Psych 3102
Introduction to Behavior Genetics

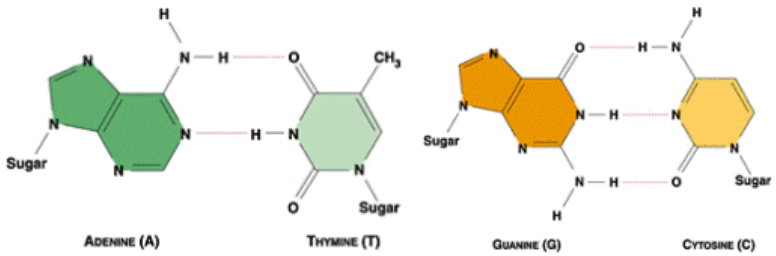
Lecture 6

Nature of the genetic material

Review: Structure and function of DNA

- Watson & Crick, 1953
- nucleic acid

  - chemical group to which RNA and DNA belong
- nucleotide
  - building block of nucleic acids
  - 3 subunits: pentose sugar
    phosphate group
    nitrogen-containing base



purines — adenine (A), guanine (G)

pyrimidines — thymine/uracil (T/U), cytosine (C)

complimentary base pairing

- double helix

Requirements for a hereditary material

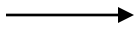
1. ability to carry information and control protein synthesis
2. ability to replicate accurately
3. capable of variation

1.How information controlling protein synthesis is carried

genetic code

- universal
- triplets of nucleotides code for single amino acids

Why a triplet?

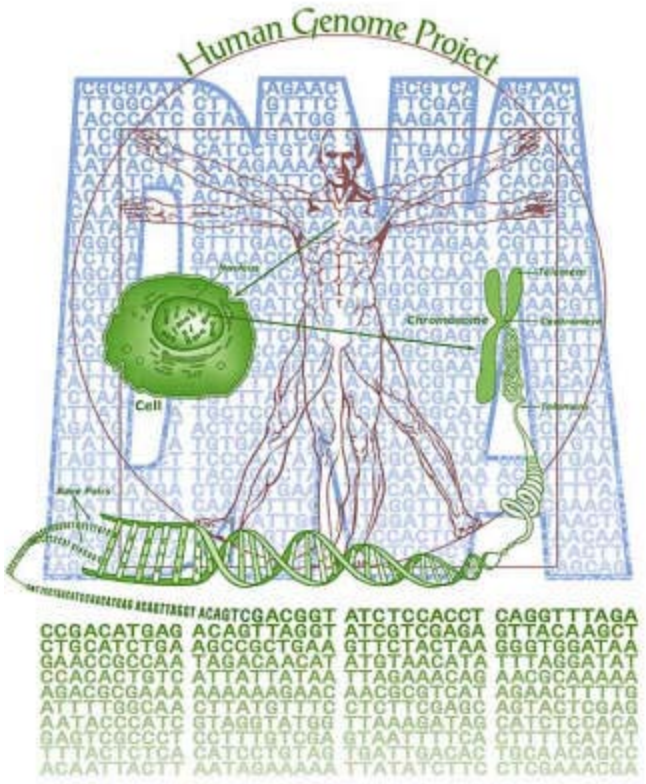


Human genome

- 3 billion base pairs
  (3000 books, 500 pages each)
- completely sequenced
  1 error/100,000 bp

- estimated 22,000 genes
  all protein kinases
  all transcription factors

- ~500 species sequenced



human/human genomes 99.9% identical
human/chimp genomes 98.7% identical
human/daffodil genomes 35% identical

haplotype map
haplotypes small DNA regions, each inherited intact (vary across human populations)

proteome all proteins able to be synthesized by a genome
ENCODE ENCyclopedia Of DNA Elements project

- less than 2% of genome is protein-coding (exon)

  but produces ½-1.5m proteins through alternative splicing

- 25% is intron, 25% recognized regulatory 48% ???

non-protein-coding RNA genes (rRNA, tRNAs, snRNAs, microRNAs involved in gene regulation)

structural motifs – stabilize DNA

relics of sequences used in past (pseudogenes), no longer produce functional proteins but may have regulatory roles (eg. may code for siRNAs)

however

- all this is based on the 7 human genomes published so far:
  1. reference genome (consensus from several individuals)
  2. Celera genome
  3. Craig Venter genome
  4. James Watson genome — all Caucasian, 3m SNPs
5. Asian (Han chinese) genome — 3m SNPs, ½ m novel
6. African (Yoruba and Nigerian) — 4m SNPs, 1m novel
7. acute myelogenous leukemia patient — normal and cancer cells (10 SNPs different)

Within and cross species differences/similarities based on surveys of SNPs and some structural variation (ie. essentially on a few million SNPs out of 3 billion)

Initial cost/genome = $100s of millions

2008 cost/genome = $10,000

Human genome and inherited disease

- 3000 (out of 20,000) human genes known to have at least 1 mutation that causes an inherited disease
- Information kept on NCBI (National Center for Biotechnology Information)
- 1/3 to ½ of all genes are expressed in the brain - more than any other organ

  reflected in large number of neurogenetic disorders

  >30% of Mendelian diseases have neurological manifestations

accurate diagnosis & counseling possible for single-gene causes with known genome location

Most genetic disorders, however, show any or several of the following:

genetic heterogeneity, variable expression, incomplete penetrance, anticipation, phenocopies, imprinting

even mitochondrial inheritance

- all complicate relating phenotype to genotype

Protein synthesis
- how the information coded into DNA is used

1. transcription

   DNA code is transcribed to form mRNA molecule

   RNA polymerase

2. RNA processing

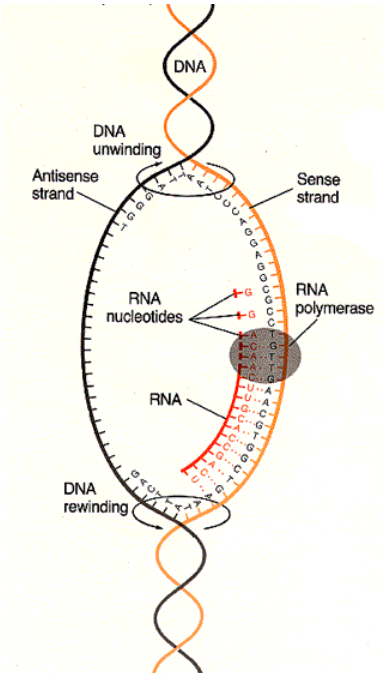
   introns spliced out leaving exons

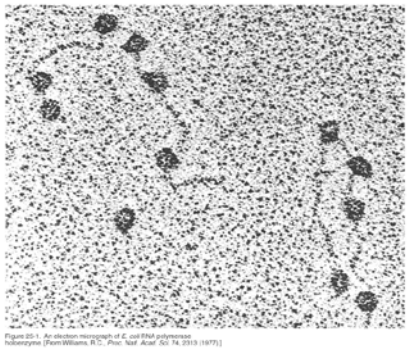
alternative splicing (+1/2 of all genes)

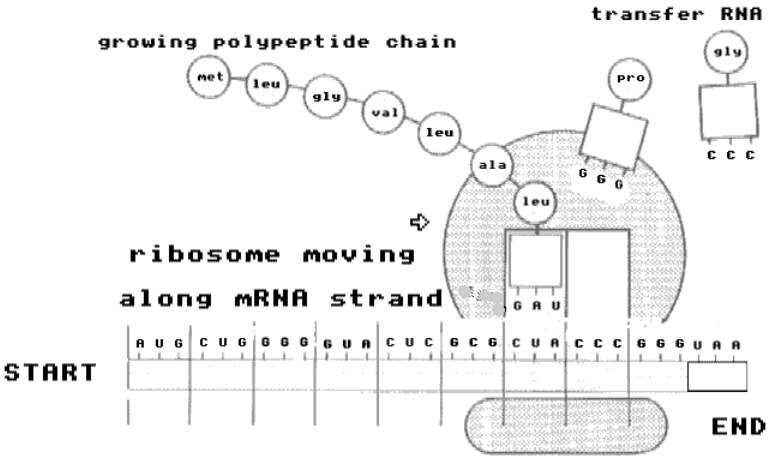
3. translation

   mRNA code is translated into sequence of amino acids to form polypeptide

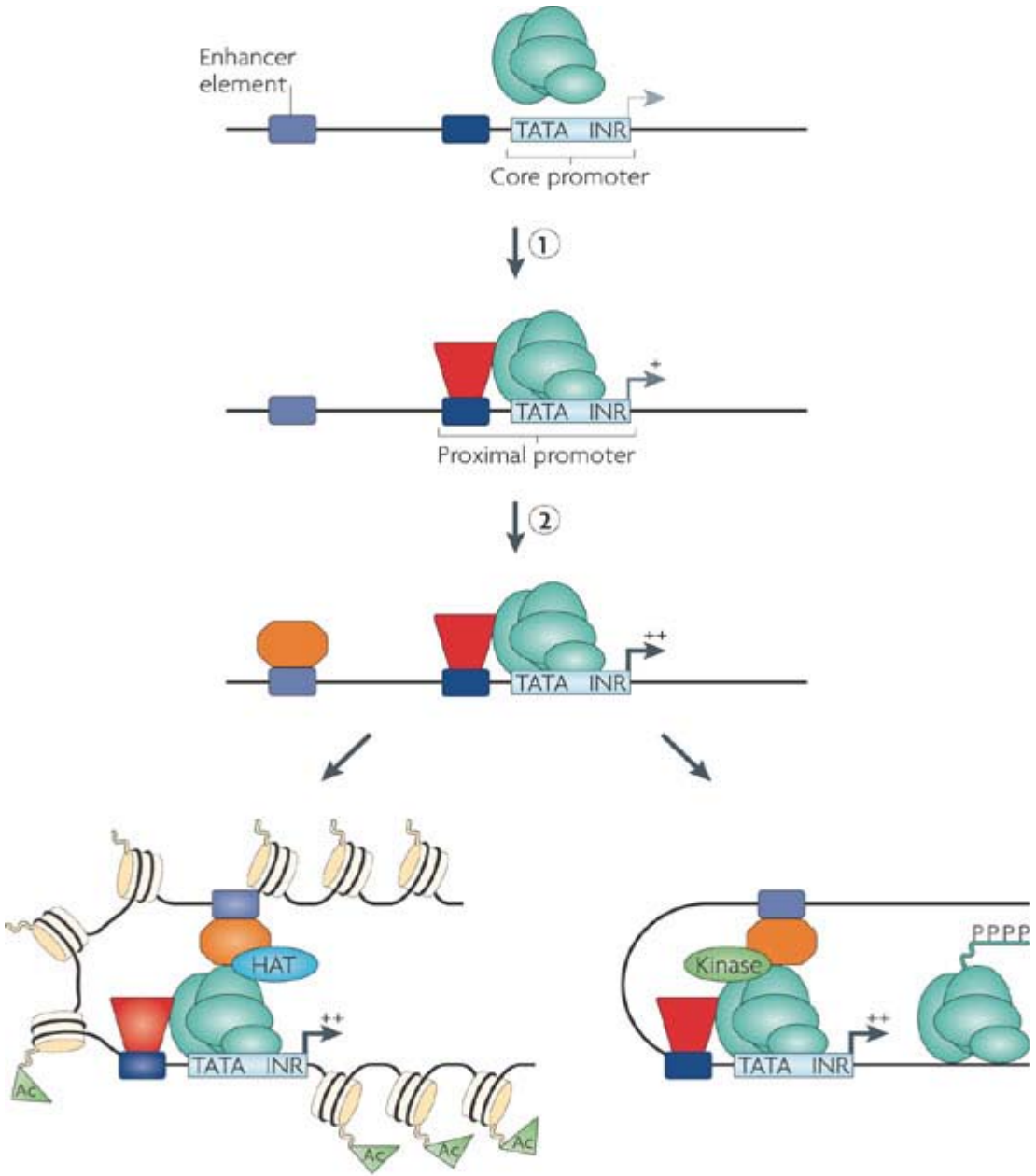
microarrays – used to study expression of many genes at once (transcriptome)







General transcription factors (green ovals) bind to core promoter regions through recognition of common elements such as TATA boxes and initiators (INR). However, these elements on their own provide very low levels of transcriptional activity owing to unstable interactions of the general factors with the promoter region. Promoter activity can be increased (represented by +) by site-specific DNA-binding factors (red trapezoid) interacting with *cis* elements (dark blue box) in the proximal promoter region and stabilizing the recruitment of the transcriptional machinery through direct interaction of the site-specific factor and the general factors (step 1). Promoter activity can be further stimulated to higher levels by site-specific factors (orange octagon) binding to enhancers (step 2). The enhancer factors can stimulate transcription by (bottom left) recruiting a histone-modifying enzyme (for example, a histone acetyltransferase (HAT)) to create a more favourable chromatin environment for transcription (for example, by histone acetylation (Ac)) or by (bottom right) recruiting a kinase that can phosphorylate (P) the carboxy-terminal domain of RNA polymerase II and stimulate elongation



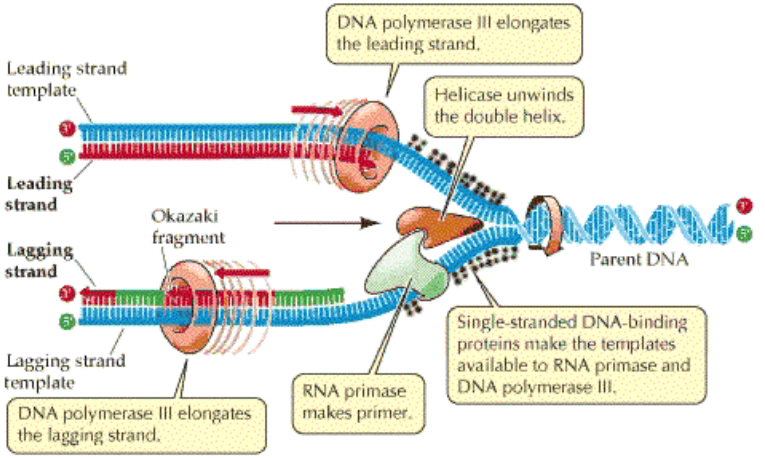
Nature Reviews Genetics **10**, 605-616 (September 2009) | doi:10.1038/nrg2636**Insights from genomic profiling of transcription factors**Peggy J. Farnham[1]

Serotonin-receptor (1A subtype) - amino acid sequence

DNA replication
- how DNA copies are produced

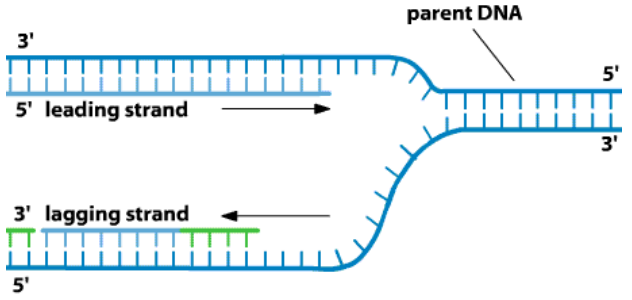
- occurs during S-phase of interphase

1. DNA double helix is unwound
2. strands are separated
3. DNA polymerase creates new strand on each template (original) strand



semi-conservative replication

http://www.youtube.com/watch?v=4PKjF7OumYo&eurl=http://io9.com/5142583/the-most-awesome-science-video-about-dna-ever-made



DNA mutations
- how DNA varies to make evolution possible

- copying errors
- somatic mutations - not passed on to offspring
- germ-line mutations – passed on to offspring
- the only way new alleles are formed
- almost always deleterious

point mutations

chromosome mutations

Point mutations

1. Substitutions

synonymous mutations (neutral, silent)

Tp53 tumor suppressor gene, codes for transcription factor that controls many genes in cell cycle

mutated in almost all cancer cells – point mutations produce change in function

but >200 point mutations occur naturally that produce NO change in function/increase in cancer risk

missense mutation

- sickle cell cystic fibrosis

PKU

codon 1 AUG → GUG start → val no product

codon 408 CGG → UGG arg → trp low activity

nonsense mutation

cystic fibrosis

10% of patients have STOP codon instead of amino acid codon in middle of gene

2. Insertions and deletions

frameshift mutation

wildtype sequence: the big boy saw the new cat eat the hot dog

point deletion: the big oys awt hen ewc ate att hen otd og_

point insertion: the big boy saw tth ene wca tea tth eho tdo g

triplet repeat mutation

- delete one codon:
  the big boy the new cat eat the hot dog 1deleted amino acid
  delete across codons:
  the big baw the new cat eat the hot dog 1 amino acid sub for 2

- triplet addition leads to additional amino acids of the same type being added

Huntington mutation CAG repeat polyglutamine

normal= 6-35 repeats

mutation=36-150 repeats

How are these mutations (polymorphisms) detected?

Fragment-length polymorphisms (and microsatellites) :

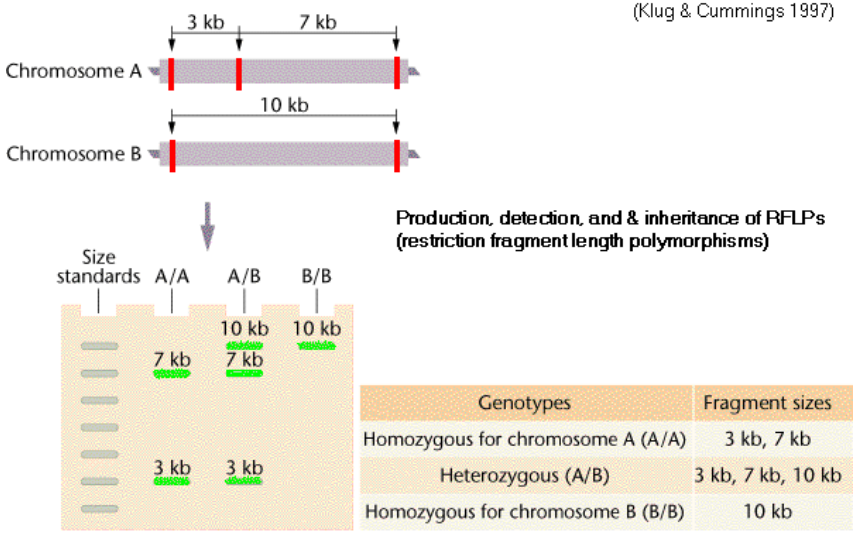
restriction enzymes

- cut DNA at specific points in the sequence

a point mutation may change the restriction point sequence

– DNA will not be cut

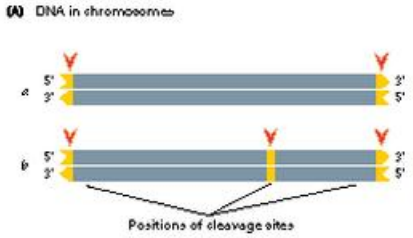
- DNA fragments of different sizes will be detected

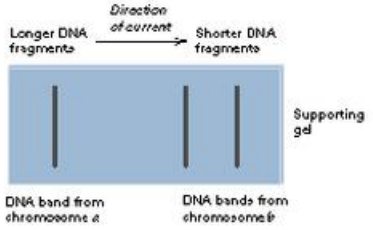


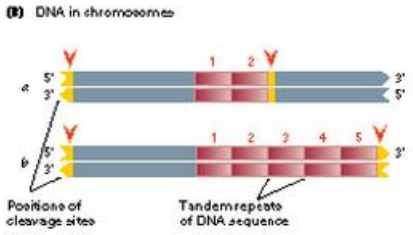
(Klug & Cummings 1997)

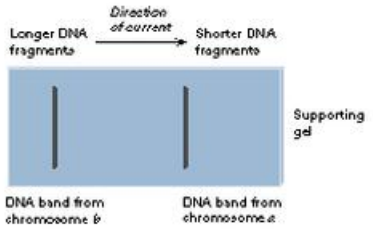
Production, detection, and & inheritance of RFLPs (restriction fragment length polymorphisms)

Genotypes	Fragment sizes
Homozygous for chromosome A (A/A)	3 kb, 7 kb
Heterozygous (A/B)	3 kb, 7 kb, 10 kb
Homozygous for chromosome B (B/B)	10 kb









How are polymorphisms detected? continued

polymerase chain reaction

- amplifies DNA sequence to be studied
- http://www.maxanim.com/genetics/PCR/pcr.swf

electrophoresis

- separates DNA fragments for genotyping or identification of markers present

To detect SNPs:

-separate DNA strands, allow to hybridize to single-stranded probe for one or the other allele, fluorescence indicates which probe has been bound and therefore which allele is present

genetic (DNA) marker

- any sequence of known location that varies from person to person, used to identify regions of DNA associated with variation for a trait

Types of polymorphisms

1. RFLPs - restriction fragment length polymorphisms

2. tandem repeat polymorphisms (microsatellites) - differences in number of copies of a repeated DNA sequence, abundant, highly polymorphic

   simple sequence repeats (SSRs):

   $^{5'}$ACACACACACAC……$^{3'}$ dinucleotide repeat

   CAGCAGCAGCAGCAGCAG… trinucleotide repeat

   variable number tandem repeats (VNTRs): - repeated unit is +10 nucleotides, easily detected, used in DNA 'fingerprinting'

3. SNPS - single nucleotide polymorphisms, only 2 alleles possible

4. copy number variants – duplication of stretches of DNA, microdeletions

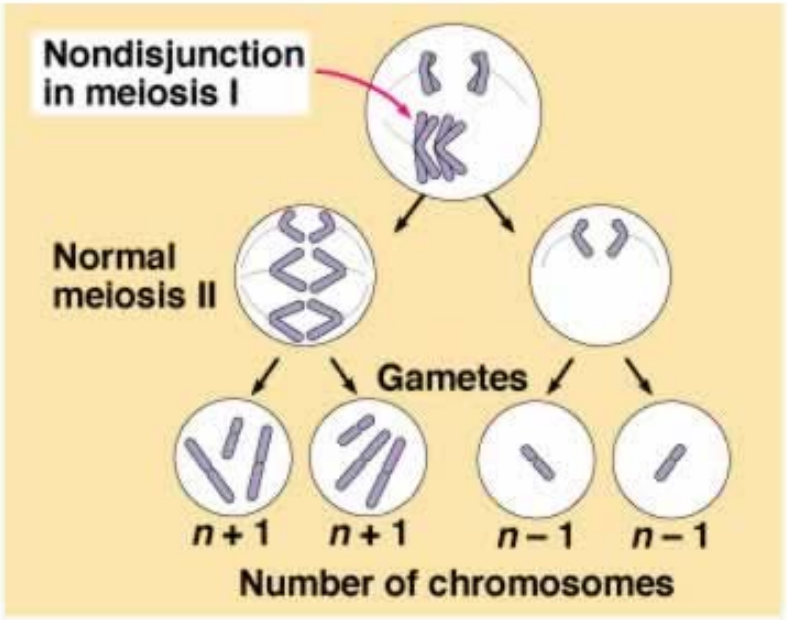
Chromosome mutations

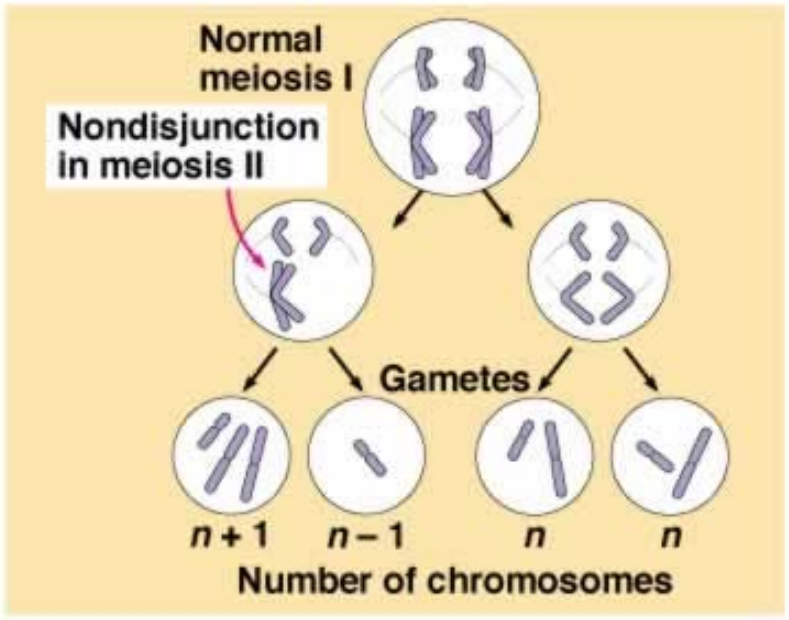
- more than one gene affected, effects on phenotype more severe

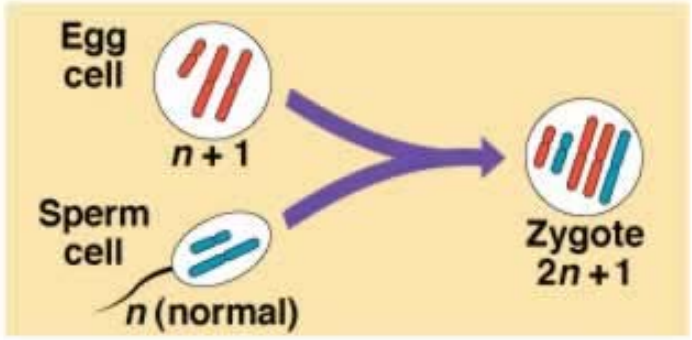
Changes in chromosome number = **aneuploidy**

non-disjunction

- process that causes aneuploidy
- failure of homologous chromosomes (or chromatids) to separate during cell division
- unpaired autosomes at meiosis are inactivated no survival of autosomal monosomies







Human chromosome aneuploidies

- no autosomal monosomies
- 3 autosomal trisomies

  all involve small chromsomes with relatively few genes

chr 21	374 genes	Down syndrome
chr 13	332 genes	Patau syndrome
chr 18	243 genes	Edward syndrome

sex chromosome aneuploidies more common

Autosomal trisomies

trisomy 21 Down syndrome

1 in 1000 (average) live births

¼ of all retarded individuals

incidence increases with age of Mo

trisomy 13 Patau syndrome

1 in 5000 live births

fatal, live ~ 3 months

trisomy 18 Edward syndrome

1 in 10,000 live births (95% die in utero)

av lifespan = 5-15 days,only 5-10% live 1 year

Sex chromosome aneuploidies

- more common, trisomies all around 1 in 1000
- less deleterious since extra X chromosomes are inactivated, Y has few genes

XXY Klinefelter male

1 in 500-1000 live births, almost 2/3 undiagnosed

incidence rising, only aneuploidy known to be 50%

paternal meiosis I

non-disjunction

some feminine features

leading cause of male

sterility

XXX Triple X female normal female

XYY normal male

Only viable human monosomy:

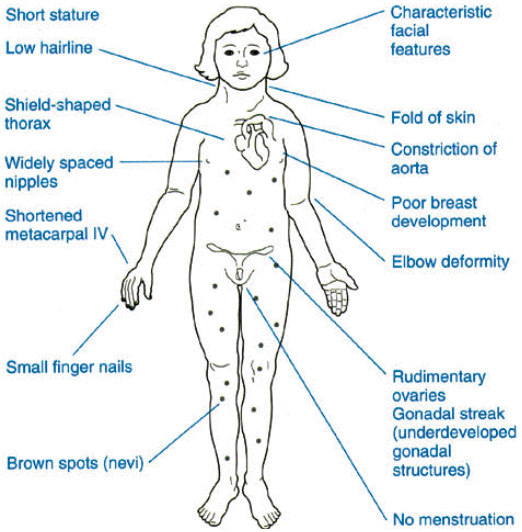
XO Turner female

1 in 3000 live births

sterile

no secondary

sex characteristics



untreated

treated

Changes in chromosome structure

- caused by breakage without correct rejoin during crossing-over, unequal crossing-over

deletion fragment of chromosome lost

duplication fragment rejoins same chromosome

inversion fragment rejoins upside down

translocation fragment joins non-homologous chromosome, may be reciprocal

cri-du-chat syndrome deletion on chromosome 5

chronic myelogenous leukemia (CML) reciprocal translocation chr 22, 9