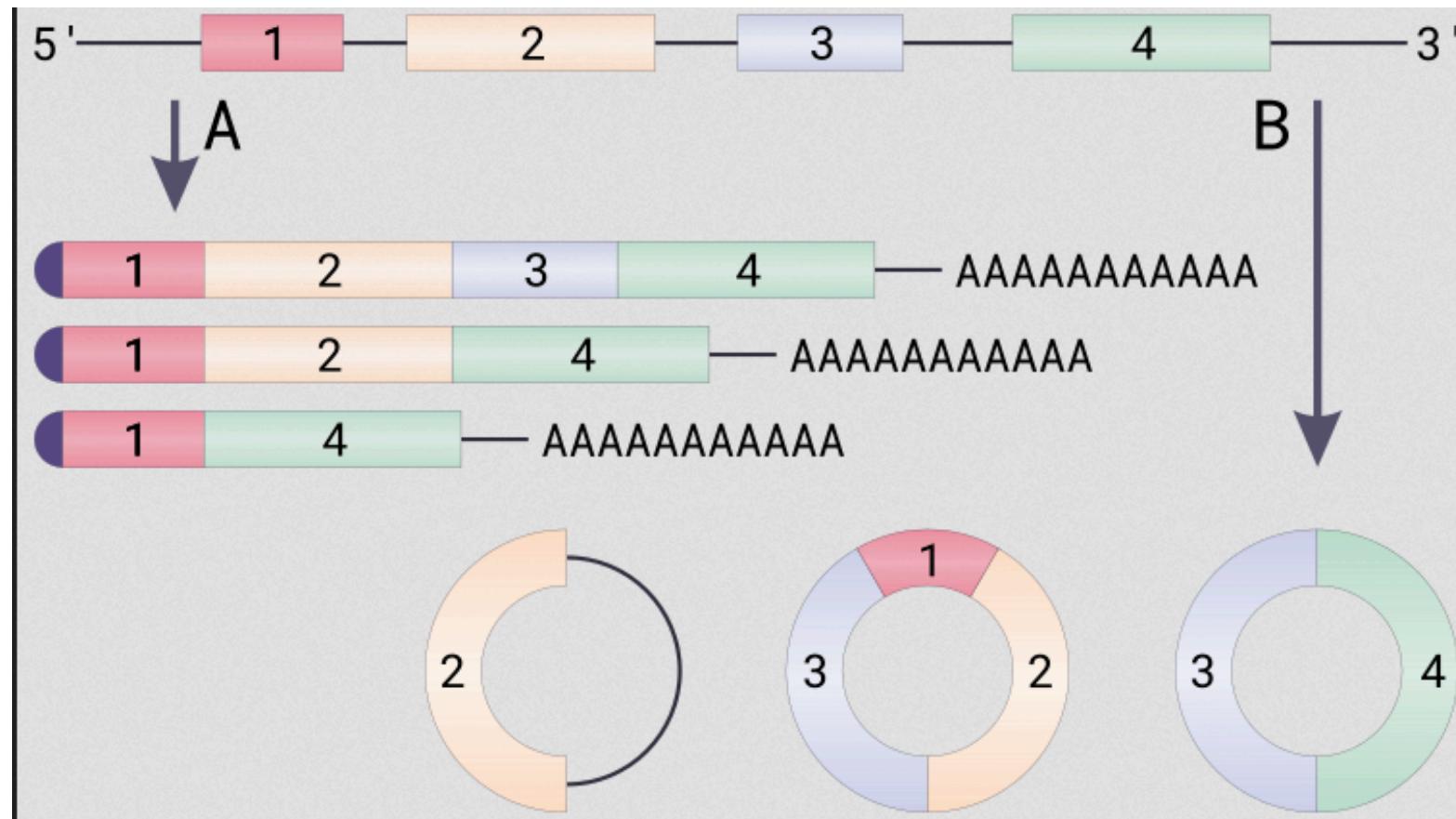
- Exons kodieren manchmal distinkte **Proteinmodule**
- Rekombination in Introns hat während der Evolution „**exon-shuffling**“ ermöglicht

Evolution des TPA-Gens durch Exon-Shuffling





# Circular RNAs (circRNA)



A normales Spleißen > lineare mRNA  
B ‚backsplicing‘ > circRNA

# Circular RNAs are a large class of animal RNAs with regulatory potency

Sebastian Memczak<sup>1\*</sup>, Marvin Jens<sup>1\*</sup>, Antigoni Elefsinioti<sup>1\*</sup>, Francesca Torti<sup>1\*</sup>, Janna Krueger<sup>2</sup>, Agnieszka Rybak<sup>1</sup>, Luisa Maier<sup>1</sup>, Sebastian D. Mackowiak<sup>1</sup>, Lea H. Gregersen<sup>3</sup>, Mathias Munschauer<sup>3</sup>, Alexander Loewer<sup>4</sup>, Ulrike Ziebold<sup>1</sup>, Markus Landthaler<sup>3</sup>, Christine Kocks<sup>1</sup>, Ferdinand le Noble<sup>2</sup> & Nikolaus Rajewsky<sup>1</sup>

Circular RNAs (circRNAs) in animals are an enigmatic class of RNA with unknown function. To explore circRNAs systematically, we sequenced and computationally analysed human, mouse and nematode RNA. We detected thousands of well-expressed, stable circRNAs, often showing tissue/developmental-stage-specific expression. Sequence analysis indicated important regulatory functions for circRNAs. We found that a human circRNA, antisense to the cerebellar degeneration-related protein 1 transcript (CDR1as), is densely bound by microRNA (miRNA) effector complexes and harbours 63 conserved binding sites for the ancient miRNA miR-7. Further analyses indicated that CDR1as functions to bind miR-7 in neuronal tissues. Human CDR1as expression in zebrafish impaired midbrain development, similar to knocking down miR-7, suggesting that CDR1as is a miRNA antagonist with a miRNA-binding capacity ten times higher than any other known transcript. Together, our data provide evidence that circRNAs form a large class of post-transcriptional regulators. Numerous circRNAs form by head-to-tail splicing of exons, suggesting previously unrecognized regulatory potential of coding sequences.

## CircRNA: functions and properties of a novel potential biomarker for cancer



CrossMark

Shujuan Meng<sup>1,2†</sup>, Hecheng Zhou<sup>1,2†</sup>, Ziyang Feng<sup>1,2†</sup>, Zhao Xu<sup>1,2</sup>, Ying Tang<sup>1,2</sup>, Peiyao Li<sup>2</sup> and Minghua Wu<sup>2\*</sup>

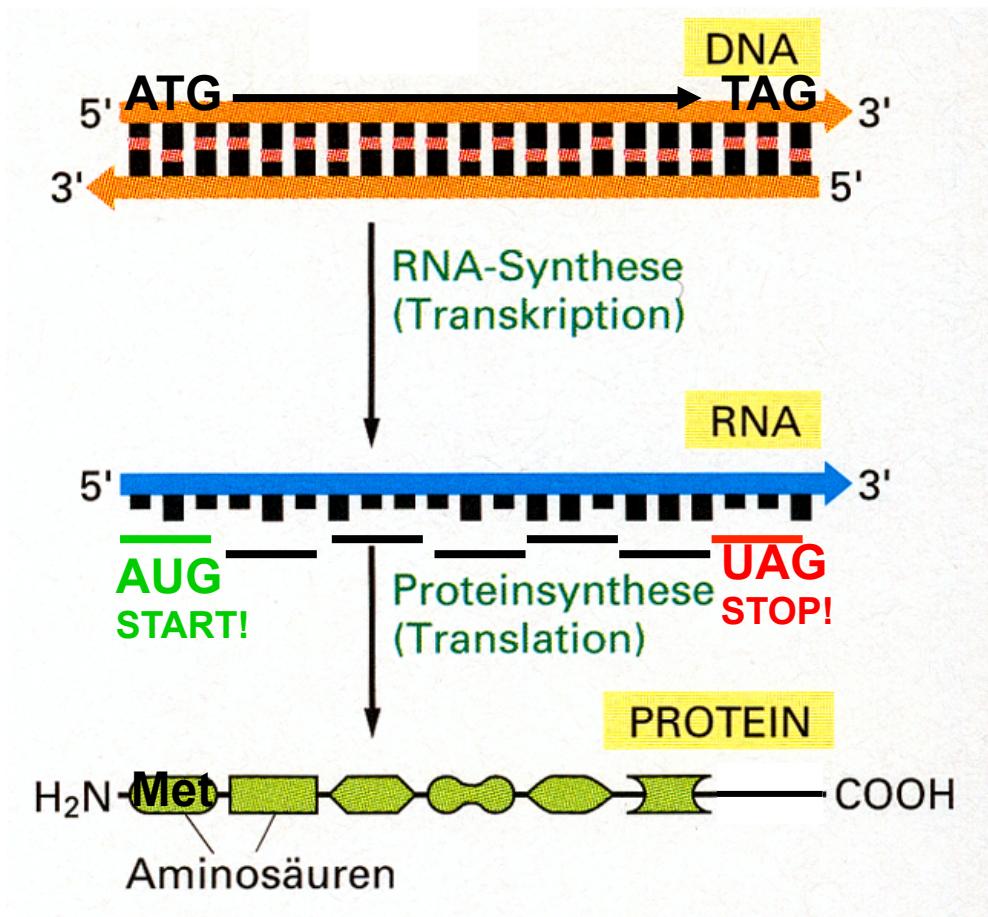
### Abstract

Circular RNAs, a novel class of endogenous noncoding RNAs, are characterized by their covalently closed loop structures without a 5' cap or a 3' Poly A tail. Although the mechanisms of circular RNAs' generation and function are not fully clear, recent research has shown that circular RNAs may function as potential molecular markers for disease diagnosis and treatment and play an important role in the initiation and progression of human diseases, especially in tumours. This review summarizes some information about categories, biogenesis, functions at the molecular level, properties of circular RNAs and the possibility of circular RNAs as biomarkers in cancers.

**Keywords:** circRNA, Non-coding RNA, microRNA sponge, Tumour, Biomarker



# Der Normalfall...



**mRNA bildet  
die DNA-Sequenz  
exakt ab!**

# RNA editing !

- RNA-Sequenz wird **NACH** der Transkription verändert !!!!
- **Konsequenz:** Aminosäuresequenz des Proteins lässt sich nicht mehr aus DNA-Sequenz ableiten
- in höheren Organismen v.a. zwei Mechanismen:
  1. **Basen-Austausch** durch ortsspezifische enzymatische Desaminierung von A oder C
  2. gezielte **Insertion** oder **Deletion** von U (gesteuert durch „guide RNAs“)

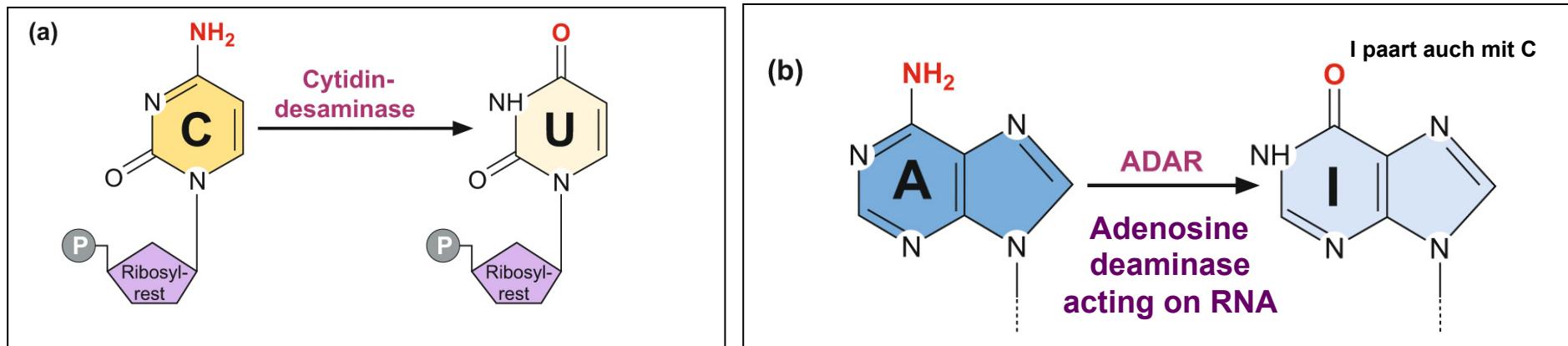
**>Genregulation; Proteinvielfalt !**



# RNA editing !

- in mRNA, tRNA, rRNA, miRNA ...
- in Eukaryoten, Viren, Archaen, Eubakterien
- in Zellkern und Cytoplasma
- in Kerngenom, Mitochondrien und Plastiden
- in Vertebraten eher noch selten, in anderen Fällen ist Mehrheit der mRNA-Basen durch ***pan-editing*** entstanden!!

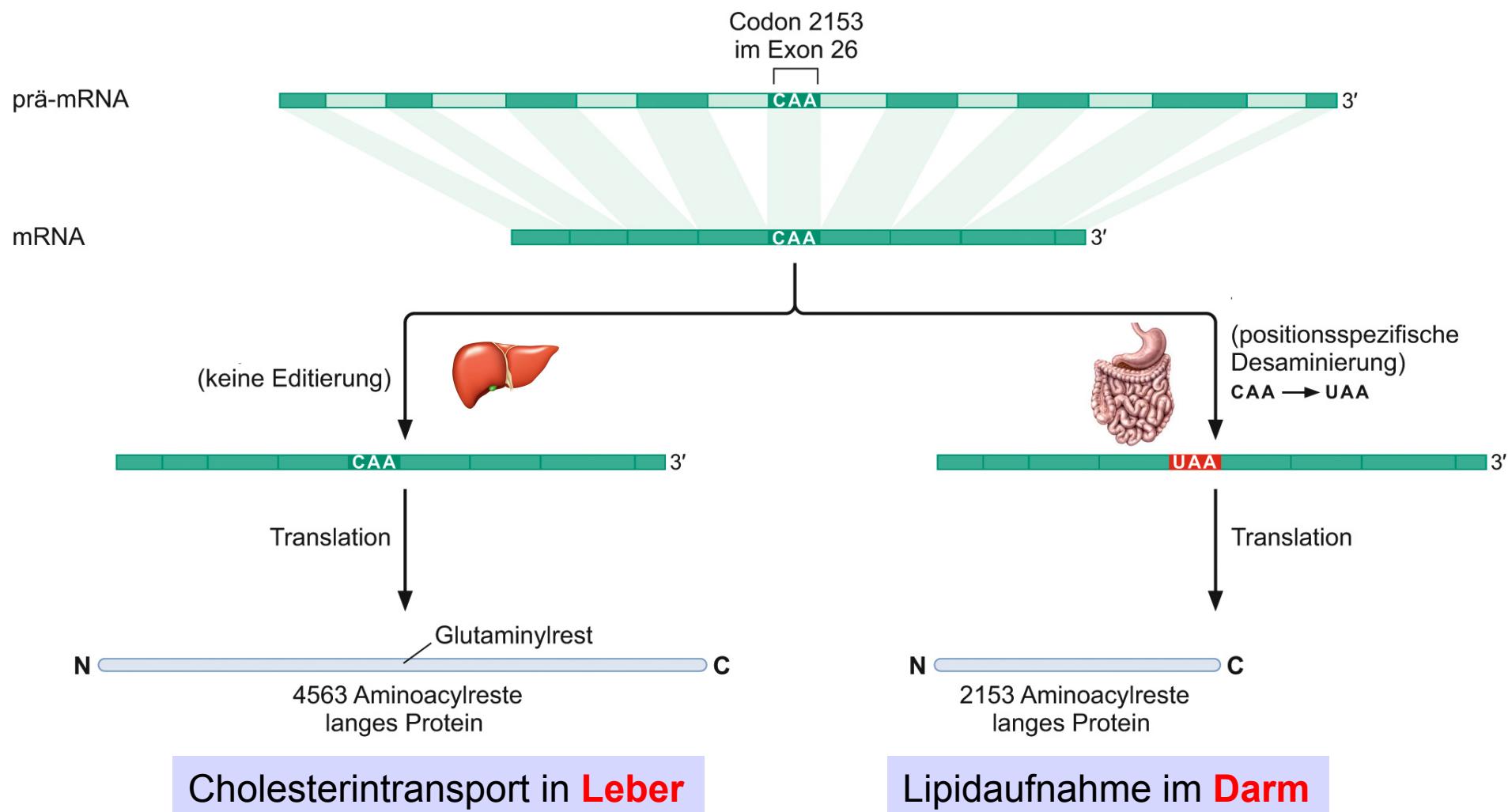
# RNA editing durch Desaminierung



- hochspezifisch reguliert
- meist nur in bestimmten Geweben oder Zelltypen
- in Säugern sind Tausende wichtiger Gene betroffen (**A>I**)
- in pflanzlichen Mitochondrien in jedem Gen (**C>U**)

# RNA editing: Desaminierung C>U

- wenige Säuger-Beispiele. Hier: Apolipoprotein B (Lipid-Transport)





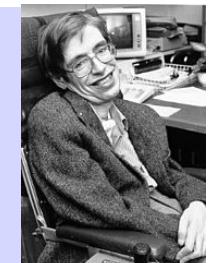
# RNA editing: Desaminierung A > I

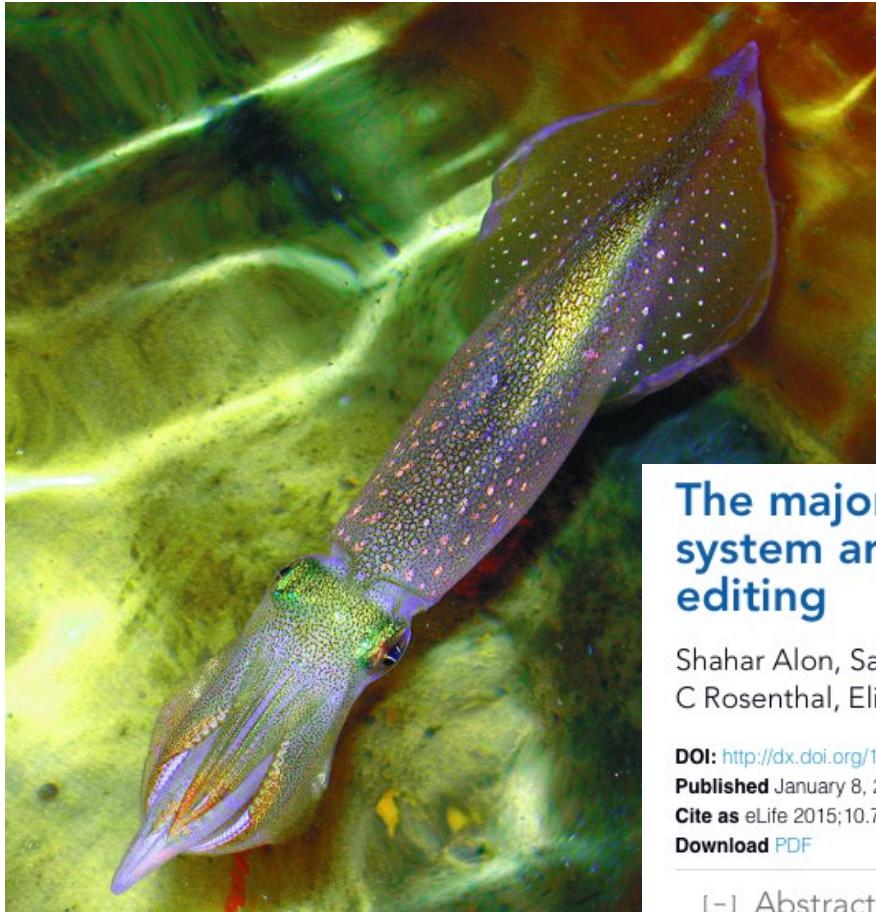
Table 1 | The known transcripts that undergo RNA editing and the functional consequence on the encoded protein

| Organism                  | Transcript  | Effects of editing   | Functional consequence   | References   |
|---------------------------|---|--|--|--|
| <b>Cytosine to uracil</b> |   |  |  |  |
| Mammals                   | ApoB<br>Nf1   | Q → STOP<br>R → STOP   | Stop codon, generation of ApoB48 isoform<br>(truncated protein)  | 2,3<br>32  |
| <b>Adenine to inosine</b> |   |  |  |  |
| Mammals                   | GluR-B<br>GluR-B<br>GluR-C<br>GluR-D<br>GluR-5<br>GluR-6<br>GluR-6<br>GluR-6<br>5-HT <sub>2c</sub> R<br>5-HT <sub>2c</sub> R<br>5-HT <sub>2c</sub> R<br>5-HT <sub>2c</sub> R<br>Adar2 | Q → R<br>R → G<br>R → G<br>R → G<br>Q → R<br>Q → R<br>I → V<br>Y → C<br>I → V, N → S, I → V<br>I → M<br>N → D<br>N → G<br>Intronic nt changed,<br>new splice acceptor site | Decreased Ca <sup>2+</sup> ion permeability<br>Increased rate of recovery receptor desensitization<br>Unknown<br>Unknown<br>Unknown<br>Increased Ca <sup>2+</sup> ion permeability<br>Modulated Ca <sup>2+</sup> ion permeability<br>Unknown<br>Reduced G-protein coupling<br>Unknown<br>Unknown<br>Unknown<br>Generation of isoform | 77,80,88<br>77,80,90<br>77,80<br>77,80<br>77,80<br>77,80<br>77,80<br>77,80<br>77,81<br>46,86 |

GluR Glutamat-Rezeptor  
5-HTR Serotonin-Rezeptor

Epilepsie  
Schizophrenie  
Depression  
Amyotrophe Lateralsklerose





## The majority of transcripts in the squid nervous system are extensively recoded by A-to-I RNA editing

Shahar Alon, Sandra C Garrett, Erez Y Levanon, Sara Olson, Brenton R Graveley, Joshua J C Rosenthal, Eli Eisenberg 

**DOI:** <http://dx.doi.org/10.7554/eLife.05198>

**Published** January 8, 2015

**Cite as** eLife 2015;10.7554/eLife.05198

**Download** [PDF](#)

---

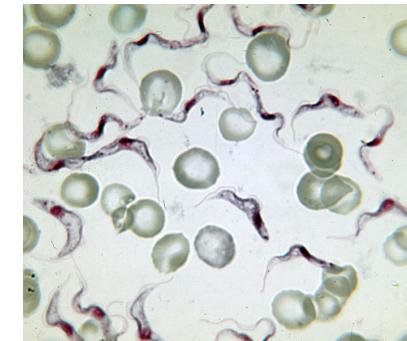
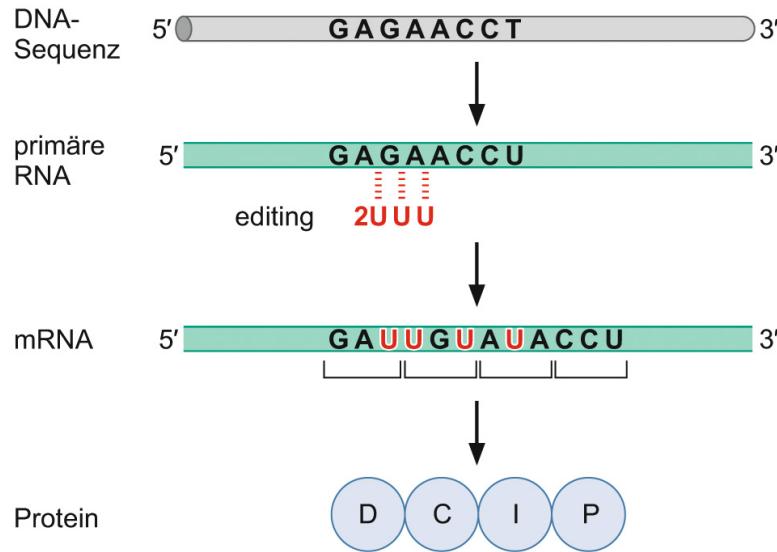
### Abstract

RNA editing by adenosine deamination alters genetic information from the genomic blueprint. When it recodes mRNAs, it gives organisms the option to express diverse, functionally distinct, protein isoforms. All eumetazoans, from cnidarians to humans, express RNA editing enzymes. However, transcriptome-wide screens have only uncovered about 25 transcripts harboring conserved recoding RNA editing sites in mammals and several hundred recoding sites in *Drosophila*. These studies on few established models have led to the general assumption that recoding by RNA editing is extremely rare. Here we employ a novel bioinformatic approach with extensive validation to show that the squid *Doryteuthis pealeii* recodes proteins by RNA editing to an unprecedented extent. We identify 57,108 recoding sites in the nervous system, affecting the majority of the proteins studied. Recoding is tissue-dependent, and enriched in genes with neuronal and cytoskeletal functions, suggesting it plays an important role in brain physiology.



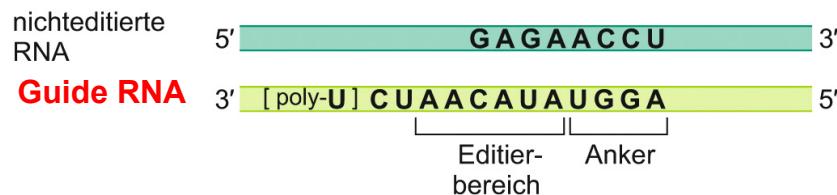
# RNA editing durch Insertion oder Deletion von U

(a)



...in Mitochondrien von Trypanosomen entdeckt.

(b)



Guide RNA spezifiziert Ort der U-Editierung.



LARGE-SCALE BIOLOGY ARTICLE

# Chemical Modifications Mark Alternatively Spliced and Uncapped Messenger RNAs in *Arabidopsis*<sup>OPEN</sup>

Lee E. Vandivier,<sup>a,b</sup> Rafael Campos,<sup>a</sup> Pavel P. Kuksa,<sup>c,d</sup> Ian M. Silverman,<sup>a,b</sup> Li-San Wang,<sup>c,d</sup> and Brian D. Gregory<sup>a,b,1</sup>

<sup>a</sup> Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

<sup>b</sup> Cell and Molecular Biology Graduate Program, University of Pennsylvania, Philadelphia, Pennsylvania 19104

<sup>c</sup> Institute for Biomedical Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania 19104

<sup>d</sup> Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania 19104

ORCID IDs: 0000-0002-1395-8571 (R.C.); 0000-0003-2248-6403 (P.P.K.); 0000-0001-7532-0138 (B.D.G.)

Posttranscriptional chemical modification of RNA bases is a widespread and physiologically relevant regulator of RNA maturation, stability, and function. While modifications are best characterized in short, noncoding RNAs such as tRNAs, growing evidence indicates that mRNAs and long noncoding RNAs (lncRNAs) are likewise modified. Here, we apply our high-throughput annotation of modified ribonucleotides (HAMR) pipeline to identify and classify modifications that affect Watson-Crick base pairing at three different levels of the *Arabidopsis thaliana* transcriptome (polyadenylated, small, and degrading RNAs). We find this type of modifications primarily within uncapped, degrading mRNAs and lncRNAs, suggesting they are the cause or consequence of RNA turnover. Additionally, modifications within stable mRNAs tend to occur in alternatively spliced introns, suggesting they regulate splicing. Furthermore, these modifications target mRNAs with coherent functions, including stress responses. Thus, our comprehensive analysis across multiple RNA classes yields insights into the functions of covalent RNA modifications in plant transcriptomes.

