Sample Genetic Maps of chromosomes from Nematodes (*C. elegans*)

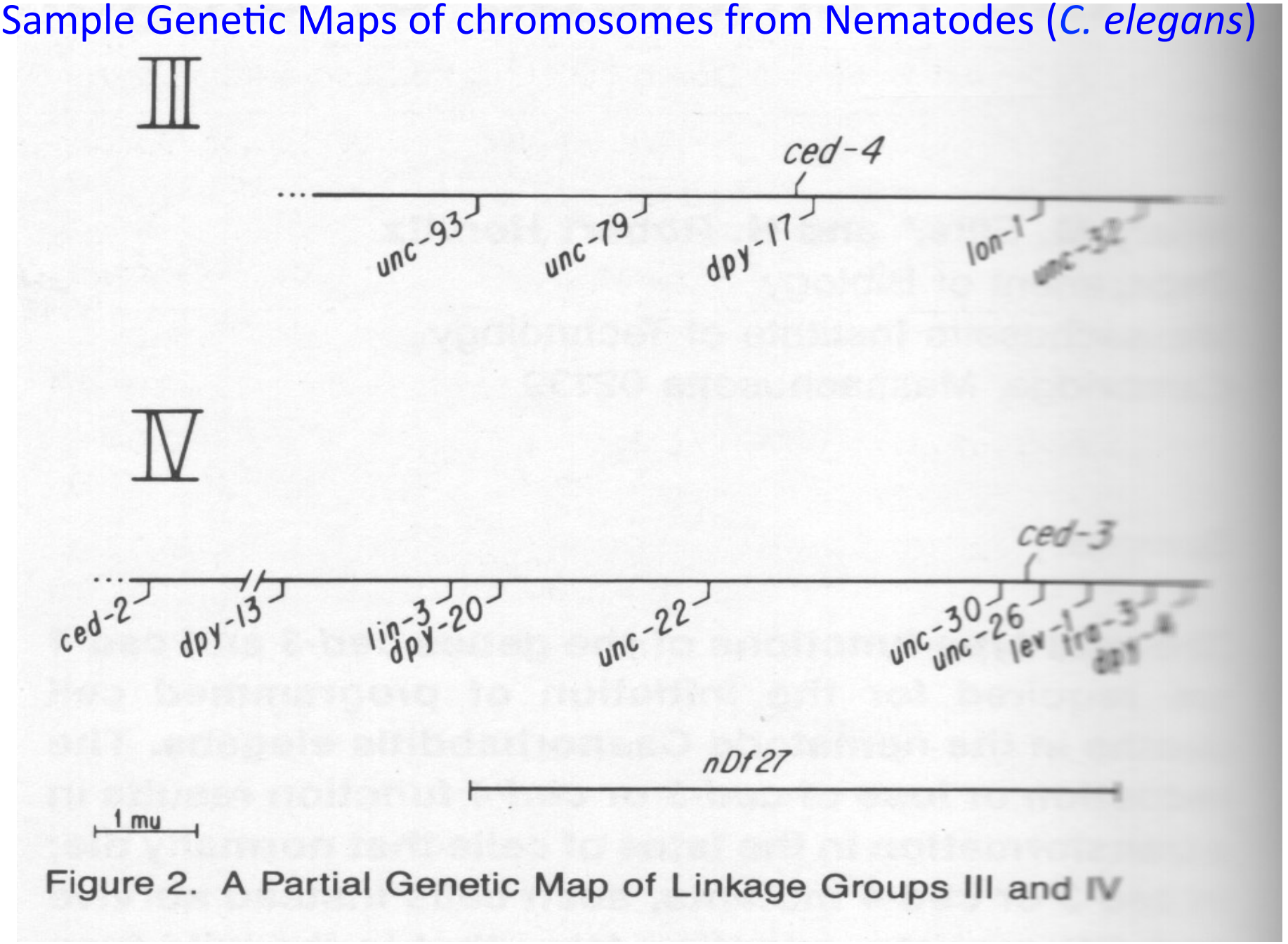
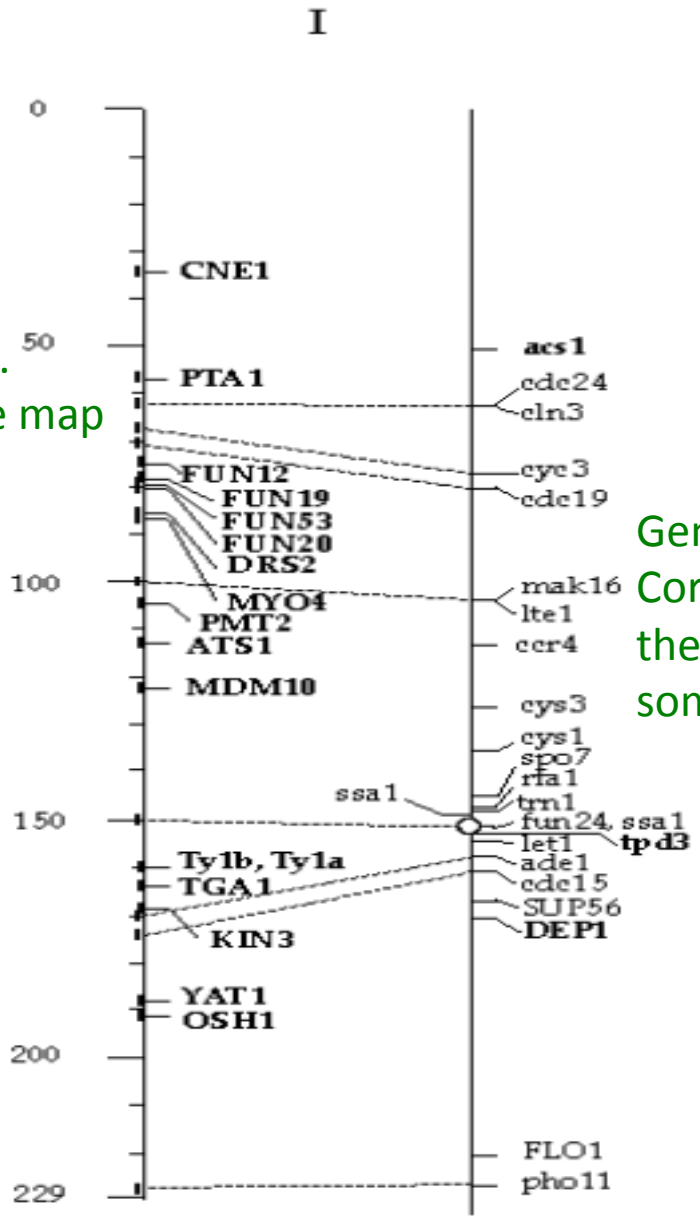


Figure 2. A Partial Genetic Map of Linkage Groups III and IV

Physical and genetic maps of yeast chromosome I

Physical map. ~250kb of DNA.
Some genes positioned on the map

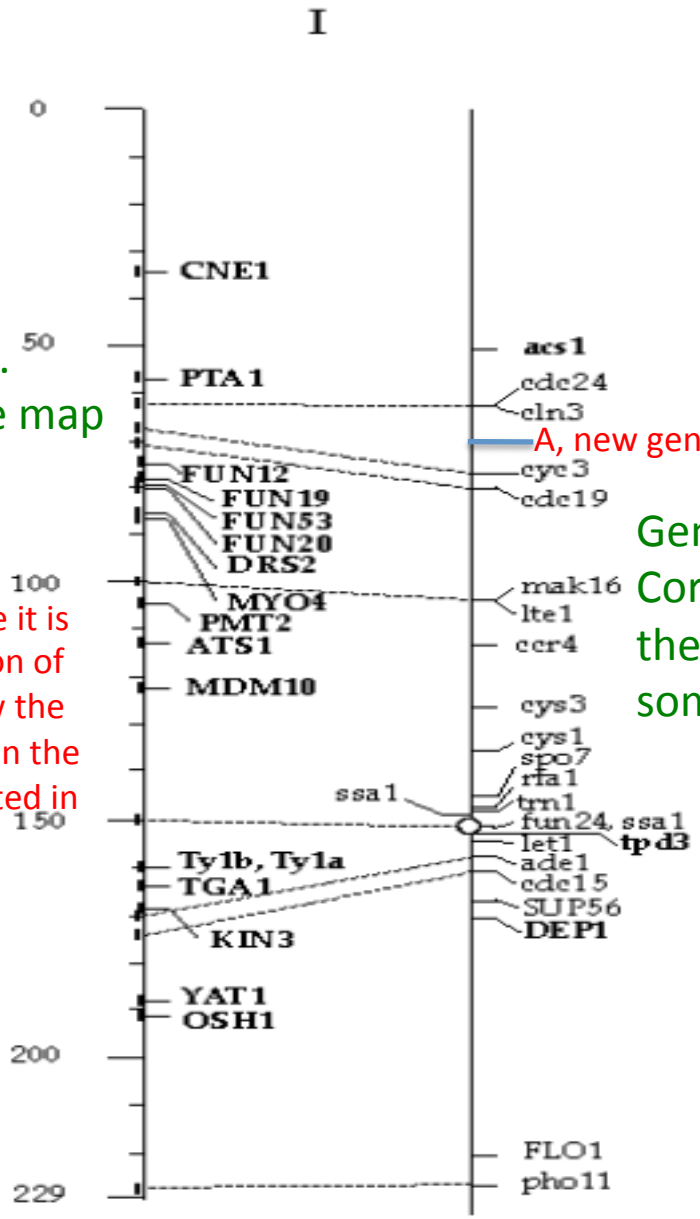


Genetic map. ~100cM.
Correspondence between the two maps made for some genes

Physical and genetic maps of yeast chromosome I

Physical map. ~250kb of DNA.
Some genes positioned on the map

By mapping the gene, you know where it is on the physical map, the representation of the DNA, which has been cloned. Now the challenge is to figure out which gene on the physical map is the one you're interested in



A, new gene you've discovered

Genetic map. ~100cM.
Correspondence between the two maps made for some genes

Lexicon and Tools for Human Genetics

One goal is to be able to clone a disease gene of interest

To do so, we need physical and genetic maps and connections between the two

Physical map: Depiction of the 3 Billion bp of DNA with location of known sequences along the map.
The sites are usually anonymous, usually without any biological meaning.
They're just identified position markers, called Sequence Tagged Sites (STS),
that can be assayed by PCR

Genetic map: Depiction of the location of sequence variation along each chromosome.
The sequence variation may be disease genes (or other physiological variation) that have been discovered. The sequence variation may also simply be at anonymous loci that happen to be polymorphic. Examples are (CA)n repeats and single nucleotide polymorphisms (SNP).

The virtues of these polymorphic markers are that they are abundant – millions in the genome - so they can be used to make a high quality genetic map and they are also STSs so they can be used to connect the genetic and physical maps

Sequence Tagged Sites

Standard STS

Useful for physical map.

Can ask whether site is present on any cloned DNA by PCR amplification

```
Ccatgatcgaagtctt…100-300nucleotides…ggtaactactgaata
Ggtactagcttcagaa…                    …ccattgatgacttat
```

Polymorphic STS

Useful for genetic (and physical) map

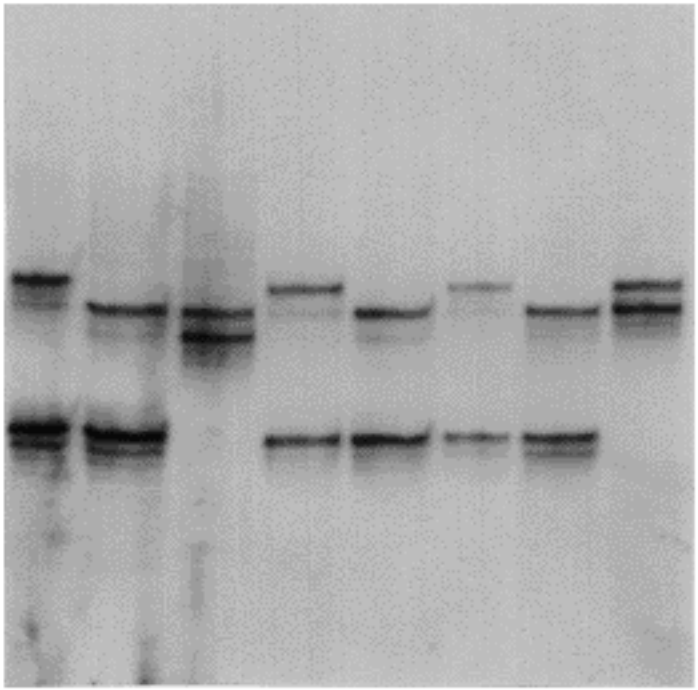
Examples include SNPs (single nucleotide polymorphisms) and simple sequence repeats, such as (CA)n repeats

```
Ccatgatcgaagtgtt…100-300nucleotides…ggtaactactgaata
Ggtactagcttcacaa…                    …ccattgatgacttat
```

(CA)n Repeats as examples of polymorphic loci

One allele: ccctgaaacgctt(CA)15tttgtaactaacggtt

Another allele: ccctgaaacgctt(CA)18tttgtaactaacggtt



How would you look for new CA repeat loci?

The (CA) repeats are flanked by unique DNA. PCR primers can be designed to the unique DNA. PCR reactions followed by gel electrophoresis allows the (CA) repeat alleles to be visualized.

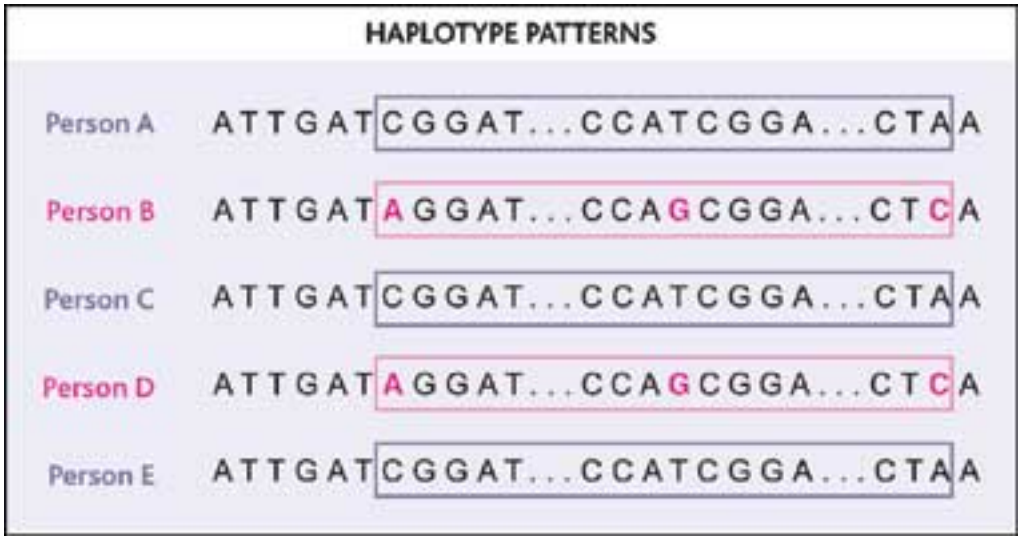
Single Nucleotide Polymorphisms

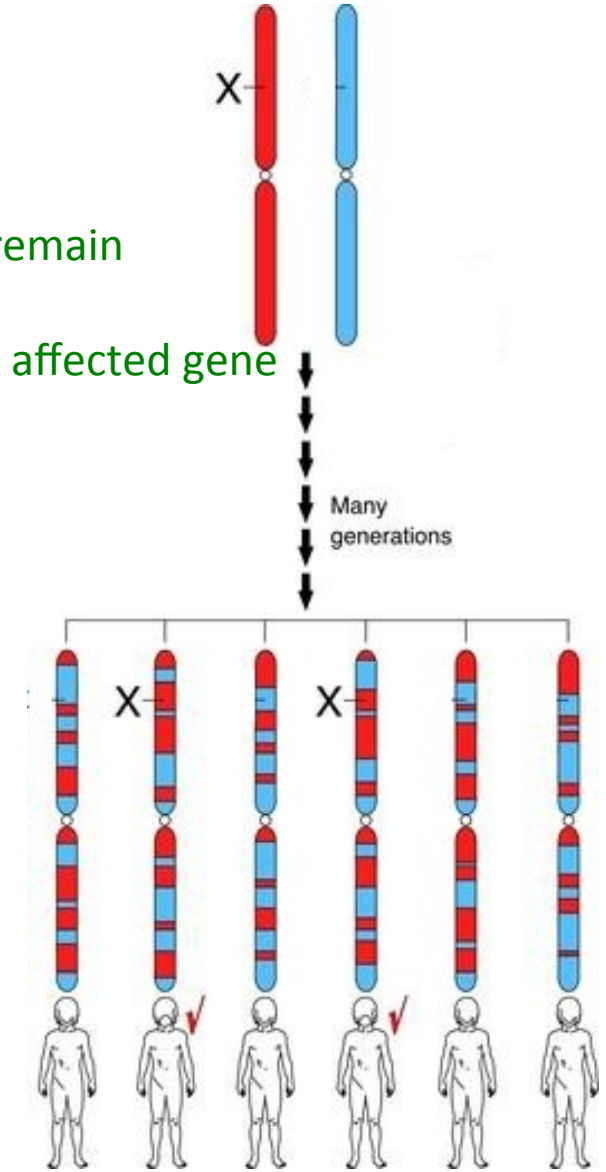
Very common. In human population, 1 every 1000 bp

Haplotypes

Nearby polymorphisms tend to be inherited together, creating “haplotype” blocks

Hence, a “founder” mutation that causes a disease will remain linked to nearby polymorphisms. Identification of these polymorphisms provides a way to clone and identify the affected gene





Human Genome Project and a Physical Map

The human genome is huge (compared to anything that had been tackled previously).
One strategy was to break it into smaller and smaller bits until DNA was a "comfortable" size.

Create libraries in vectors that can accommodate different sizes of DNA

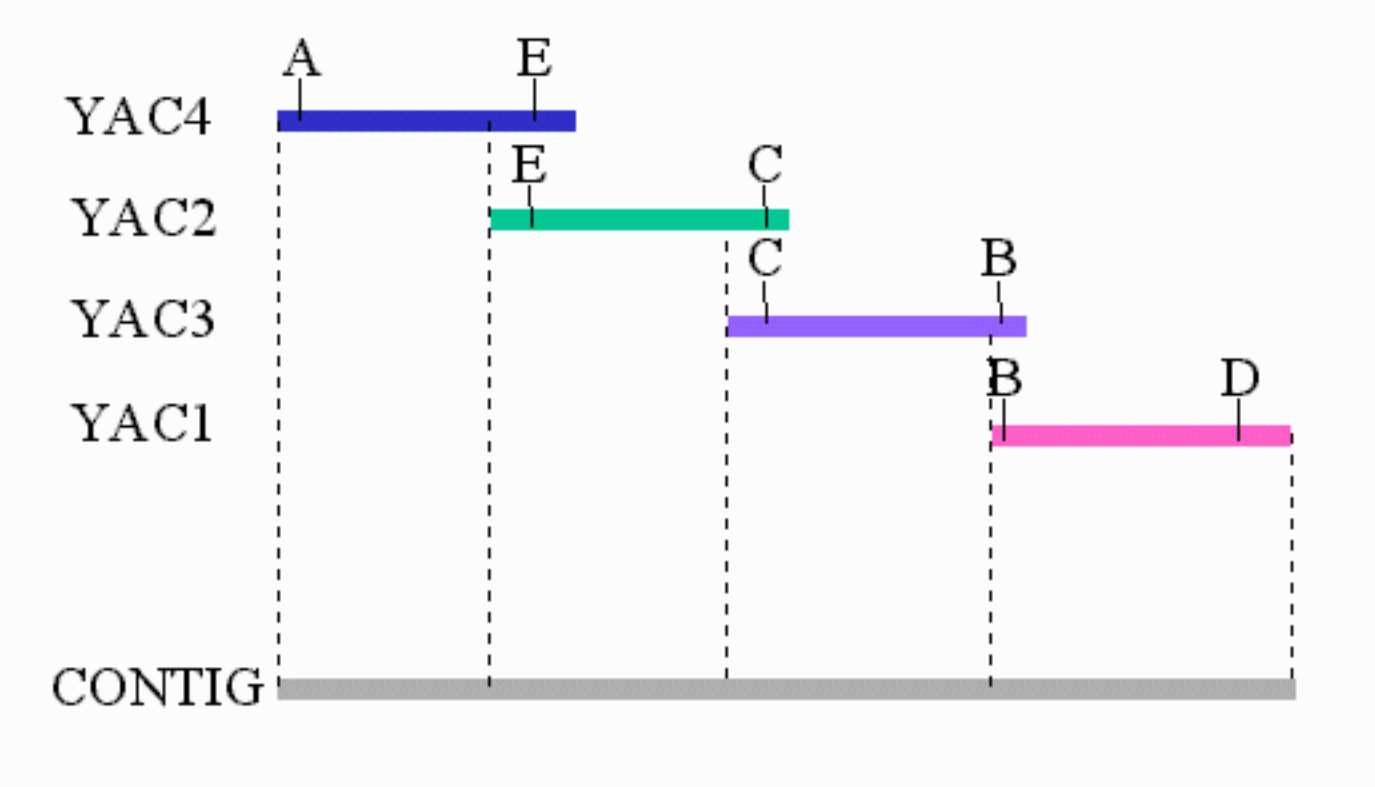
Yeast Artificial Chromosomes (YAC). 1 – 20 M bp

Bacterial Artificial Chromosomes (BAC). 100 – 500kb)

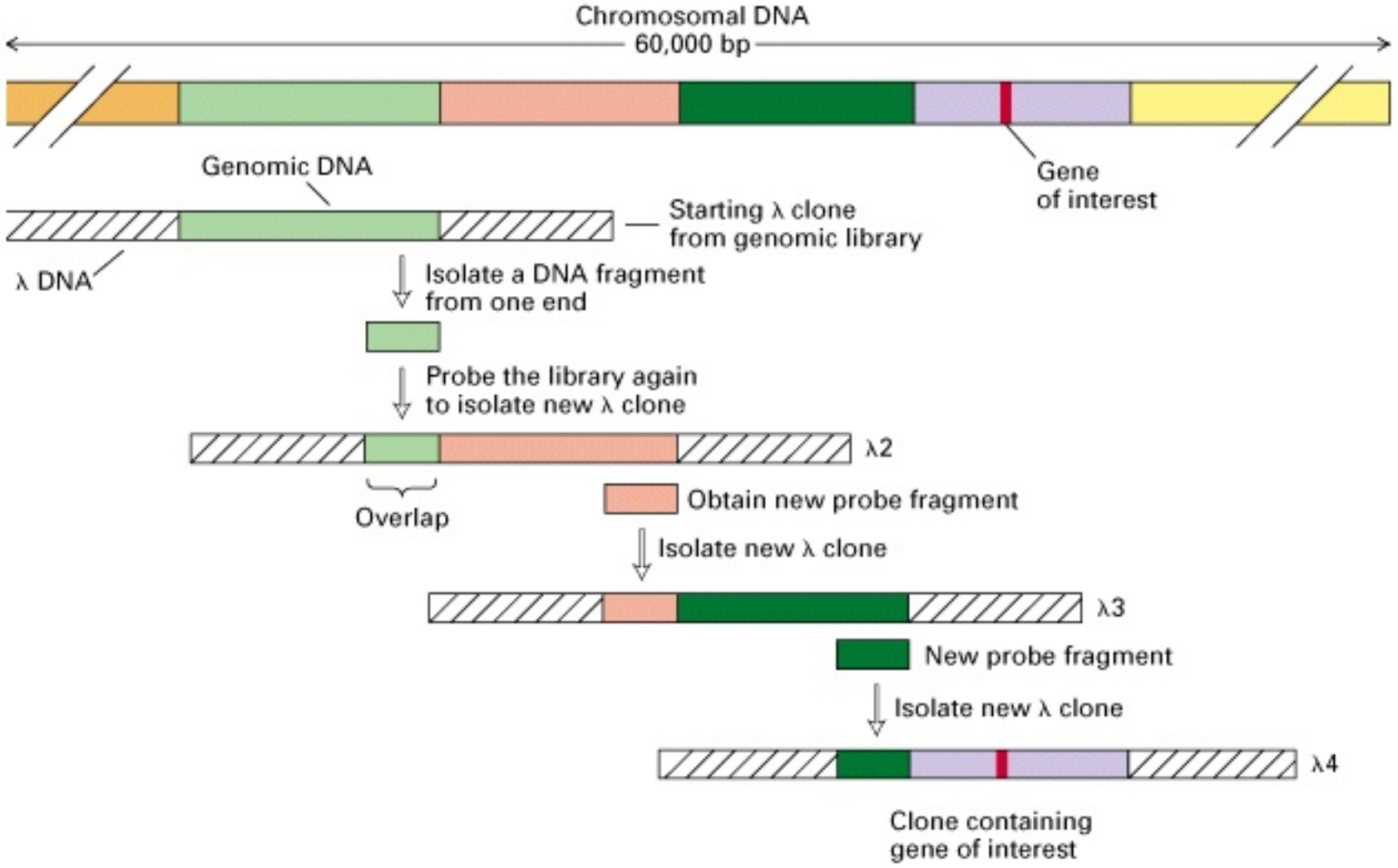
P1 Artificial Chromosomes (PAC). 50 – 100kb

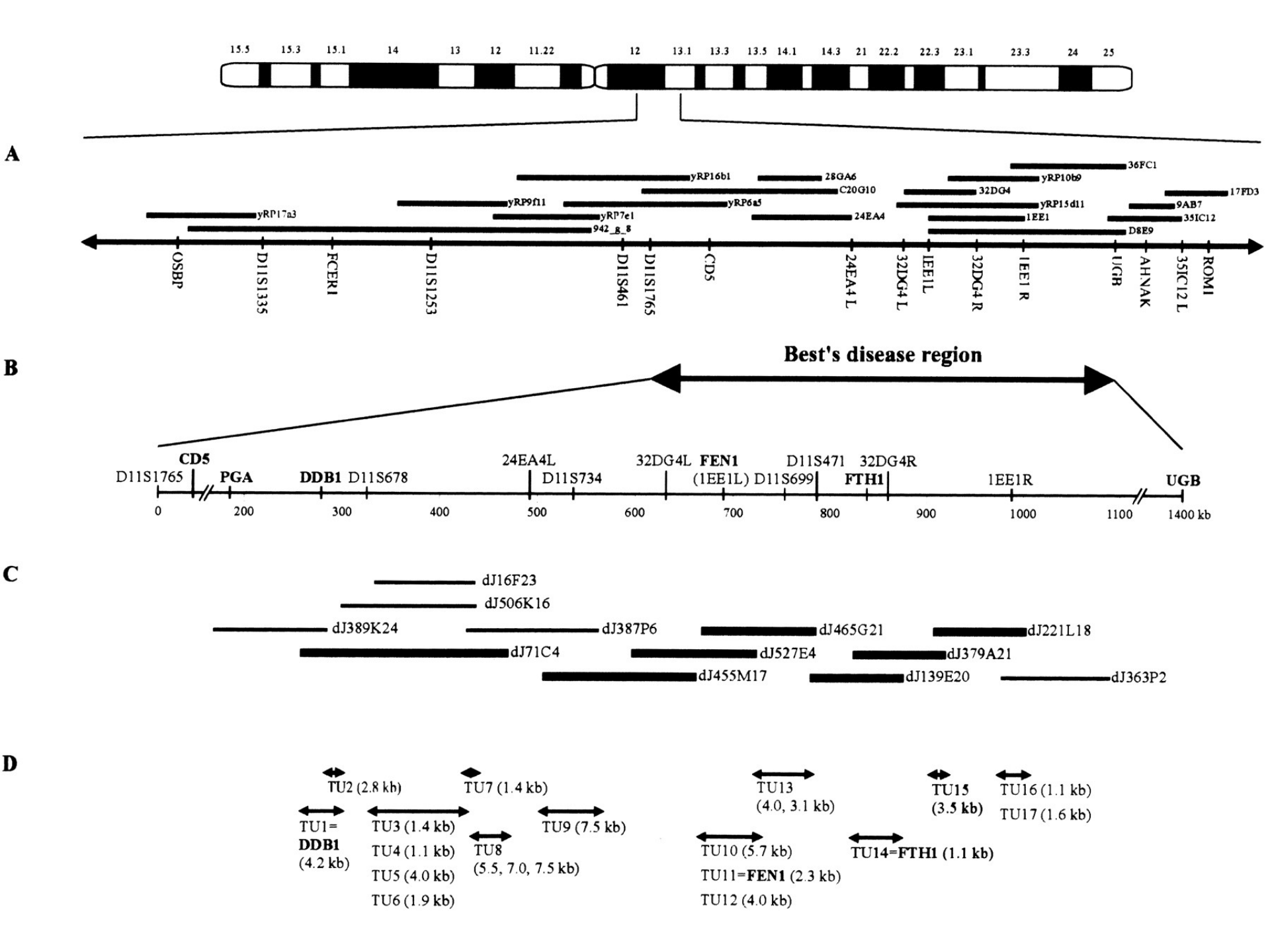
Contigs

Use the PACs, BACs, and YACs, together with STSs, to link the cloned segments together, thereby recreating the genome, at least *in silico*

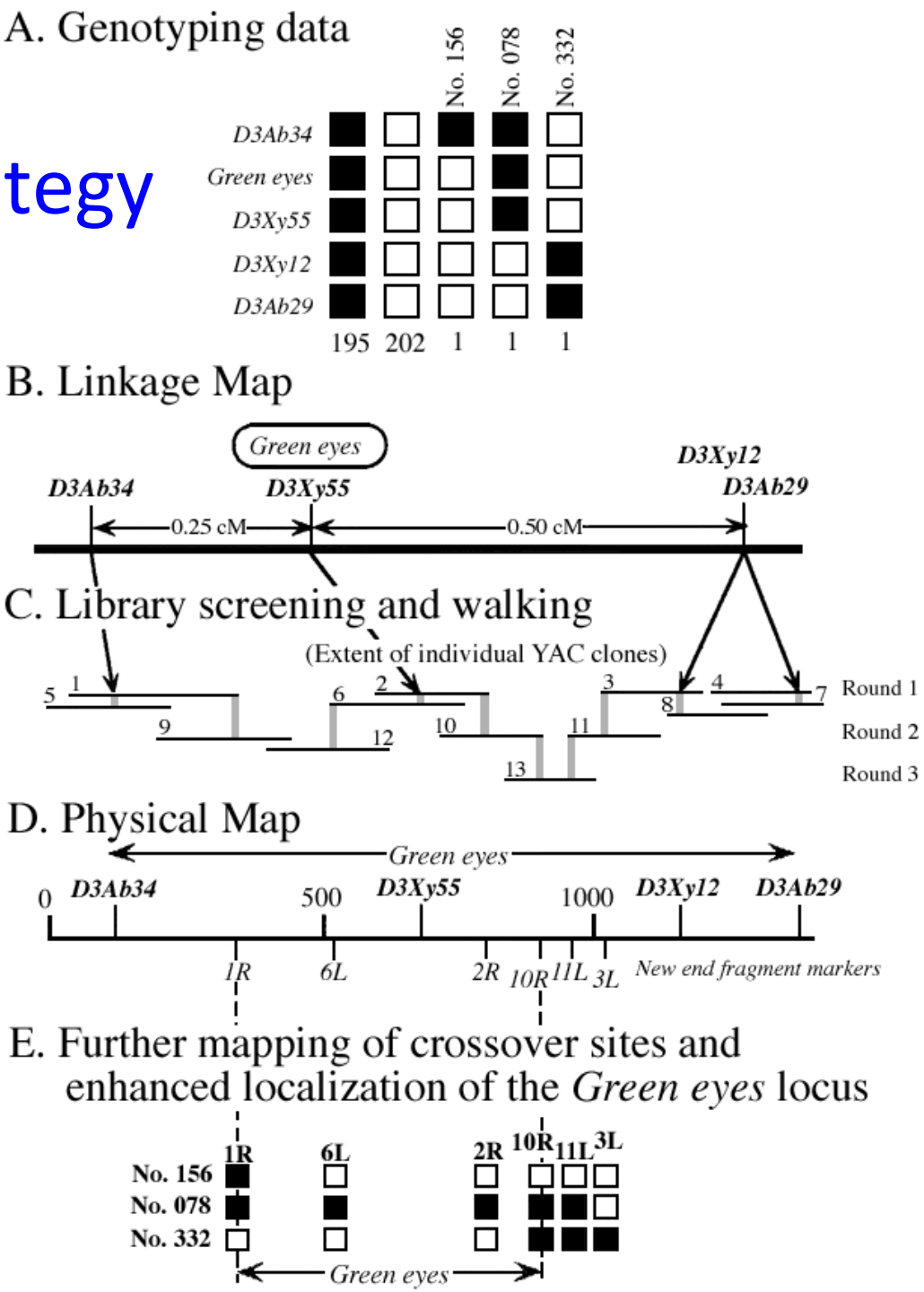


Chromosome Walking




A
15.5 15.3 15.1 14 13 12 11.22 12 13.1 13.3 13.5 14.1 14.3 21 22.2 22.3 23.1 23.3 24 25
yRP16b1
28GA6
C20G10
yRP9f11
yRP6a5
yRP17a3
yRP7e1
942_g_8
24EA4
36FC1
yRP10b9
32DG4
yRP15d11
1EE1
D8E9
17FD3
9AB7
35IC12
OSBP
D11S1335
FCER1
D11S1253
D11S461
D11S1765
CD5
24EA4 L
32DG4 L
1EE1L
32DG4 R
1EE1 R
UGB
AHNAK
35IC12 L
ROM1
B
Best's disease region
CD5
D11S1765
PGA
DDB1
D11S678
24EA4L
D11S734
32DG4L
FEN1
(1EE1L)
D11S699
D11S471
FTH1
32DG4R
1EE1R
UGB
0 200 300 400 500 600 700 800 900 1000 1100 1400 kb
C
dJ16F23
dJ506K16
dJ389K24
dJ387P6
dJ465G21
dJ221L18
dJ71C4
dJ527E4
dJ379A21
dJ455M17
dJ139E20
dJ363P2
D
TU2 (2.8 kb)
TU7 (1.4 kb)
TU1=
DDB1
(4.2 kb)
TU3 (1.4 kb)
TU4 (1.1 kb)
TU5 (4.0 kb)
TU6 (1.9 kb)
TU8
(5.5, 7.0, 7.5 kb)
TU9 (7.5 kb)
TU13
(4.0, 3.1 kb)
TU15
(3.5 kb)
TU16 (1.1 kb)
TU17 (1.6 kb)
TU10 (5.7 kb)
TU11=FEN1 (2.3 kb)
TU12 (4.0 kb)
TU14=FTH1 (1.1 kb)

The overall strategy



HNPCC

An inherited predisposition to colorectal cancer

Maps to chromosome 2p16

Lies between two genetic markers 25cM apart

1cM is roughly 1M bp, so clearly need more markers to delimit the location of HNPCC to a smaller, manageable region

Generation of new polymorphic markers

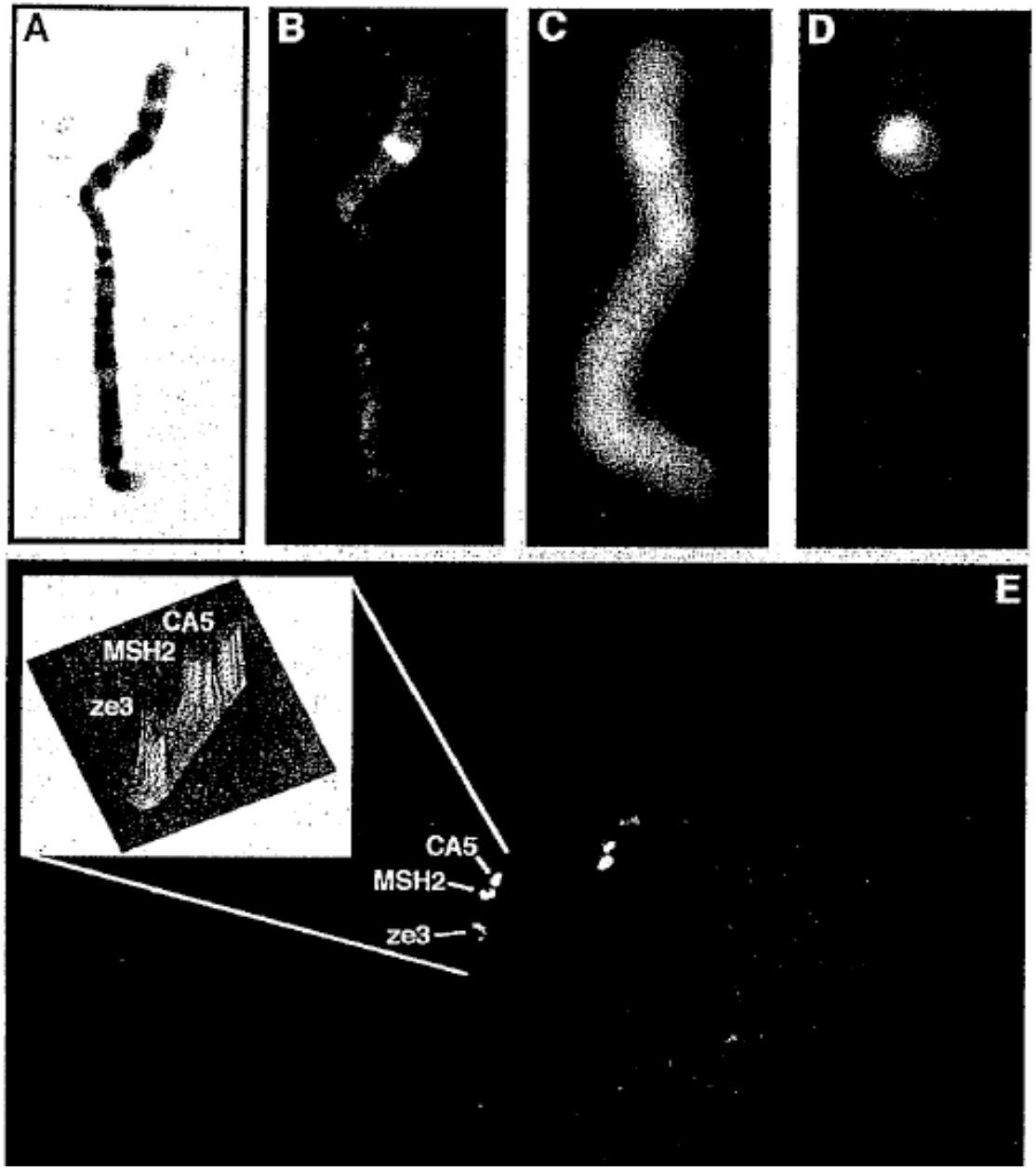


Figure 2. FISH Analysis to Determine the Proximity and Ordering of DNA Sequences within Chromosome Band 2p16

(A, B) FISH mapping of the 123 marker.
(A) G-banded metaphase chromosome 2.
(B) Identical chromosome as in (A) following FISH with a biotin-labeled P1 clone for the 123 marker. Results localize the 123 marker to chromosome band 2p16.
(C, D) Cohybridization documenting the coincident localization of a microdissection (Micro-FISH) probe from chromosome 2p16 and the 123 marker.
(C) DAPI-stained metaphase chromosome 2.
(D) Simultaneous hybridization of the biotin-labeled 123 probe (appearing as an intensely staining smaller circle) and the Spectrum–Orange labeled 2p16 Micro-FISH probe (appearing as a diffusely staining larger circle).
(E) Representative example of an interphase nucleus simultaneously hybridized with P1 clones for CA5, *hMSH2*, and ze3. The results were used to directly measure the distances between markers to establish the order and relative distance between markers (according to Trask et al., 1989).
Inset: The image processing program NIH Image was used to provide an average gray value displayed as a surface plot to support the length measurements and to graphically illustrate the relative order information. The surface plot presented defines the specified interphase chromosome and the relative order CA5–MSH2–ze3.

Can use cloned DNA from 2p16 to look for new polymorphic markers; challenge is to order them

Somatic Cell Hybrids

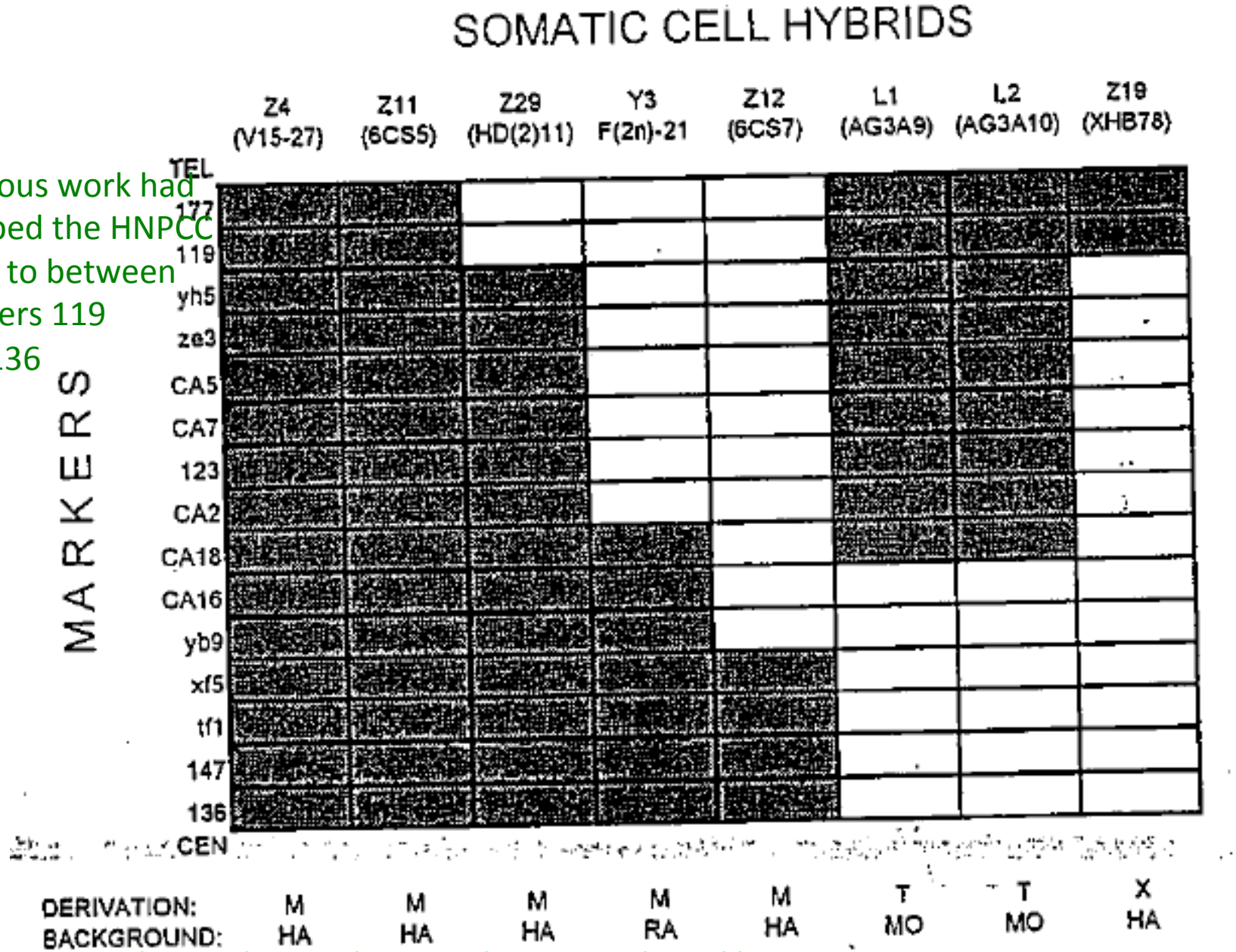
Fuse human and mice (or other rodent) cells`

Human chromosomes are unstable in such hybrids and are lost; eventually the hybrid karyotype stabilizes

Can further alter karyotype such that only part of a particular human chromosome is present, for example by X-ray treatment

Somatic Cell Panel for Quick Mapping of New Polymorphic Loci

Previous work had mapped the HNPCC locus to between markers 119 and 136



New polymorphic markers can be mapped quickly to regions by hybridization to the somatic cell panel

Linkage Analysis of Polymorphic Markers using CEPH Families

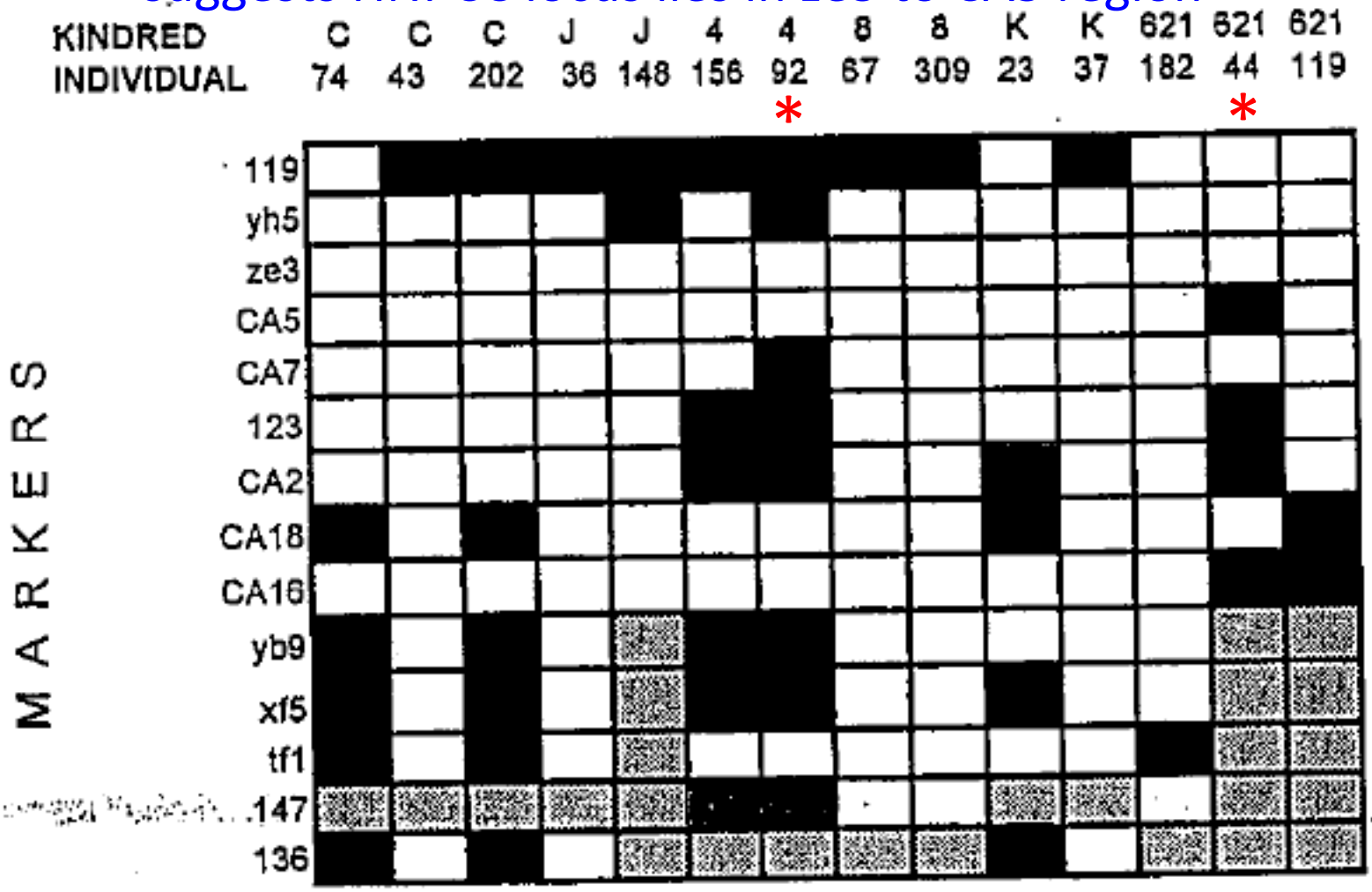
Table 1. Polymorphic Markers Used for Linkage Analysis

Marker	Derivation	cM	Lod	Heterozygosity	YAC Clones	P1 Clones
177 (AFM267zc9)	T	---		0.84		
119 (AFM077yb7)	T	6.1	5.5	0.77	11E1	406
yh5 (AFM337yh5)	T	6.4	15.4	0.76	4F4, 1E1	838, 839, 840
ze3 (AFM200ze3)	T	0.0	----	0.61	4F4,1E1, 9H6, 4A10	836, 837
CA5 (CA5)	M	2.1	4.7	0.77	7F10, 4E2, 5A11	820
CA7 (CA7)	M	1.7	3	0.78	6B8	
123 (AFM093xh3)	T	2.4	9.9	0.76	3D11, 8C7	210, 211
CA2 (CA2)	P	0.0	----	0.75	3D11, 8C7	210, 211
CA18 (CA18)	M	4.3	17.1	0.71	8E5	
CA16 (CA16)	M	0.0	----	0.69	8E5	
yb9 (AFM320yb9)	T	1.1	3.9	0.80	264	
xf5 (AFM310xf5)	T	2.7	17.6	0.76		
tf1 (AFM348tf1)	T	0.0	----	0.79		
147 (AFM199vb6)	T	2.6	9.8	0.73		
136 (AFM172xe7)	T	2		0.73		
134 (AFM168xg11)	T	1		0.76		387, 388, 389

Each of the markers listed was found by screening a genomic library with a $(CA)_n$ probe. The libraries were made from total genomic DNA (T), microdissected human chromosome 2p16 (M), or a genomic P1 clone containing marker 123 (P). The laboratory name and the formal name (in parentheses) of each marker is listed. The genetic distance between the indicated marker and the marker listed above it was determined by linkage analysis in CEPH families, as was the heterozygosity. The odds for pairwise inversion of loci is given in all cases in which these odds were greater than 1000:1. Those YAC and P1 clones obtained in this study that contained marker sequences are also listed.

Together with the somatic cell hybrids, these data give an unambiguous order for the markers

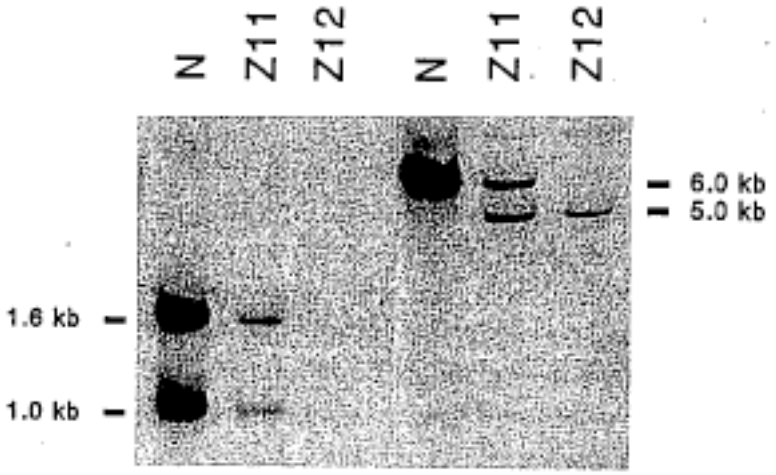
Pattern of markers in affected individuals in families suggests HNPCC locus lies in ze3 to CA5 region

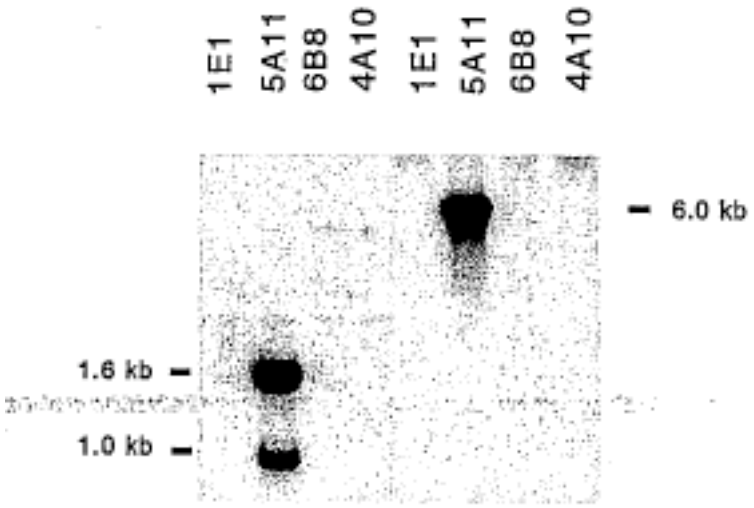


White box: presence of allele associated with disease

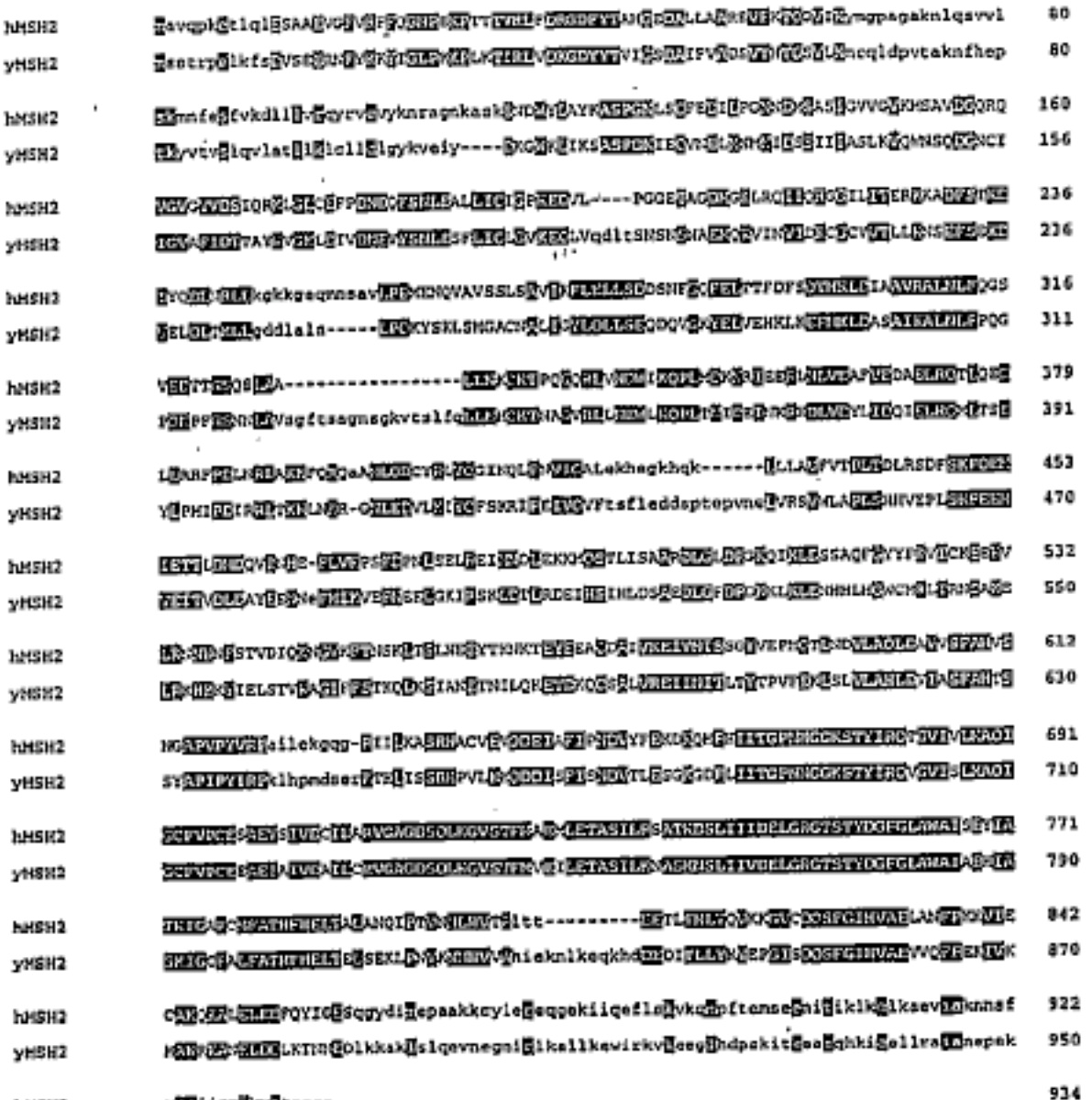
Black box: absence of allele associated with disease

Test Candidate Genes: Homolog to *MSH2* hybridizes to somatic cell hybrids and YACs





Sequence comparison of yeast and human *MSH2* genes



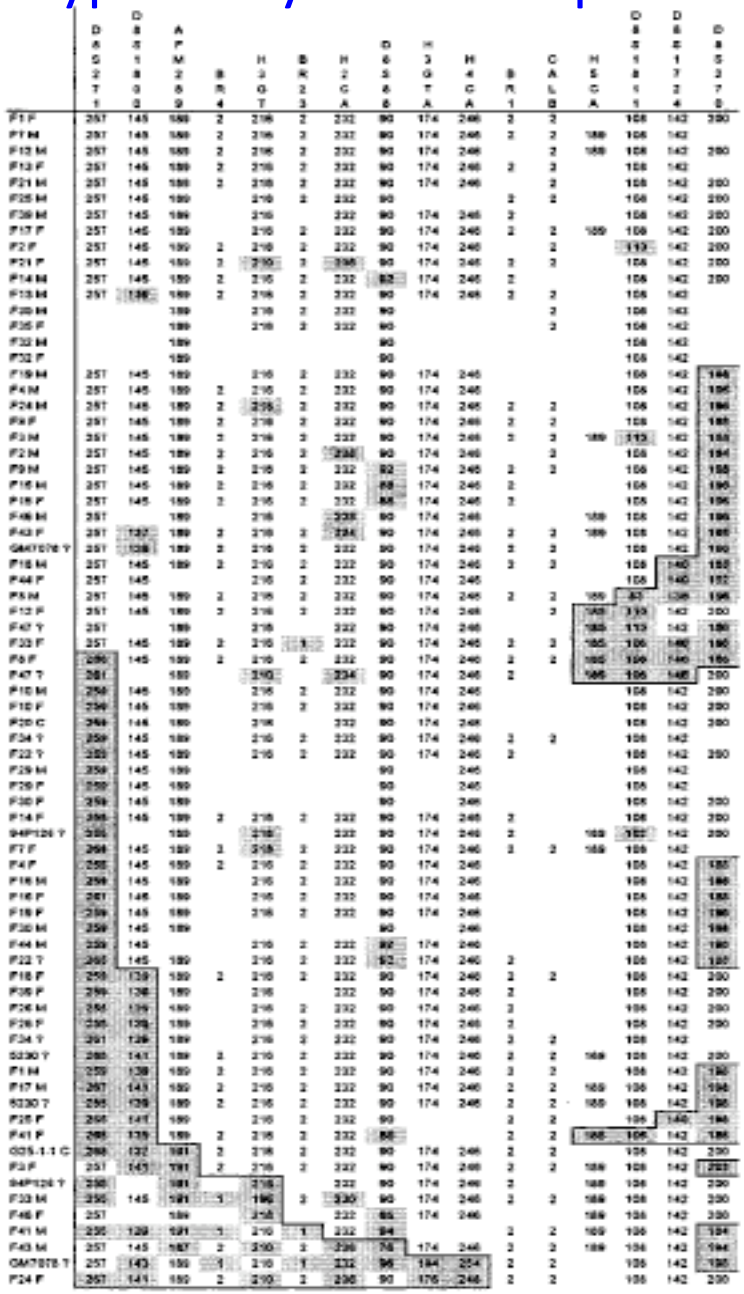
Affected individuals have mutations in the *hMSH2* gene

Table 2. Summary of Mutations

Sample	Source	Type	Codon	cDNA Nucleotide Change	Predicted Coding Change
Family J	HNPCC kindred	Germline	622	CCA to CTA	Proline to leucine
Family C	HNPCC kindred	Germline	265–314	793 to 942 deletion	In-frame deletion
Family 8	HNPCC kindred	Germline	406	CGA to TGA	Arginine to stop
Cx10	RER$^+$ tumor	Germline	639	CAT to TAT	Histidine to tyrosine
		Somatic	663	ATG to TGTG	Frameshift

Mutations were detected by sequencing PCR products, and each was confirmed by replicate PCR analyses and by examination of other affected members of the kindred (see text). The indicated germline mutations were each heterozygous; i.e., lymphocyte DNA contained one wild-type allele and one mutant allele. In Cx10, cloning of the PCR products showed that the somatic mutation at codon 663 occurred in the allele not affected by the germline mutation at codon 639.

Haplotype analysis of NBS patients

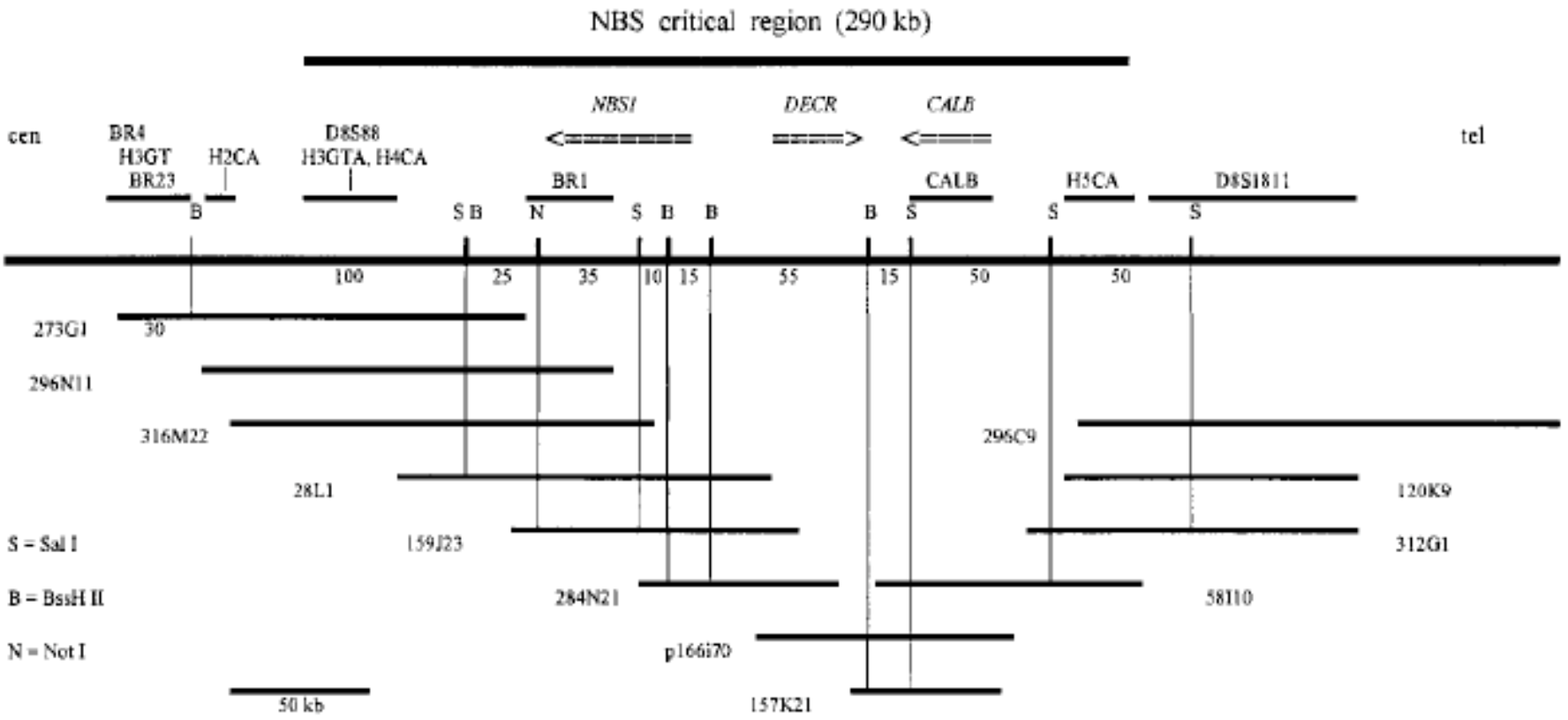


Haplotype analysis, zoomed

NBS

	D8S271	D8S180	AFM289	BR4	H3GT	BR23	H2CA	D8S88	H3GA	H4CA	BR1	CALB	H5CA	D8S1188	D8S1724	D8S270
F1 F	257	145	189	2	216	2	232	90	174	246	2	2		108	142	200
F7 M	257	145	189	2	216	2	232	90	174	246	2	2	189	108	142	
F12 M	257	145	189	2	216	2	232	90	174	246		2	189	108	142	200
F13 F	257	145	189	2	216	2	232	90	174	246	2	2		108	142	
F21 M	257	145	189	2	216	2	232	90	174	246		2		108	142	200
F25 M	257	145	189		216	2	232	90			2	2		108	142	200
F39 M	257	145	189		216		232	90	174	246	2			108	142	200
F17 F	257	145	189		216	2	232	90	174	246	2	2	189	108	142	200
F2 F	257	145	189	2	216	2	232	90	174	246		2		110	142	200
F21 F	257	145	189	2	210	2	238	90	174	246	2	2		108	142	200
F14 M	257	145	189	2	216	2	232	82	174	246	2			108	142	200
F13 M	257	138	189	2	216	2	232	90	174	246	2	2		108	142	
F35 M			189		216	2	232	90				2		108	142	
F35 F			189		216	2	232	90				2		108	142	
F32 M			189					90						108	142	
F32 F			189					90						108	142	
F19 M	257	145	189		216	2	232	90	174	246				108	142	196

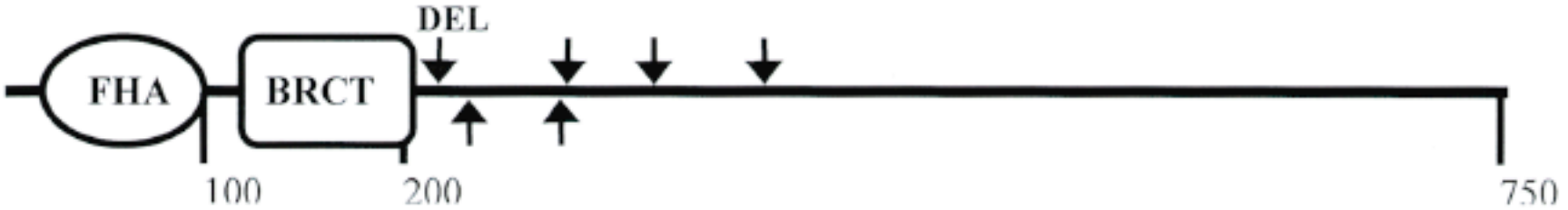
BAC/PAC contig of NBS region



NBS patients have mutations in the *NBS1* gene.
In unrelated patients, 6 different mutations were identified.

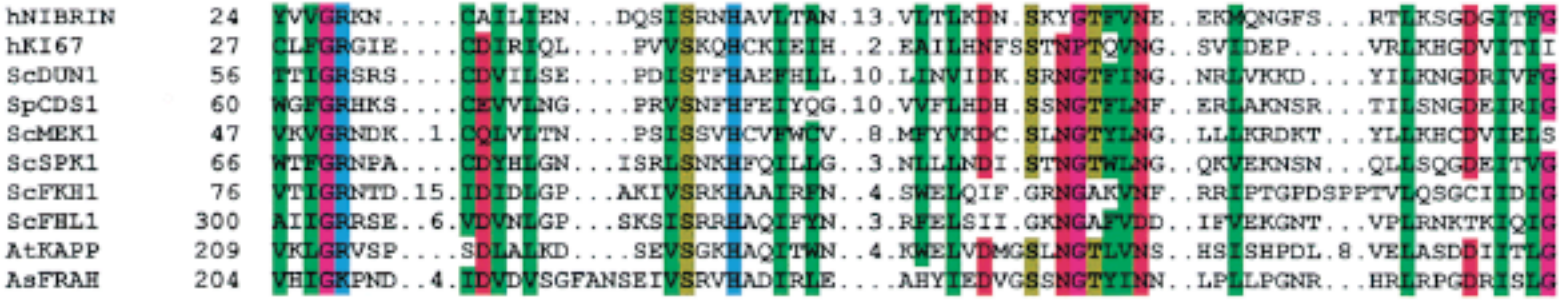
Homologs from yeasts to humans to plants

A

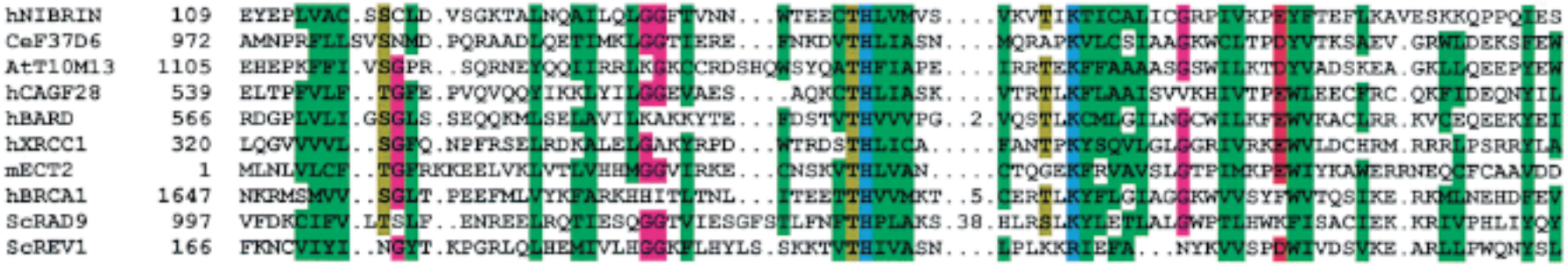


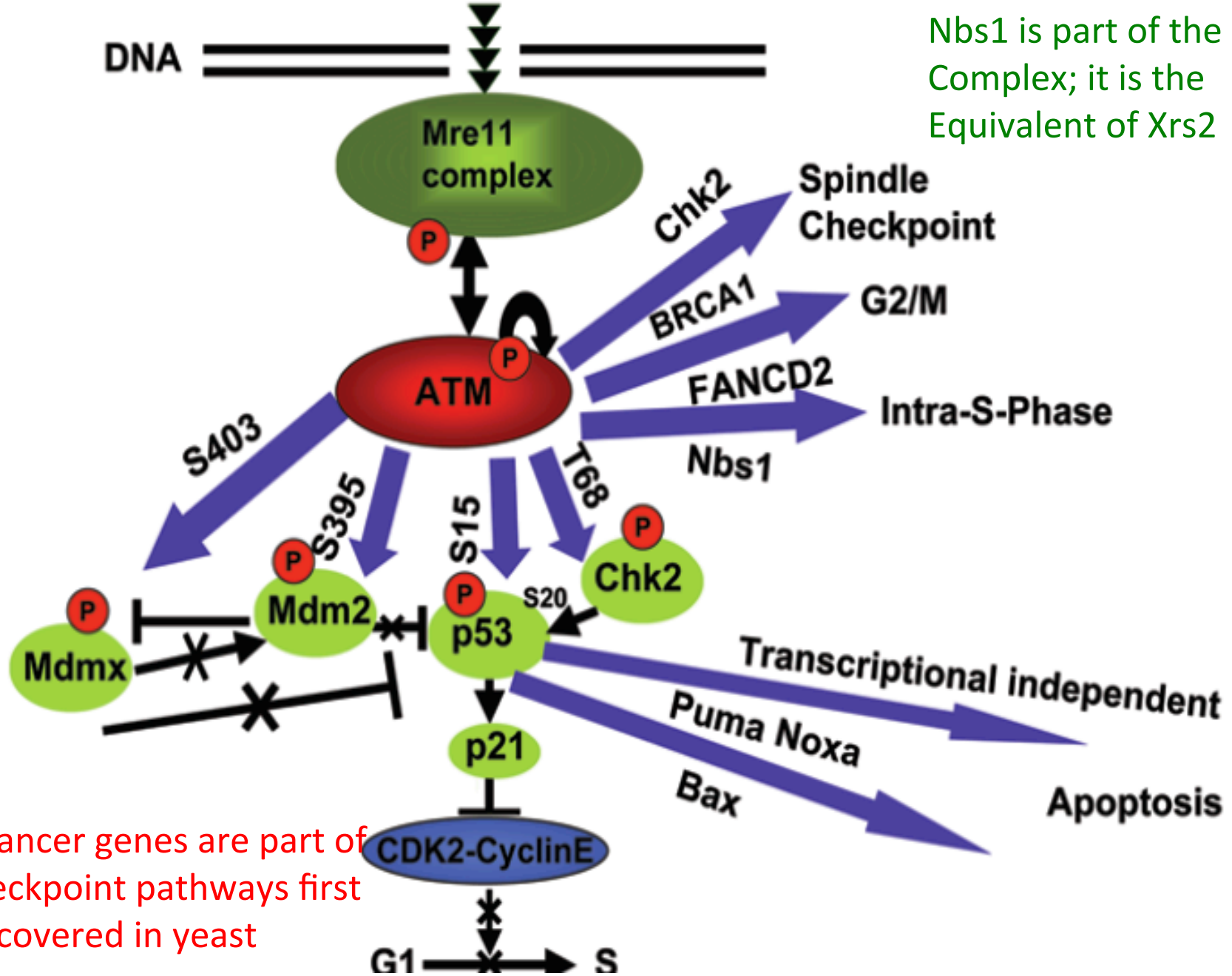
B

Arrows mark the positions of mutations in NBS patients



C





Nbs1 is part of the Mre11 Complex; it is the Equivalent of Xrs2

2 cancer genes are part of checkpoint pathways first discovered in yeast