

ATBC – OPSGEM - GENETICA (OPSGEM) - BIO-INFORMATICA

BIO-INFORMATICA

**OPSPOREN VAN GENETISCHE MUTATIES BIJ ERFELIJKE
ZIEKTES**

INTRODUCTIE
BIOLOGISCHE DATABASES

AGENDA

- Introductie Genetica (OPSGEM)
- Project Genetica (OPSGEM)
- Biologische databases
- Zoeken naar informatie in databases

EVEN VOORSTELLEN!

GENETICA (OPSGEM)

Bio-informatica

- DNA en RNA annotatie
- Biologische databases
- PCR & gelelektroforese
- Primer design en tools
- Genetisch onderzoek
- Sanger sequencing
- Exoom sequencing

Life science

- Mitose, Meiose
- De wetten van Mendel
- Genetica
- Opbouw chromosomen
- Genotype – fenotype
- Eigenschappen DNA
- Centrale dogma
- Mutaties
- Replicatie

Chemie

- Zuur-base reacties
- Redoxreacties
- Endo-/Exotherme reacties
- Enthalpie, Entropie, vrije energie
- Reactiesnelheid + effecten
- Chemisch evenwicht
- Intermoleculaire krachten
- Mengsels, oplossingen
- Osmose / Osmotische druk

Lab practicum

- Introductie practicum
- DNA isolatie
- PCR
- Gelelektroforese

Workshops

- Workshop schrijven:
 - Herhaling samenvatting
 - Materiaal & Methode
 - APA
- Professional skills 2
- Excel

Tutor

- Project Petra

TOETSING GENETICA (OPSGEM)

Naam	Code	Beoordeling	Beoordelende docent
Lab praktijk	TOETS-01	V/NV	Praktijkdocent
Individueel onderzoeksverslag	TOETS-02	Cijfer	Rick & Gonny
IPV	TOETS-03	V/NV	Tutor
Thematoets	TOETS-04	Cijfer	Emma
Kennistoets	TOETS-05/06	Cijfer	Rick
Professional skills 2	TOETS-07	V/NV	Anne-Marie
Opdracht Excel	TOETS-08	V/NV	Philip

Eindcijfer OPSGEM:

(1 x onderzoeksverslag + 2 x thematoets + 2 x kennistoets) / 5

OVERZICHT BI2B BIO-INFORMATICA

Week	Onderwerpen theorieles	Onderwerpen werkcollege
1	Introductie, biologische databases	Stamboom analyse, zoeken in OMIM
2	PCR, primer design, primer eigenschappen	Zoeken in NCBI, gen naar mRNA sequentie, primer design
3	Gelelektroforese, Sanger sequencing, genstructuur, genetische varianten	Tools voor primer design en primer analyse, blast resultaten primer analyse, bekende mutaties op het kandidaatgen
4	Introductie + toepassingen next generation sequencing & third generation sequencing, genome assemblies, genome browsers	Primer design, zoeken naar RP genen in genome browsers
5	Genetische markers, mutaties & genetische varianten	Targeted next generation sequencing data analyse, weektaak doornemen
6	Onderzoek doornemen, oefenen met de stof	Exome sequencing data analyse
7	Oefentoets, vragen	Oefentoets thematoets, weektaak afmaken, vragen

VERSCHILLENDE OF GROTE OVEREENKOMSTEN?



<https://www.shutterstock.com/>, 06-11-2021

<https://thispersondoesnotexist.com/>, 03-11-2021

https://nl.wikipedia.org/wiki/Narcis#/media/Bestand:Narcissus_flowers.jpg, 06-11-2023

VERSCHILLENDE OF GROTE OVEREENKOMSTEN?

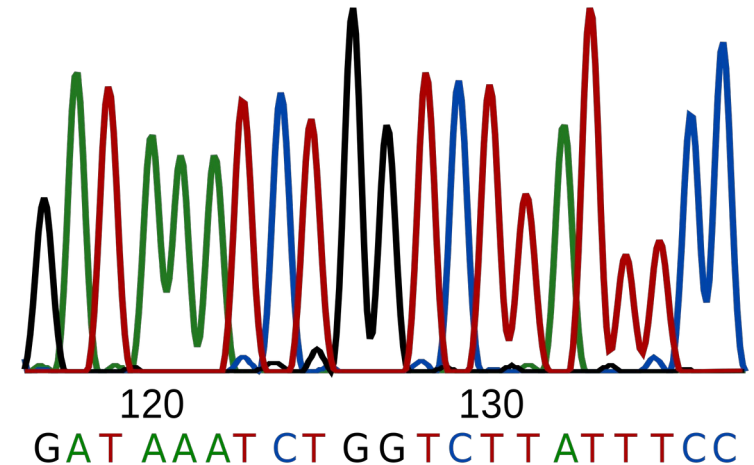


HOE ONDERZOEKEN WE HET DNA VAN ORGANISMEN?

Eerste stap: **DNA sequencen.**

Er zijn verschillende DNA sequencing methoden ontwikkeld:

- Sanger sequencing
Ontwikkeld in 1970
- Next generation sequencing
Ontwikkeld begin 21^e eeuw
- Third generation sequencing
Nu steeds meer in gebruik

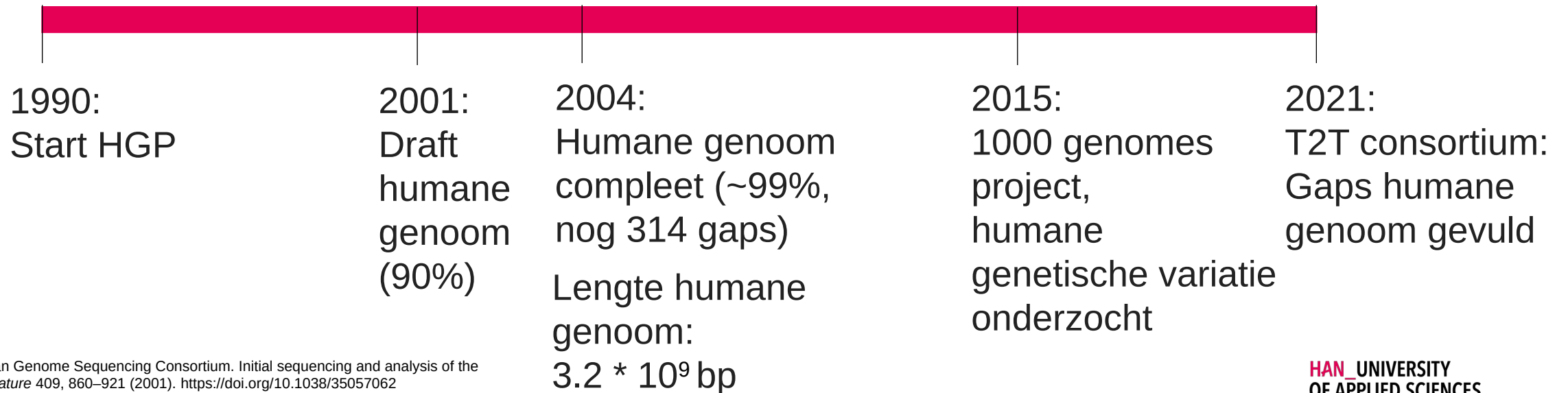


ONDERTUSSEN IS HET HELE HUMANE GENOOM GESEQUENCED

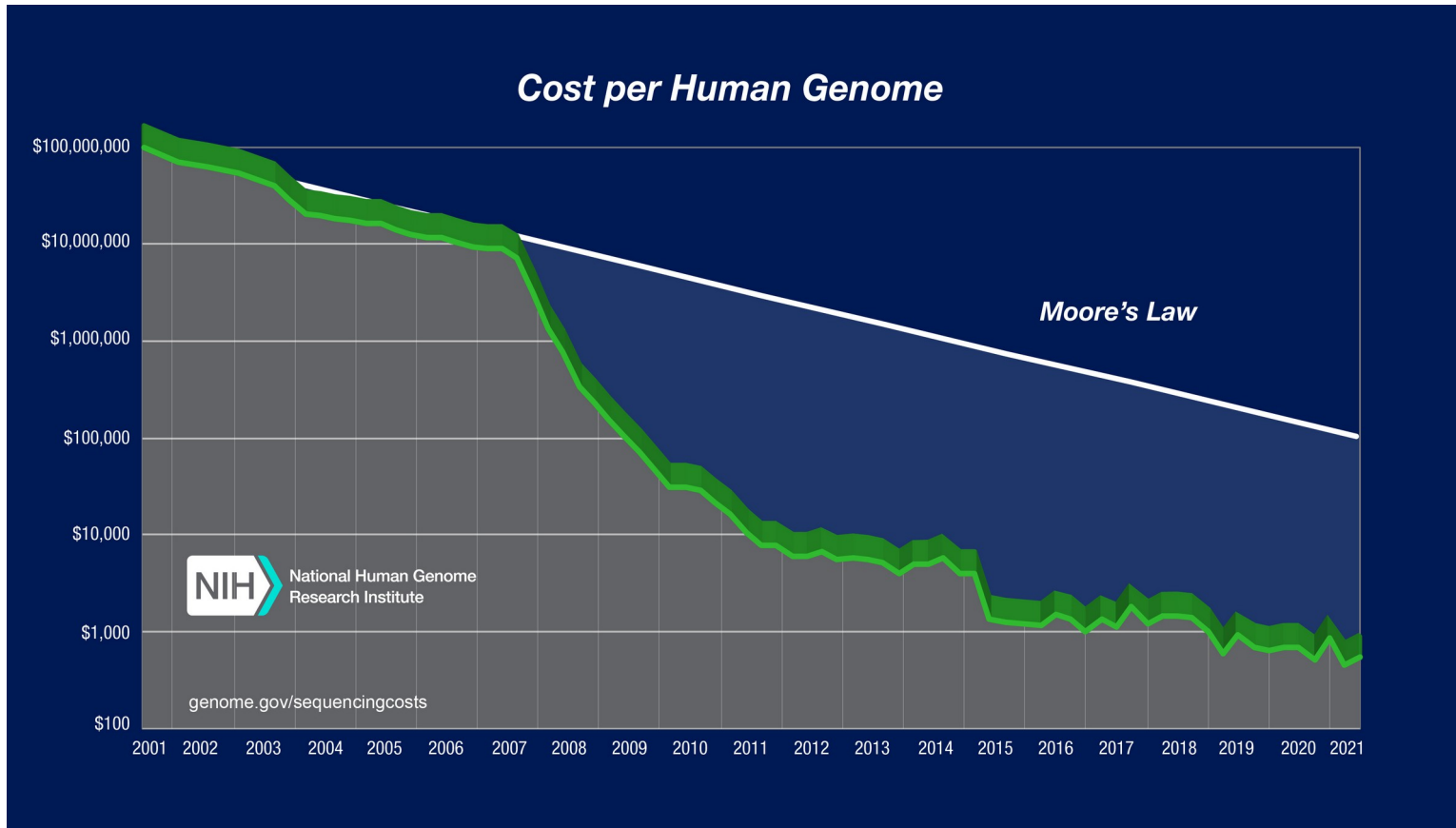
Human Genome Project:

Bepalen van DNA sequentie van het humane genoom

Totale kosten: ~3 miljard dollar



KOSTEN SEQUENCEN HUMAAN GENOOM



Het sequencen van DNA wordt steeds goedkoper.

En gaat steeds sneller!

EEN DEEL VAN DE SEQUENTIE VAN HET HUMANE GENOOM

....GCTCCACCACGGCCTCCTCGGCCACTGTCCCCTGGG
AGGGCAGCTGTGGTAAAGGCCGGAGCTCCCAGCTTTGG
GCAGGTGAGTGCCCCTGGCAGTTCTTTTCTGTGATGTAG
GTTTTTCAGACTGGGAAAAGTTGAGAGTTTCAAAGTCCA
TTGCCAGTGGGAACTGGAACCAGGCAAGCTGAACCAAG
TTCATTAGTGCTCTTGGCAACCTCAGGGCTCACCTGGTG
CACGGGGACCTTTGCAATGGCCAGGGCCTGGGGGCCAC
CCGAGCTAGGGCAAGGGGAGGGGGAGGGATGTGTTTAT
AAAATTTCTGTTTTAATTTCAAGTACAGTAATGTTGGTGGA
TAGAAACACACAAACCACAGCACTTTGATTTTGTCTAGTAA
CTCTTAAGAGTACGGAGGGTCCTGAGGCTGGGGGGGTCT
CGTGGGCACAGAGTATGATGCCTGTGAGGACGTCCTTC
CAGCCACACAGCCCGTGGACTGCAGCATTGAGTGTTCTA
TTTCCTGGGTGTCACGTCCGCAAAATCCGCTGTTTTGCA
GGGTCAGTCTTCGAGAAATGCTCTTGTAAGAACAGGTTT
TTAG...

Maar wat betekenen al deze letters nou?

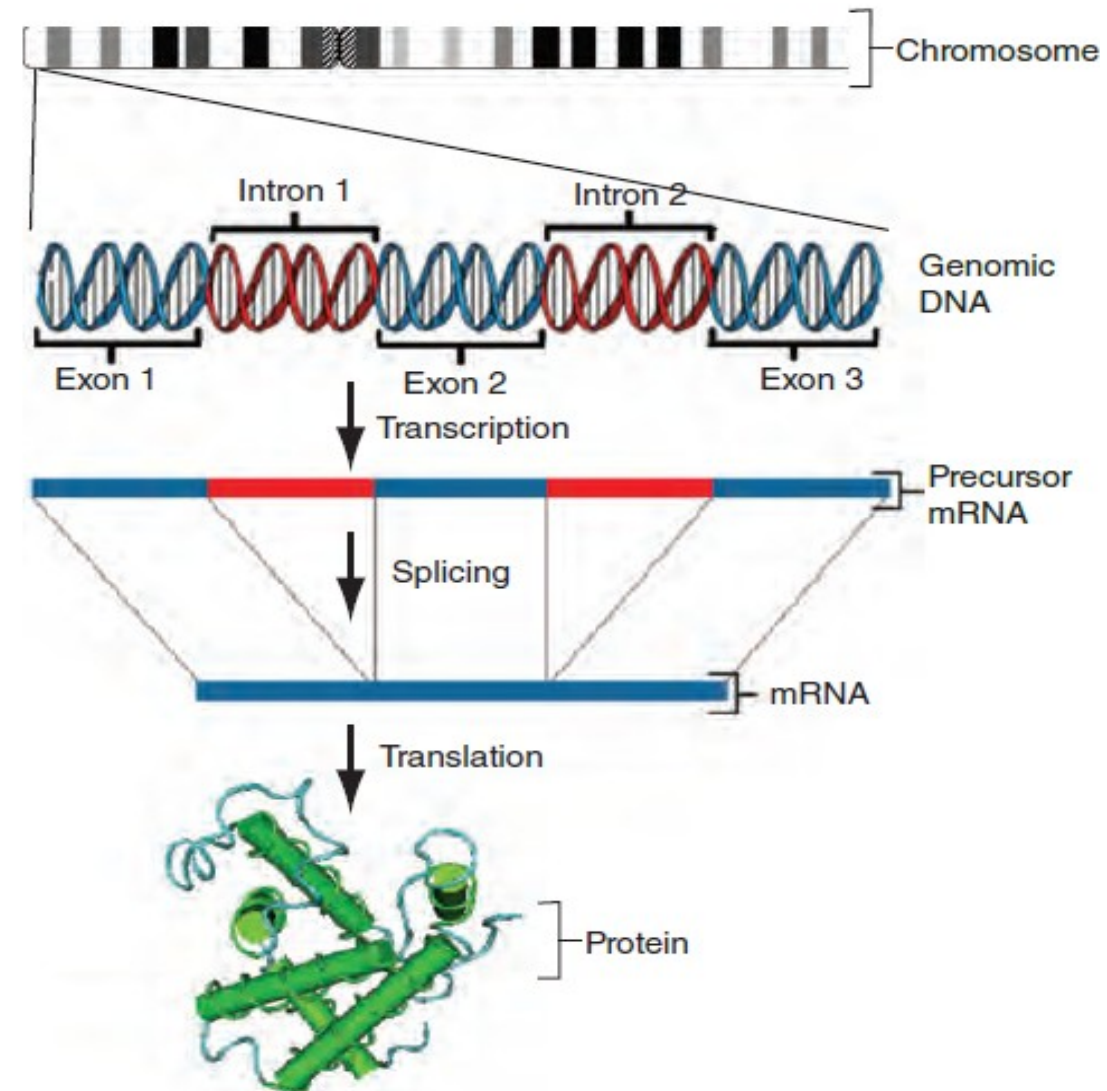
Na het verkrijgen van de sequentie komt het volgende lastige deel: uitvinden wat de DNA code betekent.

Wat kan je allemaal in een DNA sequentie vinden?

Bio-informatica voor nodig!

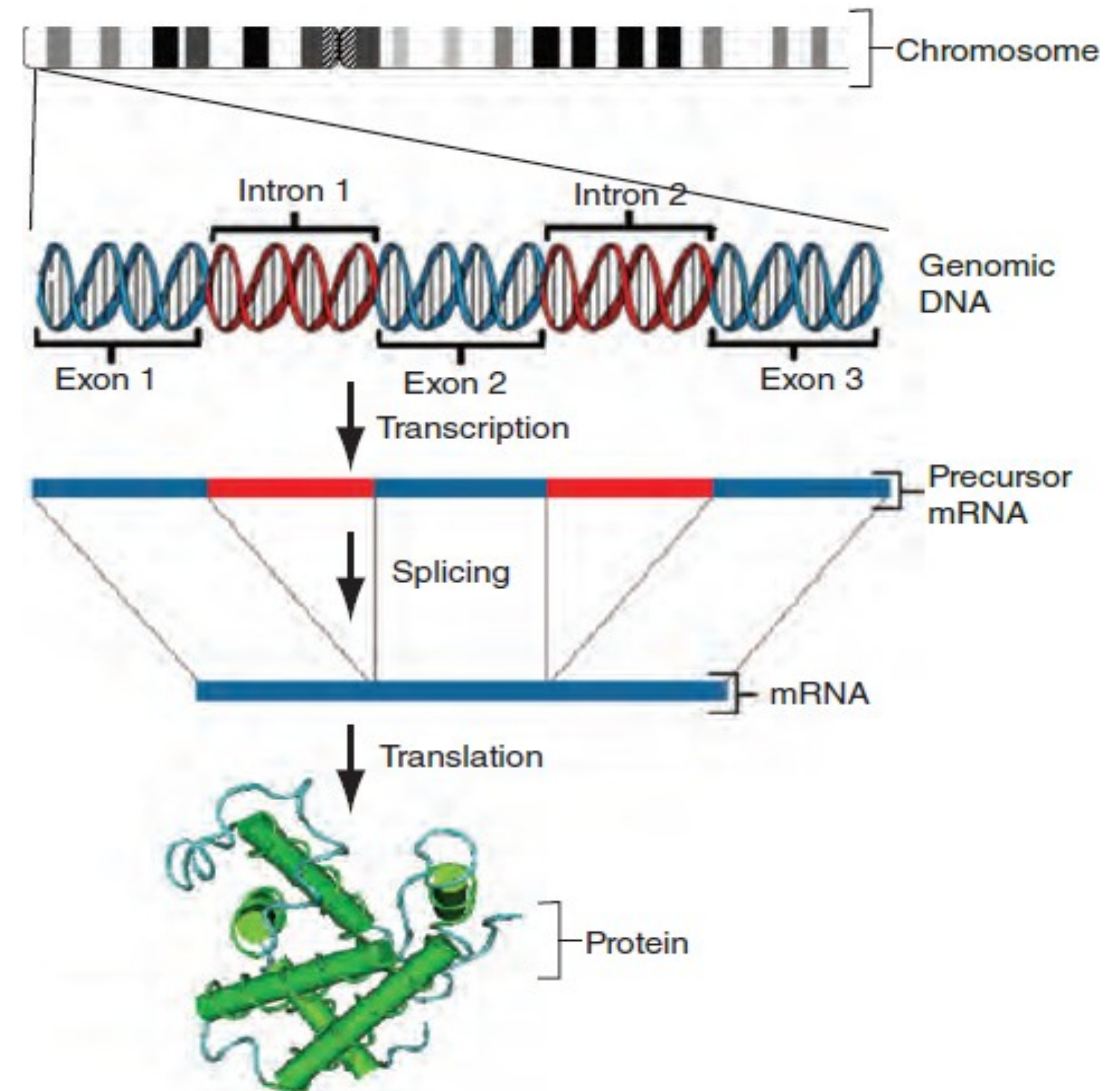
GENOOM VS. EXOOM

- Een **genoom** bestaat uit **chromosomen**
- **Chromosomen** bevatten **genen**
- **Genen** zijn opgebouwd uit **exonen** en **intronen**
- Het **exoom** bestaat uit alle **exonen**
Exoom = het eiwit coderende deel van het genoom



WAT KAN JE SEQUENCEN?

- Het hele genoom
- Het hele exoom
- Een geselecteerde set aan genen
- Één enkel gen
- Specifiek deel van het DNA



EERSTE TOEPASSING EXOOM SEQUENCING DOOR HET RADBOUDUMC



The NEW ENGLAND
JOURNAL of MEDICINE

HOME

ARTICLES & MULTIMEDIA ▾

ISSUES ▾

SPECIALTIES & TOPICS ▾

FOR AUTHORS ▾

CME »

ORIGINAL ARTICLE

Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability

Joep de Ligt, M.Sc., Marjolein H. Willemsen, M.D., Bregje W.M. van Bon, M.D., Ph.D., Tjitske Kleefstra, M.D., Ph.D., Helger G. Yntema, Ph.D., Thessa Kroes, B.Sc., Anneke T. Vulto-van Silfhout, M.D., David A. Koolen, M.D., Ph.D., Petra de Vries, B.Sc., Christian Gilissen, Ph.D., Marisol del Rosario, B.Sc., Alexander Hoischen, Ph.D., Hans Scheffer, Ph.D., Bert B.A. de Vries, M.D., Ph.D., Han G. Brunner, M.D., Ph.D., Joris A. Veltman, Ph.D., and Lisenka E.L.M. Vissers, Ph.D.
October 3, 2012 | DOI: 10.1056/NEJMoa1206524

Tutor project
opgezet in
samenwerking
met het
RadboudUMC

AGENDA

- Introductie Genetica (OPSGEM)
- Project Genetica (OPSGEM)
- Biologische databases
- Zoeken naar informatie in databases

HET PROJECT



<https://thispersondoesnotexist.com/>, 03-11-2021

Dit is Petra

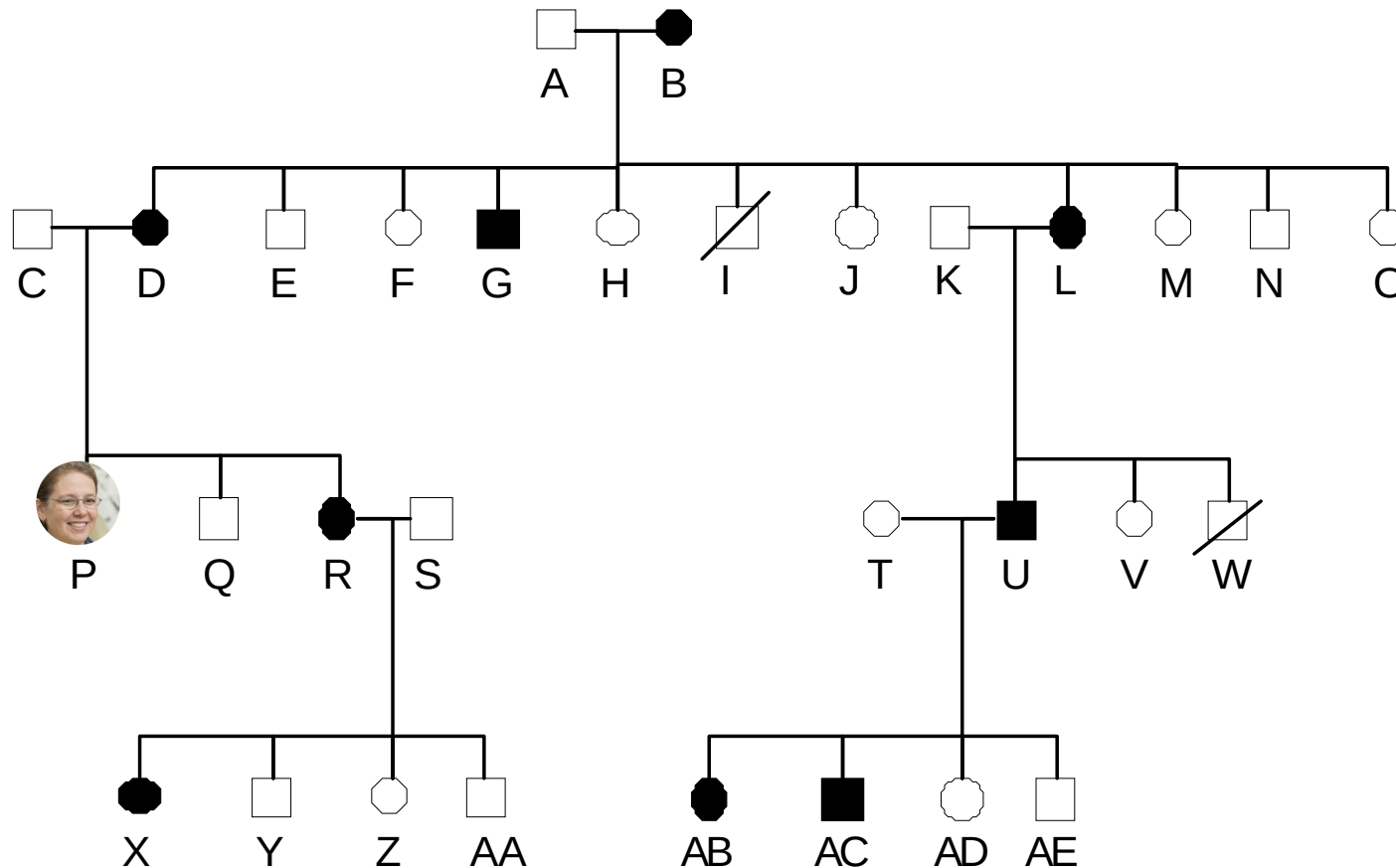
Het ziektebeeld van Petra:

- Symptomen begonnen vanaf haar 22e
- Progressieve degeneratie van het netvlies
- Slecht zien in het donker
- Langzame afname van het centrale gezichtsveld

Welke mutatie in welk gen is verantwoordelijk voor de ziekte van Petra?

→ Petra diagnosticeren door het doen van genetisch onderzoek

DE STAMBOOM VAN PETRA



□, ○ = Man, vrouw niet 'ziek'

■, ● = Man, vrouw wel 'ziek'

/ = Overleden

PETRA DIAGNOSTICEREN: WELKE ZIEKTE, WELK GEN, WELKE MUTATIE



Week 1

- Stamboom analyse
- Diagnose erfelijke ziekte Petra

Week 2

- Kandidaat gen kiezen
- Primer design voor amplificatie van kandidaat gen (handmatig)

Week 3

- Primer design voor amplificatie van kandidaat gen (m.b.v. tools)
- Uitzoeken van bekende mutaties in het kandidaatgen

Onderzoek naar één enkel kandidaat gen

PETRA DIAGNOSTICEREN: WELKE ZIEKTE, WELK GEN, WELKE MUTATIE



Week 4

Week 5

Week 6

Week 7

- Structuur van genen die te maken hebben met Petra's ziekte onderzoeken
- Verdiepen in genetische variatie en de effecten van variatie

- Genetische variatie onderzoeken in genen die te maken hebben met Petra's ziekte

- Genetische variatie onderzoeken in het exoom van Petra

- Schrijven van onderzoeksverslag

Onderzoek naar alle genen die te maken hebben met de ziekte van Petra

Onderzoek doen naar het hele exoom van Petra

HET LEVEN VAN PETRA



@petra_fan_account



AGENDA

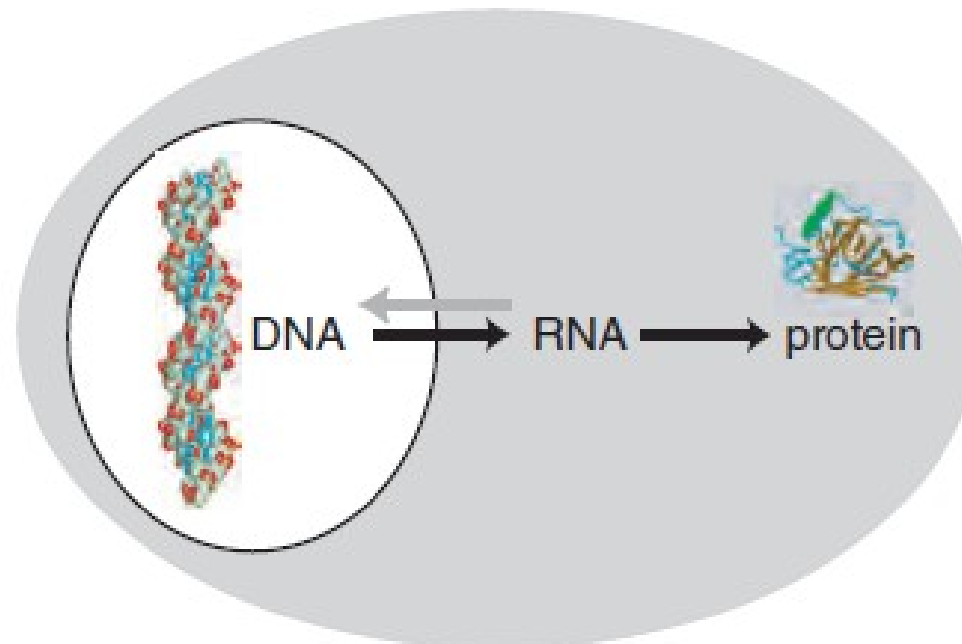
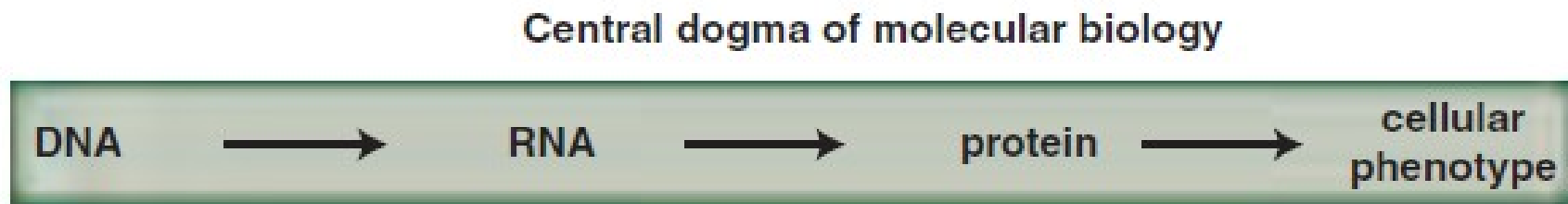
- Introductie Genetica (OPSGEM)
- Project Genetica (OPSGEM)
- Biologische databases
- Zoeken naar informatie in databases

DOELEN

Na deze week:

- Weet je in welke biologische databases je nucleotide (DNA, mRNA) sequenties kan vinden
- Weet je in welke database je informatie over ziektes kunt vinden
- Ken je verschillende vormen van data annotatie
- Weet je wat voor soort zoekvragen je aan biologische databases kunt stellen
- Kun je DNA/mRNA/eiwit sequenties uit databases verkrijgen
- Kun je informatie over ziektes en aan ziektes gerelateerde genen uit databases verkrijgen

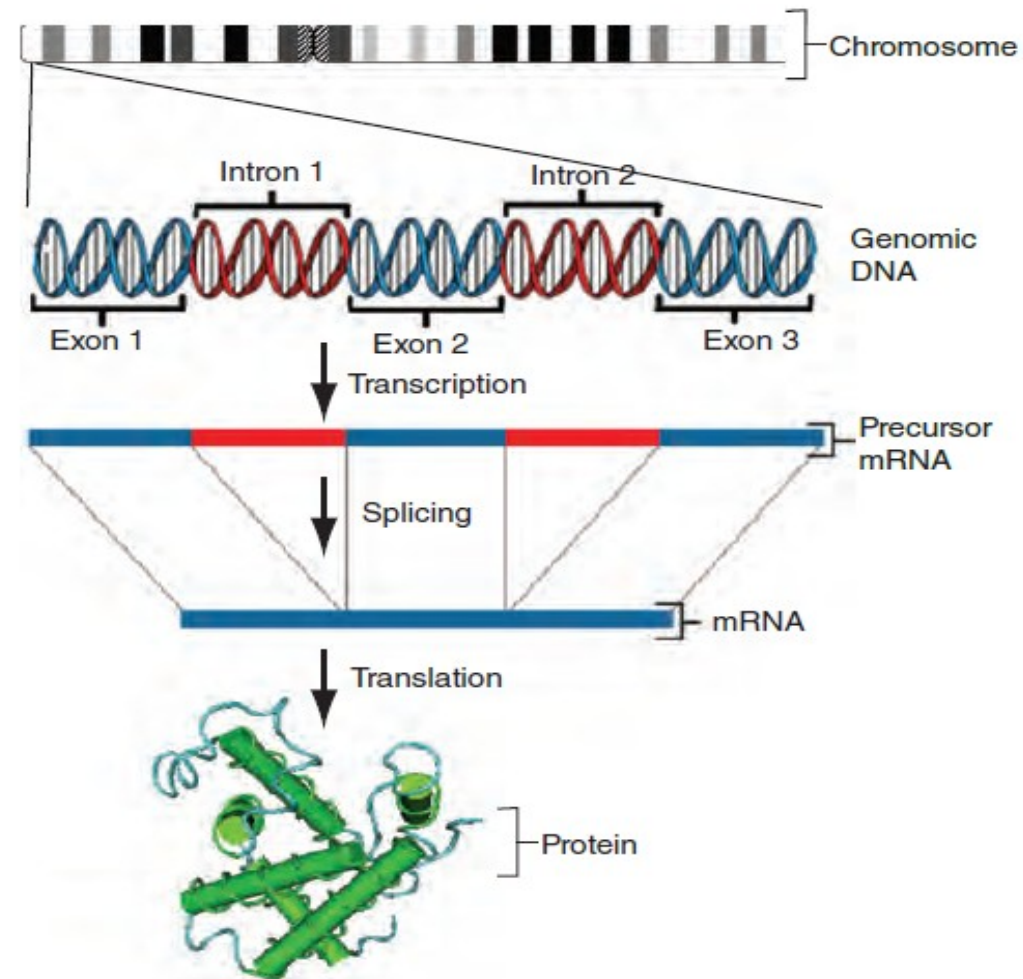
HET CENTRALE DOGMA EN BIO-INFORMATICA



DATA IN DE BIO-INFORMATICA

Verschillende typen data
zitten in verschillende
databases.

Wat is eigenlijk een
database?



Databases

Genome
dbVar

GenBank
SRA
dbGSS
dbHTGS
UniSTS
dbSNP

dbEST
UniGene
GEO profiles
GEO datasets

UniProt
Protein Data Bank
Conserved Domain Database

WAAR KAN JE NUCLEOTIDE SEQUENTIES VINDEN?

Denk eens terug aan vorig blok...

Waar hebben jullie de DNA en mRNA sequenties van de onderzochte genen vandaan gehaald?

Hoe hebben jullie dat gedaan?

WAAR KAN JE NUCLEOTIDE SEQUENTIES VINDEN?

- National Center for Biotechnology Information (NCBI):

<https://www.ncbi.nlm.nih.gov>



- European Bioinformatics Institute (EBI):

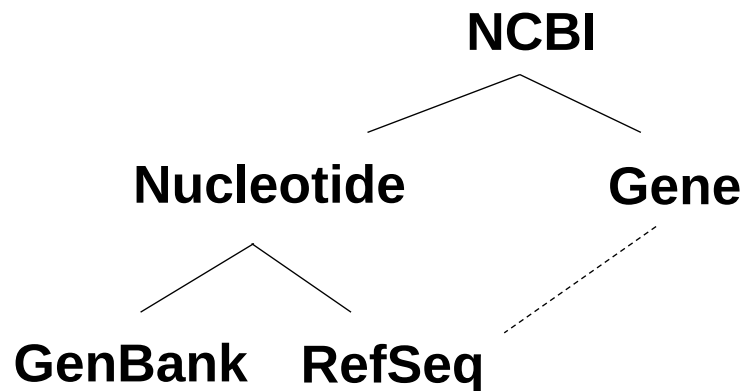
<http://www.ebi.ac.uk/>

- Database: European Nucleotide Archive (ENA)



NCBI – NUCLEOTIDE SEQUENTIE DATABASES

Nucleotide sequenties kan je in meerdere NCBI databases vinden



Nucleotide: Bevat alle bekende DNA en RNA sequenties

Genbank: Bevat alle publiekelijk beschikbare DNA en RNA sequenties. Door laboratoria beschikbaar gemaakt.

RefSeq: Bevat niet-redundante, gecureerde DNA en RNA sequenties. Oftewel referentie sequenties.

Gene: Centralizeerde gen-gerelateerde informatie over bekende gen sequenties (uit o.a. de RefSeq database)

DEMOTIJD!

Kijkje in de NCBI Nucleotide, Gene, GenBank en RefSeq databases.

Zoekvraag: Ik wil de humane DNA sequentie van het gen dat codeert voor het eiwit Presenilin 1 (een eiwit dat is geassocieerd met Alzheimer).

Welke database?

Wat voor query?

DATA EXPLOSIE

GenBank groeit
enorm snel:

Oktober 2011

132,067,413,372 nucleotiden

144,458,648 sequenties

November 2013

155,176,494,699 nucleotiden

168,335,396 sequenties

November 2015

202.237.081.559 nucleotiden

188.372.017 sequenties

November 2017

234.997.362.623 nucleotiden

201.663.568 sequenties

Oktober 2018

279.668.290.132 nucleotiden

209.656.636 sequenties

Oktober 2019

386.197.018.538 nucleotiden

216.763.706 sequenties

Oktober 2020

698.688.094.046 nucleotiden

219.055.207 sequenties

Oktober 2021

1.014.763.752.113 nucleotiden

233.642.893 sequenties

Oktober 2022

1.562.963.366.851 nucleotiden

240.539.282 sequenties

Oktober 2023

2.433.391.164.875 nucleotiden

247.777.761 sequenties

Oktober 2024

4.250.942.573.681 nucleotiden

252.347.664 sequenties

November 2025

5.676.067.778.413 nucleotiden

258.320.620 sequenties

ER WORDEN STEEDS MEER ORGANISMEN GESEQUENCED

Aantal verschillende organismen in NCBI databases:

	Nov 2013	Nov 2015	Okt 2018	Okt 2020	Nov 2021	Nov 2022	Nov 2023	Nov 2024
Archae	761	869	1066	1332	1464	1694	1883	2102
Bacteria	16483	19703	24884	29965	33177	35721	38239	40683
Eukarotya	381302	447559	542370	640982	670892	705696	733336	763484
Fungi	34358	39852	51595	62164	65680	69241	73889	78022
Metazoa	200999	241842	296121	361752	379962	399848	415365	431392
Viridiplantae	132793	151004	177065	196276	203562	213975	220573	229685
Viruses	3014	3511	5804	7604	8567	11784	11603	11959
All taxa	401627	471706	574188	679898	714120	754914	785094	818246

NCBI – ANDERE DATABASES

NCBI bevat veel meer databases dan de Nucleotide & Gene database, waaronder de volgende databases:

- Taxonomy
- PubMed
- OMIM

NCBI TAXONOMY IS ...

- Database met de voornaamste levende organismen (archaea, bacteriën, eukaryoten, virussen)
- Gecureerde classificatie en nomenclatuur van alle organismen in de sequentie databases
- Bruikbaar om eiwitten of genen van een soort te vinden

NCBI PUBMED IS ...

- Database met meer dan 34 miljoen citaties en samenvattingen van biomedische literatuur
- Bevat links naar de online journals en volledige inhoud

NCBI OMIM IS ...

- Online Mendelian Inheritance in Man
- Catalogus met humane genen en genetische aandoeningen
- Literatuur overzicht

DEMOTIJD!

Zoekvraag: Ik wil meer weten over het gen dat codeert voor het eiwit Presenilin 1 en de relatie met de ziekte van Alzheimer.

Hoe vind ik dit in OMIM?

AGENDA

- Introductie Genetica (OPSGEM)
- Project Genetica (OPSGEM)
- Biologische databases
- Zoeken naar informatie in databases

ESSENTIËLE ELEMENTEN VAN DATABASES

Wat is een database?

Elektronische gegevensbank voor data

Elke database kan z'n eigen format hebben, maar sommige data elementen zijn essentieel voor elke database:

1. Accession number of identifier
2. Naam van de data eigenaar
3. Literatuurreferenties
4. Datum waarop de data beschikbaar is gemaakt
5. De data zelf

DATA + ANNOTATIES

- **Biologische annotatie:**
Ondersteunende informatie over structuur en functie van DNA, RNA of eiwitten.
- **Bibliografische annotatie**
(metadata)

Annotatie helpt bij het interpreteren van de data.

https://www.ncbi.nlm.nih.gov/nuccore/NM_005368

Homo sapiens myoglobin (MB), transcript variant 1, mRNA

NCBI Reference Sequence: NM_005368.3

[FASTA](#) [Graphics](#)

[Go to:](#) ☒

LOCUS	NM_005368	1070 bp	mRNA	linear	PRI 05-JUL-2021
DEFINITION	Homo sapiens myoglobin (MB), transcript variant 1, mRNA.				
ACCESSION	NM_005368				
VERSION	NM_005368.3				
KEYWORDS	RefSeq; MANE Select.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1070)				
AUTHORS	Al Abdulmonem W, Aljohani ASM, Alhumaydhi FA, Mousa AHM and Rasheed Z.				
TITLE	Protective Potential of Uric Acid, Folic Acid, Glutathione and Ascorbic Acid Against the Formation of Toxic Met-Myoglobin				
JOURNAL	Protein Pept Lett 28 (3), 282-289 (2021)				
PUBMED	32957872				
REMARK	GeneRIF: Protective Potential of Uric Acid, Folic Acid, Glutathione and Ascorbic Acid Against the Formation of Toxic Met-Myoglobin.				
REFERENCE	2 (bases 1 to 1070)				
AUTHORS	Kottwitz J, Bruno KA, Berg J, Salomon GR, Fairweather D, Elhassan M, Baltensperger N, Kissel CK, Lovrinovic M, Baltensweiler A, Schmied C, Templin C, Lima JAC, Landmesser U, Luscher TF, Manka R and Heidecker B.				
TITLE	Myoglobin for Detection of High-Risk Patients with Acute Myocarditis				
JOURNAL	J Cardiovasc Transl Res 13 (5), 853-863 (2020)				
PUBMED	32006209				

DATA + ANNOTATIES

- Annotaties worden automatisch of handmatig toegevoegd
- Error rate hangt af van methodiek en menselijke interpretatie
- Error rate wordt groter naarmate de database groter wordt

https://www.ncbi.nlm.nih.gov/nuccore/NM_005368

Homo sapiens myoglobin (MB), transcript variant 1, mRNA

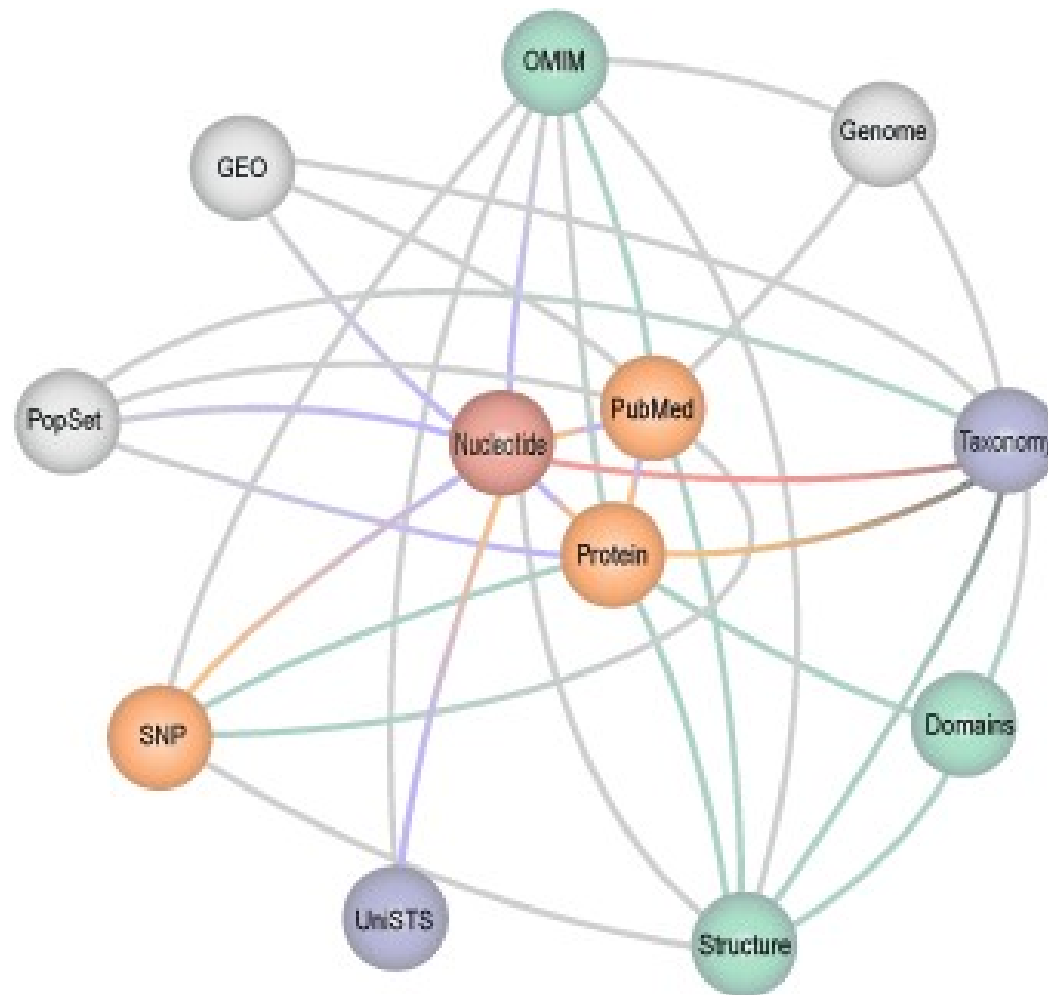
NCBI Reference Sequence: NM_005368.3

[FASTA](#) [Graphics](#)

[Go to:](#) ☐

LOCUS	NM_005368	1070 bp	mRNA	linear	PRI 05-JUL-2021
DEFINITION	Homo sapiens myoglobin (MB), transcript variant 1, mRNA.				
ACCESSION	NM_005368				
VERSION	NM_005368.3				
KEYWORDS	RefSeq; MANE Select.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1070)				
AUTHORS	Al Abdulmonem W, Aljohani ASM, Alhumaydhi FA, Mousa AHM and Rasheed Z.				
TITLE	Protective Potential of Uric Acid, Folic Acid, Glutathione and Ascorbic Acid Against the Formation of Toxic Met-Myoglobin				
JOURNAL	Protein Pept Lett 28 (3), 282-289 (2021)				
PUBMED	32957872				
REMARK	GeneRIF: Protective Potential of Uric Acid, Folic Acid, Glutathione and Ascorbic Acid Against the Formation of Toxic Met-Myoglobin.				
REFERENCE	2 (bases 1 to 1070)				
AUTHORS	Kottwitz J, Bruno KA, Berg J, Salomon GR, Fairweather D, Elhassan M, Baltensperger N, Kissel CK, Lovrinovic M, Baltensweiler A, Schmied C, Templin C, Lima JAC, Landmesser U, Luscher TF, Manka R and Heidecker B.				
TITLE	Myoglobin for Detection of High-Risk Patients with Acute Myocarditis				
JOURNAL	J Cardiovasc Transl Res 13 (5), 853-863 (2020)				
PUBMED	32006209				

DATABASES ZIJN GELINKT DOOR CROSS REFERENTIES



WELKE VRAGEN STEL JE AAN DATABASES?

Type vragen:

- Opzoeken (lookup)
- Vergelijken (compare)
- Voorspellen (predict)

WELKE VRAGEN STEL JE?

Opzoeken (lookup):

- Is het gen van mijn eiwit bekend?
- Op welk chromosoom zit het gen?
- Wat voor sequentie patronen zijn er aanwezig op mijn eiwit sequentie?
- Zijn er mutaties bekend die deze ziekte veroorzaken?
- Tot welke familie behoort mijn eiwit? Wat is er bekend over deze familie?

WELKE VRAGEN STEL JE?

Vergelijken (compare):

- Zijn er eiwit sequenties in de database die lijken op het eiwit waar ik geïnteresseerd in ben?
- Hoe kan ik leden van deze eiwitfamilie optimaal met elkaar alignen?
- Zijn deze twee sequenties vergelijkbaar?

WELKE VRAGEN STEL JE?

Voorspellen (predict):

- Kan ik de active site van dit enzym voorspellen?
- Waarom zijn deze patiënten ziek?
- Kan ik een 3D model maken van mijn eiwit?
- Kan ik een (beter) medicijn voorspellen voor deze target?
- Kan ik de thermostabiliteit van mijn eiwit verbeteren?
- Kan ik voorspellen welke genen er op dit genoom zitten?

ZOEKEN IN DE NCBI DATABASES: ENTREZ & BLAST

The screenshot shows the NCBI homepage. At the top, there's a navigation bar with 'NCBI', 'Resources', 'How To', and a 'Sign in to NCBI' link. Below this is a search bar with a dropdown menu set to 'All Databases' and a 'Search' button. A banner for 'UNITE' (A new NIH initiative to end structural racism and achieve racial equity in the biomedical research enterprise) is visible, with a 'LEARN MORE' link. On the left, a 'Resource List (A-Z)' sidebar lists various categories like 'All Resources', 'Chemicals & Bioassays', 'Data & Software', 'DNA & RNA', 'Domains & Structures', 'Genes & Expression', 'Genetics & Medicine', 'Genomes & Maps', 'Homology', 'Literature', 'Proteins', 'Sequence Analysis', 'Taxonomy', 'Training & Tutorials', and 'Variation'. The main content area is titled 'Welcome to NCBI' and includes a description of the center's mission. It features six main action buttons: 'Submit' (Deposit data or manuscripts into NCBI databases), 'Download' (Transfer NCBI data to your computer), 'Learn' (Find help documents, attend a class or watch a tutorial), 'Develop' (Use NCBI APIs and code libraries to build applications), 'Analyze' (Identify an NCBI tool for your data analysis task), and 'Research' (Explore NCBI research and collaborative projects). On the right, there's a 'Popular Resources' section listing 'PubMed', 'Bookshelf', 'PubMed Central', 'BLAST', 'Nucleotide', 'Genome', 'SNP', 'Gene', 'Protein', and 'PubChem'. Below that, an 'NCBI News & Blog' section mentions that three outdated browsers (1000 Genomes, dbGaP Data, and Get-REM) will retire in April 2022 and that the Genome Data Viewer (GDV) is now available.

- **Entrez:**
Zoek in databases
door middel van *keywords*
- **BLAST:**
Zoek in databases met behulp van
DNA, RNA of eiwit *sequenties*

ZOEKEN IN DATABASES

- Één of meerdere zoektermen per keer?
- Advanced search
- Boolean operators (AND/OR/NOT)

ZOEKEN IN DATABASES: BOOLEAN OPERATORS



* Boolean operators (AND/OR/NOT) kunnen ook in Python gebruikt worden

ADVANCED SEARCH IN ENTREZ

Voorbeeld:

Ik wil alle referentie mRNA sequenties die coderen voor het humane hemoglobine eiwit

Wat zijn belangrijke keywords?

Query:

hemoglobin[Title] **AND** homo sapiens[Organism] **AND**
biomol mrna[Properties] **AND** refseq [Filter]

Advanced search builder van de NCBI Nucleotide database:

<https://www.ncbi.nlm.nih.gov/nuccore/advanced>

SAMENVATTING

- NCBI & EBI bevatten veel databases met biologische data, waaronder sequentie databases (e.g. Nucleotide, Gene, GeneBank, RefSeq)
- Sequenzen wordt steeds makkelijker en goedkoper → aantal sequenties in deze databases neemt snel toe
- Belangrijke eigenschappen van een database:
 - Accession number/identificer
 - Biologische annotatie
 - Bibliografische annotatie

HUISWERK OPDRACHTEN

1. Je wilt de DNA en eiwit sequentie hebben van het gen en eiwit genaamd 'gag-pol' uit het HIV-1 virus (Human immunodeficiency virus 1). Hoe krijg je de DNA en eiwit sequentie?
2. Je wilt de mRNA sequentie hebben van humaan MLH1 eiwit. Hoe krijg je dit voor elkaar?
3. Je bent benieuwd naar de genen die zijn gerelateerd aan diabetes mellitus type 2. Welke genen kan je vinden?

NASLAGWERK

- Pevsner, *Bioinformatics and functional genomics*, third edition:
Hoofdstuk 2, pagina 19-26, 31-36, 38-42, 49-53 (inclusief Box 2.1 & Box 2.2)
- Campbell, *Biology*, twelfth edition: Paragraaf 13.1, 13.2, 14.1, 14.4, 15.2, 15.3



VERANTWOORDING

- In deze uitgave is géén auteursrechtelijk beschermd werk opgenomen
- Alle teksten © Ingrid Paffen en Emma Vos / HAN tenzij expliciet externe bronnen zijn aangegeven
- Screenshots op basis van eigen werk auteur en/of vernoemde sites
- Eventuele images zijn opgenomen met vermelding van bron