joanne@msl.ubc.ca

Bioinformatics

Common Tools & Tricks of the Trade

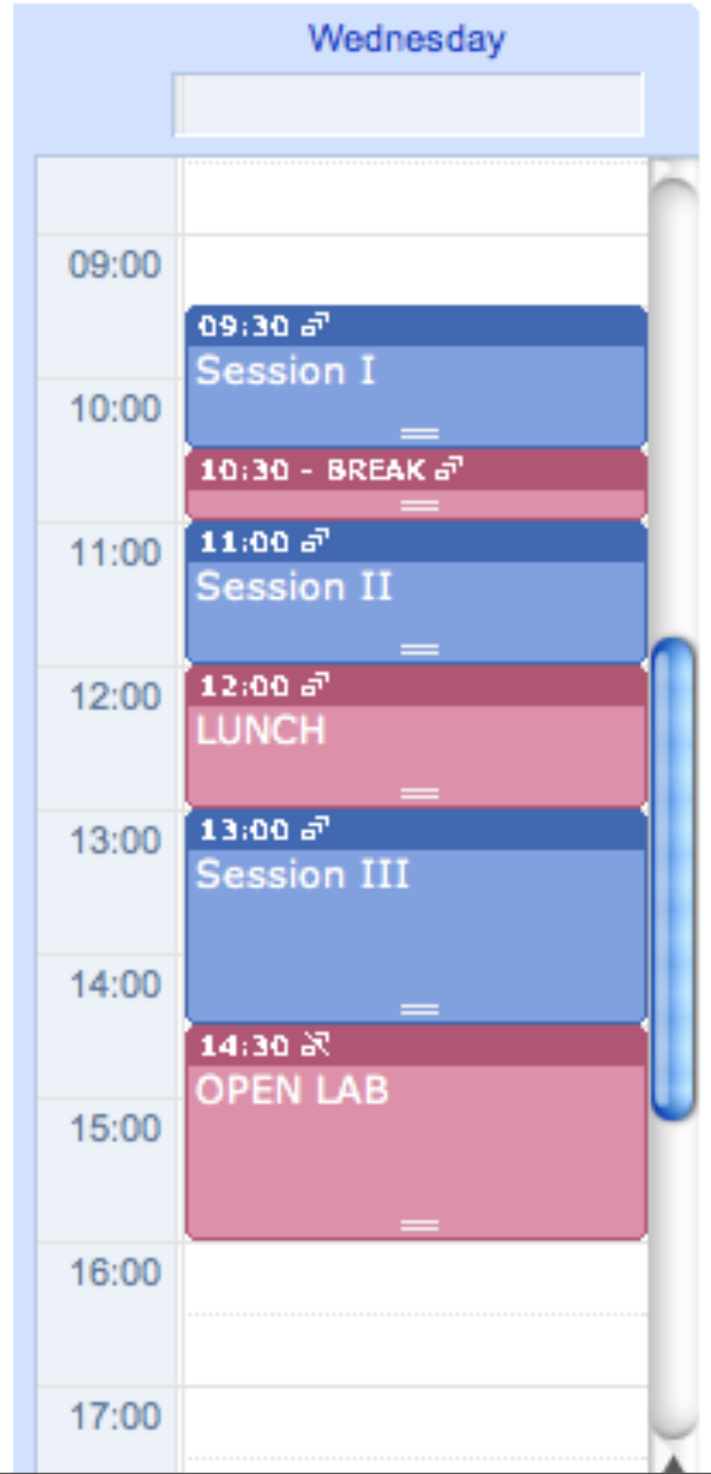
Welcome to Day 3
bioteach.ubc.ca/bioinfo2010

Workshop Schedule

- Laptops, available here for your use 9am - 4:30pm
- wireless login

  mslguest

  4myguest




Thank You
Thank You
Enjoy

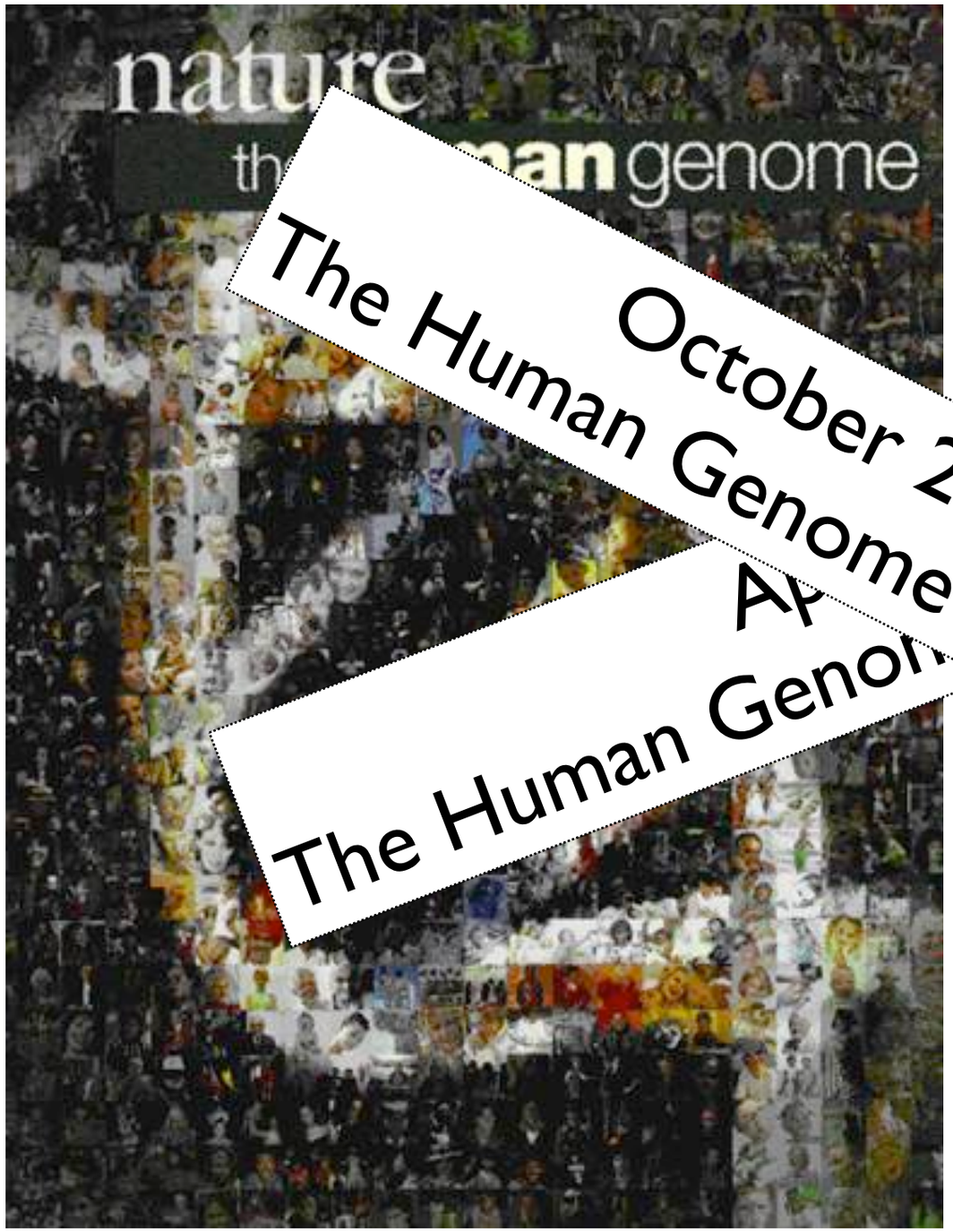
Today's Topics

- **Multiple Sequence Alignments** (refer to content from DAY2 notes)
- **Genome Browsers**, Accessing Genome Annotations.
- **PRACTICAL EXERCISES**, three different views of the BRCA1 gene
- **Pathway Resources** for Systems Biology
- **Bioinformatics Links Directory**, Conducting Research on the Web

Genome Browsers

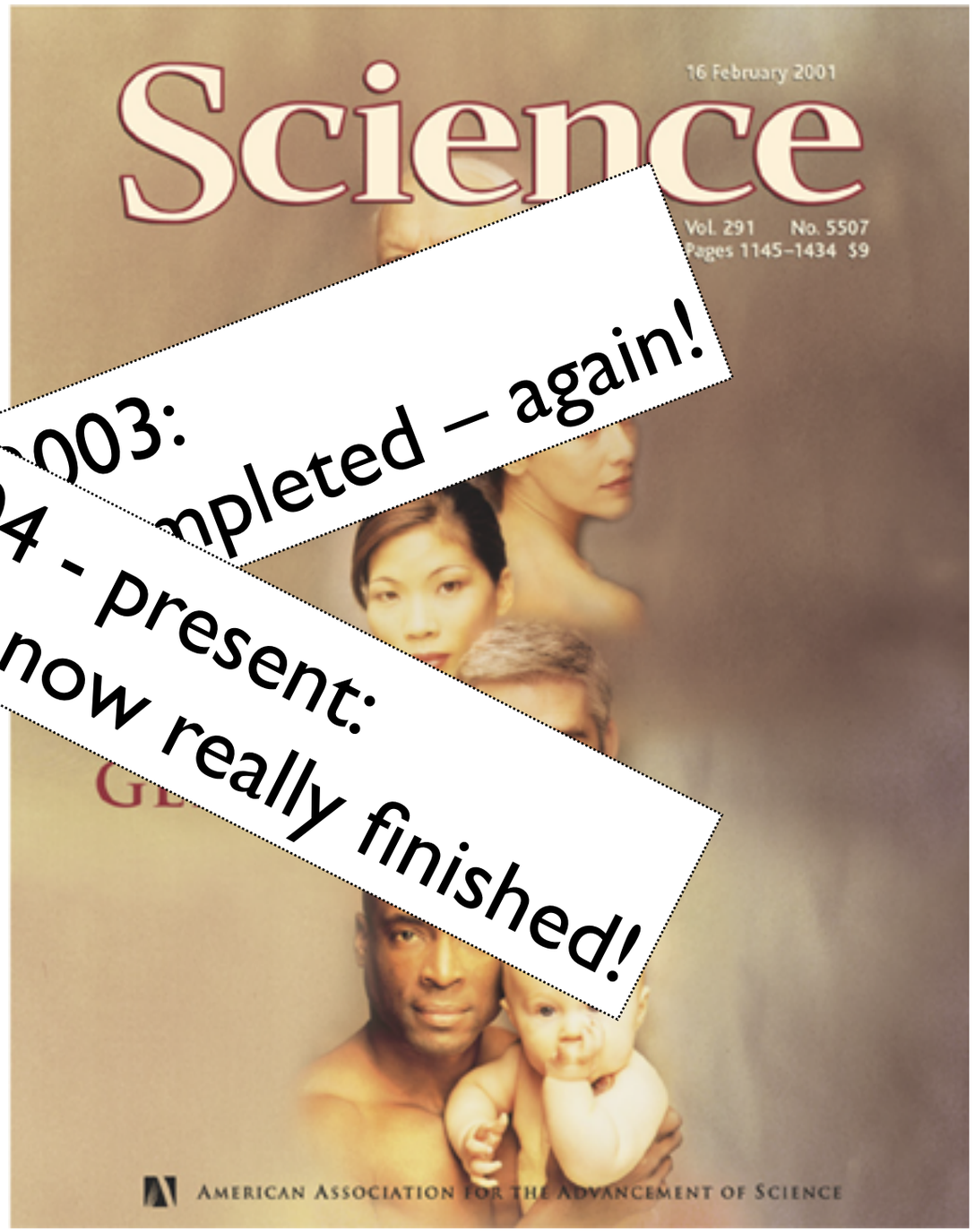
Accessing Genome Annotations & PRACTICAL EXERCISE: Three Different Views of the BRCA1 Gene

The Human Genome Project



October 2004 - present:
The Human Genome is now really finished!

003:
mpleted – again!
A
Genom



Public HGP

Celera Genomics

February 2001: Completion of the Draft Human Genome

In the Genome Race, the Sequel Is Personal

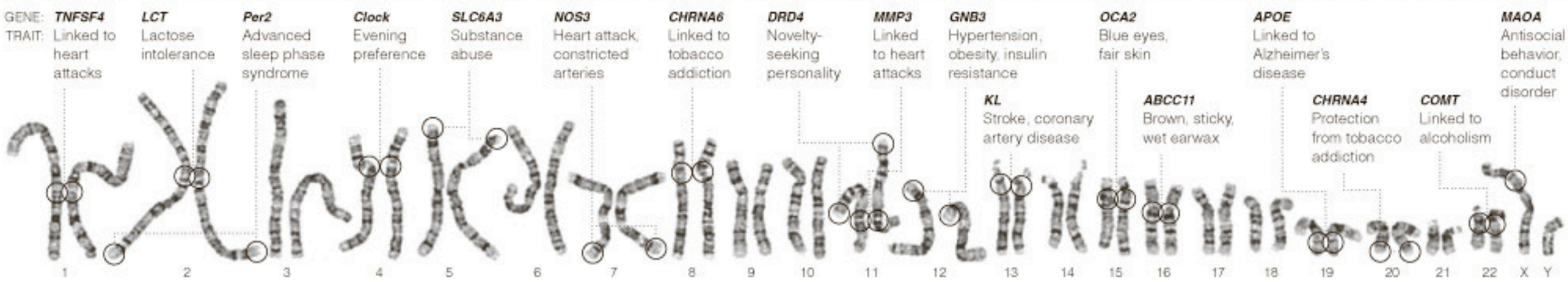
Thor Swift for The New York Times

A team led by J. Craig Venter, above, has finished the first mapping of a full, or diploid, genome, made up of DNA inherited from both parents. The genome is Dr. Venter's own.

By NICHOLAS WADE

The New York Times September 3, 2007

DECODING HIMSELF A team led by J. Craig Venter, above, has finished the first mapping of a full, or diploid, genome, made up of DNA inherited from both parents. The genome is Dr. Venter's own.



Sources: PLoS Biology; GeneCards CHROMOSOMES *Genes highlighted above are linked to specific traits but do not mean the conditions are inevitable.* THE NEW YORK TIMES

Stanley Park
Stanley Park Shuttle
Stanley Park Shuttle (seasonal)
STANLEY PARK LOOP
19
Lost Lagoon
Deadmans Island
Devonian Harbour Park
Coal Harbour
0
500
METRES
N
To North Vancouver and Lonsdale Quay →
SEABUS
SEABUS TERMINAL
WALKWAY
WATERFRONT
BURRARD
GRANVILLE
STADIUM - CHINATOWN
MAIN STREET - SCIENCE WORLD
LAGOON DR
ALBERNI
CHILCO
GILFORD
PENDRELL
DENMAN
BIDWELL
CARDERO
BROUGHTON
HARO
BARCLAY
COMOX
BUTE
THURLOW
ROBSON
GEORGIA
PENDER
HASTINGS
DAVIE
BURNABY
NICOLA
JERVIS
HARWOOD
PACIFIC
BEACH
BEACH AVE
HORNBY
HOWE
SEYMOUR
RICHARDS
HOMER
HAMILTON
SMITHE
CAMBIE
NELSON
DRAKE
DUNSMUIR
EXPO BLVD
PACIFIC BLVD
CORDOVA
WATER
POWELL
CARRALL
COLUMBIA
ABBOTT
KEEFER
BEATTY
MAIN
QUEBEC
GORE
PRIOR ST
NATIONAL
TERMINAL
GT NORTHERN
West End
Gastown
Chinatown
Yaletown
Fairview
Mount Pleasant
English Bay
False Creek
False Creek Ferry
Aquabus
West End Community Centre
Robson Market
Barclay Heritage Square
Morton Park
Alexandra Park
English Bay Beach Park
Nelson Park
St. Paul's Hospital
Coal Harbour Park
Harbour Green Park
Vancouver Convention & Exhibition Centre
CN IMAX Theatre
Vancouver Tourist InfoCentre
Sinclair Centre
Harbour Centre/SFU
Portside Park
Royal Centre
Vancouver Art Gallery
Robson Square
Pacific Centre
UBC Downtown
Orpheum Theatre
Canada Post
Vancouver Library
Queen Elizabeth Theatre
VCC
Dr Sun Yat Sen Chinese Gardens
Andy Livingstone Park
G.M. Place
B.C. Place Stadium
Plaza of Nations
Creekside Park
Science World
Pacific Central Station (Bus & Railway)
Thornton Park
Sunset Beach Park
Vanier Park
H.R. MacMillan Space Centre & Vancouver Museum
Gordon Southam Observatory
Vancouver Academy of Music
Granville Island Market
Arts Club Theatre
George Wainborn Park
David Lam Park
Coopers Park
Emily Carr Institute of Art + Design
Waterfront Theatre
Sutcliffe Park
Charleson Park
BURRARD BRIDGE
GRANVILLE BRIDGE
CAMBIE BRIDGE
CHESTNUT
BURRARD
1 AVE
2 AVE
3 AVE
4 AVE
5 AVE
6 AVE
7 AVE
8 AVE
FIR
PINE
GRANVILLE
HEMLOCK
BIRCH
ALDER
SPRUCE
OAK
LAURIEL
WILLOW
HEATHER
ASH
YUKON
ALBERTA
MANITOBA
ONTARIO
SCOTIA
BRUNSWICK
LAMEY'S MILL
BROADWAY
240.241.242.246.247.250.251
252.253.254.257.258.N24
5 N6
2.C21
32
N22
19.22 44
10.44
44
6
N6.N8.N9 N10.N15 N17.N19 N20.N35
19.20.50
5.6.8
5.6.8.20
4.7 15.N19
N22
← 50
4.7.8 →
8.10.16.20.N20
50 →
4.7.N8
50 ↑
← 4.7
3
3 →
N35
19.N19
22
C21.C23 50 →
C21.C23
3.8.19.22.N8.N19
C23 →
C21.C23
22
98
C21.N22
2.22.32.44
6.C23
C21
N6
4.7.10.16.17
50
4.6.7.10.16.17
N9.N10.N17
C21.C23
N15
15.N15
C21.C23
C23
2.22.32.44.N22
4.7.10.16.17.50.98
N9.N10.N17
44
50
4.7.84
17
+50
84
50.84
15.N15
9.16.17.99.N17
9.17.99.N9
9.50.99.N9
9.99.N9
3.8.19.N8.N19
R 09.07

maps.google.ca

joanneallisonfox@gmail.com | My Profile | Saved Locations | Help | My Account | Sign ou

Google Maps Canada

Web Images Groups News **Maps** Scholar **more »**

coffee | 2185 East Mall, Vancouver, BC | Search Businesses

Search the map | **Find businesses** | Get directions

Search Results | My Maps

Print | Email | Link to this page

Text View | Map View

Results **1-10** of about **6,156** for **coffee** near **2185 E Mall, UBC, BC V6T** - Modify search

Categories: Coffee Retail, Coffee Houses

A **Original Beanery The** - more info »
2706 Fairview Crescent, Vancouver, BC V6T 2B9
(604) 224-2326 - 0.7 km E
Category: **Coffee** Houses

B **Starbucks Coffee Co** - more info »
6190 Agronomy Road, Vancouver, BC V6T 1Z3
(604) 221-6434 - 0.4 km SE

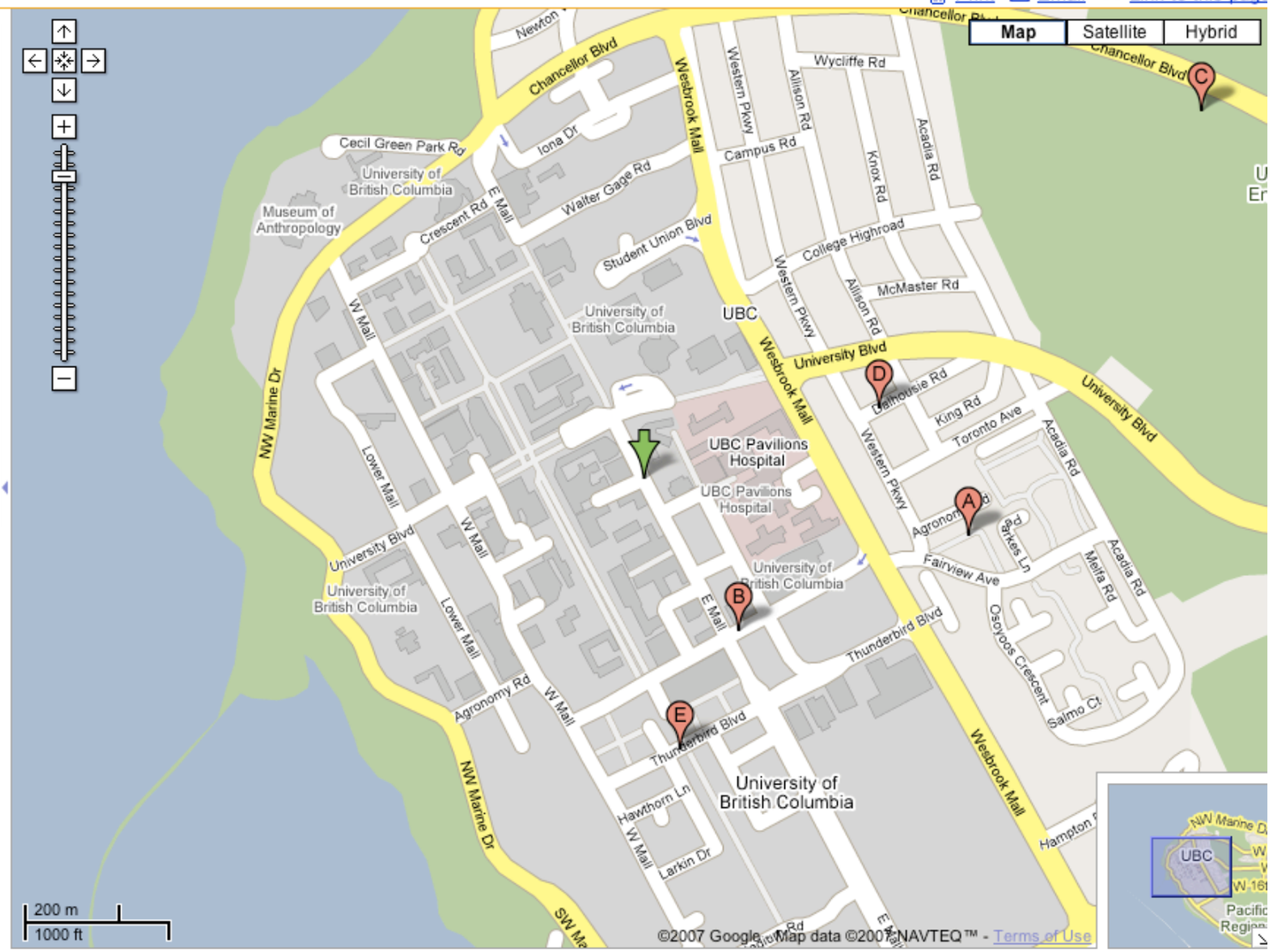
C **Blue Chip Cookies** - more info »
6138 Sub Boulevard, Vancouver, BC V6T 2A5
(604) 822-6999 - 1.4 km NE
Category: **Coffee** Houses

D **Starbucks Coffee Co** - more info »
5761 Dalhousie Road, Vancouver, BC V6T 2H9
(604) 221-0200 - 0.5 km E

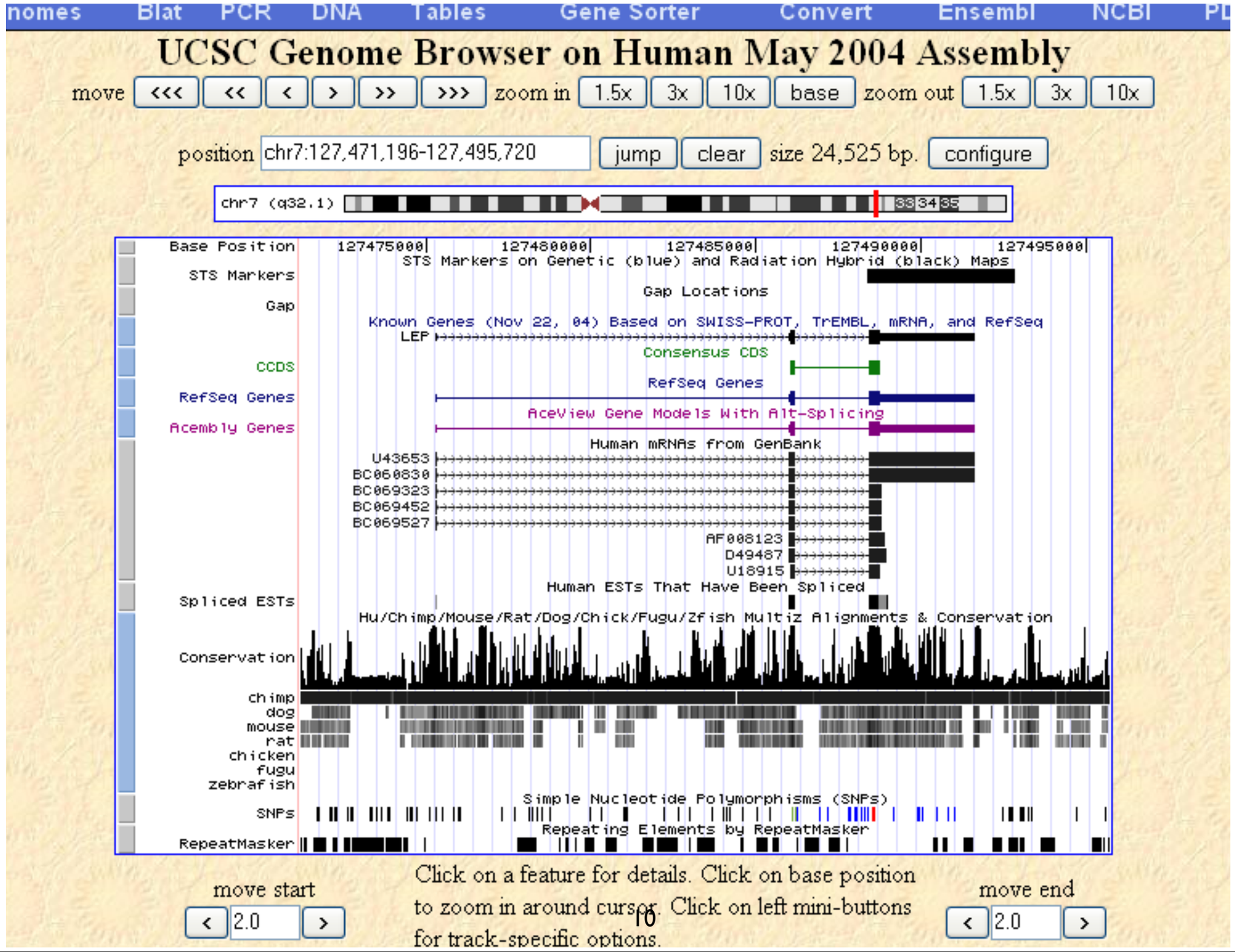
E **Salt Spring Coffee** - more info »
6308 Thunderbird Blvd, Vancouver, BC V6T 1Z4
(604) 221-6400 - 0.6 km S

F **Starbucks Coffee Co** - more info »
4580 10th Avenue West, Vancouver, BC V6R 2J1
(604) 222-1456 - 2.8 km E

G **European Crepe Cafe** - more info »
4544 West 10th Avenue, Vancouver, BC V6R 2J1



Let's Look at the Human Genome...



Objectives

- By the end of this module:

✓ You will be able to describe the following concepts: genome annotation, genome builds, and genome browsers.

✓ You will view the genomic location that contains the BRCA1 gene in the human genome using three different genome browsers.

✓ You will be able to compare and contrast the UCSC, Ensembl and MapViewer systems for visualizing genome information.

Genome Browsers

- What is a Genome Browser?
  - System for displaying, viewing, and accessing genome annotation data
- Genome annotations = knowledge attached to raw genome sequence.
  - Annotation information comes from many different sources
    - ✓ Computational pipelines
    - ✓ Research groups
    - ✓ Databases

The "Neopolitan Ice Cream" World of Genome Browsing:

- UCSC Genome Browser

  http://genome.cse.ucsc.edu/

- Ensembl

  http://www.ensembl.org/

- NCBI Map Viewer

  http://www.ncbi.nlm.nih.gov/mapview/

The underlying data is common for all three "flavors" of Genome Browsers.

- NCBI, UCSC and Ensembl use the same human genome assemblies that are generated by NCBI
  - release timing is different between sites.
- Note the version of genome assembly to which you are referring
  - available precomputed info and locations of features will be different between different assemblies.

Let's compare the view of the BRCA1 gene in all three genome browsers.

Viewing the genomic region containing BRCA1

- Common features:

✓ Coordinate system is based on the build

✓ Zoom in and out

✓ Annotations displayed – ie. Gene features

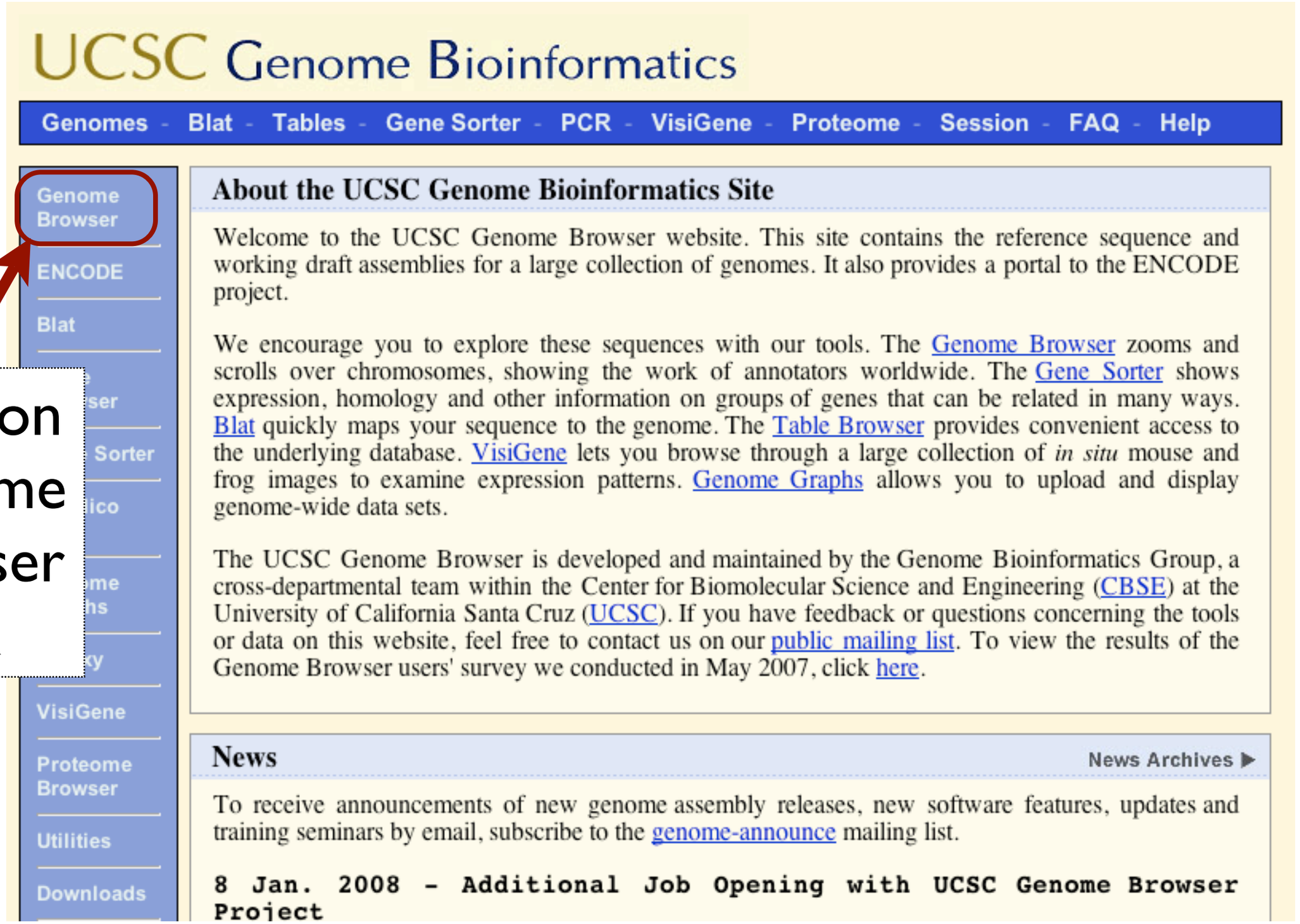
- Major Differences:

✓ Each Browser has a very different look and feel

✓ Annotation information displayed differently

✓ Different ways to navigate through the information

http://genome.cse.ucsc.edu/



Click on Genome Browser link

Home Genomes Blat Tables Gene Sorter PCR FAQ Help

Human (*Homo sapiens*) Genome Browser Gateway

The UCSC Genome Browser was created by the Genome Bioinformatics Group of UC Santa Cruz.
Software Copyright (c) The Regents of the University of California. All rights reserved.

clade	genome	assembly	position or search term	image width	
Vertebrate	Human	May 2004	BRCA1	620	submit

Click here to reset the browser user interface settings to their defaults

add your own custom tracks | configure tracks and display | clear position

About the Human May 2004 (hg17) assembly (sequences)

The May 2004 human reference sequence is based on NCBI Build 35 and was produced by the International Human Genome Se

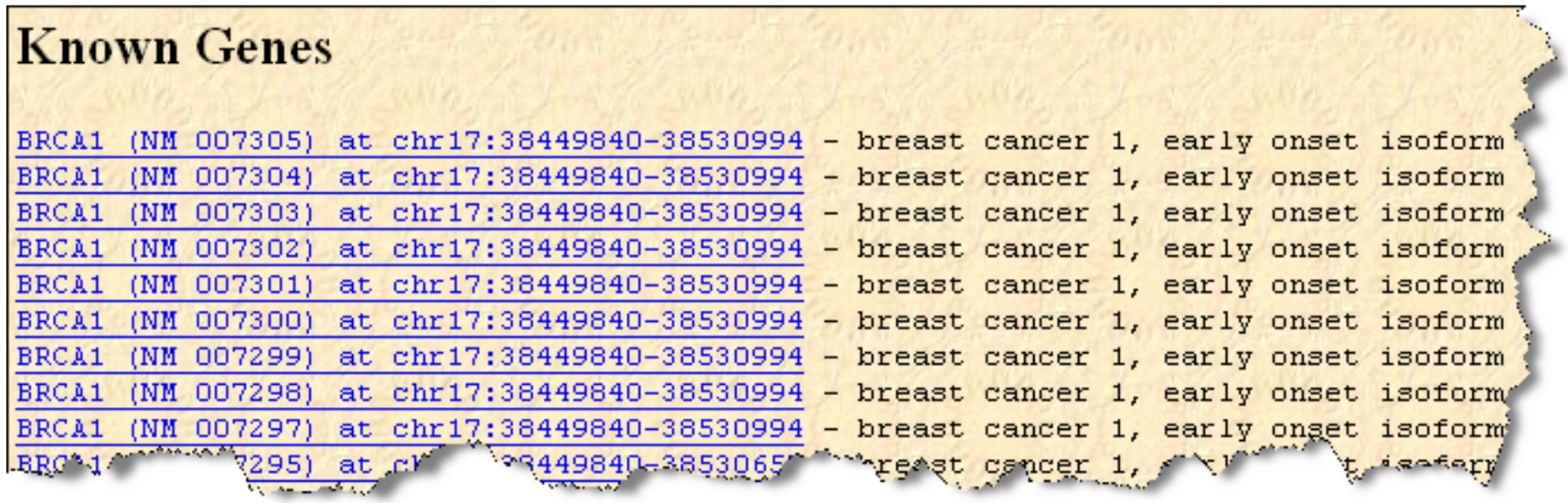
Sample position queries

A genome position can be specified by the accession number of a sequenced genomic clone, an mRNA or EST or STS m
chromosomal coordinate range, or keywords from the GenBank description of an mRNA. The following list shows examples
human genome. See the User's Guide for more information.

Request:	Genome Browser Response:
chr7	Displays all of chromosome 7
20p13	Displays region for band p13 on chr 20
chr3:1-1000000	Displays first million bases of chr 3, counting from p arm telomere
D16S3046	Displays region around STS marker D16S3046 from the Genethon/Marshfield maps. Includes 100,000 bases on each side as well.
RH18061;RH80175	Displays region between STS markers RH18061;RH80175. Includes 100,000 bases on each side as well.
AA205474	Displays region of EST with GenBank accession AA205474 in BRCA1 cancer gene on chr 17
AC008101	Displays region of clone with GenBank accession AC008101
AF083811	Displays region of mRNA with GenBank accession number AF083811
PRNP	Displays region of genome with HUGO Gene Nomenclature Committee identifier PRNP
NM_017414	Displays the region of genome with RefSeq identifier NM_017414
NP_059110	Displays the region of genome with protein accession number NP_059110
pseudogene mRNA	Lists transcribed pseudogenes, but not cDNAs

Search for BRCA1; Note sample queries

The Search Results



Known Genes

BRCA1 (NM 007305) at chr17:38449840-38530994 - breast cancer 1, early onset isoform
BRCA1 (NM 007304) at chr17:38449840-38530994 - breast cancer 1, early onset isoform
BRCA1 (NM 007303) at chr17:38449840-38530994 - breast cancer 1, early onset isoform
BRCA1 (NM 007302) at chr17:38449840-38530994 - breast cancer 1, early onset isoform
BRCA1 (NM 007301) at chr17:38449840-38530994 - breast cancer 1, early onset isoform
BRCA1 (NM 007300) at chr17:38449840-38530994 - breast cancer 1, early onset isoform
BRCA1 (NM 007299) at chr17:38449840-38530994 - breast cancer 1, early onset isoform
BRCA1 (NM 007298) at chr17:38449840-38530994 - breast cancer 1, early onset isoform
BRCA1 (NM 007297) at chr17:38449840-38530994 - breast cancer 1, early onset isoform
BRCA1 (NM 007295) at chr17:38449840-38530657 - breast cancer 1, early onset isoform 1
BRIP1 (NM 032043) at chr17:57114767-57295537 - BRCA1 interacting protein C-terminal helicase 1
CXorf53 (NM 024332) at chrX:153952904-154004542 - BRCA1/BRCA2-containing complex subunit 36
BRAP (NM 006768) at chr12:110566279-110608122 - BRCA1 associated protein
BAP1 (NM 004656) at chr3:52410067-52419049 - BRCA1 associated protein-1
CXorf53 (NM 001018055) at chrX:153952904-154004542 - BRCA1/BRCA2-containing complex subunit 36
BARD1 (NM 000465) at chr2:215301522-215382611 - BRCA1 associated RING domain 1
BRCA1 (BC072418) at chr17:38450367-38530994 - BRCA1 protein.
NBR1 (NM 031858) at chr17:38576037-38719014 - neighbor of BRCA1 gene 1
COBRA1 (NM 015456) at chr9:139269768-139287812 - cofactor of BRCA1
NBR1 (BC009808) at chr17:38576786-38717401 - Neighbor of BRCA1 gene 1 (Fragment).
BRE (AF410603) at chr2:27967164-28415087 - Brain and reproductive organ-expressed protein (Brain and reproductive organ-express
RBBP8 (BC030590) at chr18:18768711-18860331 - retinoblastoma binding protein 8
BRCA2 (NM 000059) at chr13:31787617-31871805 - breast cancer 2, early onset
NBR2 (NM 005821) at chr17:38531129-38550649 - hypothetical protein LOC10230
MDC1 (NM 014641) at chr6:30775564-30790934 - mediator of DNA damage checkpoint 1
MDC1 (NM 014641) at chr6_cox_hap1:2116161-2131531 - mediator of DNA damage checkpoint 1
MDC1 (NM 014641) at chr6_qbl_hap2:1914464-1929833 - mediator of DNA damage checkpoint 1
ZNF350 (NM 021632) at chr19:57159407-57181891 - zinc finger protein 350
CLSPN (AF297866) at chr1:35958407-36008138 - claspin homolog (Xenopus laevis)
MSH6 (NM 000179) at chr2:47863790-47887595 - mutS homolog 6
ATM (NM 138292) at chr11:107663329-107745036 - ataxia telangiectasia mutated protein isoform 2
RBBP8 (NM 002894) at chr18:18767293-18860443 - retinoblastoma binding protein 8 isoform a
ATM (NM 000051) at chr11:107598769-107745036 - ataxia telangiectasia mutated protein isoform 1
CLSPN (NM 022111) at chr1:35974407-36008138 - claspin
RAD51 (NM 133487) at chr15:38774661-38811646 - RAD51 homolog protein isoform 2
RBBP7 (NM 002893) at chrX:16772698-16798378 - retinoblastoma binding protein 7
BLM (NM 000057) at chr15:89061606-89159601 - Bloom syndrome protein
ATR (NM 001184) at chr3:143650770-143780341 - ataxia telangiectasia and Rad3 related protein
RAD51 (NM 002875) at chr15:38774661-38811646 - RAD51 homolog protein isoform 1
MLH1 (NM 000249) at chr3:37009983-37067339 - MutL protein homolog 1
SMARCA4 (NM 003072) at chr19:10932606-11033952 - SWI/SNF-related matrix-associated
MSH2 (NM 000251) at chr2:47483767-47563864 - mutS homolog 2
FANCD2 (NM 033084) at chr3:10043113-10116343 - Fanconi anemia complementation group D2 isoform
SMC1L1 (NM 006306) at chrX:53417797-53466343 - SMC1 structural maintenance of chromosomes
PMS2 (NM 001018040) at chr7:5979397-6015232 - PMS2 postmeiotic segregation increased 2 isoform
DUSP3 (NM 004090) at chr17:39199016-39211872 - dual specificity phosphatase 3
FANCD2 (NM 001018115) at chr3:10043113-10118614 - Fanconi anemia complementation group D2 isoform
PMS2 (NM 000535) at chr7:5979397-6015232 - PMS2 postmeiotic segregation increased 2 isoform
DCLRE1A (NM 014881) at chr10:115584474-115603597 - DNA-crosslink repair gene SNM1
TOPBP1 (D87440) at chr3:134802140-134863380 - topoisomerase (DNA) II binding protein 1
CHEK2 (NM 145862) at chr22:27413731-27467822 - protein kinase CHK2 isoform b
CHEK2 (NM 001005735) at chr22:27413731-27467822 - protein kinase CHK2 isoform c
CHEK2 (NM 007194) at chr22:27413731-27467822 - protein kinase CHK2 isoform a
TOPBP1 (NM 007027) at chr3:134802177-134863380 - topoisomerase (DNA) II binding protein 1
PTP4A2 (BC070182) at chr1:32146442-32157581 - protein tyrosine phosphatase type IVA, member 2
CHEK1 (NM 001274) at chr11:125001547-125030847 - CHK1 checkpoint homolog
H2AFX (NM 002105) at chr11:118469797-118471387 - H2A histone family, member X
DCLRE1C (NM 001033855) at chr10:14988879-15036100 - artemis protein isoform a
DCLRE1C (NM 001033858) at chr10:14988879-15036100 - artemis protein isoform c
RAD50 (NM 005732) at chr5:131920529-132007492 - RAD50 homolog isoform 1
MRE11A (NM 005591) at chr11:93790115-93866688 - meiotic recombination 11 homolog A isoform 1
KPNA2 (NM 002266) at chr17:63462310-63473431 - karyopherin alpha 2
KPNA2 (NM 002266) at chr17_random:1713112-1724231 - karyopherin alpha 2

RefSeq Genes

BRCA1 at chr17:38449840-38530994 - (NM_007302) breast cancer 1, early onset isoform
BRCA1 at chr17:38449840-38530994 - (NM_007297) breast cancer 1, early onset isoform
BRCA1 at chr17:38449840-38530994 - (NM_007298) breast cancer 1, early onset isoform
BRCA1 at chr17:38449840-38530994 - (NM_007296) breast cancer 1, early onset isoform 1
BRCA1 at chr17:38449840-38530994 - (NM_007294) breast cancer 1, early onset isoform 1
BRCA1 at chr17:38449840-38530657 - (NM_007295) breast cancer 1, early onset isoform 1
BRCA1 at chr17:38449840-38530994 - (NM_007305) breast cancer 1, early onset isoform
BRCA1 at chr17:38449840-38530994 - (NM_007299) breast cancer 1, early onset isoform
BRCA1 at chr17:38449840-38530994 - (NM_007303) breast cancer 1, early onset isoform
BRCA1 at chr17:38449840-38530994 - (NM_007304) breast cancer 1, early onset isoform
BRCA1 at chr17:38449840-38530994 - (NM_007300) breast cancer 1, early onset isoform
BAP1 at chr3:52410067-52419049 - (NM_004656) BRCA1 associated protein-1
BRAP at chr12:110566279-110608122 - (NM_006768) BRCA1 associated protein
BRCC3 at chrX:153952904-154004542 - (NM_001018055) BRCA1/BRCA2-containing complex, subunit 3
BRCC3 at chrX:153952904-154004542 - (NM_024332) BRCA1/BRCA2-containing complex, subunit 3
COBRA1 at chr9:139269768-139287812 - (NM_015456) cofactor of BRCA1
NBR1 at chr17:38576037-38719232 - (NM_031858) neighbor of BRCA1 gene 1
NBR1 at chr17:38576772-38719232 - (NM_005899) neighbor of BRCA1 gene 1
NBR1 at chr17:38576024-38719232 - (NM_031862) neighbor of BRCA1 gene 1
BRIP1 at chr17:57114767-57295537 - (NM_032043) BRCA1 interacting protein C-terminal helicase 1
BARD1 at chr2:215301522-215382611 - (NM_000465) BRCA1 associated RING domain 1

Non-Human RefSeq Genes

BRCA1 at chr17:38451235-38529639 - (NM_001013416) breast cancer 1, early onset
BRCA1 at chr17:38451245-38529639 - (NM_178573) breast cancer 1, early onset
Brca1 at chr17:38453196-38529662 - (NM_012514) breast cancer 1
Brca1 at chr17:38449846-38529662 - (NM_009764) breast cancer 1
BRCA1 at chr17:38451222-38529639 - (NM_001045493) breast cancer 1, early onset
Bard1 at chr2:215301576-215382511 - (NM_022622) BRCA1 associated RING domain 1
Nbr1 at chr17_random:460667-479659 - (NM_001024765) neighbor of Brca1 gene 1
Nbr1 at chr17:38531912-38718946 - (NM_001024765) neighbor of Brca1 gene 1
Bard1 at chr2:215301595-215382512 - (NM_007525) BRCA1 associated RING domain 1
Bap1 at chr3:52410186-52419027 - (NM_027088) Brca1 associated protein 1
Nbr1 at chr17_random:460667-480016 - (NM_008676) neighbor of Brca1 gene 1
Nbr1 at chr17:38576748-38718946 - (NM_008676) neighbor of Brca1 gene 1
Brip1 at chr17:57118022-57293682 - (NM_178309) BRCA1 interacting protein C-terminal helicase 1
Brap at chr12:110564693-110608153 - (NM_028227) BRCA1 associated protein
A730008L03Rik at chr9:139269714-139287204 - (NM_021393) cofactor of BRCA1
BRIP1 at chr17:57115775-57293682 - (NM_001033058) BRCA1 interacting protein C-terminal helicase 1
LOC417254 at chr9:139269780-139287192 - (NM_001006202) similar to cofactor of BRCA1; negative
BAP1 at chr3:52411344-52418952 - (NM_001030590) BRCA1 associated protein-1
BARD1 at chr2:215301762-215370076 - (NM_001031246) BRCA1 associated RING domain 1
brap at chr12:110566386-110607979 - (NM_001002466) BRCA1 associated protein
cobra1 at chr9:139269783-139286879 - (NM_200168) cofactor of BRCA1 isoform 2
cobra1 at chr9:139269783-139286879 - (NM_001020786) cofactor of BRCA1 isoform 1

Alias of STS Marker

BRCA1 at chr7:34489430-34689640 - (AFM067XA9)

Human Aligned mRNA Search Results

BC001596 - Homo sapiens BRCA1 associated protein-1 (ubiquitin carboxy-terminal hydrolase), mRNA (cDNA clone MGC:1530 IMAGE:3543
BC002999 - Homo sapiens BRCA1/BRCA2-containing complex, subunit 3, mRNA (cDNA clone MGC:3961 IMAGE:2821917), complete cds.
BC006540 - Homo sapiens BRCA1/BRCA2-containing complex, subunit 3, mRNA (cDNA clone MGC:2329 IMAGE:2821917), complete cds.
U76638 - Human BRCA1-associated RING domain protein (BARD1) mRNA, complete cds.
U64805 - Homo sapiens Brca1-delta11b (Brca1) mRNA, complete cds.
BC101472 - Homo sapiens BRCA1 interacting protein C-terminal helicase 1, mRNA (cDNA clone MGC:126521 IMAGE:8068978), complete c
BC101474 - Homo sapiens BRCA1 interacting protein C-terminal helicase 1, mRNA (cDNA clone MGC:126523 IMAGE:8068980), complete c
BC062429 - Homo sapiens breast cancer 1, early onset, mRNA (cDNA clone IMAGE:3686198), partial cds.
BC072418 - Homo sapiens breast cancer 1, early onset, mRNA (cDNA clone IMAGE:6181860), complete cds.
BC085615 - Homo sapiens breast cancer 1, early onset, mRNA (cDNA clone IMAGE:6042052), partial cds.
BC115037 - Homo sapiens breast cancer 1, early onset, mRNA (cDNA clone MGC:131629 IMAGE:7961446), complete cds.
AY597810 - Homo sapiens chromosome 17 neighbor of BRCA1 gene 2 (NBR2) mRNA, complete cds, alternatively spliced.
AY438031 - Homo sapiens BRCA1/BRCA2-containing complex subunit 45 (BRE) mRNA, complete cds.
AY438030 - Homo sapiens BRCA1/BRCA2-containing complex subunit 36 (BRCC36) mRNA, complete cds.
AK223409 - Homo sapiens mRNA for BRCA1 associated RING domain 1 variant, clone: FCC109A11.
AF464935 - Homo sapiens cofactor of BRCA1 (COBRA1) mRNA, complete cds.
AF360549 - Homo sapiens BRCA1-binding helicase-like protein BACH1 mRNA, complete cds.
AF151109 - Homo sapiens putative BRCA1-interacting protein (BRIP1) mRNA, partial cds.
AF045581 - Homo sapiens BRCA1 associated protein 1 (BAP1) mRNA, complete cds.
AF035620 - Homo sapiens BRCA1-associated protein 2 (BRAP2) mRNA, complete cds.
AB208878 - Homo sapiens mRNA for BRCA1 associated protein variant protein.
AF005068 - Homo sapiens breast and ovarian cancer susceptibility protein splice variant (BRCA1) mRNA, complete cds.
AY751490 - Homo sapiens breast and ovarian cancer susceptibility protein (BRCA1) mRNA, BRCA1-22017T/2430C/27311T/3232G/3667G/4427
Y08864 - H.sapiens mRNA for breast and ovarian cancer susceptibility protein (BRCA1), 3'UTR.
U14680 - Homo sapiens breast and ovarian cancer susceptibility (BRCA1) mRNA, complete cds.
U68041 - Human breast and ovarian cancer susceptibility protein (BRCA1) mRNA, partial cds.
AK225494 - Homo sapiens mRNA for BRCA1 associated protein-1 variant, clone: JTH08468.
BC107773 - Homo sapiens neighbor of BRCA1 gene 2, mRNA (cDNA clone MGC:104305 IMAGE:6452095), complete cds.
BC034248 - Homo sapiens neighbor of BRCA1 gene 2, mRNA (cDNA clone IMAGE:4339497), partial cds.
BC011892 - Homo sapiens cofactor of BRCA1, mRNA (cDNA clone MGC:19886 IMAGE:3941719), complete cds.
BC009808 - Homo sapiens neighbor of BRCA1 gene 1, mRNA (cDNA clone MGC:1377 IMAGE:2969212), complete cds.
BC000924 - Homo sapiens neighbor of BRCA1 gene 2, mRNA (cDNA clone MGC:5031 IMAGE:3446931), complete cds.
AK094499 - Homo sapiens cDNA FLJ37180 fis, clone BRALZ2001038, highly similar to Homo sapiens BRCA1 associated protein 1 (BAP1)
AK092725 - Homo sapiens cDNA FLJ35406 fis, clone SKNSH2007429, highly similar to Homo sapiens BRCA1 associated protein 1 (BAP1)
AF274503 - Homo sapiens breast and ovarian cancer susceptibility (BRCA1) pseudogene, partial mRNA sequence.
BC022065 - Homo sapiens neighbor of BRCA1 gene 2, mRNA (cDNA clone IMAGE:4826858), with apparent retained intron.
BC019029 - Homo sapiens cofactor of BRCA1, mRNA (cDNA clone IMAGE:4554322).
BC012847 - Homo sapiens neighbor of BRCA1 gene 1, mRNA (cDNA clone IMAGE:3858519), with apparent retained intron.

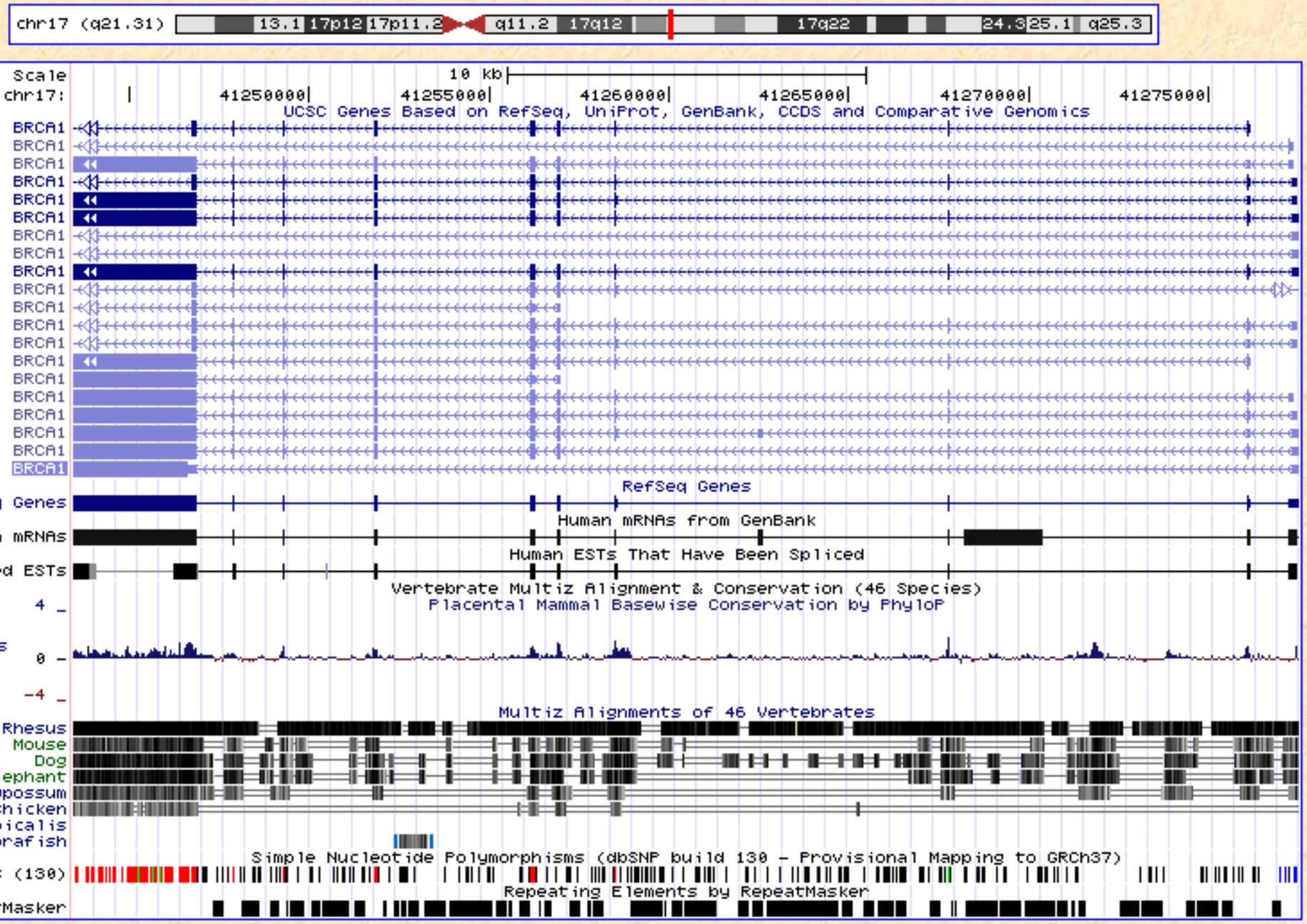
- Many BRCA1 isoforms
  - ✓ All located on chr 17
  - ✓ same chr coordinates
  - ✓ different gene structures

Home Genomes Blat Tables Gene Sorter PCR DNA Convert Ensembl NCBI PDF/PS Session Help

UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x

position/search chr17:41,243,452-41,277,500 gene jump clear size 34,049 bp. configure

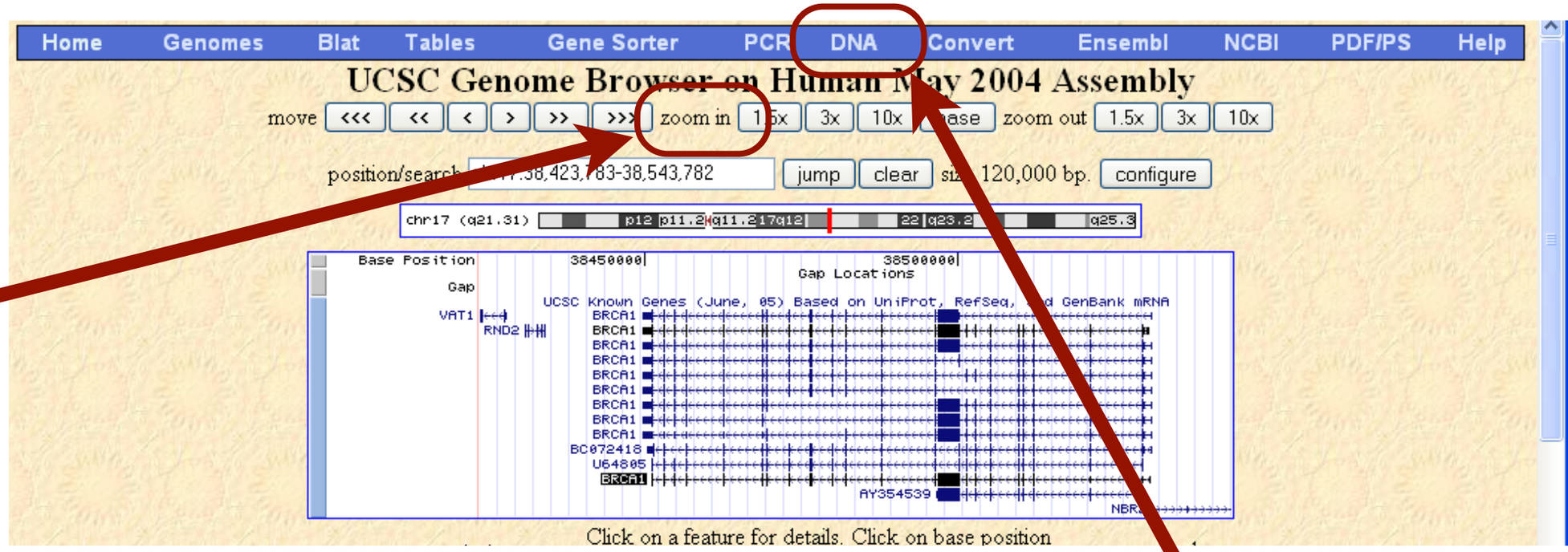


move start Click on a feature for details. Click or drag in the base position track to zoom in. move end

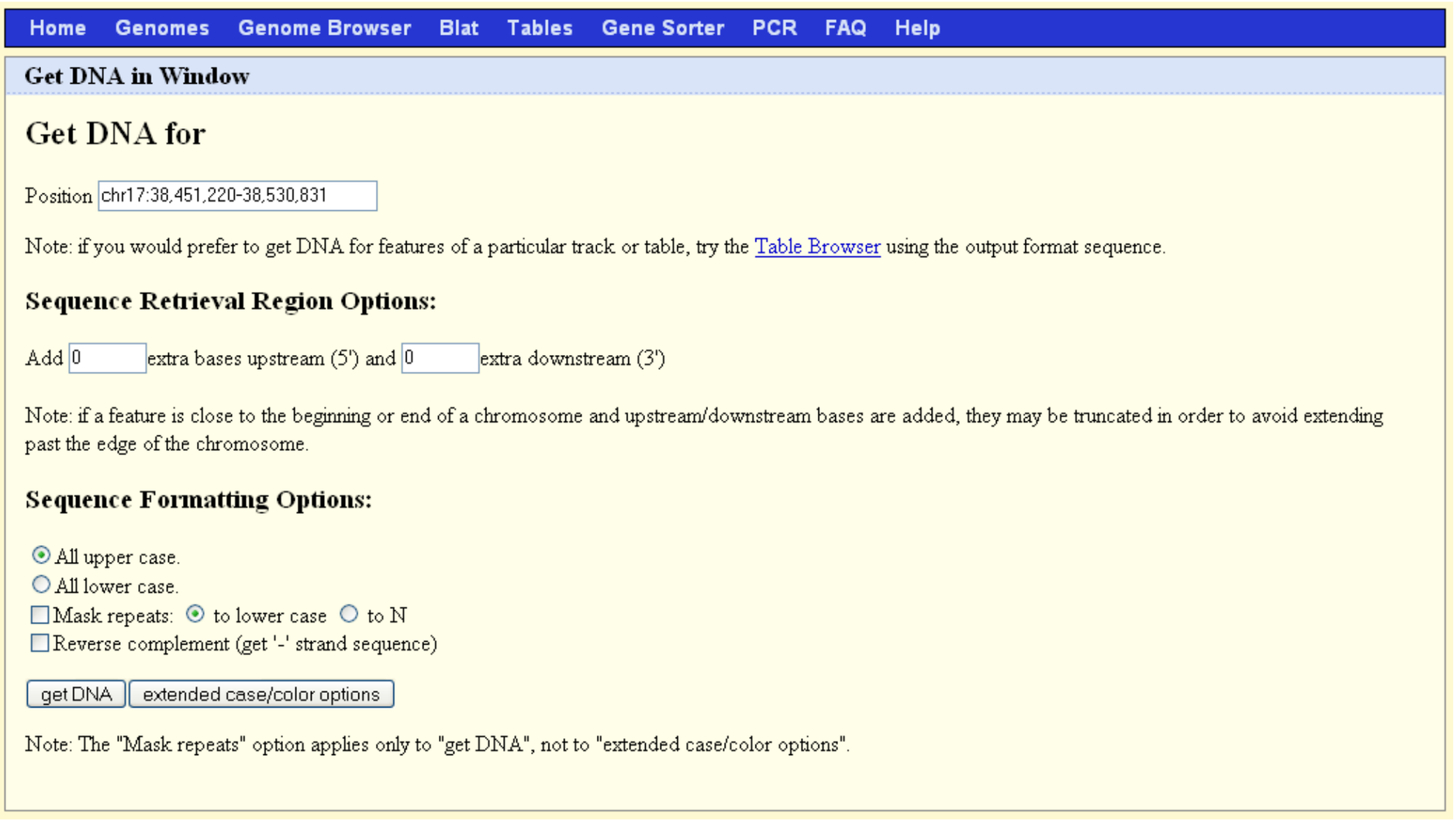
Tasks

- What genes are on either side of BRCA1 on chr 17?
- Can you figure out how to download the genomic sequence for the BRCA1 region?
- Can you figure the display to add/remove tracks that are (or are not) of interest to you?

Zoom in
Zoom out



DNA link
Download
Sequence



collapse all

Use drop-down controls below and press refresh to alter tracks displayed.
Tracks with lots of items will automatically be displayed in more compact modes.

expand all

Mapping and Sequencing Tracks refresh

Base Position: dense	Chromosome Band: hide	STS Markers: hide	FISH Clones: hide	Recomb Rate: hide	Map Contigs: hide
Assembly: hide	Gap: hide	Coverage: hide	BAC End Pairs: hide	Fosmid End Pairs: hide	GC Percent: hide
Short Match: hide	Restr Enzymes: hide				

Phenotype and Disease Associations refresh

Genes and Gene Prediction Tracks refresh

UCSC Genes: pack	Old UCSC Genes: hide	Alt Events: hide	CCDS: hide	RefSeq Genes: dense	Other RefSeq: hide
MGC Genes: pack	ORFeome Clones: hide	TransMap...: hide	Vega Genes: hide	Ensembl Genes: hide	AceView Genes: hide
SIB Genes: hide	N-SCAN: hide	CONTRAST: hide	SGP Genes: hide		
Exoniphy: hide	Augustus: hide	RNA Genes: hide	ACEScan: hide		
Pos Sel Genes: hide					

mRNA and EST Tracks

Expression refresh

Regulation refresh

Comparative Genomics refresh

Variation and Repeats refresh

Pilot ENCODE Regions and Genes refresh

Drop down controls configure the data shown in the image above

http://www.ensembl.org/

e!Ensembl West

Home | Login / Register | BLAST/BLAT | BioMart | Docs & FAQs

You've been redirected to your nearest mirror - uswest.ensembl.org

- Take me back to www.ensembl.org

Search: All species for

Go

e.g. **human gene BRCA2** or **rat X:100000..200000** or **coronary heart disease**

Browse a Genome

The Ensembl project produces genome databases for vertebrates and other eukaryotic species, and makes this information freely available online.

Click on a link below to go to the species' home page.

Popular genomes (Log in to customize this list)

Human
GRCh37

Mouse
NCBIM37

Zebrafish
Zv8

New to Ensembl?

Did you know you can:

- **Learn how to use Ensembl**
  with our video tutorials and walk-throughs
- **Add custom tracks**
  using our new Control Panel
- **Upload and**
  and save it
- **Search for a**
  sequence
  using BLAST or BLAT
- **Fetch only the data you want**
  from our public database, using the Perl API
- **Download our databases via FTP**
  in FASTA, MySQL and other formats
- **Mine Ensembl with BioMart**
  and export sequences or tables in text, html, or Excel format

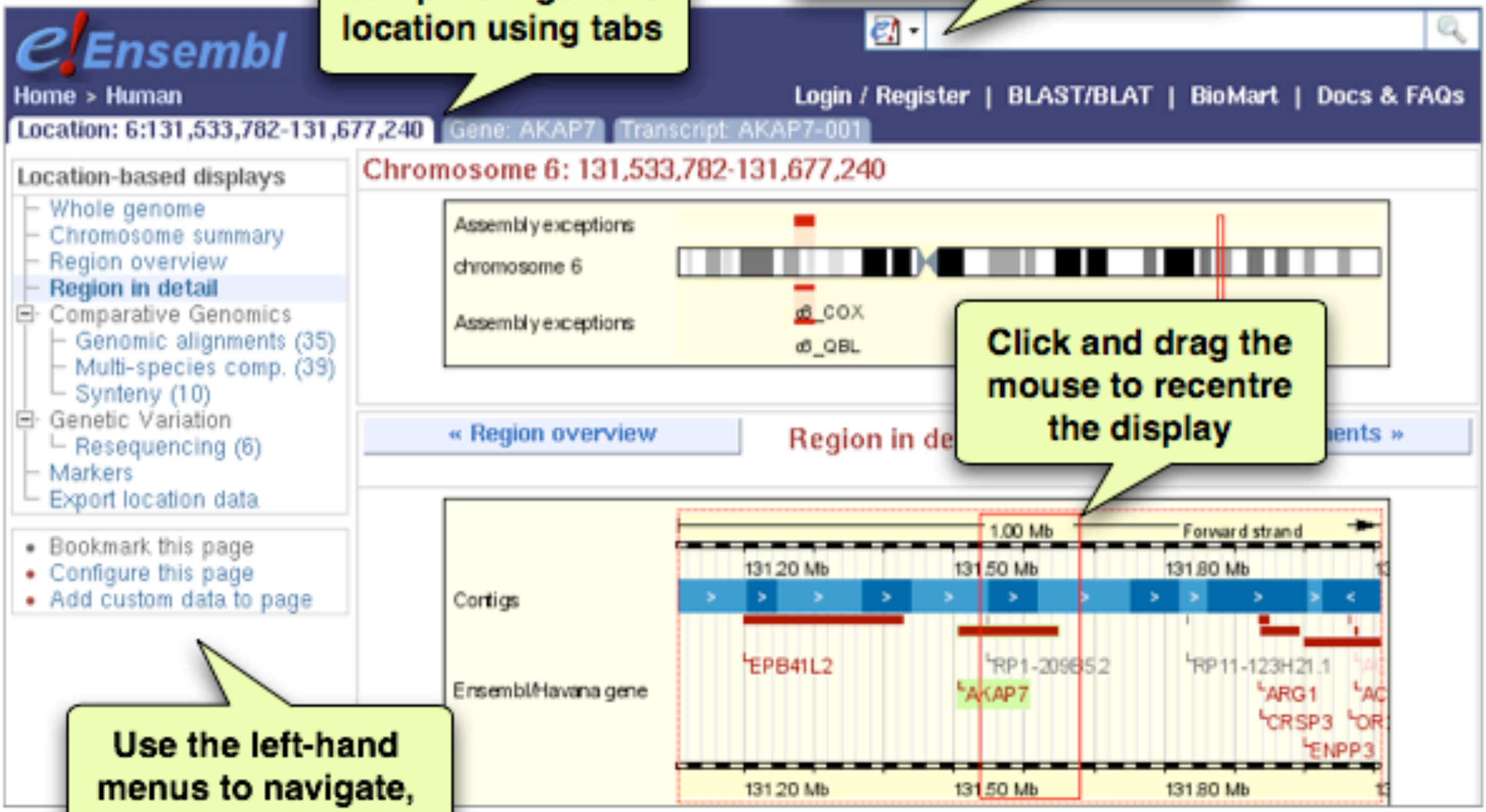
Did you know...?

A preliminary assembly of the Giant panda (Ailuropoda

NEW!

http://pre.ensembl.org/p

Click on Human


Search Ensembl, EBI or Sanger Institute
Jump from gene to location using tabs
e!Ensembl
Home > Human
Login / Register | BLAST/BLAT | BioMart | Docs & FAQs
Location: 6:131,533,782-131,677,240
Gene: AKAP7
Transcript: AKAP7-001
Location-based displays
Whole genome
Chromosome summary
Region overview
Region in detail
Comparative Genomics
Genomic alignments (35)
Multi-species comp. (39)
Synteny (10)
Genetic Variation
Resequencing (6)
Markers
Export location data
Bookmark this page
Configure this page
Add custom data to page
Chromosome 6: 131,533,782-131,677,240
Assembly exceptions
chromosome 6
Assembly exceptions
Click and drag the mouse to recentre the display
« Region overview
Region in de
ents »
1.00 Mb
Forward strand
131.20 Mb
131.50 Mb
131.80 Mb
Contigs
Ensembl/Havana gene
EPB41L2
RP1-209B5.2
AKAP7
RP11-123H21.1
ARG1
CRSP3
ENPP3
Use the left-hand menus to navigate, export data and customise the page

About this species
- Description
- Genome Statistics
  - Assembly and Genebuil
  - Top 40 InterPro hits
  - Top 500 InterPro hits
- What's New
- Sample entry points
  - Karyotype
  - Location (6:1330176
  - Gene (BRCA2)
  - Transcript (FOXP2-203
  - Variation (rs1333049)
  - Regulation (ENSR0000

Configure this page
Manage your data
Export data
Bookmark this page

You've been redirected to your nearest mirror - uswest.ensembl.org

- Take me back to www.ensembl.org

Search Ensembl Human

Search for: Go

e.g. **gene BRCA2** or **6:133017695-133161157** or **osteoarthritis**

Description

Assembly and Genebuild »

Human (*Homo sapiens*)

Assembly

This site provides a data set based on the February 2009 *Homo sapiens* high coverage assembly from the Genome Reference Consortium. The data set consists of gene models built from the genewise alignments of the human proteome as well as from alignments of human cDNAs using the cDNA2genome model of exonerate

This release of the assembly has the following properties:

- 27478 contigs.
- contig length total 3.2 Gb.
- chromosome length total 3.1 Gb.

It also includes nine haplotypic regions, mainly in the MHC region of chromosome 6.

To convert your old data from Human assembly NCBI36 to GRCh37, click on 'Manage your data' on any human page and select 'Assembly converter' from the left-hand menu.

Annotation

Ensembl West

Home > Human [GRCh37]

Login / Register | BLAST/BLAT | BioMart | Docs & FAQs

Genome

Search Ensembl

New Search

Configure this page

Manage your data

Export data

Bookmark this page

Result in Detail

brca1 corporate/tree:"Top/Species/Homo sapiens" corp Search

Your query matched 23 entries in the search database. Viewing hits 1-10

1 2 3

Ensembl protein_coding Gene: ENSG00000012048 (HGNC Symbol: BRCA1) [Region in detail]

Description: breast cancer 1, early onset [Source:HGNC Symbol;Acc:1100]

Source: e58; **Feature type**: Gene; Homo sapiens; **Species**: Homo sapiens; Gene;

Havana protein_coding Gene: OTTHUMG00000157426 (BRCA1) [Region in detail]

Description: breast cancer 1, early onset

Source: e58; **Feature type**: Gene; Homo sapiens; **Species**: Homo sapiens; Gene;

Ensembl protein_coding Gene: ENSG00000087206 (HGNC Symbol: UIMC1) [Region in detail]

action motif containing 1 [Source:HGNC Symbol;Acc:30298]

ture type: Gene; Homo sapiens; **Species**: Homo sapiens; Gene;

g Gene: ENSG00000089234 (HGNC Symbol: BRAP) [Region in detail]

Description: **BRCA1** associated protein [Source:HGNC Symbol;Acc:1099]

Source: e58; **Feature type**: Gene; Homo sapiens; **Species**: Homo sapiens; Gene;

Ensembl protein_coding Gene: ENSG00000105393 (HGNC Symbol: C19orf62) [Region in detail]

Description: **BRCA1**-A complex subunit MERIT40 (Mediator of RAP80 interactions and targeting subunit of 40 kDa)(New component of the **BRCA1**-A complex) [Source:UniProtKB/Swiss-Prot;Acc:Q9NWV8]

Click on ENSG00000012048

Ensembl West

Home > Human [GRCh37]

Login / Register | BLAST/BLAT | BioMart | Docs & FAQs

Location: 17:41,196,312-41,277,500 | Gene: BRCA1

Gene-based displays

- Gene summary
- Splice variants (32)
- Supporting evidence
- Sequence
- External references (4)
- Regulation
- Comparative Genomics
  - Genomic alignments (5
  - Gene Tree (image)
    - Gene Tree (text)
    - Gene Tree (alignment

Configure this page

Manage your data

Export data

Bookmark this page

You've been redirected to your nearest mirror - uswest.ensembl.org

- Take me back to www.ensembl.org

Gene: BRCA1 (ENSG00000012048)

breast cancer 1, early onset [Source:HGNC Symbol;Acc:1100]

Location Chromosome 17: 41,196,312-41,277,500 reverse strand.

Transcripts There are 32 transcripts in this gene

entries | Show/hide columns | Search:

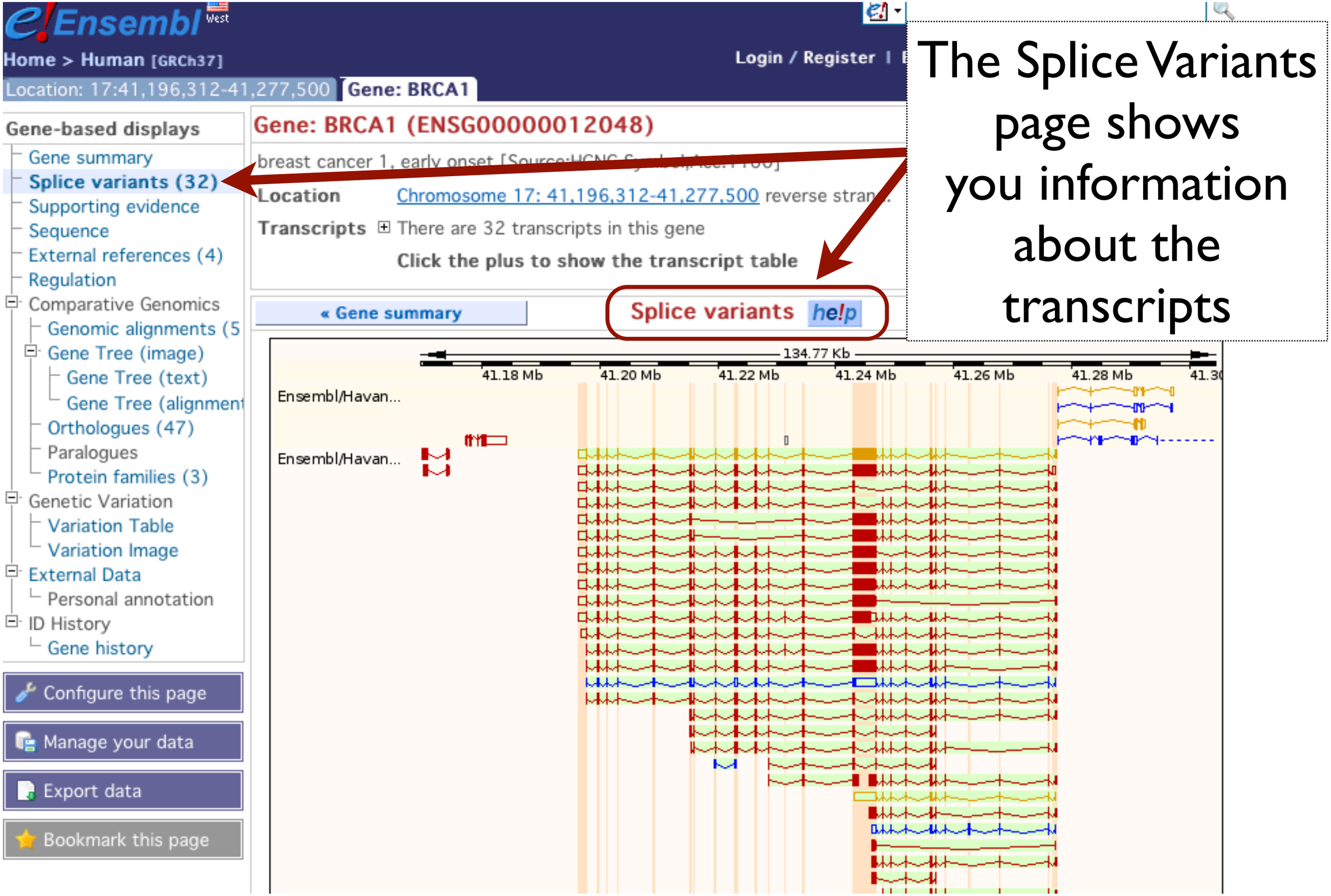
Transcript ID	Length (bp)	Protein ID	Length (aa)	Biotype
ENST00000357654	7094	ENSP00000350283	1863	Protein cod
ENST00000497488	779	ENSP00000418986	177	Protein cod
ENST00000477152	1930	ENSP00000419988	622	Protein cod
ENST00000471181	5939	ENSP00000418960	1885	Protein cod
ENST00000493795	5732	ENSP00000418775	1816	Protein cod
ENST00000		417148	699	Protein cod
BRCA1-008	ENST00000	418819	572	Protein cod
BRCA1-009	ENST00000	420412	623	Protein cod
BRCA1-011	ENST00000	419274	649	Protein cod
BRCA1-013	ENST00000	419103	473	Protein cod
BRCA1-014	ENST00000	420705	758	Protein cod
BRCA1-015	ENST00000	419481	498	Protein cod
BRCA1-016	ENST00000489037	455	ENSP00000420781	98	Protein cod
BRCA1-017	ENST00000476777	769	ENSP00000417554	222	Protein cod
BRCA1-018	ENST00000473961	958	ENSP00000420201	319	Protein cod

Gene Summary shows you information about the gene

click here to view genomic location

Tasks

- Explore the information presented in the Gene Summary views.
  - Can you figure out how to visualize the alternatively spliced isoforms for BRCA1?
  - What can you find out about known variations in this gene?
- Using the Location Based Displays, can you figure out how to download the genomic sequence for the BRCA1 region?



The Splice Variants page shows you information about the transcripts

Ensembl

ome > Human

Login / Register | BLAST/BLAT | Biol

ocation: 17:38,449,840-38,530,994 | Gene: BRCA1

ene: BRCA1

- Gene summary
- Splice variants (10)
- Supporting evidence
- Sequence
- External references (15)
- Regulation
- Comparative Genomics
  - Genomic alignments (3
- Gene Tree
  - Gene Tree (text)
  - Gene Tree (alignment
- Orthologues (28)
- Paralogues (0)
- Protein families (1)
- Genetic Variation
  - Variation Table
  - **Variation Image**
- External Data
- ID History
  - Gene history

Configure this page
Add custom data to page
Export data
Bookmark this page

Gene: BRCA1 (ENSG00000012048)

Breast cancer type 1 susceptibility protein (RING finger protein 53) Source: UniProtKB/Swiss-Prot P38398

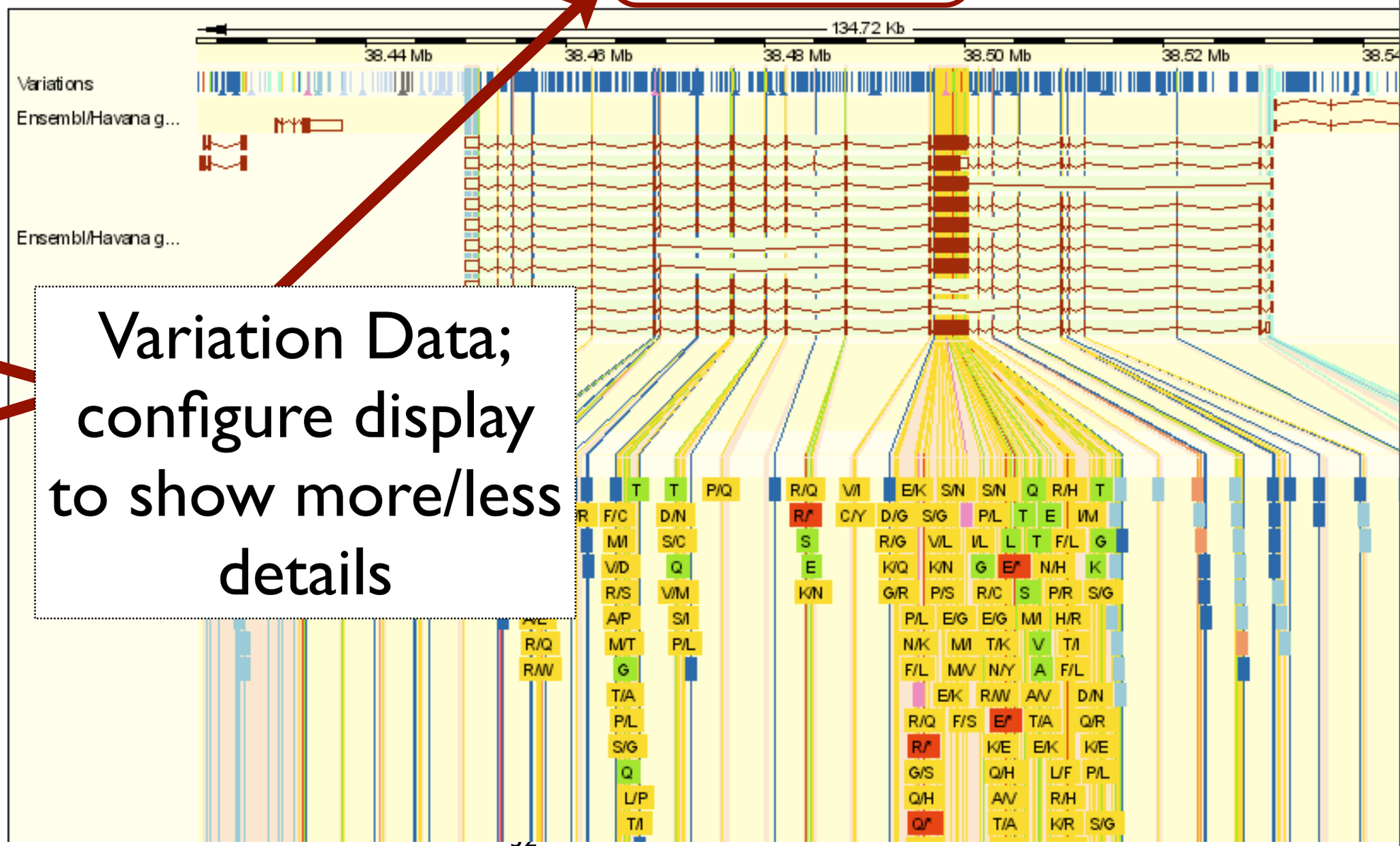
Location Chromosome 17: 38,449,840-38,530,994 reverse strand.

Transcripts There are 10 transcripts in this gene: **show transcripts**

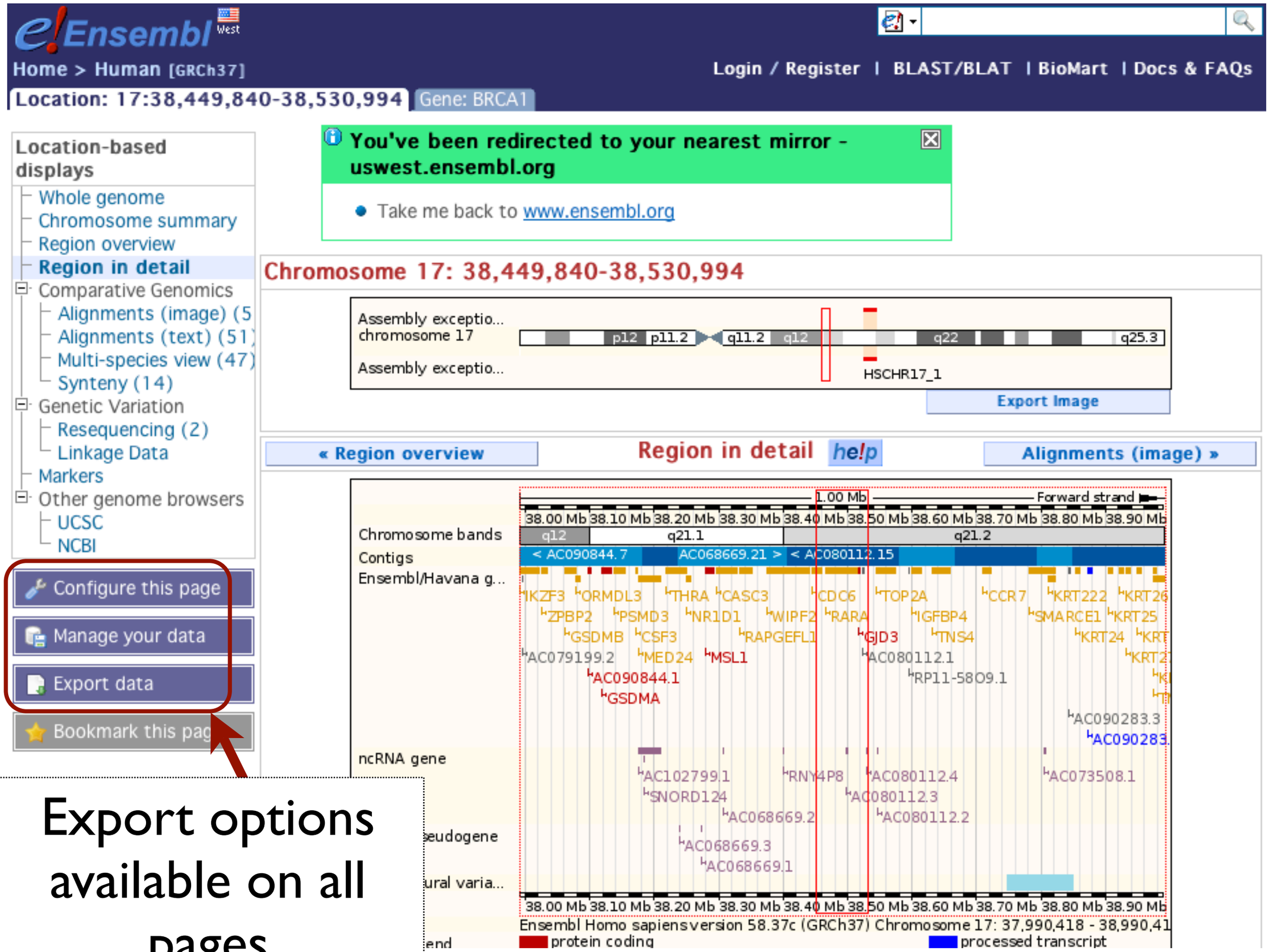
« Variation Table

Variation Image help

Ext



Variation Data; configure display to show more/less details



Export options available on all pages

http://www.ncbi.nlm.nih.gov/mapview/

NCBI Home GenBank BLAST

Map Viewer Home | Help

The Map Viewer provides a wide variety of genome mapping and sequencing data. More..

Search

Search: Select Group or Organism

for:

Go

Tools Legend

- Search or Browse the Genome
- BLAST
- Clone Finder
- Go to region on a chromosome
- Genome Resources page

News

Human build 37 released Aug 2, 200

An update to the human genome assembly and annotation is now... more

Annotation update released for human genome build 36 Mar 23, 2008

An annotation update for the human genome (NCBI Build 36.3) ... more

Show all

Related Resources

NCBI Home
NCBI Web Search
NCBI Site map
Genome Browser agreement
Genome Biology
Taxonomy
Entrez (Global Query)
BLAST

Vertebrates (17)

Mammals (14)

Primates (3)

Scientific name	Common name	Build	Tools
Homo sapiens	human	Build 37.1	Search, BLAST, Region, Clone Finder, Genome
		Build 36.3	Search, BLAST, Region, Clone Finder
Macaca mulatta	rhesus macaque	Build 1.1	Search, BLAST, Region, Genome
Pan troglodytes	chimpanzee	Build 2.1	Search, BLAST, Region, Genome

Rodents (2)

Scientific name	Common name	Build	Tools
Mus musculus	laboratory mouse	Build 37.1	Search, BLAST, Region, Clone Finder, Genome
		Build 36.1	Search, BLAST, Region
Rattus norvegicus	rat	RGSC v3.4	Search, BLAST, Region, Genome

Monotremes (1)

Marsupials (1)

Other Mammals (7)

Other Vertebrates (3)

Invertebrates (12)

Protozoa (18)

Plants (45)

Fungi (17)

Scientific name		
Aspergillus clavatus		Genome
Aspergillus fumigatus		Genome
Aspergillus niger		Genome
Candida glabrata	Build 1.1	Search, BLAST, Region
Cryptococcus neoformans	Build 2.1	Search, BLAST, Region
Debaryomyces hansenii	Build 1.1	Search, BLAST, Region
Encephalitozoon cuniculi	Build 1.1	Search, BLAST, Region
Eremothecium gossypii	Build 3.1	Search, BLAST, Region

Two builds of human; Note many genomes available

NCBI

NCBI Map Viewer

PubMed Nucleotide Protein Genome Gene Structure PopSet Taxonomy Help

Search for BRCA1 on chromosome(s) 17 assembly All Find Advanced Search

Map Viewer
- Map Viewer Home
- Map Viewer Help
- Human Maps Help
- Release Notes

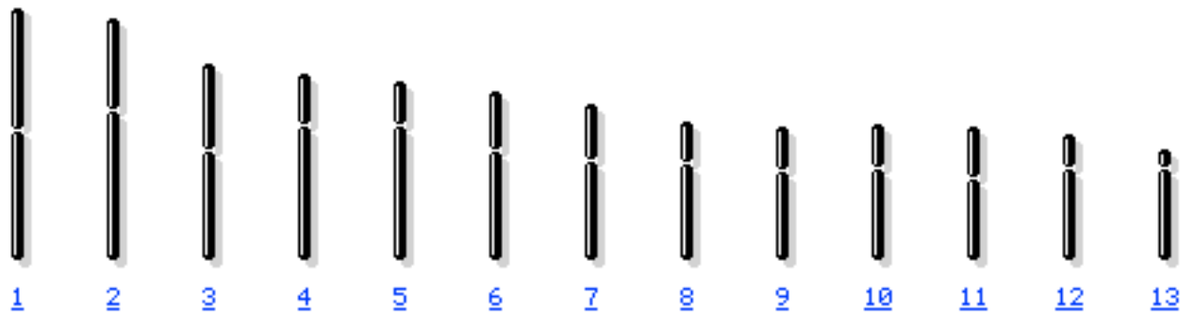
NCBI Resources
- Genome Project
- TaxPlot
- Consensus CoDing Sequence (CCDS)
- Human Genome Resources
- NCBI Handbook
- RefSeq
- Whole Genome Association (WGA)

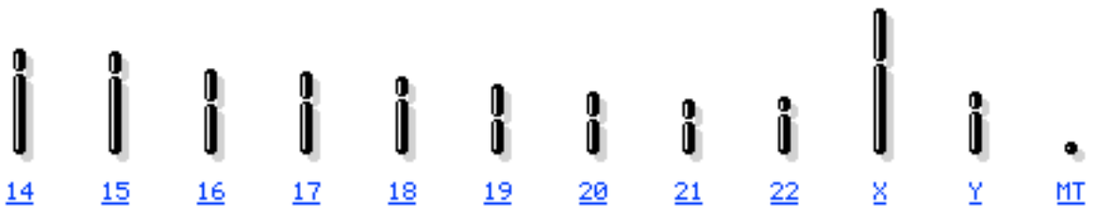
Organism Data in GenBank
- EST
- Genomic
- mRNA
- Protein

Homo sapiens (human) genome view

Build 36.2 statistics **Switch to previous build**

BLAST search the human genome





Lineage: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo; Homo sapiens

September 2006: NCBI released an annotation update for the human genome (NCBI Build 36.2); this update does not change the genome assembly. The previous version of the genome assembly, NCBI Build 35.1, can still be accessed for Map Viewer display and for BLAST. For additional information about changes, statistics, and the status of the CCDS project please refer to:

- Release Notes
- Statistics
- CCDS Project

The NCBI Map Viewer provides graphical displays of features on the human genome sequence assembly as well as

NCBI Map Viewer

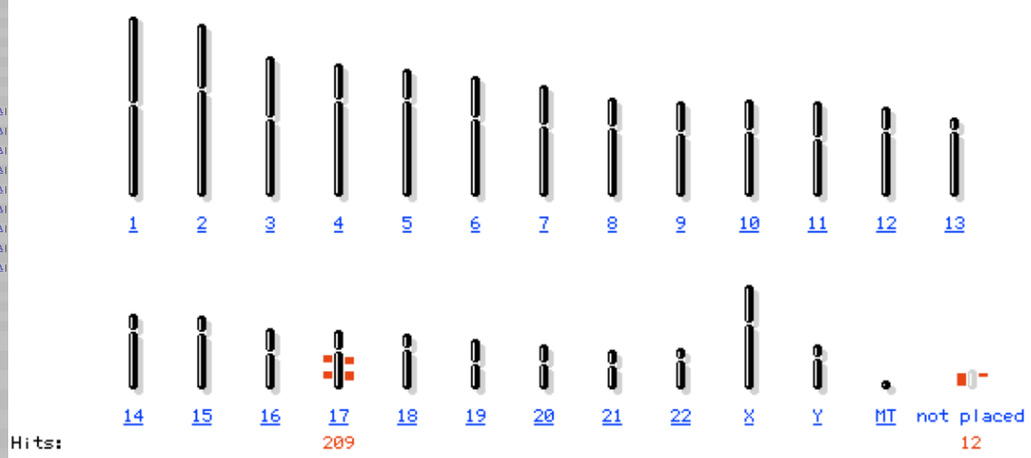
PubMed | Nucleotide | Protein | Genome | Gene | Structure | PopSet | Taxonomy | Help

Search for BRCA1 on chromosome(s) 17 assembly All Find Advanced Search

Homo sapiens (human) genome view

Build 36.2 statistics Switch to previous build

BLAST search the human genome



Hits:

Search results for query "BRCA1": 221 hits

Hits shown: 1 - 100 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y MT not placed

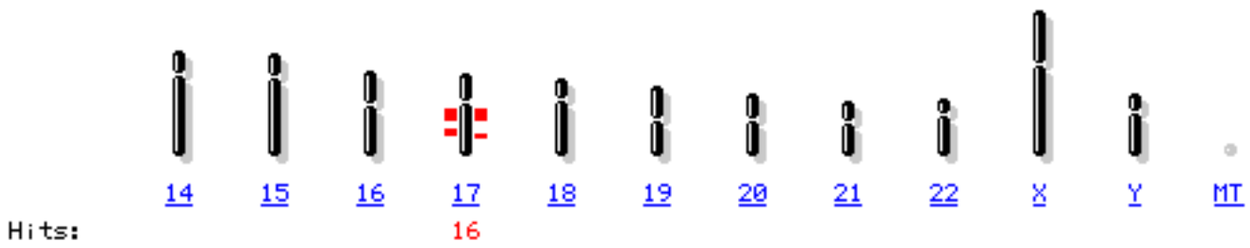
Chr	Assembly	Match	Map element	Type	Maps
17	reference	all matches			
		Neighbor of **Brca1** gene 1	Rn.94975	Rn_EST_Cl	Rn UniG
		BRCA1 interacting protein C-terminal helicase 1	Mm.186143	Mm_EST_Cl	Mm UniG
		Neighbor of **Brca1** gene 1	Mm.784	Mm_EST_Cl	Mm UniG
		Neighbor of **BRCA1** gene 2 (9 hits)	Hs.559259	Hs_EST_Cl	Hs UniG
		Neighbor of **BRCA1** gene 1 (2 hits)	Hs.546264	Hs_EST_Cl	Hs UniG
		BRCA1 interacting protein C-terminal helicase 1	Hs.532799	Hs_EST_Cl	Hs UniG
		Neighbor of **BRCA1** gene 1	Hs.373818	Hs_EST_Cl	Hs UniG
		Neighbor of **BRCA1** gene 1 (2 hits)	Hs.277721	Hs_EST_Cl	Hs UniG
		BRCA1 interacting protein C-terminal helicase 1	Gga.17801	Gga_EST_Cl	Gga UniG

Quick Filter
- ☑ Gene

Transcript :
- ☐ all
- ☐ RefSeq
- ☐ STS
- ☐ Unigene

Filter

Quick Filter
✓ Gene



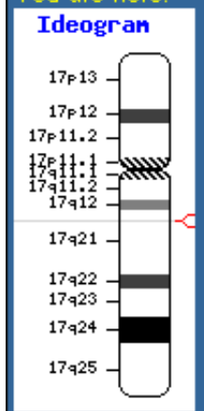
Search results for query "brca1 AND gene[obj_type]": 16 hits

Chr	Assembly	Match	Map element	Type	Maps
17	reference	all matches			
		BRCA1P1 : BRCA1 pseudogene 1	BRCA1P1	GENE	Genes_cyto | Genes_seq
		BRCA1 interacting protein C-terminal helicase 1	BRIP1	GENE	Genes_cyto | Genes_seq
		neighbor of **BRCA1** gene 2	NBR2	GENE	Genes_cyto | Genes_seq
		neighbor of **BRCA1** gene 1	NBR1	GENE	Genes_cyto | Genes_seq
		BRCA1 : breast cancer 1, early onset	BRCA1	GENE	Genes_cyto | Genes_seq
		BRCA1 : ENSG00000012048	BRCA1	GENE	ensGenes
17	Celera	all matches			
		BRCA1P1 : BRCA1 pseudogene 1	BRCA1P1	GENE	Genes_seq
		BRCA1 interacting protein C-terminal helicase 1	BRIP1	GENE	craGenes | Genes_seq
		neighbor of **BRCA1** gene 2	NBR2	GENE	craGenes | Genes_seq
		neighbor of **BRCA1** gene 1	NBR1	GENE	craGenes | Genes_seq
		BRCA1 : breast cancer 1, early onset	BRCA1	GENE	craGenes | Genes_seq
17	HuRef	all matches			
		BRCA1P1 : BRCA1 pseudogene 1	BRCA1P1	GENE	Genes_seq
		BRCA1 interacting protein C-terminal helicase 1	BRIP1	GENE	Genes_seq
		neighbor of **BRCA1** gene 2	NBR2	GENE	Genes_seq
		neighbor of **BRCA1** gene 1	NBR1	GENE	Genes_seq
		BRCA1 : breast cancer 1, early onset	BRCA1	GENE	Genes_seq

Human genome overview page (Build 36.2)
Human genome overview page (Build 35.1)
Map Viewer Home

Map Viewer Help
Human Maps Help
FTP
Data As Table View
Maps & Options
Compress Map
Region Shown:
38,389K
38,592K
Go

You are here:



default
master

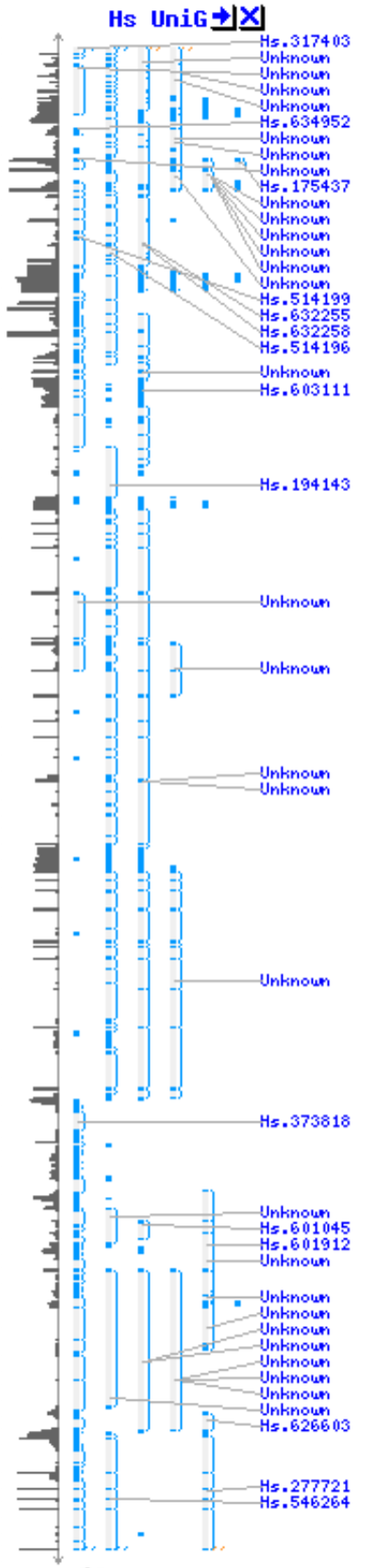
Master Map: Genes On Sequence

Summary of Maps

Maps & Optio

Region Displayed: **38,389K-38,592K bp**

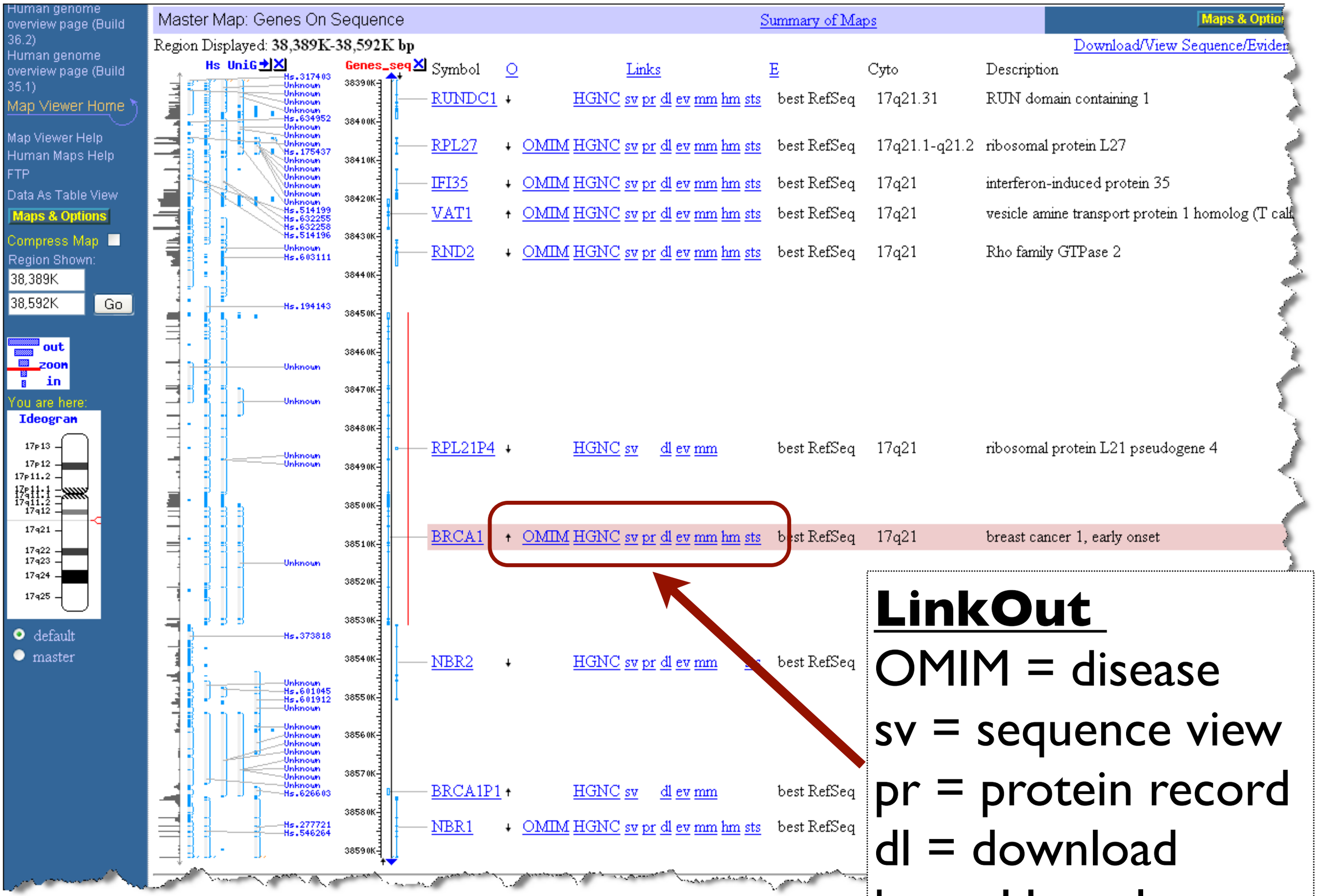
Download/View Sequence/Evider



Symbol	O	Links	E	Cyto	Description
RUNDC1	↓	HGNC sv pr dl ev mm hm sts	best RefSeq	17q21.31	RUN domain containing 1
RPL27	↓	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21.1-q21.2	ribosomal protein L27
IFI35	↓	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21	interferon-induced protein 35
VAT1	↑	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21	vesicle amine transport protein 1 homolog (T cali
RND2	↓	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21	Rho family GTPase 2
RPL21P4	↓	HGNC sv dl ev mm	best RefSeq	17q21	ribosomal protein L21 pseudogene 4
BRCA1	↑	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21	breast cancer 1, early onset
NBR2	↓	HGNC sv pr dl ev mm sts	best RefSeq	17q21	neighbor of BRCA1 gene 2
BRCA1P1	↑	HGNC sv dl ev mm	best RefSeq	17q21	BRCA1 pseudogene 1
NBR1	↓	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21.31	neighbor of BRCA1 gene 1

Two tasks

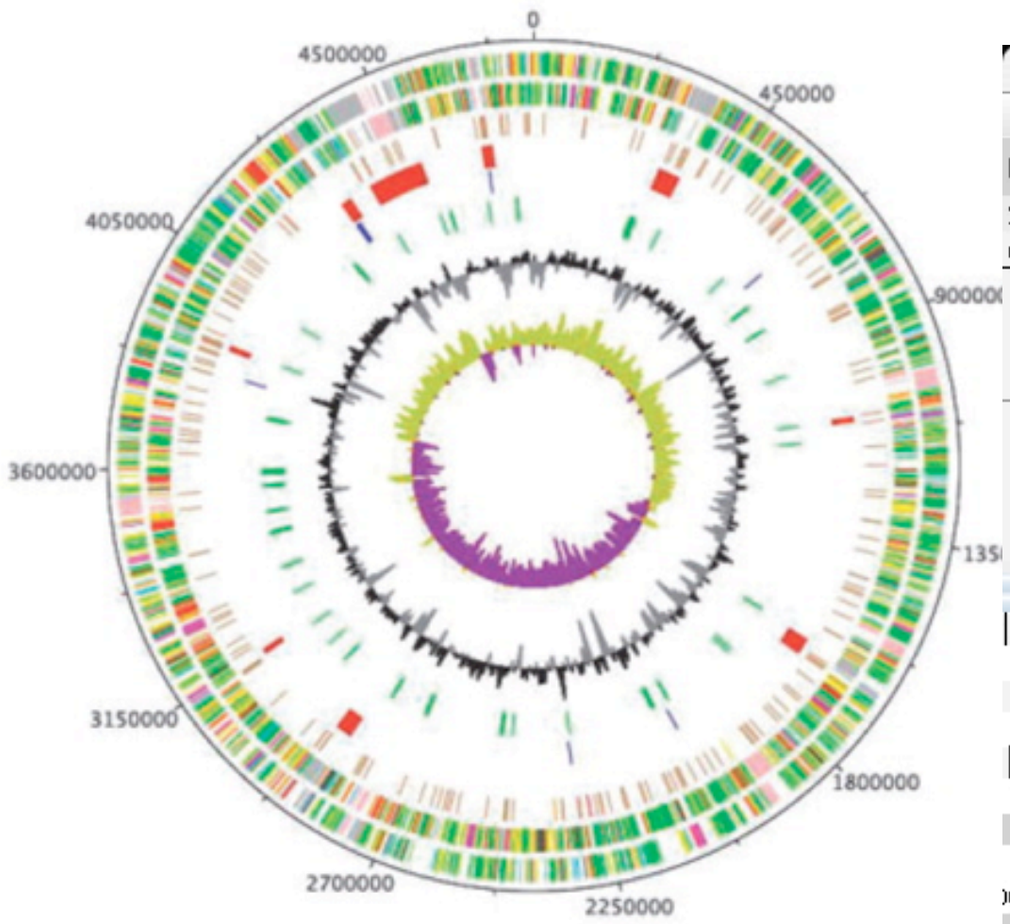
- Can you figure out how to LinkOut to the OMIM and/or Homologene entries for BRCA1?
- Can you figure out how to download the genomic sequence for the BRCA1 region?


Human genome overview page (Build 36.2)
Human genome overview page (Build 35.1)
Map Viewer Home
Map Viewer Help
Human Maps Help
FTP
Data As Table View
Maps & Options
Compress Map
Region Shown:
38,389K
38,592K
Go
out
zoom
in
You are here:
Ideogram
17p13
17p12
17p11.2
17p11.1
17q11.1
17q11.2
17q12
17q21
17q22
17q23
17q24
17q25
default
master
Master Map: Genes On Sequence
Summary of Maps
Maps & Options
Region Displayed: 38,389K-38,592K bp
Download/View Sequence/Evidence
Hs UniG
Genes_seq
Symbol
O
Links
E
Cyto
Description
RUNDC1 HGNC sv pr dl ev mm hm sts best RefSeq 17q21.31 RUN domain containing 1
RPL27 OMIM HGNC sv pr dl ev mm hm sts best RefSeq 17q21.1-q21.2 ribosomal protein L27
IFI35 OMIM HGNC sv pr dl ev mm hm sts best RefSeq 17q21 interferon-induced protein 35
VAT1 OMIM HGNC sv pr dl ev mm hm sts best RefSeq 17q21 vesicle amine transport protein 1 homolog (T cali
RND2 OMIM HGNC sv pr dl ev mm hm sts best RefSeq 17q21 Rho family GTPase 2
RPL21P4 HGNC sv dl ev mm best RefSeq 17q21 ribosomal protein L21 pseudogene 4
BRCA1 OMIM HGNC sv pr dl ev mm hm sts best RefSeq 17q21 breast cancer 1, early onset
NBR2 HGNC sv pr dl ev mm best RefSeq
BRCA1P1 HGNC sv dl ev mm best RefSeq
NBR1 OMIM HGNC sv pr dl ev mm hm sts best RefSeq
LinkOut
OMIM = disease
sv = sequence view
pr = protein record
dl = download
hm = Homologene

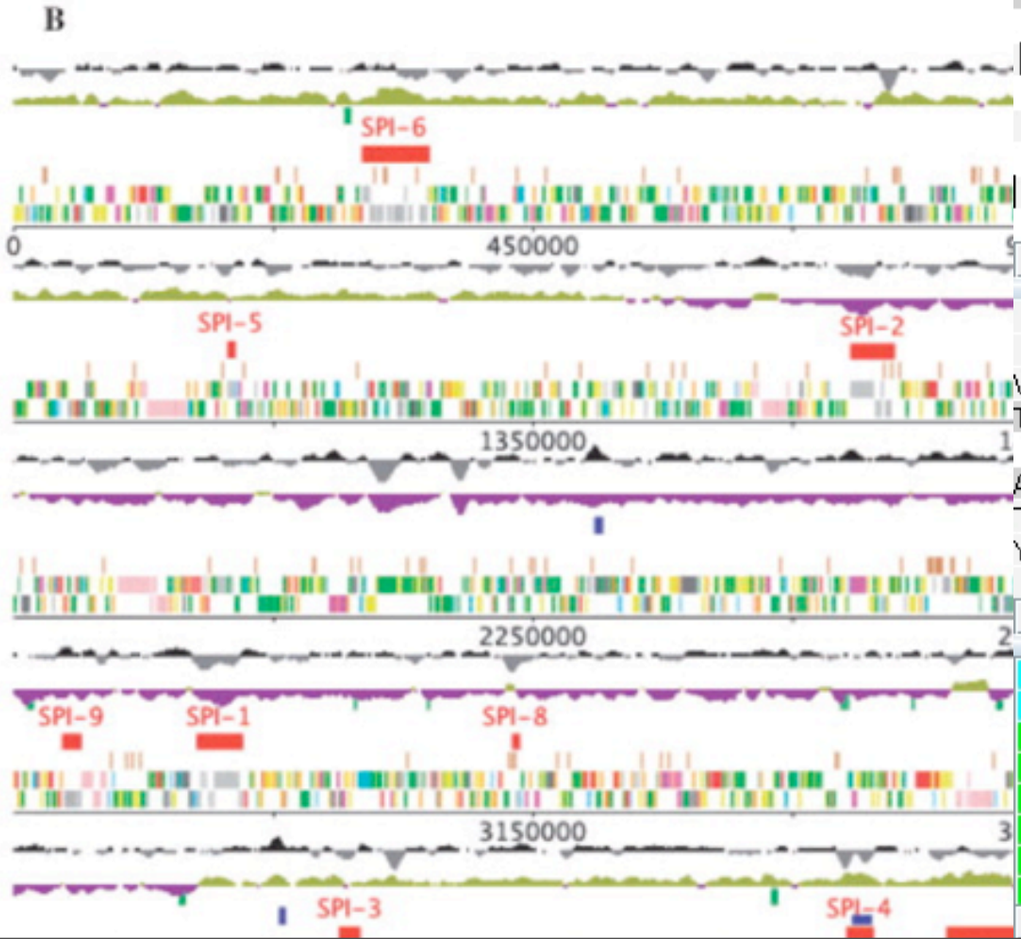
Artemis: DNA Sequence Viewer and Annotation Tool

- Artemis is a free genome viewer and annotation tool that allows visualisation of sequence features and the results of analyses within the context of the sequence, and also its six-frame translation.
- http://www.sanger.ac.uk/resources/software/artemis/

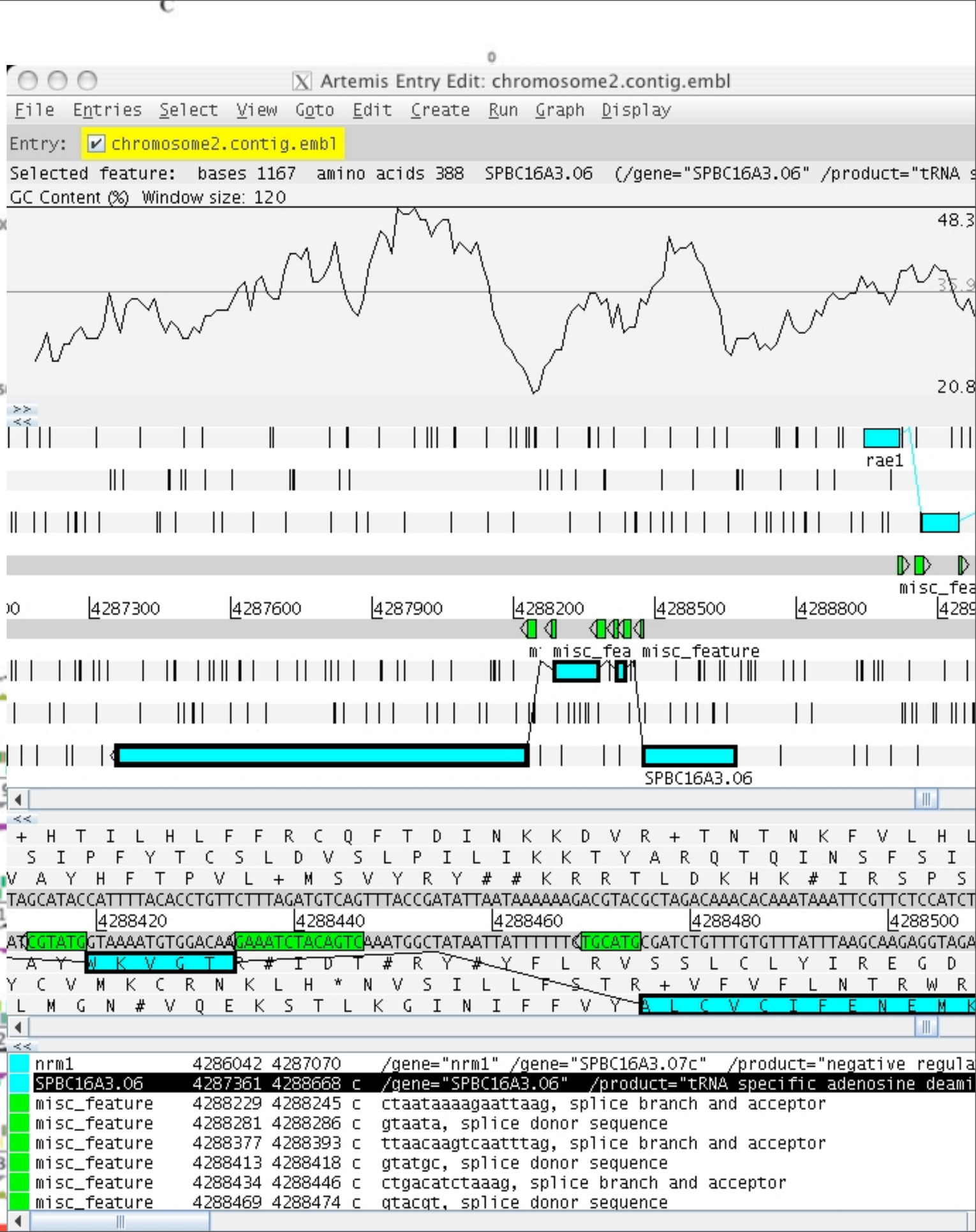
A


0
450000
900000
1350000
1800000
2250000
2700000
3150000
3600000
4050000
4500000

B


SPI-6
0
450000
SPI-5
SPI-2
1350000
2250000
SPI-9
SPI-1
SPI-8
3150000
SPI-3
SPI-4

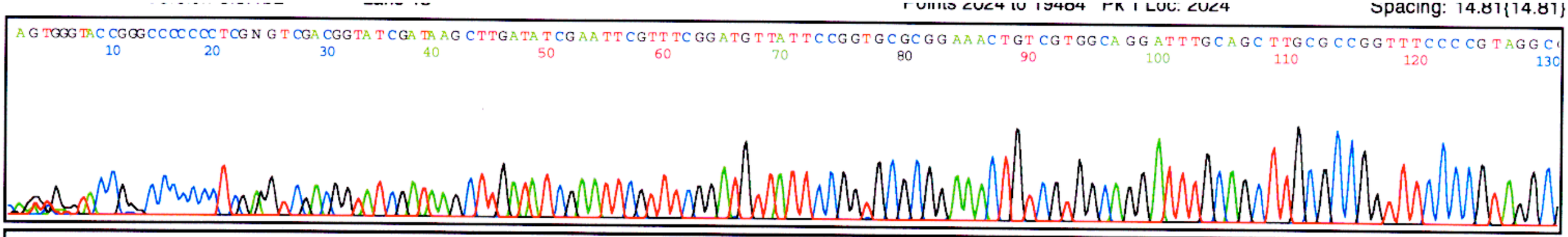
C

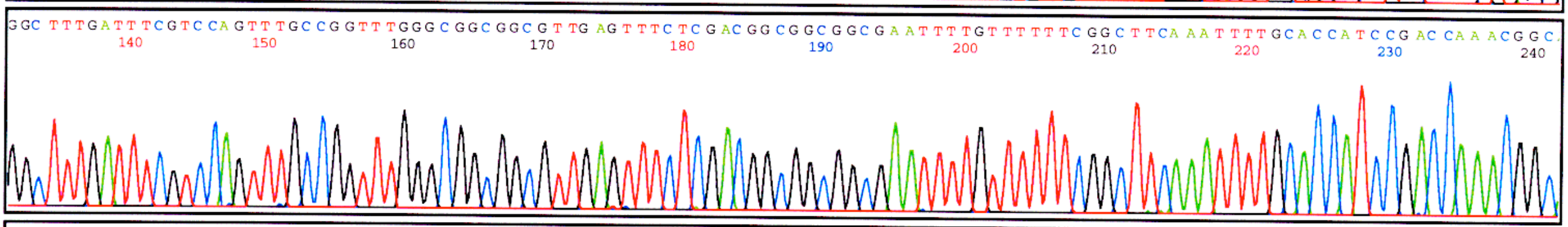

Artemis Entry Edit: chromosome2.contig.embl
File Entries Select View Goto Edit Create Run Graph Display
Entry: chromosome2.contig.embl
Selected feature: bases 1167 amino acids 388 SPBC16A3.06 (/gene="SPBC16A3.06" /product="tRNA
GC Content (%) Window size: 120
48.3
20.8
rae1
misc_fea
4287300
4287600
4287900
4288200
4288500
4288800
misc_feature
SPBC16A3.06
4288420
4288440
4288460
4288480
4288500
nrm1 4286042 4287070 /gene="nrm1" /gene="SPBC16A3.07c" /product="negative regula
SPBC16A3.06 4287361 4288668 c /gene="SPBC16A3.06" /product="tRNA specific adenosine deami
misc_feature 4288229 4288245 c ctaataaaagaattaag, splice branch and acceptor
misc_feature 4288281 4288286 c gtaata, splice donor sequence
misc_feature 4288377 4288393 c ttaacaagtcaatttag, splice branch and acceptor
misc_feature 4288413 4288418 c gtatgc, splice donor sequence
misc_feature 4288434 4288446 c ctgacatctaaag, splice branch and acceptor
misc_feature 4288469 4288474 c gtacgt, splice donor sequence

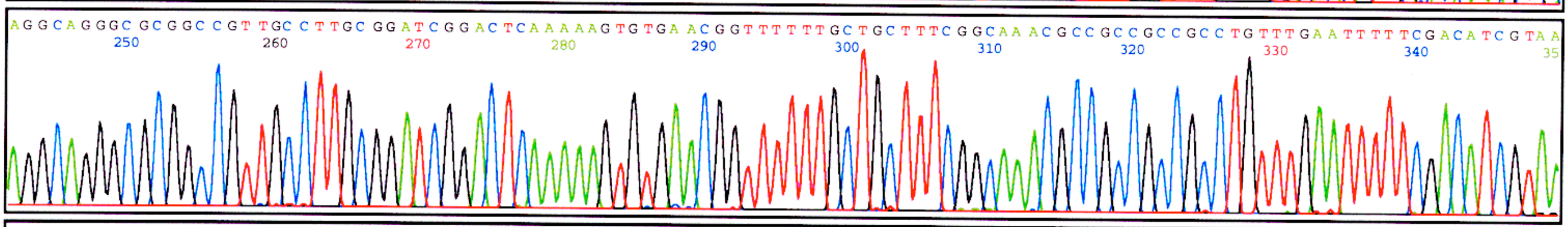
Bioinformatics

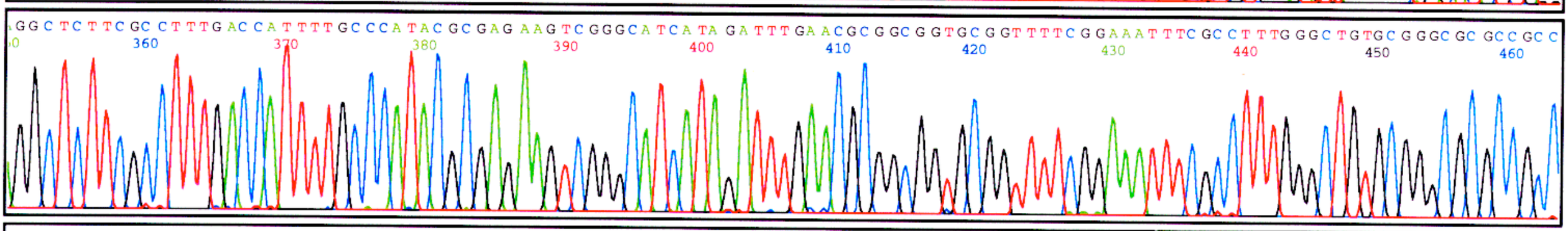
Session 3.2 - Pathway Resources for Systems Biology

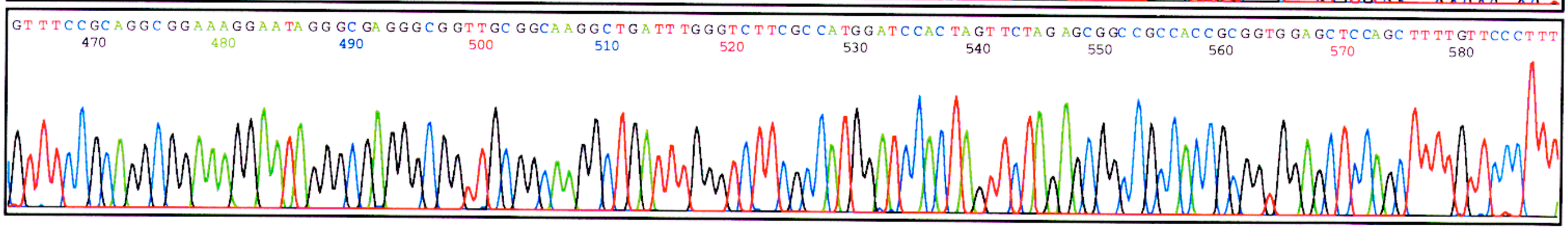
ggagcctcgg gaggtggtgg agtgacctgg ccccagtgct gcgtccttat cagccgagcc
ggtcccagct cttgctcctg cctgtttgcc tggaaatggc cacgcttctc cttctccttg
gggtgctggt ggtaagccca gacgctctgg ggagcacaac agcagtgcag acacccacct
ccggagagcc tttggtctct actagcgagc ccctgagctc aaagatgtac accacttcaa
taacaagtga ccctaaggcc gacagcactg gggaccagac ctcagcccta cctccctcaa
cttccatcaa tgagggatcc cctctttgga cttccattgg tgccagcact ggttcccctt
tacctgagcc aacaacctac caggaagttt ccatcaagat gtcatcagtg ccccaggaaa
cccctcatgc aaccagtcat cctgctgttc ccataacagc aaactctcta ggatcccaca
ccgtgacagg tggaaccata acaacgaact ctccagaaac ctccagtagg accagtggag
cccctgttac cacggcagct agctctctgg agacctccag aggcacctct ggacccccctc
ttaccatggc aactgtctct ctggagactt ccaaaggcac ctctggaccc cctgttacca
tggcaactga ctctctggag acctccactg ggaccactgg acccccctgtt accatgacaa
ctggctctct ggagccctcc agcggggcca gtggacccca ggtctctagc gtaaaactat
ctacaatgat gtctccaacg acctccacca acgcaagcac tgtgcccttc cggaacccag
atgagaactc acgaggcatg ctgccagtgg ctgtgcttgt ggccctgctg gcggtcatag
tcctcgtggc tctgctcctg ctgtggcgcc ggcggcagaa gcggcggact ggggccctcg
tgctgagcag aggcggcaag cgtaacgggg tggtggacgc ctgggctggg ccagcccagg
tccctgagga gggggccgtg acagtgaccg tgggagggtc cgggggcgac aagggctctg
ggttccccga tggggagggg tctagccgtc ggcccacgct caccactttc tttggcagac
ctggctctct ggagccctcc agcggggcca gtggacccca ggtctctagc gtaaaactat
ctacaatgat gtctccaacg acctccacca acgcaagcac tgtgcccttc cggaacccag
atgagaactc acgaggcatg ctgccagtgg ctgtgcttgt ggccctgctg gcggtcatag


AGTGGGTACCGGGCCCCCCCTCGNGTCGACGGTATCGATAAGCTTGATATCGAATTCGTTTCGGATGTTATTCCGGTGCGCGGAAACTGTCGTGGCAGGATTTGCAGCTTGCGCCGGTTTCCCCGTAGGC
10 20 30 40 50 60 70 80 90 100 110 120 130


GGCTTTGATTTCGTCCAGTTTGCCGGTTTGGGCGGCGGCGTTGAGTTTCTCGACGGCGGCGGCGAATTTTGTTTTTTCGGCTTCAAATTTTGCACCATCCGACCAAACGGC
140 150 160 170 180 190 200 210 220 230 240


AGGCAGGGCGCGGCCGTTGCCTTGCGGATCGGACTCAAAAAGTGTGAACGGTTTTTTGCTGCTTTCGGCAAACGCCGCCGCCGCCTGTTTGAATTTTTCGACATCGTAA
250 260 270 280 290 300 310 320 330 340 350


GGCTCTTCGCCTTTGACCATTTTGCCCATACGCGAGAAGTCGGGCATCATAGATTTGAACGCGGCGGTGCGGTTTTCGGAAATTTCGCCTTTGGGCTGTGCGGGCGCGCCGCC
360 370 380 390 400 410 420 430 440 450 460


GTTTCCGCAGGCGGAAAGGAATAGGGCGAGGGCGGTTGCGGCAAGGCTGATTTGGGTCTTCGCCATGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTT
470 480 490 500 510 520 530 540 550 560 570 580

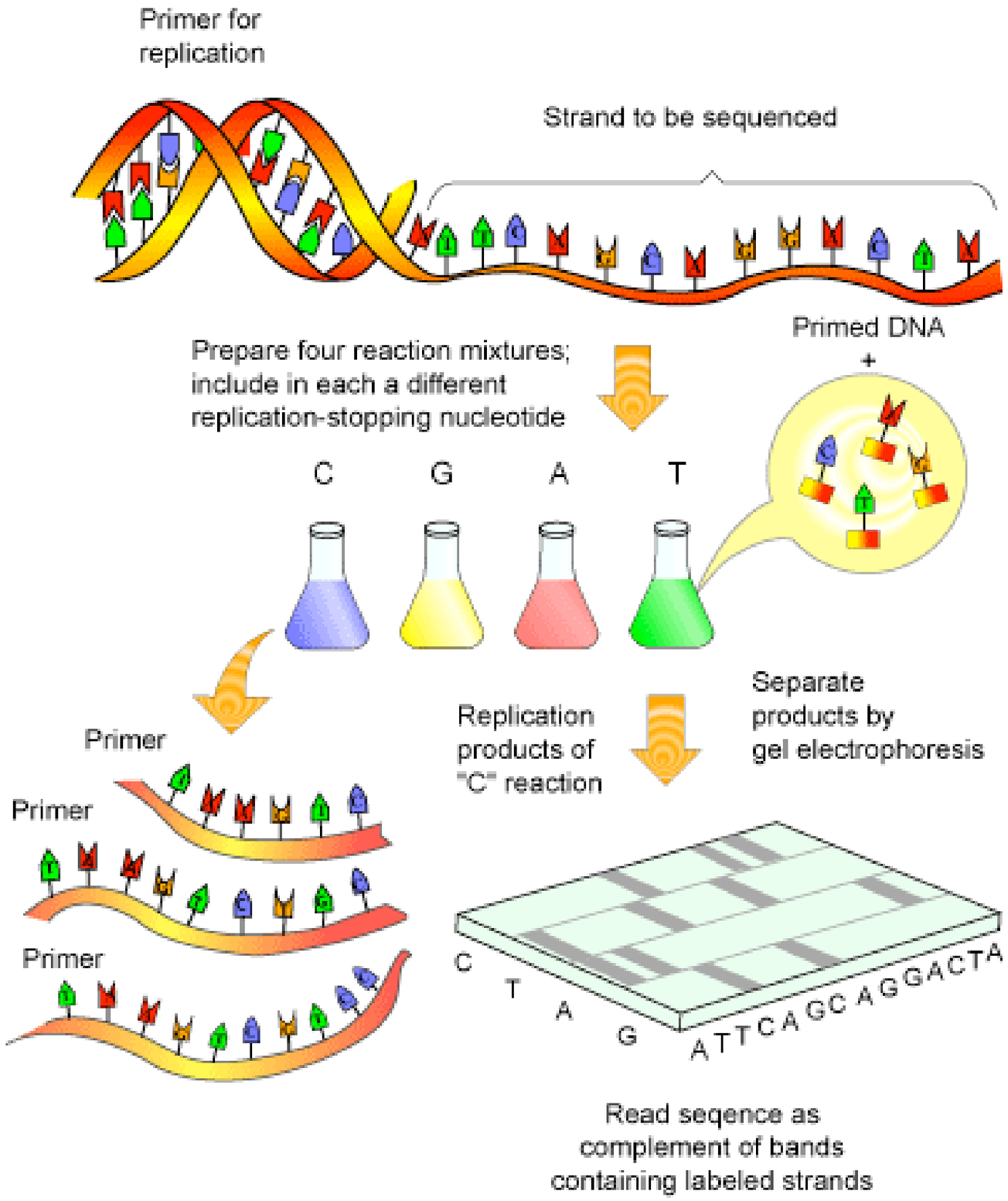
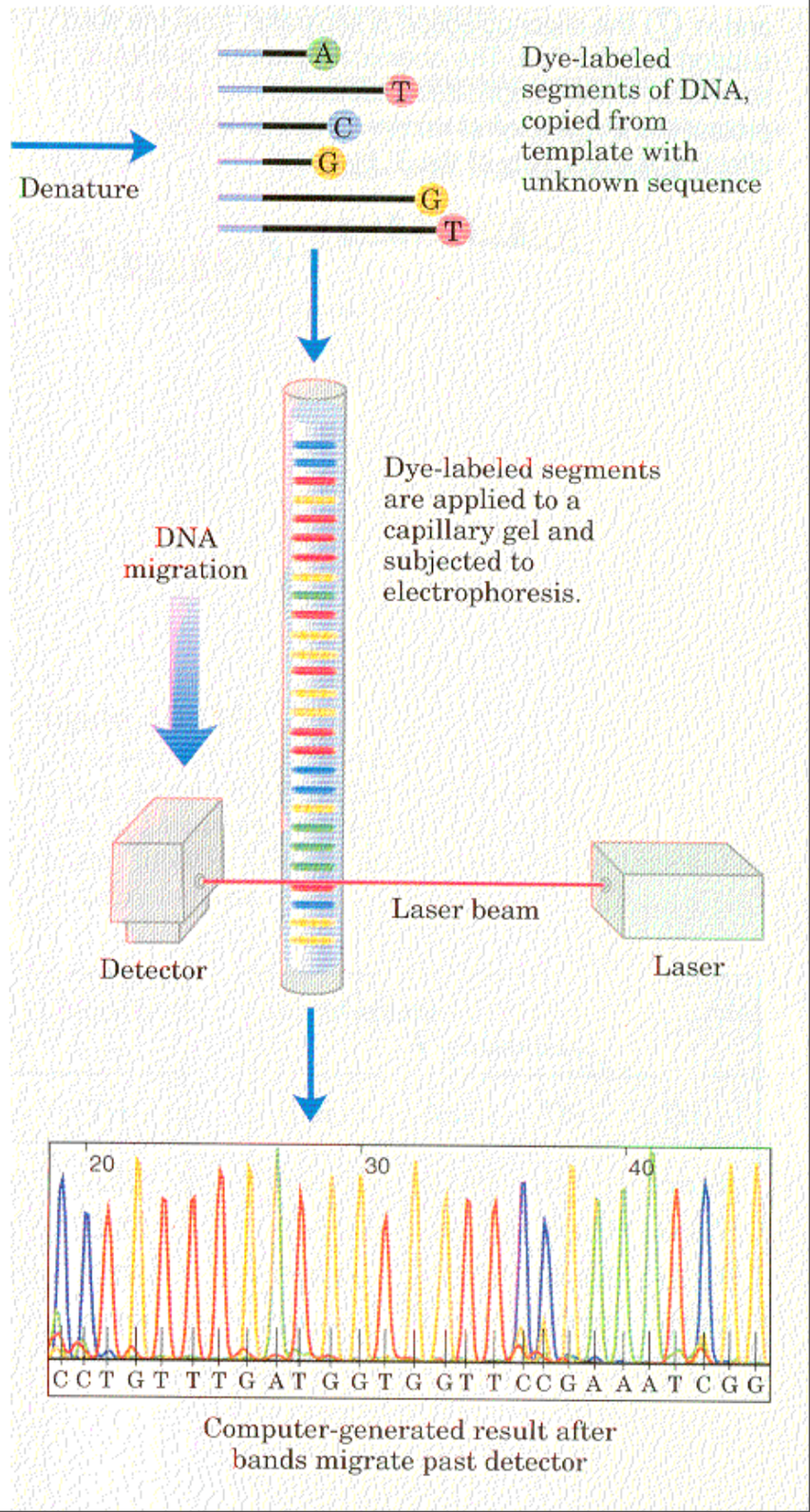


Figure 1. *The Sanger sequencing reaction. Single stranded DNA is amplified in the presence of fluorescently labelled ddNTPs that serve to terminate the reaction and label all the fragments of DNA produced. The fragments of DNA are then separated via polyacrylamide gel electrophoresis and the sequence read using a laser beam and computer.*

source: http://www.scq.ubc.ca/genome-projects-uncovering-the-blueprints-of-biology/

Nucleotide Sequencing

- Dideoxy termination (Sanger)
  - –Separate by gel or capillary electrophoresis
  - –Fluorescence detection
    - Rapid and reliable sequences
    - No information on chemical modifications



New Sequencing Technology

Every few years, a new technology comes along that dramatically changes how fundamental questions in biology are addressed. The impact of the technology is not always appreciated at first ...

- Stanley Fields

Solexa Technology

- DNA sequencing by synthesis
- approach built around very large number of short sequence reads

- key points:
  - ✓ solid phase amplification = no cloning necessary
  - ✓ reversible chemistry
  - ✓ data generated by imaging
  - ✓ read lengths 30-50 bp
  - ✓ < 1% cost, ultra high-throughput

10

After laser excitation, collect the image data as before. Record the identity of the second base for each cluster.

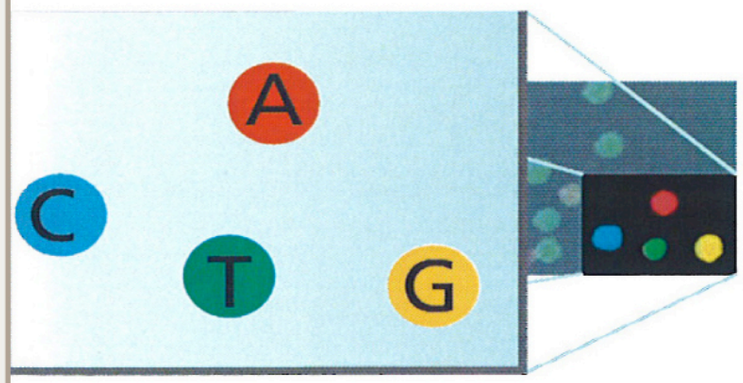
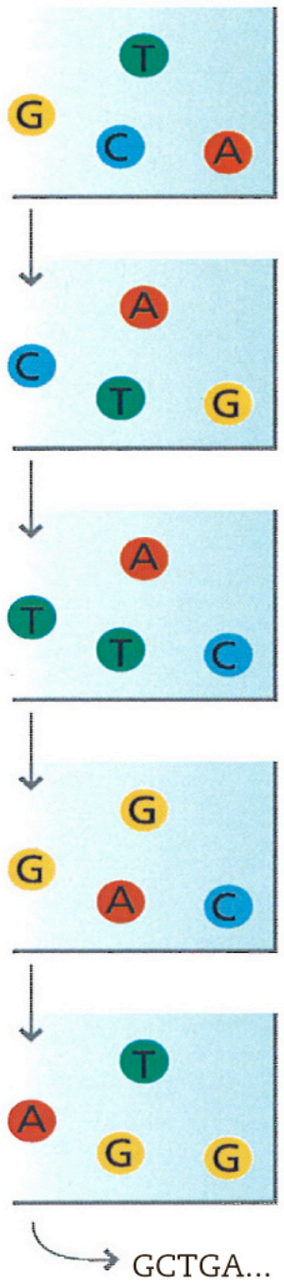


IMAGE OF SECOND CHEMISTRY CYCLE IS CAPTURED BY THE INSTRUMENT.

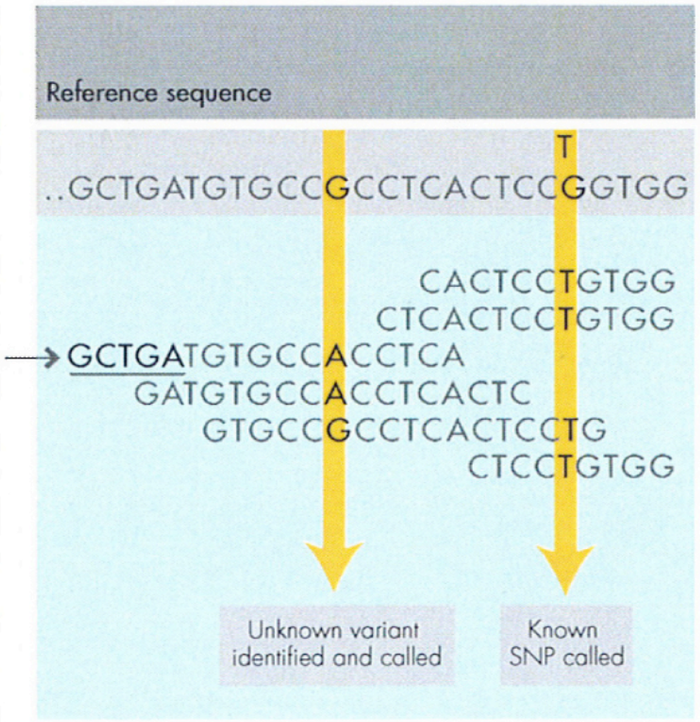
11

Repeat cycles of sequencing to determine the sequence of bases in a given fragment a single base at a time.



SEQUENCE READS OVER MULTIPLE CHEMISTRY CYCLES

12

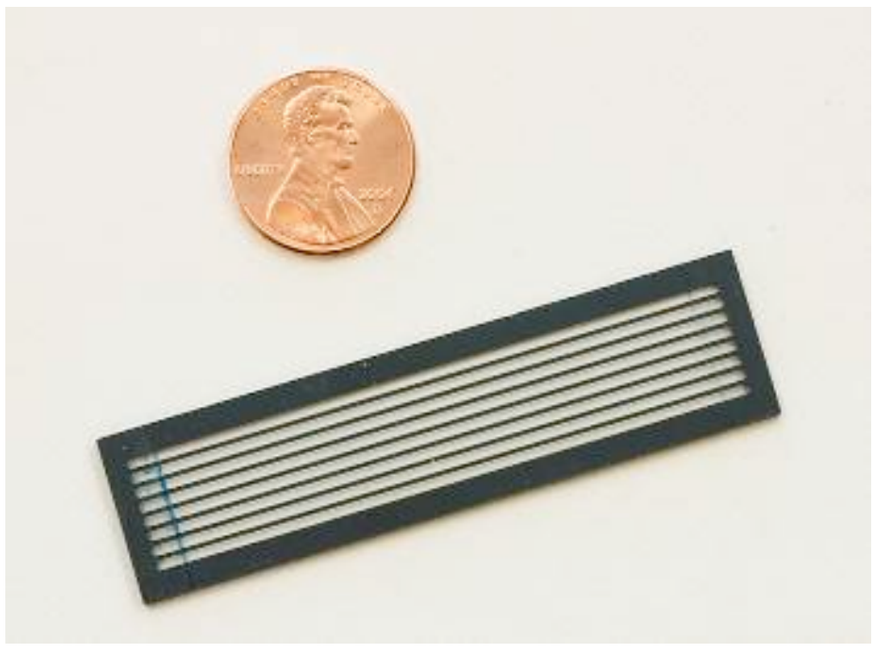


ALIGN THE NEW DATA TO A REFERENCE AND IDENTIFY SEQUENCE DIFFERENCES.

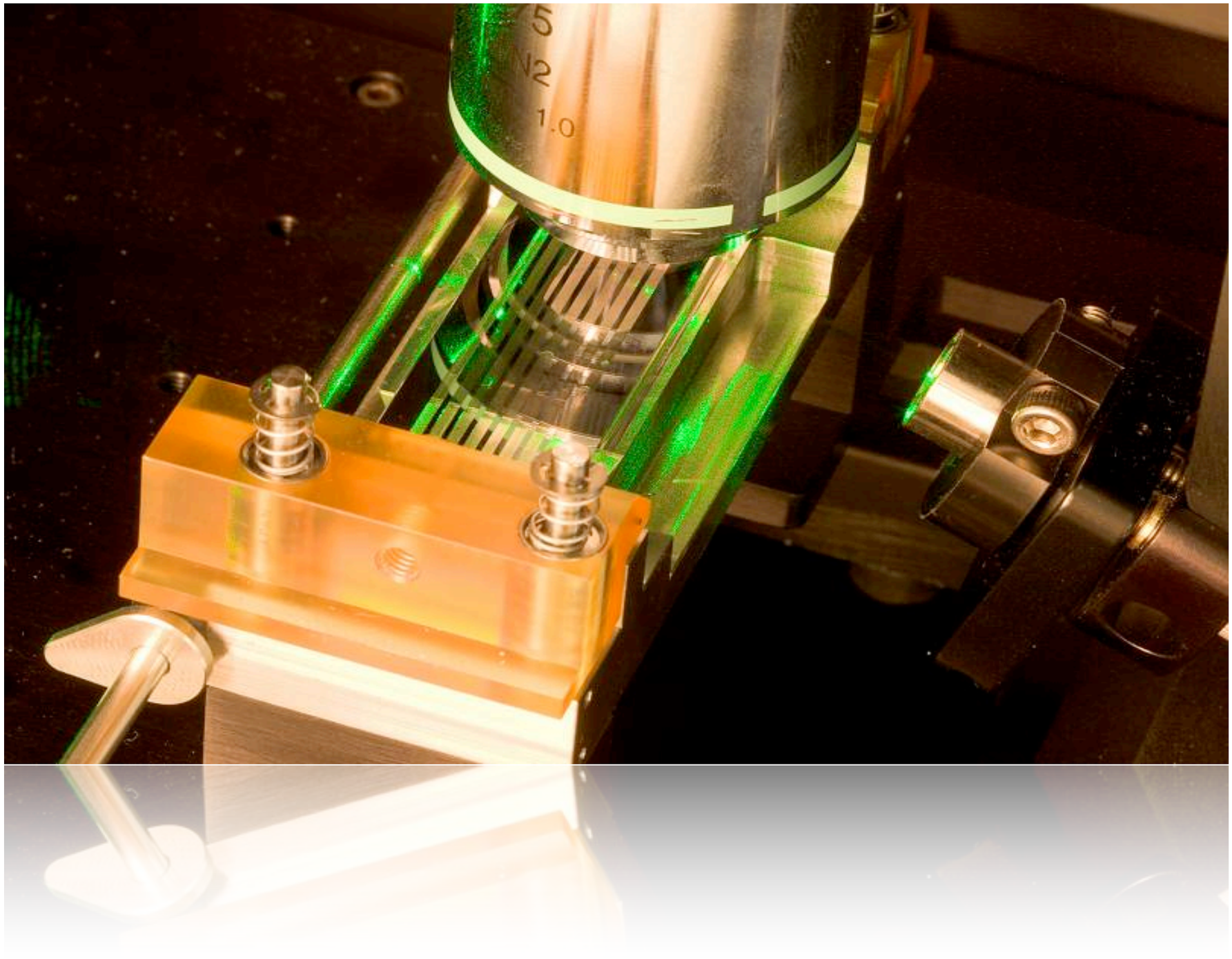
source: http://www.illumina.com/

Illumina/Solexa instrument

- Laser-based Optics
- 4-colour Detection
- CCD camera
- 8-channel flow cell
- 1 Gb / run at launch



source: Dr. Steven Jones, GSC



Data acquisition

source: Dr. Steven Jones, GSC

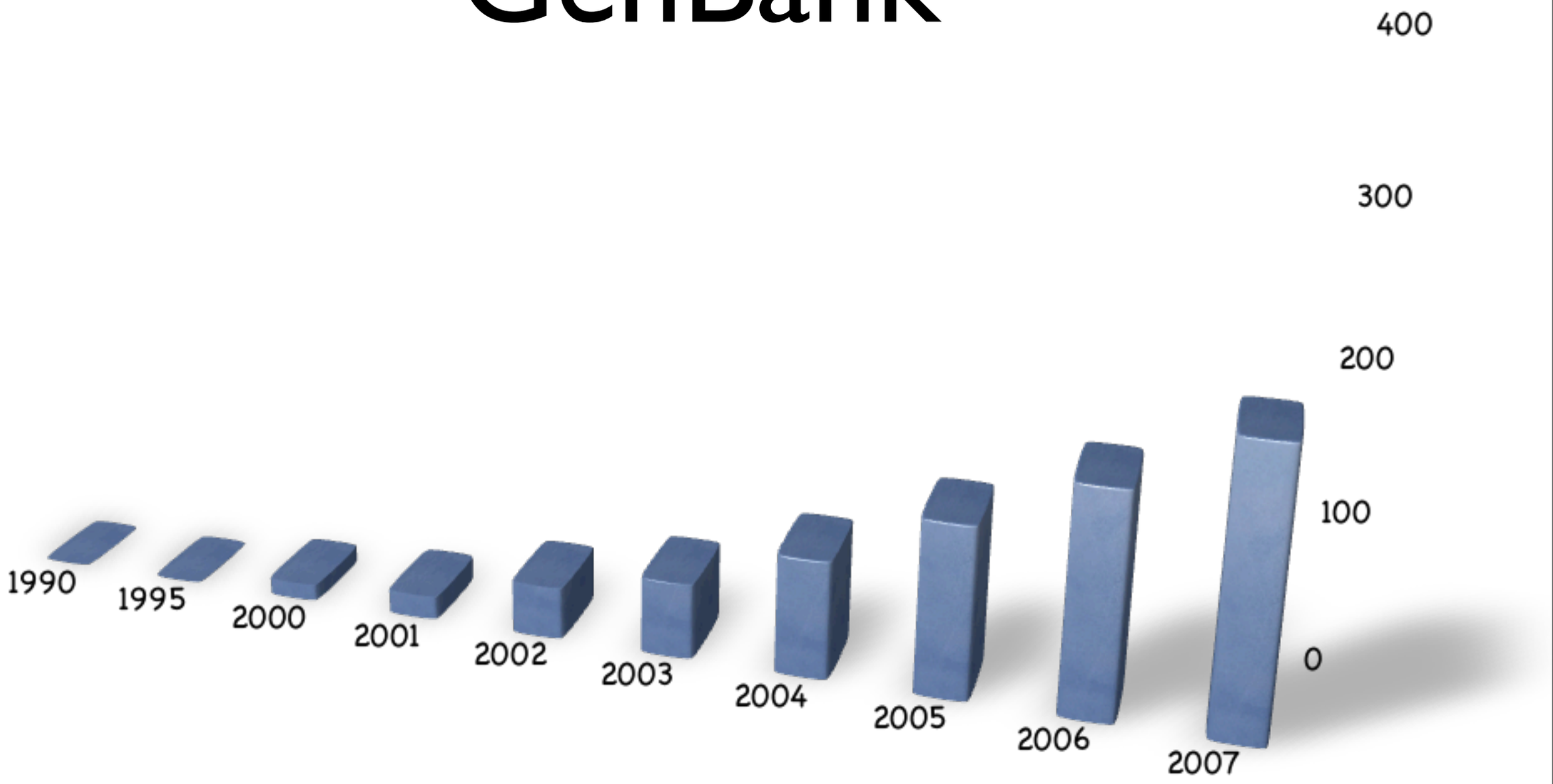
Clonal Single Molecule Array™ Technology

Up to 40 million clusters per flow cell

100 microns

20 microns

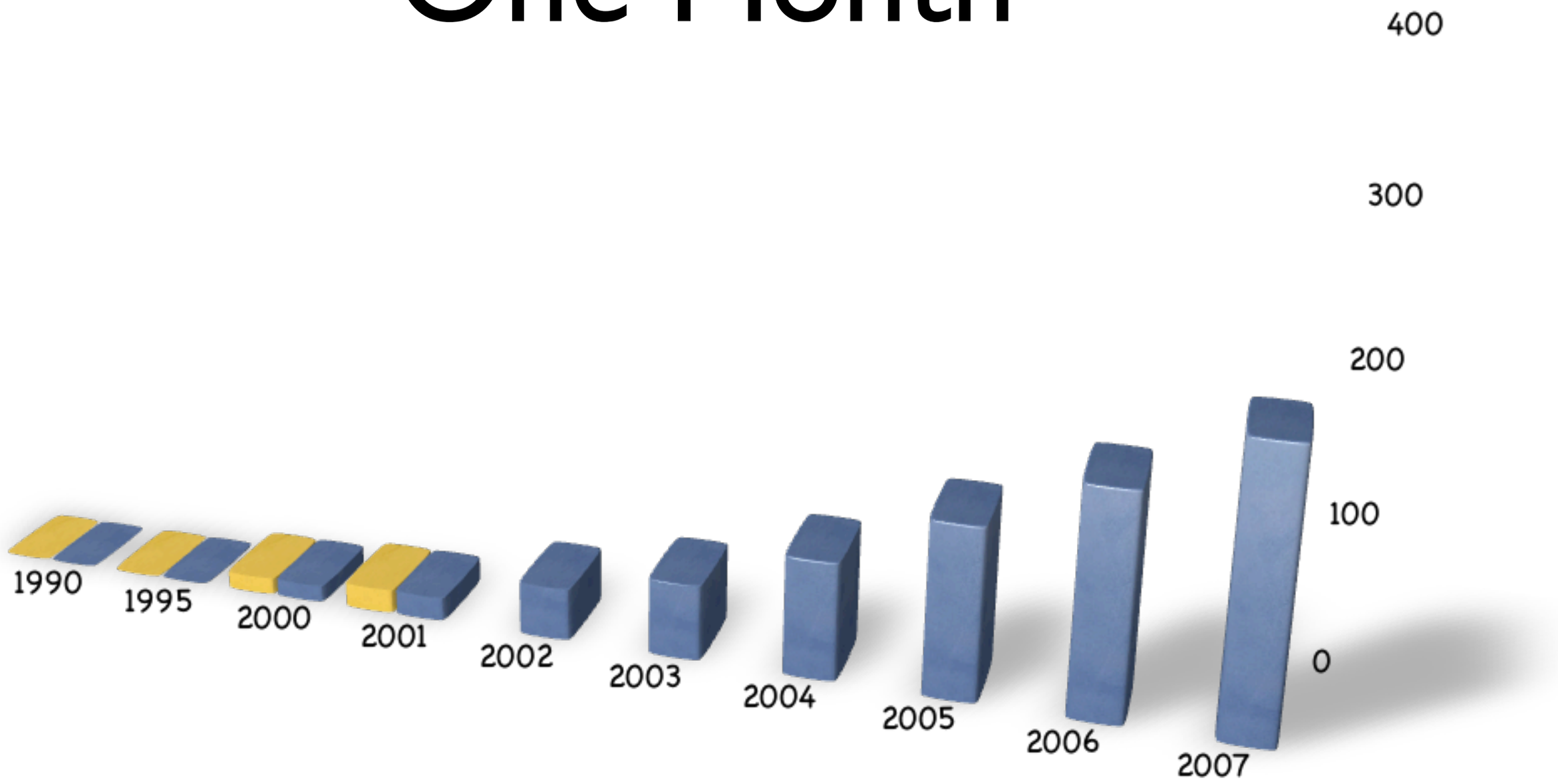
GenBank


400
300
200
100
0
1990
1995
2000
2001
2002
2003
2004
2005
2006
2007

Two Days

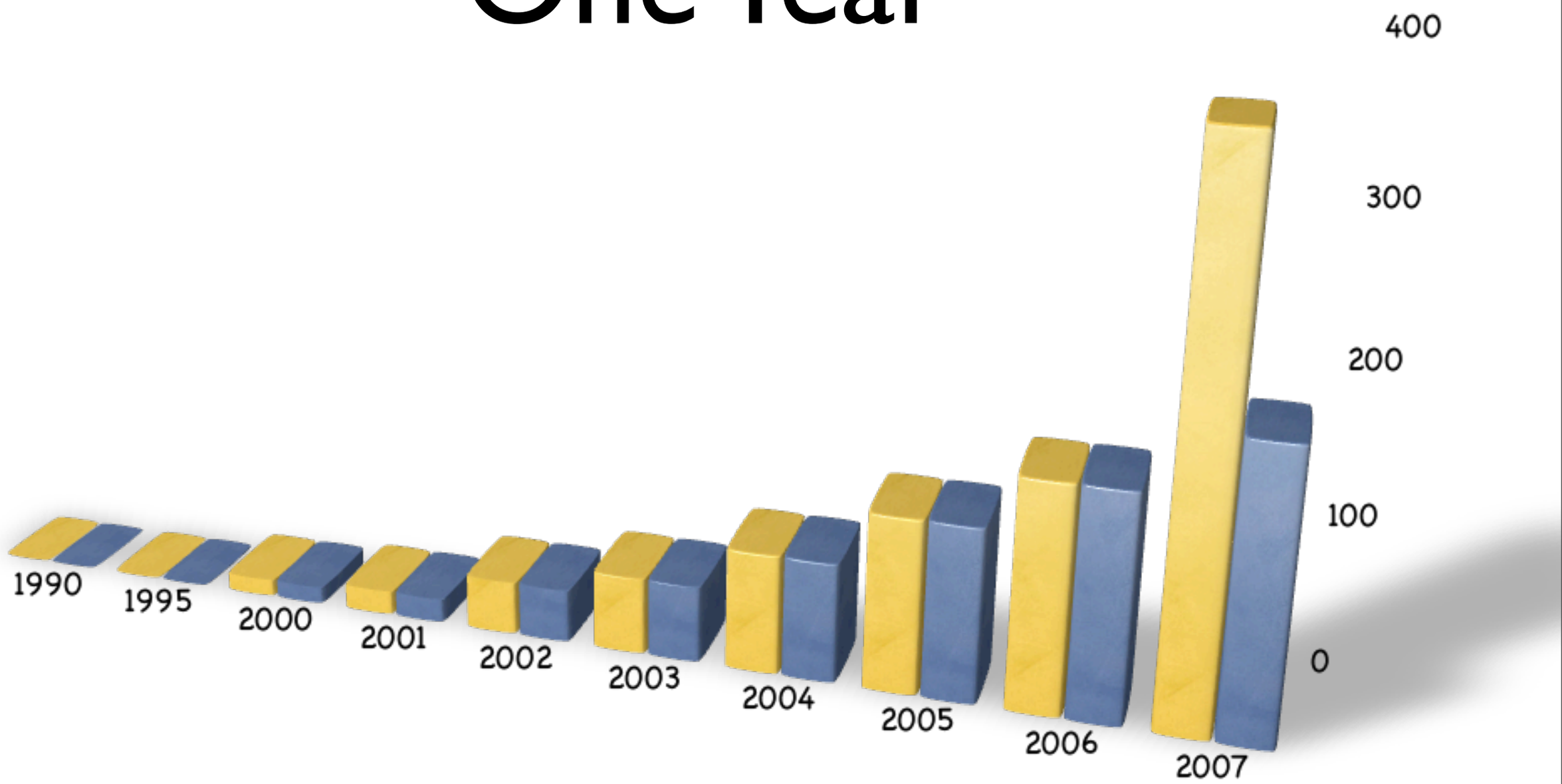
One machine, running for two days, can generate ~1 Gb data.

One Month

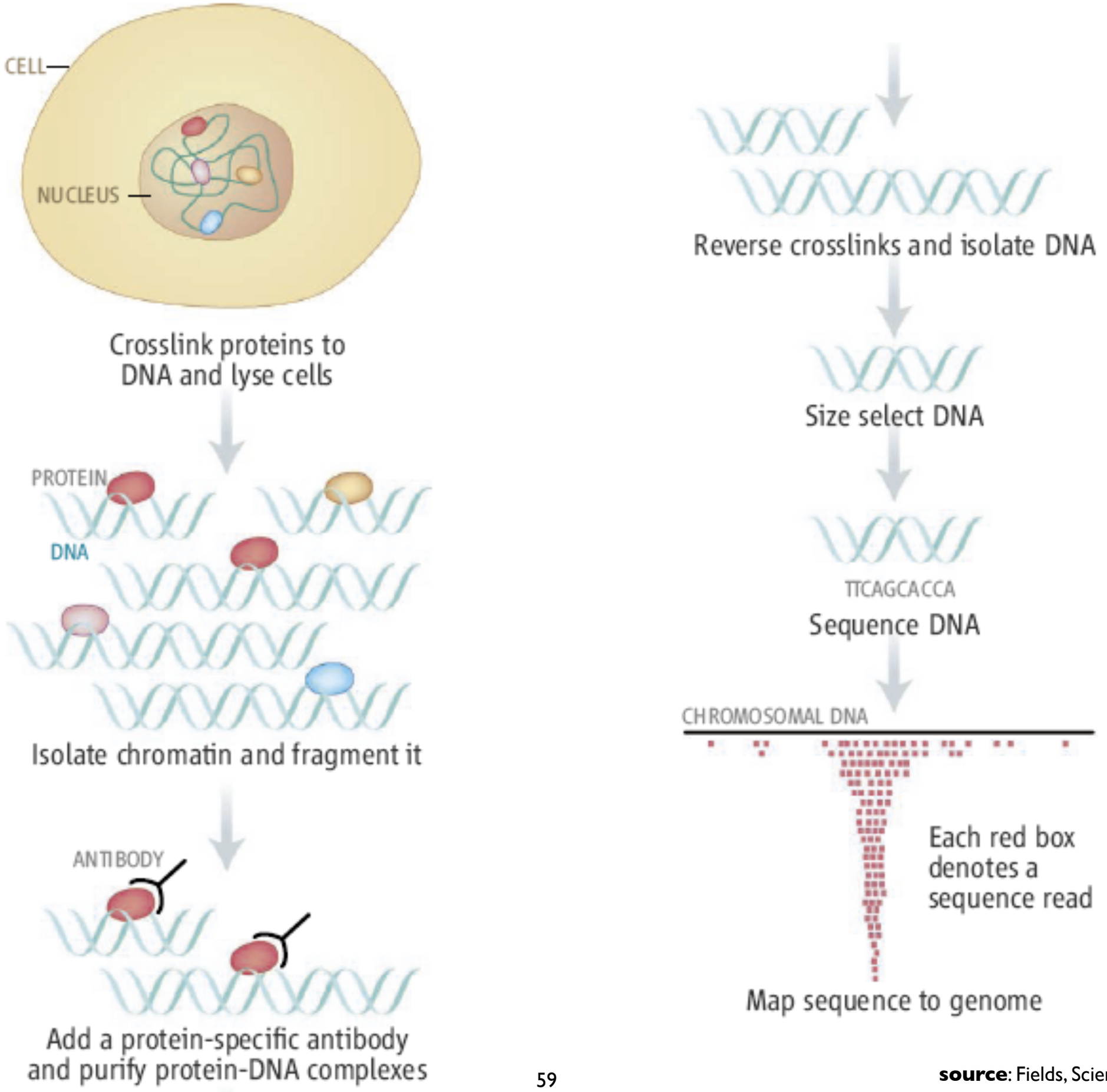


One machine, running for one month, could generate ~15 Gb data.

One Year



Two machines, running for one year, could generate ~365 Gb data.



source: Fields, Science 2007, p1441

Frontiers in Bioinformatics

- ultra high-throughput sequencing
  - DNA binding site identification
  - genome re-sequencing; SNPs, expression
    - ✓ low cost
    - ✓ whole genome
    - ✓ any genome

Credits & References

- Technology Spotlight on DNA Sequencing with Solexa Technology:

http://www.illumina.com/downloads/SS_DNAsequencing.pdf

- Dr. Steven Jones, GSC

several slides/images used with permission

- Stanley Fields, "Site-Seeing by Sequencing", Science, 8 June 2007

Bioinformatics

Session 3.2 - Pathway Resources for Systems Biology

Proteomics

- How large is the human proteome, anyway?

Class	Size	Description
Non Redundant Proteins	20,000-25,000	representative protein from every gene locus
Variants	50,000-500,000	different proteins obtained by splicing or proteolysis
Combinatorial Variants	>10,000,000	different proteins generated by somatic DNA rearrangements
Protein Species	>100,000	proteins that differ in chemical composition due to PTM
Protein Alleles	75,000-150,000	proteins that differ by genetic variation (coding SNPs)

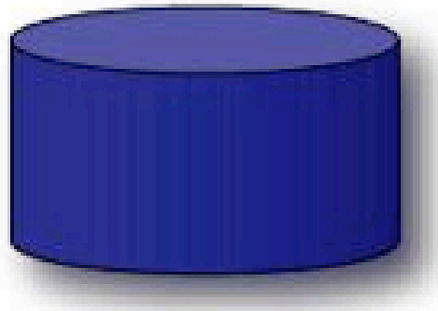
source: Uhlen M et al, Molecular & Cellular Proteomics (2005)

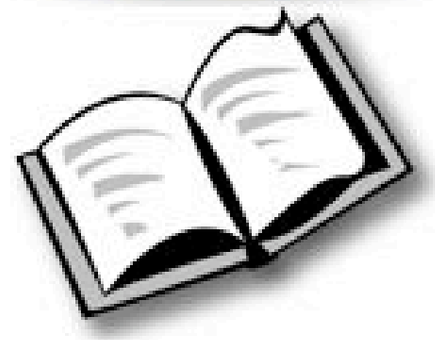
Cellular Pathways

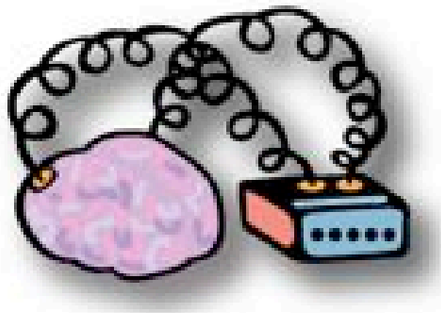
- A striking similarity between intracellular signaling pathways and the Tokyo subway system

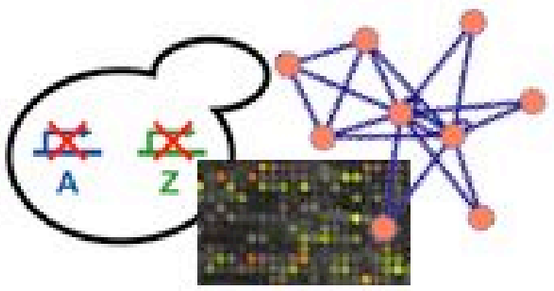
Pathway Information

- Databases
  - Fully electronic
  - Easily computer readable
- Literature
  - Increasingly electronic
  - Human readable
- Biologist's brains
  - Richest data source
  - Limited bandwidth access
- Experiments
  - Basis for models









http://www.pathguide.org/

Home | BioPAX | cBio | MSKCC

Pathguide » the pathway resource list

Navigation

- Protein-Protein Interactions
- Metabolic Pathways
- Signaling Pathways
- Pathway Diagrams
- Transcription Factors / Gene Regulatory Networks
- Protein-Compound Interactions
- Genetic Interaction Networks
- Protein Sequence Focused
- Other

Search

Organisms: All

Availability: All

Standards: All

Reset | Search

Analysis

- Statistics
- Database Interactions

Contact

Complete Listing of All Pathguide Resources

Pathguide contains information about **325** biological related resources and molecular interaction related resources. Click on a link to go to the resource home page or 'Details' for a description page. Databases that are free and those supporting BioPAX, CellML, PSI-MI or SBML standards are respectively indicated.

If you know of a pathway resource that is not listed here, or have other questions or comments, please send us an e-mail.

News

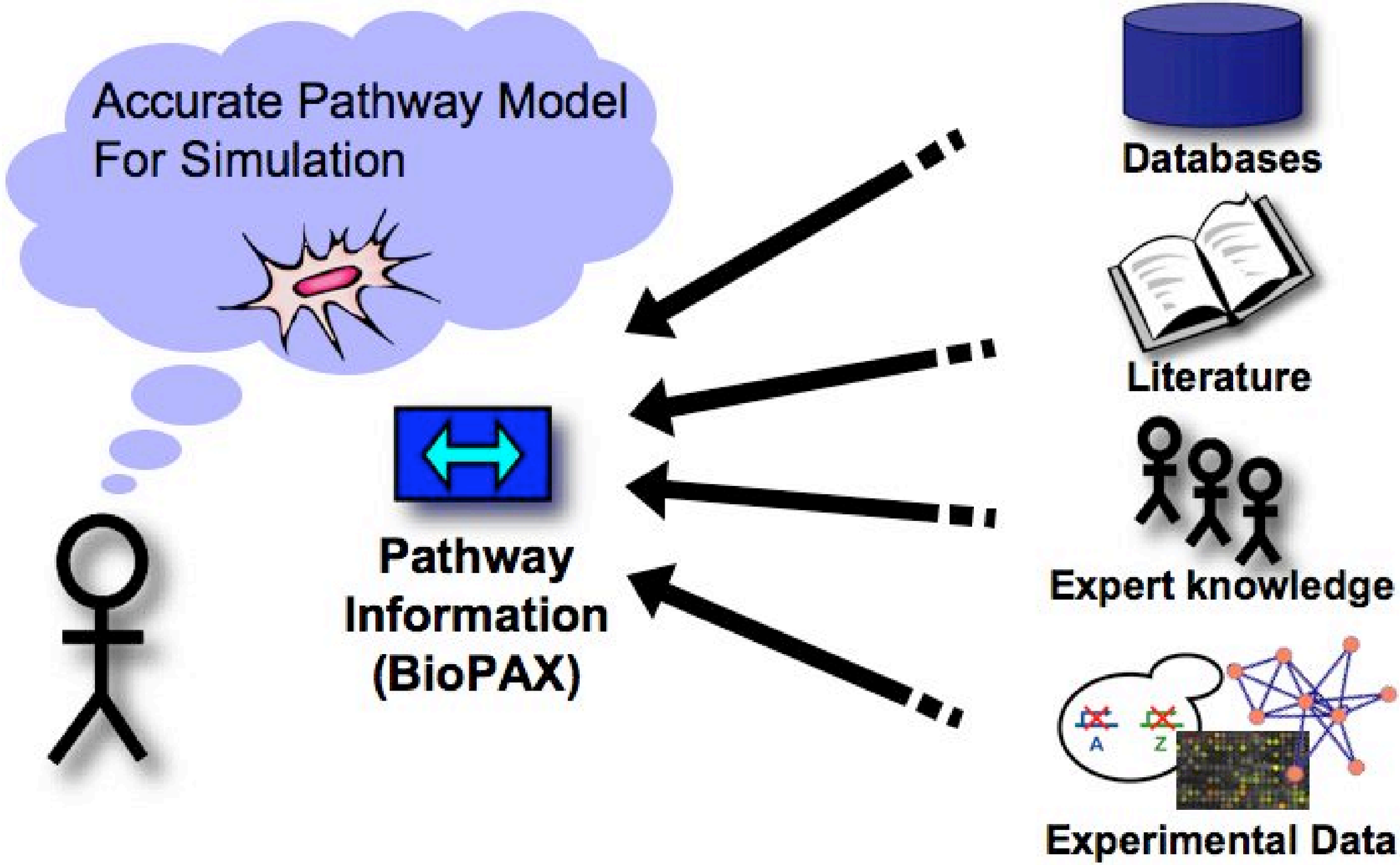
New visual navigation
May.2010
Click the 'Database interactions' link on the left menu to access.

Major update
All resources were recently reviewed and many new ones were added

Protein-Protein Interactions

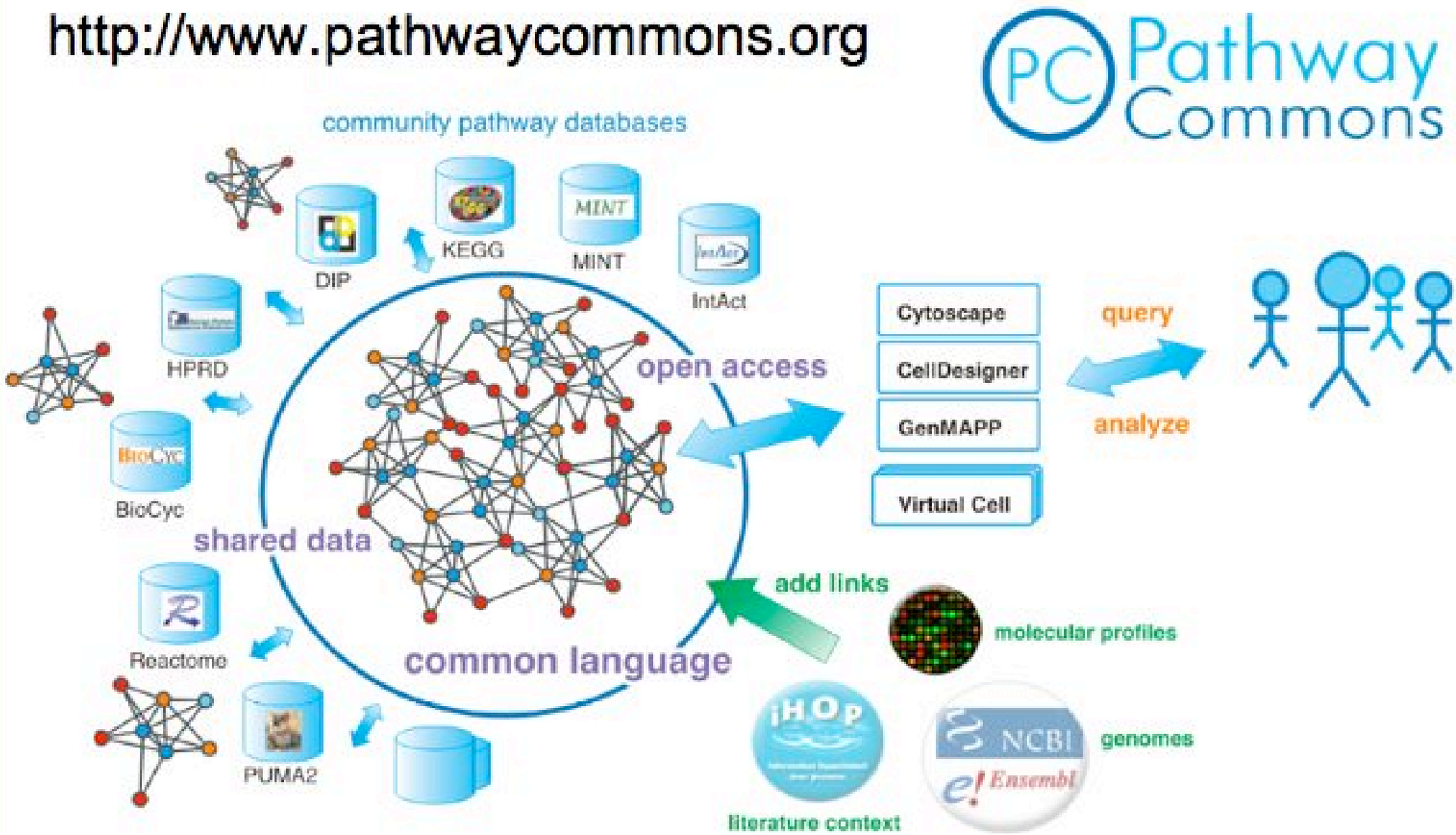
Database Name (Order: alphabetically | by web popularity)	Full Record	Availability	Standards
3DID - 3D interacting domains	Details	Free	
ADAN - Prediction of protein-protein interaction of modular domains	Details	Free	
AllFuse - Functional Associations of Proteins in Complete Genomes	Details	X	
aMAZE - Protein Function and Biochemical Pathways Project	Details	Free	
APID - Agile Protein Interaction DataAnalyzer	Details	Free	
ASEdb - Alanine Scanning Energetics Database	Details	Free	
ASPD - Artificial Selected Proteins/Peptides Database	Details	Free	
AtPID - Arabidopsis thaliana Protein Interactome Database	Details	Free	
BID - Binding Interface Database	Details	Free	
BIND - Biomolecular Interaction Network Database	Details	Free	PSI-MI
BioGRID - Biological General Repository for Interaction Datasets	Details		PSI-MI
BRITE - Biomolecular Relations in Information Transmission and Expression	Details	Free	
CA1Neuron - Pathways of the hippocampal CA1 neuron	Details	Free	
Cancer Cell Map - The Cancer Cell Map	Details	Free	BioPAX

Using Pathway Information



Aim: Convenient Access to Pathway Information

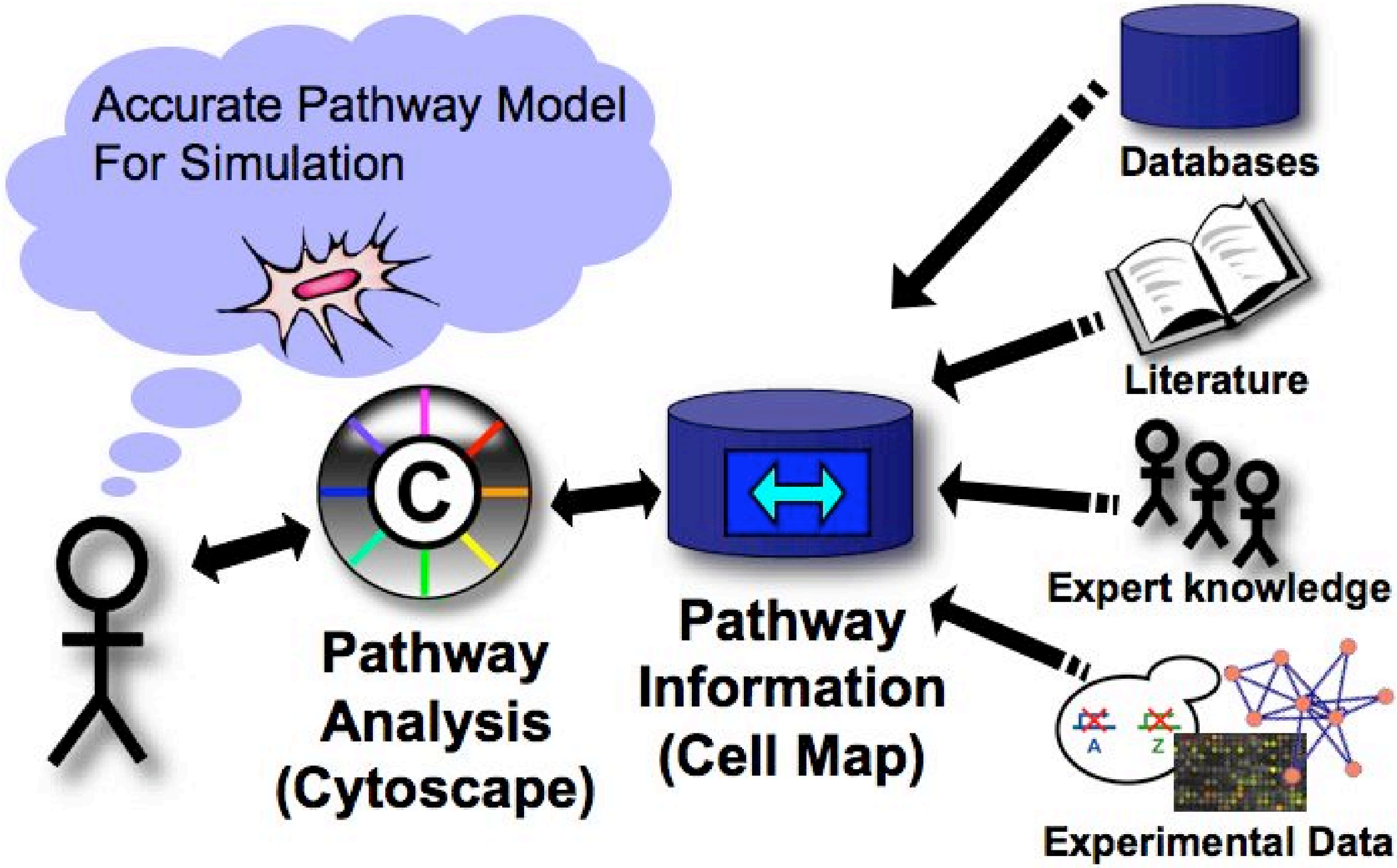
http://www.pathwaycommons.org



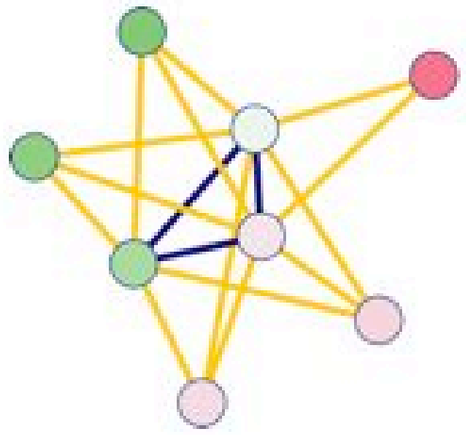
Facilitate creation and communication of pathway data
Aggregate pathway data in the public domain
Provide easy access for pathway analysis

Long term: Converge to integrated cell map

Using Pathway Information



Cytoscape - Network Visualization and Analysis



http://cytoscape.org

- Freely-available (open-source, java) software
- Visualizing biological networks (e.g. molecular interaction networks)
- Analyzing networks with gene expression profiles and other cell state data

UCSD, ISB, Agilent, MSKCC, Pasteur, UCSF, UToronto

Other software: Osprey, BioLayout, VisANT, Navigator, PIMWalker, ProViz

Cytoscape Desktop (New Session)

File Edit View Select Layout Plugins Help

Sample2 Search: yal003w

CytoPanel 1

Network Editor

Network	Nodes	Edges
galFiltered.sif	331(19)	362(35)

galFiltered.sif

CytoPanel 2

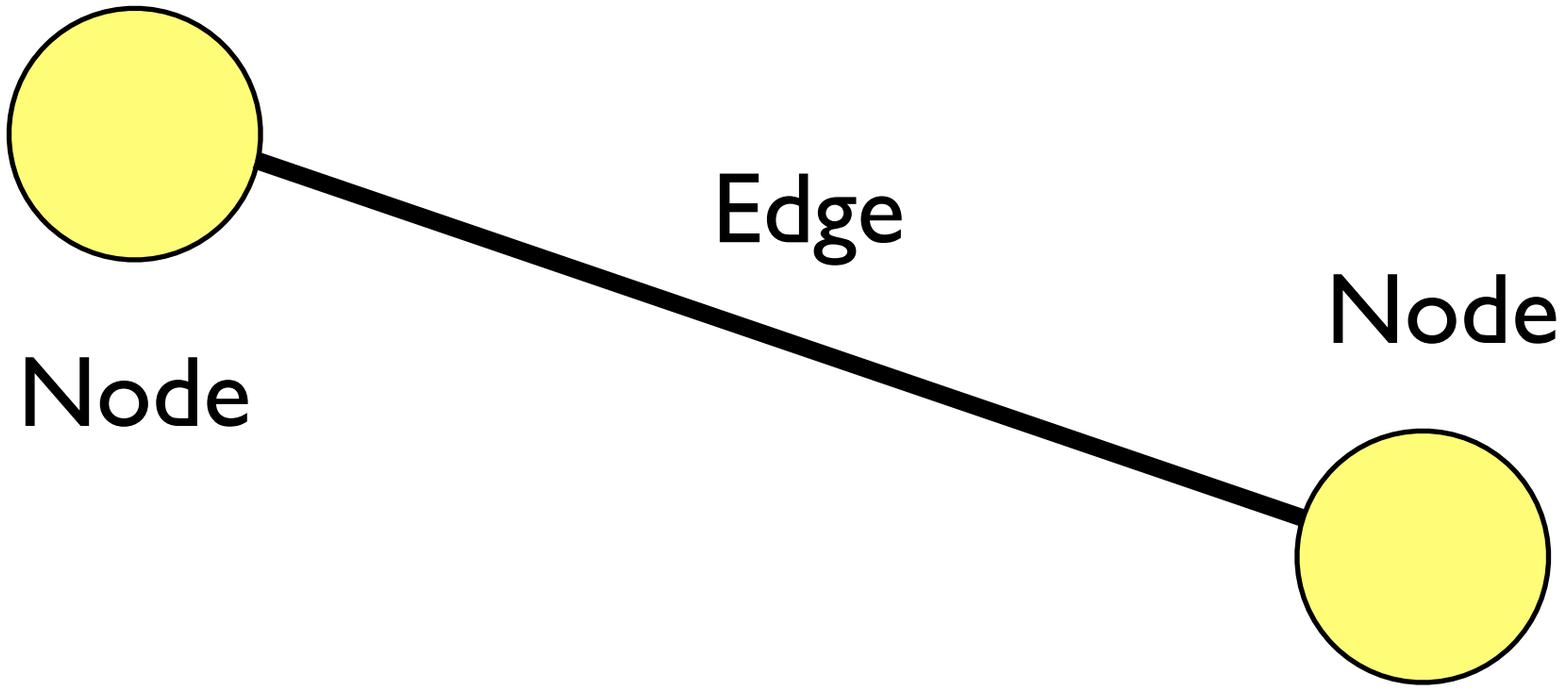
Node Attribute Browser (galFiltered.sif)

ID	gal1RGexp	gal1RGsig	gal4RGexp	gal4RGsig	gal80Rexp	gal80I
YGL008C	-0.352	1.0007E-5	-0.282	7.1366E-4	-0.573	1.2622E
YCL067C	0.169	0.0012873	-0.085	0.11481	0.301	0.0027
YNL145W	-0.764	3.148E-11	-0.098	0.05338	-1.237	1.1916E
YMR043W	-0.183	0.0035372	-0.654	4.2514E-6	0.457	2.4112E

Node Attribute Browser | Edge Attribute Browser | Network Attribute Browser

Welcome to Cytoscape 2.4.0-b1 Right-click + drag to ZOOM Middle-click + drag to PAN

Pathway Graphs

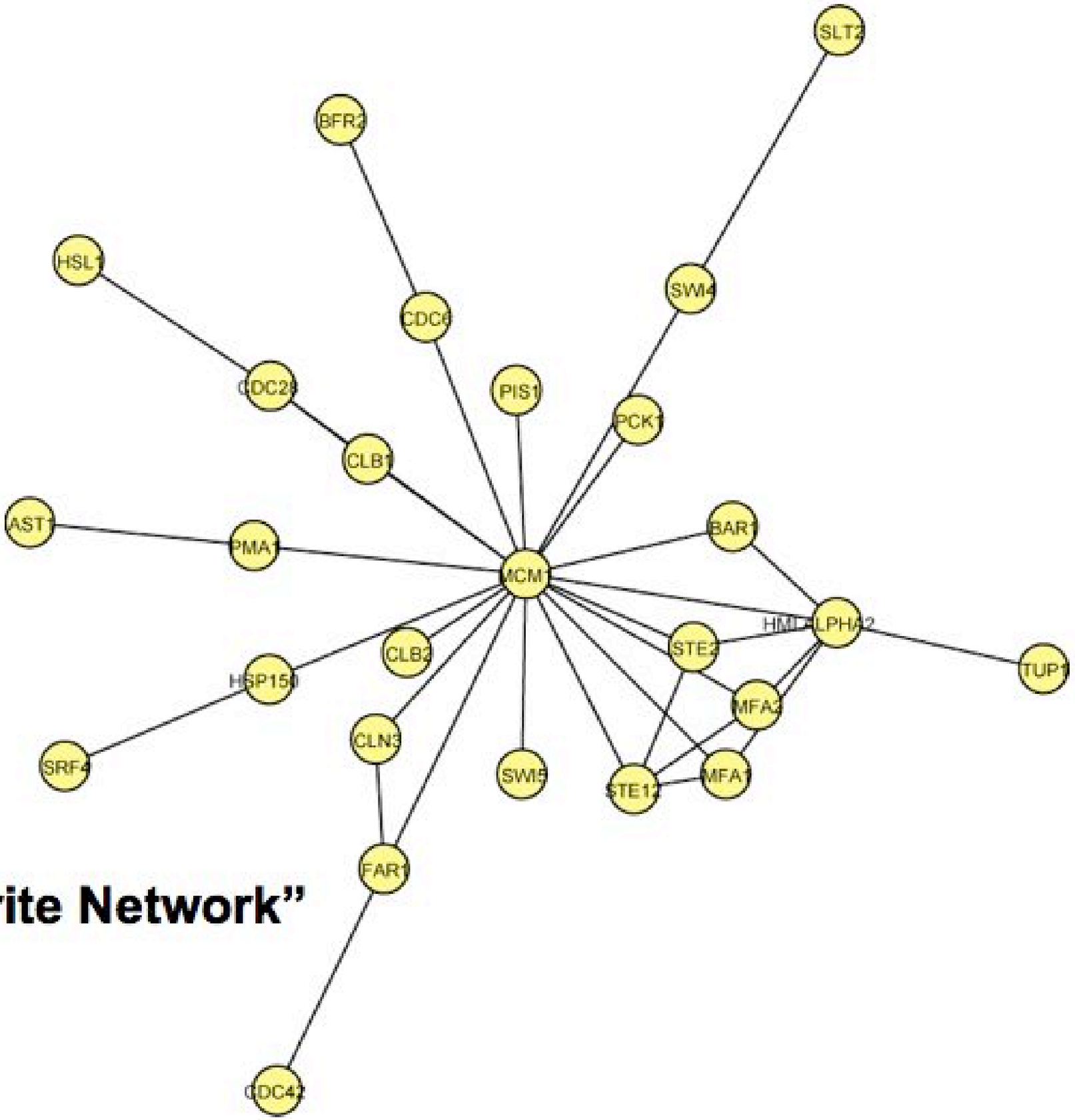


- In addition to describing the network topology, nodes and edges can each have their own attributes

Visual Style

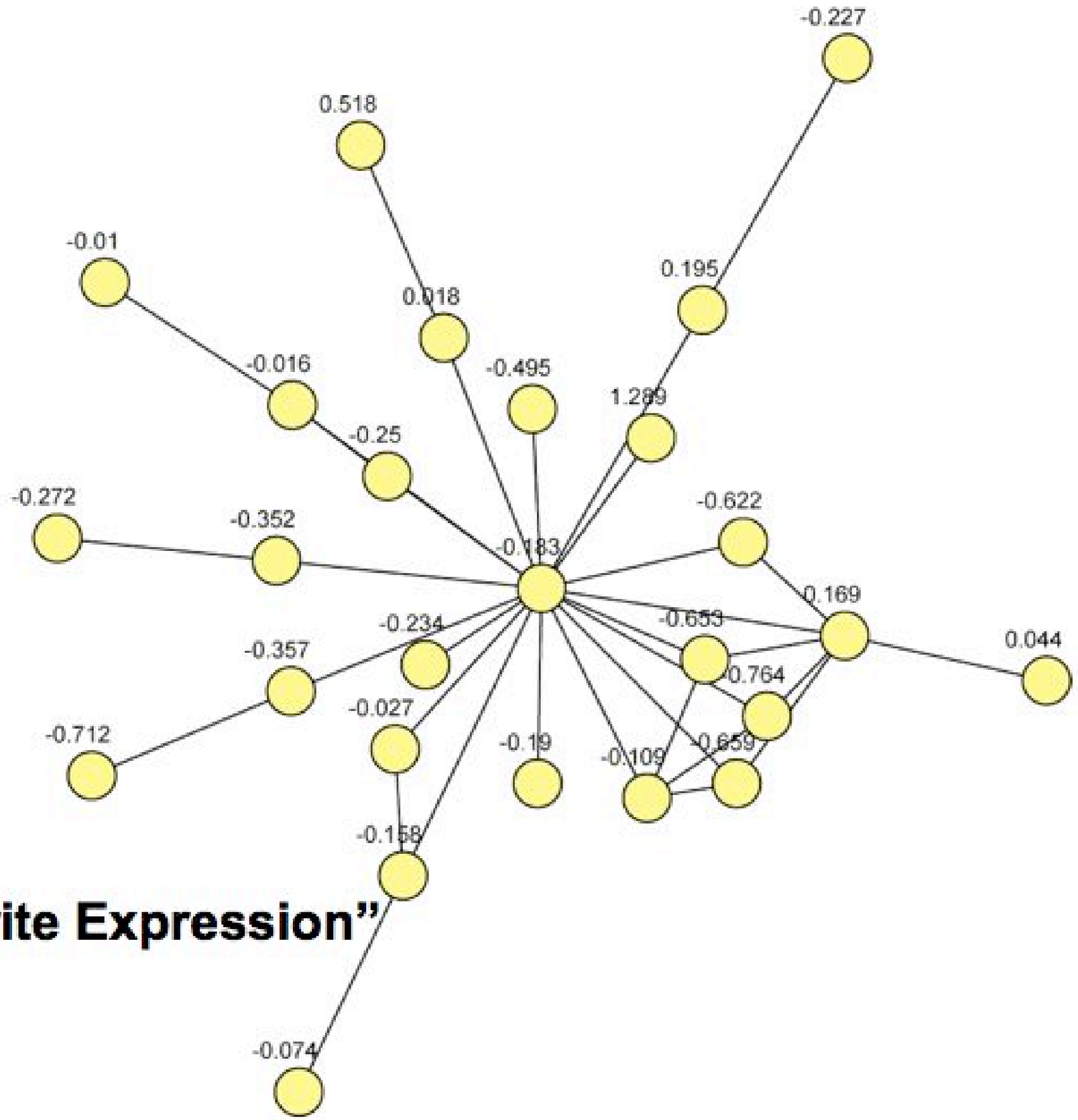
- Customized views of experimental data in a network context
- Network has node and edge attributes
  - E.g. expression data, interaction type, GO function
- Mapped to visual attributes
  - E.g. node/edge size, shape, colour…
- E.g. Visualize gene expression data as node colour gradient on the network

Visual Style



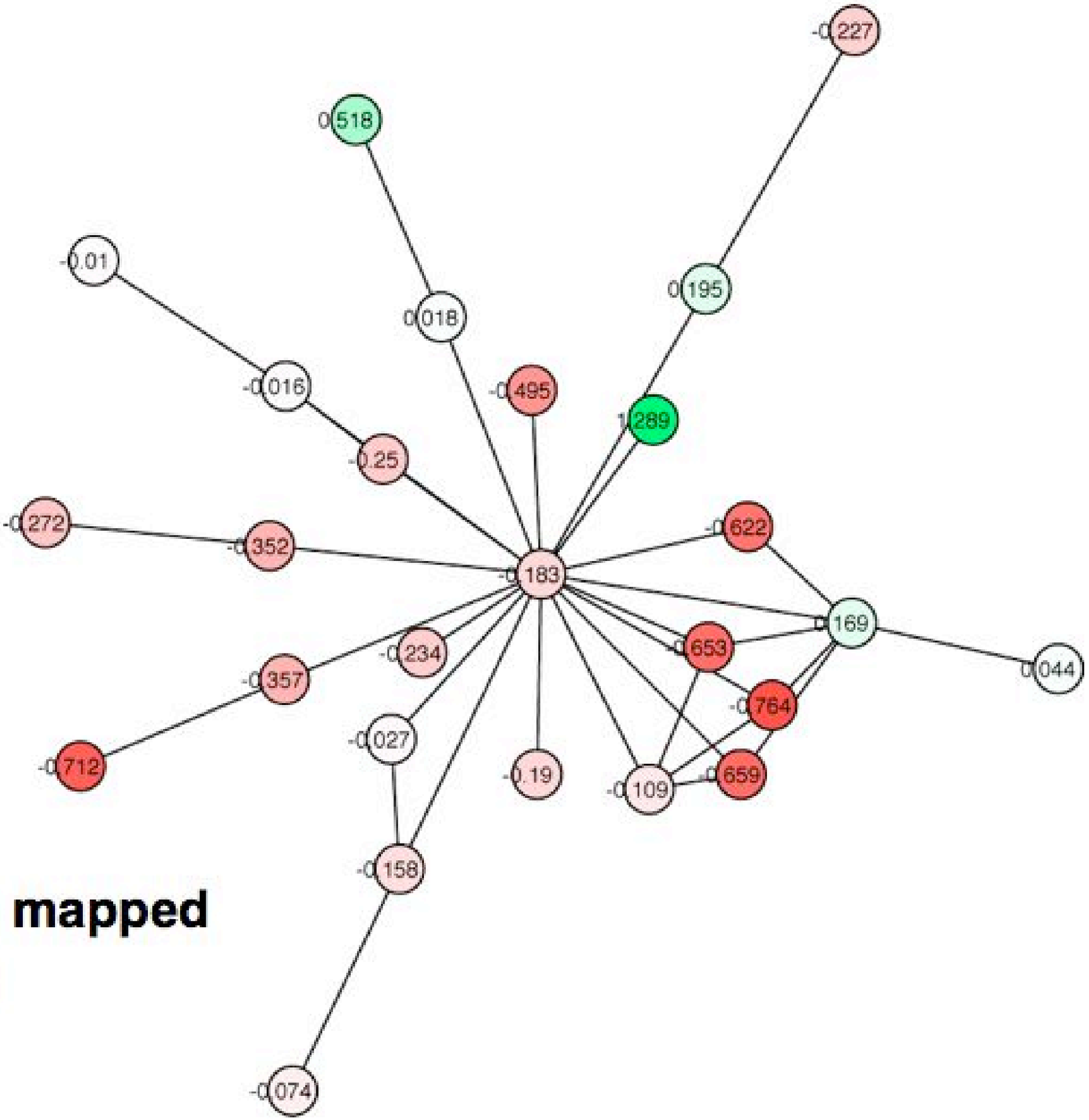
Load “Your Favorite Network”

Visual Style



Load "Your Favorite Expression" Dataset

Visual Style



Expression data mapped to node colours

Systems Biology

- Goals:
  - ✓ integrating diverse data types, pathways
  - ✓ cellular simulations
- Community approaches:
  - ✓ pathguide, pathway commons, cytoscape
- Open data exchange key to success

Credits & References

- Dr. Gary Bader, DCCBR, UofT

  slides/images used with permission

- Cary MP, Bader GD, Sander C "Pathway Information for Systems Biology", FEBS Letters (2005)

Bioinformatics Links Directory

Finding online tools & resources for Life Sciences research

Conducting Research on the Web: 2007 Update for the Bioinformatics Links Directory

Joanne A. Fox, Scott McMillan and B. F. Francis Ouellette*

UBC Bioinformatics Centre (http://bioinformatics.ubc.ca/), I
Vancouver, British Columbia, Canada, V6T 1Z4

Received May 18, 2007; Accepted May 22, 2007

ABSTRACT

The Bioinformatics Links Directory, http://bioinformatics.ca/links_directory, is an actively maintained compilation of servers published in this and previous issues of *Nucleic Acids Research* issues together with many other useful tools, databases and resources for life sciences research. The 2007 update includes the 130 websites highlighted in the July 2007 Web Server issue of *Nucleic Acids Research* and brings the total number of servers listed in the Bioinformatics Links Directory to just under 1200 links. In addition to the updated content, the 2007 update of the Bioinformatics Links Directory includes new features for improved navigation, accessibility and open data exchange. A complete listing of all links listed in this *Nucleic Acids Research* 2007 Web Server issue can be accessed online at, http://bioinformatics.ca/links_directory/narweb2007. The 2007 update of the Bioinformatics Links Directory, which includes the Web Server list and summaries is also available online, at the *Nucleic Acids Research* web site, http://nar.oupjournals.org.

COMMENTARY

With the publication of the 2007 *Nucleic Acids Research* Web Server issue, we have a chance to reflect on how the web has transformed the way we conduct scientific

W2–W4 Nucleic Acids Research, 2008, Vol. 36, Web Server issue
doi:10.1093/nar/gkn399

Keeping pace with the data: 2008 update on the Bioinformatics Links Directory

Michelle D. Brazas[1], Joanne A. Fox[2], Timothy Brown[1], Scott McMillan[3] and B. F. Francis Ouellette[1,*]

[1]Ontario Institute for Cancer Research, 101 College St, Suite 800, Toronto, Ontario, Canada M5G 0A3, [2]University of British Columbia, Michael Smith Laboratories and [3]University of British Columbia, Office o Learning and Technology, Vancouver, British Columbia, Canada

Received June 3, 2008; Revised and Accepted June 5, 2008

ABSTRACT

The Bioinformatics Links Directory, http://bioinformatics.ca/links_directory/, is an online resource for public access to all of the life science research web servers published in this and previous issues of *Nucleic Acids Research*, together with other useful tools, databases and resources for bioinformatics and molecular biology research. Dependent on community input and development, the Bioinformatics Links Directory exemplifies an open access research tool and resource. The 2008 update includes the 94 web servers featured in the July 2008 Web Server issue of *Nucleic Acids Research*, bringing the total number of servers listed in the Bioinformatics Links Directory to over 1200 links. A complete list of all links listed in this *Nucleic Acids Research* 2008 Web Server issue can be accessed online at http://bioinfomatics.ca/links_directory/narweb2008/. The 2008 update of the Bioinformatics Links Directory, which includes the Web Server list and summaries, is also available online at the *Nucleic Acids Research* website, http://nar.oxfordjournals.org/.

networks at play in a given disease or biological
or ask questions that explore the commonalities
iations between large data sets from different r
lecules, species or organisms.

Keeping pace with these advances in techno
data output has been the number of specialized
vers and bioinformatic resources developed or
to meet these new data intensive research nee
2004, *Nucleic Acids Research* has peer-reviewed
lished in their Web Server issue, a compendiu
latest web servers and freely available online bioi
tools to keep researchers abreast of the deluge of
matic resources available to them. This year's W
issue introduces an additional 94 bioinforma
molecular biology web servers, 10 of which ar
(Table 1). Along with the long-standing Datab
(1), the special Web Server issues represent an i
source of bioinformatic tools and resources for
national life-science research community. The
listing of URLs cited in the 2008 Web Server
be accessed online at the *Nucleic Acids Research*
http://nar.oxfordjournals.org/, as well as at http:
matics.ca/links_directory/narweb2008/.

The Bioinformatics Links Directory, http://bioin
ca/links_directory/, is a public, curated collectio

http://bioinformatics.ca/links_directory/

bioinformatics.ca
links directory

Bioinformatics Links Directory

The Bioinformatics Links Directory features curated links to molecular resources, tools and databases. The links listed in this directory are selected on the basis of recommendations from bioinformatics experts in the field. We also rely on input from our community of bioinformatics users for suggestions. Starting in 2003, we have also started listing all links contained in the NAR Webserver issue.

Search Directory

Computer Related (64)

This category contains links to resources relating to programming languages often used in bioinformatics. Other tools of the trade, such as web development and database resources, are also included here.

DNA (488)

This category contains links to useful resources for DNA sequence analyses such as tools for comparative sequence analysis and sequence assembly. Links to programs for sequence manipulation, primer design, and sequence retrieval and submission are also listed here.

Education (72)

Links to information about the techniques, materials, people, places, and events of the greater bioinformatics community. Included are current news headlines, literature sources, educational material and links to bioinformatics courses and workshops.

Expression (358)

Links to tools for predicting the expression, alternative splicing, and regulation of a gene sequence are found here. This section also contains links to databases, methods, and analysis tools for protein expression, SAGE, EST, and microarray data. Expression analysis of next-generation sequencing data sets is also covered.

Human Genome (160)

This section contains links to draft annotations of the human genome in addition to resources for sequence polymorphisms and genomics. Also included are

Literature (52)

Links to resources related to published literature, including tools to search for articles and through literature abstracts. Additional text mining resources,

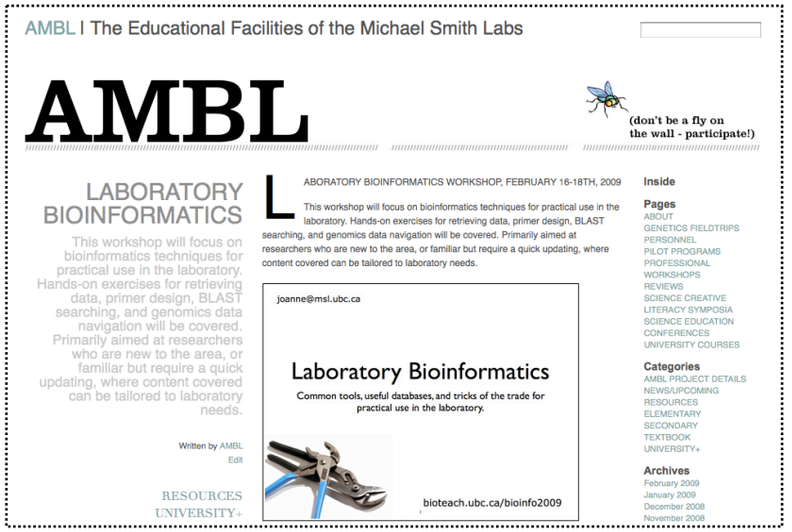
- Main Page
- Citations
- Acknowledgements
- News
- Suggest a URL
- NAR Collaboration
- RSS Feeds
- Support

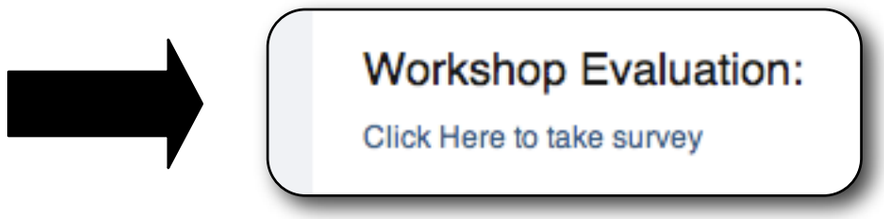
bioinformatics.ca

Return to BioInformatics.ca

Your Feedback is Important!

navigate to:
bioteach.ubc.ca/bioinfo2010





Thanks for attending the AMBL Laboratory Bioinformatics Workshop

- Questions? Please contact:

Dr. Joanne Fox
Michael Smith Laboratories
joanne@msl.ubc.ca