**Stichwort: ORF, Leserahmen bzw. Leseraster**

Als **offener Leserahmen** (OLR) oder **offenes Leseraster** (als Übersetzung von engl. *open reading frame*, ORF) wird in der Genetik derjenige Bereich der DNA bzw. mRNA bezeichnet, dessen Leserahmen zwischen einem Start-Codon und einem Stopp-Codon liegt. Der offene Leserahmen kodiert potenziell für die Aminosäuresequenz eines Peptids bzw. Polypeptids (wenn eine kurze Sequenz vorliegt) oder Proteins (wenn eine lange Sequenz vorliegt). Verbindungen mit 10 bis 100 Aminosäuren werden häufig als Polypeptide bezeichnet, ab 100 Aminosäuren spricht man im Allgemeinen von einem Protein.

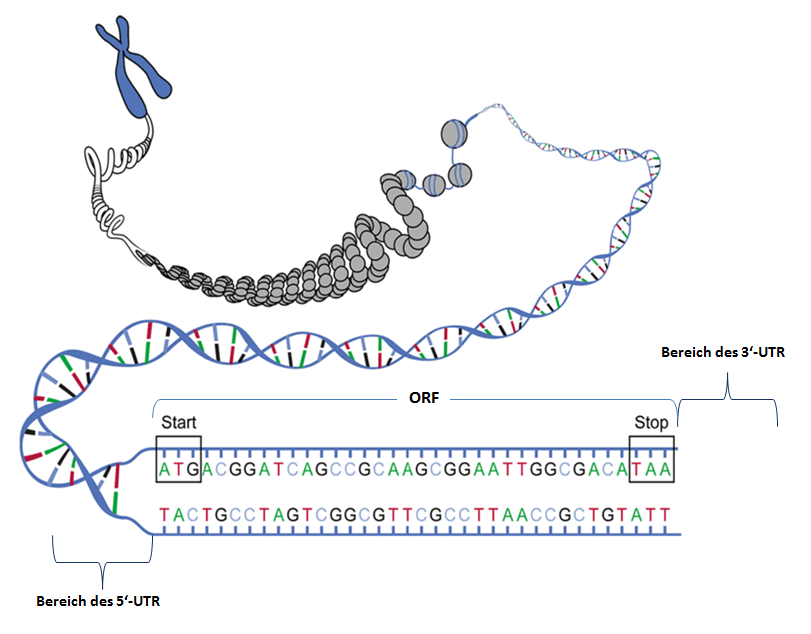


Abb.1 Schematische Darstellung von DNA mit einem ORF zwischen Start- und Stopp-Codon; dieser Bereich wird im Allgemeinen nach der Transkription auch translatiert. Dieser Protein-kodierende Bereich wird von nicht translatierten Bereichen (untranslated regions, *UTR*) flankiert, welche ebenfalls in mRNA überschrieben werden.

Offene Leserahmen werden von nicht-codierenden Bereichen des Gens umgeben, dem 5' UTR-Bereich und dem 3' UTR-Bereich. UTR steht für **u**n**t**ranslated **r**egion. Dabei handelt es sich um Regionen eines Gens, die zwar bei der Transkription in mRNA transkribiert werden, bei der Translation jedoch nicht für eine Aminosäuresequenz codieren. In diesen Bereichen liegen wichtige regulatorische Informationen für die Translation des offenen Leserahmens.

**Gensuche**

**Warum ist das Wissen über mögliche ORFs innerhalb einer DNA-Sequenz wichtig?**

Aus Sicht der Bioinformatik ist eine DNA-Sequenz eine Zeichenkette, in der sinntragende, kodierende Teile „eingestreut“ sind. Eine wichtige Aufgabe der Bioinformatik ist es, diese DNA-Abschnitte, die Gene, zu finden. Ziel dieses Prozesses muss es sein, mit höchster Genauigkeit diese informationstragenden Teilsequenzen (ORFs) aufzuspüren und in eine Aminosäuresequenz zu übersetzen, um beurteilen zu können, ob es sich um ein interessantes, unbekanntes oder aber ein bereits bekanntes Polypeptid oder Protein handelt. Nicht jeder ORF, der eine kurze Peptid- oder Proteinsequenz liefert, ist aber auch ein Gen! Dazu werden auch noch zusätzliche regulierende DNA-Abschnitte vor und hinter dem Gen benötigt, wie z.B. eine Promotorsequenz und eine Terminatorsequenz. Nicht jeder ORF ist also ein Gen – aber jedes Gen enthält mindestens einen ORF! Gene sind diejenigen ORFs, die tatsächlich von der Zelle in Polypeptide oder Proteine übersetzt werden.

Hat man erst neue interessante Gene entdeckt, können viele biologische Prozesse besser verstanden werden.

Ein ORF ist ein Stück DNA, welches von einem Start- und einem Stoppcodon flankiert wird und eine, ganzzahlig durch 3 teilbare Anzahl von Basen (die Codonen, englisch *codons*) umfasst. Da jeder der sechs möglichen Leserahmen codieren kann, überlappen sich ORFs häufig. Es ist die hohe Kunst der Genidentifikation, aus der großen Menge der ORFs die tatsächlich verwendeten Gene herauszufiltern.

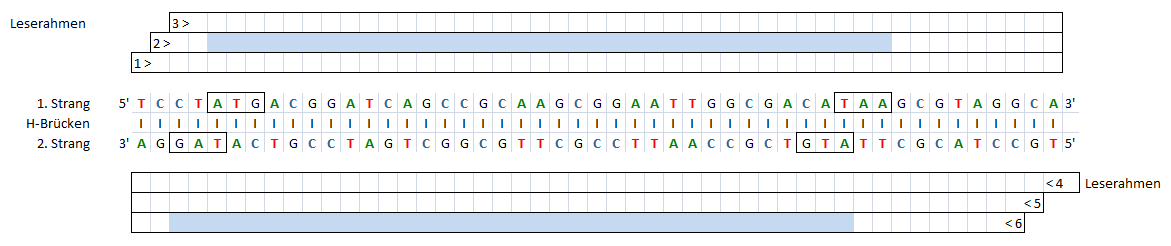


Abb.2 Schematische Darstellung der sechs möglichen Leserahmen; nur in zwei der sechs Leserahmen finden sich ORFs. Diese beginnen beide mit einem Start-Codon (ATG) und enden mit einem der drei möglichen Stopp-Codons (TGA, TAG, TAA).

Ein vielversprechender ORF, der mit hoher Wahrscheinlichkeit einem Gen zuzuordnen ist, kodiert für eine Aminosäuresequenz, die mit einem Methionin beginnt und nach einer *gewissen Länge* der Kette mit einem STOPP (oft als \* dargestellt) endet:

z. B. **M** E I N P R O T E I N **\***

oder **M** E I N L A N G E S L A N G E S P R O T E I N **\***

aber nicht **M** E**\***

Ein ORF kann – nach theoretischer Übersetzung – auch innerhalb einer Sequenz gefunden werden, die ohne ein Methionin, beginnt:

z. B. A E S L P C **M** E I N P R O T E I N **\*** G S A E H

oder entsprechend Abbildung 2:

2. Leserahmen P **M** T D Q P Q A E L A T **\*** A **\*** A

oder

6. Leserahmen P T L **M** S P I P L A A D P S **\***

Eine Reihe von online verfügbaren Programmen ermöglichen die ORF-Suche in frei wählbaren DNA-Sequenzen (z. B. Translate, siehe Abb.3).

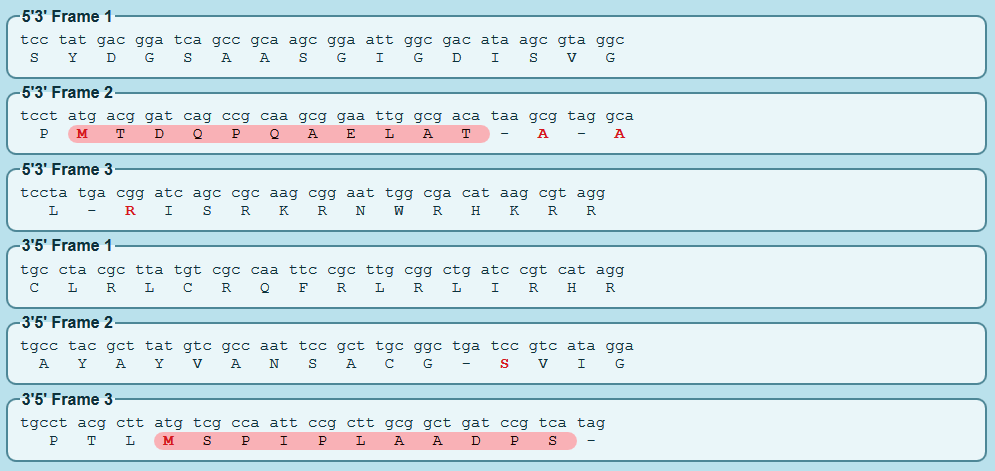


Abb.3 Darstellung der ORF-Suche mit dem Programm Translate (<https://web.expasy.org/translate/>) für die Sequenz 5‘ TCCTATGACGGATCAGCCGCAAGCGGAATTGGCGACATAAGCGTAGGCA 3‘ (siehe Abb.2)