

Eine RNA-Menagerie: miRNAs und andere kodierende und nichtkodierende RNAs

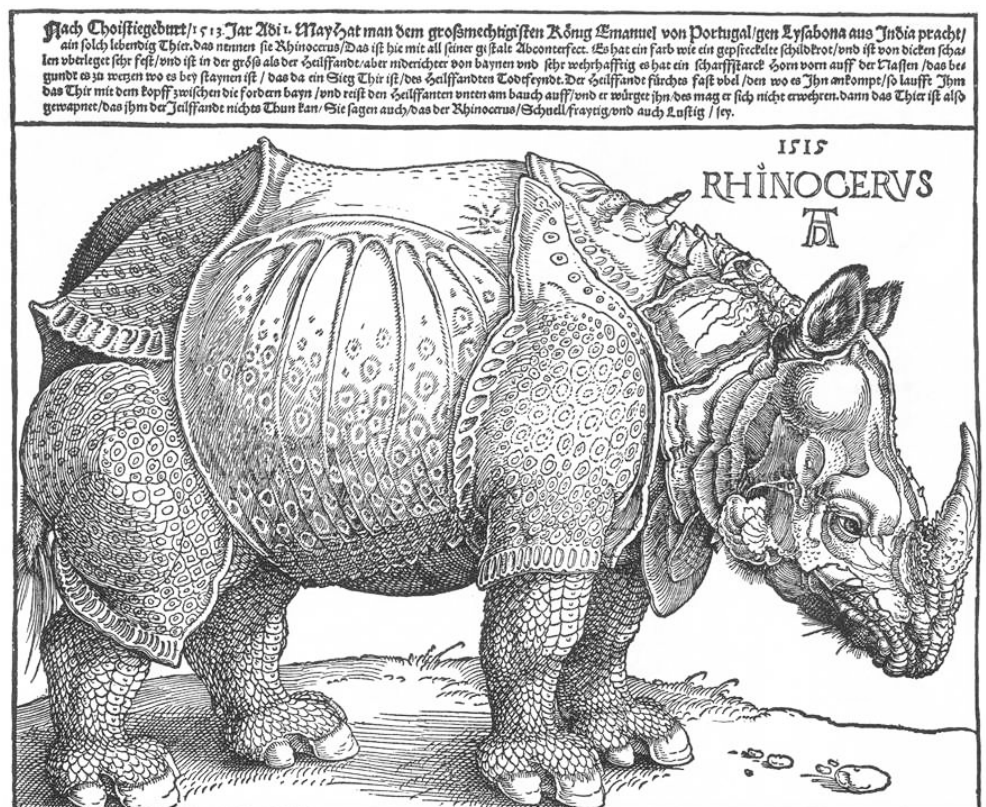
Peter N. Robinson

Institut für medizinische Genetik
Charité Universitätsmedizin Berlin

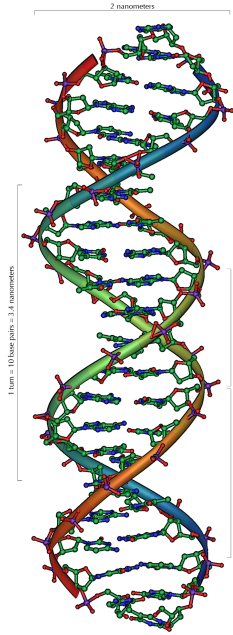
23. Juni 2008

Eine RNA-Menagerie

- mRNA
- tRNA
- rRNA
- snRNA
- snoRNA
- miRNA
- XIST-RNA
- piRNA

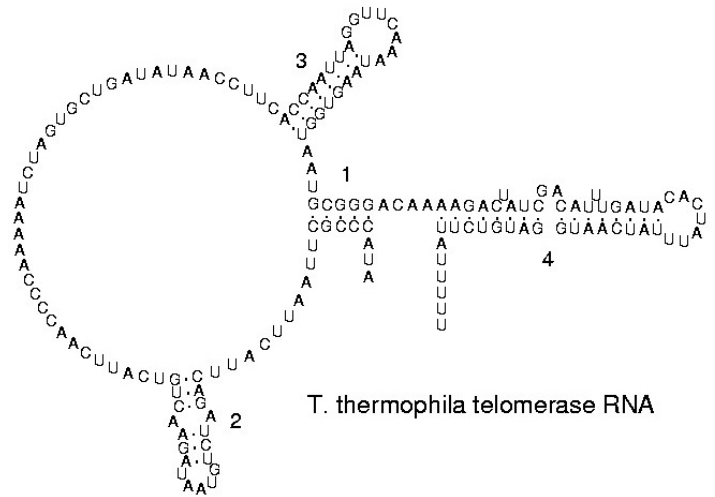


RNA vs. DNA (1): 1 vs 2 Stränge



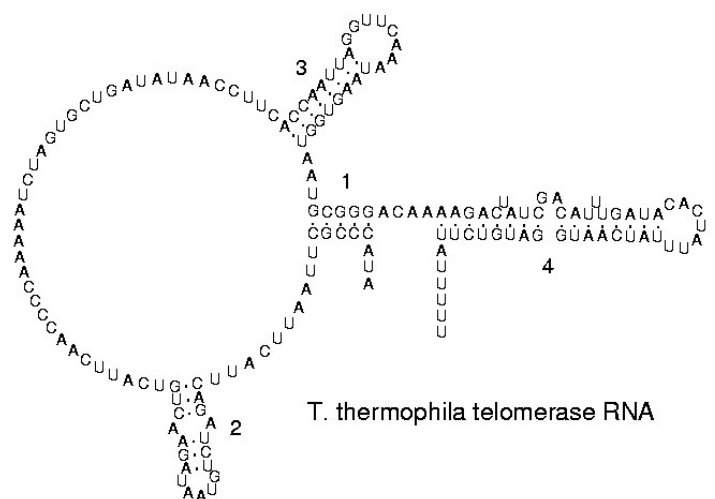
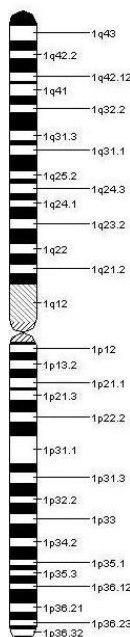
- DNA: I.d.R. doppelsträngig

Bildquelle: Wikipedia



- RNA: I.d.R. einzelsträngig, oft Sekundärstrukturen durch intramolekulare Wasserstoffbrücken

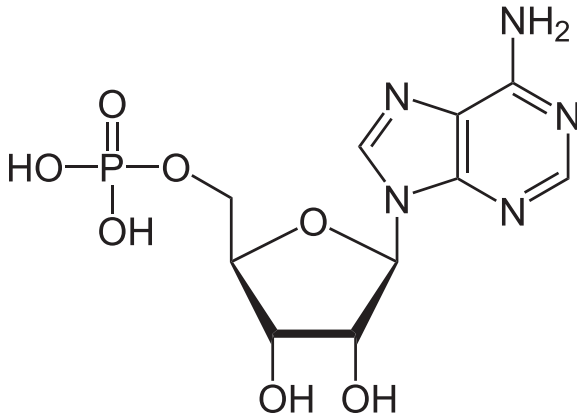
RNA vs. DNA (2): Länge



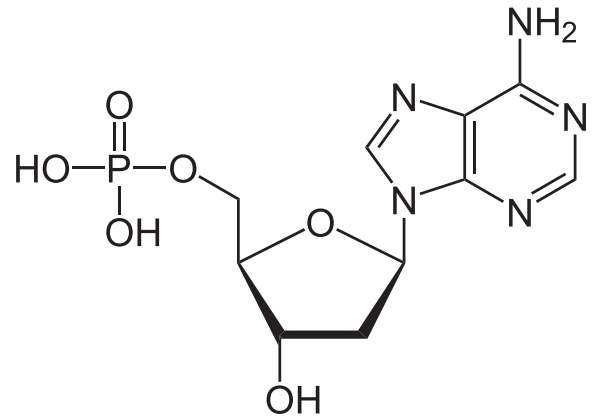
- DNA: Millionen von Basenpaaren
- RNA: ~20 bis mehrere Tausend Nukleotide

Bildquelle: Wikipedia

RNA vs. DNA (3): Biochemie



Adenosinmonophosphat (Wikipedia)



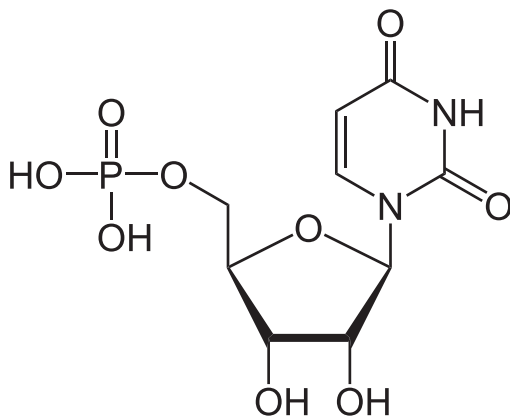
Desoxyadenosinmonophosphat (Wikipedia)

- RNA: Ribose

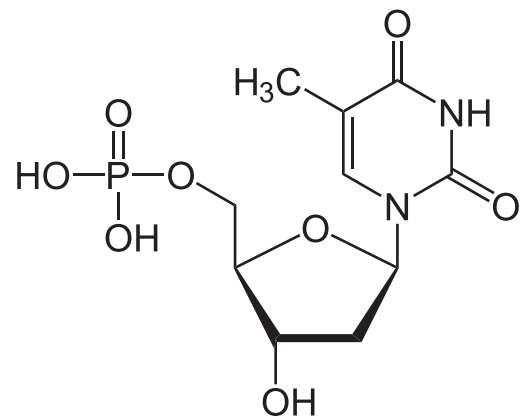
- DNA: 2-Deoxyribose

Die zusätzlich 2-Hydroxylgruppe macht die RNA weniger stabil als die DNA, da sie leichter hydrolysiert werden kann

RNA vs. DNA (4): Biochemie



Uridinmonophosphat (Wikipedia)



Desoxythymidinmonophosphat (Wikipedia)

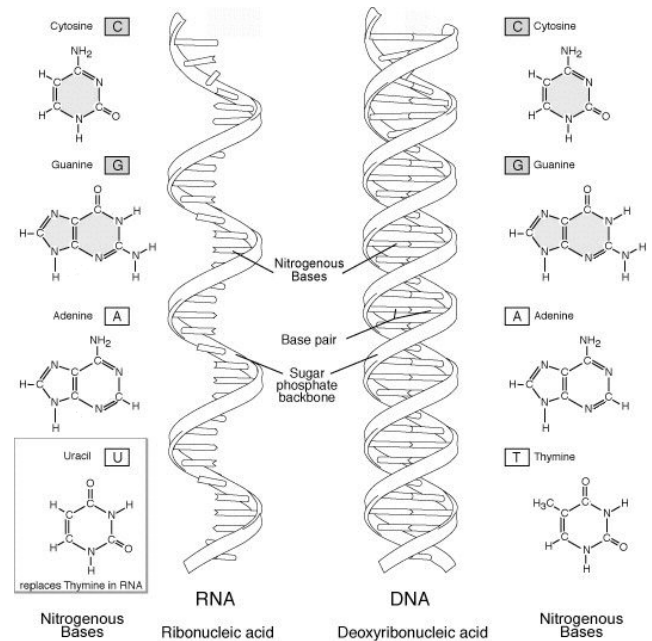
- RNA: Uracil

- DNA: Thymin

In der RNA ist nicht Thymin (T) sondern Uracil (U) zu Adenin (A) komplementär

RNA vs. DNA (5): Biologische Rollen

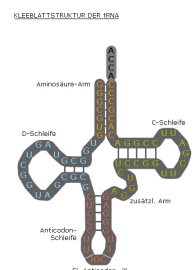
- DNA
 - ▶ Trägerin der Erbinformation.
- RNA
- Im Gegensatz zur DNA spielt die Struktur der RNA bei deren Funktion eine wesentliche Rolle
- 3D-Struktur aus mehreren kürzeren Helices, ähnlich wie Proteine
- Katalyse wie bei Enzyme
- Zahlreiche unterschiedliche Funktionen ...



Wikipedia

Ebenen der RNA-Struktur

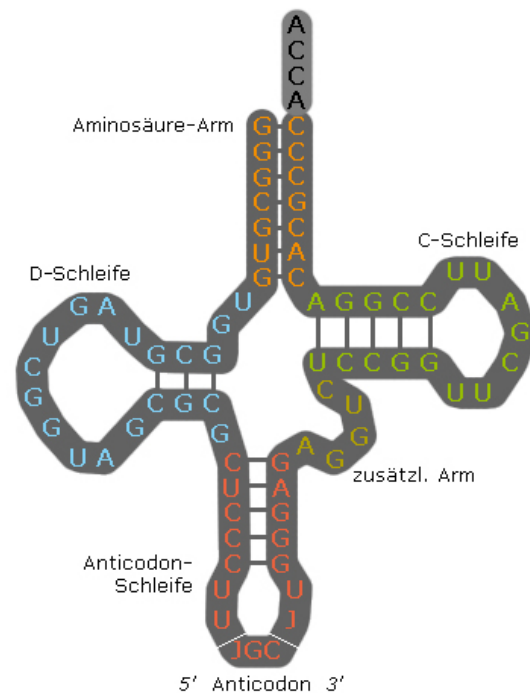
- Primärstruktur: Die Nukleotidsequenz
 - ▶ z.B. die Sequenz CUCUCGGUAAGCUUAGGUACCA
- Sekundärstruktur: Paare von Nukleotiden, welche eine Wasserstoffbrückenbildung miteinander eingehen
 - ▶ G–C: drei Wasserstoffbrücken
 - ▶ A–T: zwei Wasserstoffbrücken
 - ▶ G–U: Eine Wasserstoffbrücke ("wobble pair" → Wackelpaar)
- Hairpin- und Stemloop-Strukturen, Helixstrukturen sowohl Einzelstrang- als Doppelstrangbereiche.
- Tertiärstruktur (3D)
- Quarternärstruktur (Beziehung zu anderen RNAs/Proteinen im Komplex)



RNA-Sekundärstruktur

KLEEBLATTSTRUKTUR DER tRNA

- G–C: drei Wasserstoffbrücken
- A–T: zwei Wasserstoffbrücken
- G–U: Eine Wasserstoffbrücke ("wobble pair" → Wackelpaar)



Wikipedia

RNA-Sekundärstruktur

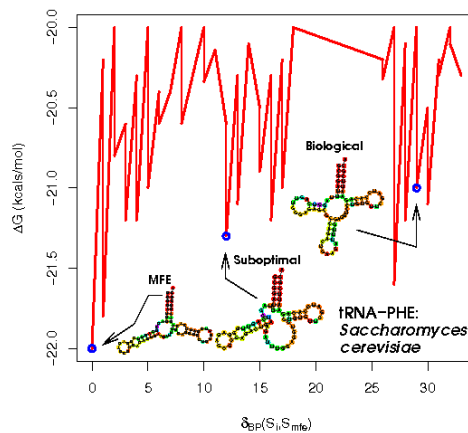
- RNA-Sekundärstruktur: häufige Motive wie Hairpin, Helix, Stem loop, bulge loop, interior loop, multiple loop
- RNA-Strukturbestimmung experimentell schwierig, daher ein wichtiges Thema für die Bioinformatik

Urheberrechtlich geschütztes Bild entfernt

Hairpin loop (Haarnadel)

Strukturvorhersage durch Maximierung von Basenpaarungen

- Minimieren der freien Energie
- Wasserstoffbrücken sind eine Schlüsselkomponente der RNA-Stabilität
- Viele Algorithmen versuchen daher, die Anzahl der Wasserstoffbrücken zu maximieren



S. cerevisiae tRNA-PHE: Energien alternativer Strukturen (Wikipedia)

Strukturvorhersage durch Maximierung von Basenpaarungen

Primary Proximity Constraint

Bilden Nukleotide i und j eine Wasserstoffbrücke, dann $|i - j| > 3$

Diese Bedingung ergibt sich aus der Tatsache, dass eine RNA-Kette nicht ausreichend flexibel ist, damit sich eine Wasserstoffbrücke zwischen eng benachbarten Nukleotiden bilden könnte.

Strukturvorhersage durch Maximierung von Basenpaarungen

Nesting Constraint

Sind (i, j) und (p, q) zwei Wasserstoffbrücken (Paare von Nukleotiden), wobei $i < p < j$, dann gilt $q < j$

Diese "Nistungsbedingung" verbietet überkreuzte Wasserstoffbrücken, erlaubt dagegen genistete Wasserstoffbrücken. Überkreuzte Wasserstoffbrücken, so genannte *Pseudoknoten*, kommen relativ selten vor. Algorithmen, welche Pseudoknoten zulassen, sind wesentlich weniger effizient als solche, die sie verbieten.

Strukturvorhersage durch Maximierung von Basenpaarungen

- Beispiel: UUGACAUCG
- Ziel: die Sekundärstruktur mit der maximalen Anzahl an Basenpaaren finden, wobei zwischen zwei paarenden Basen mindestens eine Ungepaarte stehen soll



RNA-Struktur: Klammern und Punkte

- Wir können die RNA-Struktur als Strings mit balanzierten Klammern und Punkten mit der entsprechenden Nistungsebene (nesting level) darstellen

UUGACAUCG	UUGACAUCG
(. .) (. . .)	(. (. . . .))
011001110	011222210

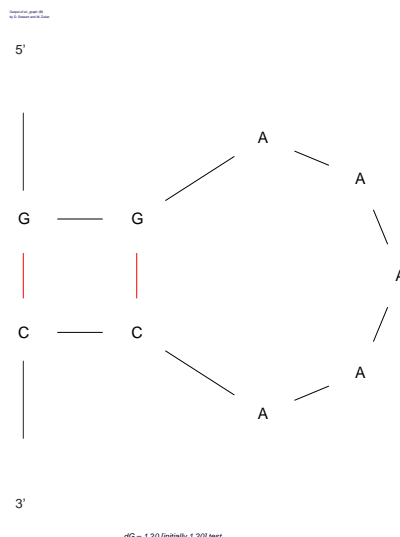
UUGACAUCG 2+3=5

UUGACAUCG 1+3=4

RNA-Struktur: Klammern und Punkte

- Die Sekundärstruktur kann von der Sequenz und Klammerndarstellung ermittelt werden
- Beispiel

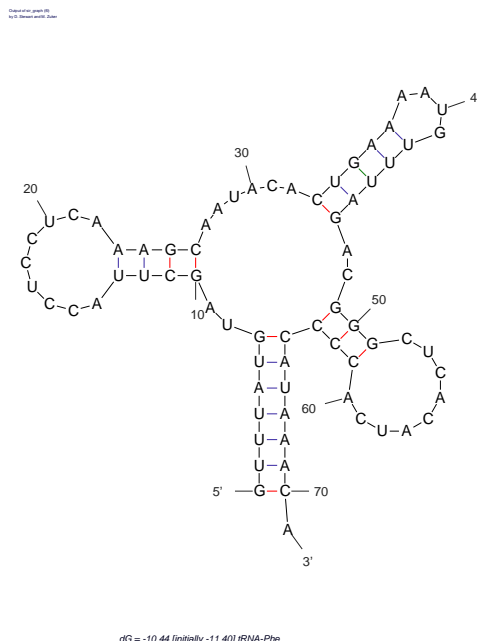
GGAAAAACC
 ((.))



Übung für zu Hause

- die DNA-Sequenz für das menschliche Mitochondriengenom aufrufen (Accessionnummer: 17981852)
- <http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=nucleotide&id=17981852>
- in der FEATURES-Liste nach dem Gen für tRNA-Phe suchen (Position 579–649), die Sequenz für dieses Gen im neuen Fenster aufrufen
- Display auf FASTA einstellen, die FASTA-formatierte Sequenz kopieren
- Die Sequenz mit dem mfold-Programm analysieren
- <http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1.cgi>

Übung für zu Hause



GUUUUAUGUAGCUUACCUCUCAAAGCAAUACACUGAAAAUGUUUAGACGGGCUCACAUACCCCAUAAACA
(((((((...((((.....)))))).....((((.....))))))..((((.....)))))))).

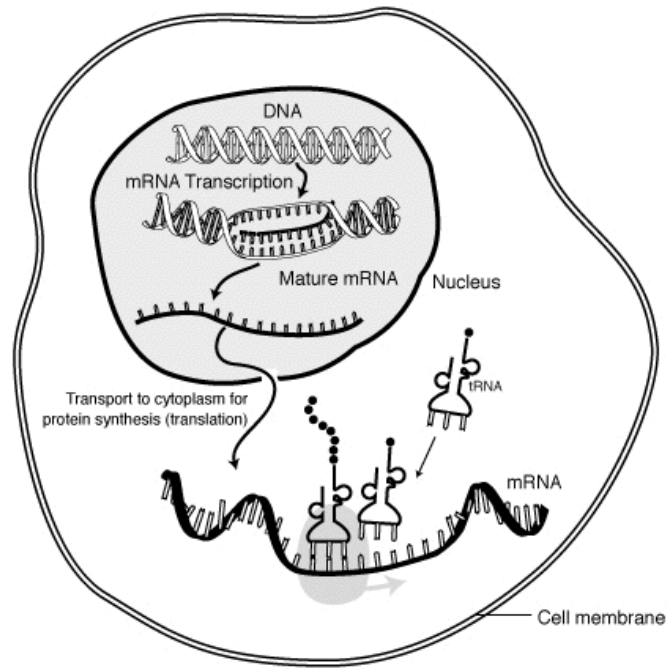
- Wieviele Strukturen werden vorhergesagt? Worin unterscheiden sie sich?

- Zahlreiche Algorithmen
- Dynamic programming
- Freie Energie
- Einzelheiten in späteren Semestern

Eine RNA-Menagerie

- Zahlreiche Klassen von RNA
- Es folgt zunächst ein Überblick über mRNA, tRNA, rRNA, snRNA, die Sie bereits kennen (sollen)

Urheberrechtlich geschütztes Bild entfernt

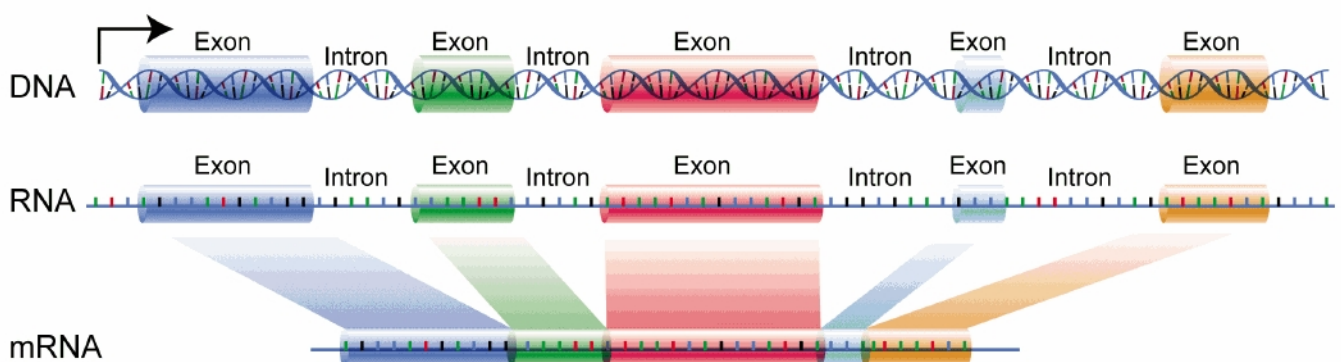


Wikipedia commons

mRNA: Spleißen

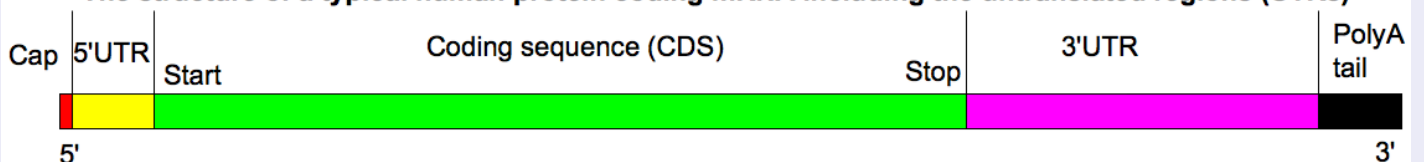
Wikipedia

mRNA-Spleißen



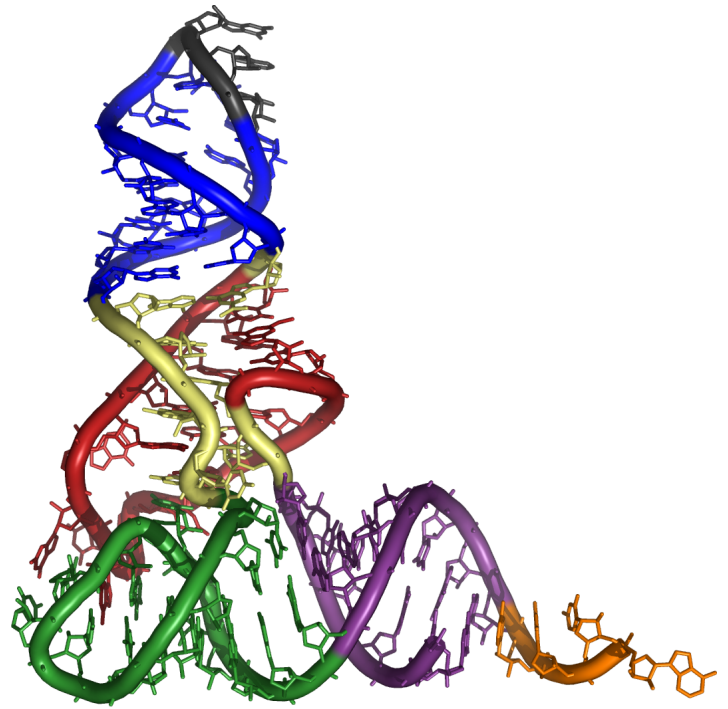
Struktur einer reifen mRNA

The structure of a typical human protein coding mRNA including the untranslated regions (UTRs)



tRNA

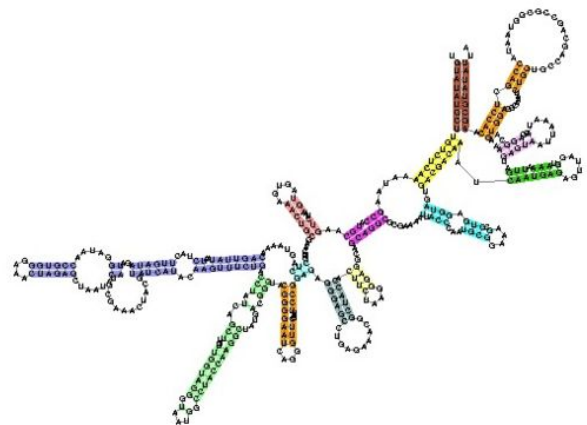
- ~ 85 Nukleotide lang
- Struktur etwa wie der Buchstabe "L"
- Drei-Nukleotid Anticodon auf der Spitze des L bindet an komplementäres Codon in mRNA
- Der "Fuß" des L bindet an eine der 20 Aminosäuren



Anticodonarm: blau, Anticodon: schwarz (Wikipedia commons)

rRNA

- ribosomale RNA
- zusammen mit den ribosomalen Proteinen am Aufbau und der enzymatischen Aktivität des Ribosoms und damit an der Proteinsynthese beteiligt.
- 60S Untereinheit (28S, 6S, 5S rRNA) und 40S Untereinheit (18S rRNA)



5' Domäne der kleinen rRNA (Wikipedia commons)

- small nuclear RNA
- Immer mit spezifischen Protein assoziiert: *small nuclear ribonucleoproteins* (snRNP)
- Spleißen
- Regulation von Transkriptionsfaktoren (7SK RNA)
- Aufrechterhaltung der Telomere

Das Spleißosom Urheberrechtlich geschütztes Bild

entfernt

miRNAs

- Wir werden uns in der verbleibenden Zeit mit micro-RNAs (miRNAs) beschäftigen
- miRNAs sind 1993 in *C. elegans* entdeckt worden
- Die große Bedeutung von miRNAs in einer Reihe von biologischen Prozessen auch bei Säugern ist wohl seit Anfang des Jahrtausends nach und nach klar geworden, zahlreiche Aspekte des miRNA-Metabolismus sind noch nicht geklärt
- Wichtiges Thema für die Bioinformatik: Beitrag der miRNAs zur Genregulation verstehen

- Sehr kurze RNA-Moleküle (~ 22 nt)
- Antisense-Regulatoren anderer Gene
- miRNAs entstehen aus Vorstufen mit ~ 70 nt, welche eine umgekehrte Wiederholungssequenz enthalten, die die Bildung einer Haarnadelstruktur ermöglicht
- Mindestens \sim Tausend miRNAs beim Menschen
- eine miRNA reguliert i.d.R. bis zu ein paar Hundert proteinkodierende Gene
- Grundsätzlich eine *negative* Regulation

Biogenese der miRNAs (1)

- miRNA-Vorstufen ([pri-miRNAs](#), primary microRNAs) werden als unabhängige miRNA-Gene transkribiert oder stellen in anderen Fällen Abschnitte von Introns proteinkodierender Gene
- eine pri-miRNA kann Sequenzen mehrerer miRNAs enthalten
- pri-miRNAs falten als Haarnadelstrukturen mit imperfekter Basenpaarung

Urheberrechtlich geschütztes Bild entfernt

Biogenese der miRNAs (2)

- pri-miRNAs werden dann durch die Endonuclease **Drosha** verarbeitet^a
- Das Ergebnis sind ~ 70 nt Haarnadeln namens **prä-miRNAs**

Urheberrechtlich geschütztes Bild entfernt

^aDrosha bindet an das Produkt des DiGeorge syndrome critical region Gens 8 (*DGCR8*)

Biogenese der miRNAs (3)

- In einigen Fällen entsprechen ausgespleißte Introns^b genau einer prä-miRNA-Sequenz (= **Mirtron**)
- Mitrons benötigen daher Drosha-DGCR8 nicht

Urheberrechtlich geschütztes Bild entfernt

^bMirtrons kommen bei *Caenorhabditis elegans*, *D. melanogaster* und Säugetieren vor.

Biogenese der miRNAs (4)

- prä-miRNAs werden durch Exportin-5 ins Zytoplasma transportiert

Urheberrechtlich geschütztes Bild entfernt

Biogenese der miRNAs (5)

- Im Zytoplasma werden die prä-miRNAs durch **Dicer**^a gespalten
- Das Ergebnis ist ein ~ 20 bp miRNA-Duplex

Urheberrechtlich geschütztes Bild entfernt

^aDicer bildet einen Komplex mit TAR RNA binding protein (TRBP).

Biogenese der miRNAs (6)

- Nach der Verabreichung durch Dicer werden die miRNAs in einen Ribonukleoproteinkomplex namens micro RNP (miRNPs) bzw. **miRNA-induzierte Silencing-Komplex (miRISCs)**
- Der Aufbau des miRISC ist an die prä-miRNA-Verarbeitung gekoppelt
- Ein Strang des Duplex bildet die reife miRNA, der andere wird abgebaut
- Die wichtigsten Proteinbestandteile des miRISC sind Proteine der **Argonaut-Familie**

Urheberrechtlich geschütztes Bild entfernt

Biogenese der miRNAs (7)

- Zusammengefasst...

Urheberrechtlich geschütztes Bild entfernt

Wie steuern miRNAs die Genexpression?

- miRNAs steuern die Expression von jeweils bis einigen Hundert Genen posttranskriptionell
- Verminderung der mRNA-Stabilität
- Verminderung der mRNA-Translation

Wie steuern miRNAs die Genexpression?

- Basenpaarung an die 3'-UTR der Ziel-mRNAs
- Perfekte Basenpaarung in der Saatregion (Nukleotide 2–8 der miRNA)
- Die Saatregion initiiert die miRNA-mRNA-Assoziation
- Fehlpaarung in der mittleren Region
- (Imperfekte) Basenpaarung in der 3'-Region der miRNA

Wie steuern miRNAs die Genexpression?

- miRNA-Bindungsstellen sind in der 3'-UTR der Ziel-mRNA gelegen
- Mehrfache Bindungsstellen sind in der Regel für eine wirksame Repression der Genexpression benötigt
- Synergistische Wirkung insbesondere von Bindungsstellen, die nah beieinander gelegen sind (10–40 nt)

Wie steuern miRNAs die Genexpression?

Urheberrechtlich geschütztes Bild entfernt

Urheberrechtlich geschütztes Bild entfernt

Chen K, Rajewsky N. (2007) *Nat Rev Genet.* 8:93–103.

miRNA-Funktionen

miRNA-Gene stellen ca. 1–2% aller Gene bei Eukaryonten dar. Die Funktionen der meisten miRNAs sind noch unbekannt.

- Regulatoren des zeitlichen Ablaufs der Entwicklung der Larvenstadien (lin-4, lin-7, *C. elegans*)
- Links-rechts-Asymmetrie der Chemorezeptorexpression (Isy-6, *C. elegans*)
- Apoptose, Fettstoffwechsel (miR-14, *D. melanogaster*)
- Hämatopoietische Differenzierung (miR-181a, Maus)
- Spaltung von Hox-B8-Transkripten (miR-196, Maus)
- Rolle bei Krebs, neurologischen Krankheiten, ...,?

miRNA vs. siRNA

Urheberrechtlich geschütztes Bild entfernt

- (A) dsRNA von Transposons, (B) Viren (C) miRNAs werden von Dicer prozessiert

miRNA vs. siRNA

Urheberrechtlich geschütztes Bild entfernt

- siRNA/miRNA bilden den RNA-induzierten Silencing-Komplex (RISC)
- Ist die Bindung an die Ziel-mRNA (nahezu) Perfekt, wird die Ziel-mRNA gespalten (eher der Fall bei siRNA)
- Ist die Bindung nicht perfekt, wird die mRNA destabilisiert bzw. die Translation gehemmt (eher der Fall bei miRNA)

- Piwi-Unterfamilie der Argonautenproteine
- Ziel: Transposons, Retroelemente
- Spaltung der Ziel-mRNA erzeugt eine neue piRNA
- Rolle vor allem im Keimgewebe, um neue Insertionen von Transposons zu verhindern
- Chapman, Carrington, Nature Genetics Reviews 2007

Urheberrechtlich geschütztes Bild entfernt

Bioinformatik der miRNAs

- Zwei Problemfelder
 - 1 miRNA-Gene vorhersagen
 - 2 miRNA-Zielgene vorhersagen

Vorhersage von miRNAs

- Experimentelle Identifikation von miRNAs
 - ▶ Klonierung/Sequenzierung der ~ 22-nt-Fraktion der Gesamt-RNA
 - ▶ Expression von miRNAs jedoch extrem variabel in unterschiedlichen Geweben bzw. Entwicklungsstadien
 - ▶ Expression von vielen miRNAs zu gering für einfache Detektion
 - ▶ Daher sind viele bekannte miRNAs stark bzw. ubiquitär exprimiert
- Bioinformatische Identifikation von miRNAs
 - ▶ Analyse von DNA-Sequenzen, vorhergesagten RNA-Sekundärstrukturen, Homologien
 - ▶ Vorhersagen können gezielt geprüft werden

Vorhersage von miRNA-Zielgenen

- miRNA-Mikroarrays mit derzeit über 700 menschlichen miRNA-Sonden
- Heute ist relativ wenig über miRNA-Zielgene (targets) bekannt
- Experimentelle Validierung von miRNA-targets kompliziert

Grundlagen der Vorhersage von miRNA-Zielgenen: Das 5'-Ende

- 5'-Ende der miRNA: 'Saat' (seed)
- 6–8 Nukleotide, perfekte Basenpaarung mit Zielgen
- Das 5'-Ende ist typischerweise ungepaart, oder beginnt mit Uracil und enthält i.d.R. keine G:U-Wobble-Paare
- Die Bindungsstelle im Zielgen ist oft von Adenosinresten flankiert.

Grundlagen der Vorhersage von miRNA-Zielgenen: Das 3'-Ende

- Perfekte Basenpaarung am 3'-Ende kann eine imperfakte Basenpaarung am 5'-Ende kompensieren
- Viele miRNA-Bindungsstellen haben einen 'Buckel' (Bulge) im mittleren Bereich

3' GAUGGUAUCCCAUU---UUGGUGAc 5' hsa-miR-140-5p	Alignment score: 161.0
	PhastCons score: 0.6432655
164:5' CUACCCUAGUUCAAAACAACCACUu 3' FBN1	Energy: -16.02

Vorhergesagte Bindungsstelle für miRNA-140 in *FBN1*

Vorhersage von miRNA-Zielgenen

- Konservierte miRNA-Saatsequenzen, oft mit flankierenden A's, deuten auf miRNA-Zielgene. Alignment der orthologen Sequenzen der 3'-UTR des HIC-Gens mit konservierter Saatsequenz für miR-23a¹

Urheberrechtlich geschütztes Bild entfernt

¹ Lewis BP, Burge CB, Bartel DP. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*. **120**:15-20.

Vorhersage von miRNA-Zielgenen

- Konservierung für 3227 menschliche Gene mit 14301 Zielsequenzen

Urheberrechtlich geschütztes Bild entfernt

Lewis et al (2005)

Vorhersage von miRNA-Zielgenen

- Andere Algorithmen verwenden ein Model der miRNA-Ziel-mRNA-Bindung
- Proximale Binding ≥ 7 nt, das 5'-Nukleotid kann an mRNA binden oder auch nicht
- α : Zentrale Schleife in beiden Sequenzen (2–3 nt)
- β : Zentrale Schleife (bulge) in Ziel-mRNA (2–5 nt)
- γ : Zentrale Schleife (bulge) im miRNA (6–9 nt)
- Distale Bindung ≥ 5 nt, bulges zugelassen

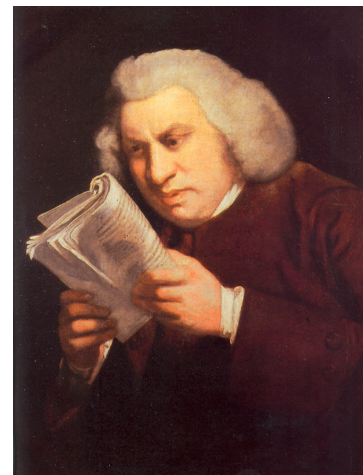
Kiriakidou et al (2004)

Genes Dev. 2004;**18**:1165-78.

Urheberrechtlich geschütztes Bild entfernt

The End of the Lecture as We Know It

- Diese Vorlesungsdias stehen unter der GNU-Lizenz für freie Dokumentation^b
- Kontakt:
peter.robinson@charite.de
- Chen K, Rajewsky N (2007) The evolution of gene regulation by transcription factors and microRNAs. *Nat Rev Genet.* 8:93-103
- Filipowicz W, Bhattacharyya SN, Sonenberg N. (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet.* 9:102-14.



Lectures were once useful; but now, when all can read, and books are so numerous, lectures are unnecessary. If your attention fails, and you miss a part of a lecture, it is lost; you cannot go back as you do upon a book... People have nowadays got a strange opinion that everything should be taught by lectures. Now, I cannot see that lectures can do as much good as reading the books from which the lectures are taken. I know nothing that can be best taught by lectures, except where experiments are to be shown. You may teach chymistry by lectures. You might teach making shoes by lectures!

Samuel Johnson, quoted in Boswell's Life of Johnson (1791).

^b<http://www.gnu.org/licenses/fdl.txt>