Sept 2. Structure and Organization of Genomes

Today: Genetic and Physical Mapping

Assignments: Gibson & Muse, pp.4-10
Brown, pp. 126-160
Olson *et al*., Science 245: 1434

New homework:Due, before class, on Sept 9
(please submit answers on-line)

Sept 9. Forward and Reverse Genetics

Assignments: Gibson & Muse, pp.212-234
Brown, pp. 198-206
Hutchison *et al*., Science 286: 2165

Genetic and Physical Mapping

The ultimate goal of mapping is to identify the gene(s) responsible for a given phenotype or the mutation responsible for a specific variant.

The initial steps in mapping are to:

1. establish the proximity of genes or traits to one another
2. assign the genes to a particular chromosome

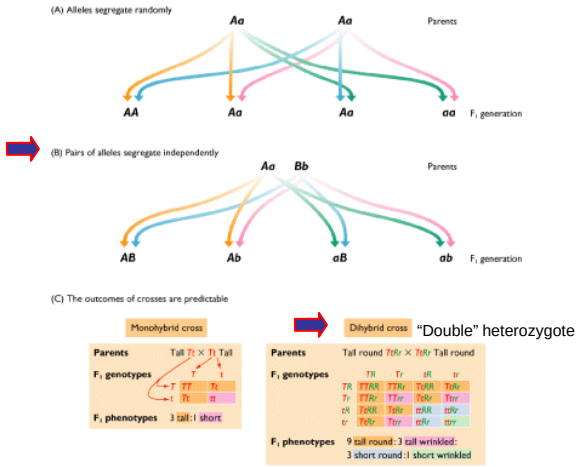
What is the difference between a genetic and a physical map?

Genetic maps depict relative positions of loci based on the degree of recombination. This approach studies the inheritance/assortment of traits by genetic analysis.

Physical maps show the actual (physical) distance between loci (in nucleotides). This approach applies techniques of molecular biology.

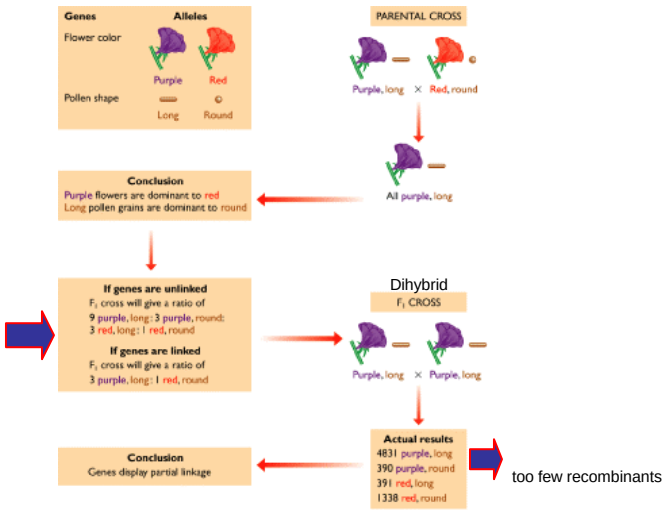
Rules for the first genetic (*i.e.*, linkage) map were uncovered by Mendel.

Mendel's Law of Independent Assortment

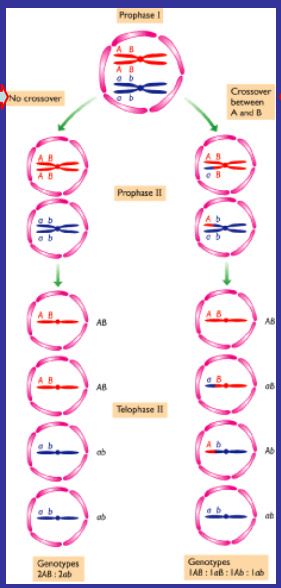


*If traits segregate independently of one another, they are "**unlinked**"*

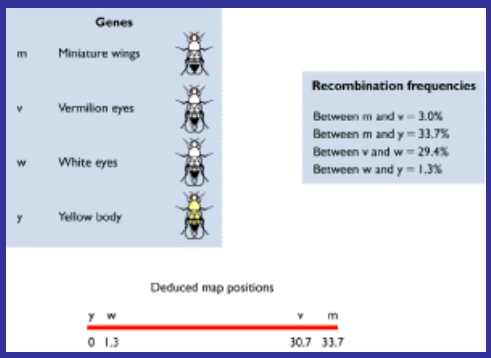
Enter Punnett, Bateson and Saunders (1905), who examined two other traits (flower color & pollen shape) in pea plants.



*If traits are co-inherited more often than expected by chance, they are **"linked"***



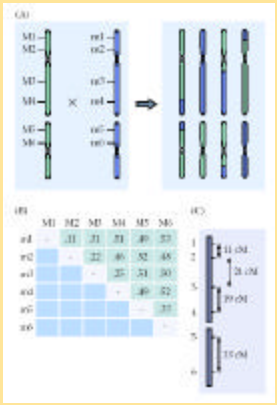
Morgan & Sturtevant (1913): Recombination frequency is a measure of the distance between two genes



See Box 5.2, Brown p.141; example of a multi-point cross

Genetic mapping via Linkage Analysis

By doing series of crosses, increasing the number of genetic markers (and examining large numbers of progeny to detect rare crossing-over events):



- Various linkage relationships appear (complete, partial, independent assortment) Between two loci, range from 0 - 0.5 (0 - 50 cM)
- The number of linkage groups emerge

What is the relationship between number of linkage groups and chromosomes?

Genetic mapping via Linkage Analysis

By doing series of crosses, increasing the number of genetic markers (and examining large numbers of progeny to detect rare crossing-over events):

- Various linkage relationships appear
  (complete, partial, independent assortment)
- The number of linkage groups can resolved

What is the relationship between number of linkage groups and chromosomes?

Need they always be the same?

How is it possible to assign traits (or a linkage groups) to a specific chromosome?

Genetic mapping via Linkage Analysis

By doing series of crosses, increasing the number of genetic markers (and examining large numbers of progeny to detect rare crossing-over events):

- Various linkage relationships appear
  (complete, partial, independent assortment)
- The number of linkage groups can resolved

What is the relationship between number of linkage groups and chromosomes?

Need they always be the same?

How is it possible to assign traits (or a linkage groups) to a specific chromosome?

Easy for sex-linked traits, but what about autosomes ?

What is the advantage of a test cross over a dihybrid cross for linkage mapping?

Genetic mapping relies on polymorphic markers/traits

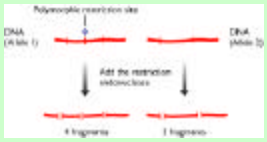
Am J Hum Genet. 1980 May;32(3):314-31. Related Articles, Links

Construction of a genetic linkage map in man using restriction fragment length polymorphisms.

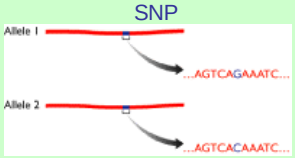
Botstein D, White RL, Skolnick M, Davis RW.

We describe a new basis for the construction of a genetic linkage map of the human genome. The basic principle of the mapping scheme is to develop, by recombinant DNA techniques, random single-copy DNA probes capable of detecting DNA sequence polymorphisms, when hybridized to restriction digests of an individual's DNA. Each of these probes will define a locus. Loci can be expanded or contracted to include more or less polymorphism by further application of recombinant DNA technology. Suitably polymorphic loci can be tested for linkage relationships in human pedigrees by established methods; and loci can be arranged into linkage groups to form a true genetic map of "DNA marker loci." Pedigrees in which inherited traits are known to be segregating can then be analyzed, making possible the mapping of the gene(s) responsible for the trait with respect to the DNA marker loci, without requiring direct access to a specified gene's DNA. For inherited diseases mapped in this way, linked DNA marker loci can be used predictively for genetic counseling.

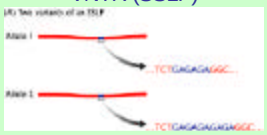
RFLP



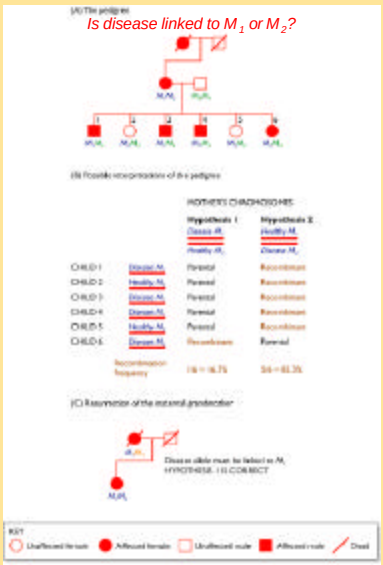
SNP



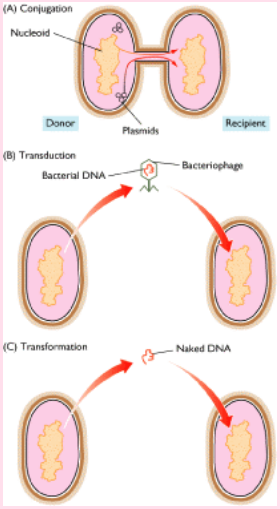
VNTR (SSLP)

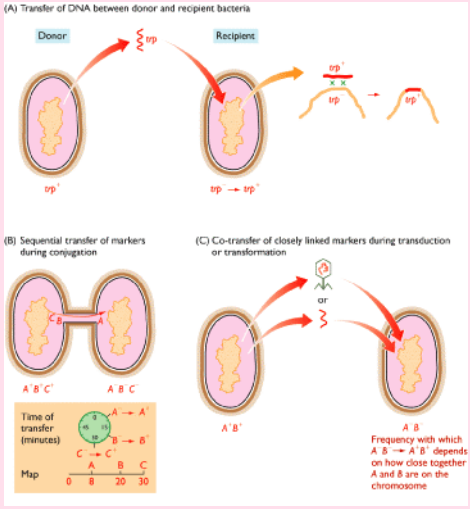


Linkage Mapping in Humans:
Association of disease state with a minisatellite (M_N) polymorphism



Genetic Mapping in Bacteria





What are the limitations to constructing a genetic map?

- Access to polymorphic traits or markers
- Need for a large number of progeny and/or multiple generations
- Best performed in model organisms subject to selective breeding
- Crossing-over does not occur at random (maps of limited accuracy)

In contrast, some form of physical map can be constructed for any organism.

Physical Mapping of Genomes

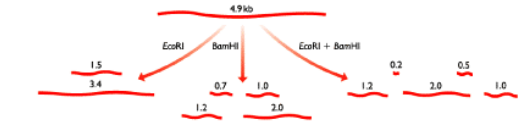
Physical maps plot the actual location of DNA sequences in the genome

- **Restriction maps**: locate the positions of and distances between endonuclease recognition sites on a DNA molecule
- **Long-range restriction maps**: locate the positions of rare-cutting endonuclease recognition sites on a DNA molecule by PFGE
- **Clone (contig) maps**: consist of libraries of overlapping clones where the relationship of each clone to other clones has been resolved
- **Fluorescent *in situ* hybridization** (**FISH**): locates the position of a marker by hybridizing a labeled probe to intact chromosomes
- **Optical maps**: visually inspects and measures the positions of endonuclease recognition sites on a DNA molecule
- **EST** (expressed sequence tags) **maps**: plot the location of transcribed sequences
- **STS content maps**:(tbd)

Restriction mapping:

finding the distance between restriction enzyme recognition sites in a DNA fragment

1. Combination of single & double digests



INTERPRETATION OF THE DOUBLE RESTRICTION

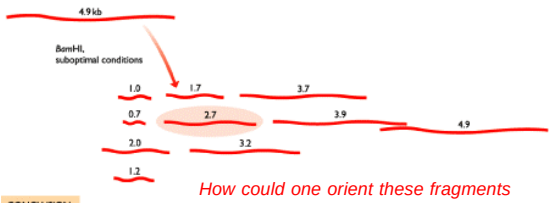
Fragments	Conclusions
0.2 kb, 0.5 kb	These must derive from the 0.7 kb *Bam*HI fragment, which therefore has an internal *Eco*RI site:
1.0 kb	This must be a *Bam*HI fragment with no internal *Eco*RI site. We can account for the 1.5 kb *Eco*RI fragment if we place the 1.0 kb fragment thus:
1.2 kb, 2.0 kb	These must also be *Bam*HI fragments with no internal *Eco*RI sites. They must lie within the 3.4 kb *Eco*RI fragment. There are two possibilities:

MAP I MAP II

PREDICTED RESULTS OF A PARTIAL *Bam*HI RESTRICTION

If Map I is correct, then the partial restriction products will include a fragment of 1.2 + 0.7 = 1.9 kb
If Map II is correct, then the partial restriction products will include a fragment of 2.0 + 0.7 = 2.7 kb

2. Partial digestion such that DNA is not cut to completion



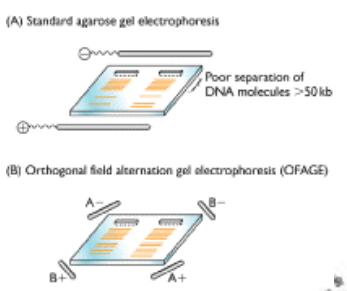
CONCLUSION
Map II is correct

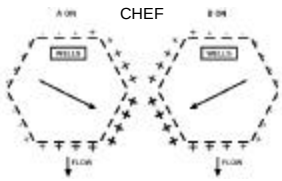
How could one orient these fragments without relying on complete digests?

Long-range restriction mapping

Two innovations:

1. Pulsed-field gel electrophoresis (for separating large DNA fragments)





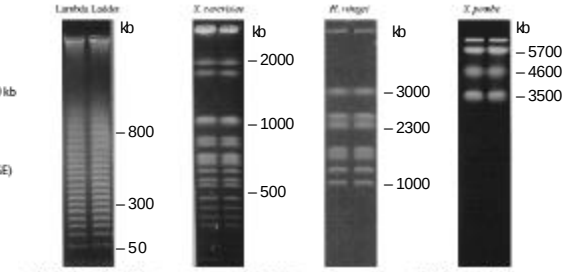


Fig. 4.2. **A. Lambda ladder** (catalog number 170-3635) was separated on a 1.0% Molecular Biology Certified Agarose (catalog number 162-0133) gel in 0.5x TBE, recirculated at 14 °C. The run time was 22 hours at 6 V/cm with a 60 to 90 second switch time ramp.

B. *Saccharomyces cerevisiae* Strain YNN295. (Catalog number 170-3605). Chromosomes were separated on a 1.0% Pulsed Field Certified Agarose (catalog number 162-0137) gel in 0.5x TBE, recirculated at 14 °C. The run time was 24 hours at 6 V/cm with a 60 to 120 second switch time ramp.

C. *Hansenula wingei* Strain YB-4002-VIA. (Catalog number 170-3667). Chromosomes were separated on a 0.8% Molecular Biology Certified Agarose gel in 1.0x TAE, recirculated at 14 °C. The run time was 50 hours at 3 V/cm with a 250 to 900 second switch time ramp.

D. *Schizosaccharomyces pombe* Strain 972 h-. (Catalog number 170-3633). Chromosomes were separated on a 0.6% Chromosomal Grade Agarose (catalog number 162-0135) gel in 1.0x TAE, recirculated at 14 °C. The run time was 72 hours at 2 V/cm with a 20 to 30 minute switch time ramp.

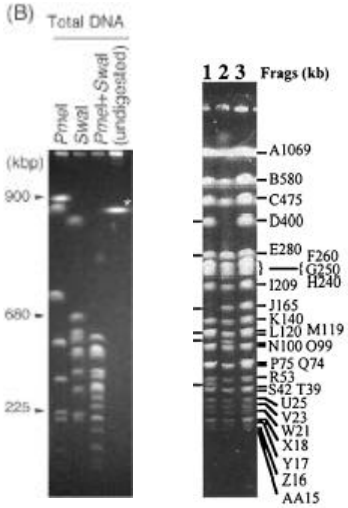
"Molecular Karyotype"

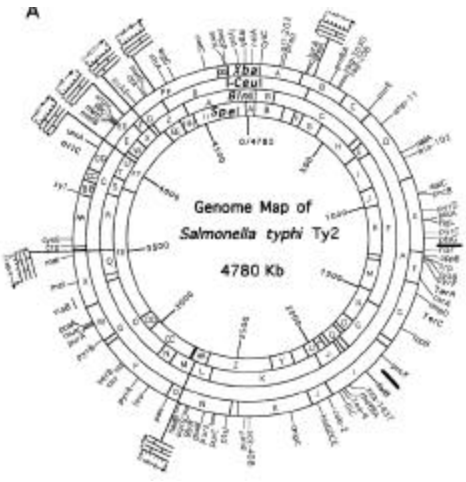
Long-range restriction mapping

2. "Rare-cutting" restriction enzymes

8-bp recognition sites: *e.g.*, *Not*I (GC^GGCCGC); *Swa*I (ATTT^AAAT); *Pme*I (GTTT^AAAC)

Homing (intron- or intein-encoded) endonucleases: *e.g.*, I-*Ceu*I (TAACTATAACGGTC^CTAA)





Optical Mapping of Chromosomes I

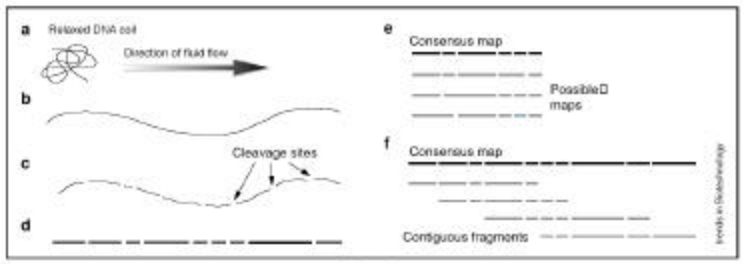
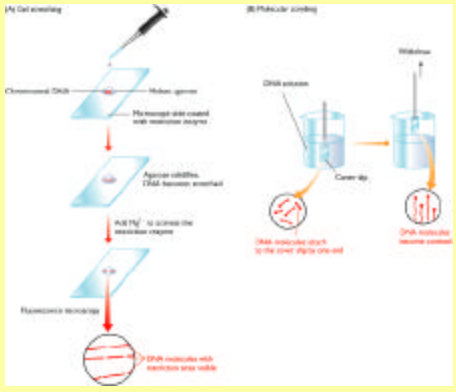
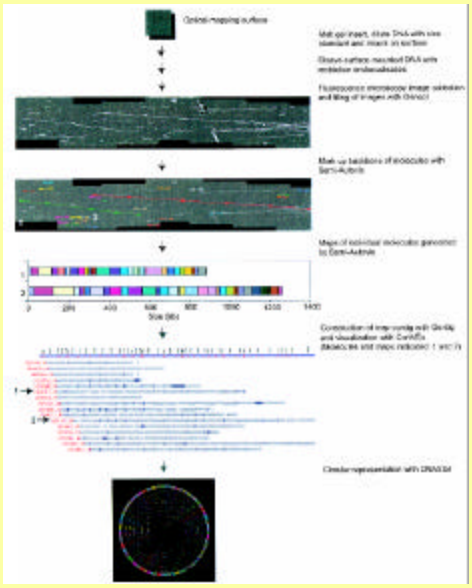


Figure 1

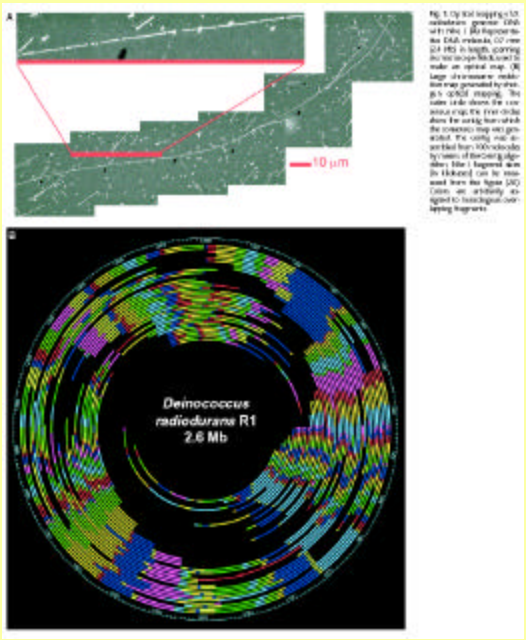
Steps in optical mapping. Large DNA molecules, comprising large-insert clones[1,12] (bacterial artificial chromosomes, yeast artificial chromosomes) or genomic DNA[13] (J. Lin et al., unpublished), are mounted on silanized glass surfaces (**a**). Fluid flows elongate the DNA molecules (**b**) and charge interactions hold the stretched DNA molecules on the surface (shown here looking down onto the surface). The DNA molecules are digested with a restriction enzyme; cleavage sites are visible as small gaps between DNA fragments, which retain their original order (**c**). The DNA is stained with an intercalating dye and visualized by fluorescence microscopy[14,21]. Digital images are recorded with a charge-coupled-device camera. The fragments are sized by measuring the integrated fluorescence intensity and maps of single molecules are constructed[8,10] (**d**). Fragments as small as 800 bp can be imaged and sized. Consensus maps of cloned DNA are created using a probabilistic scheme[22] (**e**) and maps of genomic DNA are aligned into contigs using the Gentig software (**f**)[23] (J. Lin et al., unpublished). Consensus maps are subsequently used as scaffolds for sequence assembly and verification.

Optical Mapping of Chromosomes II

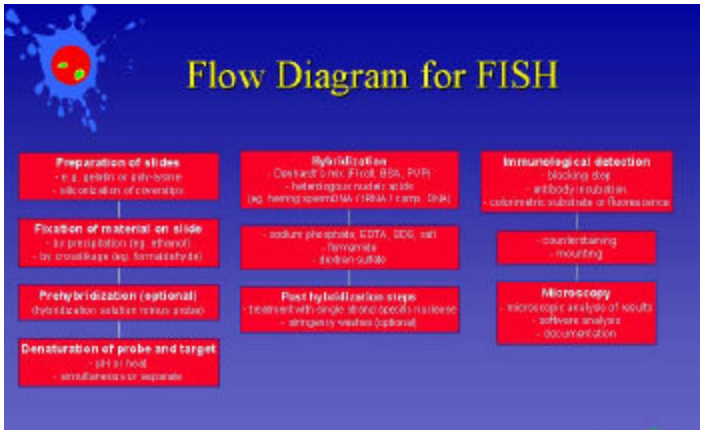


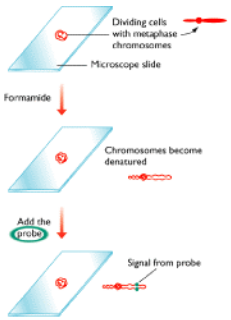


Optical Mapping of Chromosomes III



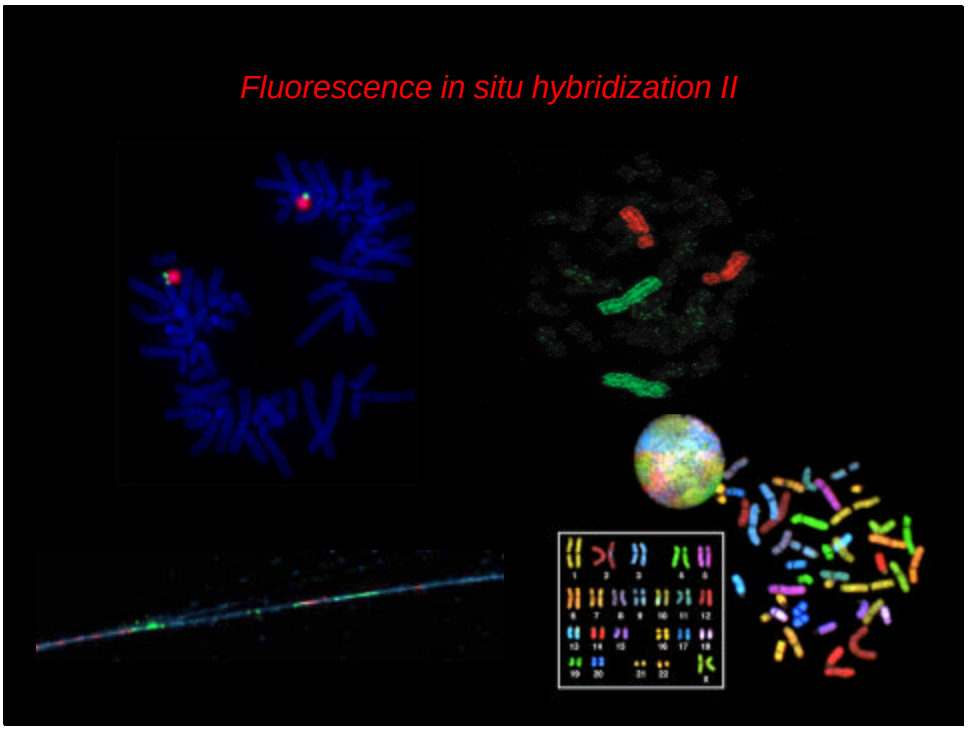
Fluorescence in situ hybridization I





First *in situ* detection - Bauman *et al*. 1980, *Exp. Cell Res*. 128: 485
Two-color detection: Hopman *et al*. 1986, *Histochemistry* 85: 1
Three-color detection: Nederlof *et al*. 1989, *Cytometry* 10: 20
Combinatorial color-coding (mFISH): Nederlof *et al*. 1990, *Cytometry* 11: 126

Fluorescence in situ hybridization II



Somatic cell hybrids and radiation hybrid mapping

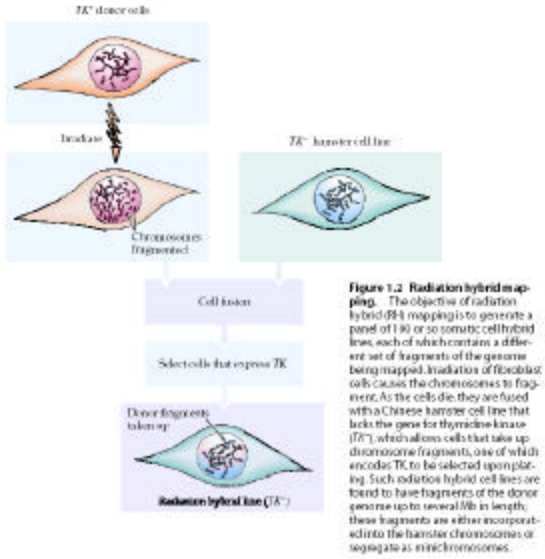
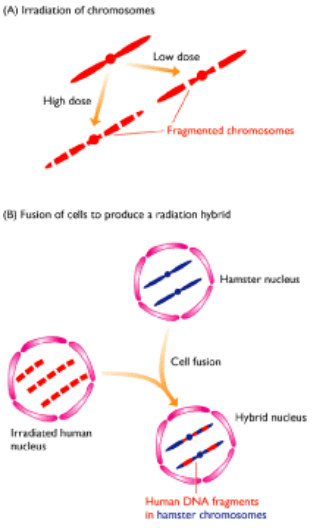


Figure 1.2 Radiation hybrid mapping. The objective of radiation hybrid (RH) mapping is to generate a panel of 100 or so somatic cell hybrid lines, each of which contains a different set of fragments of the genome being mapped. Irradiation of fibroblast cells causes the chromosomes to fragment. As the cells die, they are fused with a Chinese hamster cell line that lacks the gene for thymidine kinase (TK⁻), which allows cells that take up chromosome fragments, one of which encodes TK, to be selected upon plating. Such radiation hybrid cell lines are found to have fragments of the donor genome up to several Mb in length; these fragments are either incorporated into the hamster chromosomes or segregate as mini-chromosomes.



A radiation hybrid map of the human genome

Gabor Gyapay[1,*], Karin Schmitt[2], Cécile Fizames[1], Hywel Jones[2], Nathalie Vega-Czarny[1], Dominique Spillett[2], Delphine Muselet[2], Jean-François Prud'Homme[1], Colette Dib[1], Charles Auffray[3], Jean Morissette[1,4], Jean Weissenbach[1] and Peter N. Goodfellow[2]

We have developed a panel of whole-genome radiation hybrids by fusing irradiated diploid human fibroblasts with recipient hamster cells. This panel of 168 cell lines has been typed with microsatellite markers of known genetic location. Of 711 AFM genetic markers 484 were selected to construct a robust framework map that spans all the autosomes and the X chromosome. To demonstrate the utility of the panel, 374 expressed sequence tags (ESTs) previously assigned to chromosomes 1, 2, 14 and 16 were localized on this map. All of these ESTs could be positioned by pairwise linkage to one of the framework markers with a LOD score of greater than 8. The whole genome radiation hybrid panel described here has been used as the starting material for the Genebridge4 panel that is being made widely available for genome mapping projects.

INTRODUCTION

Radiation hybrids, produced by fusing irradiated donor cells with recipient rodent cells, can be used for constructing genetic maps that are complementary to both recombination maps and physical maps based on contigs. Recombination maps are limited to polymorphic markers and are constrained by the recombination rate; in contrast, radiation hybrid methods exploit differences between species and can be used to map both polymorphic markers and non-polymorphic markers such as sequence tagged sites (STS) and expressed sequence tags (EST). The resolution of radiation hybrid mapping is a function of both fragment size and retention frequencies. The fragment size can be varied by altering the radiation dose and it is possible to construct panels designed either for map continuity with few markers or for high resolution with large numbers of markers. Information on localization using

SCIENCE VOL 291 16 FEBRUARY 2001

A High-Resolution Radiation Hybrid Map of the Human Genome Draft Sequence

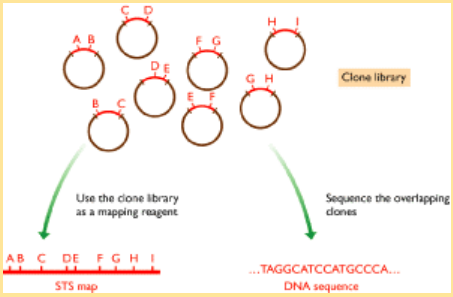
Michael Olivier,[1] Amita Aggarwal,[1] Jennifer Allen,[1] Annalisa A. Almendras,[1] Eva S. Bajorek,[1] Ellen M. Beasley,[1*] Shannon D. Brady,[1] Jannette M. Bushard,[1] Valerie I. Bustos,[1] Angela Chu,[1] Tisha R. Chung,[1] Annick De Witte,[1] Mirian E. Denys,[1] Rakly Dominguez,[1] Nicole Y. Fang,[1] Brian D. Foster,[1] Robert W. Freudenberg,[1] David Hadley,[1] Libby R. Hamilton,[1] Tonya J. Jeffrey,[1] Libusha Kelly,[1] Laura Lazzeroni,[1] Michelle R. Levy,[1] Saskia C. Lewis,[1] Xia Liu,[1] Frederick J. Lopez,[1] Brent Louie,[1] Joseph P. Marquis,[1] Robert A. Martinez,[1] Margaret K. Matsuura,[1] Nedda S. Misherghi,[1] Jolanna A. Norton,[1] Adam Olshen,[1] Shanti M. Perkins,[1] Amy J. Perou,[1] Chris Piercy,[1] Mark Piercy,[1] Fawa Qin,[1] Tim Reif,[1] Kelly Sheppard,[1] Vida Shokoohi,[1] Geoff A. Smick,[1] Wei-Lin Sun,[1] Elizabeth A. Stewart,[1†] J. Fernando Tejeda,[1] Nguyet M. Tran,[1] Tonatiuh Trejo,[1] Nu T. Vo,[1] Simon C. M. Yan,[1] Deborah L. Zierten,[1] Shaying Zhao,[2] Ravi Sachidanandam,[2] Barbara J. Trask,[4] Richard M. Myers,[1] David R. Cox[1‡]

We have constructed a physical map of the human genome by using a panel of 90 whole-genome radiation hybrids (the TNG panel) in conjunction with 40,322 sequence-tagged sites (STSs) derived from random genomic sequences as well as expressed sequences. Of 36,678 STSs on the TNG radiation hybrid map, only 3604 (9.8%) were absent from the unassembled draft sequence of the human genome. Of 20,030 STSs ordered on the TNG map as well as the assembled human genome draft sequence and the Celera assembled human genome sequence, 36% of the STSs had a discrepant order between the working draft sequence and the Celera sequence. The TNG map order was identical to one of the two sequence orders in 80% of these discrepant cases.

Olson, Hood, Cantor, Botstein (1989) Science 245: 1434
A common language for physical mapping the human genome

1. *What is an STS?*
2. *How many people read the assigned (2-page) article?*
3. *Why does an STS need to be a unique sequence?*
4. *How will STS technology "solve the problem of merging data from many sources"? (And what kind of data are they taking about?)*
5. *How does one find an STS in the genome?*
6. *Technically, how are STSs recovered and assayed?*
7. *What are some of the problems in developing contig maps?*
8. *How will STSs assist in the assembly of contig maps?*
9. *What are some of the disadvantages of restriction maps?*
10. *What are the advantages of using STSs as genomic landmarks?*
11. *How many STSs are needed to be useful?*

Assembling clone contigs by STS content mapping



STS content mapping assays for the presence of known sequences (STS) in DNA fragments from any source (clones, RH, etc.) and can thereby align physical maps derived by different methods

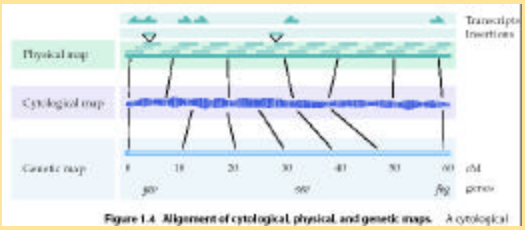


Figure 1.4 Alignment of cytological, physical, and genetic maps. A cytological

Another method for mapping a gene/clone is to obtain pure preparations of a particular chromosome by flow cytometry or fluorescence-activated chromosome sorting.