Op zoek naar een gen in een databank

Tijdens het DNA-lab 'Lees de taal van de tumor' hebben jullie drie genen onderzocht.

Welke genen hebben jullie onderzocht? Wat was de functie van deze genen?

	Naam	Functie
Gen 1		
Gen 2		
Gen 3		

Eén van deze genen gaan jullie nu verder onderzoeken, met behulp van databanken. Een databank is een verzameling gegevens die digitaal zijn opgeslagen. Onderzoekers over de hele wereld kunnen met behulp van een computer en internet deze gegevens raadplegen en zelf nieuwe gegevens toevoegen.

Deze databanken zijn voor iedereen toegankelijk, dus ook voor jullie. Je gaat nu zelf aan de slag met databanken om meer over één van de onderzochte genen te weten te komen. Je krijgt van het bedrijf 'DNA-lab support' de opdracht een rapport te maken. Het rapport vind je op de laatste bladzijde. Aan de hand van de vragen en opdrachten in de handleiding kun je dit rapport invullen.



Opdracht

Maak een rapport over het gen p53.

Om specifiekere medicijnen tegen kanker te kunnen ontwikkelen, is het belangrijk dat we over informatie beschikken van de genen die gemuteerd zijn. Het gen p53 is één van de meest belangrijke genen in kankeronderzoek. Zoek daarom het volgende uit over het gen p53:

- Op welke locatie in het DNA ligt het p53 gen?
- Hoe groot is het p53 gen?
- Welke genen liggen er naast p53?

Achtergrondinformatie

- De officiële aanduiding van het p53 gen is 'Tp53'.
- Databanken zijn altijd in het Engels, zodat ze begrijpelijk zijn voor onderzoekers over de hele wereld. Een databank heet dan ook 'database'.
- Tp53 komt niet alleen bij mensen voor, maar ook bij andere organismen. In databanken wordt het menselijke Tp53 gen aangeduid met 'human' ('humaan' in het Nederlands).
- Al het DNA van één organisme noemt men het 'genoom' (samenvoeging van 'gen' en 'chromosoom'), de Engelse vertaling is 'genome'.

Handleiding

Er bestaan verschillende databanken. Aan deze databanken wordt bijna iedere dag nieuwe informatie toegevoegd. Ga naar <u>www.ebi.ac.uk/genomes</u>

EMBL-EBI			Services Research Training Industry About us		
Genom	es Pages	- At the	e EBI		
* Complete genomes	Databases + Ruclestide + The	Durapean Nucleokke Archiv	n > Conjulte Genomes		
= Archaea					
* Archaeal virus	Access to Completed	Genomes	GENOME		
* Eukaryota * Organelle	The first completed genome sequences of as many geno including archaes. Factoria s	s from <u>viruses</u> , <u>phages</u> a mes as possible combi	and organeties were deposited into the EMBL Database in the early 1980's. Since their, molecular biology's shift to obtain the complete read with major developments in sequencing technology resulted in hundreds of complete genome sequences being added to the database, it can be not access to a large number of committed menomes being validable to decreme the failabase.		
+ Phage	Whole Genome Shotoun Ser	suences (MGS)	e prime and a second of a sufficient of second s		
= Plasmid	Matter and a second second second				
* Viroid	wearons asing wagle devou	ie suodinu gata ale dise	o w gan a raige amount of genome coverage for an organism. Proc usia for a growing number of organisms are being submitted to DUBUEMBLUGENBARK.		
* Virus	More information about WOS	projects_			
 Links 					
WOS info	Last 40 Genome Entries				
EnsemblGenomes	Date		Description		
Genome Reviews	10-JUL-2013	A8775659.1	Sapovirus HuNagoyaNGV-1/2012/JFN		
Integra (proteomes)	10-JUL-2013	AD017729.1	Persinenon viroki 2		
Fasta33 Server	10-JUL-2013	CP003347.1	Mycobacterium yangonerose 05-1390		
Ensemul	10-JUL-2013	CP002646.1	Acinetobacter baunanna BJAB07104		
	10-JUL-2013	<u>CP003847.1</u>	Acinetobacter baumannii BJA80715		
	10-JUL-2013	CP003649.1	Acinetobacter baumannii BJAB0868		
	10-JUL-2013	CP004845.1	Atteromonis macleodii str. 'English Channel 615'		
	10-JUL-2013	CP004848.1	Ateromons macleodi str. 'Aegean Sea MED64'		
	10-JUL-2013	CP004849.1	Ateronome macleodi str. tenian Sea U41		
	10-JUL-2013	GP004851.1	Atteromonas macleodii str. 'tonian Sea UP'		
	10-JUL-2013	CP004852.1	Ateronome macleodi str. tonian Sea U8'		
	10-JUL-2013	CP004853.1	Aberomonse macleodii str. tenian Sea UMP		
	10-JUL-2013	CP004855.1	Alteronoma macleodi str. 'anian Sea UM40'		
	10-JUL-2013	CP005484.1	Lactobacilius rhamnosus LOCK900		
	10-JUL-2013	CP005485.1	Lactobacilius rhannosus LOCK908		
	10-JUL-2013	CP005486.1	Lactobacilius casei LOCK919		
	10-JUL-2013	H0259069.1	Abisera fylioides mitochandrion		
	10-JUL-2013	30417196.1	Helix aspersa from Chile mitochondrion		
	10-JUL-2013	KC778721.1	Turnip rosette virus isolate TRoV-2		
	08-JUL-2013	KF148055.1	Salmonella phage Jersey		
	08-JUL-2013	KE246569.1	Fur seal faeces associated circular DNA virus isolate as50		
	07-11-2013	AR626608.1	Paintonia stuana 85530		

Je komt op een pagina met meer informatie over de genomen die recent zijn toegevoegd aan databanken. Je ziet een lijst met data en daarachter welk genoom(deel) er toegevoegd is. Deze lijst verandert met de dag. De lijst in de handleiding zal dus verschillen met wat je zelf ziet op je beeldscherm.

Vraag 1

Op welke datum is het nieuwste genoom in de bank geplaatst? Zoek op van welk organisme dit is.

Datum:

Organisme:

Op de site van het European Bioinformatics Institute (EBI) staat een overzicht van databanken. Ga naar de homepage van het EBI: <u>www.ebi.ac.uk</u>



Klik vervolgens op 'Services' (in de screenshot op de vorige pagina rood omcirkeld).



Je ziet nu een overzicht van wat voor typen databanken er zijn.

Vraag 2

In databanken is over veel onderwerpen informatie te vinden. Noem 3 onderwerpen die je herkent.
1
2
3
Noem een onderwerp dat je nog niet kent, en zoek op wat het inhoudt.
Onderwerp:
Betekenis:

Klik nu op 'DNA & RNA'.



Je ziet een overzicht van verschillende databanken waarmee je meer informatie over het genoom kunt opzoeken. Databanken verwijzen naar elkaar!

Eén van de meest gebruikte databanken heet 'Ensembl', hierin kun je genomen van verschillende organismen opzoeken. Klik op '**Ensembl**' om naar de databank Ensembl te gaan.



In de databank Ensembl is informatie te vinden over de mens, maar ook over andere organismen. Bijvoorbeeld de muis, zebravis, rat, C. elegans (worm), fruitvlieg of mug. Deze informatie kunnen onderzoekers gebruiken om te kijken of bepaalde genen in verschillende organismen van elkaar verschillen of juist sterk op elkaar lijken. Een gen dat in zowel de mens, de muis als de zebravis precies hetzelfde eruit ziet, zal waarschijnlijk een belangrijke functie hebben!

Klik op '**Human'** om naar het menselijke genoom te gaan. Je krijgt de overzichtspagina van de humane databank te zien.



Klik op 'View karyotype' (hierboven rood omcirkeld). In een karyotype staan de chromosomen van een organisme afgebeeld, geordend op grootte.

	7/BLAT BlaMart Tools Downloads Help & Documentat	ion Blog Mirars	Login/Register 🗑 - Search Human Q
Location-based displays Whole genome Chromosome summary Region overview Region in detail	Location		
	whole genome 😡		
- Alignments (mage)	\$ \$ < @		
Genetic Version System Genetic Version System Genetic Version Genetic Versio Genetic Versio Genetic Version Genetic Version G			
	Crick on the image above to jump to a chromosome, or ch	ruk anu uray to sereti a region	
	Summary		
	Assembly: GI	RCh37.p11, Feb 2009	
	Base Pairs: 32	.ər 123.950.079	
	Golden Path Length: 3,	101,804,739	
	Genebuild by: Er	sembl	
	Genebuild method: Fu	ll genebuild	
	Genebuild started: Ju	2010	
	Genebuild released: Ap Genebuild last undated/extched:	r 2011 - 2013	
	Gene counts (Primary assembly)	1.2013	

Je krijgt nu op je scherm een overzicht van het humane genoom te zien, dus een overzicht van alle verschillende menselijke chromosomen. Onder het karyotype staat een samenvatting van het humane genoom.

Vraag 3

Uit hoeveel baseparen bestaat het humane genoom?

.....

.....

Vraag 4

Hoeveel eiwitcoderende genen bevat het humane genoom?

Klik op chromosoom **17**. Er verschijnt een pop up. Klik op '**Chromosome summary**' om naar een uitgebreid overzicht van het chromosoom te gaan.

CEnsembl 🗤	ST/BLAT BioMart Tools Downloads Help & Documentation Blog Mirrors	Login/Register
Human (GRCh37) V Location: 1	772 917 574 71 001 574	
Location-based displays	Chromosome 17: 72,943,674-73,043,674	
- Regisin convolver - Regisin in detail Comparative Genomics - Alagoments (insag) (64) - Alagoments (insag) (64) - Region Comparison (78) - Systemy (16) So Cenetic Variation - Resequencing (20) - Linkage Data - Markers Mother genome browsers	Construction Assembly exceptions	
- UCSC - NCBI Vega	Chromosome summary 0	
Configure this page	수 최 < Change Chromosome: 17	O 0
💦 Add your data	Dhronosome Protein Coding Genes Non Coding Genes Pseudogenes X DC 17 Chromosome Statistics	
🖞 Export data	□r ^{10,3}	5,210
Bookmark this page	r55-2 Protein cading gene count 1,205	
-	Non coding gene count 1,179	
Share this page	Pseudogene count 531	
	Variations 1,618	869

Nu is het chromosoom weergegeven, de (bekende) genen in het chromosoom zijn aangegeven en je kunt chromosoom statistieken zien.

Vraag 5

Hoe groot is chromosoom 17?

..... baseparen

Vraag 6

Hoeveel eiwitcoderende genen van chromosoom 17 zijn bekend?

.....

In je biologieboek staat de volgende informatie over chromosomen:

Een centromeer deelt een chromosoom in twee armen: een lange (q) en een korte (p). Een chromosoom bestaat uit dubbelstrengs DNA. Een gen ligt niet op allebei de DNA-strengen: Slechts één streng van het DNA codeert voor het gen (en dus voor het eiwit dat wordt gecodeerd door dit gen).



Typ nu **'Tp53'** in de zoekbalk rechts bovenin je scherm, om te zoeken naar het gen p53 in de Ensembl databank. Klik vervolgens op '**Human'** en dan op '**gene**' om het menselijke p53 gen te vinden.

ClEnsembl 🚥	ST/BLAT BioMart Tools Dewnload		🗑 - Search Human	Login/Register
Human (GRCh37) 🔻				i i
Search Ensembl	Result in Detail			
🔅 Configure this page	15 Genes match your query (Tp53) in Human			
💦 Add your data	Showing results 1.10			
🖞 Export data	1 2 Next s			
🖗 Bookmark this page	TP53			
< Share this page	Description	tumor protein p53 [Source:HGNC Symbol;Acc:11998] [Type: protein coding Ensembl/Havana merge]		
	Gene ID	ENSG0000141510		
	Location	<u>17:7565097-7590856:-1</u>		
	Variations	Variation Table		
	Source	872		
	TP53TG3D			
	Description	TP53 target 3D [Source:HGNC Symbol;Acc:44657] [Type: protein coding Ensembl/Havana merge]		
	Gene ID	ENSG0000205456		
	Location	<u>16:32264040-32267240:1</u>		
	Variations	Variation Table		
	Source	972		
	TRIAP1P1			
	Description	TP53 regulated inhibitor of apoptosis 1 pseudogene 1 [Source: HGNC Symbol;Acc: 31659] [Type: pseudogene Havana]		
	Gene ID	ENSG0000237956		
	Location	10.27667067-27667291-1		
	Variations	Variation Table		
	Source	072		
	TRIAP1P1			
	Description	TP53 regulated inhibitor of apoptosis 1 pseudogene 1 [Source:HGNC Symbol;Acc:31659] [Type: pseudogene Projected Havana]		
	Gene ID	ENSG00000263055		
	Location	HSCHR10_1_CTG2:27714387-27714591:-1		

Klik op de link die het best bij de zoekopdracht past (**TP53**). Net zoals een zoekopdracht bij Google, komt de best passende link bovenaan in de lijst te staan. Klik bij '**Location**' op de link naar het chromosoom.

CEnsembl 🗤	ST/BLAT BioMart T	iools Downloads Help & Documentation	Blog Mirrors		<u>.</u>	Login/Register Search Human Q
Human (SRUhsz) Location: 12 Location-based displays Whole genome Chromesome summary	Chromosome	Gene 1753 17: 7,565,097-7,590,856				
- Region overview						
Region in detail Comparative Genomics Alignments (image) (64) Alignments (itex) (64) Alignments (itex) (64) Region Comparison (78) Systemy (16) Genetic Variation Resequencing (20) Linkage Data Markers	Assembly exceptions Chr. 17 Assembly exceptions	1333 3322 (333 0)2 (4999, Mr OH HSN17, Mr OH HSN17, Mr OH	р112 а1 Нозиг, мл сн нажа	2 012 012 012 012 00 10700 - 10700 00 10700 - 1080, MO 1080, MO 1080, MO 1080, MO 1080, 12, 00 1080, 1	192 (2115) (122 (233) 7,3 (201)	242 943 961 963 19,970 HOR12,2,074 HOR12,1,074 HOR1,970 HOR2, HOR2,970 HOR2,
B' Other genome browsers UCSC - NCBI Vega	Region in deta	ail O		HOILE	EJNTOH	
😒 Configure this page	� < ⊞ ല					
Add your data de Export data Re Bookmark this page	Chromosome bands Contigs Genes (Merged E	7.10 Mb 7.20 Mb Act2000779 > Accossod 34 >	7,30 M6 7.40 M6 AC113189 11 >	100 M6 7.50 Mb 013 1 < AC007422 11 CAC007422 11	7,70 Mb 7,80 Mb Accessito 23 > Accessite 22 >	Forward strand == 7,50 Mb 8,00 Mb AC1204926 >
< Share this page		K edicali 1 ≤ INC2AS > Doc211	**************************************	<pre>\cynue.com = Accorectly = Years =</pre>	Чедер>, Чедино, Чедина, тренев - констрание са - Чедино, Чедино, Чедино, 12 - Чедино, Чедино, Чедино, 12 - Чедино, Чедино, Чедино, чедино,	 Констала - Констала - Констала
				~< AC113189.5		
		Ensembl Homo sapiens version 72.37 (GRCh37	7.30 Mb 7.40 Mb Chromosome 17: 7,077,977 - 8,077,9	7.50 Mb 7.60 Mb	7.70 Mb 7.80 Mb	7.90 Mb 8.00 Mb
	Gene Legend	protein coding merged Ensembl/Havana pseudogene		processed RNA gene	transcript	
		5007 7500055	C		(m)	

Met behulp van een rood kader wordt aangegeven waar het gen zich bevindt op het chromosoom. Bovenaan de pagina staat achter het chromosoomnummer in cijfers de start- en stoppositie van het gen weergegeven, aan de hand van het aantal nucleotiden.

Vraag 7

Op welk chromosoom ligt het p53 gen?

Vraag 8

Op welke chromosoomarm ligt het p53 gen?

.....

.....

Vraag 9

Uit hoeveel baseparen is p53 opgebouwd? Schrijf op hoe je dit berekend hebt.

Aantal baseparen:

Berekening:

.....

Als je naar beneden **scrollt**, vind je meer gedetailleerde informatie over de genen op chromosoom 17. Zichtbaar is het p53 gen en de 'buurgenen' die op hetzelfde chromosoom liggen. In het onderste kader is het chromosoom met een horizontale blauwe balk weergegeven. In rood zijn de RNA transcripten van de genen weergegeven.



Zoom nu helemaal uit, zoals op het onderstaande plaatje:



Scroll naar het chromosoom in het onderste kader.



In rood zijn de RNA transcripten van de genen weergegeven. Alles wat boven het chromosoom staat wordt afgelezen van de ene DNA-streng, en alles wat onder het chromosoom staat wordt afgelezen van de andere DNA-streng.

Als je zoekt naar het p53 RNA, vind je negen p53 RNA's (dit houdt in dat van het p53 gen negen verschillende RNA transcripten gemaakt kunnen worden). Ook zie je dat er links en rechts van p53 nog heel veel andere genen op het chromosoom liggen.

Klik met de linkermuisknop op een willekeurig buurgen (met een rode kleur). Er verschijnt dan een pop-up met informatie over het gen. Bij 'Location' vind je de positie van het gen.

Vraag 10

Zoek op welk gen er aan de linkerkant (op dezelfde streng) zo dicht mogelijk tegen p53 aan ligt. Welk gen is dit?

.....

Vraag 11

Hoe groot is de ruimte tussen het p53 gen en zijn linker buurgen? Gebruik voor je berekening de startpositie van p53, die je ook in vraag 9 hebt gebruikt.

.....

Vraag 12

Zoek nu op welk gen er aan de rechterkant (op dezelfde streng) zo dicht mogelijk tegen p53 aan ligt, en hoe groot de ruimte is tussen het p53 gen en zijn rechter buurgen.

Buurgen:

Ruimte:

Je hebt nu alle informatie gevonden die je nodig hebt om het rapport op de volgende bladzijde in te vullen.

Je hebt waarschijnlijk nog nooit eerder op deze manier met een databank gewerkt. Wat vond je bijzonder om te ontdekken? Je mag meerdere opties omcirkelen.

- 1. Dat aan de databanken bijna elke dag nieuwe informatie wordt toegevoegd.
- 2. Dat onderzoekers voor hun werk deze databanken gebruiken.
- De grootte van het humane genoom.
 Het aantal genen in het humane genoom dat codeert voor een eiwit.
- 5. Dat een gen niet van beide DNA strengen afgelezen wordt, maar slechts vanaf één streng.
- 6. Het aantal genen dat op chromosoom 17 ligt.
- 7. Dat er tussen genen stukken DNA zitten die niet coderen voor een eiwit.
- 8. Benoem zelf een ontdekking:

.....



Onderzoeksrapport

Datum:

Ik heb het menselijke p53 gen onderzocht in de Ensemble databank en ik ben tot de volgende conclusies gekomen:

• De locatie van het p53 gen is op:

chromosoom,

op de arm,

van baseparen.

• Het p53 gen:

is baseparen lang.

• Op hetzelfde chromosoom:

ligt aan de linkerkant van p53 ook het	gen,
de ruimte tussen p53 en dit gen is	baseparen.
Aan de rechterkant van p53 ligt het	gen,
de ruimte tussen p53 en dit gen is	baseparen.

Met vriendelijke groet,

.....