Colorectal Cancer (CRC)

~5% of Americans will be afflicted with CRC

(1 out of 20 people will get CRC)

•153,000 new cases each year

•56,000 deaths each year

10% to 25% of those cases will be a consequence of an inherited form of the disease

Inherited Colorectal Cancer

Hereditary Non-Polyposis Colorectal Cancer (HNPCC) – the most common (5-13% of all colorectal cancers)

1895
Seamstress to Aldred Warthin predicted her early death to cancer because many in her family had died of the disease.

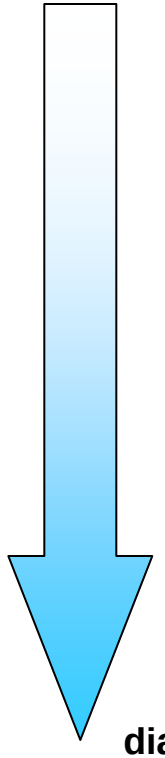

Family G

1913
Using her family, Warthin was the first to document the hereditary nature of this cancer syndrome

Inherited Forms of Colorectal Cancer

- Familial Adenomatous Polyposis (FAP)
- Hereditary Non-Polyposis Colon Cancer (HNPCC) or Lynch Syndrome

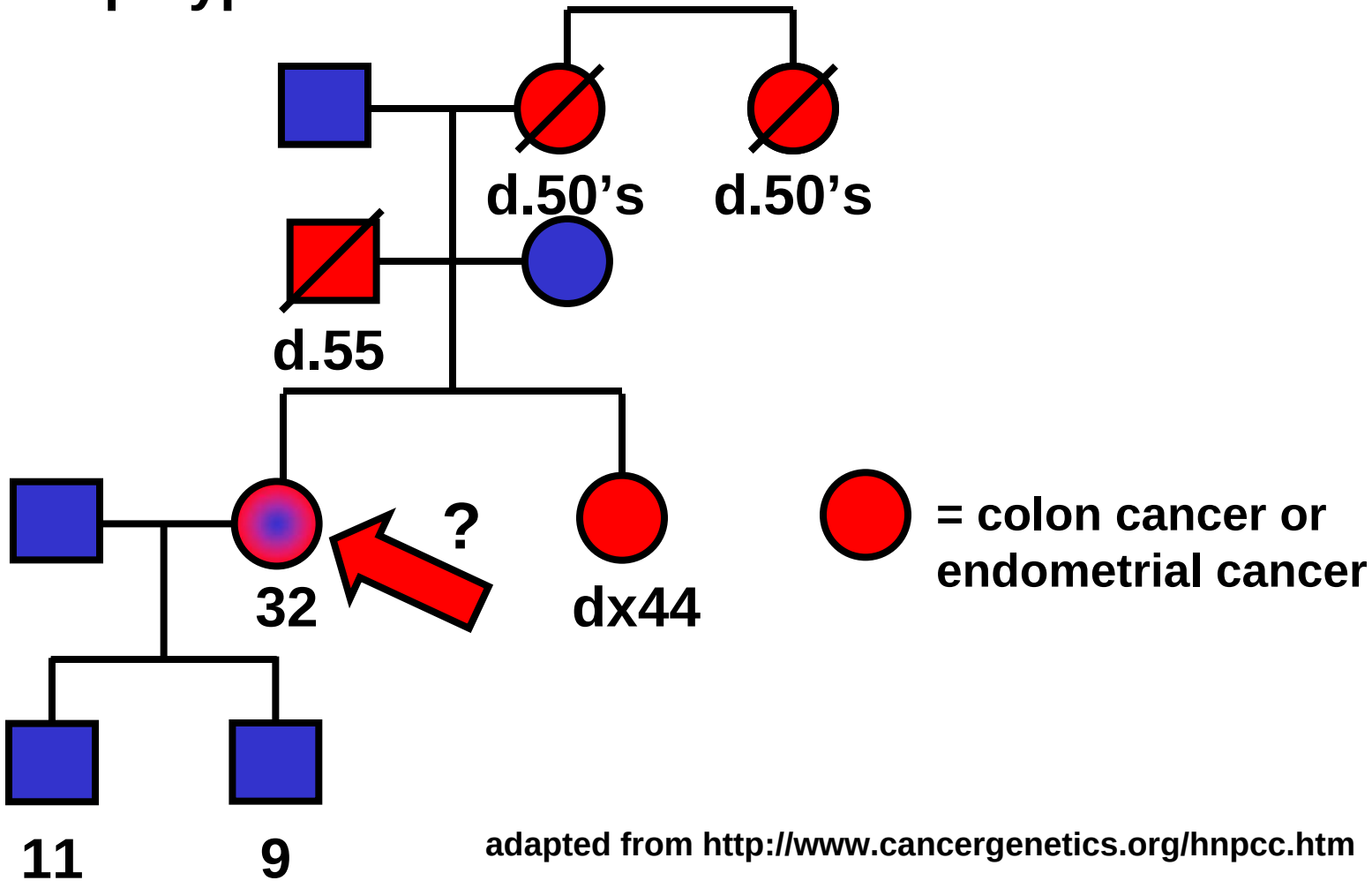
Clinical Feature	HNPCC	FAP
% of all CRC	5-10%	1%
Autosomal Dominance	+	+
Penetrance	80-90%	90-100%
Average age of CRC	<50 y	40 y
Tumor Location	proximal (right)	distal (left)
Polyposis	-	+



diagnostic

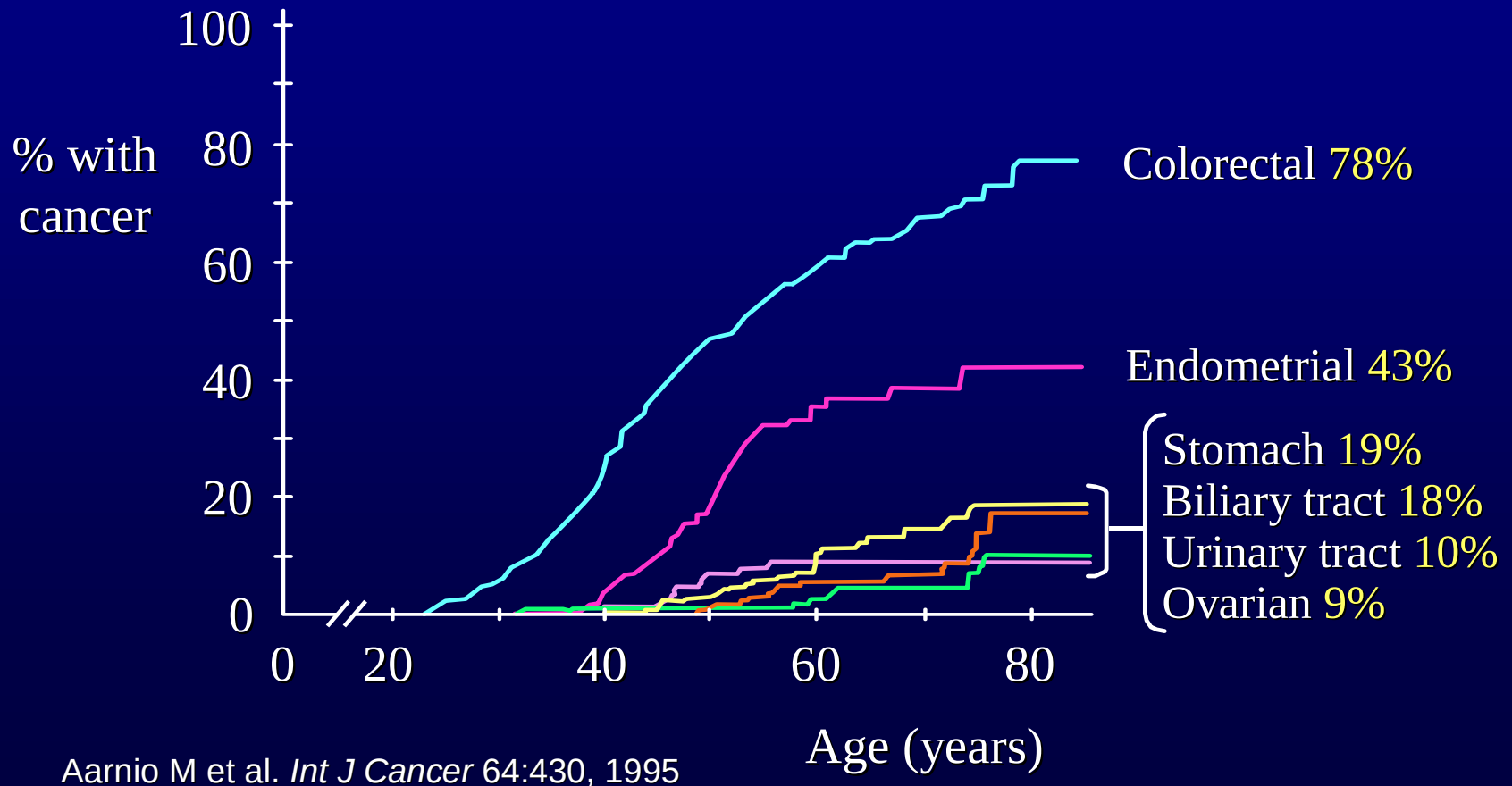
Amsterdam Criteria for Inheritance: HNPCC

- 3 close relatives effected
- 2 generations effected
- 1 relative with the disease before 50 yrs
- Rule out polyposis

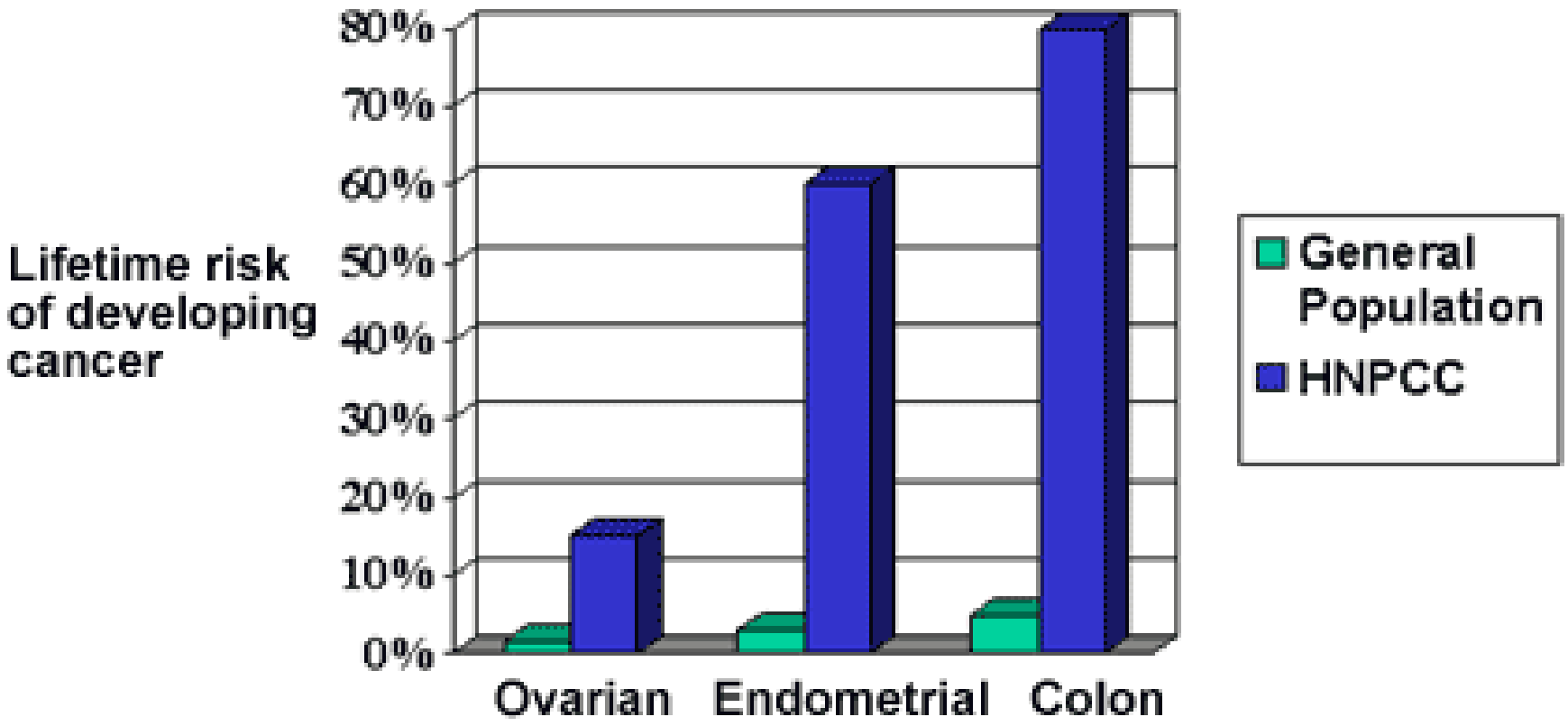


adapted from http://www.cancergenetics.org/hnpcc.htm

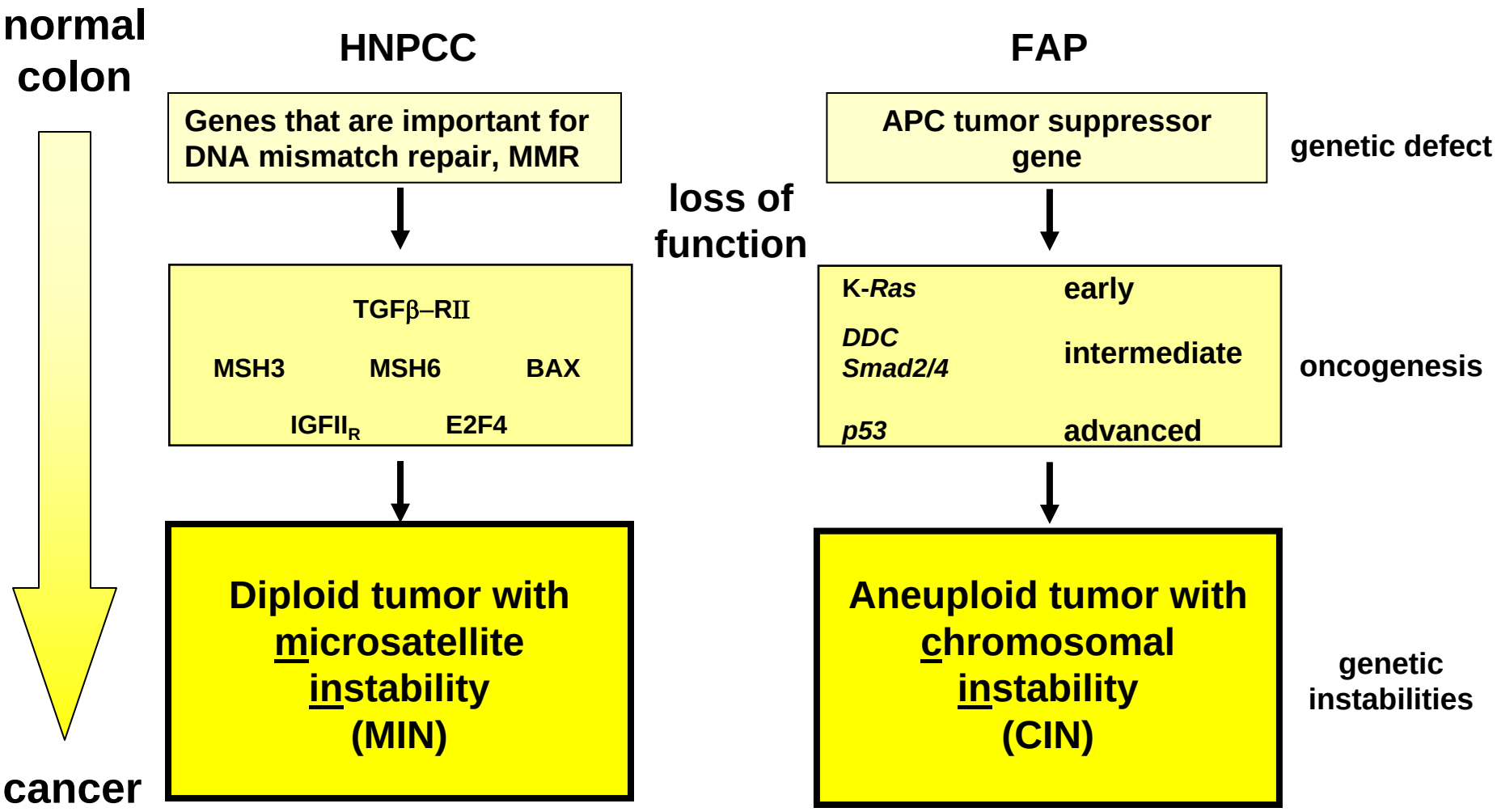
Cancer Risks in HNPCC



Aarnio M et al. *Int J Cancer* 64:430, 1995


Lifetime risk of developing cancer
80%
70%
60%
50%
40%
30%
20%
10%
0%
Ovarian
Endometrial
Colon
General Population
HNPCC

Differences in FAP and HNPCC: genetic defect, oncogenesis and genetic instabilities (MIN vs. CIN)



http://www.hosppract.com/genetics/9711gen.htm
and Duval et al. 2002. Annales de Genetique

Microsatellite Instability

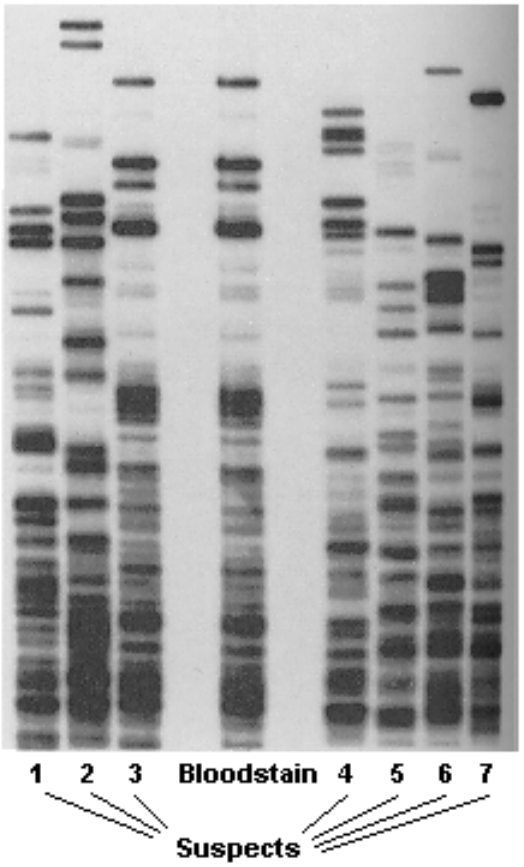
Repetitive DNA (microsatellite DNA) is normally more unstable than other regions of the genome

found in all organisms

function is controversial

used in mapping differences in individuals:

- forensics
- medical pedigree analyses
- paternity cases



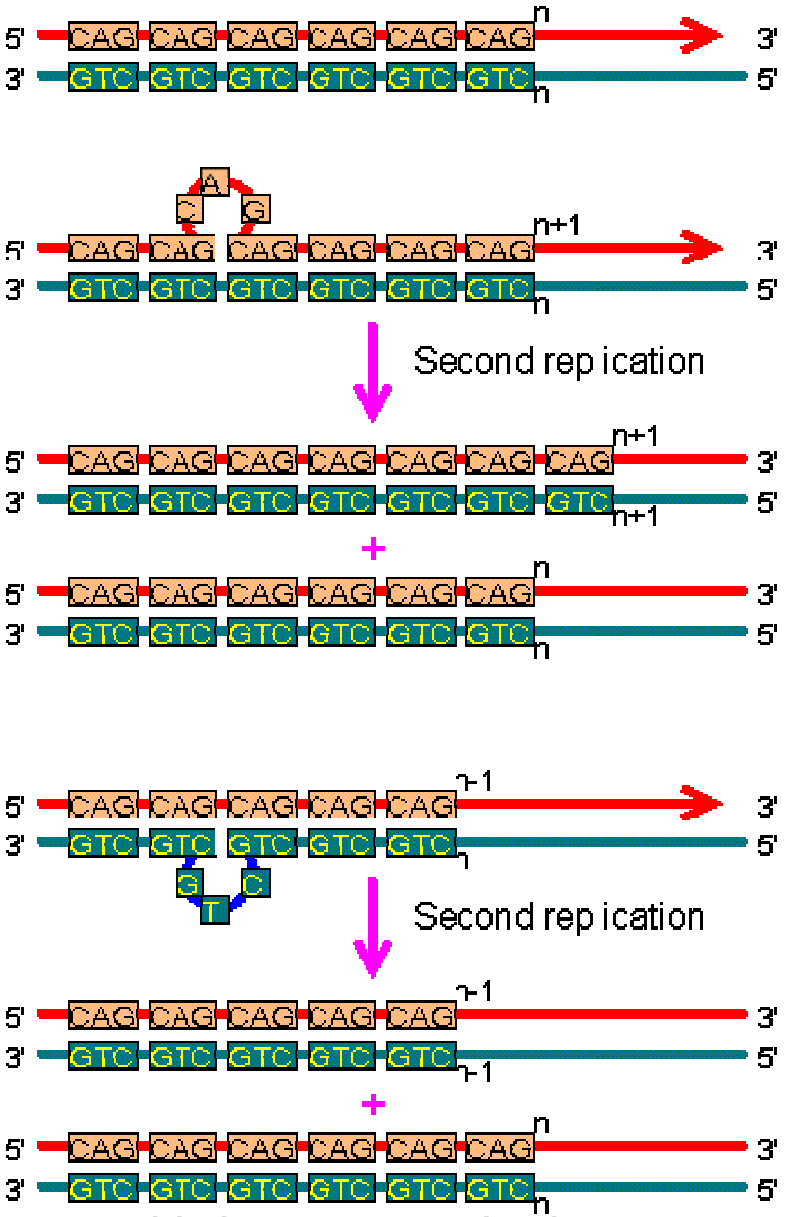
http://www.people.virginia.edu/~rjh9u/gif/forens1.gif

Why is repetitive DNA more unstable?

•During replication repetitive DNA elements can result in slippage of the polymerase.

•As a consequence the new strand of DNA has either more or less of the repetitive DNA encoded compared to the template strand.

•If not repaired, this will become a stable insertion or deletion after the next round of replication.

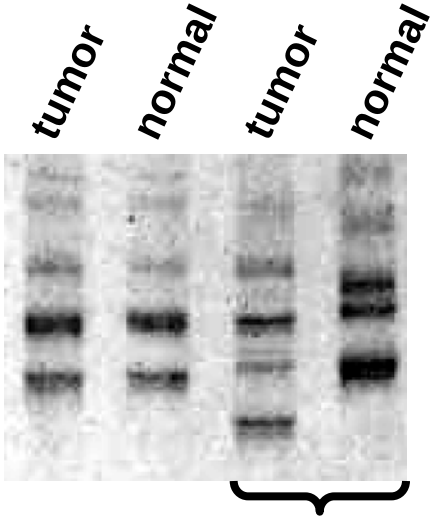


http://www.web-books.com/MoBio/Free/Ch7F3.htm http://www.web-books.com/MoBio/Free/Ch7F3.htm

HNPCC afflicted individuals show an increase in microsatellite instability in tumor vs. normal tissue (100-fold)

Parsons *et al.* 1993

Microsatellite Instability (MIN) Assay: Tumor *vs.* Normal Tissue Samples



RER+ (Replication Error +)

Photo from:
http://zapruder.pds.med.umich.edu/users/Frank/MIN.html

Tumors from sporadic cancers also show MIN or RER+ genetic instabilities (13-25%). First shown by Ionov *et al.* 1993

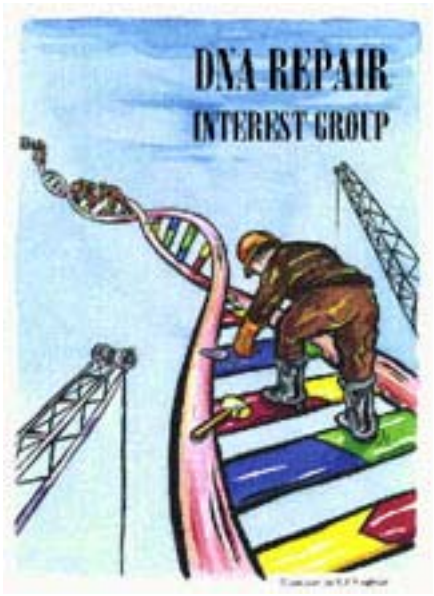
Discovery that HNPCC was a consequence of a failure in the DNA mismatch repair machinery

1960's and 1970's

- **DNA mismatch repair was characterized in bacteria. Important genes were identified: *mutH, mutS* and *mutL*** (Siegel and Bryson, 1964; Cox et al. 1972)

Late 1980's

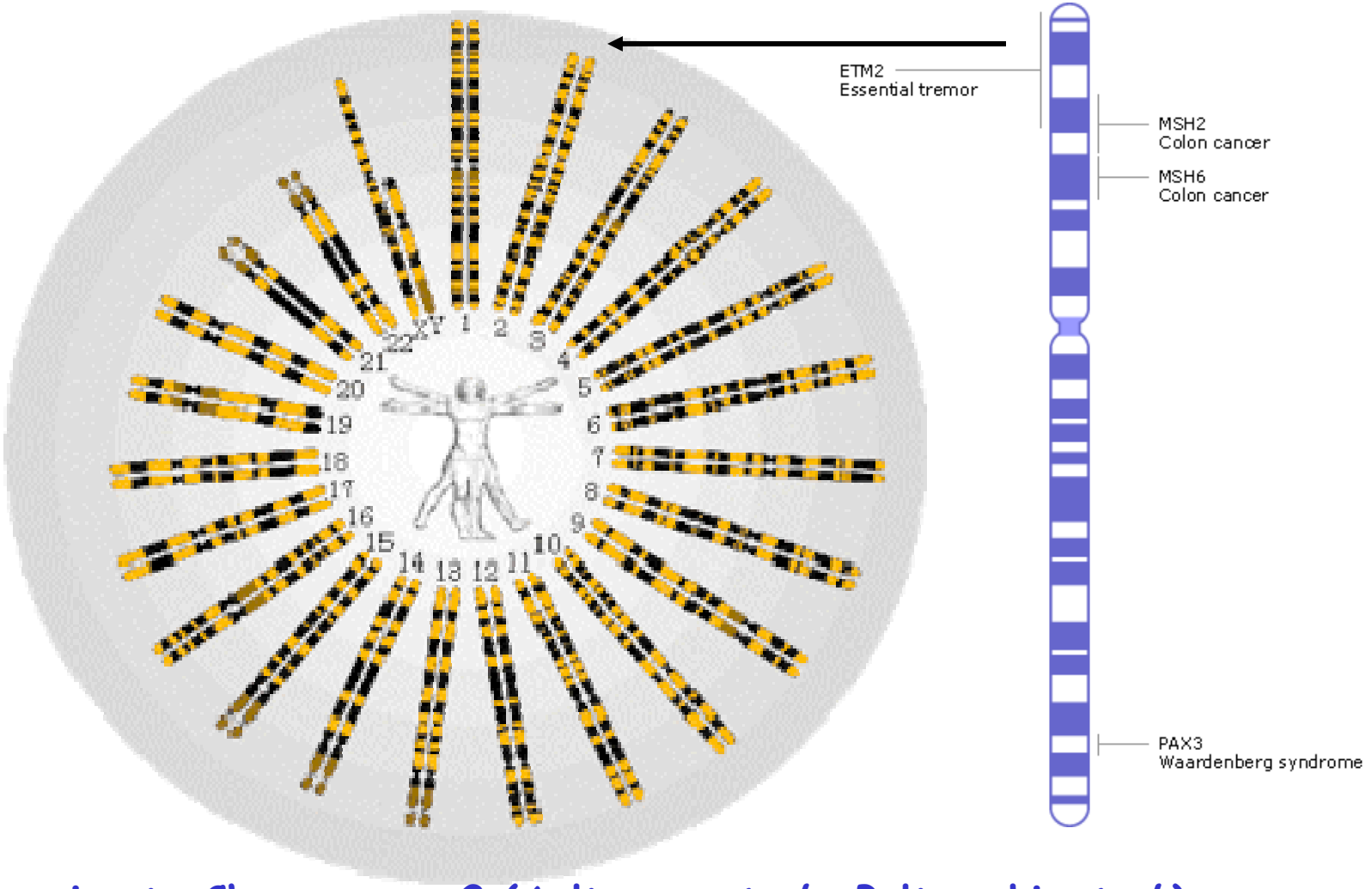
- **Levinson and Gutman noticed *mutS* mutants exhibited increased dinucleotide repeat instability**
- **DNA mismatch repair systems were found in other organisms (e.g. yeast and drosophila)**



Early 1990's

- **DNA mismatch repair homologs were found in yeast (*MSH1-6*)** (Reenan and Kolodner 1993)
- **Strand *et al.* showed that dinucleotide instability was increased in mismatch repair mutant yeast strains.**

1993, exactly 80 years after Warthin's publication the faulty gene in Family G was identified



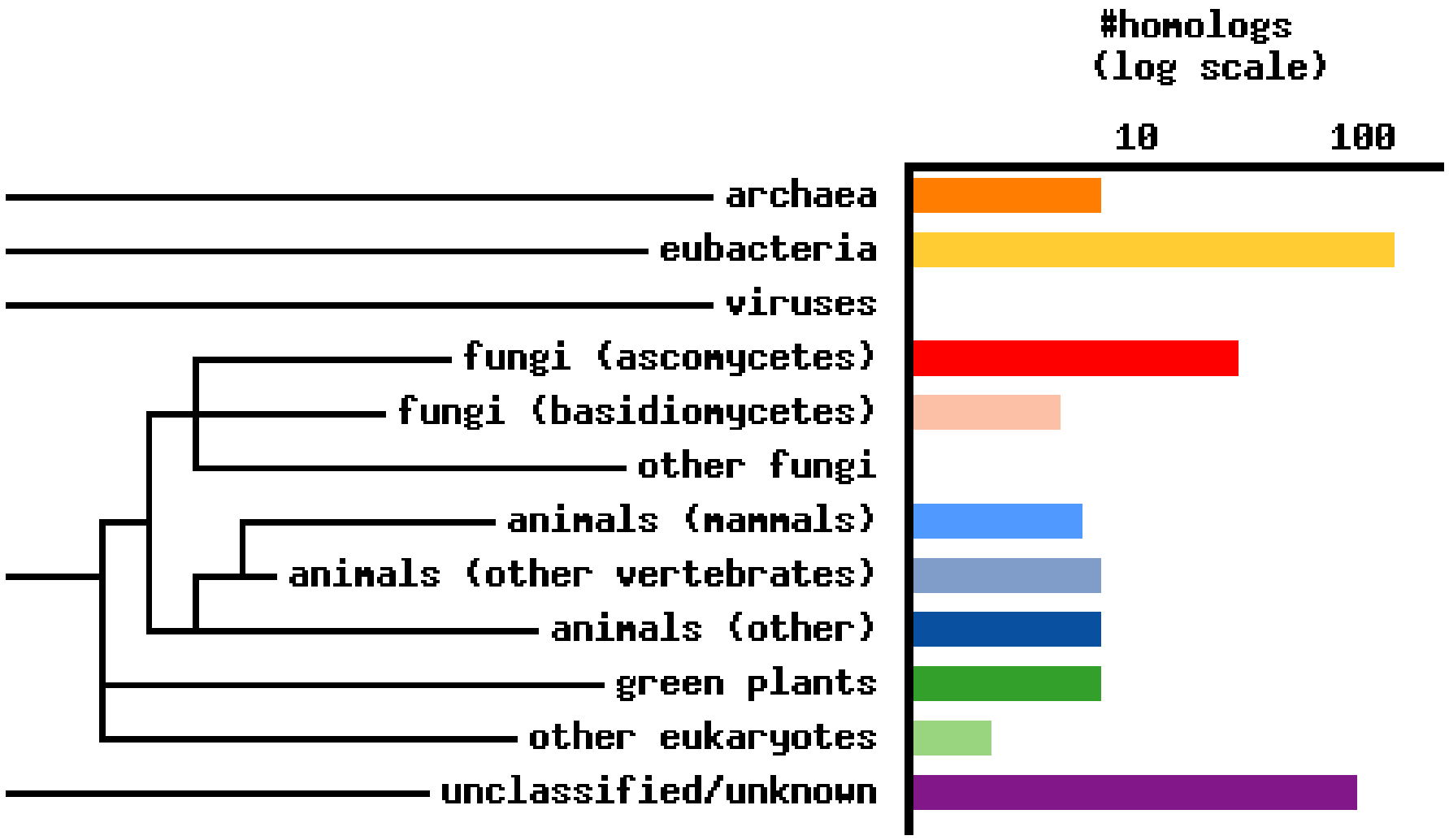
Classical mapping to Chromosome 2 (Aaltonen *et al.*; Peltomaki *et al.*)

Positional cloning (Leach *et al.*) based on mapping data.

Candidate gene approach (Fishel *et al.*) based on homology to y*MSH2* and *mutS*.

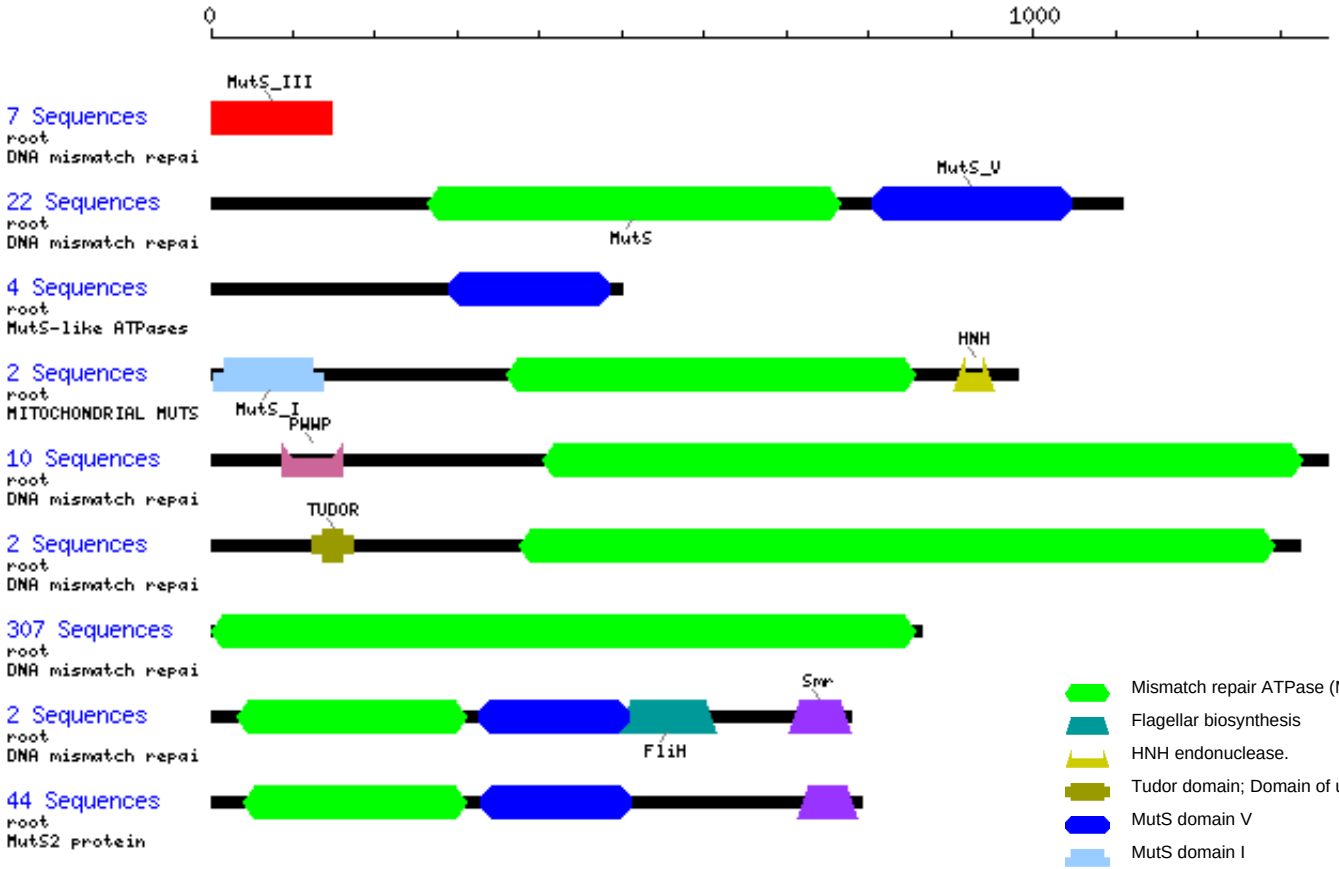
http://www.ornl.gov/sci/techresources/Human_Genome/launchpad/

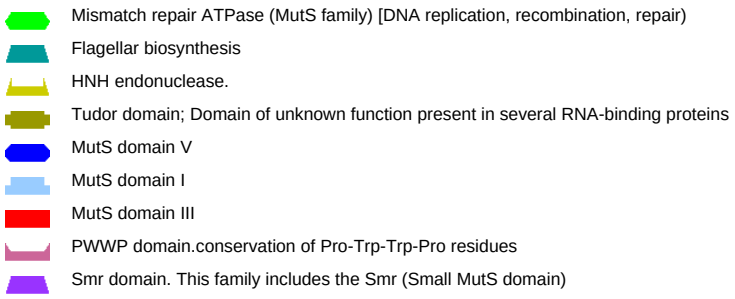
MSH, or MutS homologs found in all branches of the phylogenetic tree



http://db.yeastgenome.org/cgi-bin/homolog/uniprotHomolog?locus=YOL090W

The homology is based on amino acid sequence and function

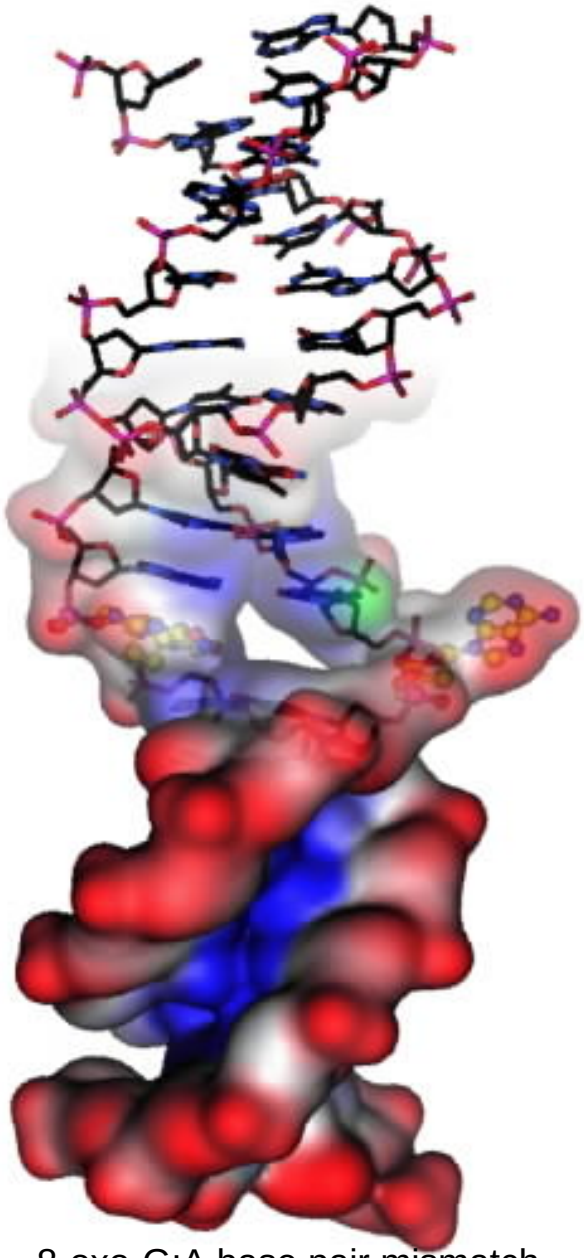




http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi?cmd=cdd&uid=18015

The DNA mismatch repair complex recognizes helical distortions arising from

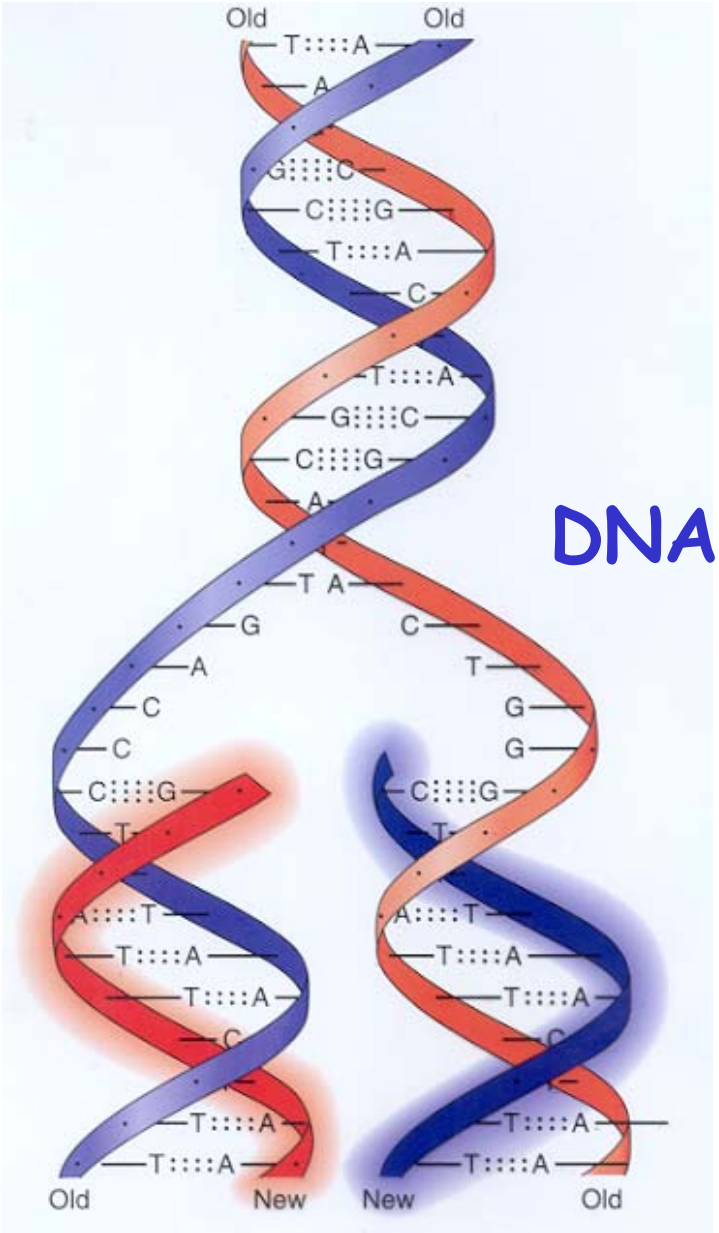
- replication errors
- recombination
- DNA damage



8-oxo-G:A base pair mismatch
from oxidative stress

http://www.chemistry.mcmaster.ca/faculty/berti/DNA_flipped_A_web.JPG

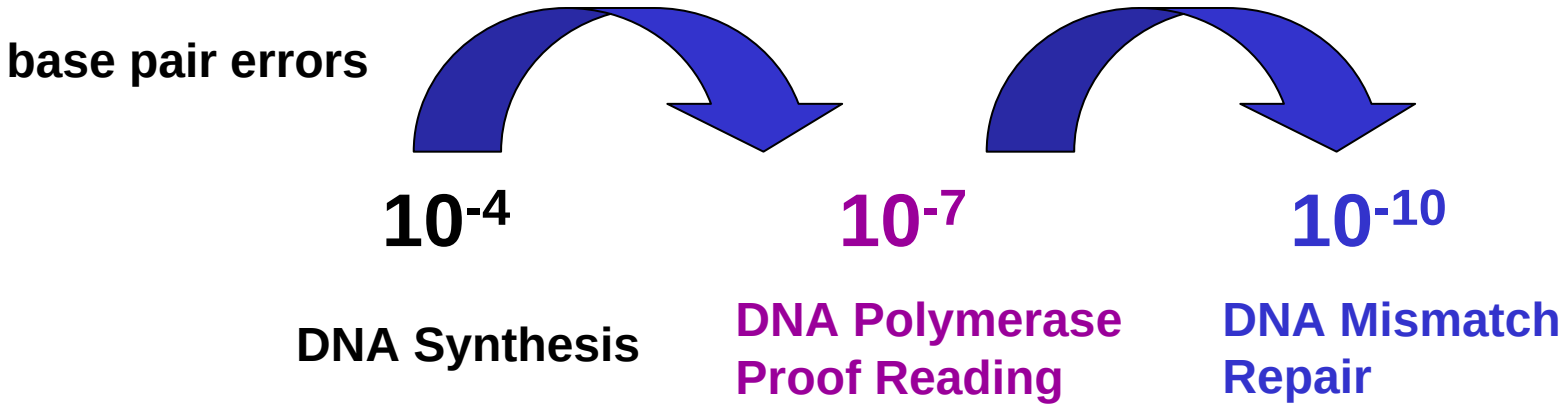
Most common cause of mismatches



DNA replication

DNA Mismatch Repair (MMR)

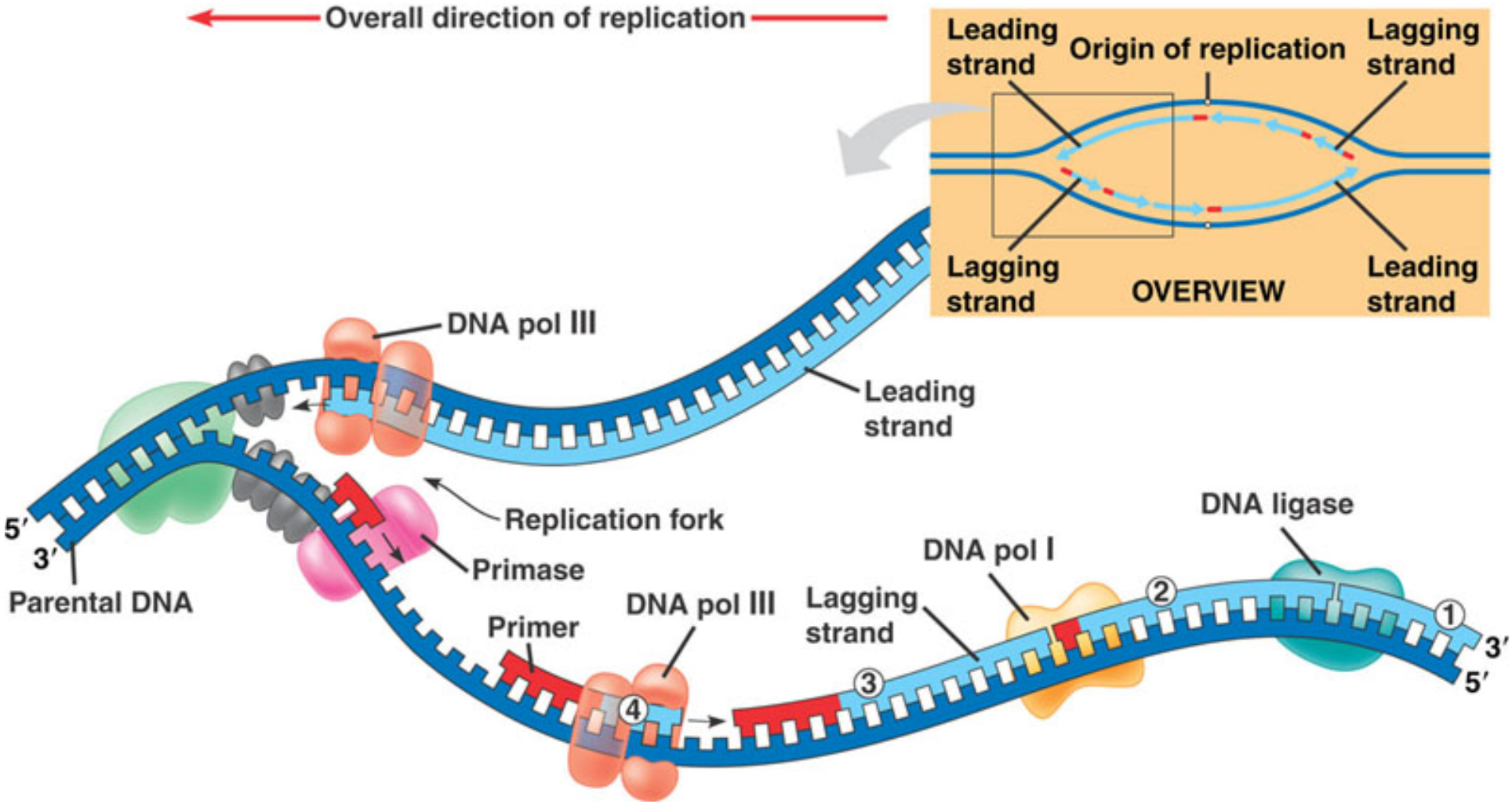
12 billion nucleotides must be replicated per cell division in humans



If the error rate is 10^{-4} then there will be ~1,200,000 errors per cell division

If the error rate is 10^{-7} then there will be ~1,200 errors per cell division

If the error rate is 10^{-10} then there will be ~1 error per cell division


Overall direction of replication
Leading strand
Origin of replication
Lagging strand
Lagging strand
OVERVIEW
Leading strand
DNA pol III
Leading strand
5′
3′
Parental DNA
Replication fork
Primase
Primer
DNA pol III
Lagging strand
DNA pol I
DNA ligase
1
2
3
4
3′
5′

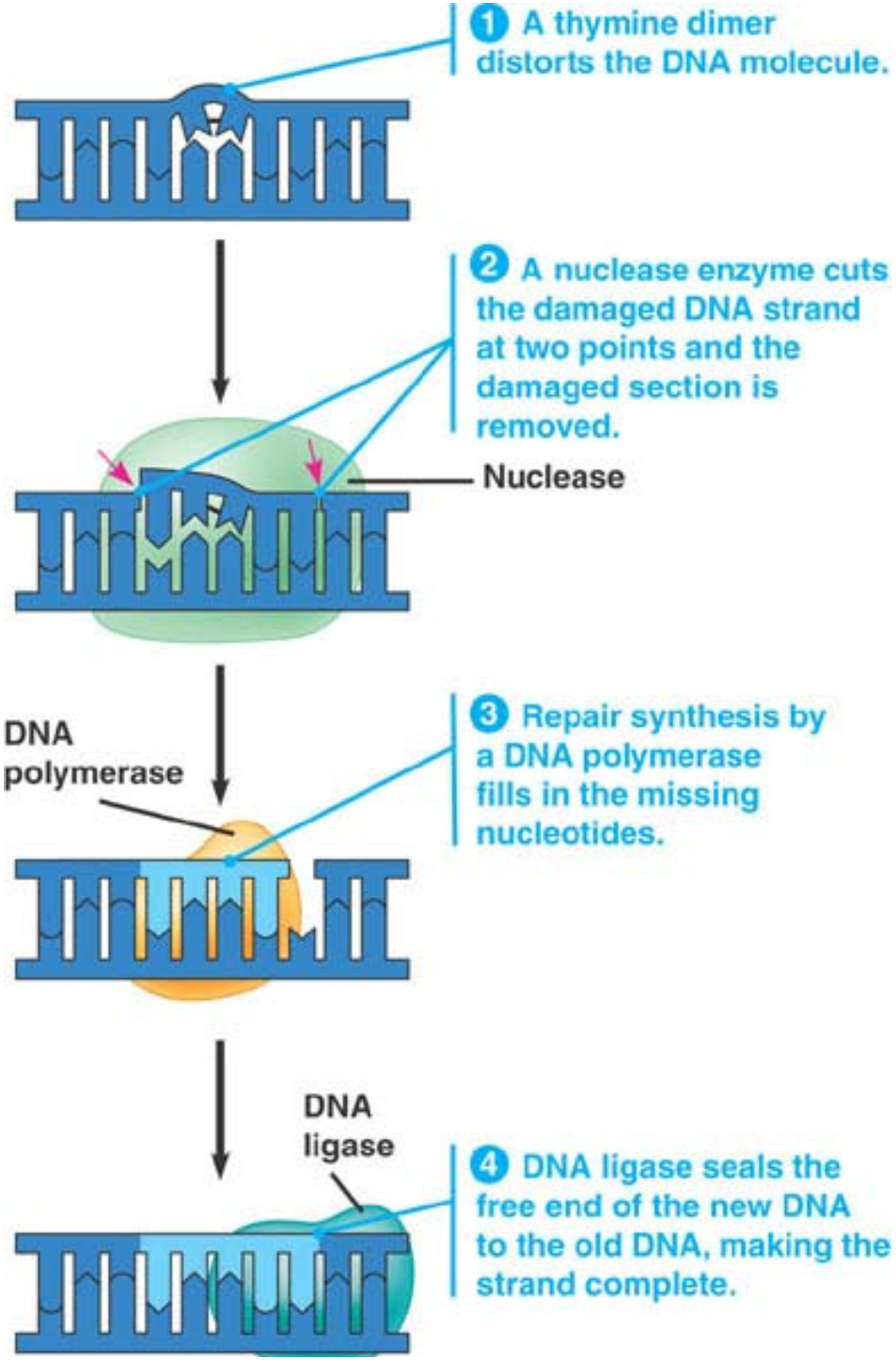
All of the proteins we need

Table 16.1 Bacterial DNA replication proteins and their functions

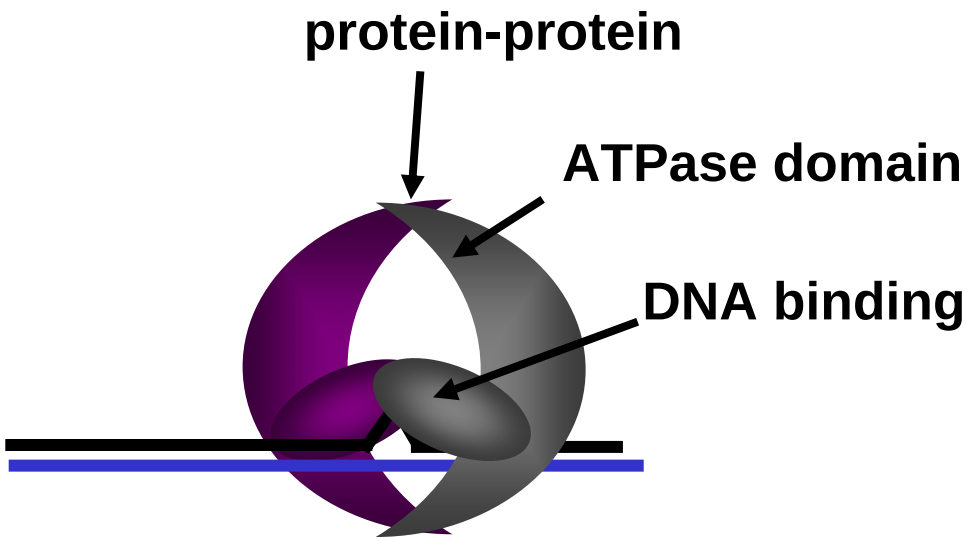
Protein	Function for Leading and Lagging Strands	
Helicase	Unwinds parental double helix at replication forks	
Single-strand binding protein	Binds to and stabilizes single-stranded DNA until it can be used as a template	
Topoisomerase	Corrects "overwinding" ahead of replication forks by breaking, swiveling, and rejoining DNA strands	
	Function for Leading Strand	**Function for Lagging Strand**
Primase	Synthesizes a single RNA primer at the 5′ end of the leading strand	Synthesizes an RNA primer at the 5′ end of each Okazaki fragment
DNA pol III	Continuously synthesizes the leading strand, adding on to the primer	Elongates each Okazaki fragment, adding on to its primer
DNA pol I	Removes primer from the 5′ end of leading strand and replaces it with DNA, adding on to the adjacent 3′ end	Removes the primer from the 5′ end of each fragment and replaces it with DNA, adding on to the 3′ end of the adjacent fragment
DNA Ligase	Joins the 3′ end of the DNA that replaces the primer to the rest of the leading strand	Joins the Okazaki fragments

Nucleotide excision repair of DNA

- Thymine intrastrand binding → ERROR
- Nuclease
  - Removes the error
- Polymerase
- DNA ligase

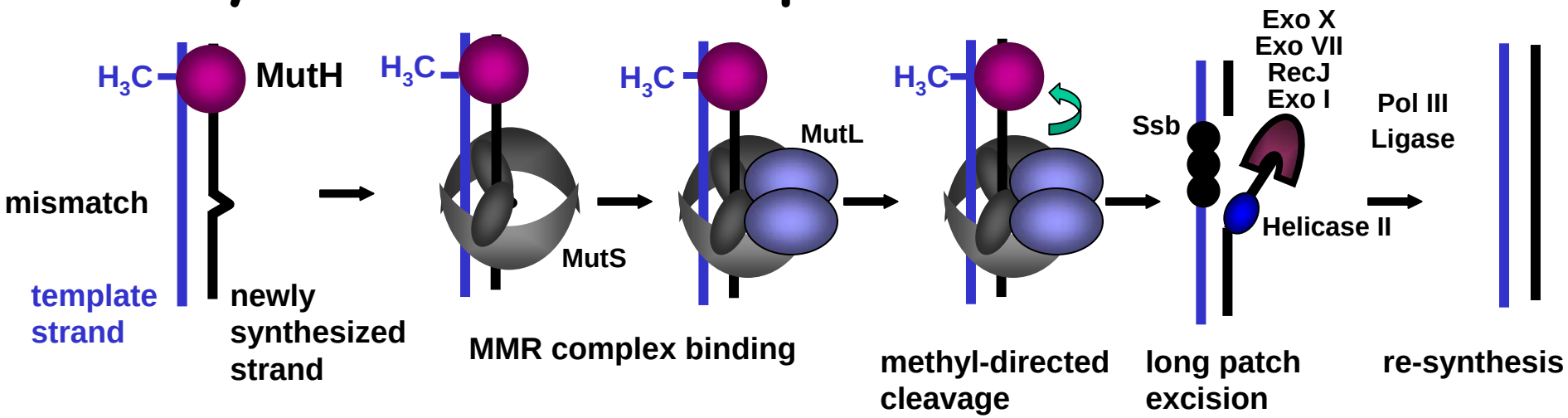


Msh2p Important Domains

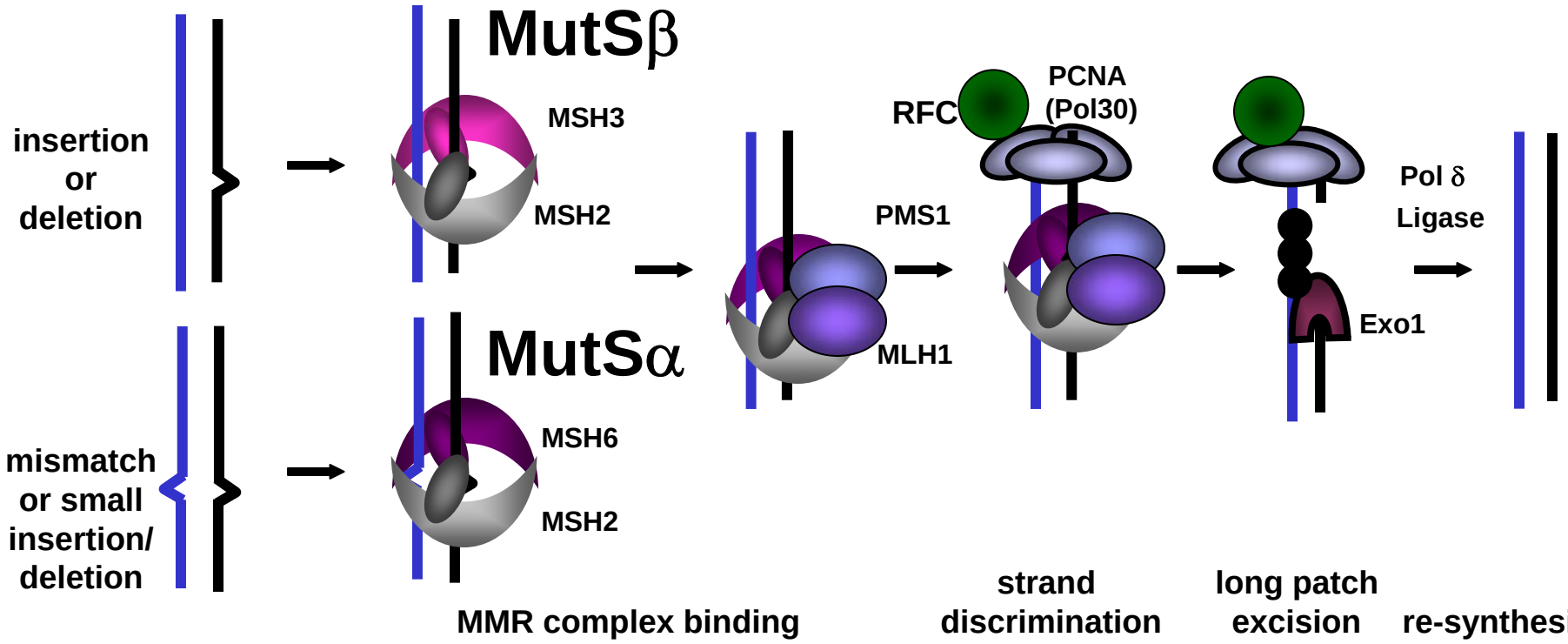


- Structural integrity residues (throughout)
- DNA binding region
- Protein-protein interacting regions
- ATPase domain

Prokaryotic Mismatch Repair

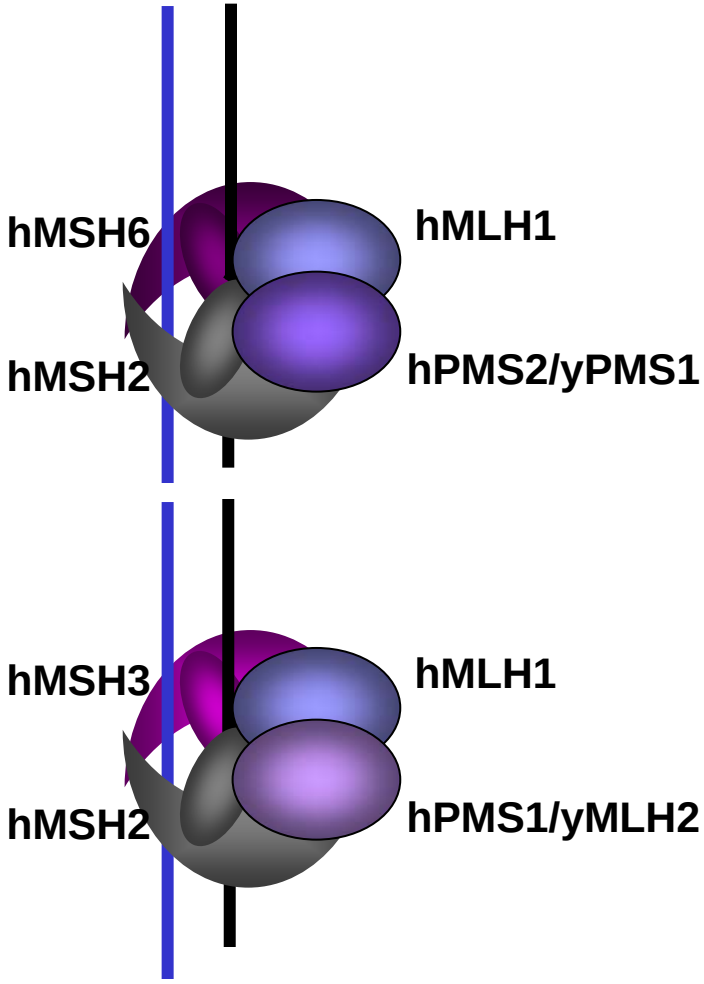


Eukaryotic Mismatch Repair



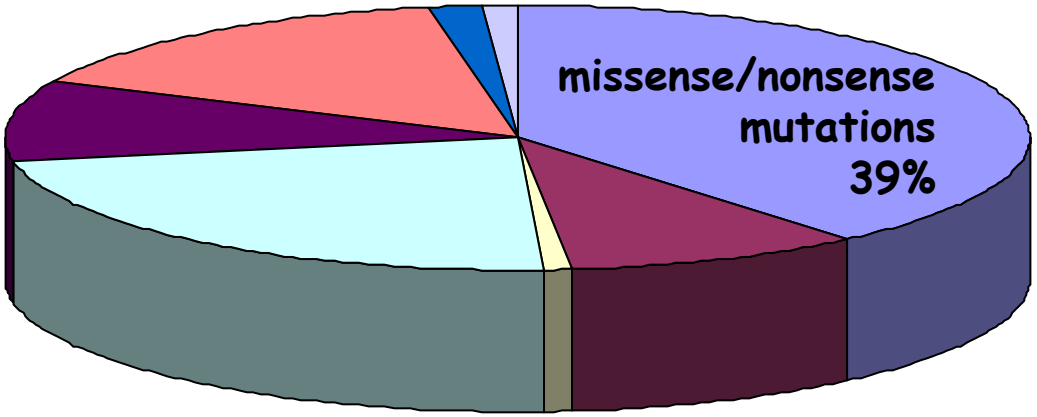
Human germline lesions in DNA mismatch repair genes associated with HNPCC

Human Gene	% of total HNPCC disease alleles
MLH1	33-35%
MSH2	30-31%
PMS1	<1%
PMS2	<5%
MSH6	2-5%
unknown	*25-30%*



Adapted from Stern and Lagarde *Can J Surg* (1998) 41:345-9 and Lui et al. Nature Medicine (1996) 92:169-174

Clinically Identified Mutations in hMSH2



- Nucleotide substitutions (missense / nonsense)
- Nucleotide substitutions (splicing)
- Nucleotide substitutions (regulatory)
- Small deletions
- Small insertions
- Gross deletions
- Gross insertions & duplications
- Complex rearrangements (including inversions)

http://uwcmml1s.uwcm.ac.uk/uwcm/mg/search/203983.html

Many afflicted families have mutations in MSH2
345 mapped (as of Jan 4, 2007)

~75% cause an obvious change

NO PROTEIN PRODUCED

~25% cause a very small change

MISSENSE MUTATIONS

Can we be sure that the "little" changes cause the disease?

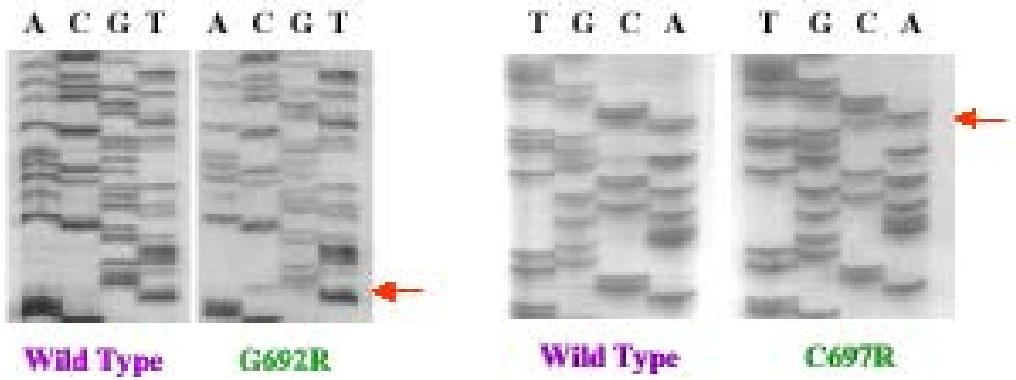
we must be sure for genetic counseling

Codon	Nucleotide	Amino acid	Phenotype
1	cATG-CTG	Met-Leu	Colorectal cancer, non-polyposis
8	ACG-ATG	Thr-Met	Colorectal cancer, non-polyposis
44	ACG-ATG	Thr-Met	Colorectal cancer, non-polyposis
45	GCG-GTG	Ala-Val	Colorectal cancer, non-polyposis
107	aGCT-CCT	Ala-Pro	Colorectal cancer, non-polyposis
127	AAT-AGT	Asn-Ser	Colorectal cancer, non-polyposis
145	ATTg-ATG	Ile-Met	Colorectal cancer, non-polyposis
167	gGAT-CAT	Asp-His	Colorectal cancer, non-polyposis
198	GAA-GGA	Glu-Gly	Colorectal cancer, non-polyposis
199	aTGT-CGT	Cys-Arg	Glioma
305	tGCA-ACA	Ala-Thr	Colorectal cancer, non-polyposis
322	GGC-GAC	Gly-Asp	Colorectal cancer, non-polyposis ?
323	TCT-TGT	Ser-Cys	Colorectal cancer, non-polyposis
333	gTGT-CGT	Cys-Arg	Colorectal cancer, non-polyposis
333	TGT-TAT	Cys-Tyr	Colorectal cancer, non-polyposis
336	cCCT-TCT	Pro-Ser	Colorectal cancer, non-polyposis
390	aCTT-TTT	Leu-Phe	Colorectal cancer, non-polyposis
440	CTT-CCT	Leu-Pro	Colorectal cancer, non-polyposis
506	gGAC-TAC	Asp-Tyr	Colorectal cancer, non-polyposis
524	CGT-CCT	Arg-Pro	Colorectal cancer, non-polyposis
562	GAG-GTG	Glu-Val	Colorectal cancer, non-polyposis
622	CCA-CTA	Pro-Leu	Colorectal cancer, non-polyposis
636	aGCA-CCA	Ala-Pro	Colorectal cancer, non-polyposis
639	CAT-CGT	His-Arg	Colorectal cancer, non-polyposis
639	gCAT-TAT	His-Tyr	Colorectal cancer, non-polyposis
647	tGAA-AAA	Glu-Lys	Colorectal cancer, non-polyposis
674	GGT-GAT	Gly-Asp	Colorectal cancer, non-polyposis
688	ATGg-ATA	Met-Ile	Colorectal cancer, non-polyposis
692	tGGG-CGG	Gly-Arg	Colorectal cancer, non-polyposis
697	aTGT-CGT	Cys-Arg	Colorectal cancer, non-polyposis
697	TGT-TTT	Cys-Phe	Colorectal cancer, non-polyposis
751	gGGA-AGA	Gly-Arg	Colorectal cancer, non-polyposis
834	tGCT-ACT	Ala-Thr	Colorectal cancer, non-polyposis
845	tAAA-GAA	Lys-Glu	Colorectal cancer, non-polyposis
886	GAG-GGG	Glu-Gly	Colorectal cancer, non-polyposis
905	ACA-AGA	Thr-Arg	Colorectal cancer, non-polyposis
923	GTA-GAA	Val-Glu	Colorectal cancer, non-polyposis
930	ATAa-ATG	Ile-Met	Colorectal cancer, non-polyposis ?

Listing of hMSH2 missense mutations in the Human Gene Mutation Database

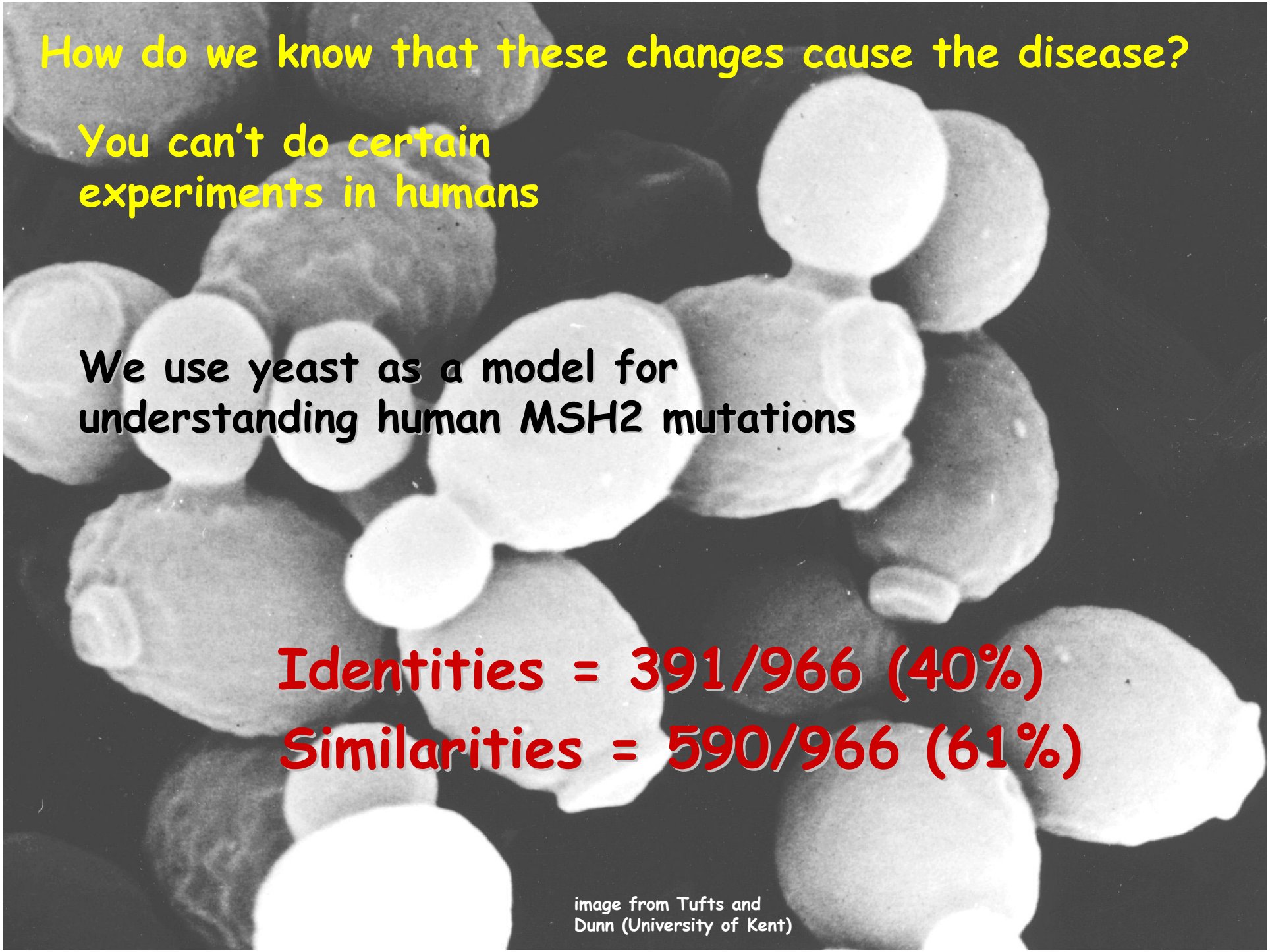
International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC) Database

Human Gene Mutation Database (in association with Celera)



Isidro *et al.* 1999

Yellow have identical amino acids in yeast Msh2



How do we know that these changes cause the disease?

You can't do certain experiments in humans

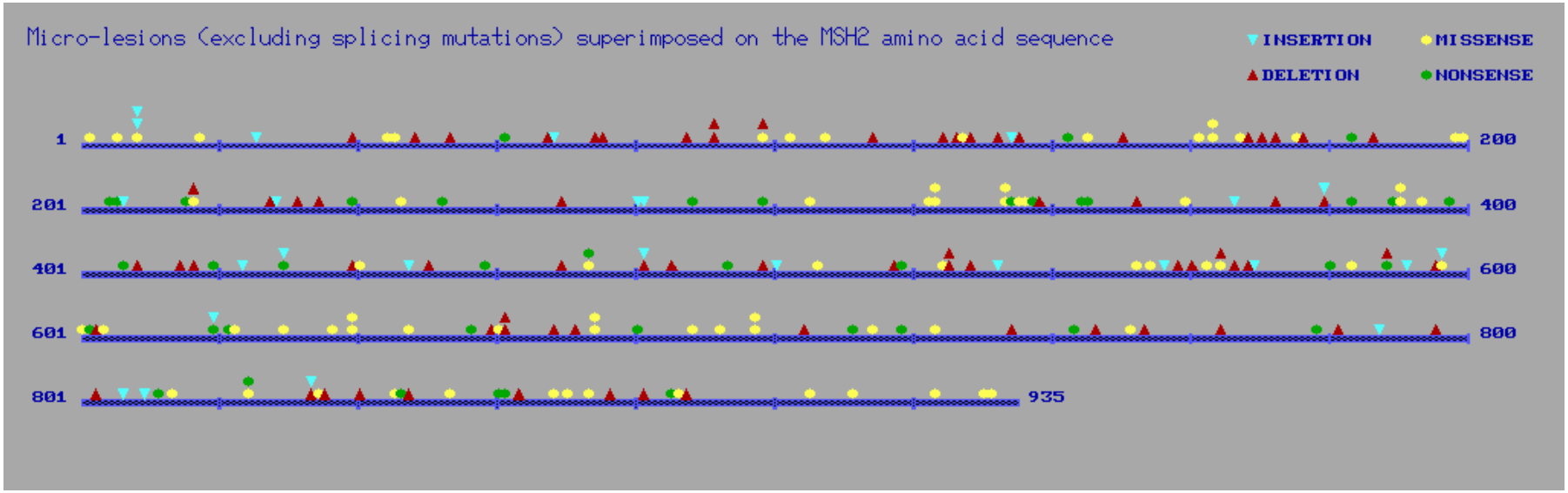
We use yeast as a model for understanding human MSH2 mutations

Identities = 391/966 (40%)

Similarities = 590/966 (61%)

image from Tufts and Dunn (University of Kent)

Project Overview



Manipulate the yeast *MSH2* gene to determine which human missense mutations are likely to be benign or pathogenic in nature.

Examine the defect at a molecular level to determine why the Msh2 variants are dysfunctional.

What effect might the single amino acid change have on the structure and function of the Msh2 protein?

Crystal Structure of MutS homodimer

Obmolova *et al.* (2000) *Nature* 407:703
Lamers *et al.* (2000) *Nature* 407:711

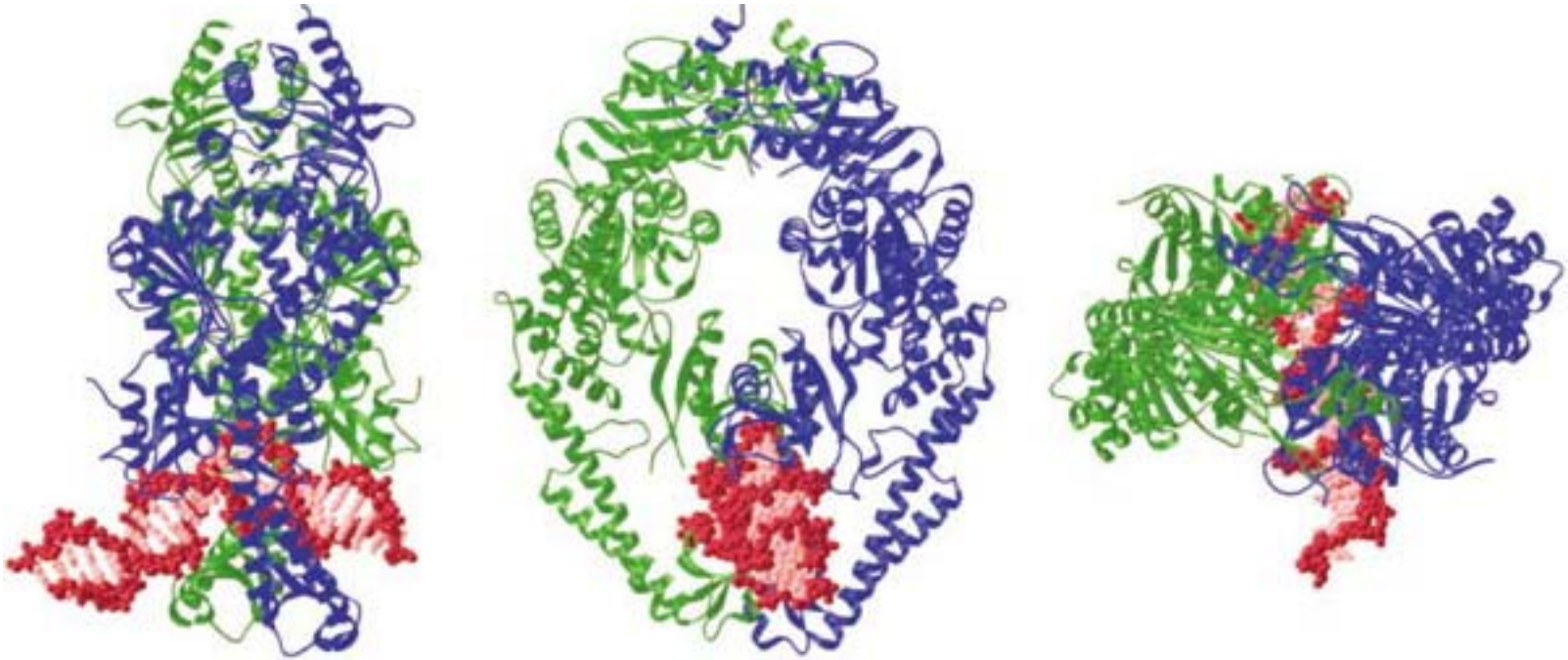


Image from Obmolova *et al.* (2000) Nature 407:703