

BIOL3120 –Human Genetics and Evolutionary Medicine

Faculty of Science and Engineering



Lecturer profile

DR OLIVER GRIFFITH

Dr Oliver Griffith

Evolutionary Genomics Lab

Department of Biological Sciences
Faculty of Sciences and Engineering

Email: oliver.Griffith@mq.edu.au



My path to Macquarie

Bachelor of Engineering UNSW
(dropped out after 2 years)

Bachelor of Science at Usyd
- Biology Major

PhD at Usyd – evolution of pregnancy in reptiles

Postdoc at Yale University – marsupial pregnancy

Postdoc at University of Melbourne
- Maternal-fetal communication

Lecturer at Macquarie – Evolutionary Genomics Lab



MACQUARIE
University

Human Genetics

With great power comes great responsibility

- Human Genome Project
- Next Generation Sequencing
- Genetic Testing
- Pre-implantation screening
- Genome Wide Associated Studies
- Personalised Medicine
- Genome Editing



BIOL3120 –Human Genetics & Evolutionary Medicine

LEARNING OBJECTIVES



On successful completion of this unit, you will be able to:

- **ULO1:** Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources.
- **ULO2:** Interpret and demonstrate understanding of the primary scientific literature.
- **ULO3:** Explain the importance of new techniques in human genetics for understanding human disease.
- **ULO4:** Explain the principles of evolutionary biology and their role in human health and disease.
- **ULO5:** Learn basic bioinformatic skills, including handling of genetic sequence data.

BIOL3120 –Human Genetics & Evolutionary Medicine

iLEARN

BIOL3120 Unit Information

Welcome to Human Genetics and Evolutionary Medicine. This online space is your virtual classroom, where we can work together to enhance your learning and experience in the unit. Embracing this online space is an important part of this unit. There are a number of things that contribute to online community building and one of them is establishing a visual online presence. Please take the time (if you haven't already) to setup your iLearn profile and include a photograph of yourself (or an avatar if you like!) that way we can bring a little personality into this online space. [Here's a guide to help.](#)

We'll be running this unit on a Flipped Classroom model. Therefore, it's really important that you watch the online lecture content *before* attending your practical class. You'll get much more out of the unit this way and we can answer your specific questions in class. We'll also be available on zoom during your scheduled lecture slot to answer all of your questions. Please see the timetable information for the zoom link.

This unit has the following Learning Outcomes. On completion of this unit you should be able to:

1. Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources
2. Interpret and demonstrate an understanding of the primary scientific literature
3. Explain the importance of new techniques in human genetics for understanding human disease
4. Explain the principles of evolutionary biology and their role in human health and disease
5. Learn basic bioinformatic skills, including handling of genetic sequence data

Week	Lectures	Practical Class	Assessments	Unit Learning Outcome
1	Introductory Lecture and Overview Intro to Evolutionary Medicine	No Practical Classes in Week 1		Explain the principles of evolutionary biology and their role in human health and disease
2	Oliver's Research Problem solving in genetics	No Practical Classes in Week 2		Interpret and demonstrate understanding of the primary scientific literature

BIOL3120 –Human Genetics & Evolutionary Medicine

iLEARN

Timetable: Lectures and Practicals



Activity	Day	Time	Location	Lecturer/Tutor
Lecture	Released Monday Morning	Watch before tutorial	iLearn	Oliver, Emily, or a Guest Lecturer
Drop In Session	Wednesday	1-2 pm	Zoom	Oliver
Practical 1	Wednesday	9-11 am	14 Sir Christopher Ondaatje Ave - 163 Active Learning Space	Erin
Practical 2	Thursday	3-5 pm	01CC 103 Active Learning Space	Oliver
Practical 3	Thursday	12-2 pm	01CC 214 Active Learning Space	Oliver
Practical 4	Wednesday	2-4 pm	01CC 218 Active Learning Space	Callum
Practical 5	Wednesday	11 am - 1 pm	Zoom	Erin

Last modified: Wednesday, 16 February 2022, 12:23 PM

BIOL3120 –Human Genetics & Evolutionary Medicine

iLEARN

 Announcements

▼ Open all **^ Close all**

BIOL3120 Unit Information

Communications

Please don't discuss problem sets which are still open, as they are worth marks. I will open discussion forums for problem sets after they close.

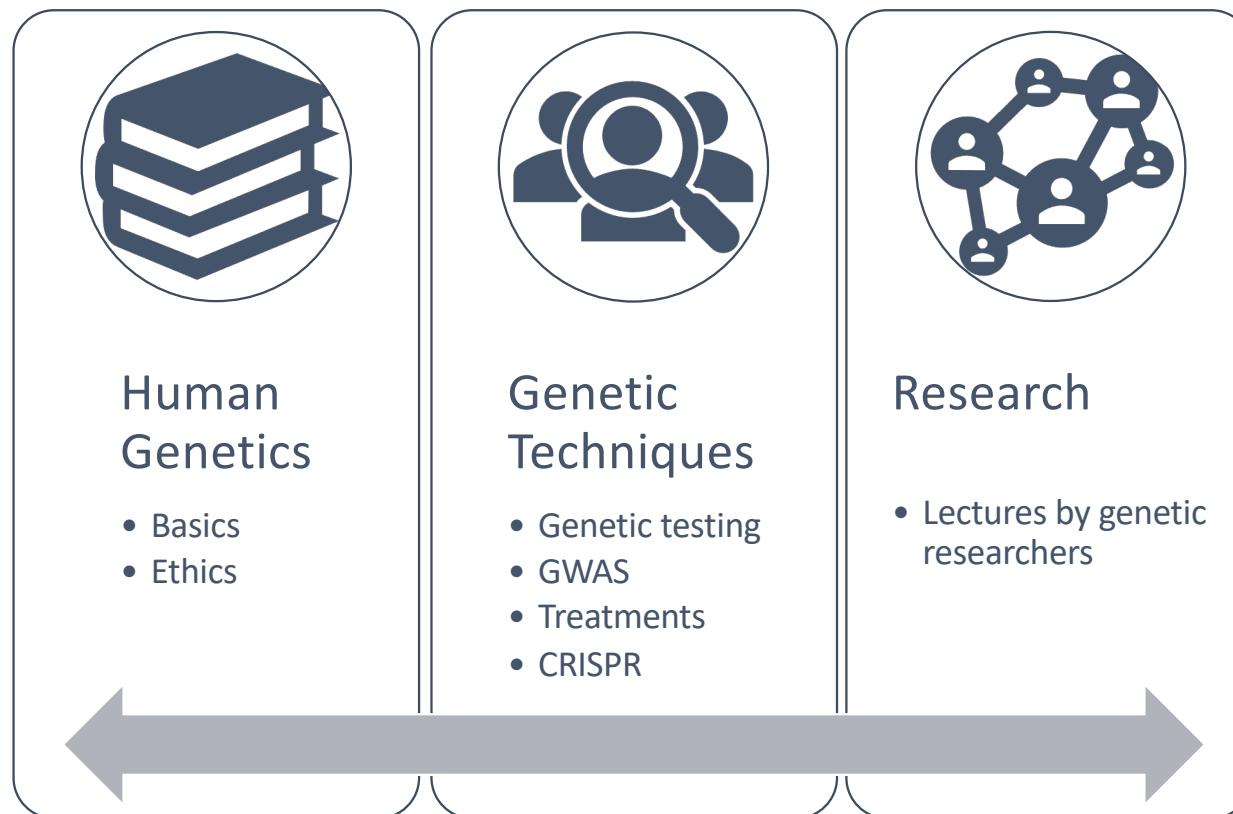
 Unit administration and general information discussion

 Lecture and unit content discussion

 Exam revision discussion

BIOL3120 –Human Genetics & Evolutionary Medicine

What to expect?



BIOL3120 –Human Genetics & Evolutionary Medicine Assessments

	Type of Assessment	Value	Due	Learning Outcome
AT1	Problem Sets and Practical Classes	25%	Internal Students Problem Set 1 (Week 3, Sunday 13th March) Problem Set 2 (Week 4, Sunday 20th March) Problem Set 3 (Week 5, Sunday 27th March) Problem Set 4 (Week 6, Sunday 3rd April) Problem Set 5 (Week 7, Sunday 10th April)	Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources Interpret and demonstrate understanding of the primary scientific literature Explain the principles of evolutionary biology and their role in human health and disease Learn basic bioinformatic skills, including handling of genetic sequence data

BIOL3120 –Human Genetics & Evolutionary Medicine Assessments

	Type of Assessment	Value	Due	Learning Outcome
AT2	Literature Review	25%	Week 13 (11:59 pm, Friday 3rd June)	<p>Interpret and demonstrate understanding of the primary scientific literature</p> <p>Explain the importance of new techniques in human genetics for understanding human disease</p>
AT3	Final Exam	50%	Exam Period	<p>Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources</p> <p>Interpret and demonstrate understanding of the primary scientific literature</p> <p>Explain the importance of new techniques in human genetics for understanding human disease</p> <p>Explain the principles of evolutionary biology and their role in human health and disease</p> <p>Learn basic bioinformatic skills, including handling of genetic sequence data</p>



What's next?

BIOL3120 –Human Genetics & Evolutionary Medicine

What's next?

Week	Lectures	Practical Class	Assessments	Unit Learning Outcome
1	Introductory Lecture and Overview Intro to Evolutionary Medicine	No Practical Classes in Week 1		Explain the principles of evolutionary biology and their role in human health and disease
2	Oliver's Research Problem solving in genetics	No Practical Classes in Week 2		Interpret and demonstrate understanding of the primary scientific literature
3	The Human Genome Modes of Inheritance and Population Genetics	Problem Set 1	Problem Set 1 (5%)	Explain the principles of evolutionary biology and their role in human health and disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources
4	Heritability and Polygenics Chromosomal Mutations	Problem Set 2	Problem Set 2 (5%)	Explain the principles of evolutionary biology and their role in human health and disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources

Flipped Classroom



BIOL3120 –Human Genetics & Evolutionary Medicine

A week in the life of a BIOL3120 student

Week 1 content

Welcome to Week 1 of Human Genetics and Evolutionary Medicine. This week we're going to start slowly by introducing the unit and the concept of evolutionary medicine. Whilst we don't have practical classes this week, we've included a module on genetics that you may wish to review to check your background knowledge. Please don't hesitate to get in touch if you have some questions for us, either via the forums for general knowledge or via email for private matters.

Lecture Material

-  [Lecture 01 -Introductory Lecture and Overview Recording -being updated](#)
-  [Lecture 01 -Introduction and Overview Slides -being updated](#) 3.6MB PDF document
-  [Lecture 02 -Introduction to Evolutionary Medicine Recording](#)
-  [Lecture 02 -Intro to Evolutionary Medicine Slides](#) 13.4MB PDF document

Practical Class Preparation and Resources

Whilst we don't have practical classes this week. You may wish to review this genetics module to check your background knowledge.

-  [Assumed Genetics Knowledge -being updated](#)

Your Feedback is Important to Us

Hidden from students

-  [Unit Difficulty](#)

BIOL3120 –Human Genetics & Evolutionary Medicine

A week in the life of a BIOL3120 student

The image shows a video player interface. At the top right, there is a small video thumbnail of a person with glasses and a blue top. To the left of the thumbnail is the logo for MacEwan University, which consists of a red shield with a white castle and the text 'MAC Univ'. The main video frame displays the title 'BIOL3120 –Human Genetics and Evolutionary Medicine' and 'Faculty of Science and Engineering' in white text on a dark background. Below the title is a photograph of a modern building with a glass facade and a colorful, geometric mural on its side. The mural features large, overlapping rectangles in shades of red, orange, yellow, and purple. A progress bar at the bottom of the video frame indicates the video is at 0:06 / 4:32. Below the video frame is a horizontal control bar with several icons: a blue play button, a volume icon, a full-screen icon, a three-dot menu icon, a star icon, and a settings icon.

BIOL3120 –Human Genetics & Evolutionary Medicine

LEARNING OBJECTIVES



On successful completion of this unit, you will be able to:

- **ULO1:** Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources.
- **ULO2:** Interpret and demonstrate understanding of the primary scientific literature.
- **ULO3:** Explain the importance of new techniques in human genetics for understanding human disease.
- **ULO4:** Explain the principles of evolutionary biology and their role in human health and disease.
- **ULO5:** Learn basic bioinformatic skills, including handling of genetic sequence data.



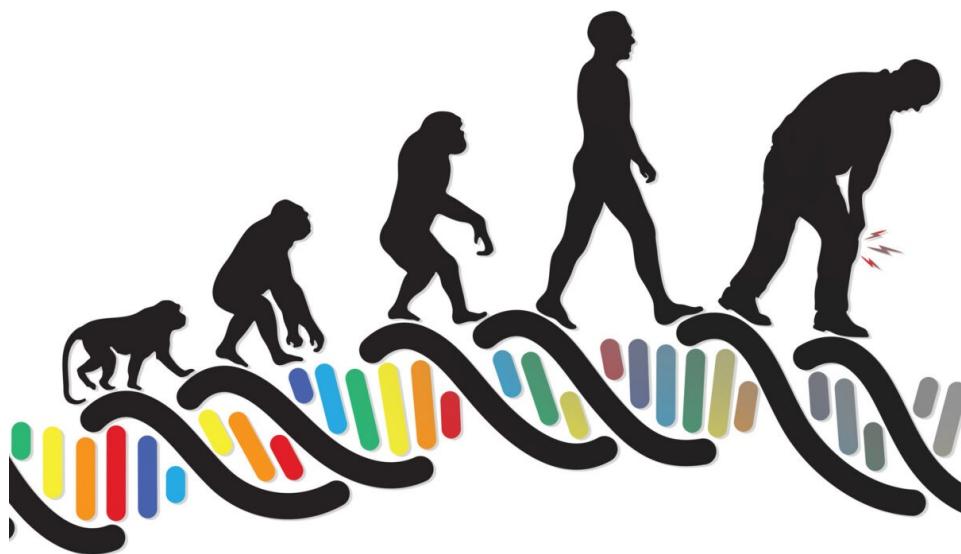
MACQUARIE
University
SYDNEY · AUSTRALIA

Introduction to Evolutionary Medicine

BIOL3120 – LECTURE 2



What is evolutionary medicine?



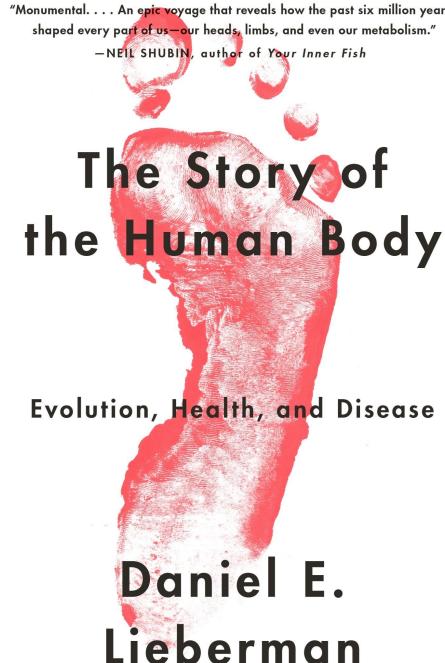
- Uses evolutionary biology to understand, prevent, and treat disease
- Evo Med, uses the knowledge that we and our pathogens are the product of evolution, to understand how we were put together, how we work, and how disease manifests

Book recommendation



- Accessible book that looks at human evolution and how it impacts on human health
- Great read if you have some down time through the semester

"Monumental. . . An epic voyage that reveals how the past six million years shaped every part of us—our heads, limbs, and even our metabolism."
—NEIL SHUBIN, author of *Your Inner Fish*



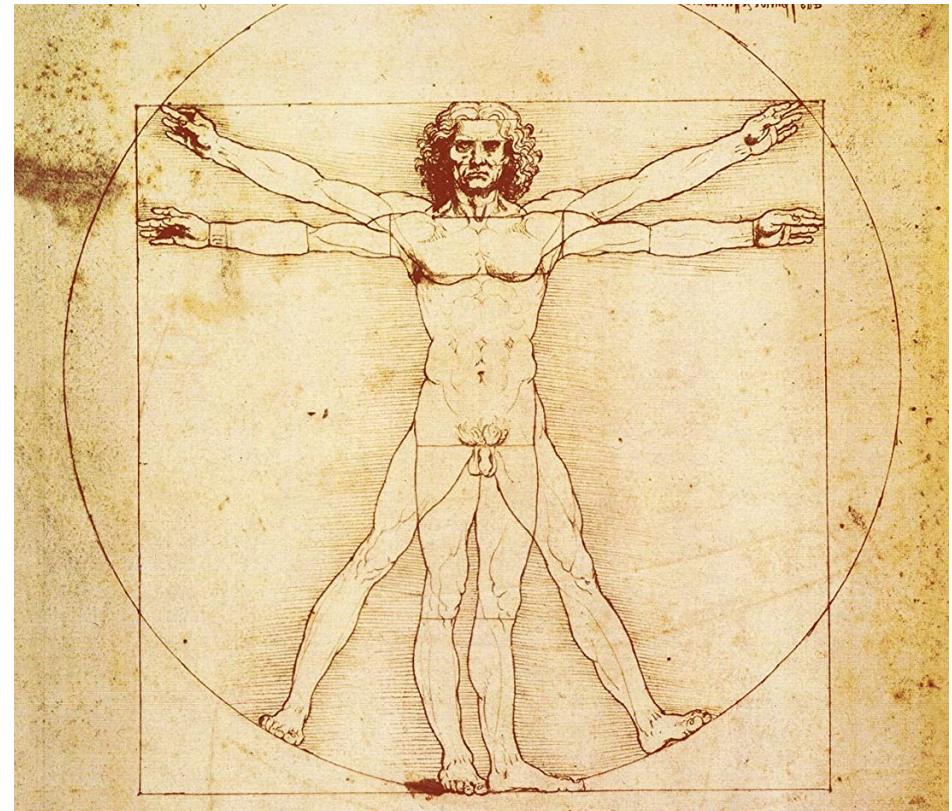
Evo Med – ted talk

Ultimate vs. proximate understanding of disease

Some evolutionary principles

- Natural selection acts slowly
- There are limitations to what evolution can produce
- We are optimized for reproductive success, not health

Constant arms race with pathogens



Evolution, Medicine, and Public Health [2018] pp. 13–23
doi:10.1093/emph/eox025

ORIGINAL
RESEARCH
ARTICLE

Core principles of evolutionary medicine

A Delphi study

Daniel Z. Grunspan,^{1,2} Randolph M. Nesse,^{1,2} M. Elizabeth Barnes² and
Sara E. Brownell*^{1,2}



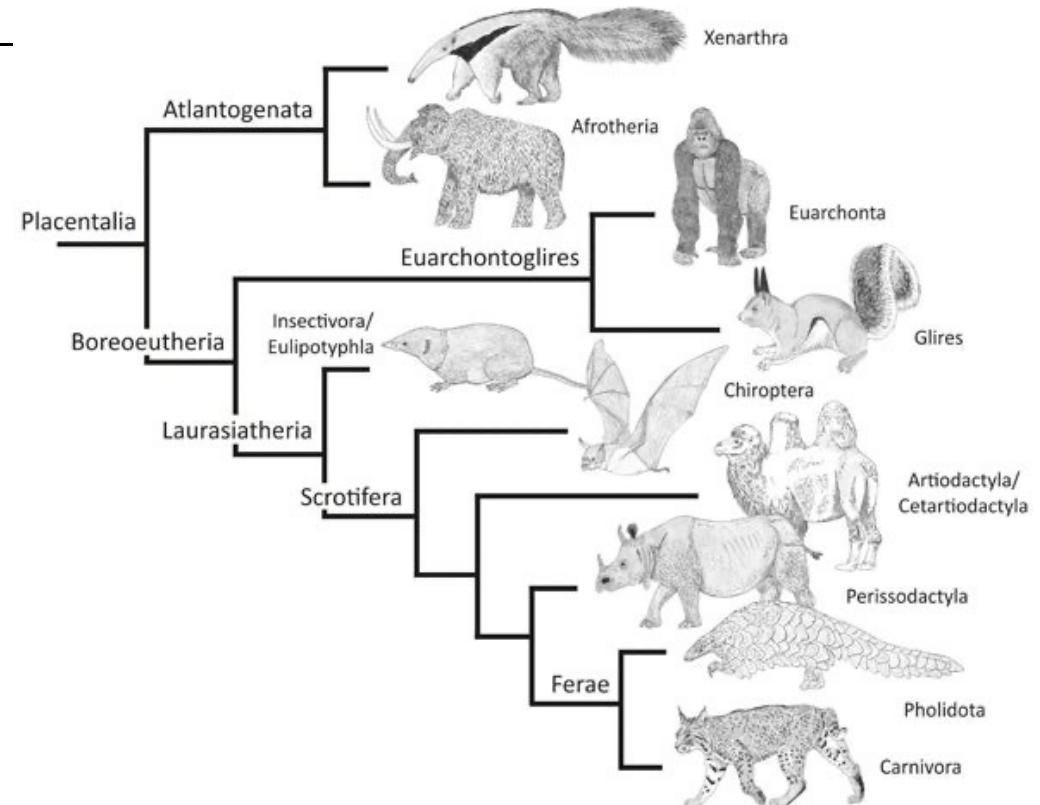
EVOLUTION,
MEDICINE, &
PUBLIC HEALTH



Topic	Core principle
Types of explanation (question framing)	Both proximate (mechanistic) and ultimate (evolutionary) explanations are needed to provide a full biological understanding of traits, including those that increase vulnerability to disease.
Evolutionary processes (evolution I)	All evolutionary processes, including natural selection, genetic drift, mutation, migration and non-random mating, are important for understanding traits and disease.
Reproductive success (evolution I)	Natural selection maximizes reproductive success, sometimes at the expense of health and longevity.
Sexual selection (evolution I)	Sexual selection shapes traits that result in different health risks between sexes.
Constraints (evolution I)	Several constraints inhibit the capacity of natural selection to shape traits that are hypothetically optimal for health.
Trade-offs (evolutionary trade-offs)	Evolutionary changes in one trait that improve fitness can be linked to changes in other traits that decrease fitness.
LHT (evolutionary trade-offs)	Life history traits, such as age at first reproduction, reproductive lifespan and rate of senescence, are shaped by evolution, and have implications for health and disease.
Levels of selection (evolution II)	Vulnerabilities to disease can result when selection has opposing effects at different levels (e.g. genetic elements, cells, organisms, kin and other levels).
Phylogeny (evolution II)	Tracing phylogenetic relationships for species, populations, traits or pathogens can provide insights into health and disease.
Coevolution (evolution II)	Coevolution among species can influence health and disease (e.g. evolutionary arms races and mutualistic relationships such as those seen in the microbiome).
Plasticity (evolution II)	Environmental factors can shift developmental trajectories in ways that influence health and the plasticity of these trajectories can be the product of evolved adaptive mechanisms.
Defenses (reasons for vulnerability)	Many signs and symptoms of disease (e.g. fever) are useful defenses, which can be pathological if dysregulated.
Mismatch (reasons for vulnerability)	Disease risks can be altered for organisms living in environments that differ from those in which their ancestors evolved.
Cultural practices (culture)	Cultural practices can influence the evolution of humans and other species (including pathogens), in ways that can affect health and disease (e.g. anti-biotic use, birth practices, diet, etc.).

Core principles of evolutionary medicine

- Evolutionary understanding is required to fully understand traits including disease susceptibility
- Evolutionary processes shape traits and disease
- Natural selection maximises for reproductive success not health
- Evolution has trade offs
- Many signs and symptoms of disease are useful defenses (e.g. fever)



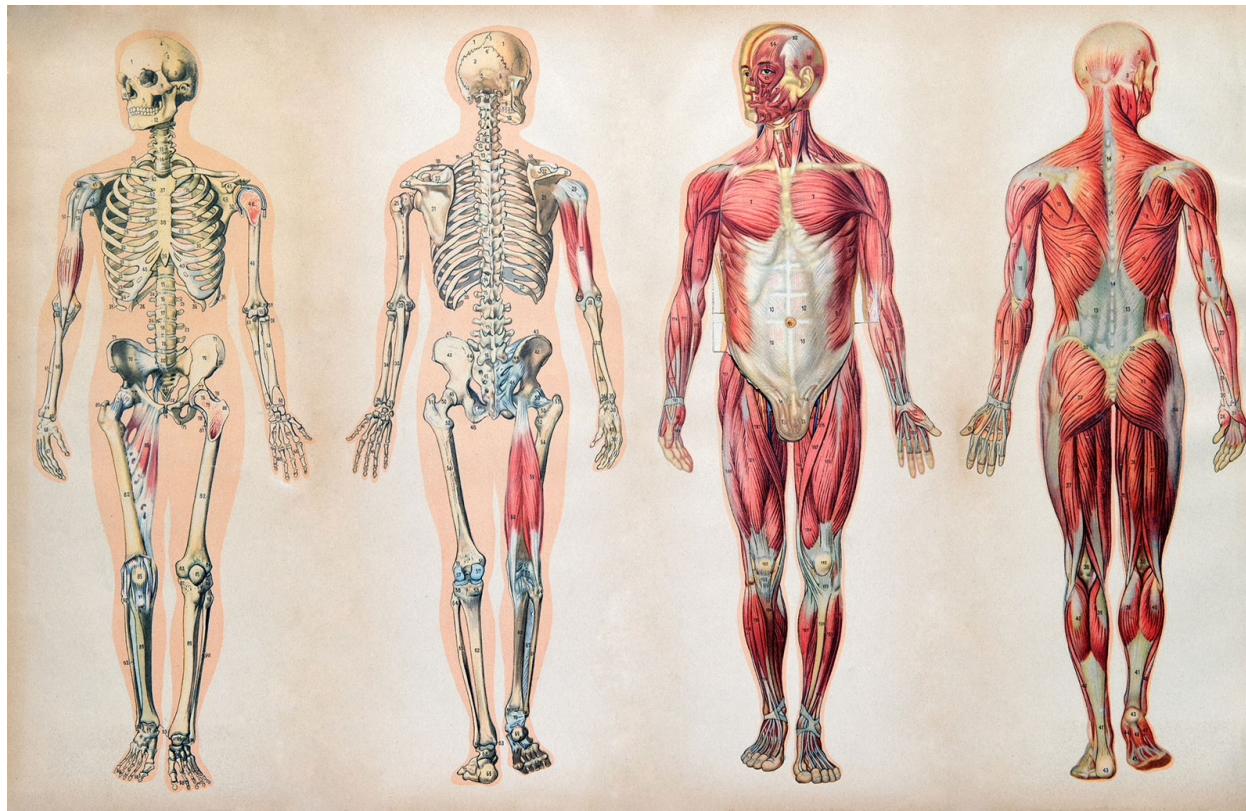
Darren Naish

Evolutionary understanding is required to fully understand traits including disease susceptibility

- Our bodies are the product of evolution, so to truly understand health, we must know how we were put together
- Ultimate vs. proximate
- Type II Diabetes
 - **Proximate cause:** fat and muscle become insulin resistant and pancreas can't produce enough insulin
 - **Ultimate cause:** changes in diet and exercise regime after agriculture and industrialisation of society



Evolutionary processes shape traits and disease



Natural selection maximises for reproductive success not health

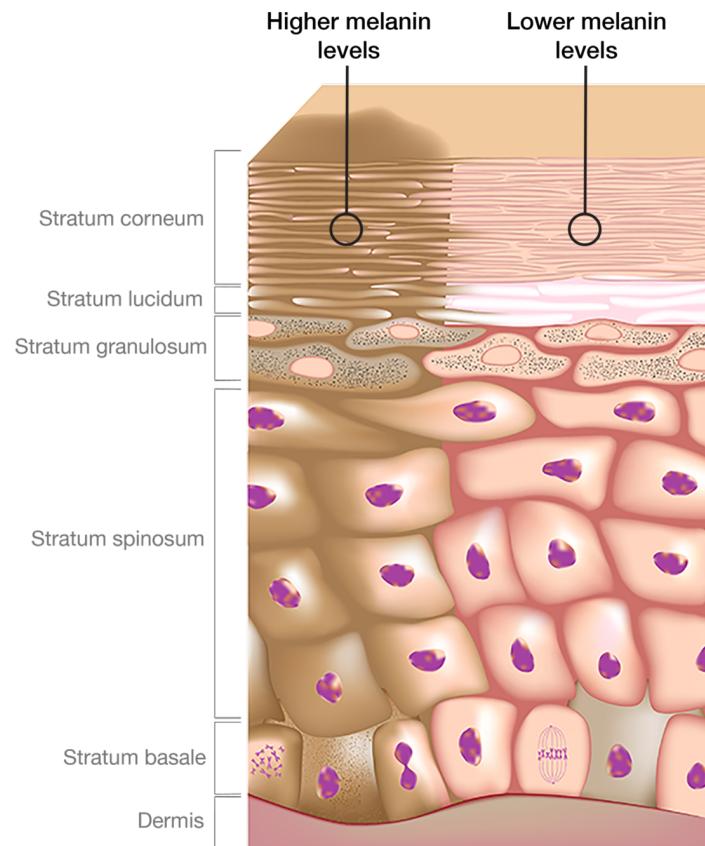
- No where is this more evident than in the brown antechinus
- Males live for just 1 year
- They can re-produce in a short 2 week breeding season each year
- During the breeding season males spend all their time fighting to gain access to females
- This produces so much stress hormones that their body gives up the ability to fight disease and otherwise function
- At the end of the breeding all males in the population die



Evolution has trade offs

- Fitness advantages often come at a cost in other areas of our physiology
- Loss of melanin leads to greater vitamin D synthesis in cooler climates, but leads to increased susceptibility to skin cancer

SKIN PIGMENTATION

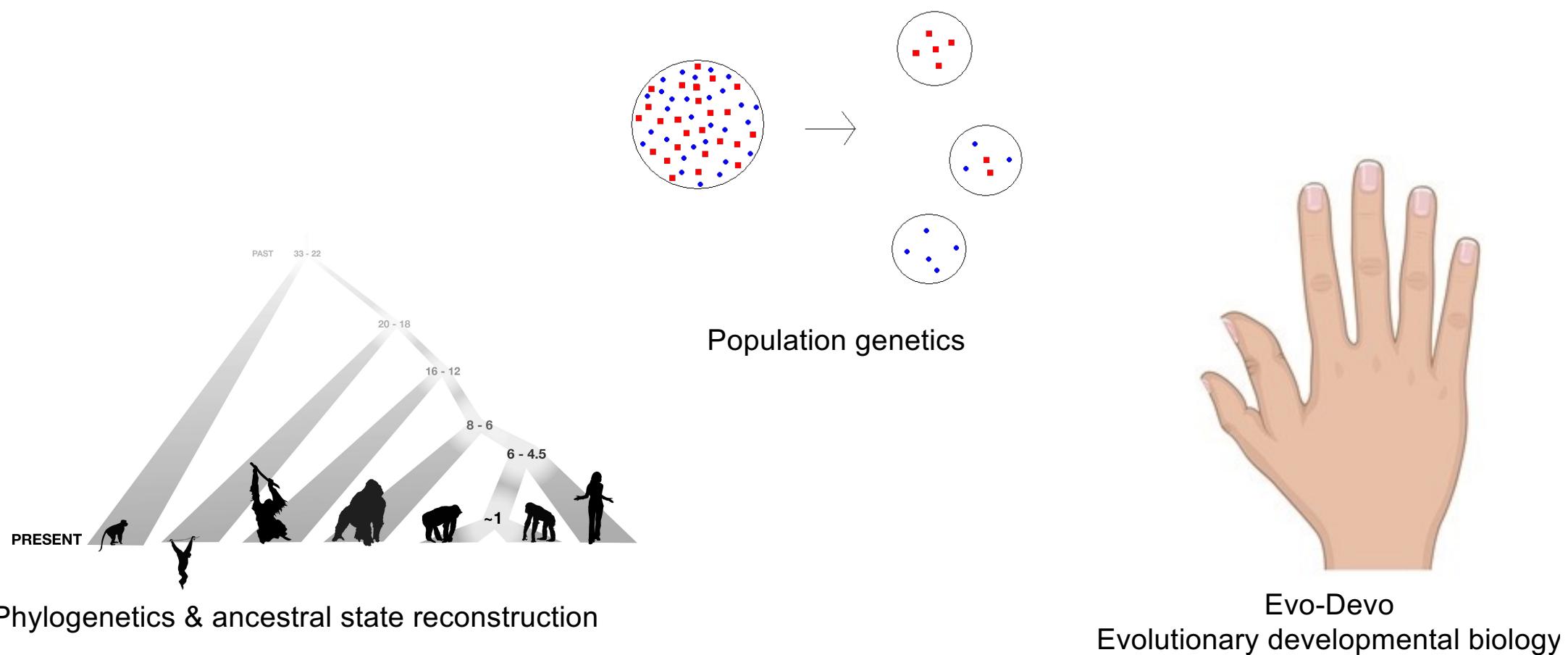


Many signs and symptoms of disease are useful defenses (e.g. fever)

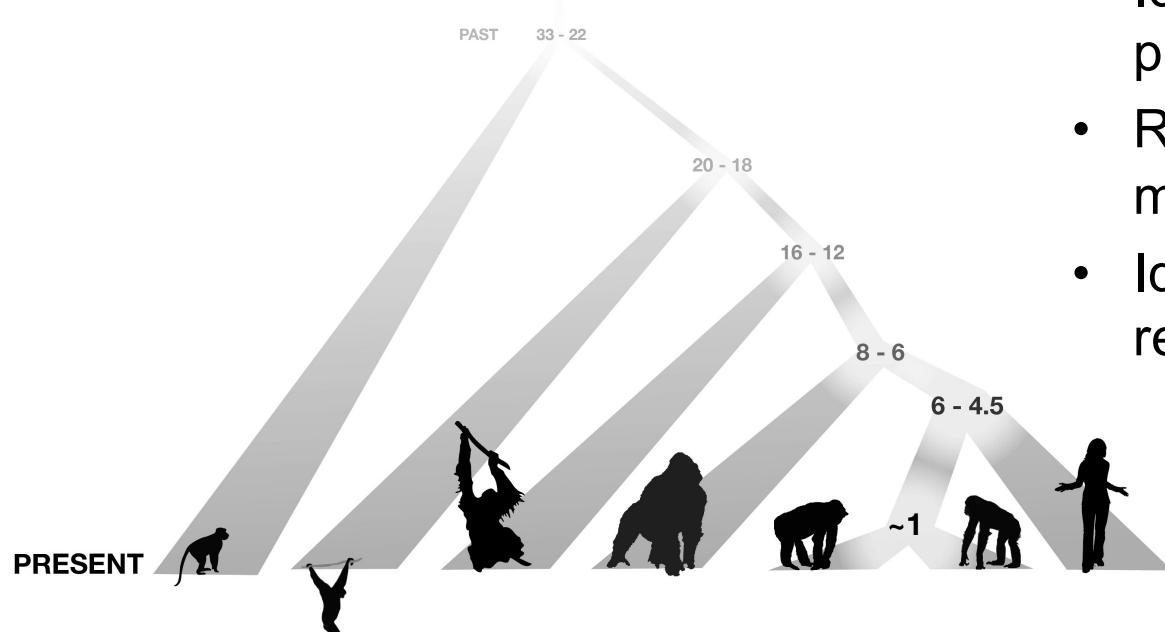
- Tissue damage and infection results in pro-inflammatory signals
- Inflammation recruits the immune system, leads to swelling, and fever
- Inflammation is painful and uncomfortable
- Managing inflammation is often done to decrease discomfort
- Inflammation is a necessary part of immune system to promote healing and eliminate pathogens



Tools of evolutionary biology



Applications of evolutionary tools



Phylogenetics

- Identify evolutionary relationship of species
- Identify the timeline of evolutionary processes
- Required to evaluate animal models
- Identify the evolutionary relationship of pathogens

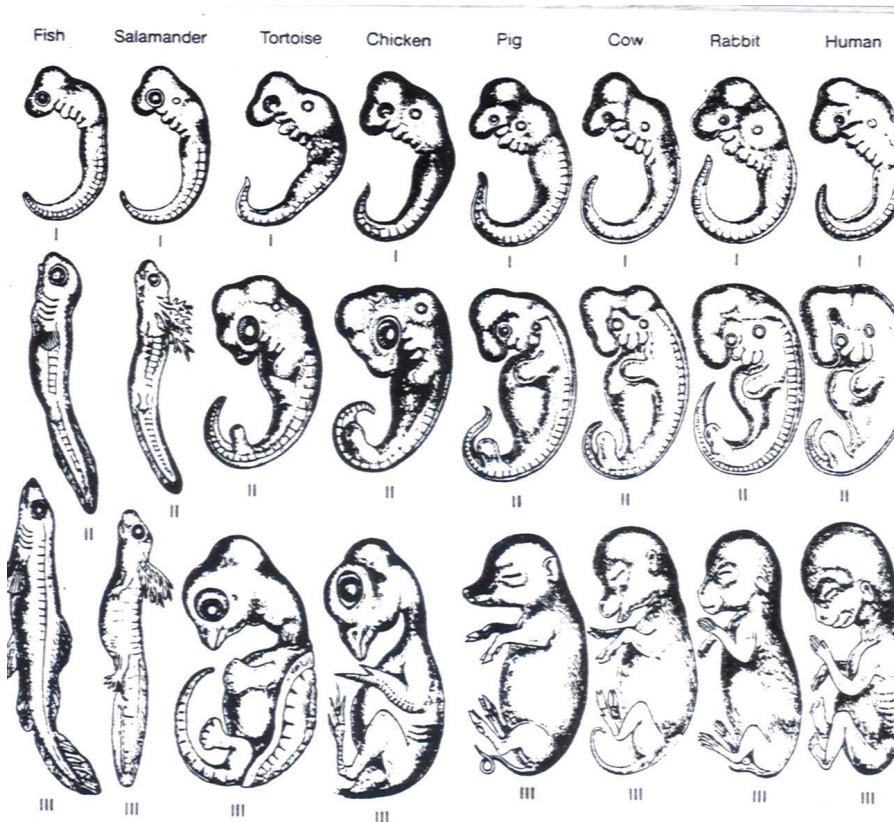
Applications of evolutionary tools

Population genetics

- Field of genetics that identifies frequency of genes within and between populations
- Required to identify genes associated with disease phenotypes
- Identify difference in disease susceptibility between populations



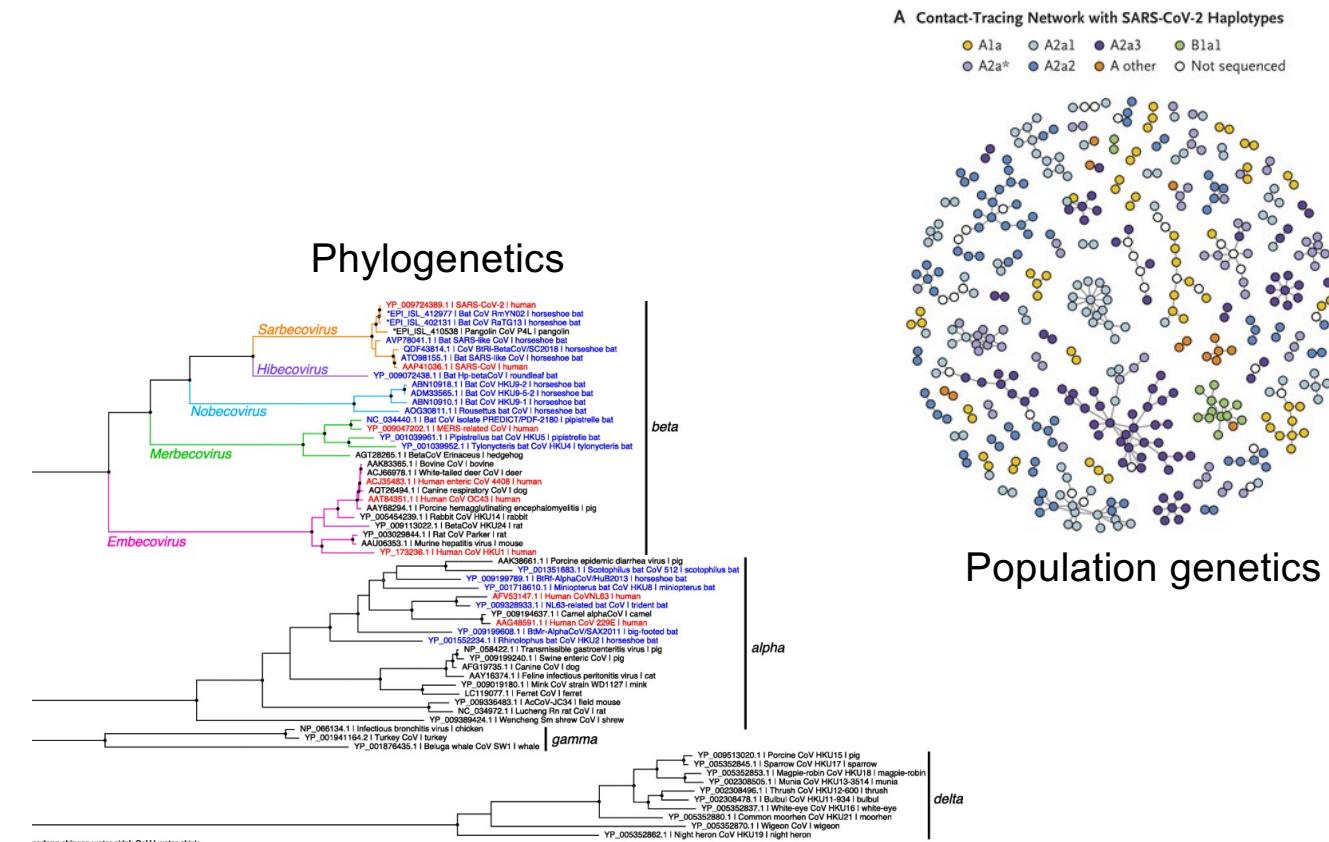
Applications of evolutionary tools



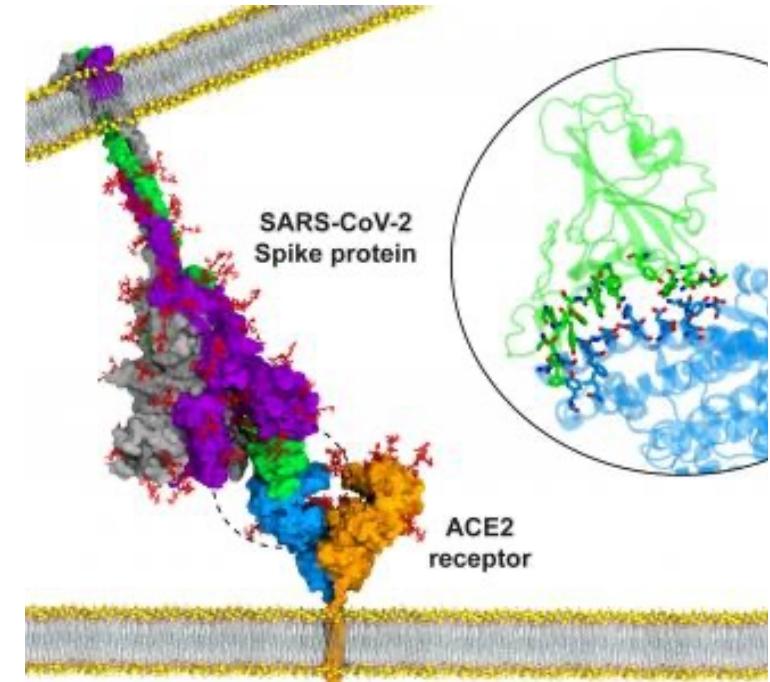
Evo devo

- Compares developmental process between organisms, to identify how development has shaped our evolutionary history
- Vulnerabilities of development
- Understand vestigial structures
- Understand quirks of human anatomy

SARS-CoV-2



Evo-Devo



Complexities of implantation: a case for evolutionary medicine

Oliver Griffith



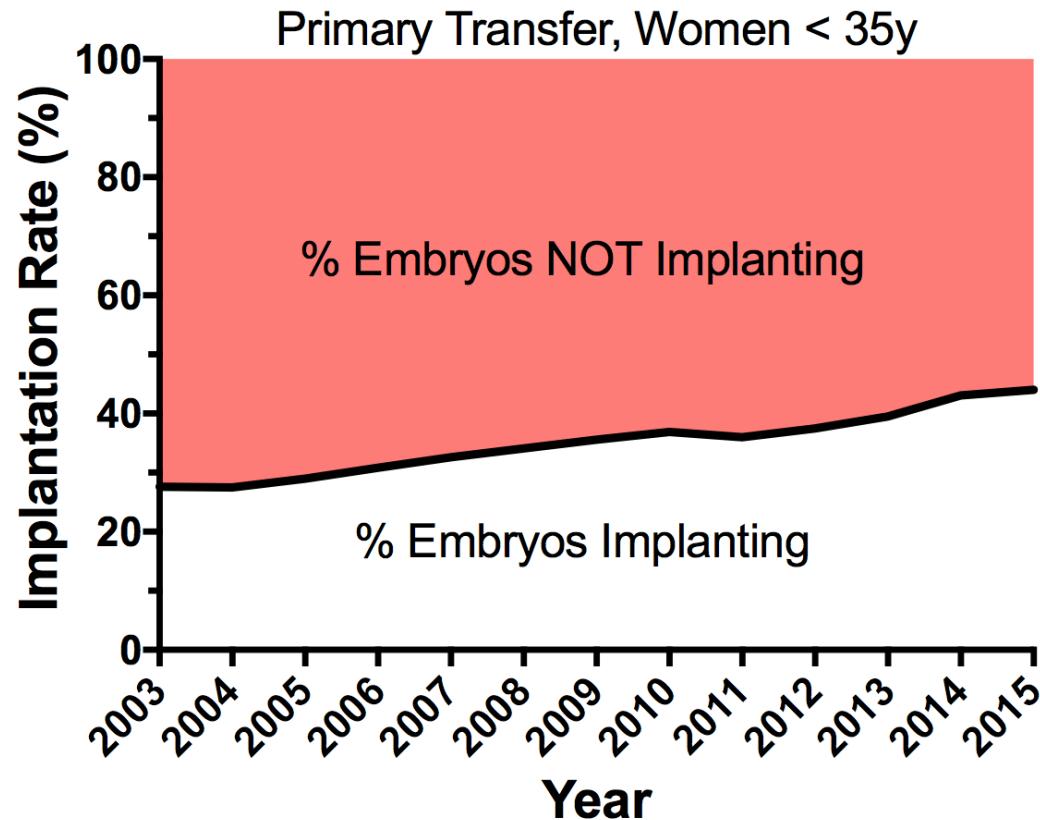
@oligriffith



Implantation failure

- 75% of pregnancies are lost at implantation
- If implantation is imperfect, this can lead to complications further in pregnancy
- While gains in IVF success rates have been made, these are predominantly in increasing embryonic growth, and screening embryo developmental success

Implantation in infertile couples

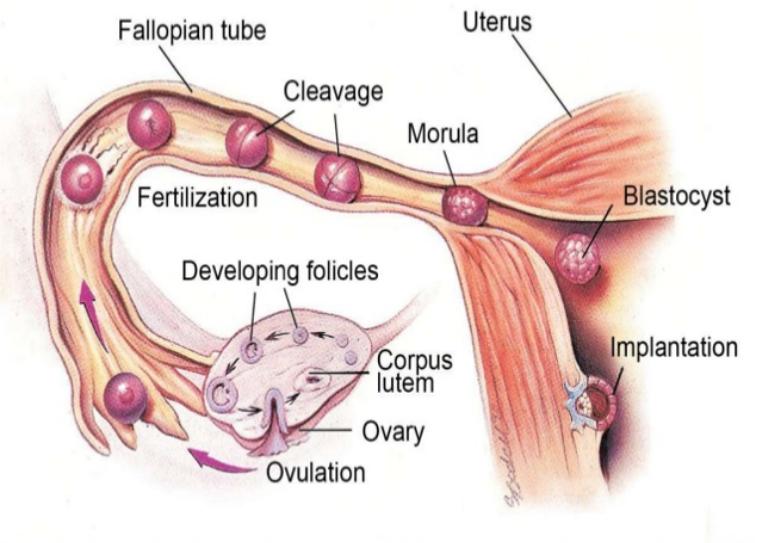


About 60% of
Embryos Fail to
Implant in Young
Women

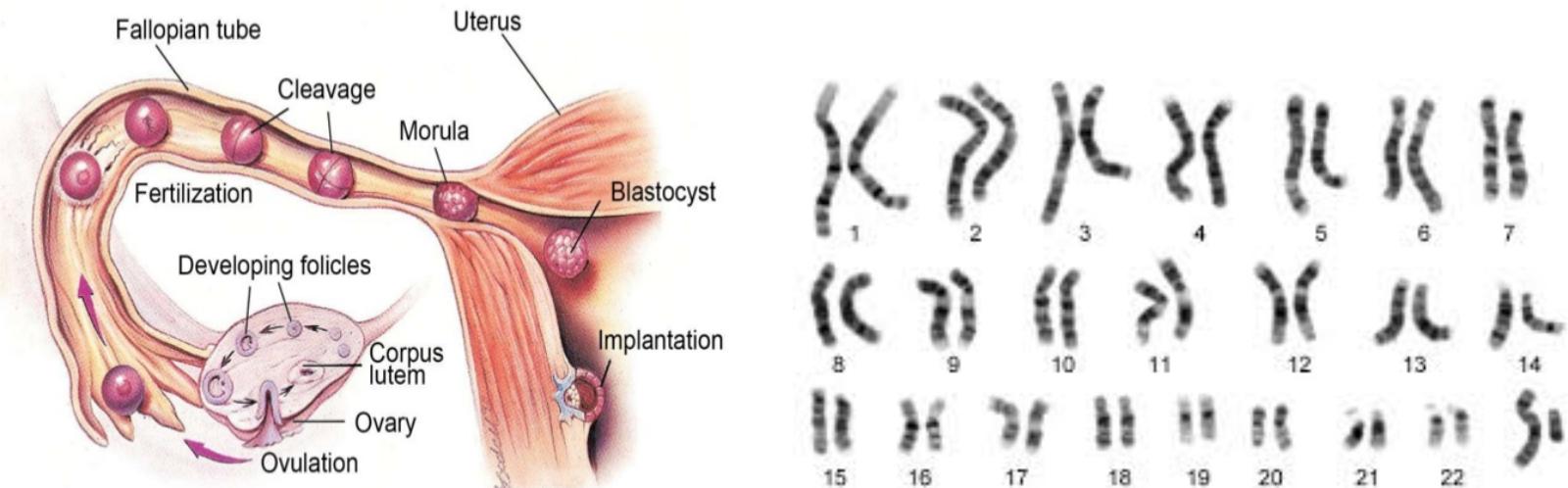
We need a better understanding of the endometrial changes that support implantation so we can better characterize, diagnose, and treat implantation disorders in women.



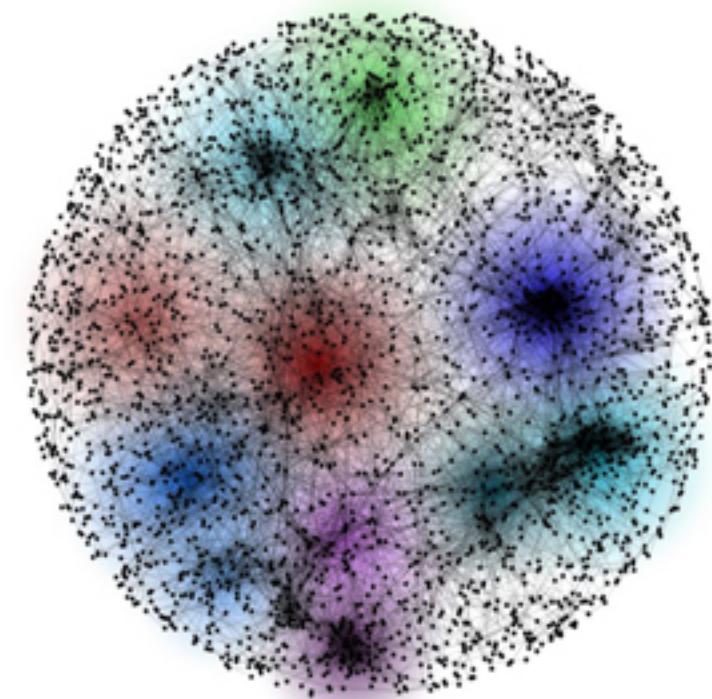
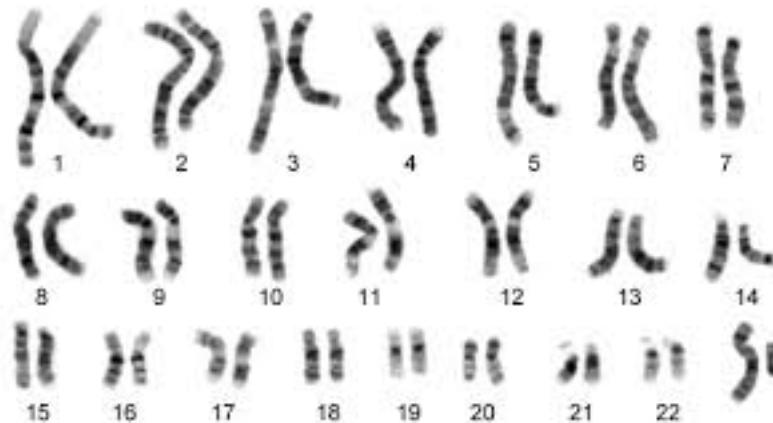
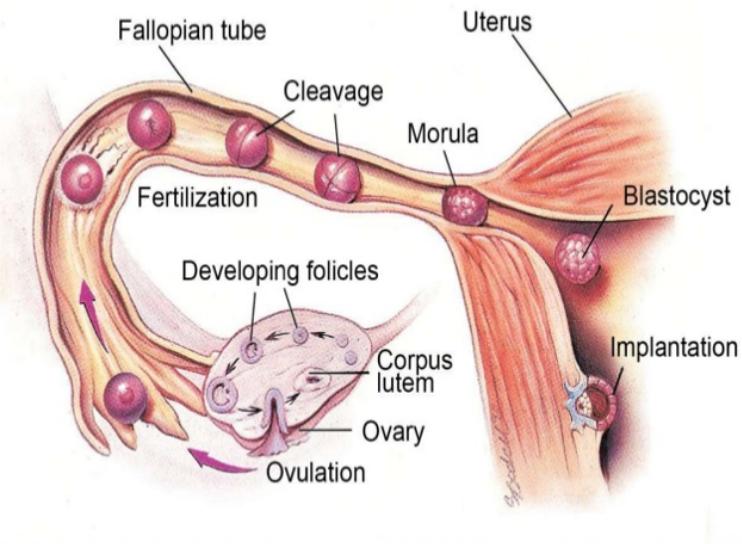
Biology is complex



Biology is complex



Biology is complex



Evolutionary comparisons allow us to understand how things were put together

Evolutionary comparisons allow us to understand how things were put together

- We can identify changes associated with the evolution of a trait
 - Identify candidate genes that underpin the process

Evolutionary comparisons allow us to understand how things were put together

- We can identify changes associated with the evolution of a trait
 - Identify candidate genes that underpin the process
- We can identify conserved components of a process
 - Most critical components for the physiological outcome

Evolutionary comparisons allow us to understand how things were put together

- We can identify changes associated with the evolution of a trait
 - Identify candidate genes that underpin the process
- We can identify conserved components of a process
 - Most critical components for the physiological outcome
- We can identify how different animals do things differently
 - Alternative strategies for treatment

Pregnancy and the immune system

The immune system has been widely thought of as being dangerous for pregnancy, because it may 'reject' the fetus as foreign tissue for containing, foreign, paternal genetic material.



Controlling the immune system during implantation

The screenshot shows a top navigation bar with a green progress bar. Below it is a purple header bar containing links: About Fertility, Fertility Treatment, Success Rates, Specialists & Clinics, About Us, Fees, and Resources. Underneath is a grey banner with the breadcrumb trail: Home > Fertility Treatment > Fertility Treatments.

- Getting Started >
- Fertility Treatments >
- Ovulation Cycle Tracking >
- Ovulation Induction (OI) >
- Artificial Insemination/IUI >
- IVF Treatment >
- ICSI Treatment >
- IMSI Procedure >
- Frozen Embryo Transfer >

Natural Killer Cell Testing

Immune cells in the uterus are important in the early detection and elimination of foreign cells, such as infections or cancer. These immune cells are normally present in every person as part of their immune system.

'Natural Killer cells' (NK Cells) play an important role in the immune system responses to viral infections as well during implantation. Some women who have fertility problems, and specifically **recurrent miscarriage** or failed IVF, are more likely to have higher levels of activity of these NK cells than other women. What this means for treatment is still not clear, but a number of different treatments are being trialled.

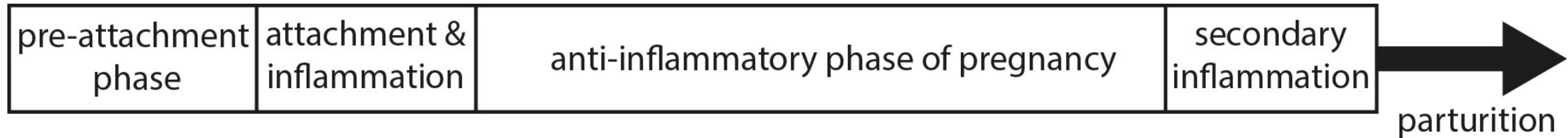


WHAT IS NATURAL KILLER CELL TESTING?

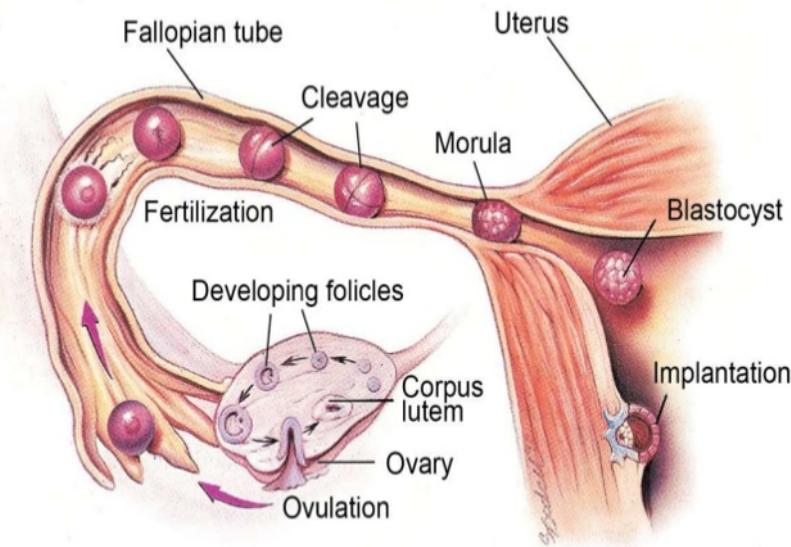
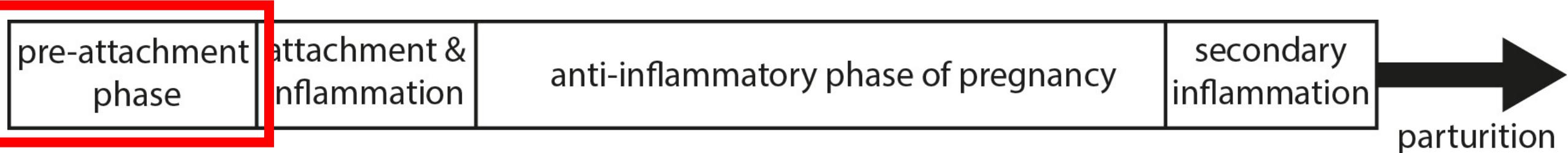
HOW DO I GET THE NATURAL KILLER CELL TEST?

MORE ABOUT NATURAL KILLER CELLS

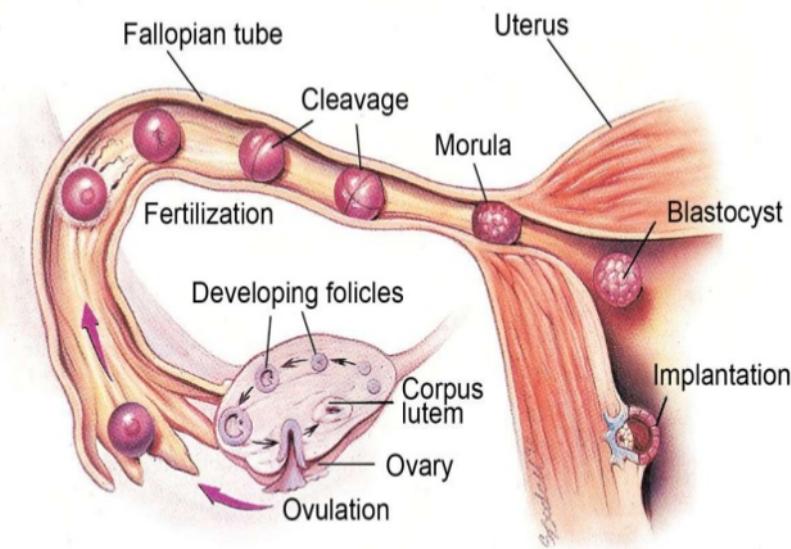
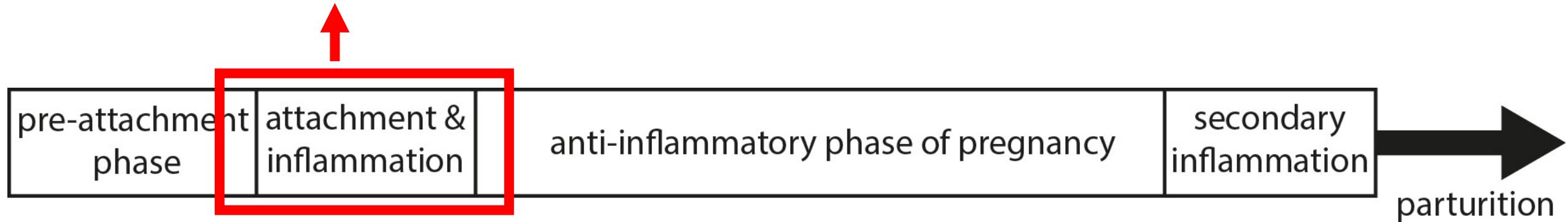
The inflammation paradox



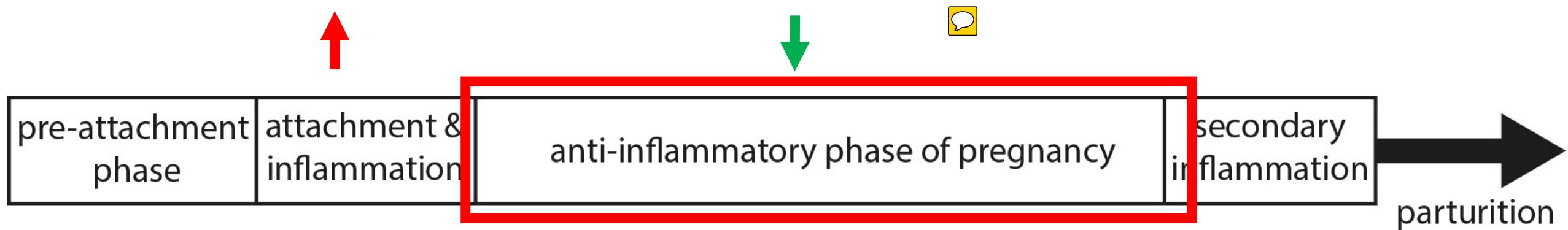
The inflammation paradox



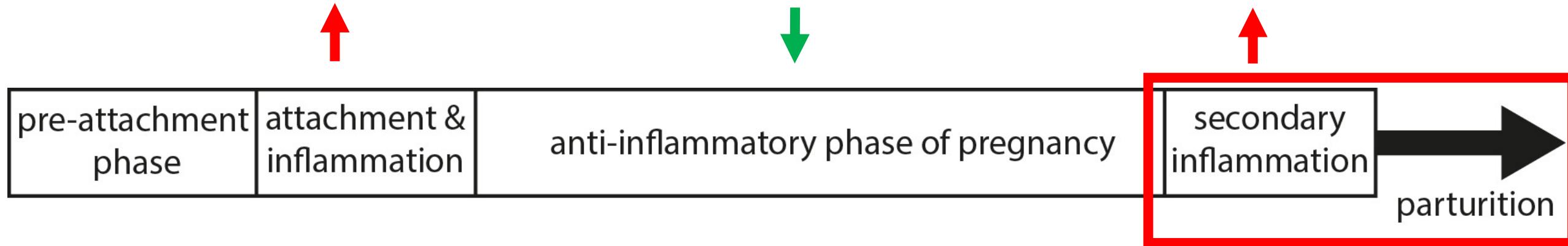
The inflammation paradox



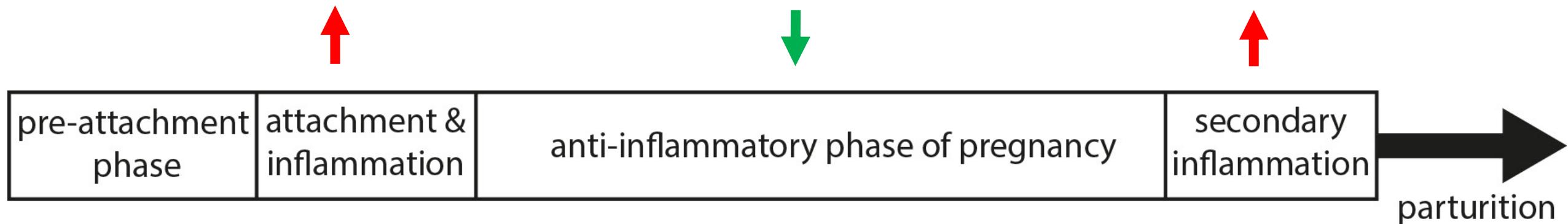
The inflammation paradox



The inflammation paradox



The inflammation paradox



Why is inflammation used to regulate normal physiological processes when it is also the biggest threats to pregnancy during most of development?

Talk outline

- An evolutionary model of inflammation in pregnancy

Talk outline

- An evolutionary model of inflammation in pregnancy
- Hypothesise about the role of inflammation in maternal recognition of pregnancy?

Talk outline

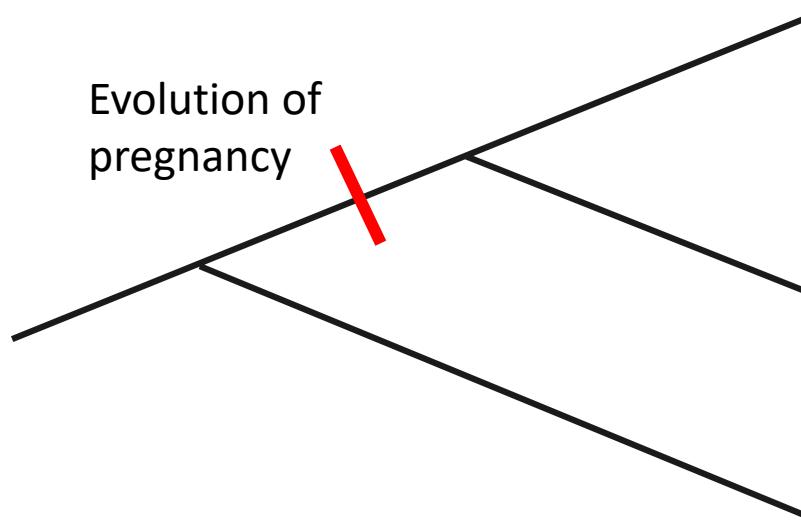
- An evolutionary model of inflammation in pregnancy
- Hypothesise about the role of inflammation in maternal recognition of pregnancy?
- Interrogating the model, how can we learn more about our own pregnancy?

The first mammals laid eggs

- Major transition in therian mammals was the shift from egg laying to producing live young



Phylogeny of mammals



Eutherians



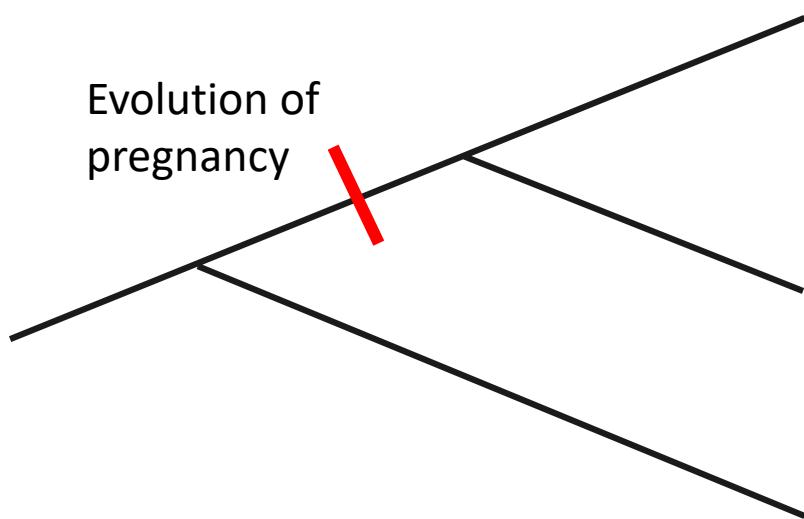
Marsupials



Monotremes



Phylogeny of mammals



Eutherians



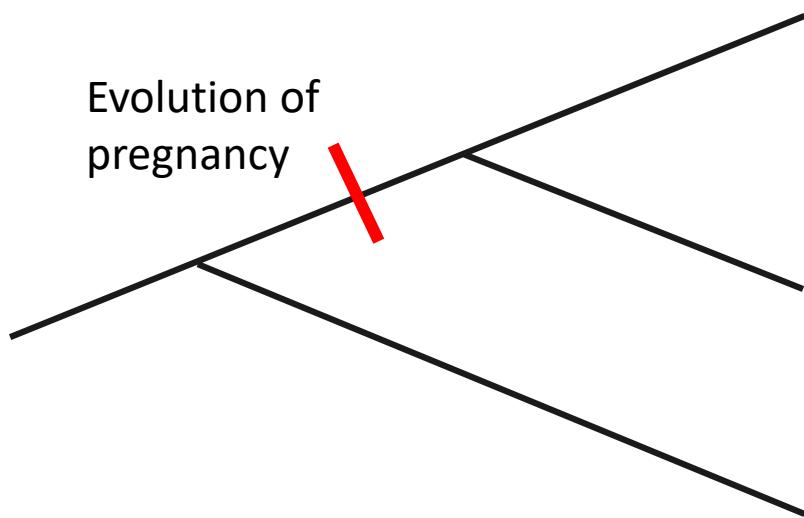
Marsupials



Monotremes



Phylogeny of mammals



Eutherians



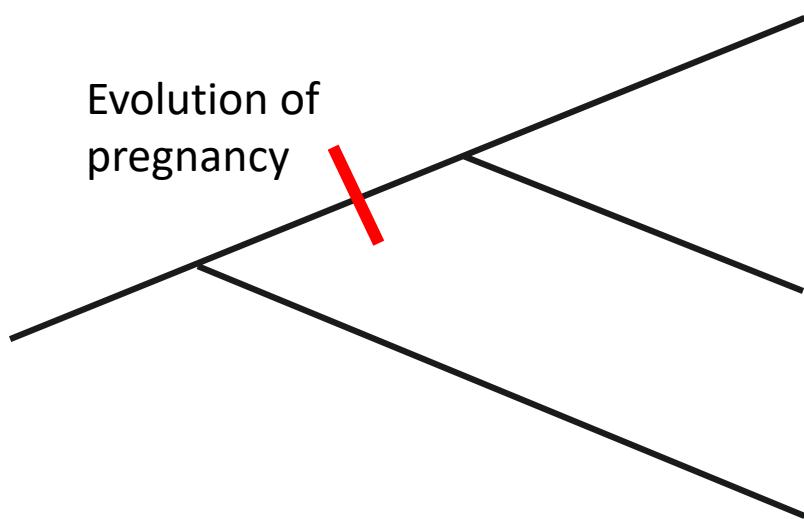
Marsupials



Monotremes



Phylogeny of mammals



Eutherians



Marsupials

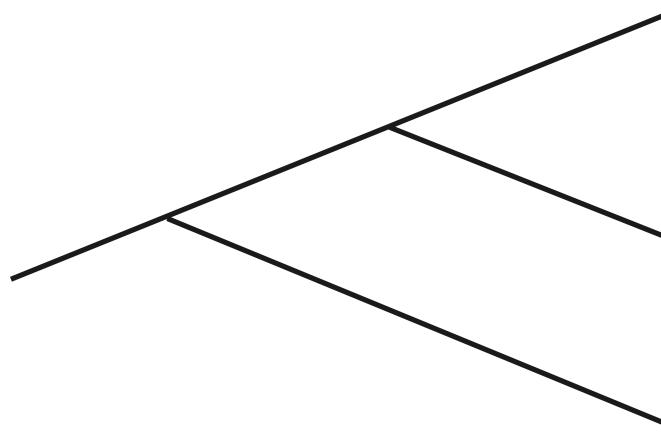


Monotremes



Marsupials share features with both monotreme and eutherian reproduction

- Monotreme like:
 - Embryonic development is short
 - Very precocious young
- Eutherian like:
 - Viviparous
 - Formation of a placenta



Eutherians



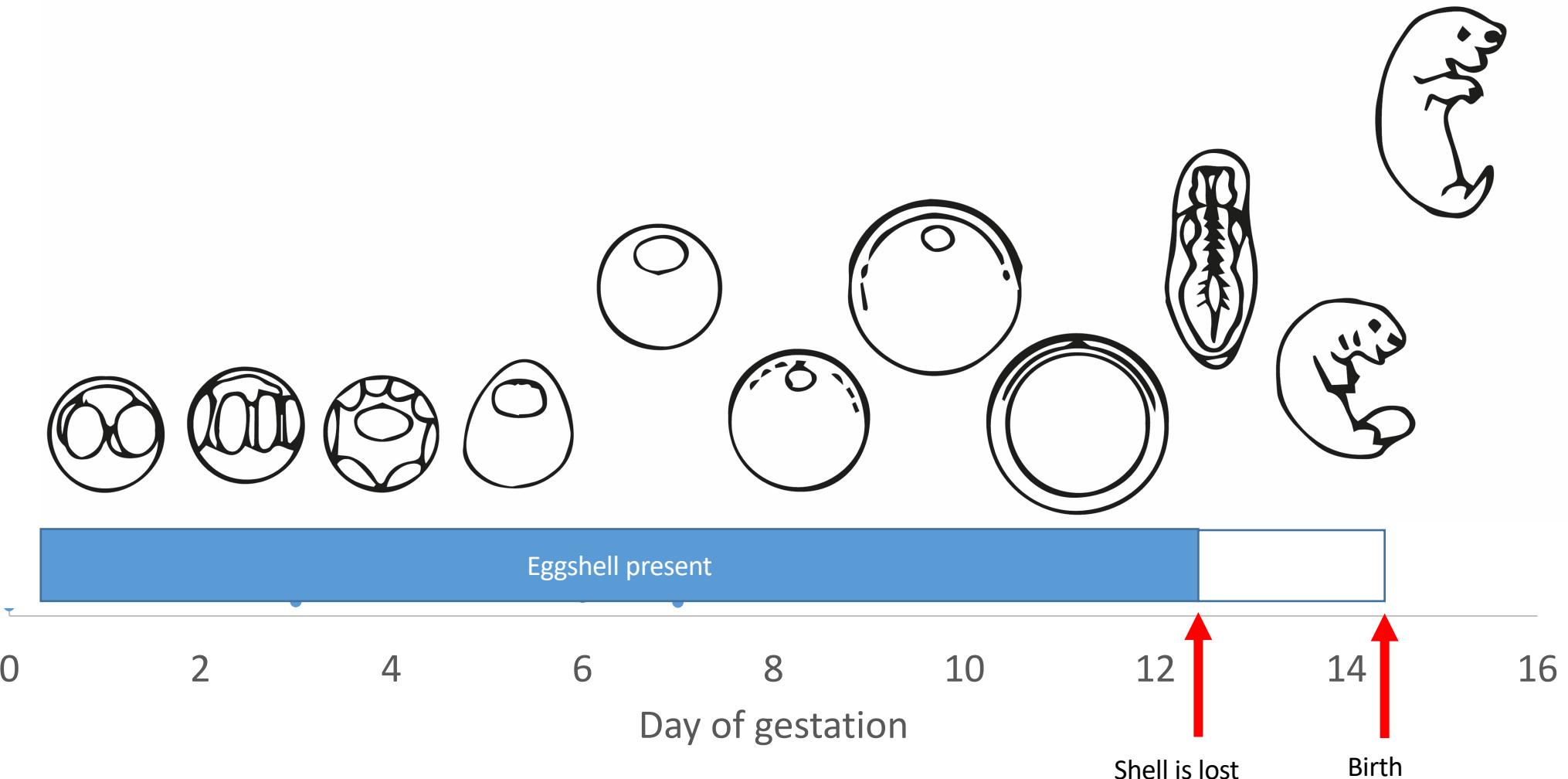
Marsupials



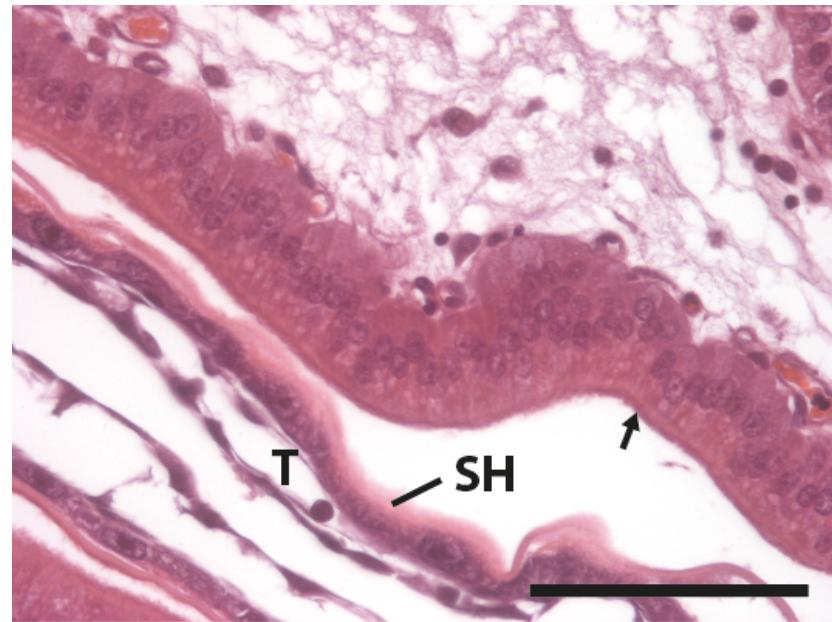
Monotremes



Embryonic development of marsupials

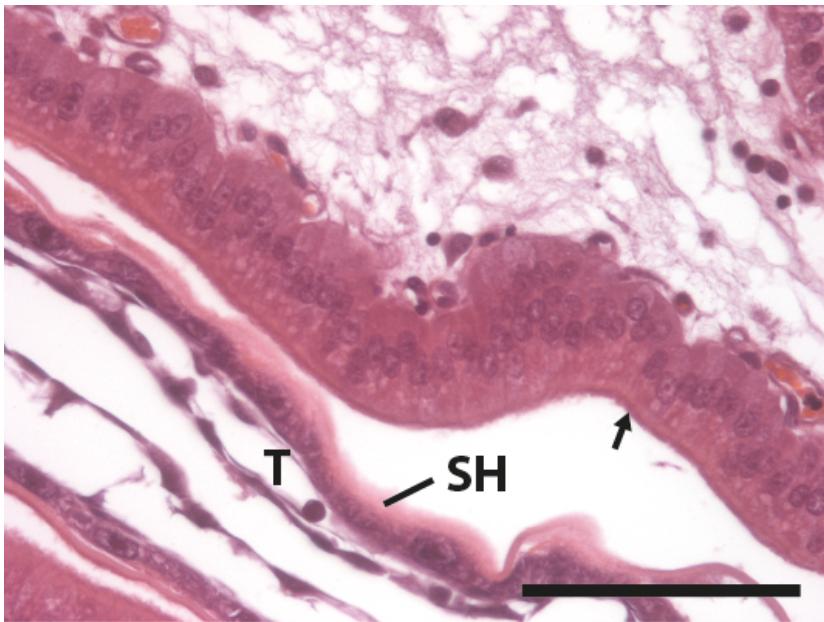


Eggshell is lost on day 12 of a 14 day pregnancy

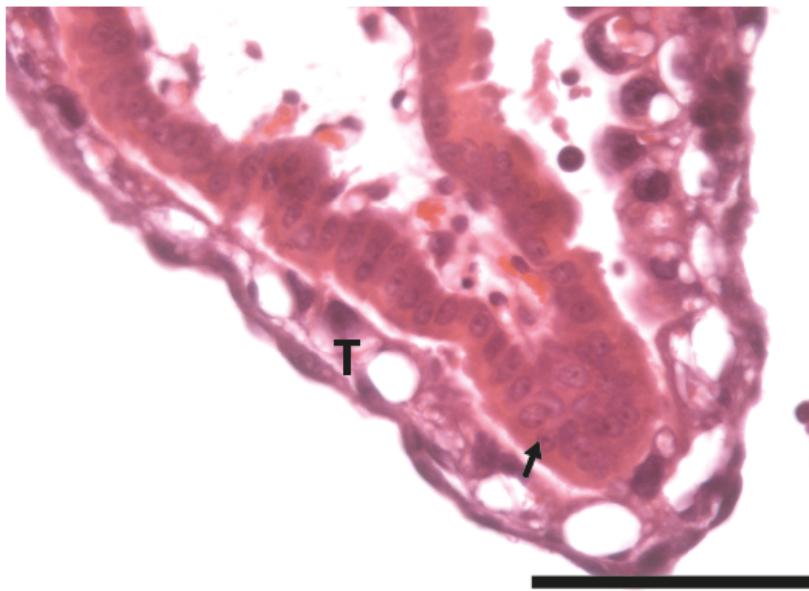


Day 11.5

Eggshell is lost on day 12 of a 14 day pregnancy

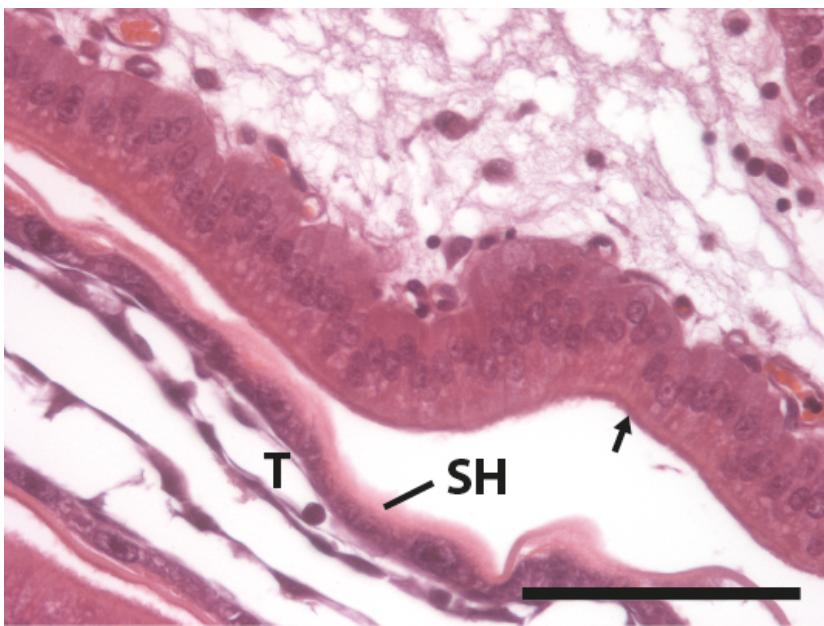


Day 11.5

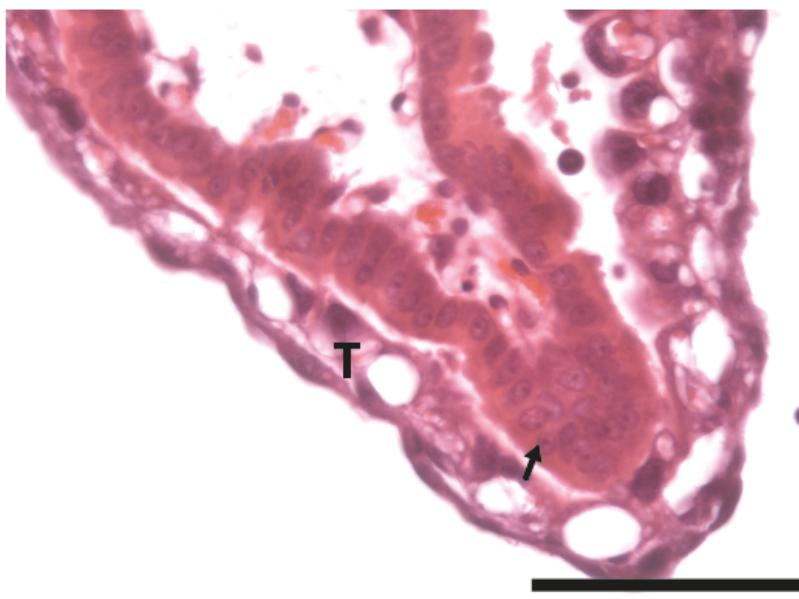


Day 12.5

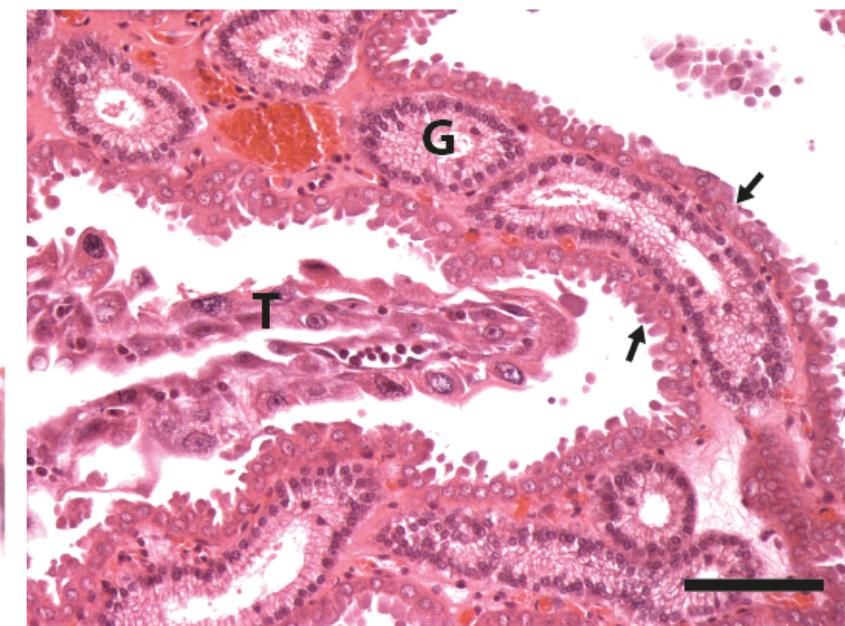
Eggshell is lost on day 12 of a 14 day pregnancy



Day 11.5

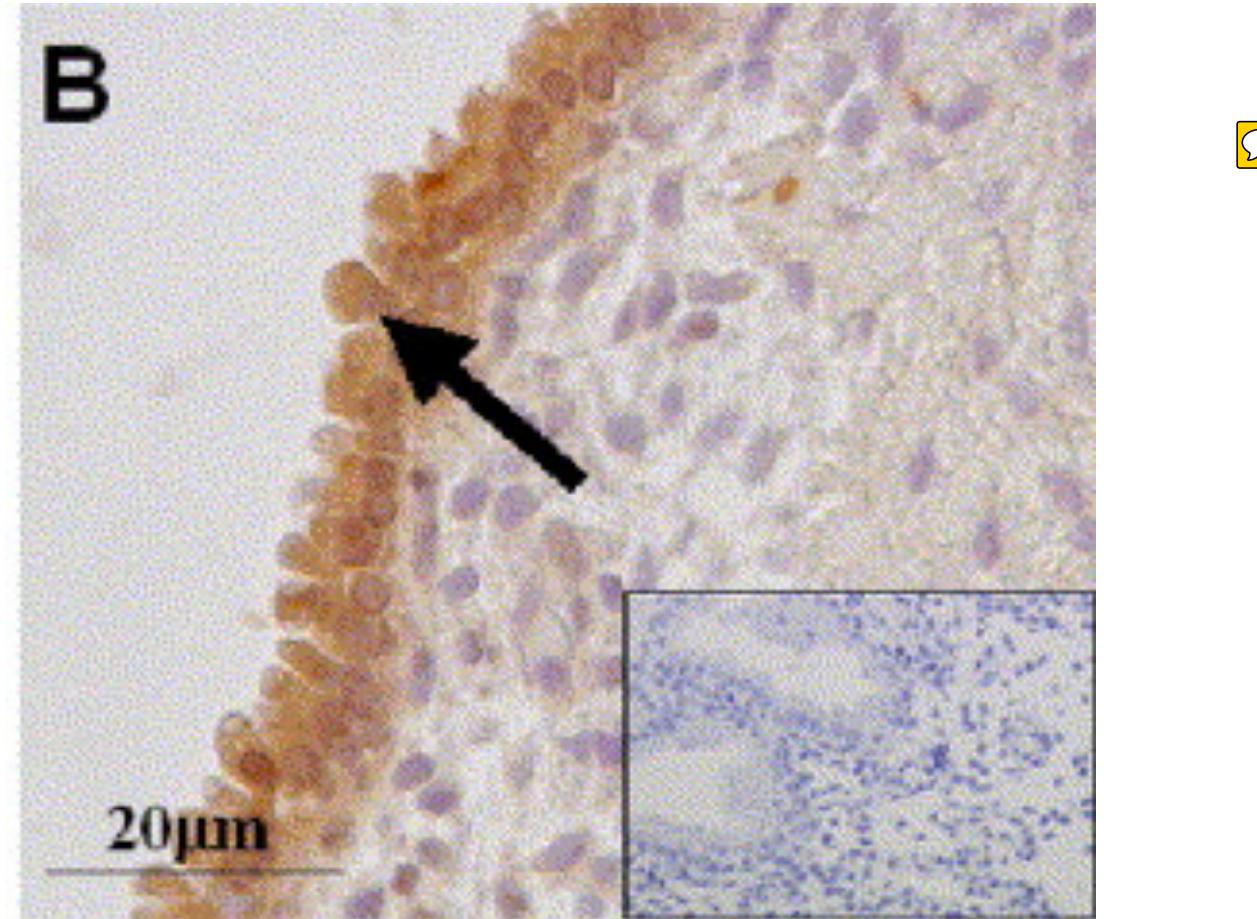


Day 12.5



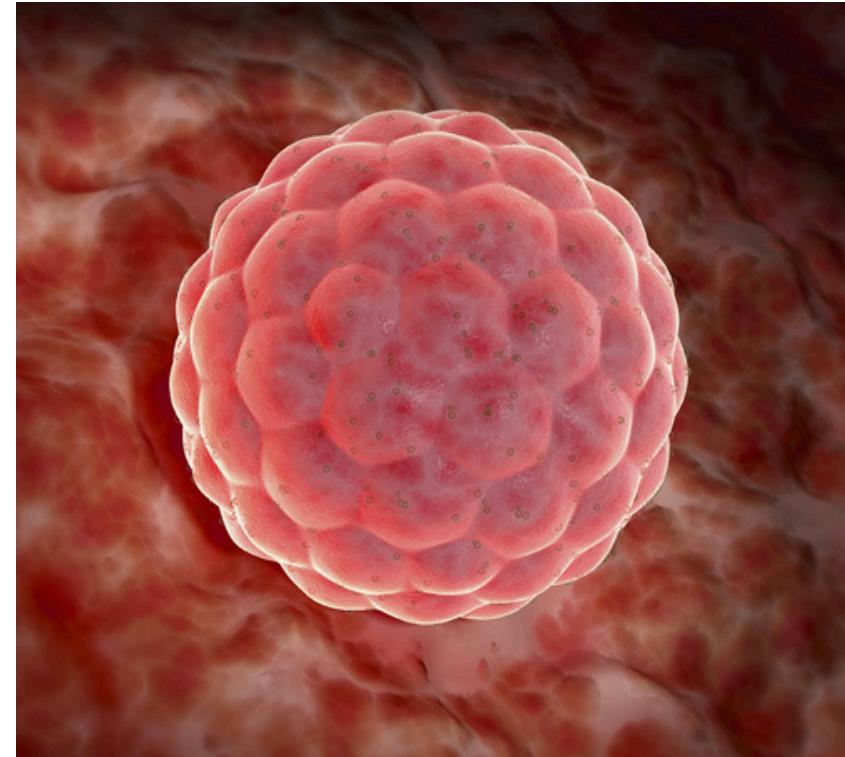
Day 13.5

Late gestation is morphologically similar to implantation in eutherian mammals



What is implantation?

- Apposition
- Attachment
- Invasion



Is term pregnancy in marsupials
similar to implantation at the
molecular level?

Grey short tailed opossum

- Found in south America
- Typical marsupial mode of pregnancy



Is term pregnancy in opossum
similar to implantation at the
molecular level?

- Inflammation
- Transcriptome wide similarities

Central Dogma of Biology

'Omics

DNA → Genomics

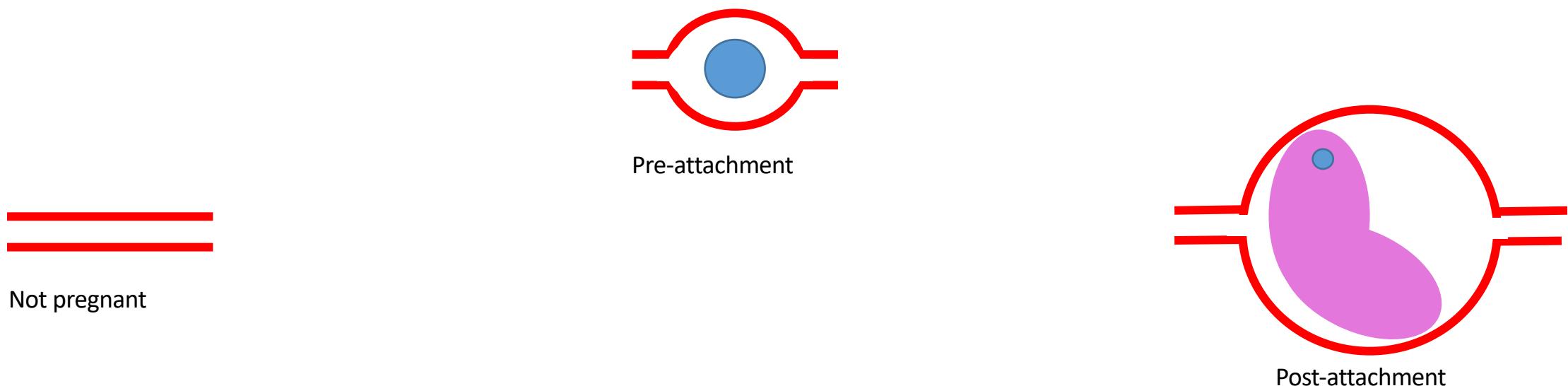
Transcription
↓

RNA → Transcriptomics

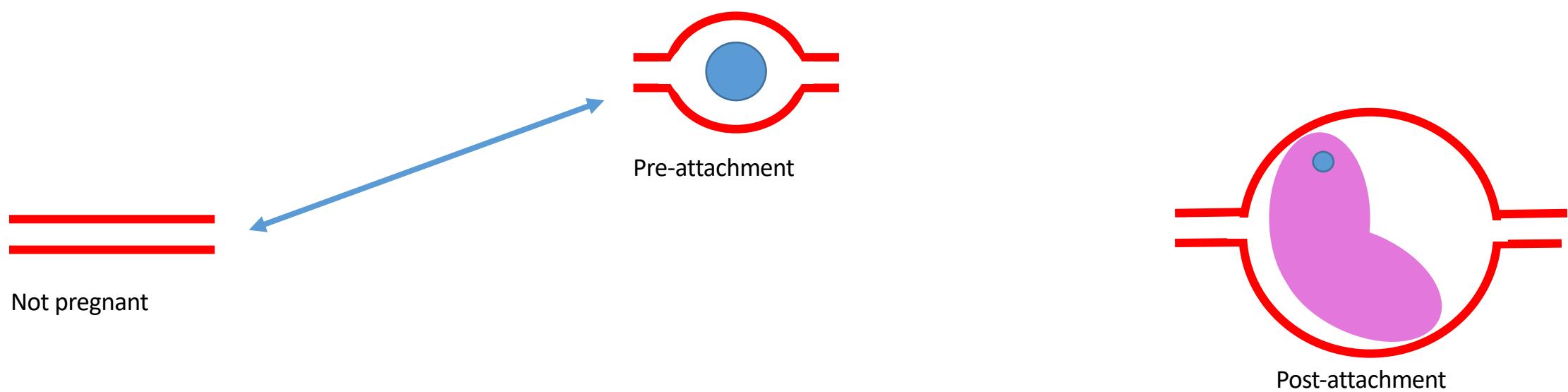
Translation
↓

Protein → Proteomics

Transcriptomics of the endometrium through the reproductive cycle



Non-pregnant vs pre-attachment

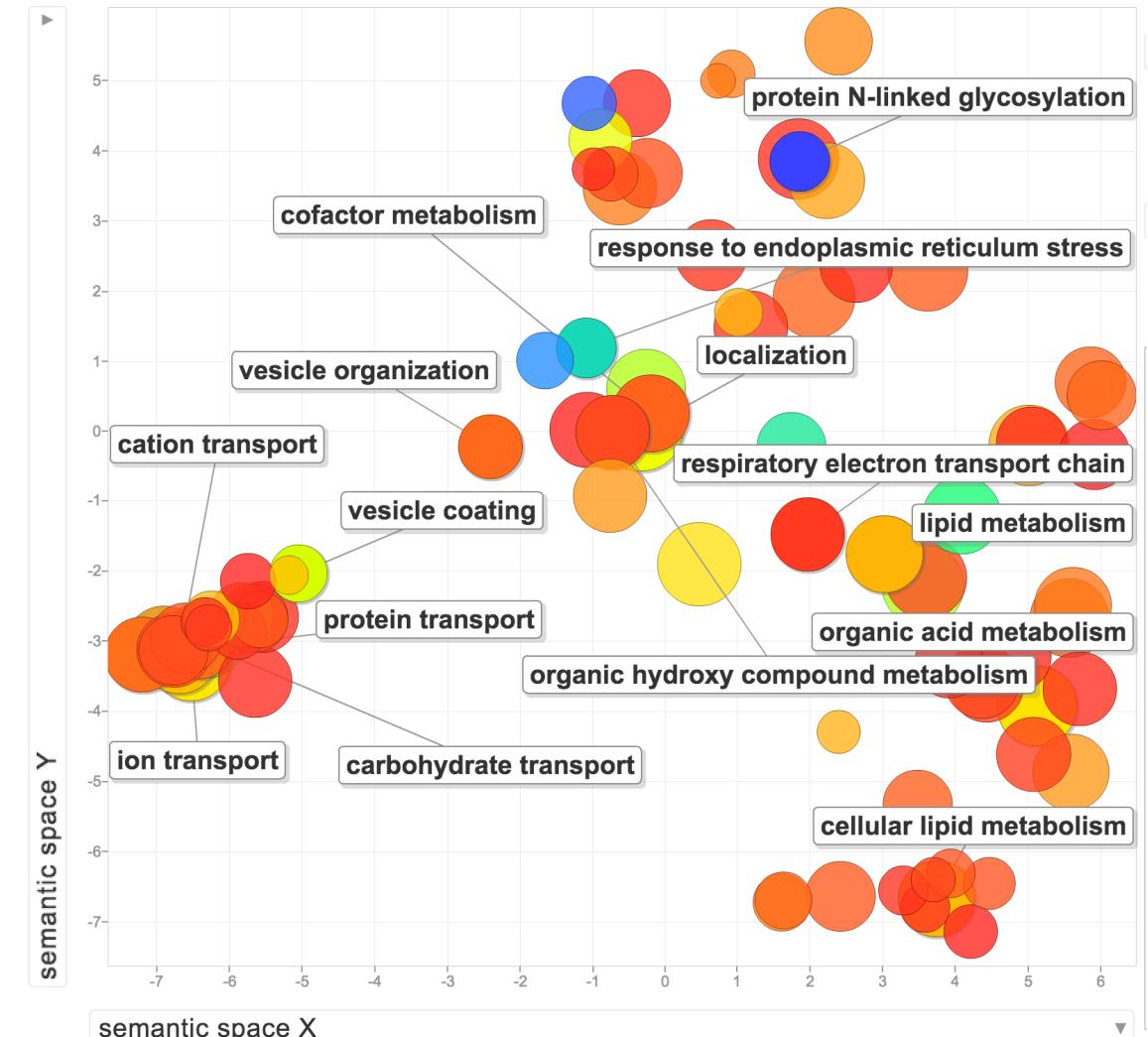


Non-pregnant Vs pre-attachment

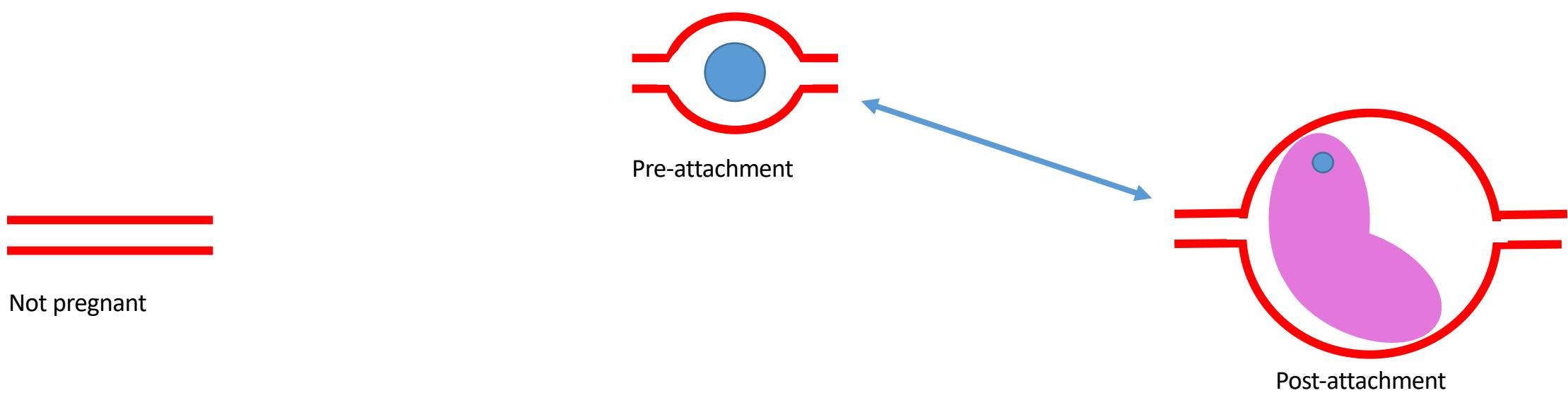
- 1358 Upregulated genes
- Gene ontology: 

 - Transport
 - Biosynthesis
 - Metabolism

- NO immune related terms
 - - inflammation
 - - immune
 - - cytokine

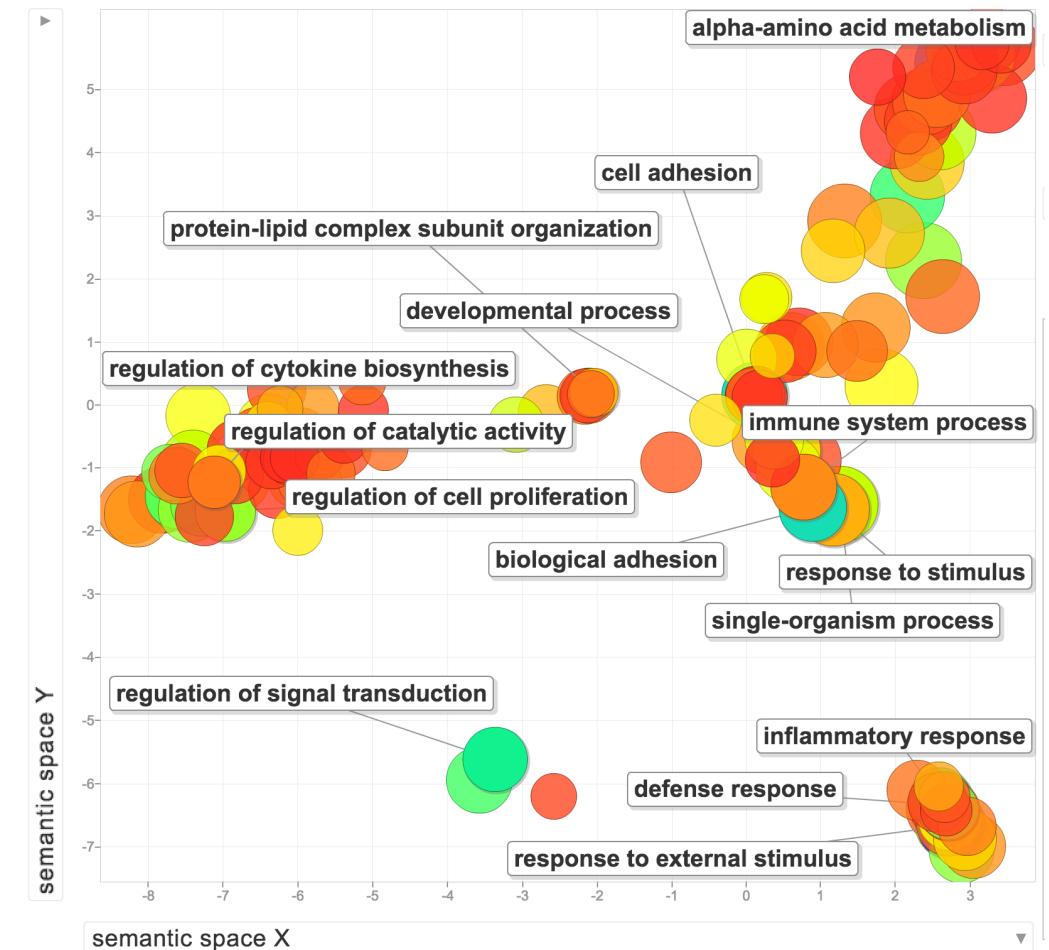


Pre-attachment vs post-attachment



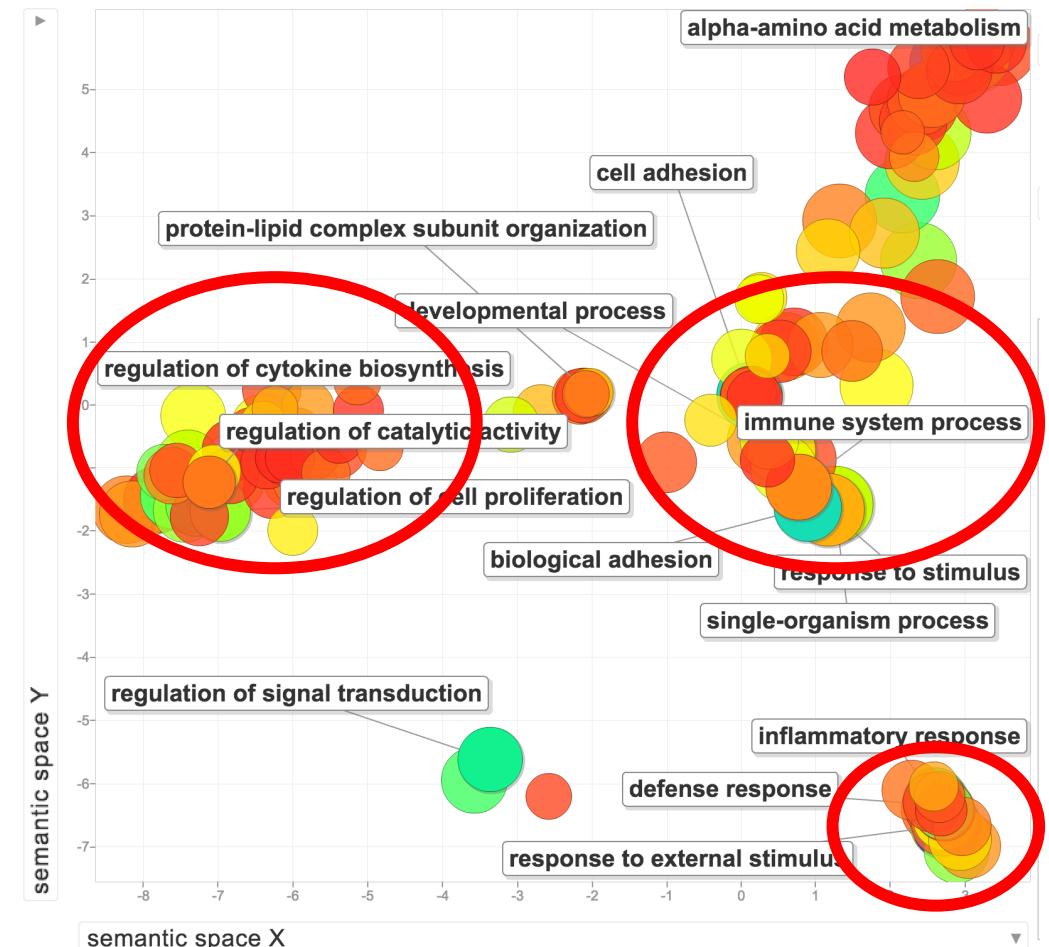
Pre-attachment Vs post-attachment

- 2056 genes higher in late gestation
- Gene ontology:
 - Transport
 - Biosynthesis
 - Metabolism
- Immune related terms (11 GO terms)
 - Immune system process (~250 genes)
 - Acute immune response
 - Inflammatory response



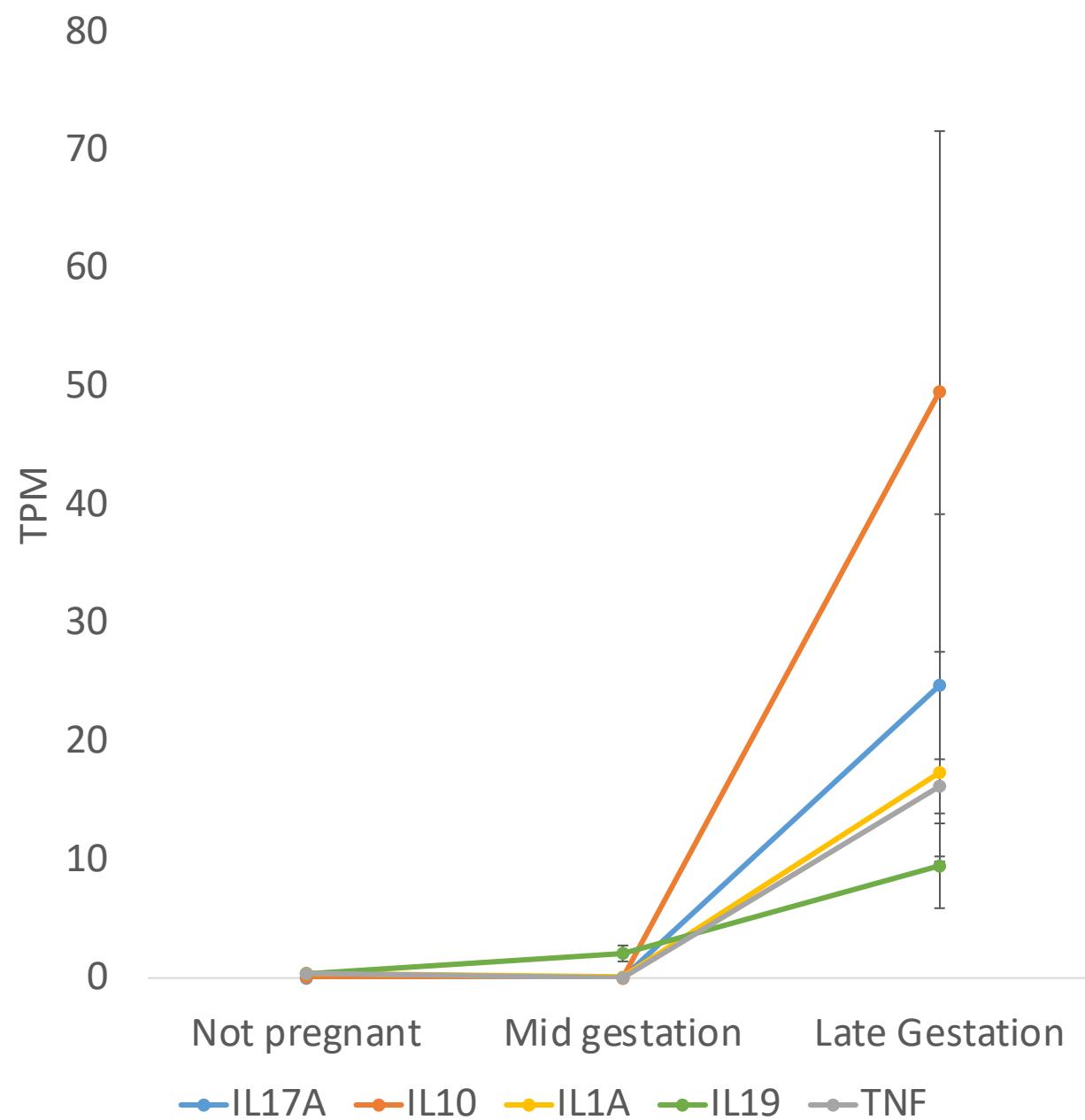
Pre-attachment Vs post-attachment

- 2056 genes higher in late gestation
- Gene ontology:
 - Transport
 - Biosynthesis
 - Metabolism
- Immune related terms (11 GO terms)
 - Immune system process (~250 genes)
 - Acute immune response
 - Inflammatory response

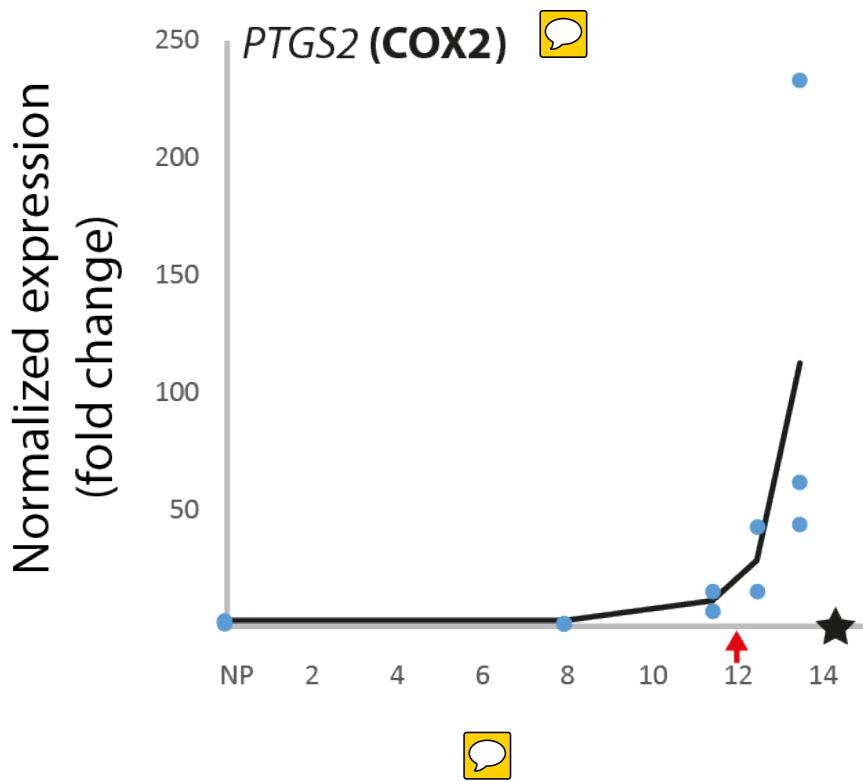


Up-regulation of cytokines

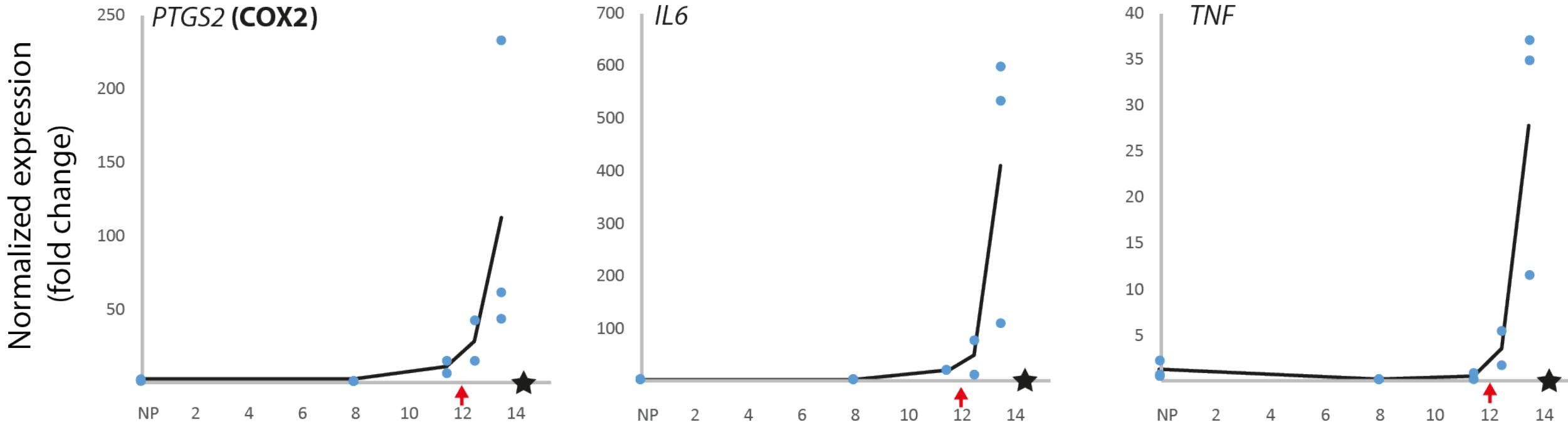
- Pro-inflammatory cytokines
 - IL17A 
 - IL1A
 - IL19
 - TNF
- Anti-inflammatory cytokine
 - IL10



How does the expression of inflammatory markers temporally correlate with the attachment reaction?



How does the expression of inflammatory markers temporally correlate with the attachment reaction?



COX2
Opossum
not pregnant

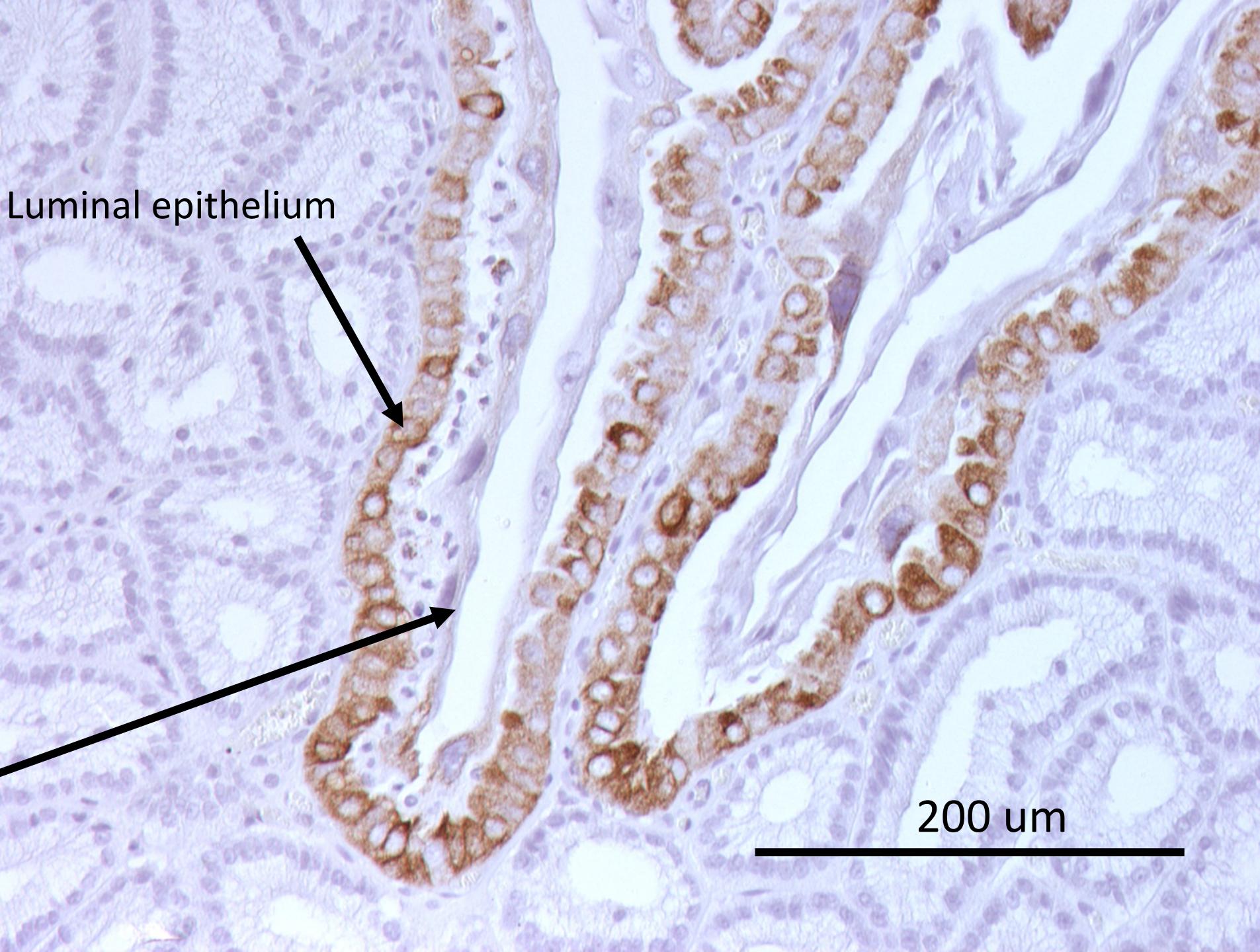


200 μ m

COX2 Opossum
post-
attachment

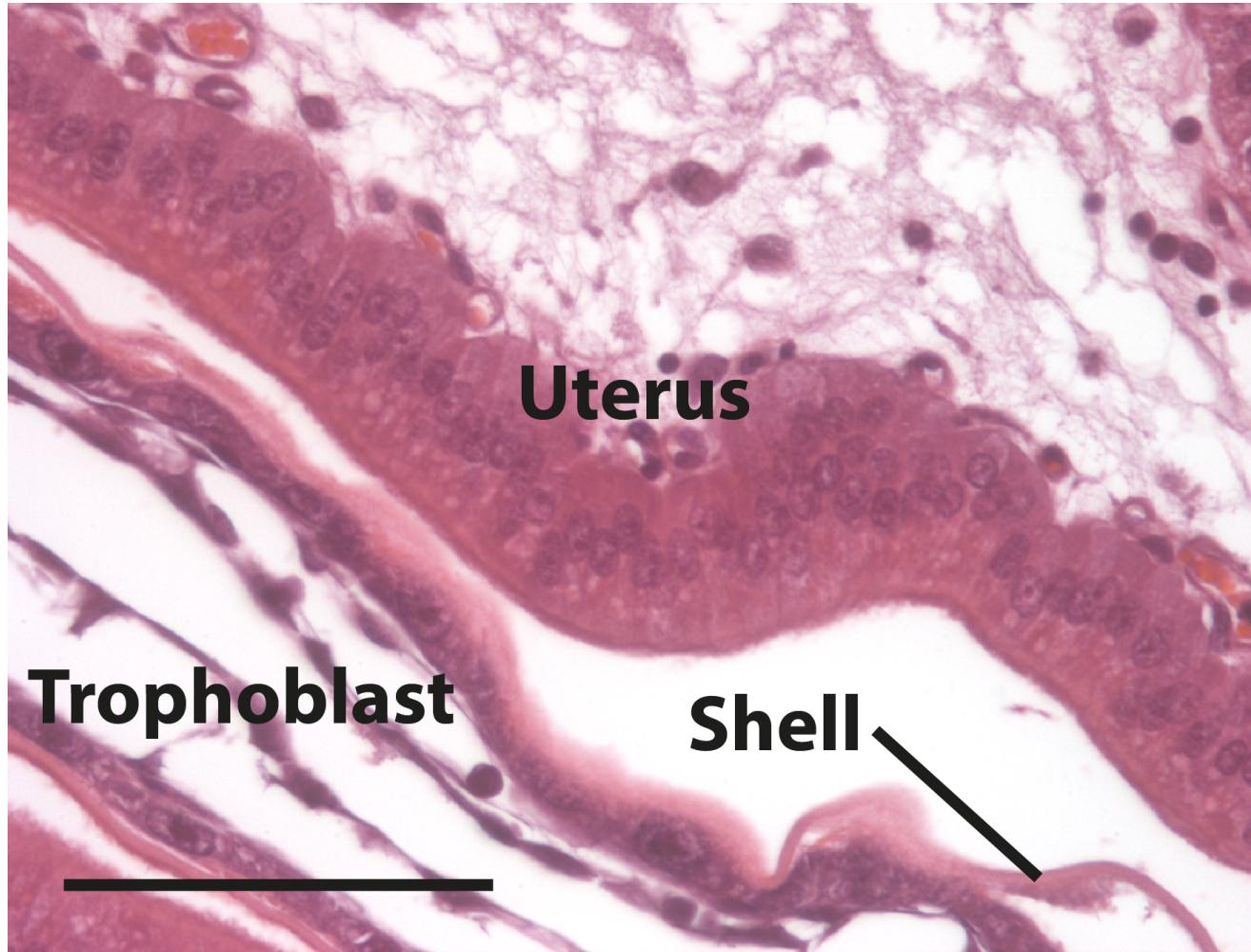


trophoblast



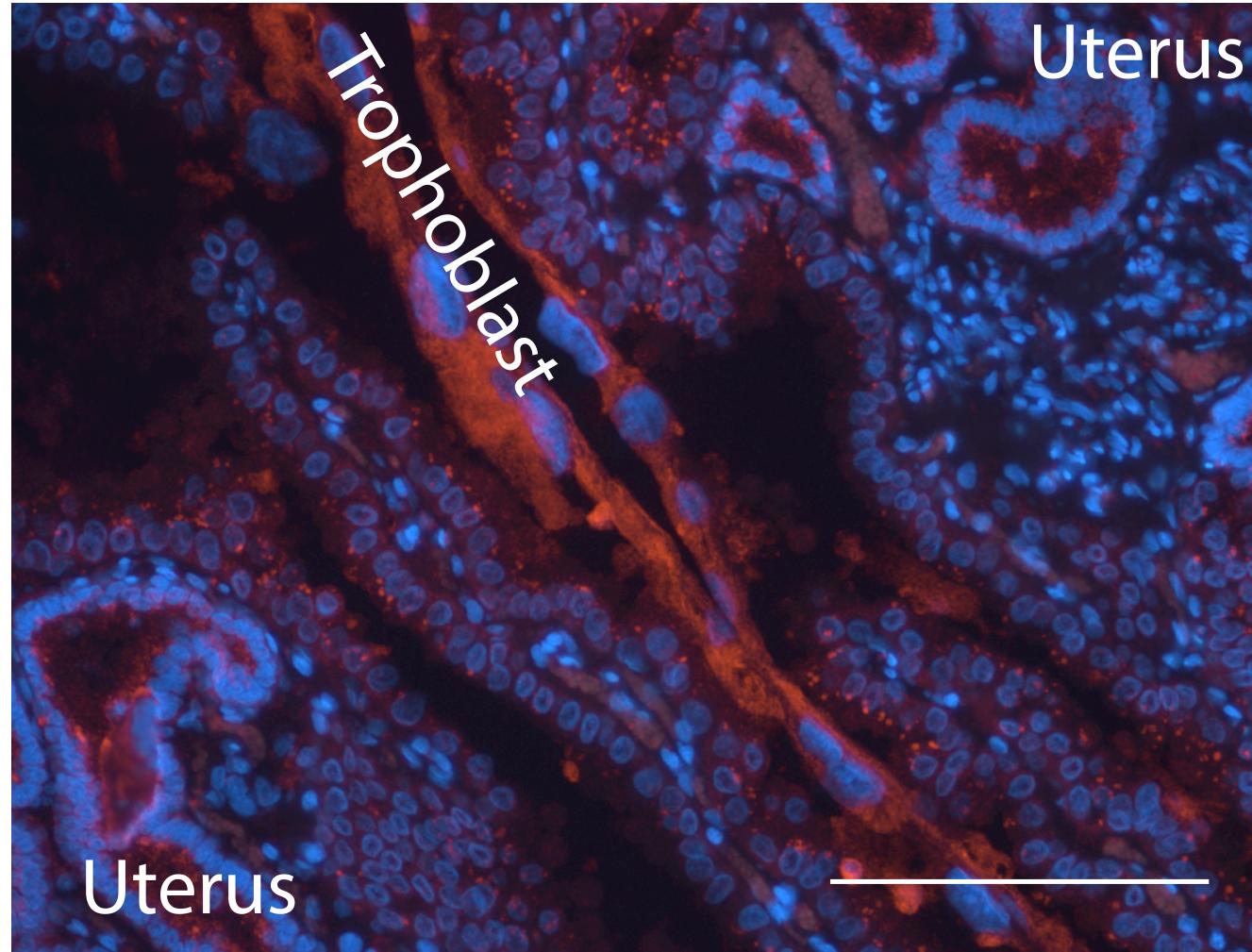
200 um

Shell coat is eroded by the embryo



Embryo produces proteases that degrade shell and then potentially irritate the uterine lining

Day 14.5



DAPI
PRSS8

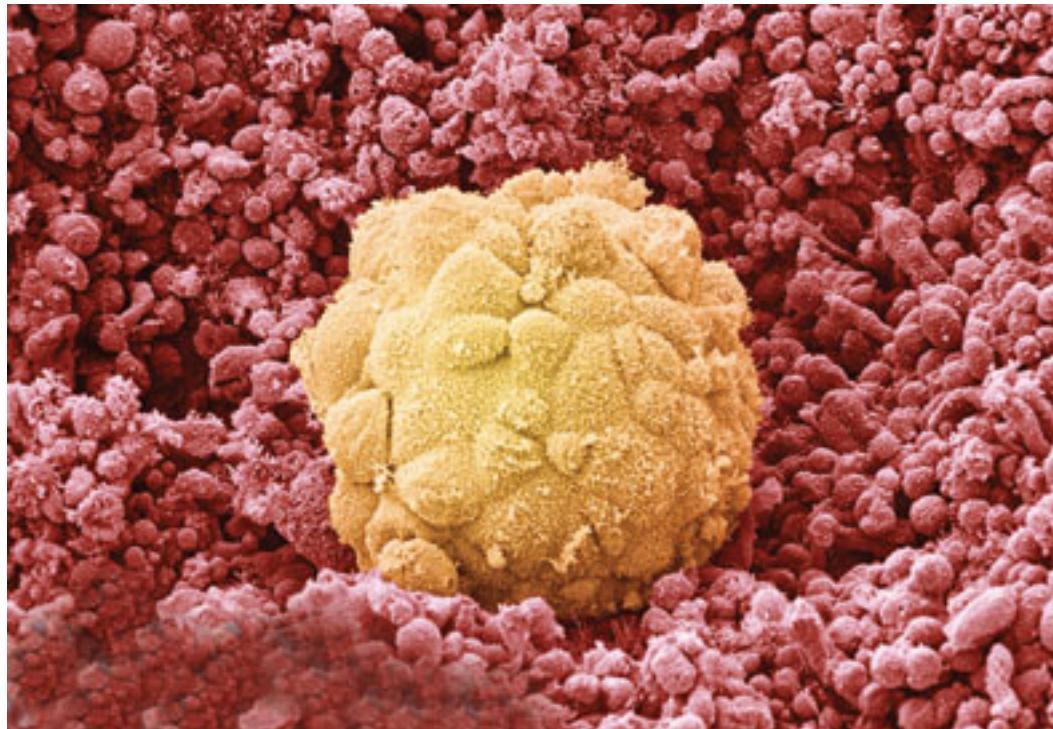


Inflammation is an important component of implantation in eutherians

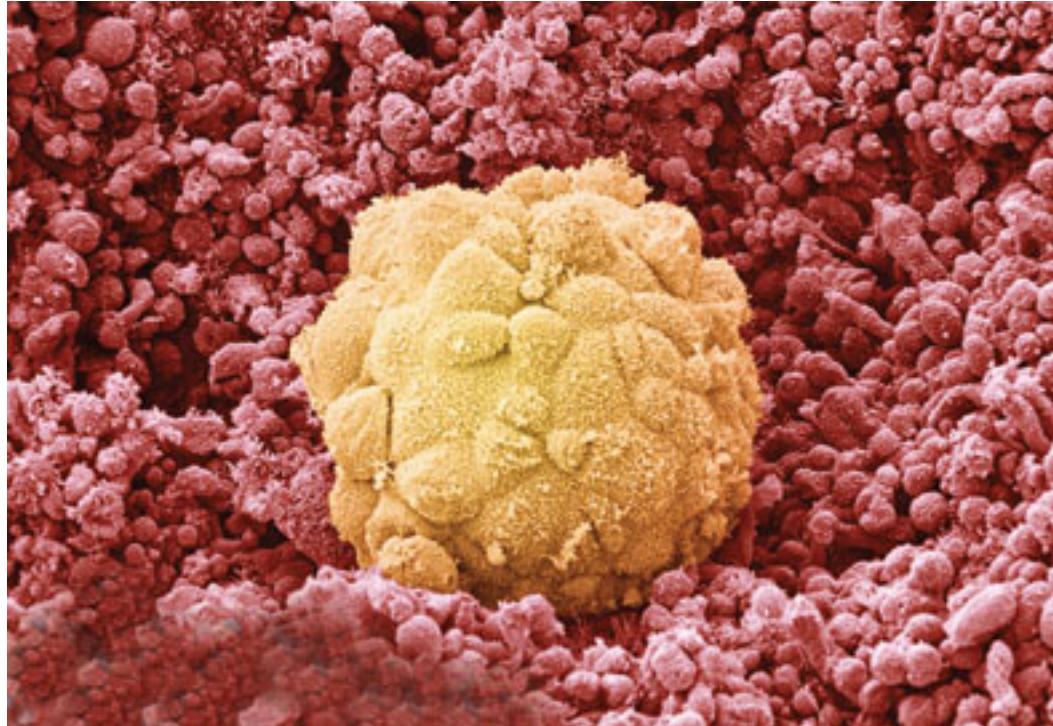
Inflammation is an important component of implantation in eutherians

In opossums inflammation is spatially and temporally correlated with the attachment reaction but precedes birth 

Transcriptome wide similarity with human implantation



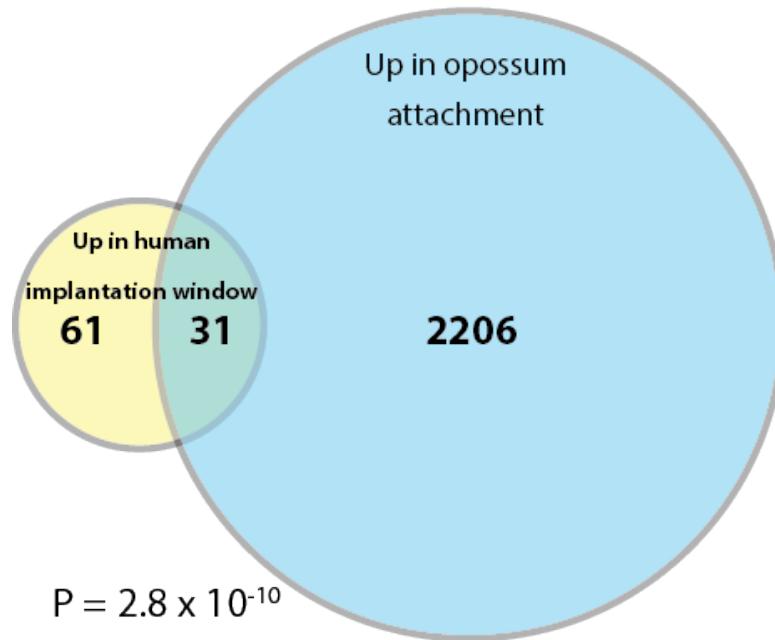
Transcriptome wide similarity with human implantation



Look at uterine biopsies either within or outside implantation window

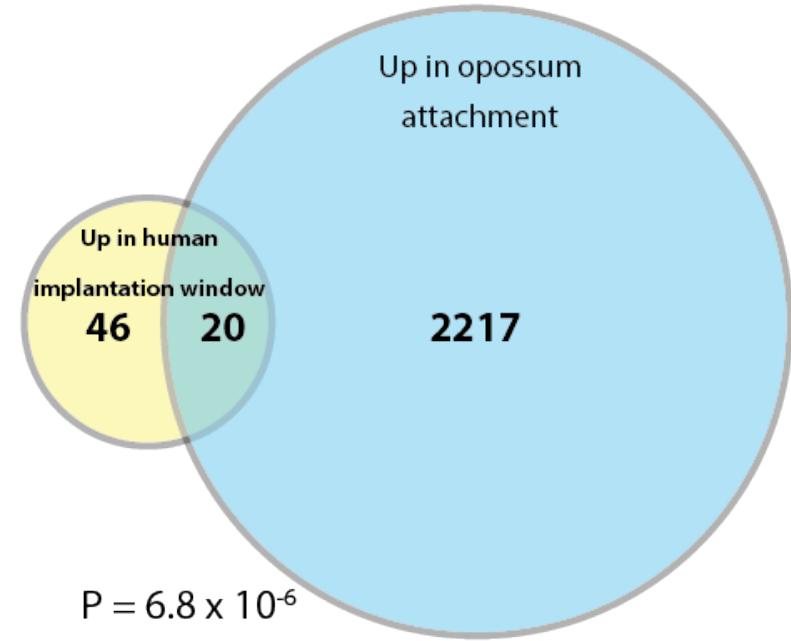
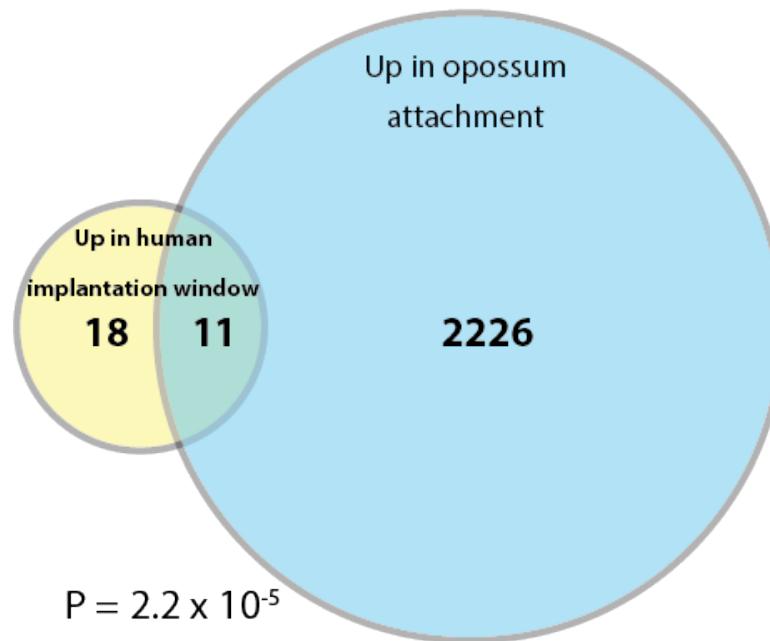
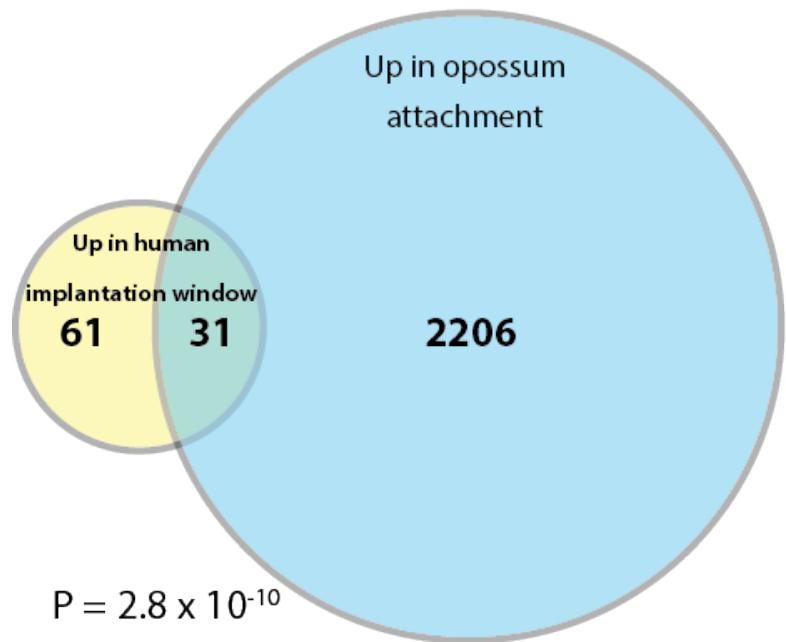
Transcriptome wide similarity with human implantation

Comparison with uterine biopsies either within or outside implantation window



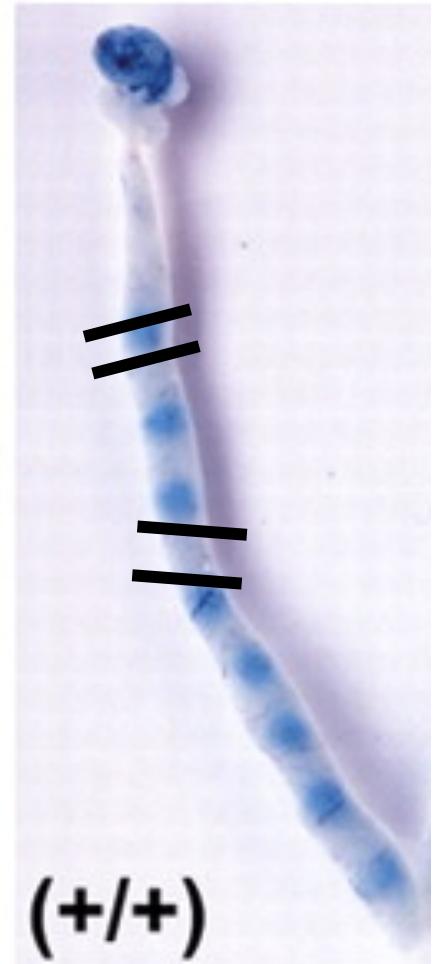
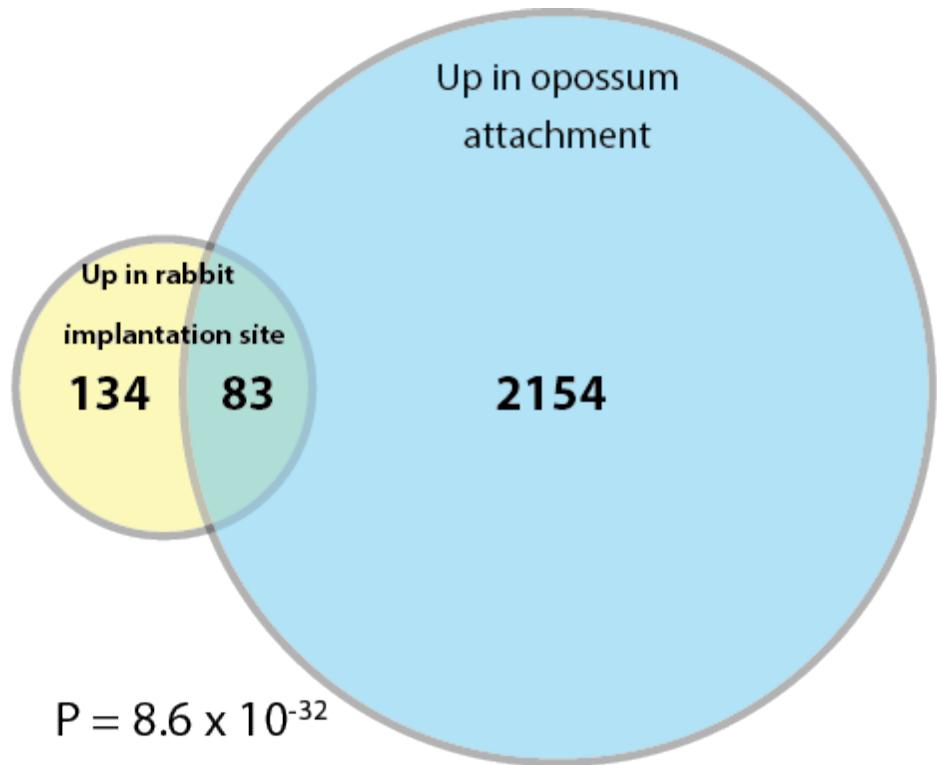
Transcriptome wide similarity with human implantation

Comparison with uterine biopsies either within or outside implantation window

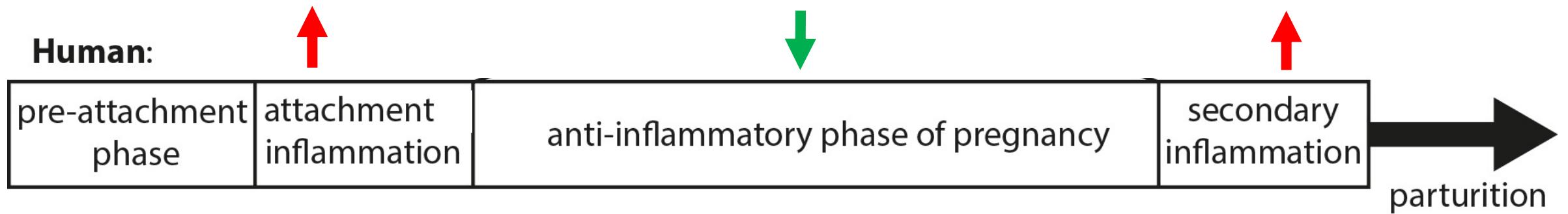


Transcriptome wide similarity with eutherian implantation

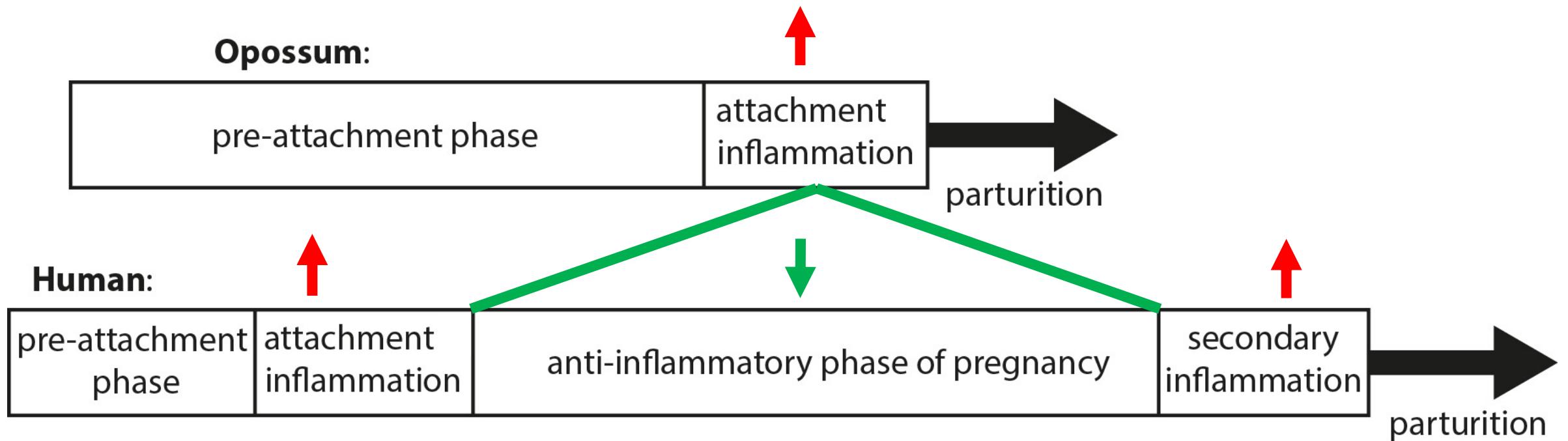
Rabbit vs. opossum comparison



Inflammation paradox



Inflammation paradox



Hypothesis:
Inflammation was an early mechanism for
recognising pregnancy

- 1) Inflammation is a direct consequence of exposure of the uterine epithelium to the yolk-sac membrane following breakdown of the eggshell barrier

Hypothesis:

Inflammation was an early mechanism for recognising pregnancy

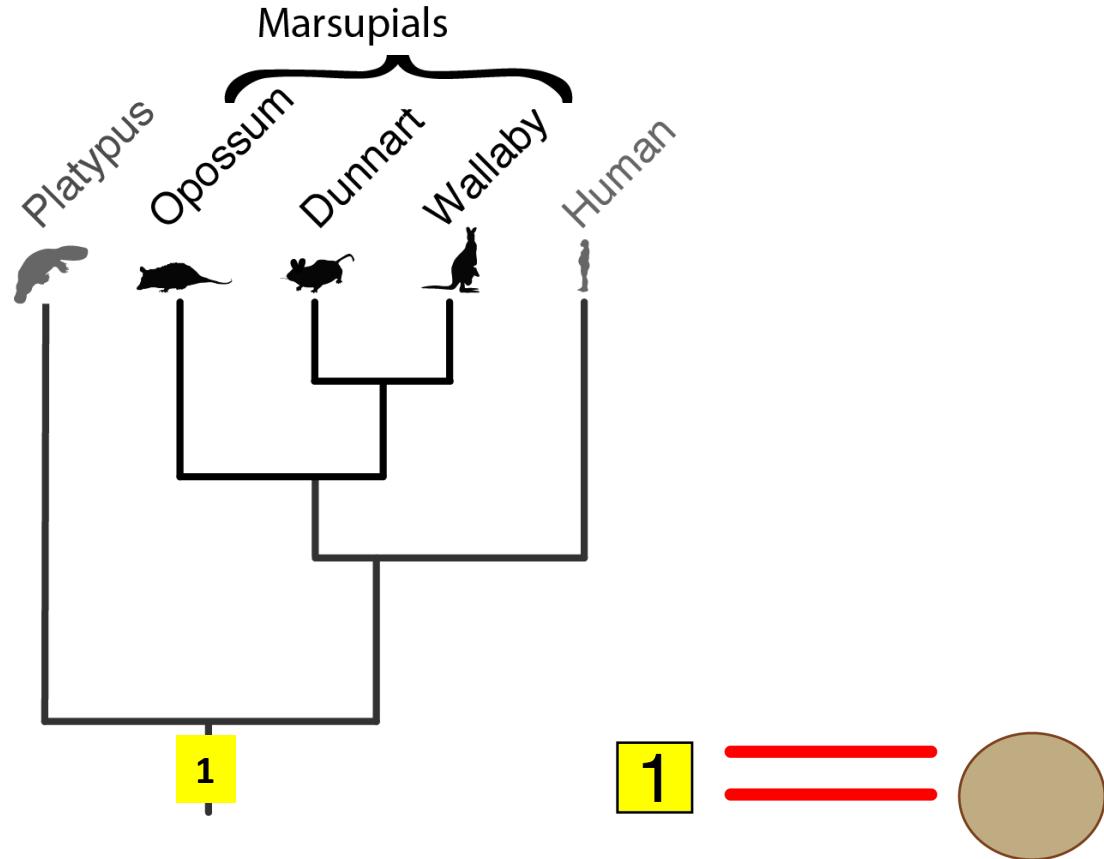
- 1) Inflammation is a direct consequence of exposure of the uterine epithelium to the yolk-sac membrane following breakdown of the eggshell barrier
- 2) Inflammation results in endometrial changes that are advantageous to the developing fetus
 - angiogenesis
 - vascular leakage
 - oedema

Hypothesis:

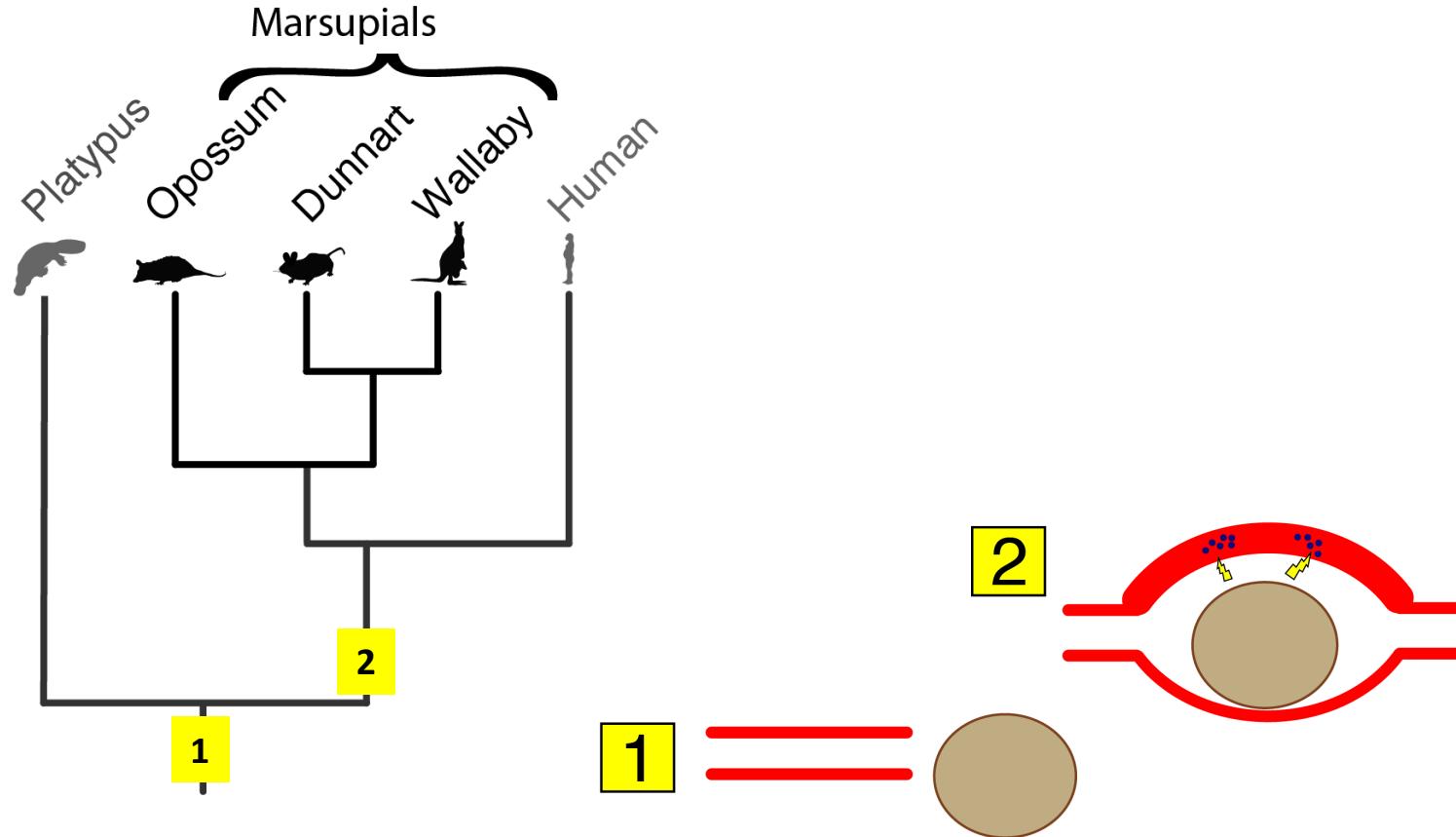
Inflammation was an early mechanism for recognising pregnancy

- 1) Inflammation is a direct consequence of exposure of the uterine epithelium to the yolk-sac membrane following breakdown of the eggshell barrier
- 2) Inflammation results in endometrial changes that are advantageous to the developing fetus
 - angiogenesis
 - vascular leakage
 - oedema
- 3) Inflammatory signalling may have further supported the parturition reaction which occurs shortly after the recognition of pregnancy in the opossum and likely the common ancestor of today's therian mammals

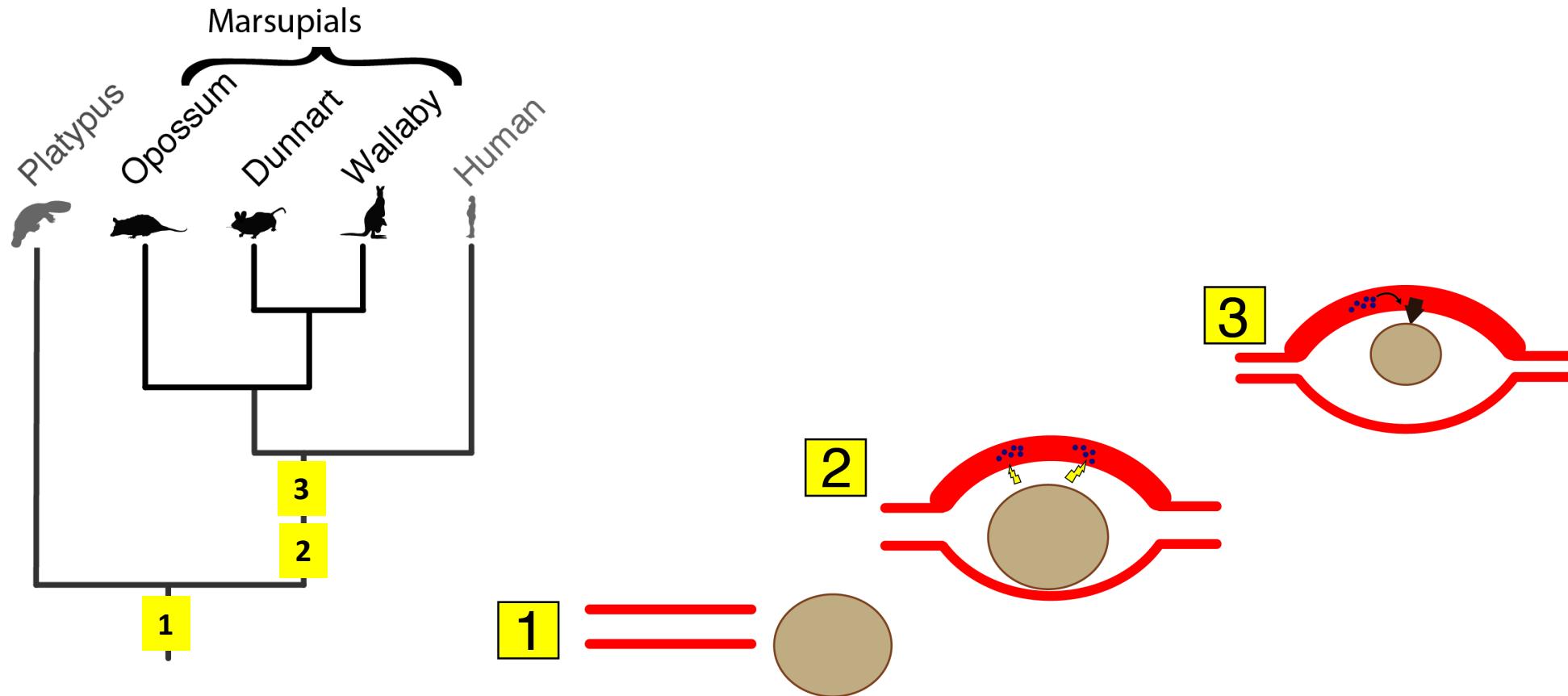
Inflammation as an early form of maternal fetal signalling in mammals



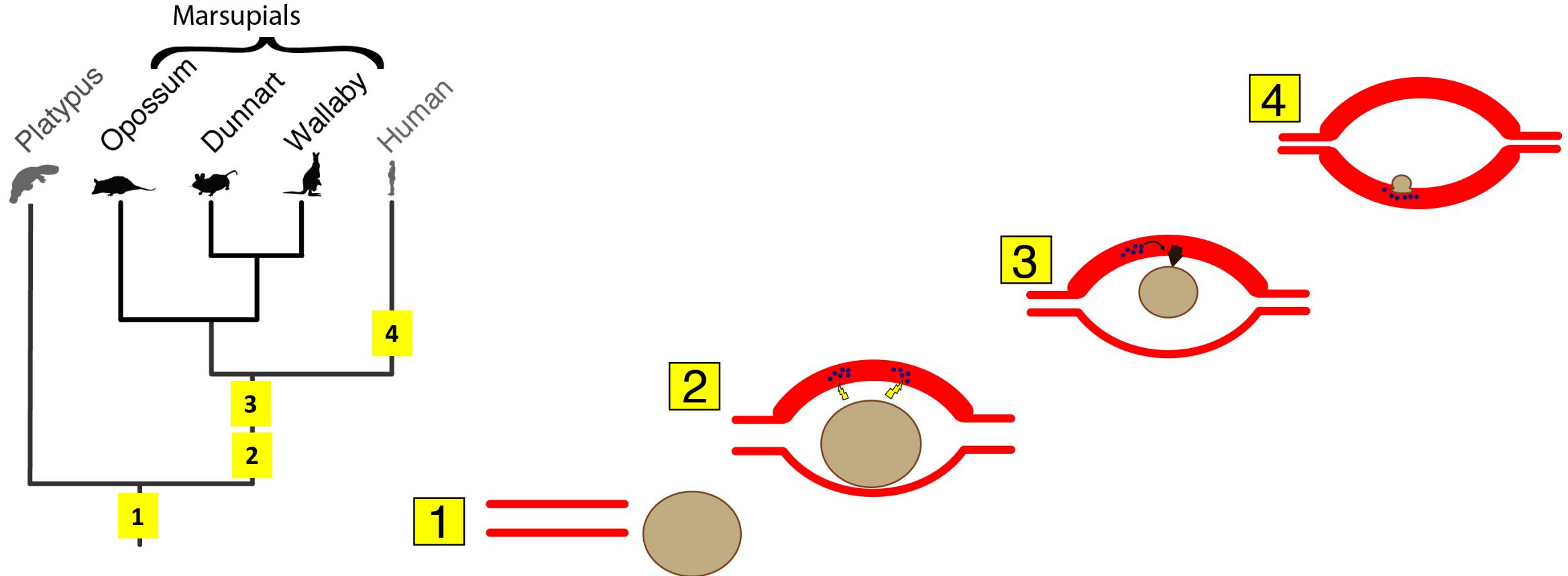
Inflammation as an early form of maternal fetal signalling in mammals



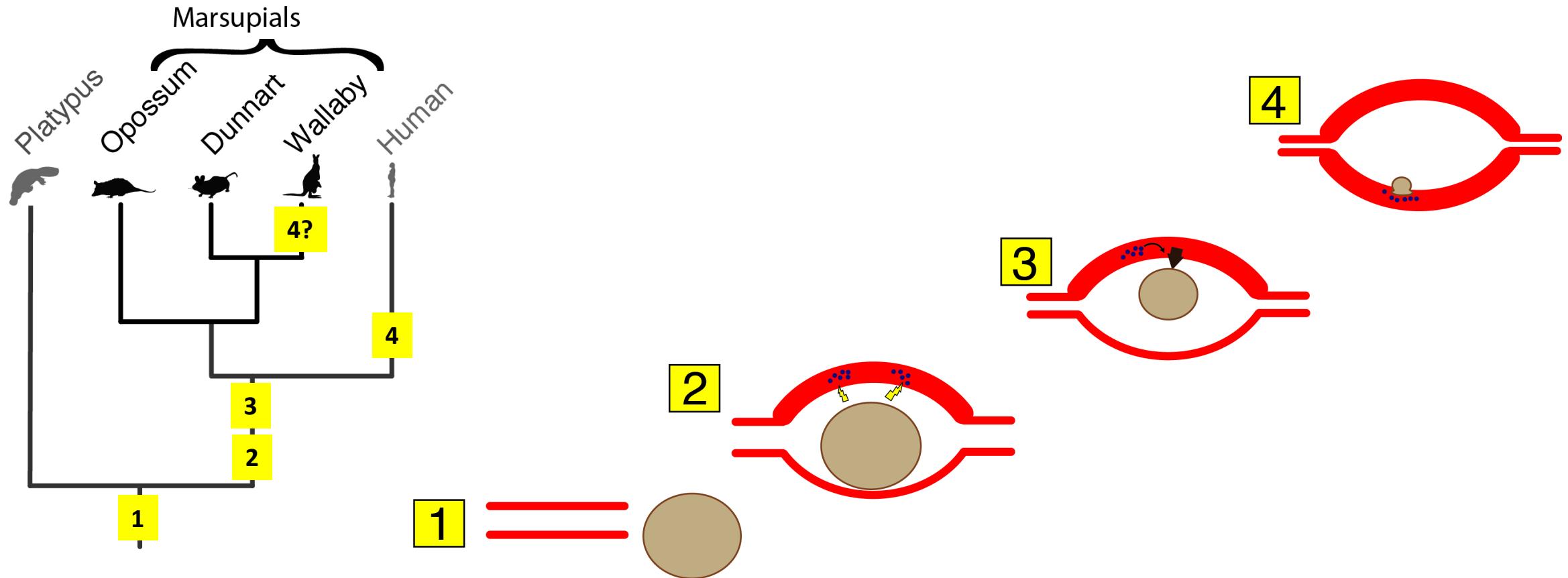
Inflammation as an early form of maternal fetal signalling in mammals



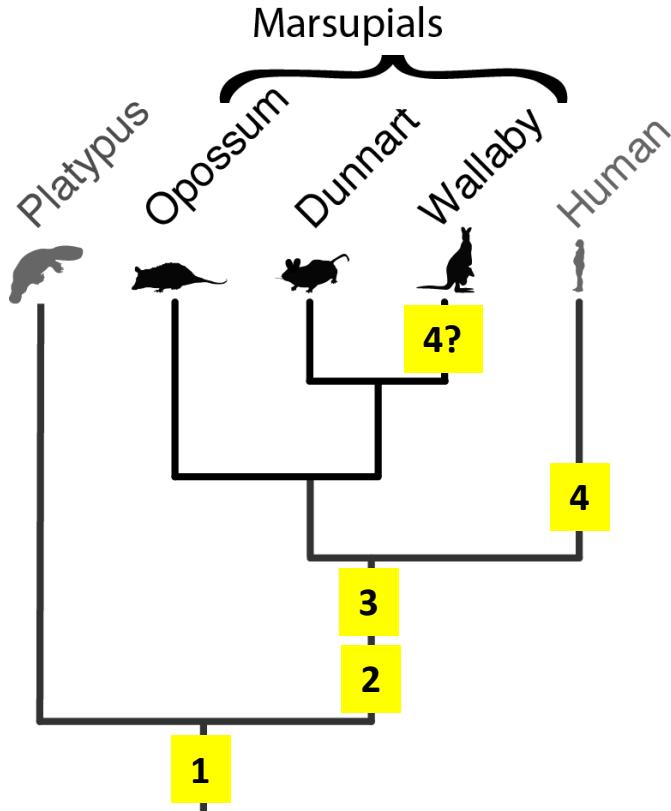
Inflammation as an early form of maternal fetal signalling in mammals



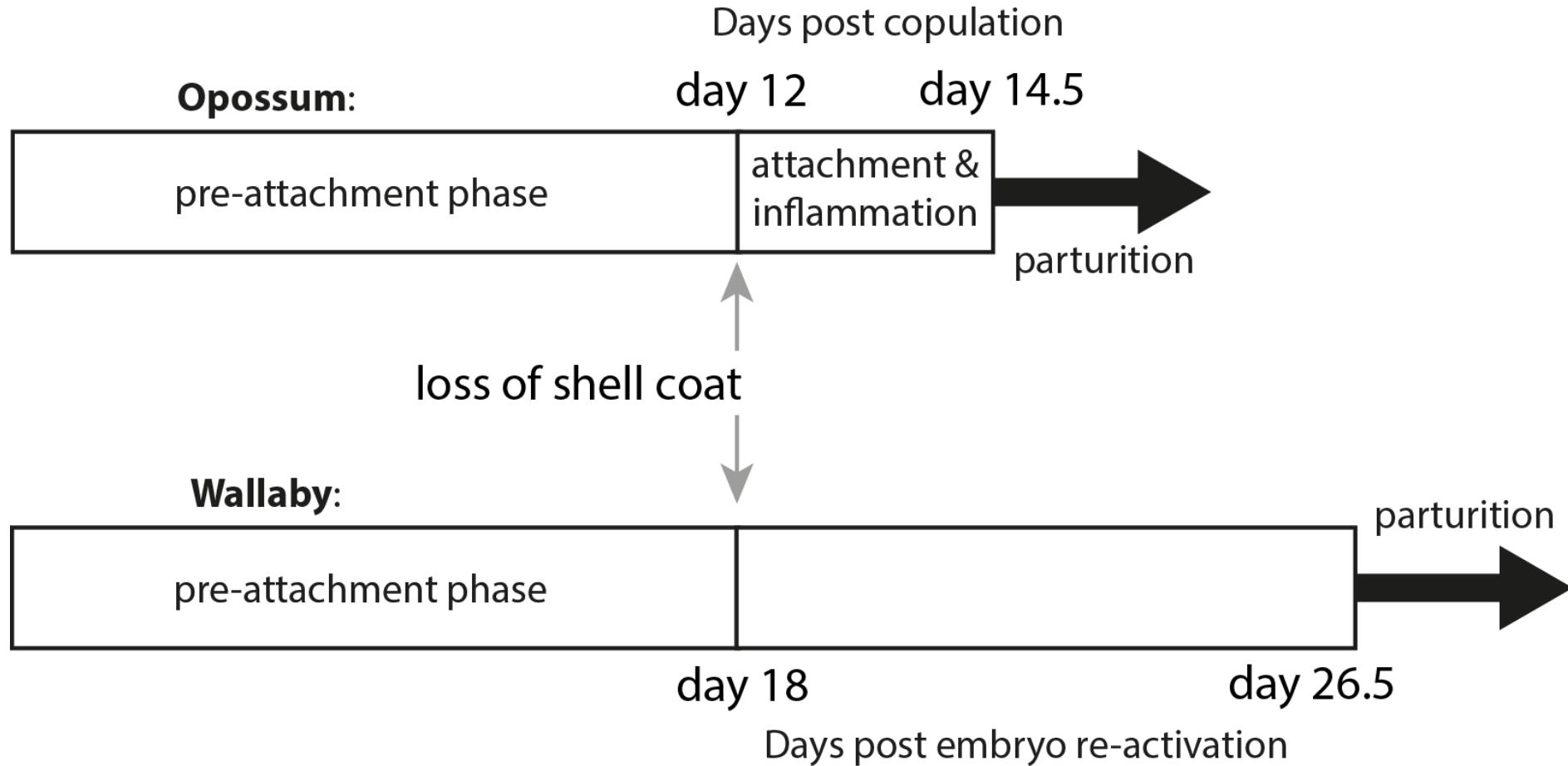
Are eutherians the only mammals with an extended period of placentation?



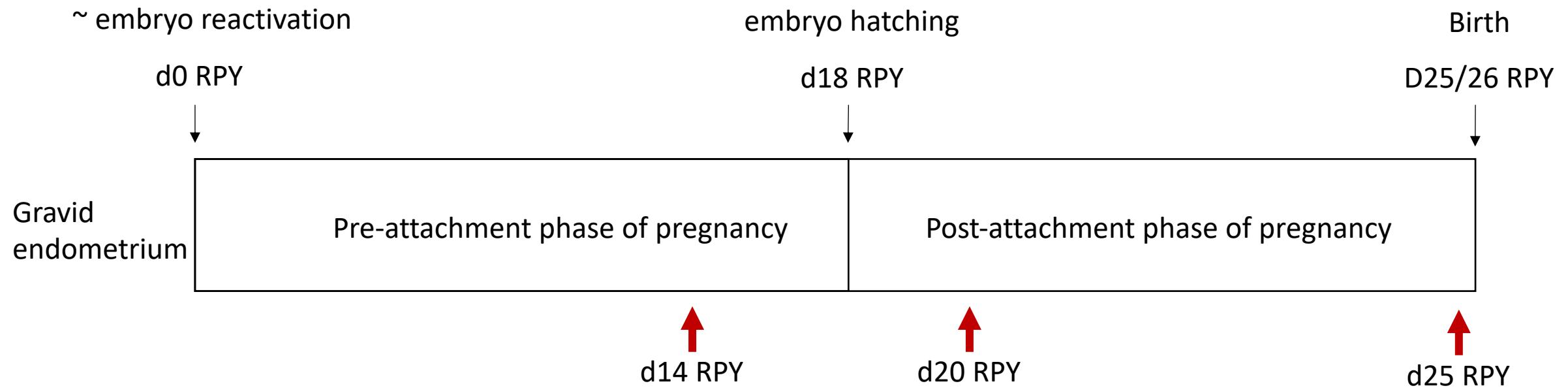
Are eutherians the only mammals with an extended period of placentation?



How do wallabies contribute to understanding implantation?

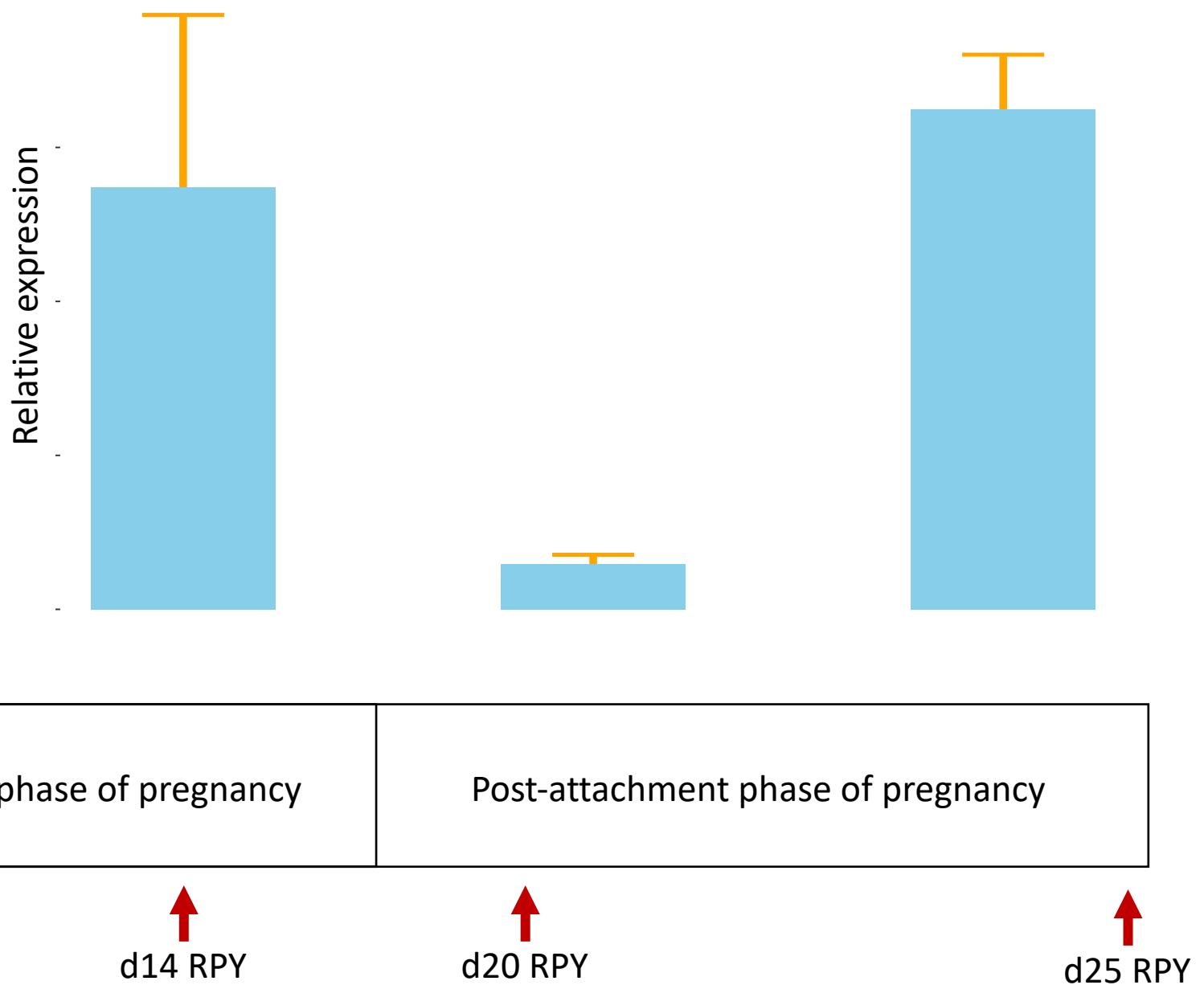
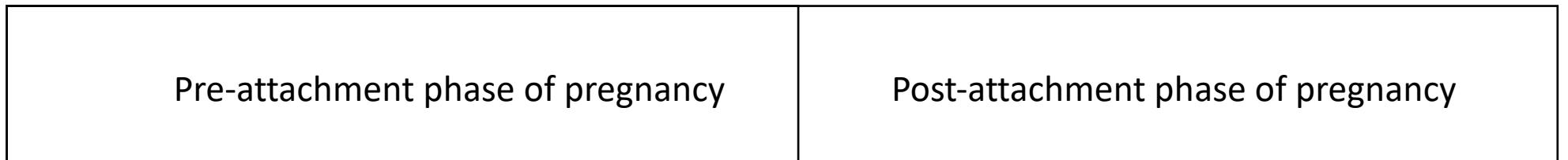


Molecular changes through pregnancy in the wallaby

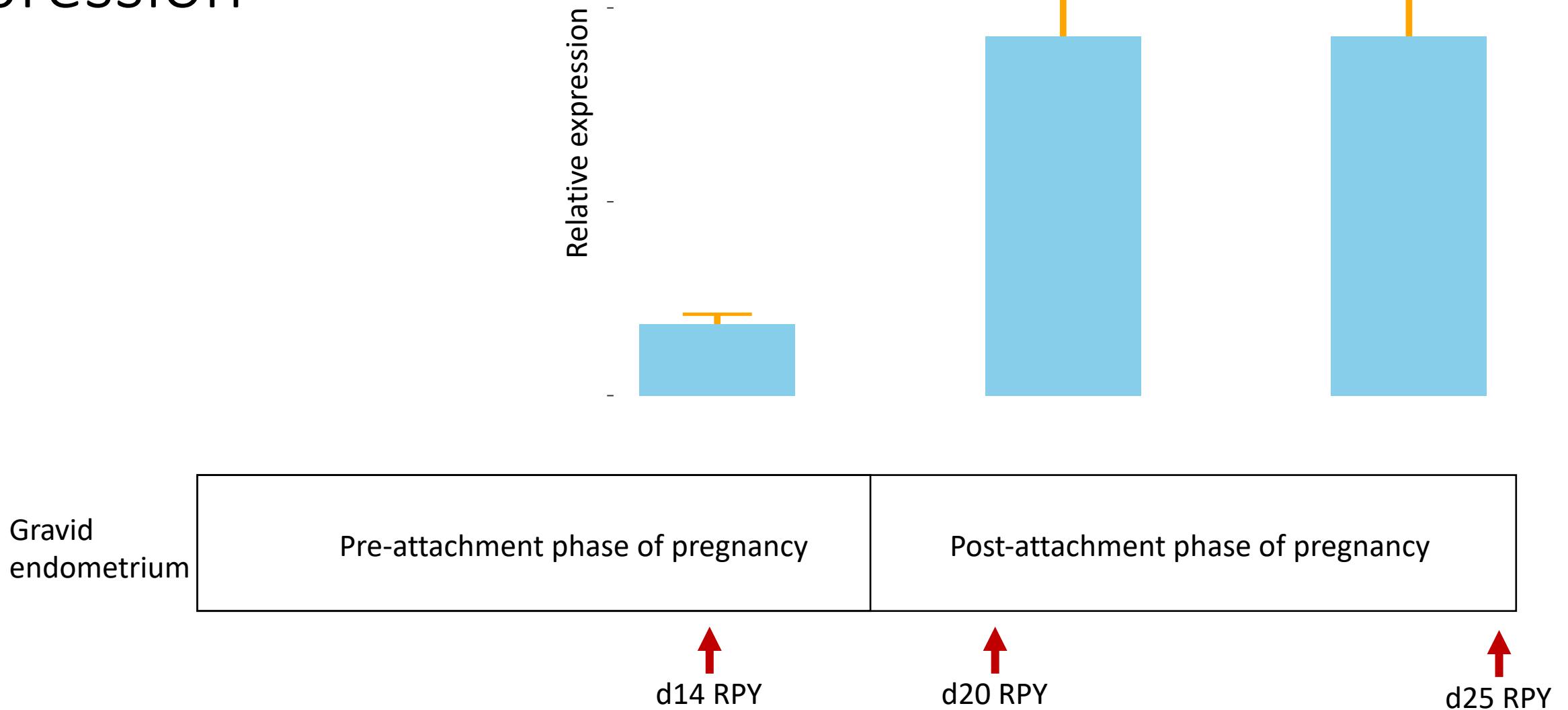


Pattern of inflammatory gene expression

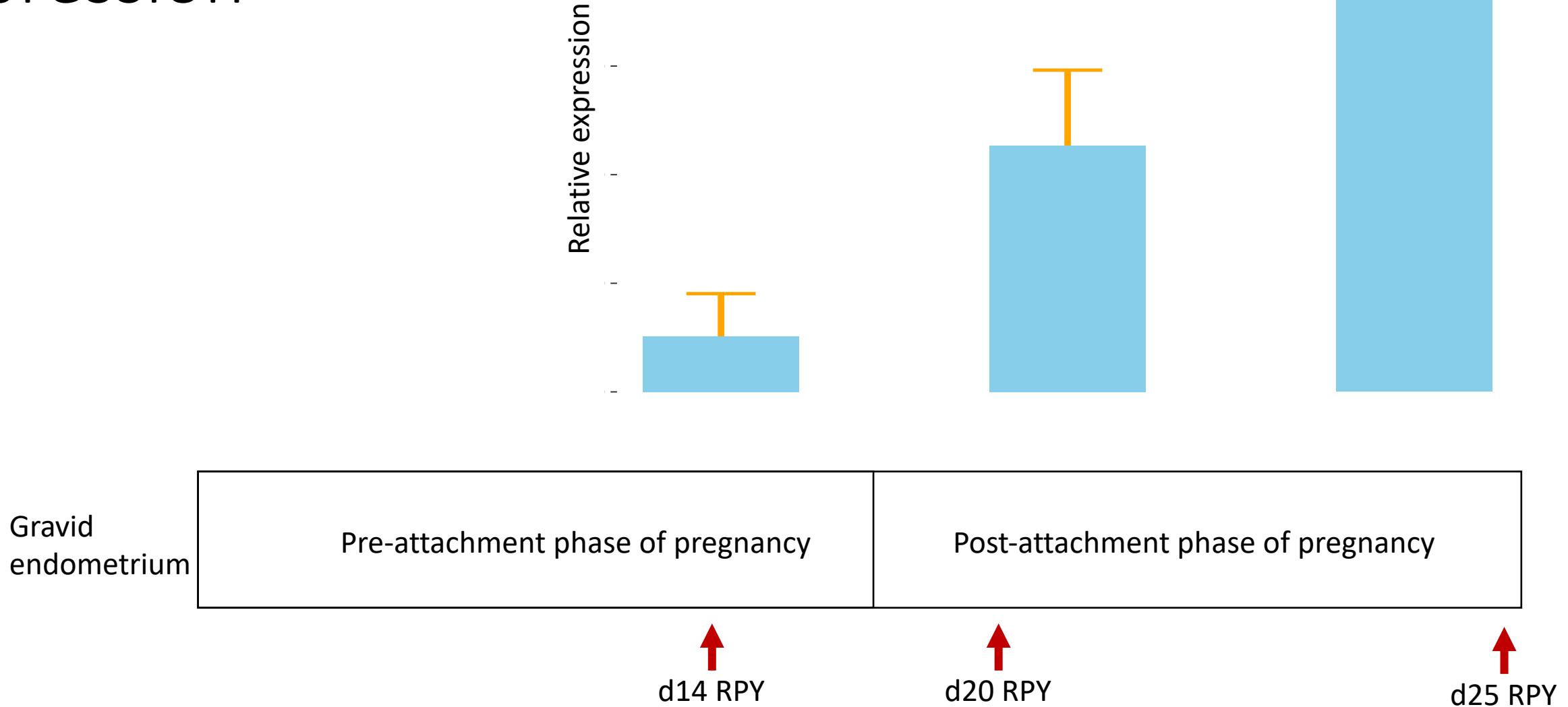
Gravid endometrium



Pattern of inflammatory gene expression



Pattern of inflammatory gene expression



A close-up photograph of a wallaby, showing its brown fur and large ears. A small joey is visible nestled in its pouch. The background is blurred greenery.

Wallaby pregnancy uses inflammatory signalling in a dynamic way through placentation

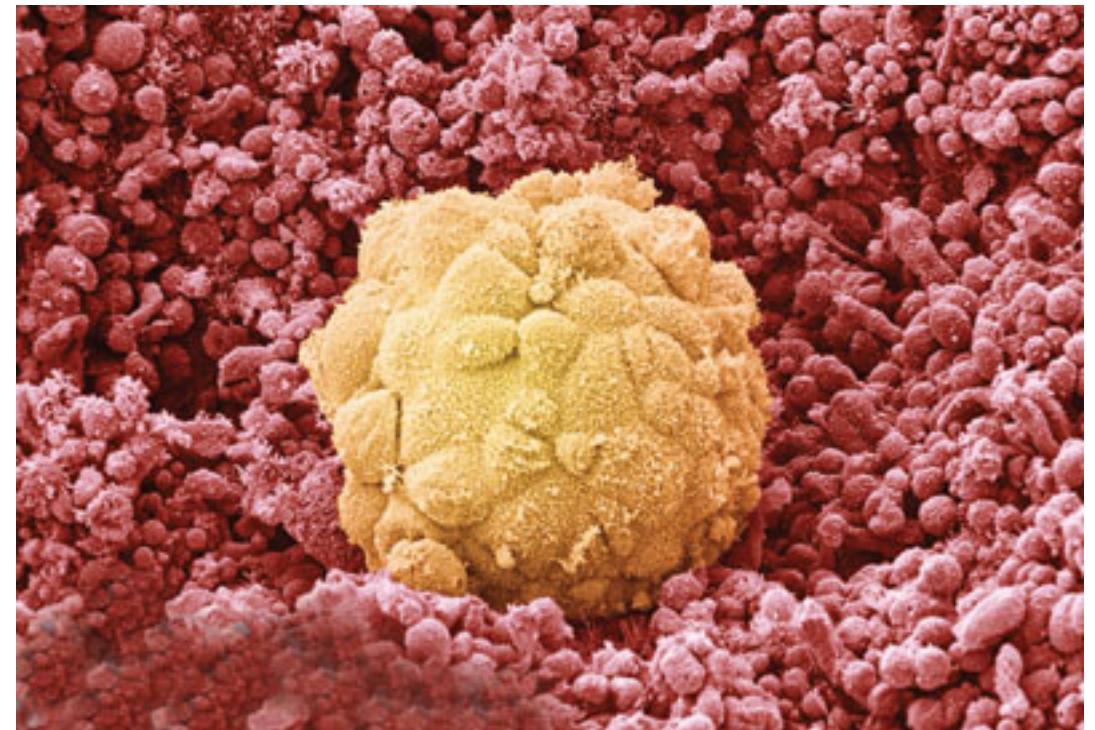
Consistent with a derived state of maternal-fetal interaction

Hopefully I have convinced you

- Inflammation at implantation is a modified maternal response to the presence of the embryo
- The extension of pregnancy has involved the modification of this inflammatory pathway perhaps to support derived aspects of pregnancy
- We can use this evolutionary framework to identify key aspects of implantation biology

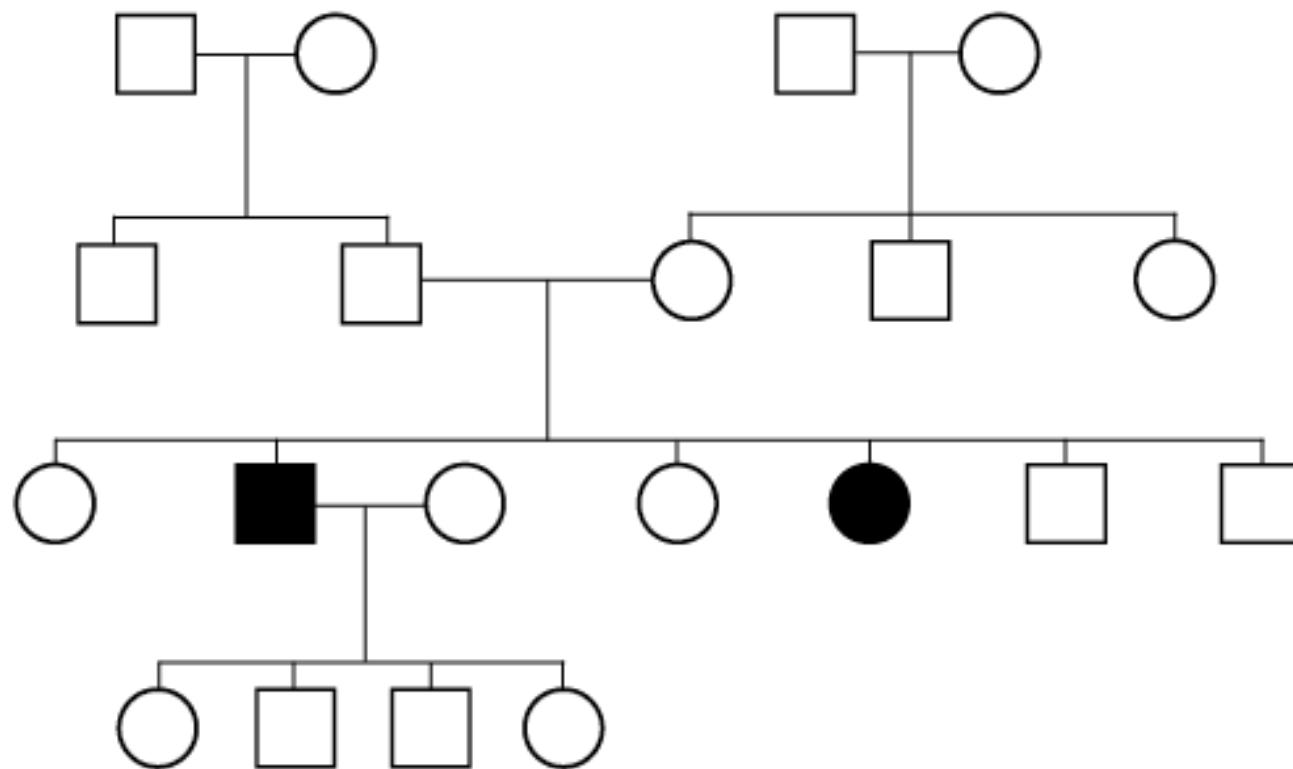
Why is an evolutionary perspective important?

- Narratives drive research agendas
- This explanation re-frames our understanding of inflammation
 - Shifting from “dangerous due to paternal genetic material” to a normal outcome of maternal-fetal interaction



BIOL3120

Lecture 4: Modes of Inheritance and Population genetics



Some terms/ideas to recall

- **Gene** = many different meanings depending on context.
 - In genetics we might talk about the gene for a specific trait. i.e. The gene associated with red hair.
 - In genomics we talk about genes as being a part of the genome associated with a specific transcribed unit of genetic information
 - In evolutionary biology we talk about genes as a unit of inheritance
- **Locus** = specific, fixed position on a chromosome where a particular gene or genetic marker is located
- **Allele** = a variant of a particular gene locus

Some terms/ideas to recall

- **Dominant** = Having one copy of particular allele will show this condition/trait
- **Recessive** = Need two copies of recessive alleles for condition/trait to show
- **Genotype** = the alleles (versions of gene) you have at a particular gene or locus (genetic location)
 - Often shown with upper and lower case letters
 - **A** = dominant allele
 - **a** = recessive allele
 - Humans have two copies of most genes: **AA** or **Aa** or **aa**
- **Homozygote** = having two copies of the same allele at a particular gene/locus
- **Heterozygote** = having two different alleles at a particular gene/locus
 - **AA** is homozygous dominant genotype
 - **aa** is homozygous recessive genotype
 - **Aa** is heterozygote genotype
 - Homo = same, hetero = different, zygote = things coming together

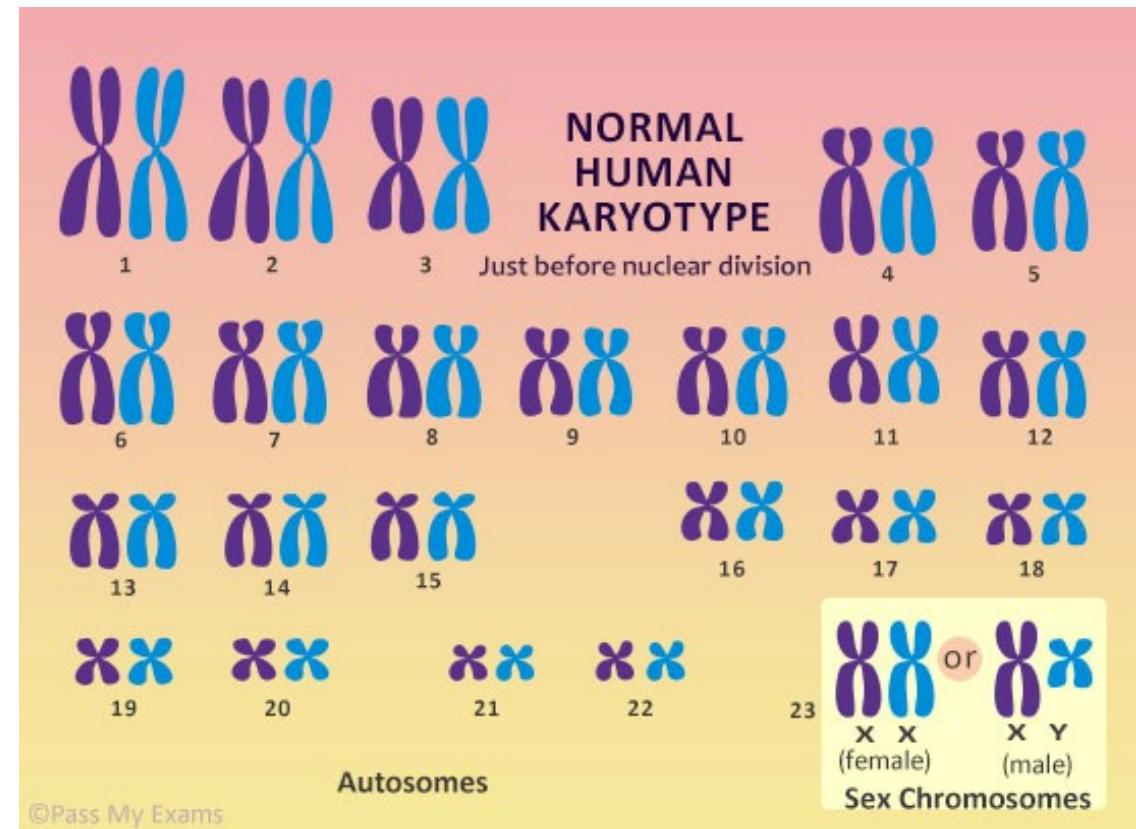
More terms/ideas to recall

Autosomal = on a ‘normal’ chromosome (chromosomes 1-22)

- ie not X or Y chromosomes which determine your sex
- Generally affects males/females equally

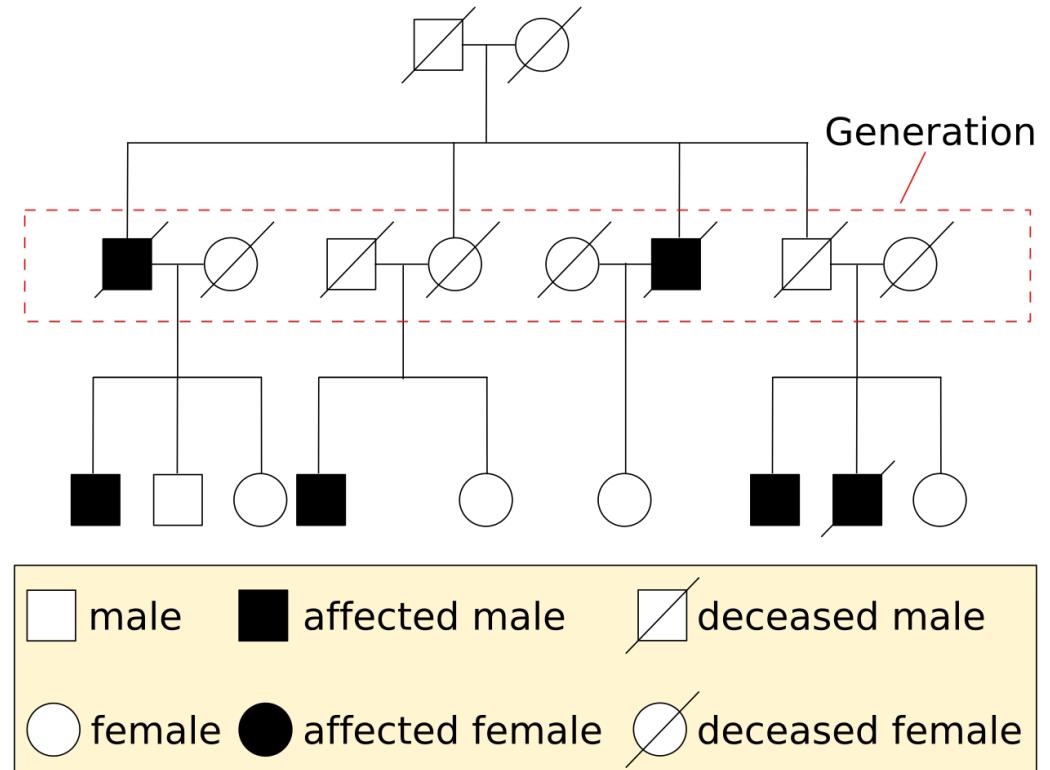
Sex-linked = on the X or Y chromosome

- Will affect males/females differently
- Female karyotype = $22^*2 + XX$
- Male karyotype = $22^*2 + XY$

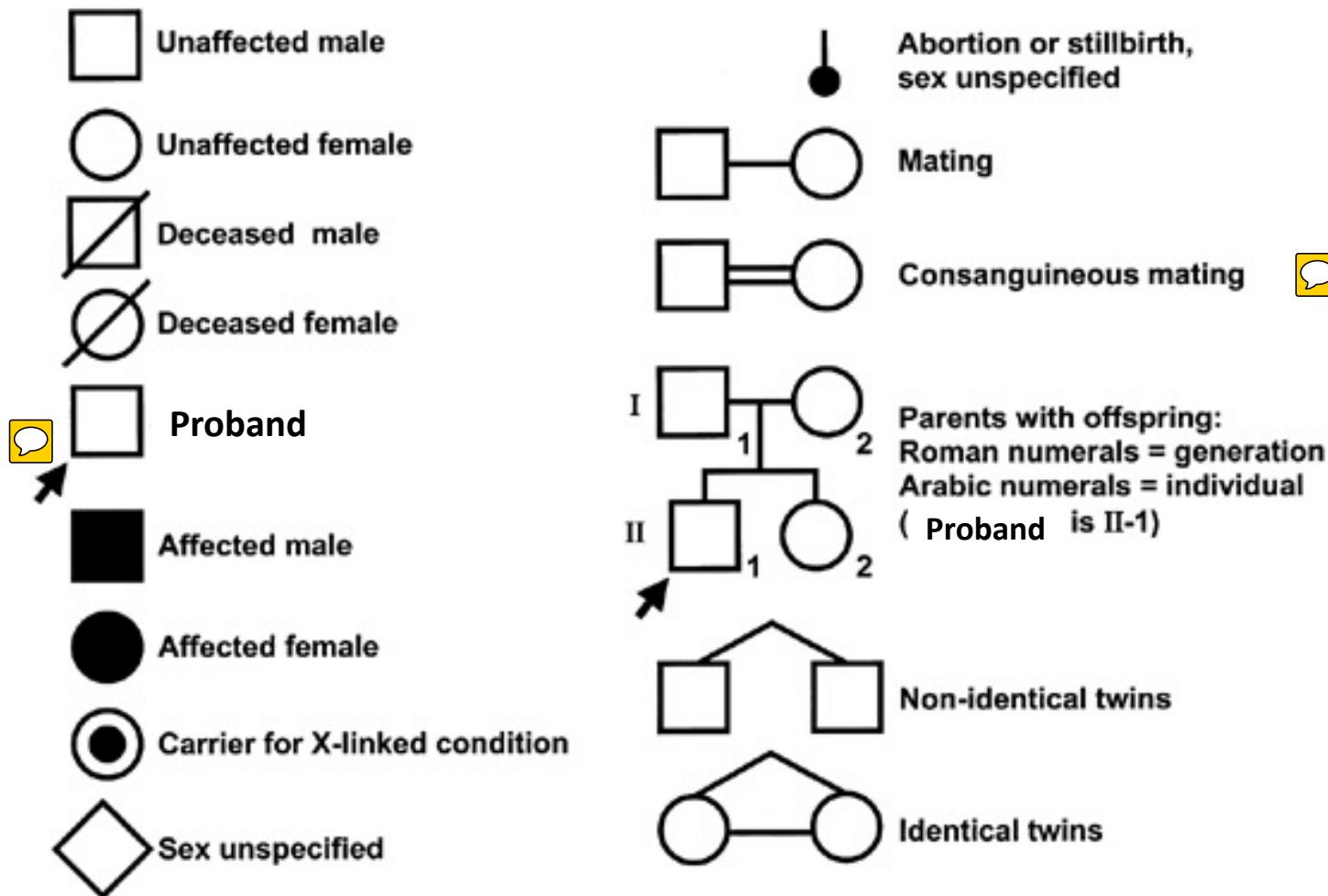


Inheritance in humans

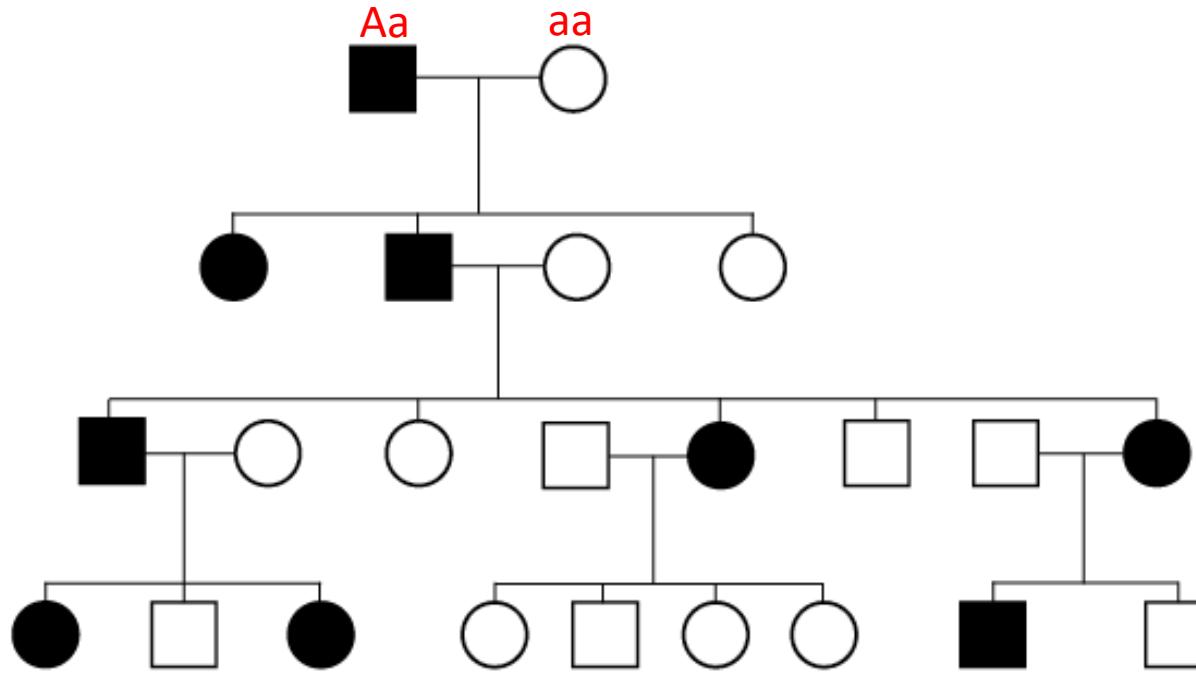
1. Autosomal dominant inheritance
2. Autosomal recessive inheritance
3. X-linked recessive inheritance
4. X-linked dominant inheritance
5. Y-linked inheritance
6. Mitochondrial inheritance



Human pedigree key

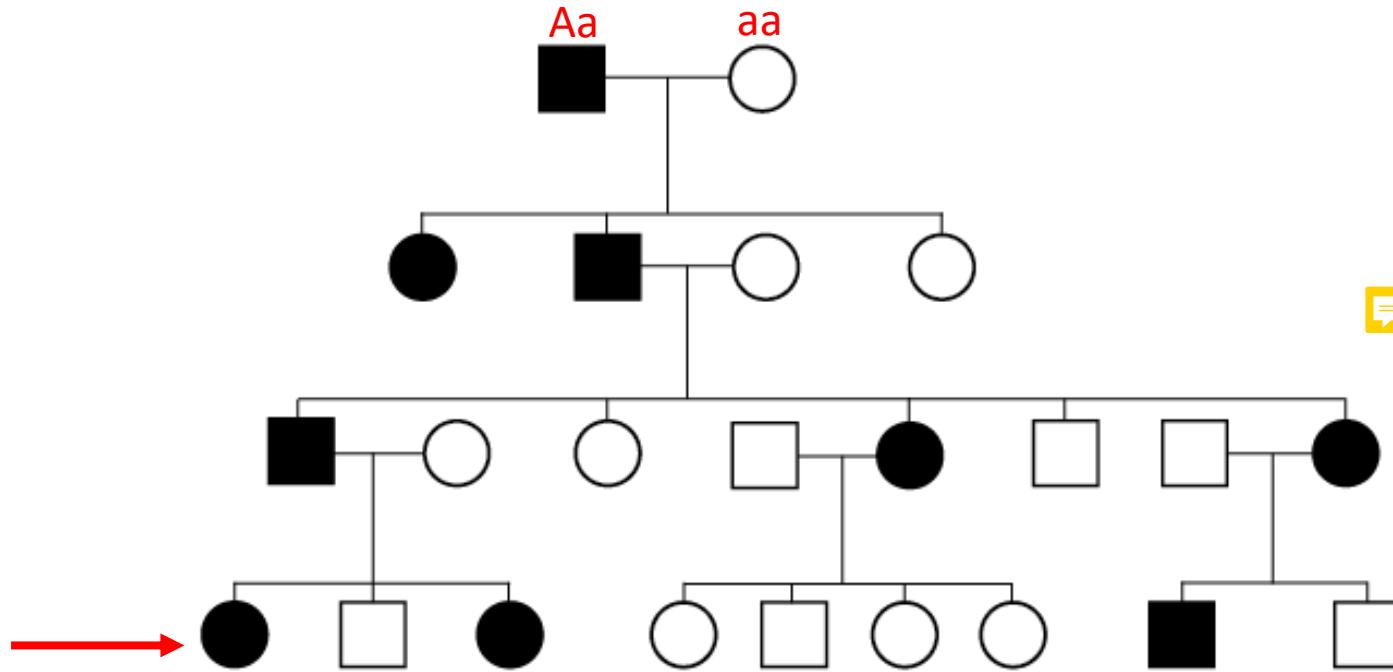


Autosomal dominant inheritance



- One faulty copy of gene causes the condition/trait
- 50% chance of passing on to each child
- An affected person has at least one affected parent
- Approximately half the offspring of an affected parent will be affected
- Affects both genders evenly

Autosomal dominant inheritance



- If this woman has children with an unaffected man, what are the chances of their first child being affected?
- Her genotype Aa, therefore 50% chance child would be affected

Autosomal dominant inheritance

Exceptions to autosomal dominant pattern of inheritance:

Incomplete penetrance

- When person carries mutant gene but does not show signs of disease
- Example: mutations in BRCA1 gene → 80% risk of breast cancer
- Pedigrees show skipping generations

Variable Expression

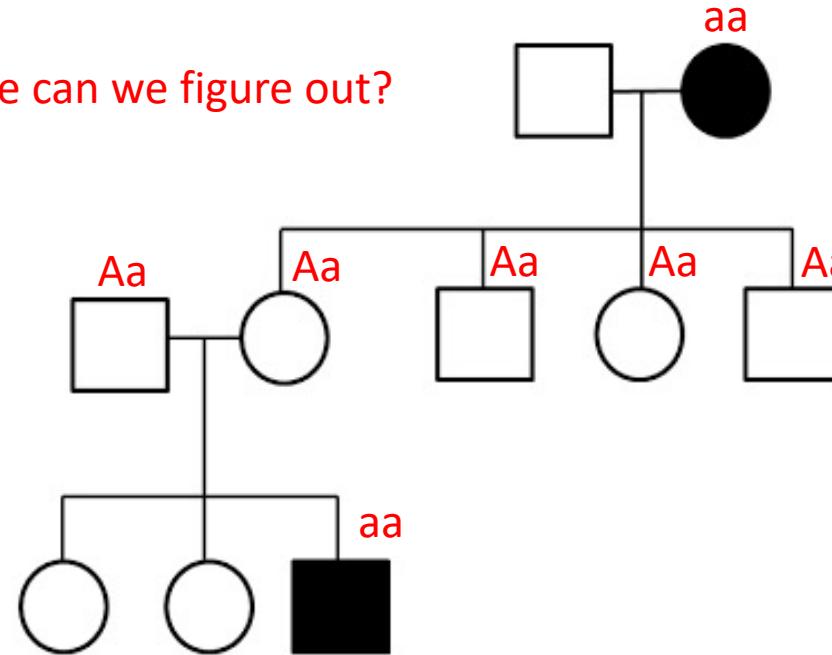
- Severity and manifestations of phenotype vary between individuals
- Example: mutation in FBN1 gene → Marfan syndrome is highly variable (disorder of connective tissue)

Other situations that might not look dominant

- De novo (new) mutation: sporadic cases without affected parents
- Delayed onset: appears to skip a generation, consider an early death

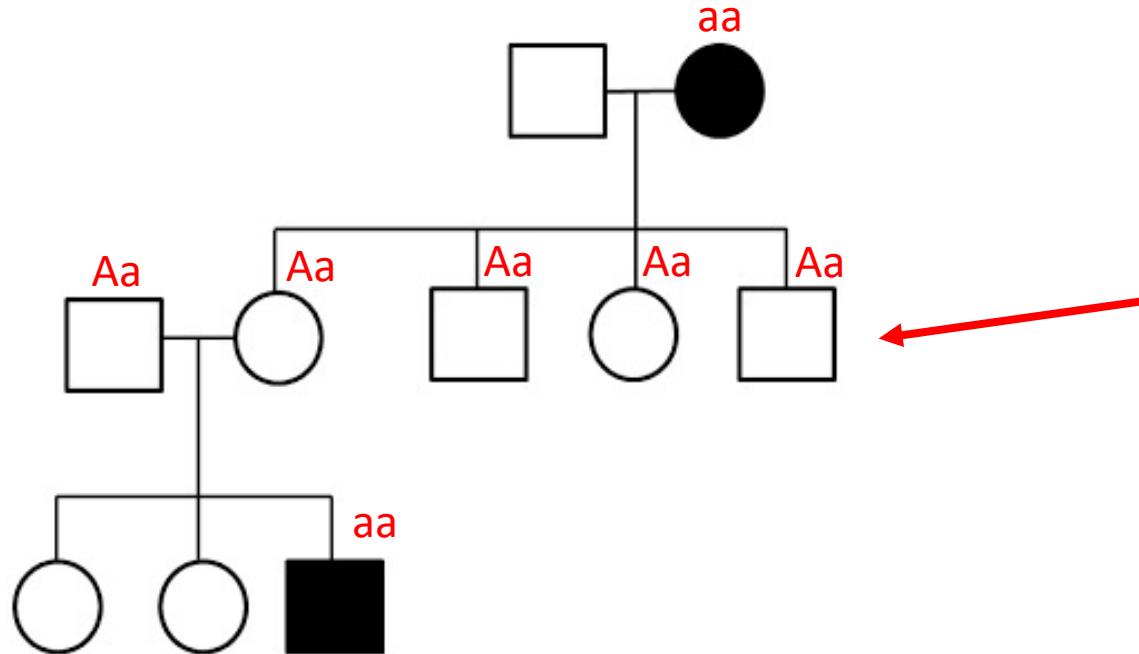
Autosomal recessive inheritance

Who else's genotype can we figure out?



- Two copies of gene required for the condition/trait (homozygous recessive)
- Can skip generations
- Two unaffected parents can have an affected child (indicates parents are carriers)
- Affects genders equally

Autosomal recessive inheritance



- If this person has a child with a carrier woman, what are the chances their child would be affected?
- If both parents carriers
 - $\frac{1}{4}$ of offspring affected aa
 - $\frac{2}{4}$ of offspring carriers Aa
 - $\frac{1}{4}$ of offspring normal AA

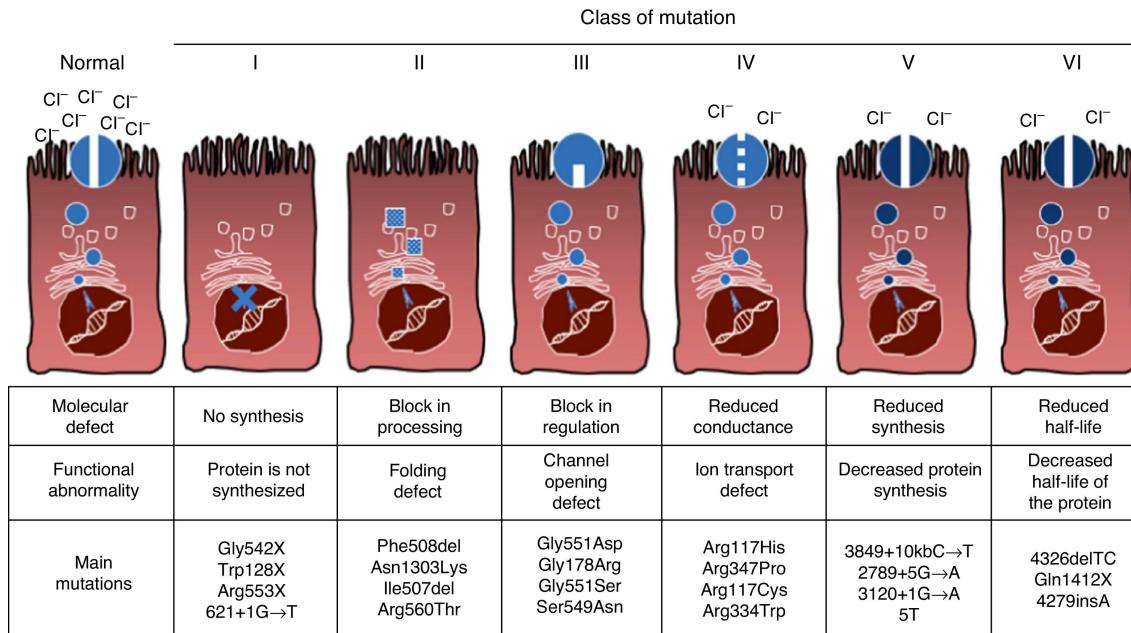
	A	a
A	AA	Aa
a	Aa	aa

Allelic heterogeneity =

Usually many mutations in a gene which can cause a condition

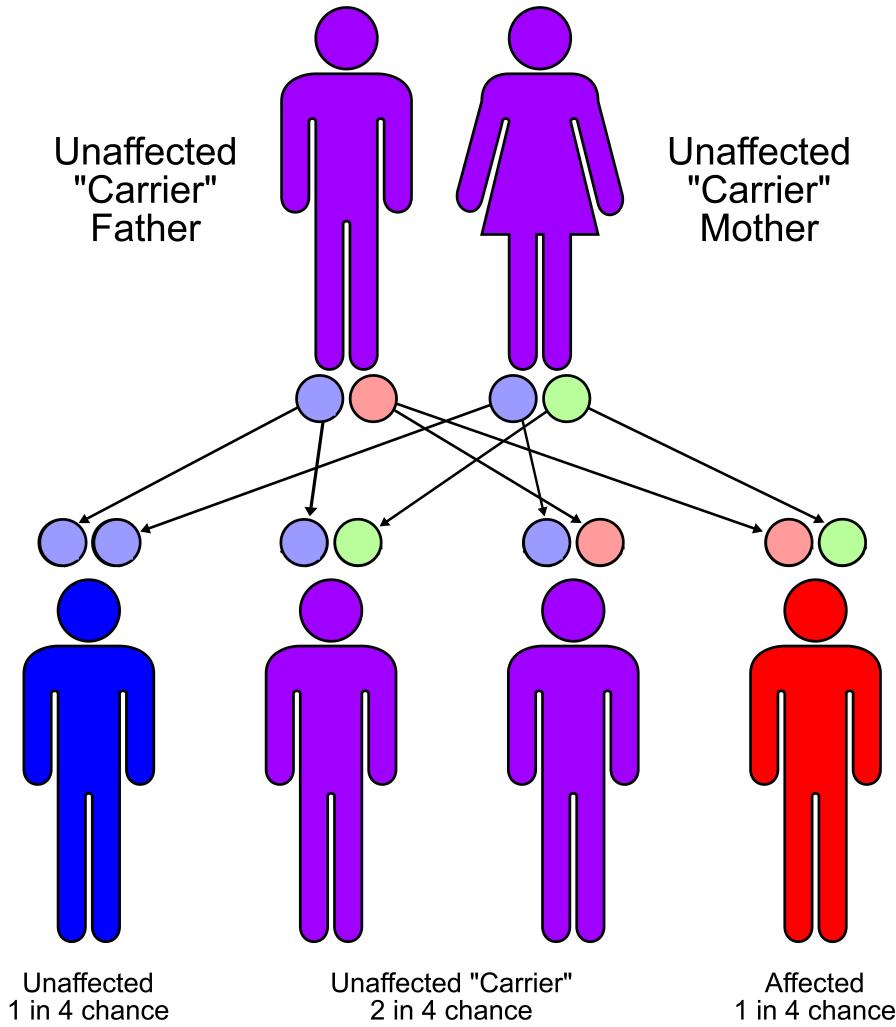
Cystic fibrosis = autosomal recessive genetic disorder

- Caused by mutations in CFTR gene (cystic fibrosis transmembrane conductance regulator)
- over 2,000 mutations in the CFTR gene associated with cystic fibrosis disease
- varying degrees of frequency within the disease carrying population
- different mutations produce varying degrees of disease phenotypes
- can also work in combinations to produce additive phenotypic effects.

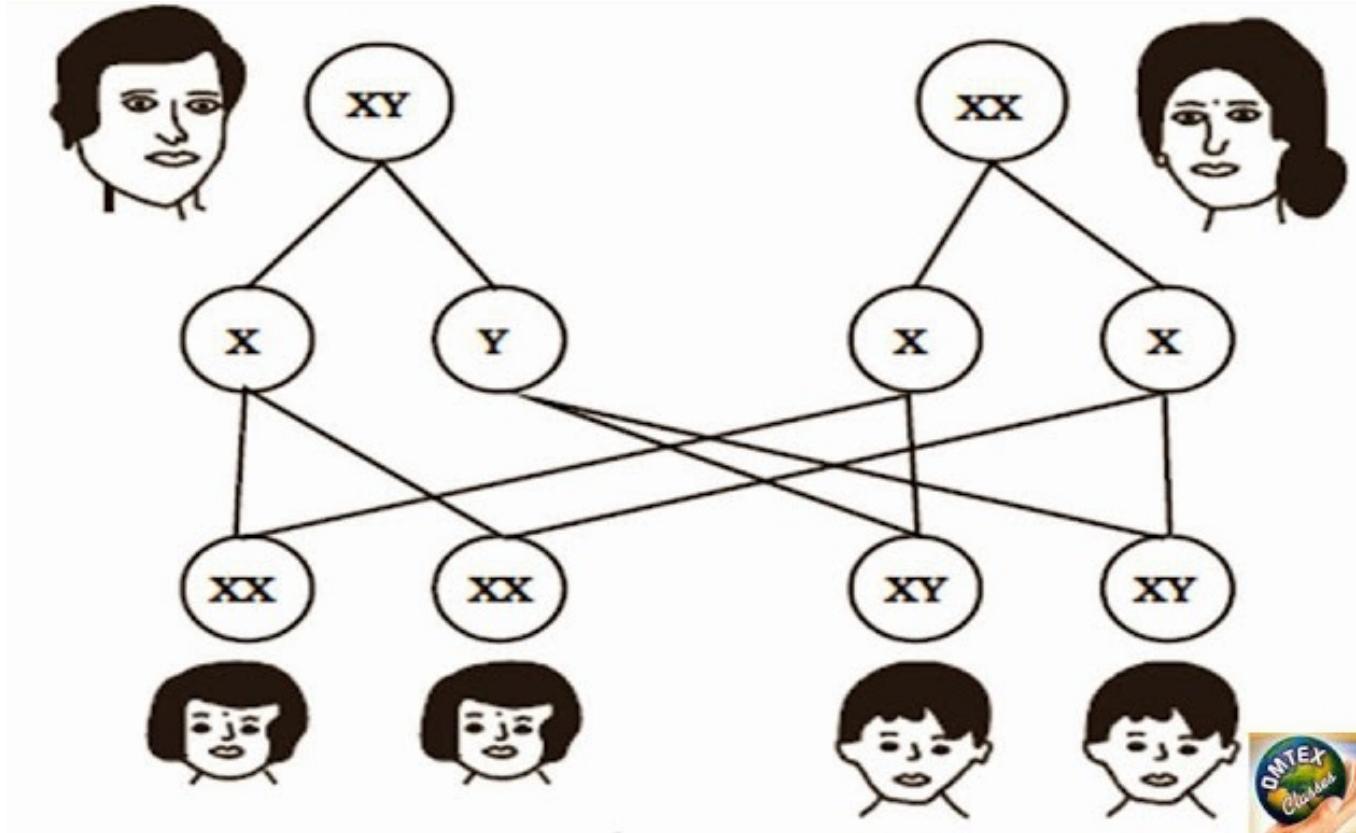


Compound heterozygote

- The presence of two different mutant alleles at a particular gene locus
- Both are ‘recessive’



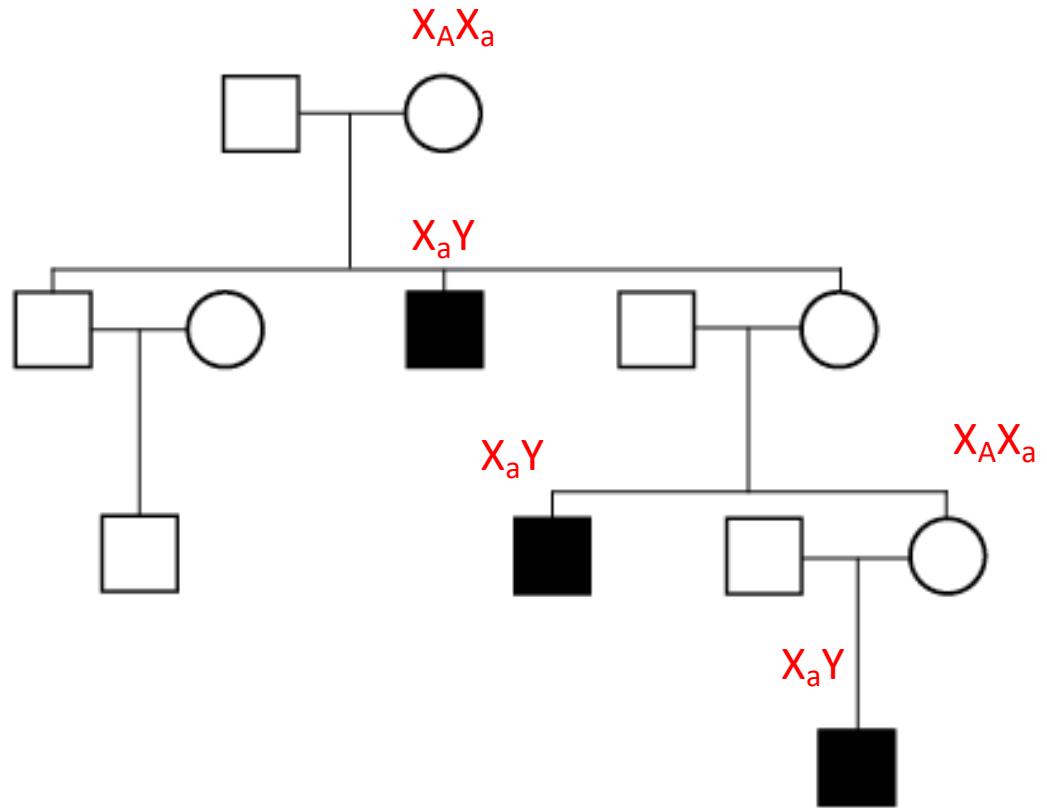
Quick revision of sex determination



- Mothers always pass on one of their two X chromosomes
- Fathers pass on:
 - Y to sons
 - X to daughters

X-linked recessive inheritance

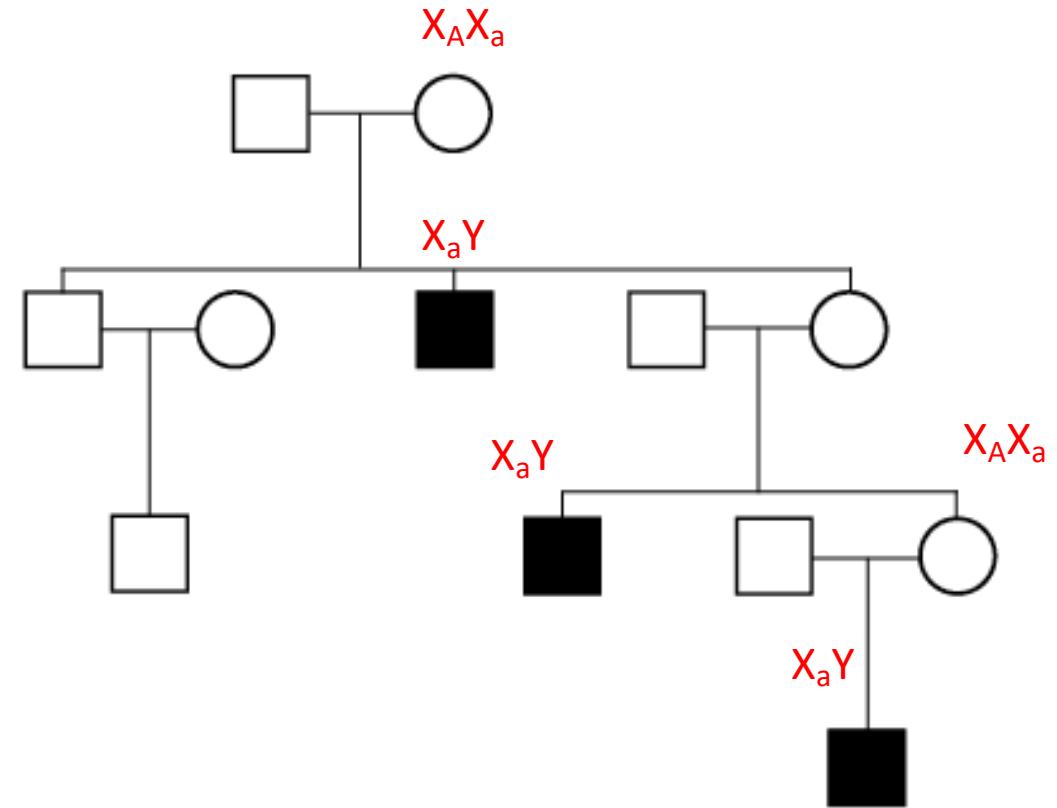
- Males hemizygous (only one copy) = affected or not, can't be carriers
 - Males more often affected by X-linked conditions
 - Males get their X from mother, so affected males have carrier mothers
- Females can be carriers, affected females much lower frequency since two copies of X
- Carrier mothers pass on either healthy X or mutant X
 - 50% of sons of carrier mothers will be affected
 - 50% of daughters of carriers mothers will be carriers
- Affected fathers pass on either X or Y
 - Y = boy – affected males can't pass condition onto a son
 - X = girl – all daughters of affected males will be carriers



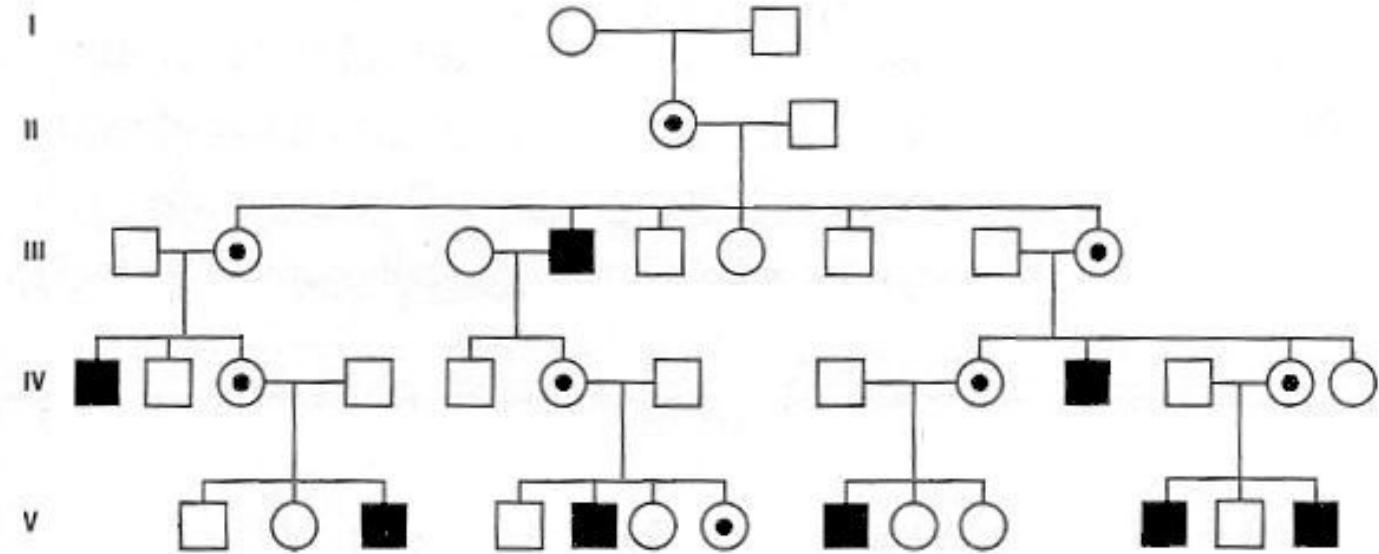
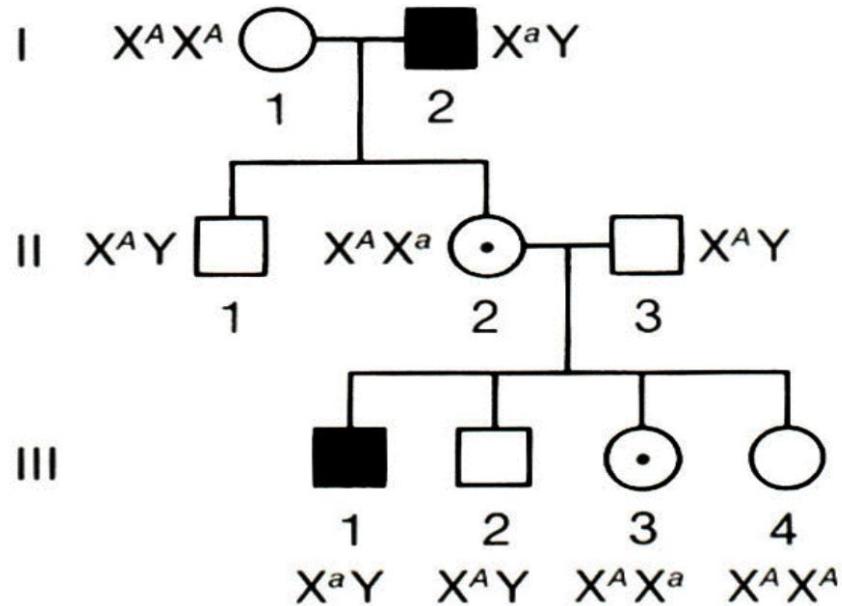
X-linked recessive inheritance

End result of all this:

- Only / mostly males affected
- Can skip generations

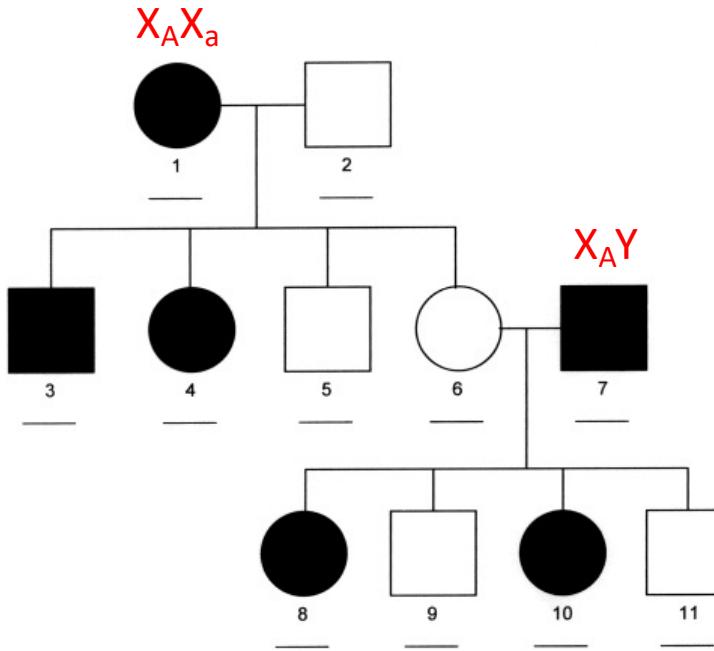


X-linked recessive inheritance



- Sometimes dots or circles used to indicate carrier women in X-linked
- But don't assume you will get this information

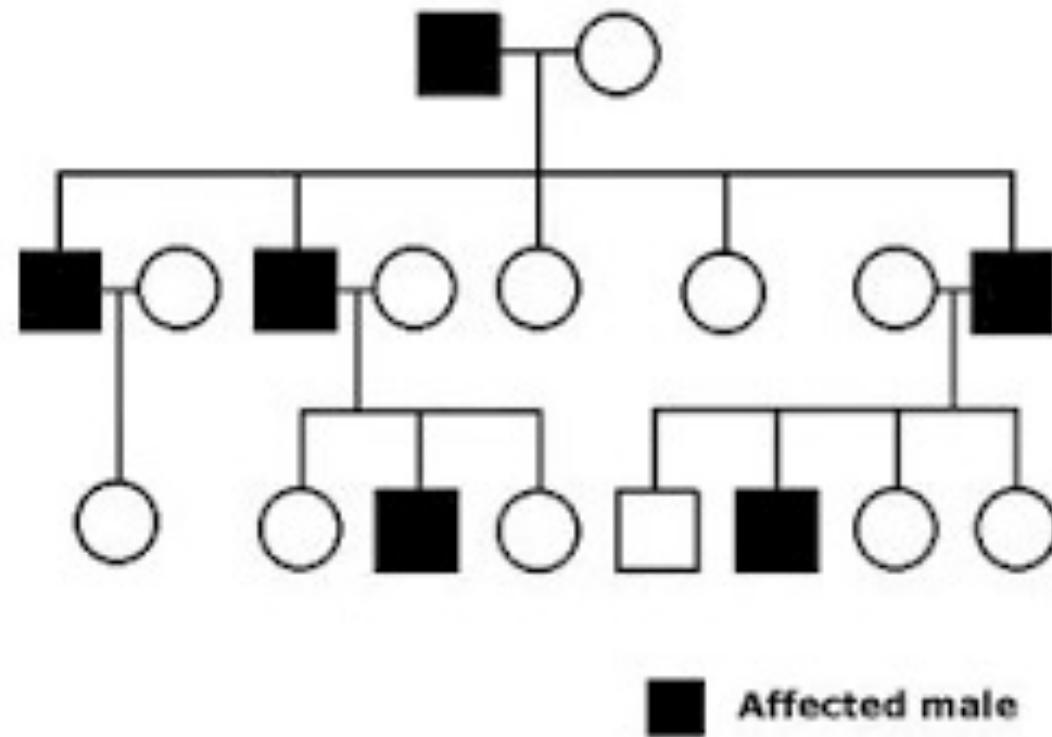
X-linked dominant inheritance



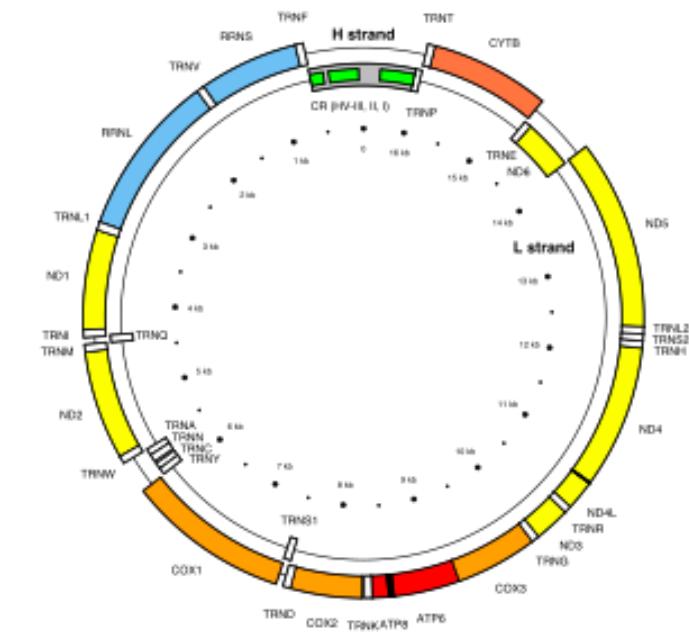
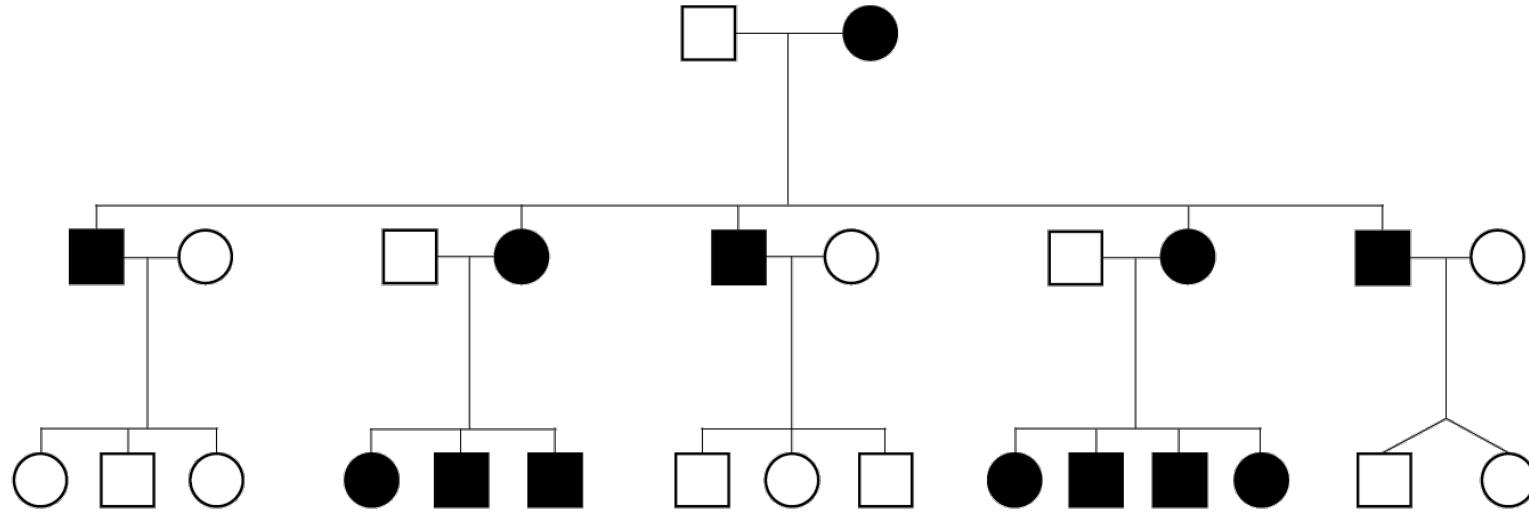
- No male-to-male inheritance since Y being passed on
- Affected males transmit to 100% of daughters but 0% of sons
- Affected females transmit to 50% of their offspring, regardless of sex
 - ratio = 2 affected females : 1 affected male

Y-linked inheritance

- All sons of an affected father are affected
 - Very rare for expressed traits
 - Only males can be affected
 - Dominance is irrelevant
-
- Far fewer Y-linked genetic disorders compared to X-linked
 - Y chromosome smaller and fewer genes compared to X
 - The SRY (sex determining region of the Y-chromosome) develops the testes and can be associated with Y-linked disorders

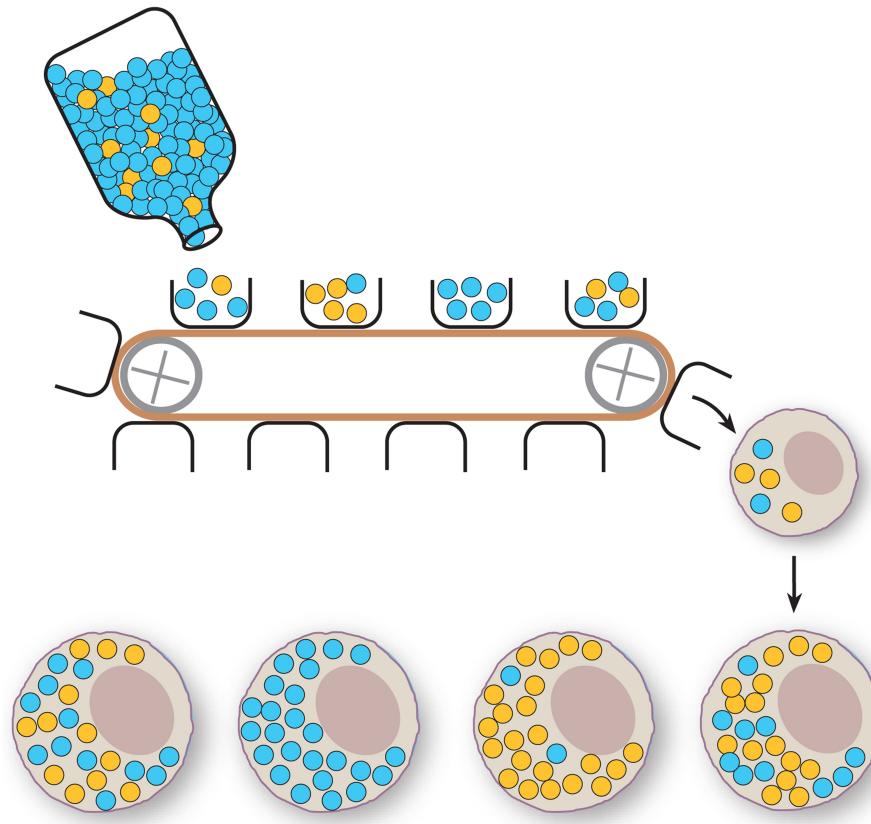


Mitochondrial inheritance



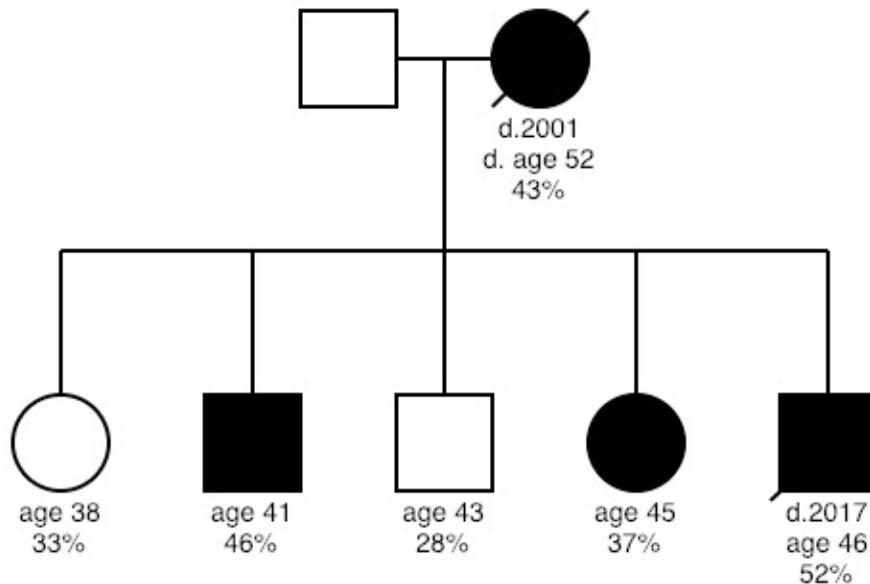
- Mitochondrial genome (16,569 base pairs) encodes genes of mitochondrial oxidative phosphorylation pathway and mitochondrial tRNAs
 - Mutations can cause structurally and functionally abnormal mitochondria
 - Usually maternally inherited (mitochondria in egg cell, not sperm cells)
 - Rare evidence of father passing on mitochondrial disease
 - Theoretically: every child of affected mother is affected

Reality: Mitochondrial heteroplasmy



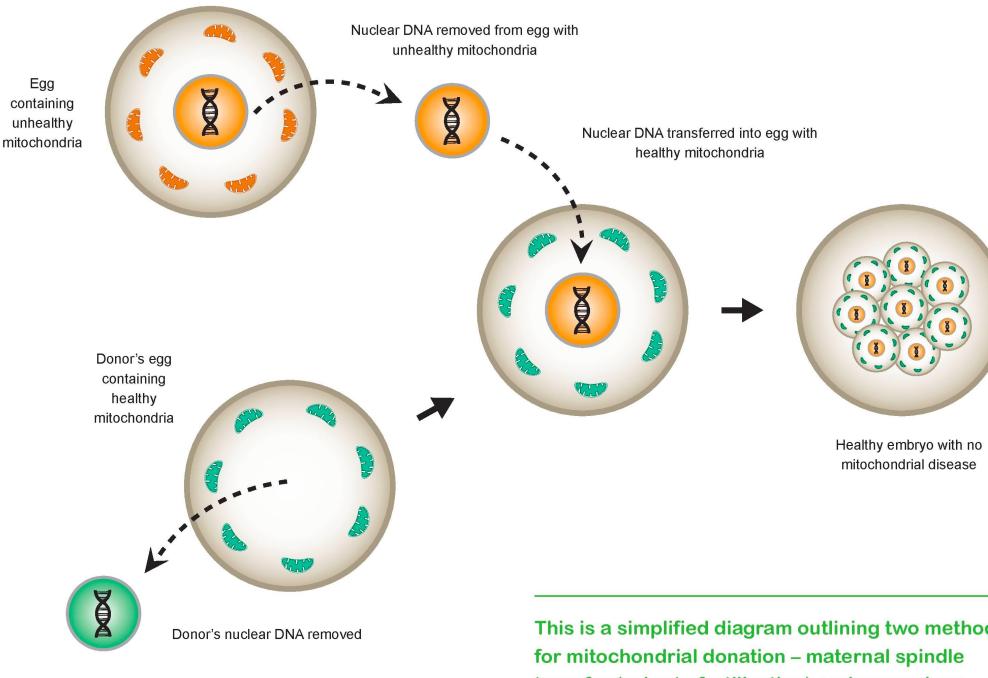
- Heteroplasmy = having a mix of mitochondrial genomes within cells/organism
- Heteroplasmy levels vary for each child of affected mum

MELAS: Mitochondrial Encephalopathy, Lactic acidosis and Stroke-like episodes



- m.3243A>G in majority of affected people
- Can be a ‘threshold level’ before disease onset
- Severity and age of onset often vary with heteroplasmy level

Mitochondrial Replacement Therapy



- Nuclear DNA removed from a fertilised egg of affected mother
- Transplanted into the fertilised egg of a female donor whose nuclear material has been removed
- Legal in UK from 2018, legal in Australia soon?

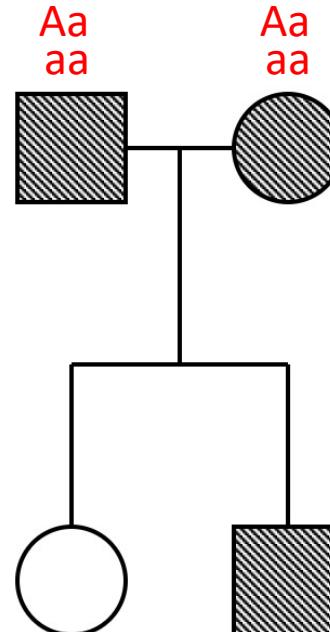
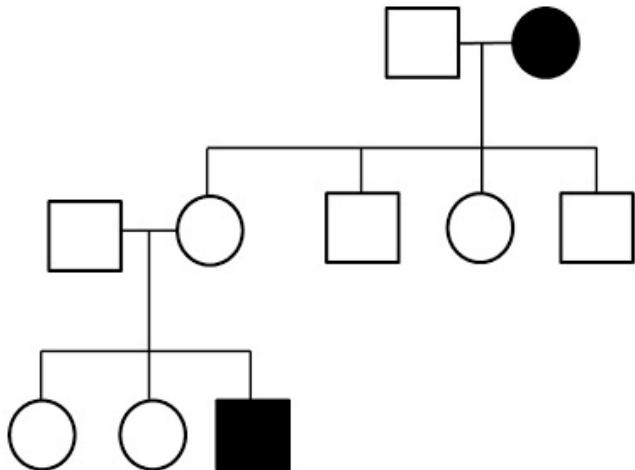
Multiple modes of inheritance for same disease:

Kallman syndrome / normosmic isolated gonadotrophin-releasing hormone deficiency

	Gene ^{1, 2}	% of IGD Attributed to Pathogenic Variants in This Gene ³
X-R	<i>ANOS1</i> (<i>KALI</i>)	5%-10% (KS)
AD	<i>CHD7</i>	5%-10% (KS or nIGD)
AD	<i>FGFR1</i>	~10% (KS or nIGD)
AD or AR	<i>GNRHR</i>	5%-10% (nIGD)
AR	<i>IL17RD</i>	2%-5% (KS or nIGD)
AR	<i>PROKR2</i>	~5% (KS or nIGD)
AD	<i>SOX10</i>	2%-5% (KS)
AR	<i>TACR3</i>	~5% (nIGD)

Pedigrees help on a family by family basis, not necessarily a disease by disease basis

General pedigree tips



- If you know the mode of inheritance, you know the genotype of anyone who is affected
 - You might be able to infer the genotypes of an affected person's parents
- You may be able to rule out a particular mode of inheritance
 - Father passing onto son = can't be X-linked (though mum may be a carrier)
 - Two affected people having an unaffected child = can't be recessive
 - Try to genotype everyone assuming a particular mode of inheritance – it just won't work for some modes

Population genetics: Hardy Weinberg equilibrium

$$p + q = 1$$

Allele	A	a
Frequency	p	q

$$p^2 + 2pq + q^2 = 1$$

Genotype	AA	Aa	aa
Frequency	p^2	$2pq$	q^2

Assumptions of Hardy-Weinberg equilibrium

1. Large population (avoids random chance having a big impact on allele frequencies).
2. No selection (ie no allele is necessarily more likely to get passed on than another)
3. Mating is random (ie any individual has an equal chance of mating with any other individual).
4. Mutation either does not occur or is in equilibrium.
5. Immigration and emigration do not occur.

The consequences of abiding by the Hardy-Weinberg Law are that allele frequencies remain constant from generation to generation, so we can calculate allele/genotype frequencies

Do these assumptions fit for humans?

Hardy Weinberg equilibrium

$$p + q = 1$$

Allele	A	a
Frequency	p	q

$$p^2 + 2pq + q^2 = 1$$

Genotype	AA	Aa	aa
Frequency	p^2	$2pq$	q^2

In Caucasian populations, approximately 1 in 2,500 people are affected with cystic fibrosis

Theoretically, what is the carrier rate in this population?

$$q^2 = 1/2500$$

$$q = \text{square root } 1/2500 = 0.02$$

$$p = 1 - q = 0.98$$

$$\text{Carrier rate} = 2pq = 2 * .02 *.98 = .0392$$

One more...

Consider two alleles A and a for sickle cell disease

	AA	Aa	aa
Observed:	25,374	5,482	67
Expected:	25,562	5,106	255

Allele frequencies $p = 0.91$ and $q = 0.09$

d	0.05	0.01	0.001
1	3.841	6.635	10.828
2	5.991	9.210	13.816
3	7.815	11.345	16.266
4	9.488	13.277	18.467
5	11.070	15.086	20.515
6	12.592	16.812	22.458
7	14.067	18.475	24.322
8	15.507	20.090	26.125
9	16.919	21.666	27.877

Chi-squared = 167.7, p-value < 0.001 (very significantly different)

Excess of heterozygotes, deficiency of homozygotes

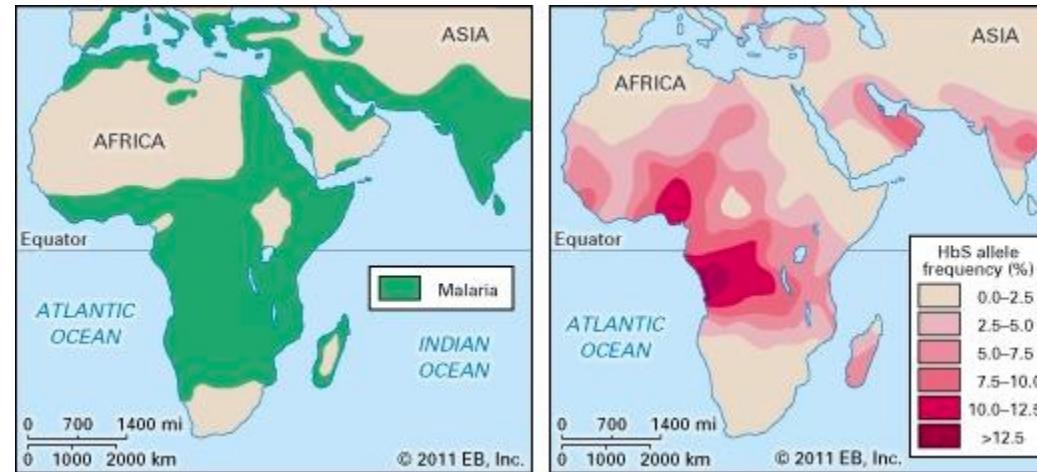
Why might this population be so different from HWE?

Positive selection in humans

AA – no sickle cell, but relatively susceptible to malaria

Aa – sickle cell carrier, not clinically important

aa – die early because of sickle cell anaemia



in equatorial Africa, up to 40% of people are carriers of this mutated gene

Read a recent paper to find out more:

<http://science.sciencemag.org/content/334/6060/1283.full>



MACQUARIE
University

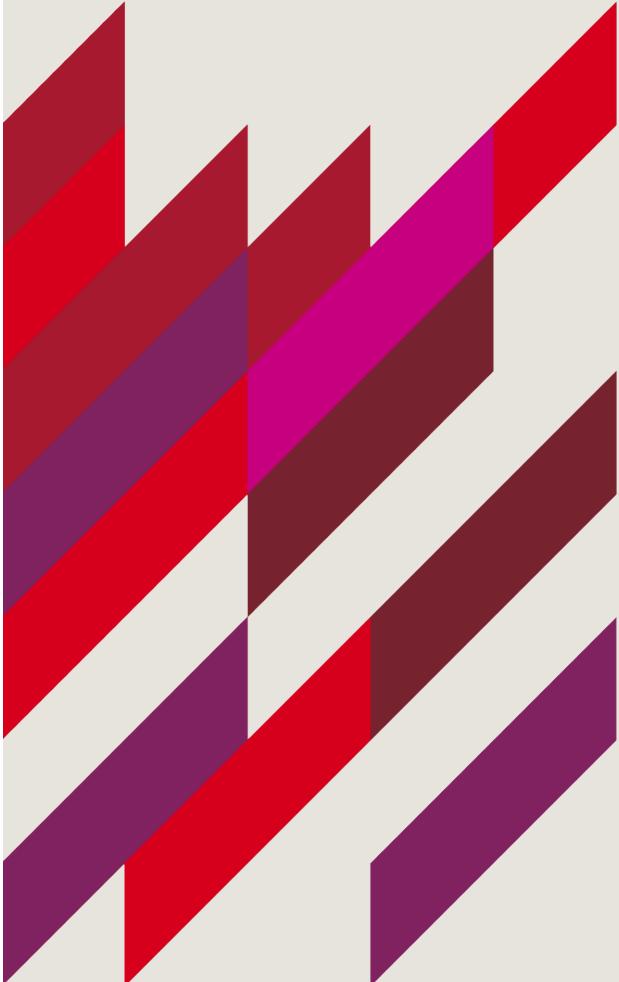
The Human Genome

BIOL3120 Human Genetics and Evolutionary Medicine



The Human Genome

LEARNING OBJECTIVES



At the end of this lecture you should be able to:

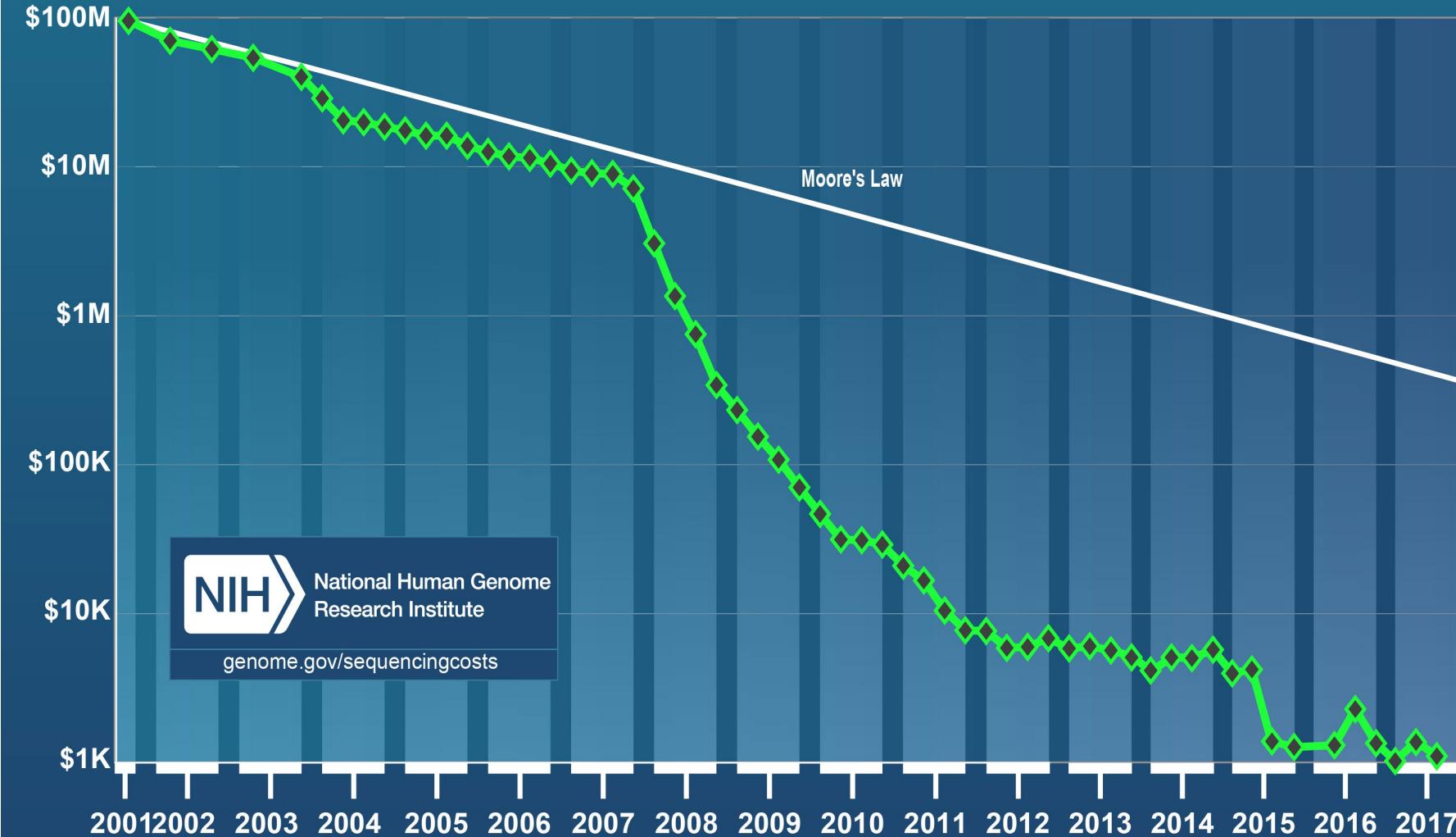
- Define the terms used to describe the human genome
- Understand the structure of chromosomes
- Understand the coding regions of the human genome

The human genome

- First draft sequence February 2001
- Complete draft 2003
- Cost ~ 2.7 billion USD



Cost per Genome



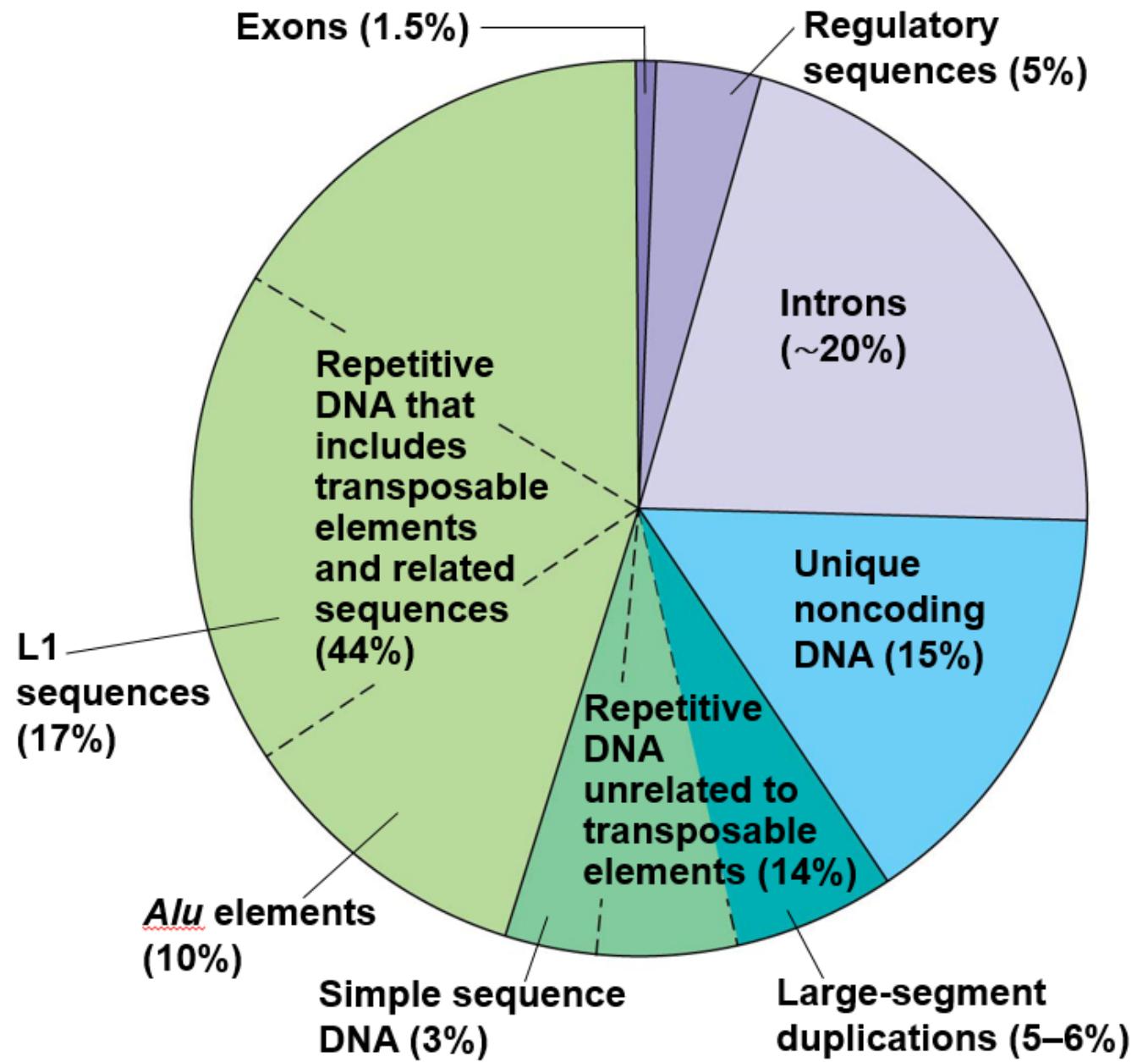
The human genome

- First draft sequence February 2001
 - Complete draft 2003
 - Cost ~ 2.7 billion USD
-
- Sequence determined but little understanding of its meaning



Quick overview

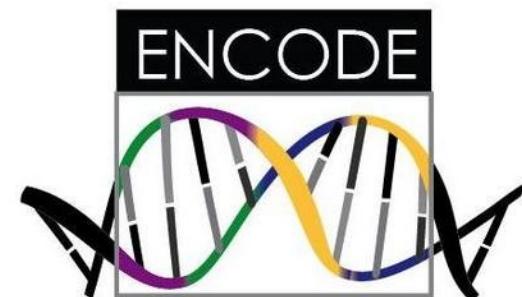
- 3.2 billion basepairs x 2
- Over 23 chromosome pairs
- ~20,000 protein-coding genes



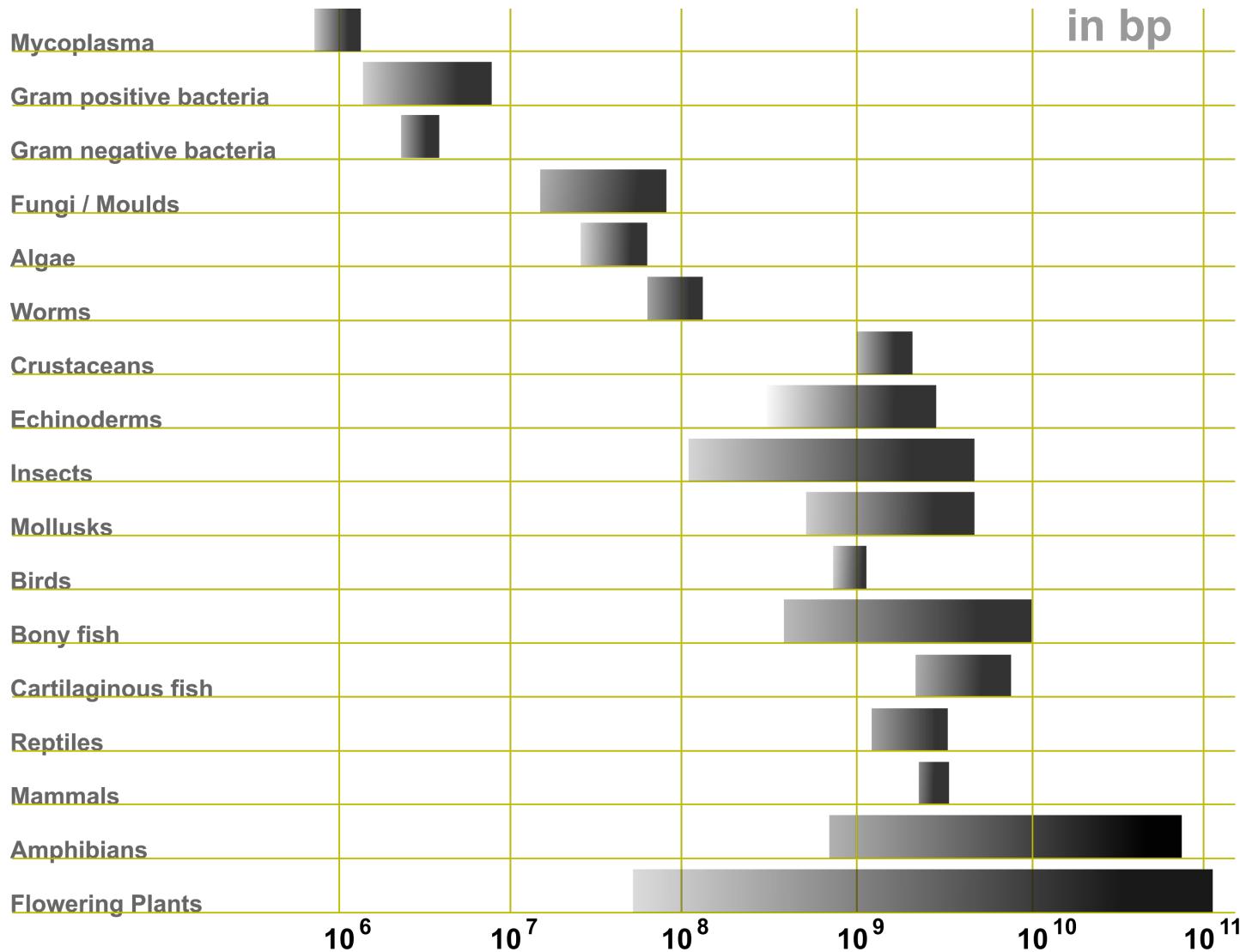
Junk DNA?

- Historically, non-coding DNA referred to as junk DNA
- ENCODE project (2012):
 - “The vast majority (80.4%) of the human genome participates in at least one biochemical RNA- and/or chromatin-associated event in at least one cell type”
 - (think about gene definition)
- <https://www.nature.com/articles/nature11247>

National Human Genome Research Institute



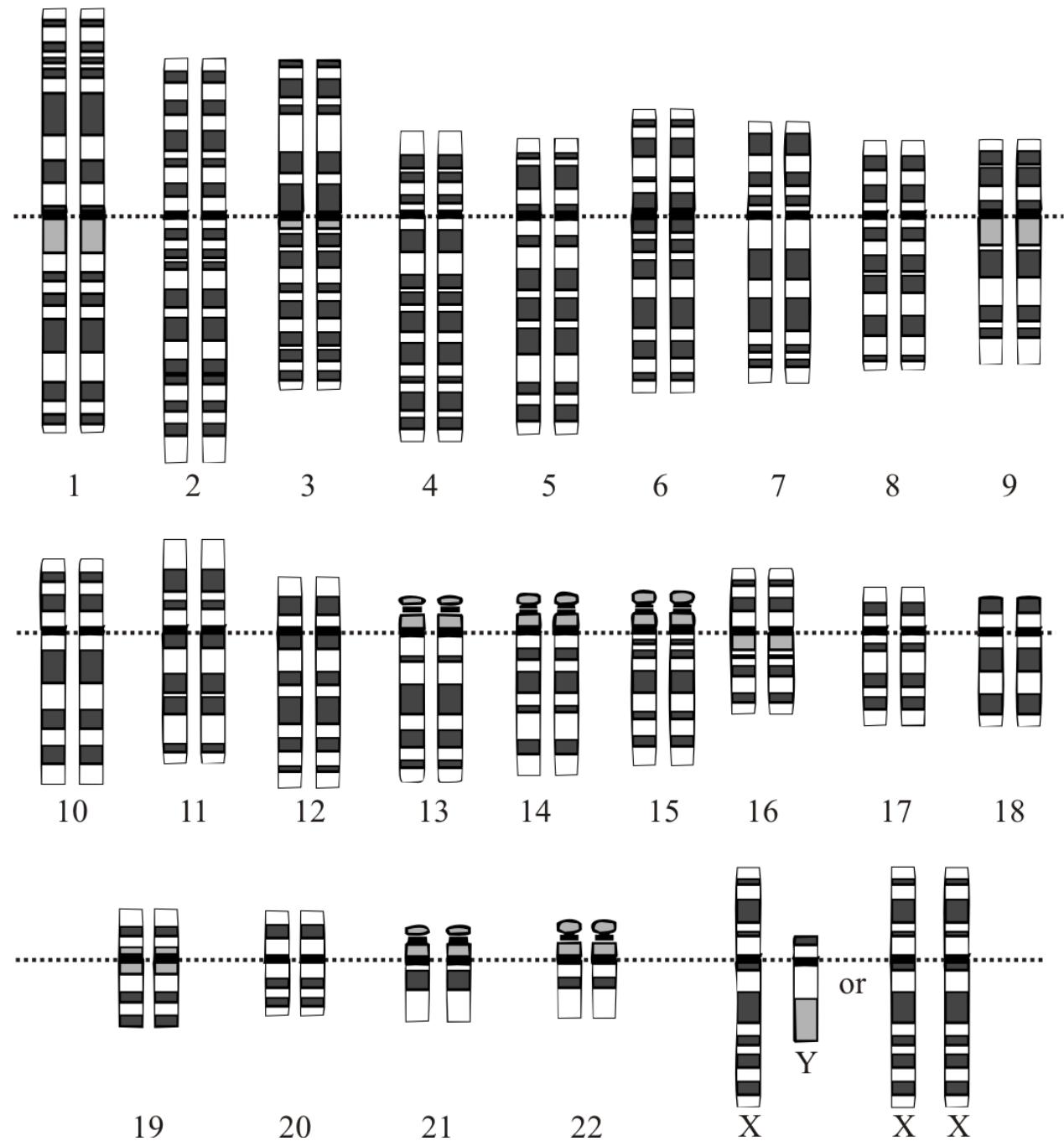
Example genome sizes





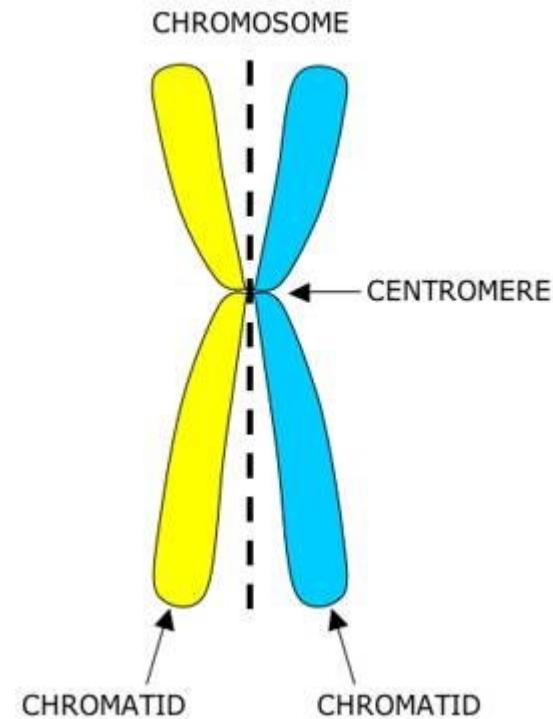
Terms used to describe the human genome

Chromosomes



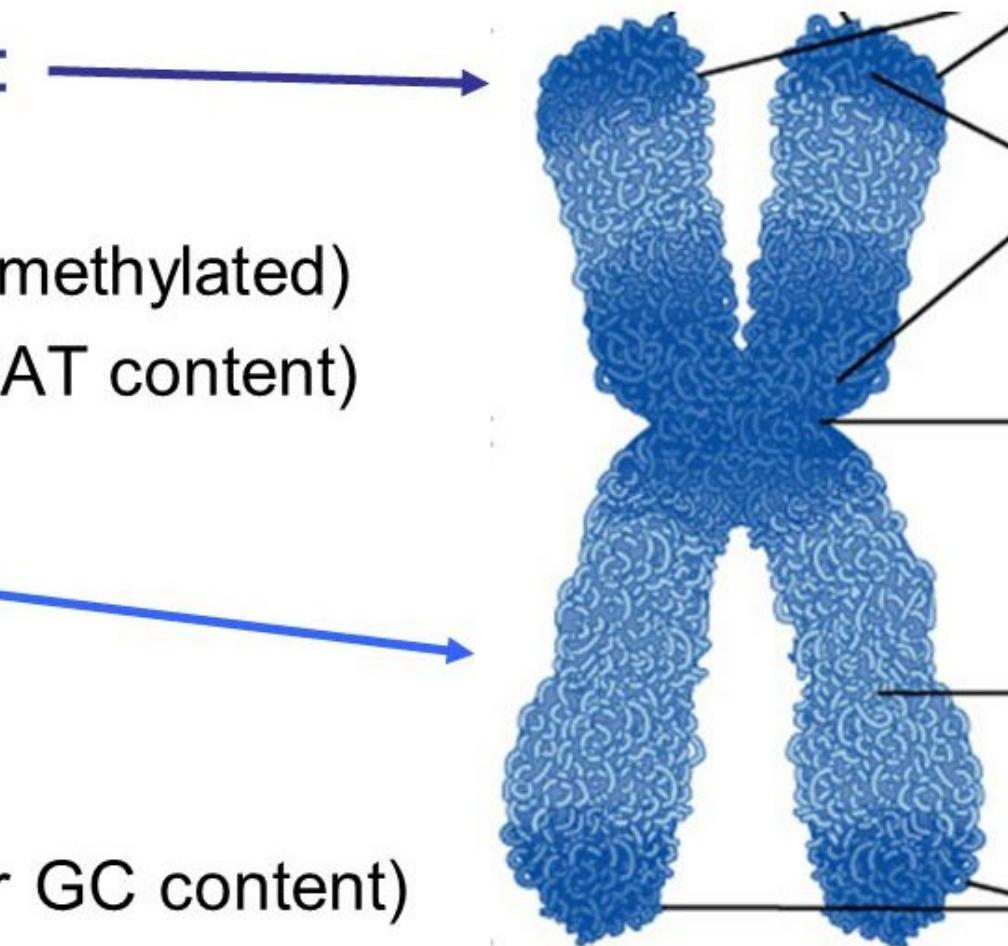
Chromosome related nomenclature

- Chromatin: the material of which chromosomes are made.
 - complex of DNA, and proteins (histones and non-histones)
- Chromatid: one of the two copies of a chromosome
 - joined at the centromere
 - two homologous chromatids are called sister chromatids.
- Centromere is DNA normally in a heterochromatin state
 - Site of spindle attachment to chromosomes (important in meiosis)

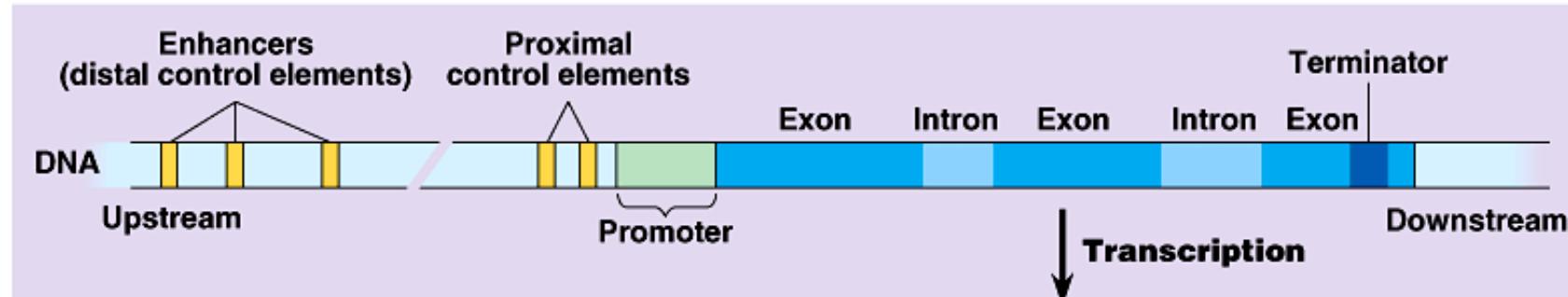


Chromosome parts

- **Heterochromatin:**
 - More condensed
 - Silenced genes (methylated)
 - Gene poor (high AT content)
 - Stains darker
- **Euchromatin:**
 - Less condensed
 - Gene expressing
 - Gene rich (higher GC content)
 - Stains lighter



DNA & Gene parts



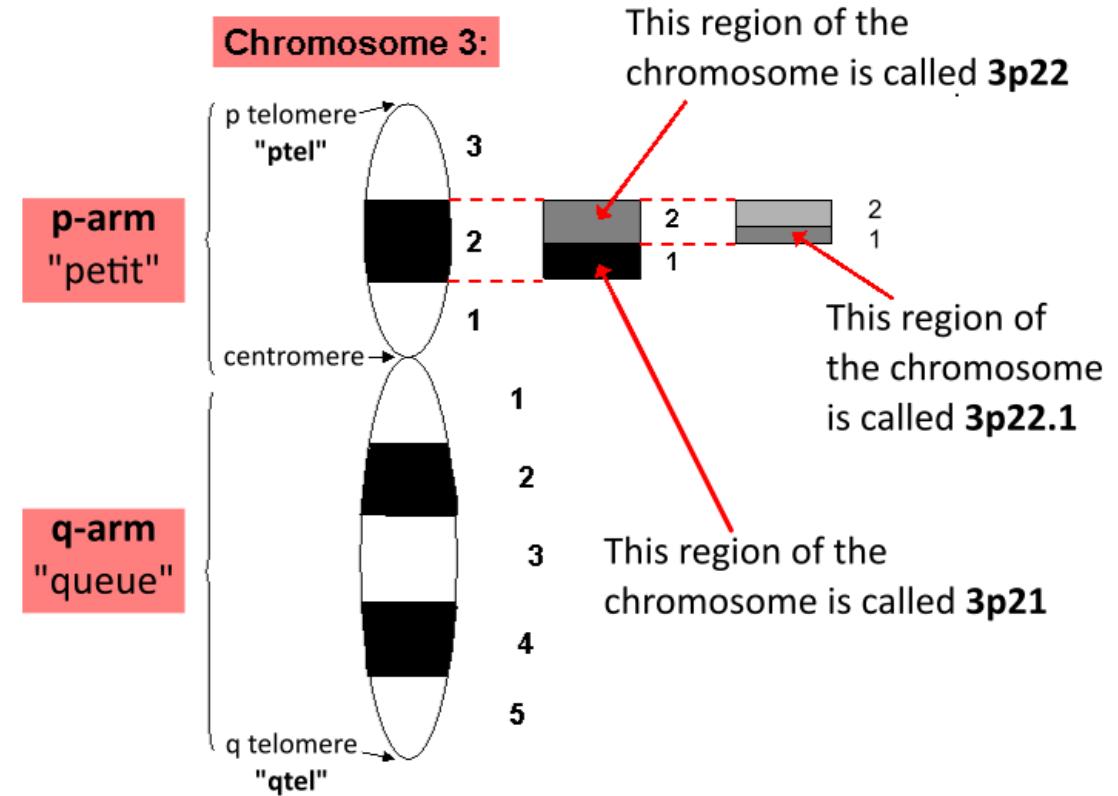
- Enhancers: genomic DNA (gDNA) which can alter the expression of a gene
- Promoters: gDNA region where transcription factors bind to initiate transcription of a gene
- Exons: gDNA that is the coding regions of a gene
- Introns: a portion of a gene that does not code for an amino acid
- Gene: a sequence of nucleotides in DNA that codes for a molecule that has a function



The Structure of Chromosomes

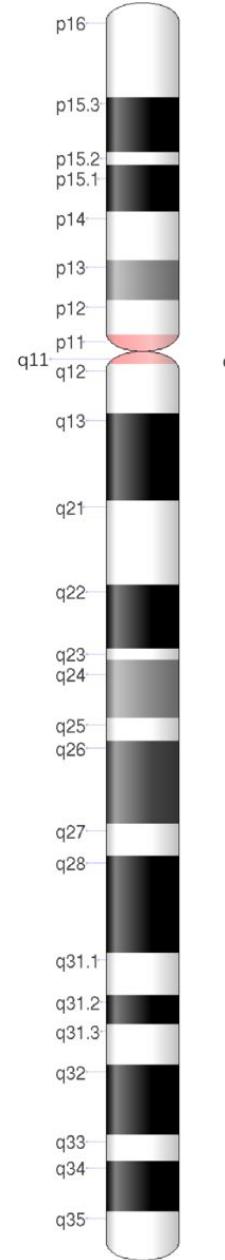
Chromosome numbering

- 1-22 Arranged in order of decreasing size
- q = long arm; p = short arm



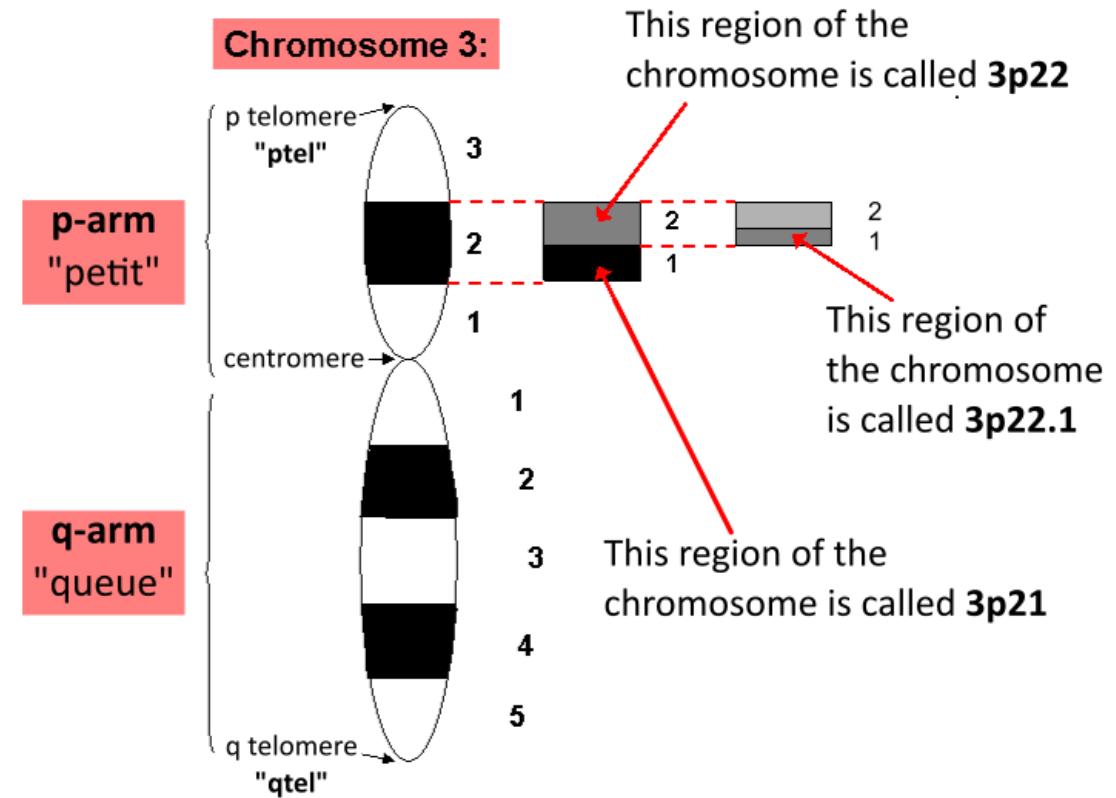
G-banding example: Chromosome 4

- Giemsa stain
 - darker on heterochromatin (AT rich) regions, less accessible
 - Lighter on euchromatin (GC rich) regions, more accessible and active
 - 92% of genes in euchromatin in humans



Chromosome numbering

- 1-22 Arranged in order of decreasing size
- q = long arm; p = short arm
- Black and white stripes are because of staining: G-banding
- ‘Bands’ have the same appearance on homologous chromosomes so identification became easier
- Banding allows:
 - Unequivocal identification of each human chromosome
 - Detection of rearrangements (e.g. translocations)

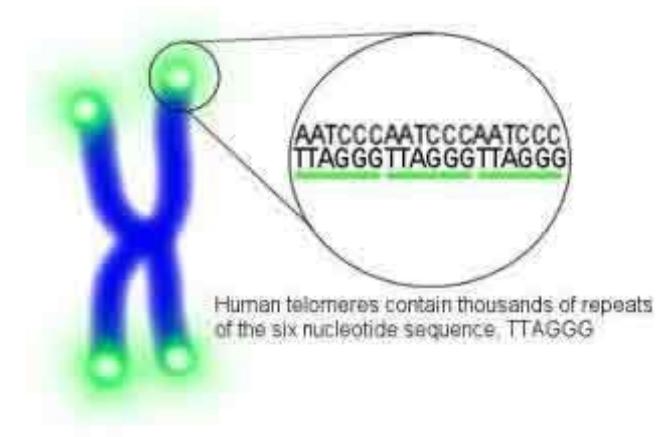
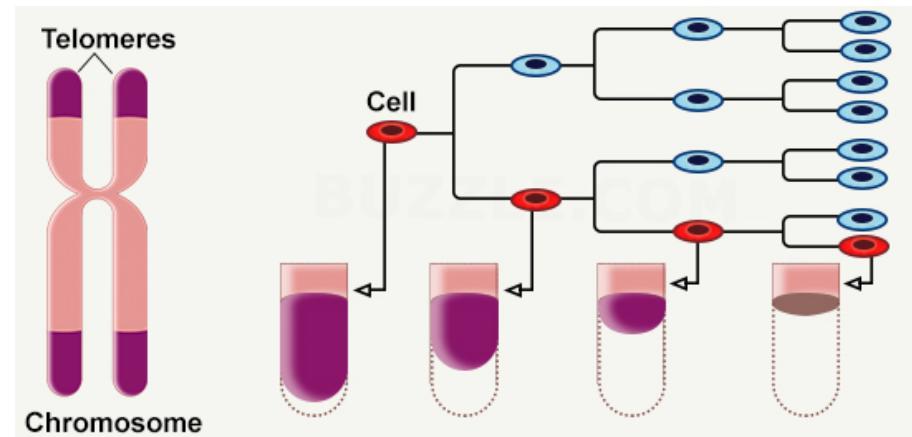


Chromosomes



Telomeres

- protect functional regions (containing genes) by capping the ends with a 'buffer' sequence
- Telomeres consist of variable numbers of a repetitive sequence (TTAGGG)
- Telomere length is reduced every time a cell divides
- average telomere length declines from ~ 11 kilobases at birth to less than 4 kilobases in old age, decline greater in men
- Marker of cell age – signal for apoptosis (cell self-destruction)
- “telomeric DNA is **lost** at an **average rate** of '25.7–27.7 base pairs' per year



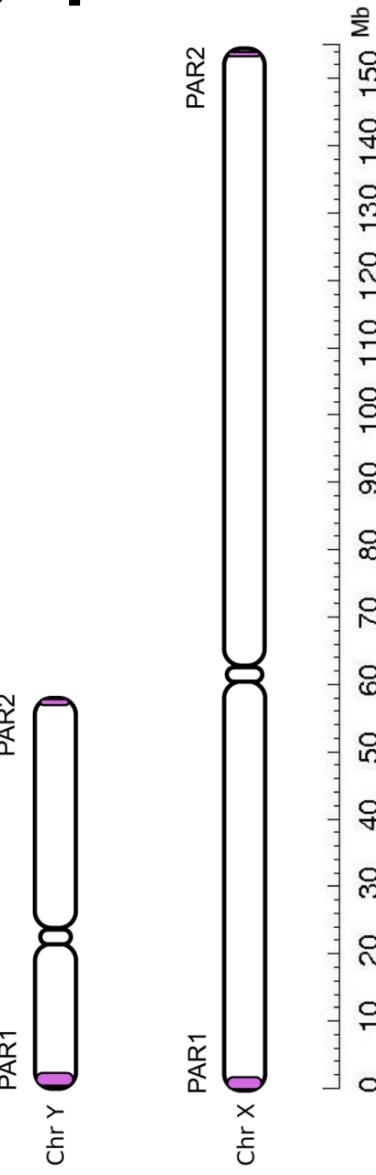
Telomeres

- Various premature ageing disease characterised by critically short telomeres
- Cancer cells often do not show telomere length shortening - telomerase protein not normally active has become active in many cancer cell types
- Telomerase and telomere-binding proteins are potential targets for anti-cancer and anti-ageing therapies



Pseudoautosomal regions on X & Y chromosomes

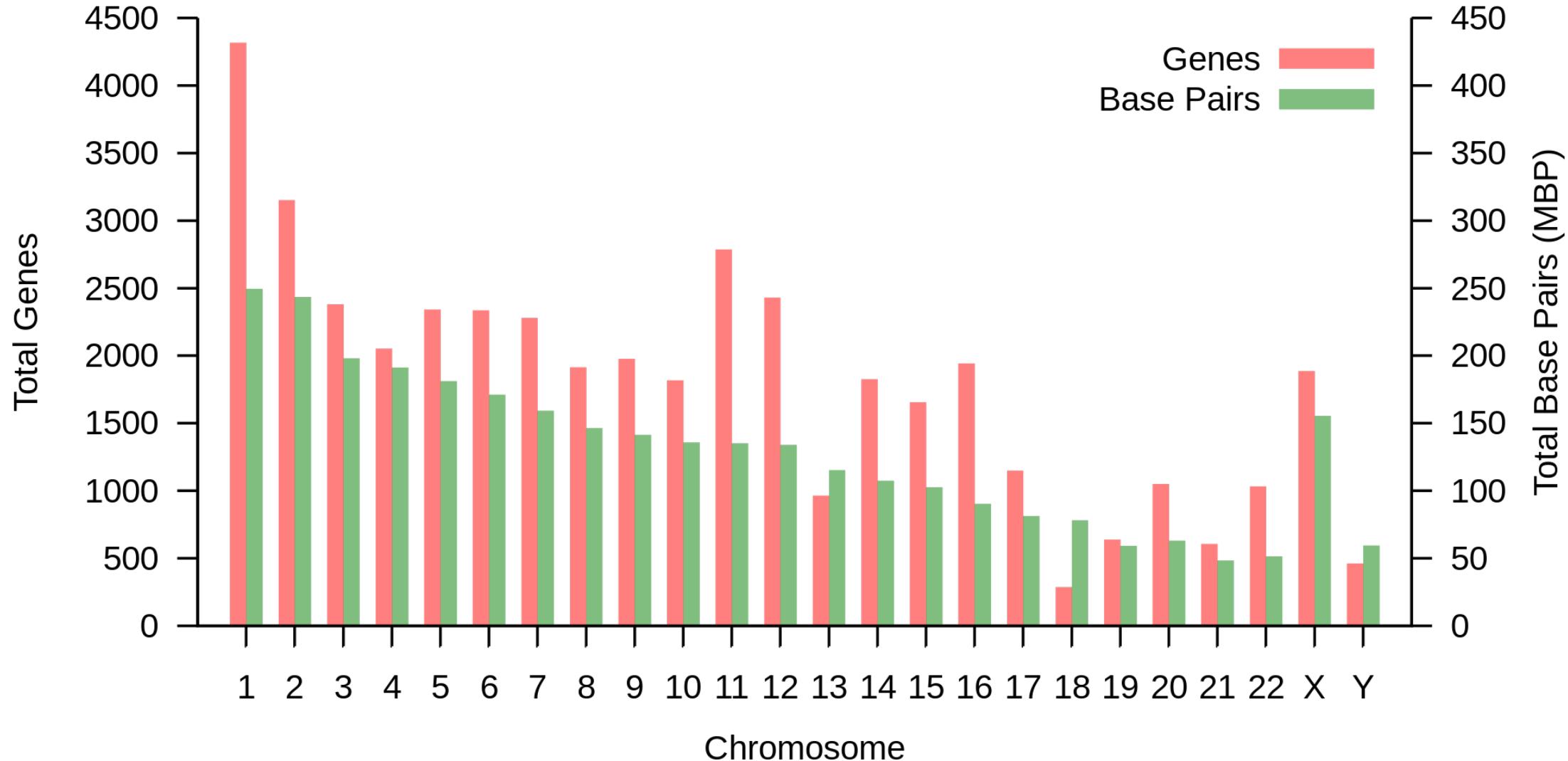
- Homologous regions at ends of X and Y chromosomes
 - Allows pairing during meiosis
 - Regions tend not to be silenced in X-inactivation
 - Inherited in an autosomal manner
 - These regions are called PAR1 and PAR2



Number of genes on each chromosome

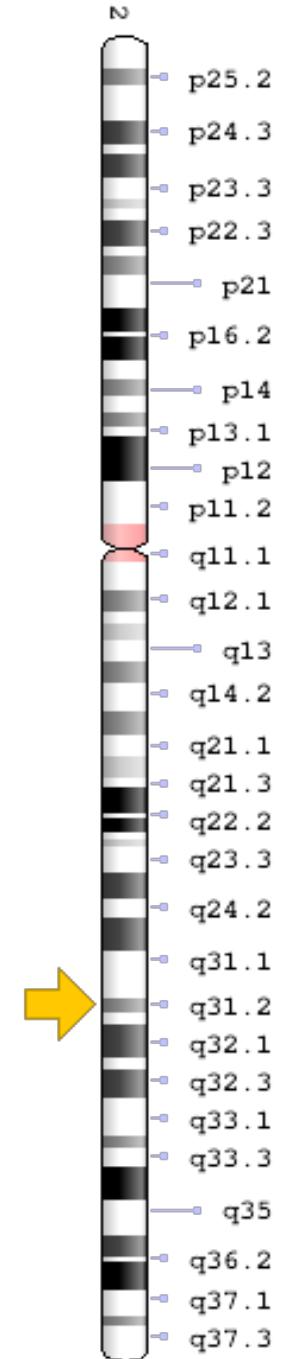
Chromosome	Length (mm)	Base pairs	Variations	Protein-coding genes	Pseudo-genes	Total long ncRNA	Total small ncRNA	miRNA	rRNA	snRNA	snoRNA	Misc ncRNA	Links	Centromere position (Mbp)	Cumulative (%)
1	85	248,956,422	12,151,146	2058	1220	1200	496	134	66	221	145	192	EBI	125	7.9
2	83	242,193,529	12,945,965	1309	1023	1037	375	115	40	161	117	176	EBI	93.3	16.2
3	67	198,295,559	10,638,715	1078	763	711	298	99	29	138	87	134	EBI	91	23
4	65	190,214,555	10,165,685	752	727	657	228	92	24	120	56	104	EBI	50.4	29.6
5	62	181,538,259	9,519,995	876	721	844	235	83	25	106	61	119	EBI	48.4	35.8
6	58	170,805,979	9,130,476	1048	801	639	234	81	26	111	73	105	EBI	61	41.6
7	54	159,345,973	8,613,298	989	885	605	208	90	24	90	76	143	EBI	59.9	47.1
8	50	145,138,636	8,221,520	677	613	735	214	80	28	86	52	82	EBI	45.6	52
9	48	138,394,717	6,590,811	786	661	491	190	69	19	66	51	96	EBI	49	56.3
10	46	133,797,422	7,223,944	733	568	579	204	64	32	87	56	89	EBI	40.2	60.9
11	46	135,086,622	7,535,370	1298	821	710	233	63	24	74	76	97	EBI	53.7	65.4
12	45	133,275,309	7,228,129	1034	617	848	227	72	27	106	62	115	EBI	35.8	70
13	39	114,364,328	5,082,574	327	372	397	104	42	16	45	34	75	EBI	17.9	73.4
14	36	107,043,718	4,865,950	830	523	533	239	92	10	65	97	79	EBI	17.6	76.4
15	35	101,991,189	4,515,076	613	510	639	250	78	13	63	136	93	EBI	19	79.3
16	31	90,338,345	5,101,702	873	465	799	187	52	32	53	58	51	EBI	36.6	82
17	28	83,257,441	4,614,972	1197	531	834	235	61	15	80	71	99	EBI	24	84.8
18	27	80,373,285	4,035,966	270	247	453	109	32	13	51	36	41	EBI	17.2	87.4
19	20	58,617,616	3,858,269	1472	512	628	179	110	13	29	31	61	EBI	26.5	89.3
20	21	64,444,167	3,439,621	544	249	384	131	57	15	46	37	68	EBI	27.5	91.4
21	16	46,709,983	2,049,697	234	185	305	71	16	5	21	19	24	EBI	13.2	92.6
22	17	50,818,468	2,135,311	488	324	357	78	31	5	23	23	62	EBI	14.7	93.8
X	53	156,040,895	5,753,881	842	874	271	258	128	22	85	64	100	EBI	60.6	99.1
Y	20	57,227,415	211,643	71	388	71	30	15	7	17	3	8	EBI	12.5	100
mtDNA	0.0054	16,569	929	13	0	0	24	0	2	0	0	0	EBI	N/A	100
total		3,088,286,401	155,630,645	20412	14600	14727	5037	1756	532	1944	1521	2213			

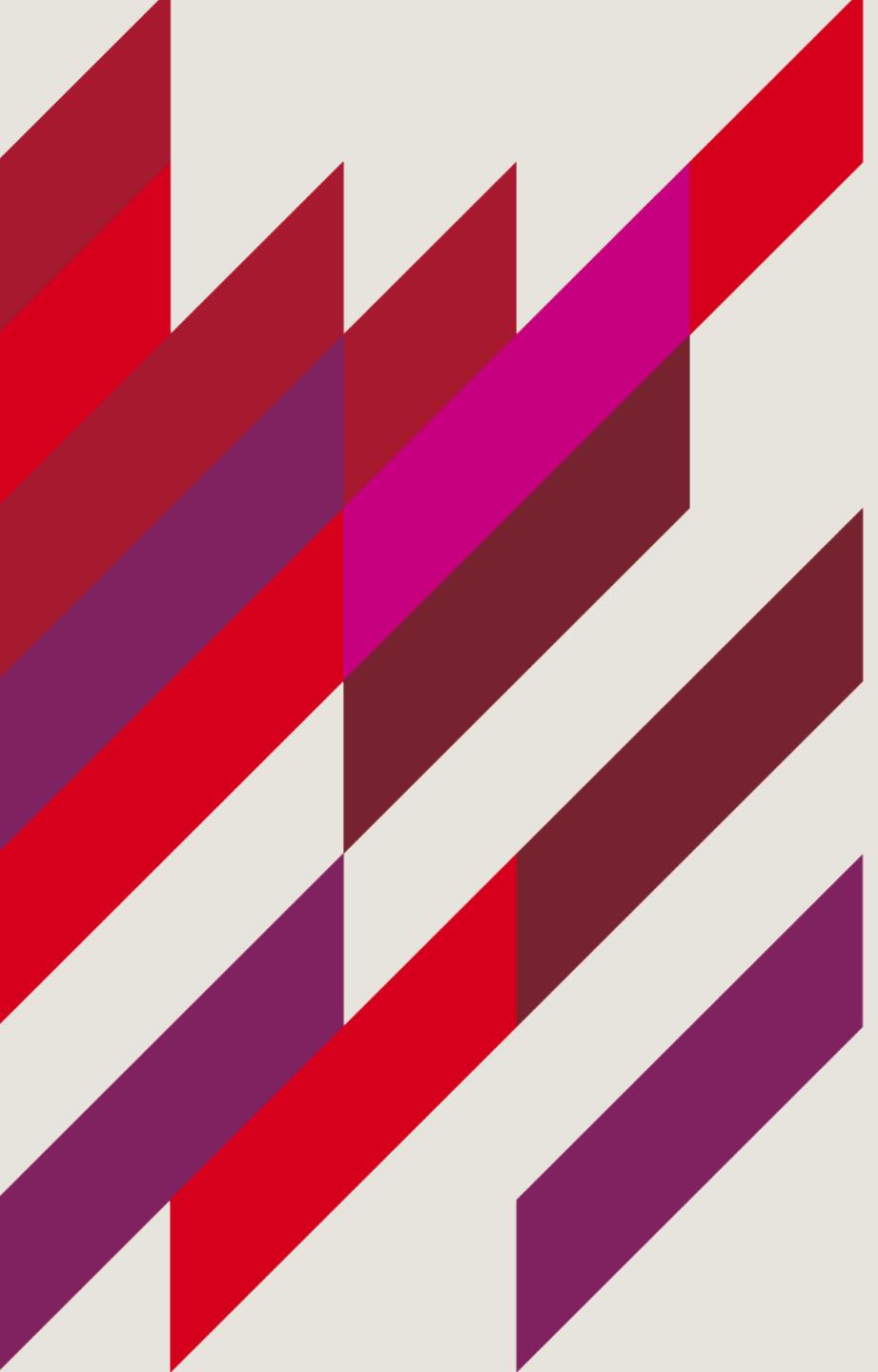
Chromosome size vs number of genes



Gene size in humans

- Titin gene (TTN)
 - 365,719bp total gene size (including introns)
 - Exons: 80,781bp across 365 exons
- DMD gene (Duchenne muscular dystrophy)
 - 2,220,390 bp including introns
 - ~14kb mRNA
- Average gene size ~8kb



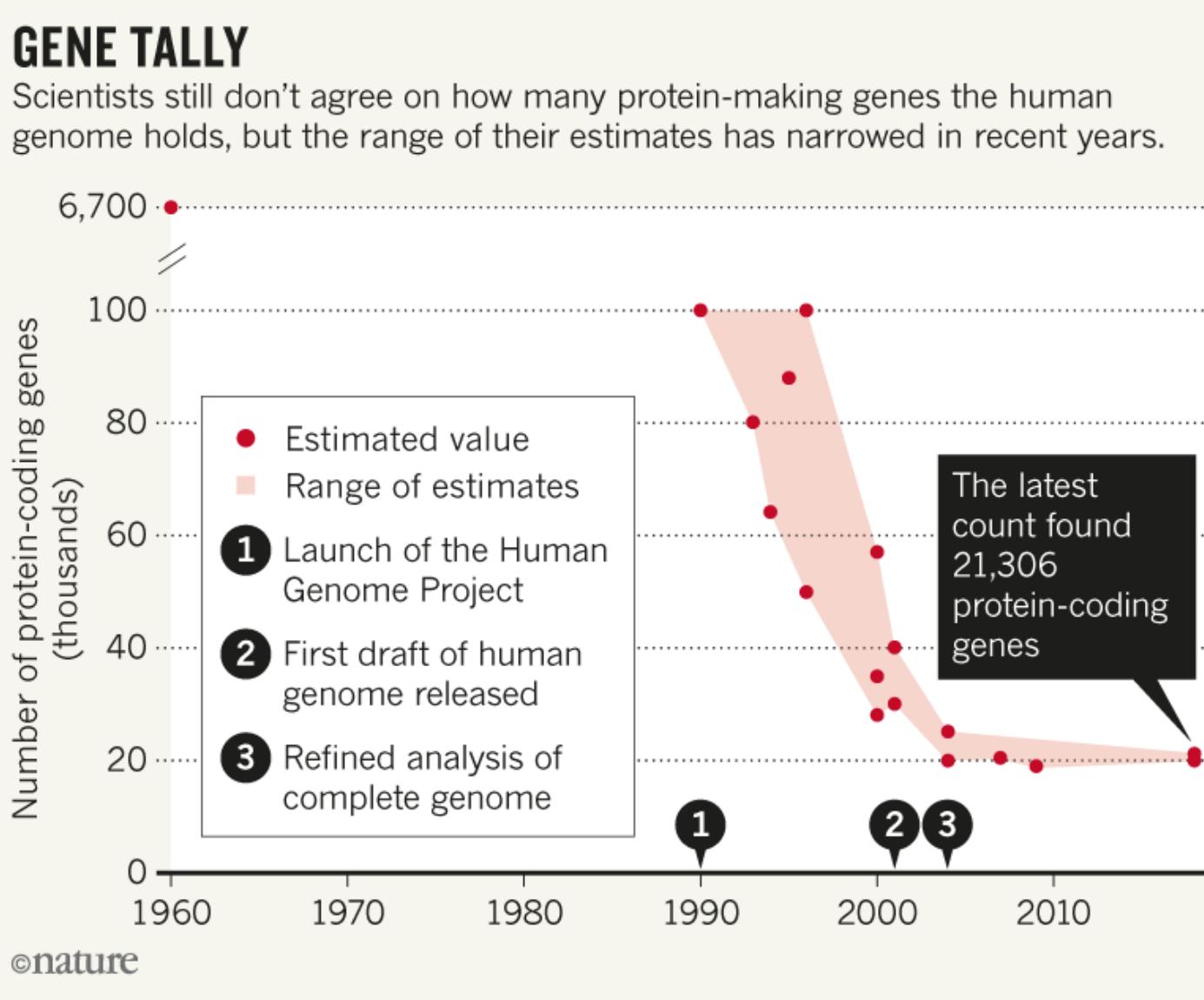


Coding Regions Human Genome

Human genome summary

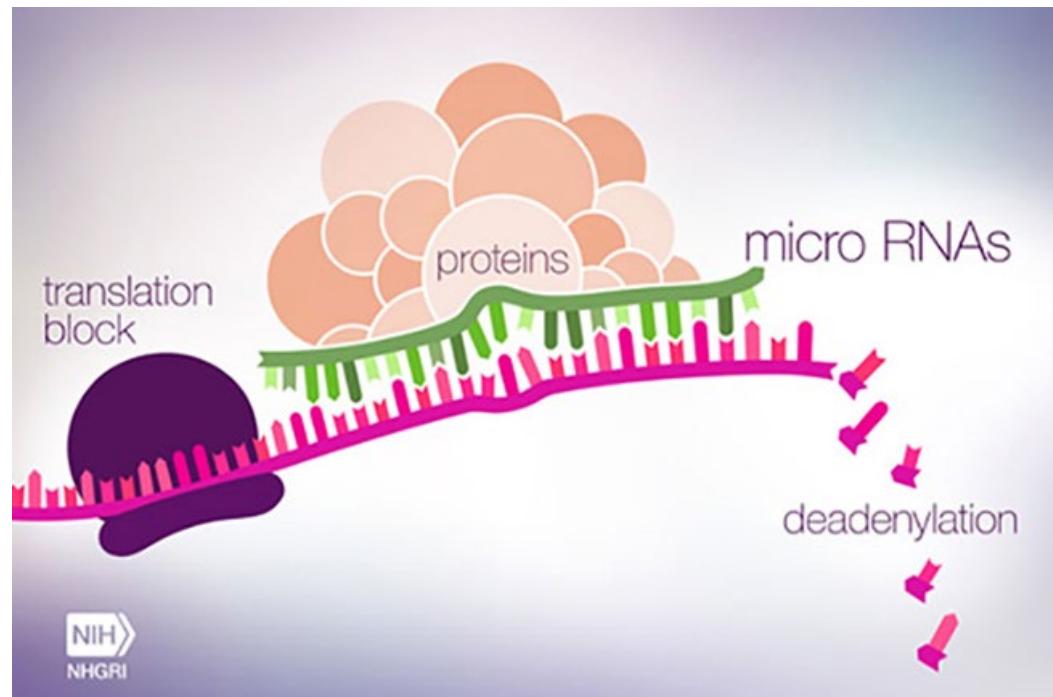
- 3.2 billion base pairs x 2 across 23 chromosome pairs
- ~20,000 protein coding genes, but other functional molecules
- Only 1-2% codes for protein, but up to 80% of genome regulatory function

How many genes?

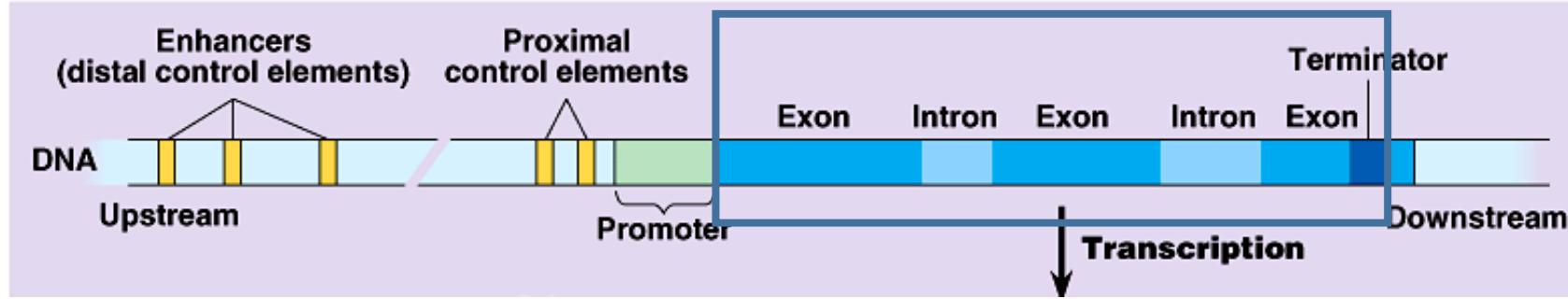


What is a gene?

- Classically, region of DNA which codes for a protein
- Nowadays:
 - a sequence of nucleotides in DNA that codes for a molecule that has a function
 - This might code for a protein, or just be transcribed into RNA e.g. t-RNAs, non-coding RNAs

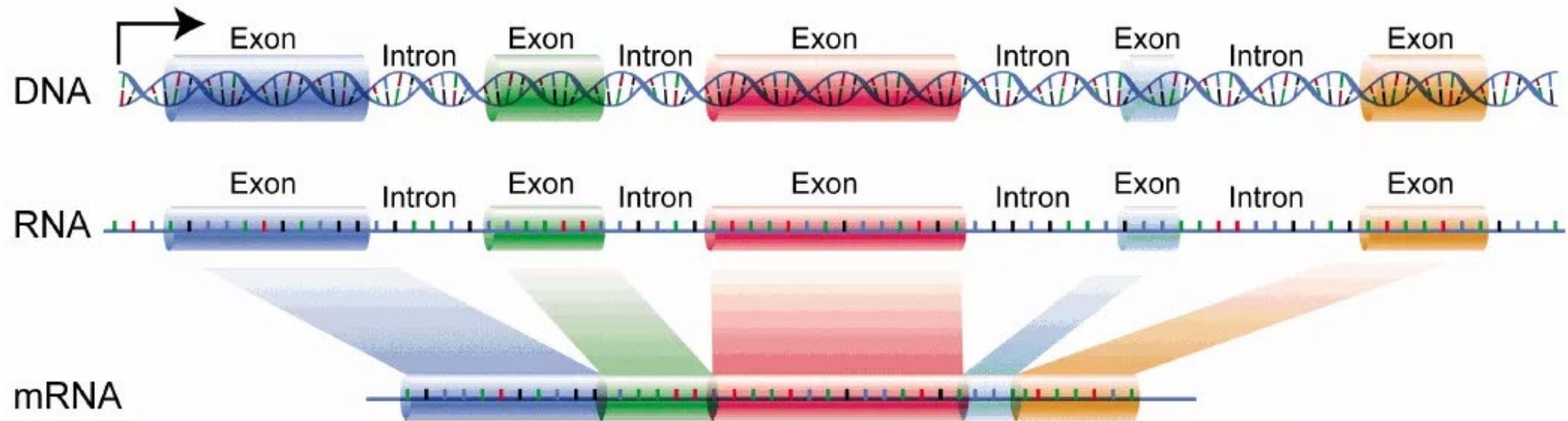


What is a gene?



Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

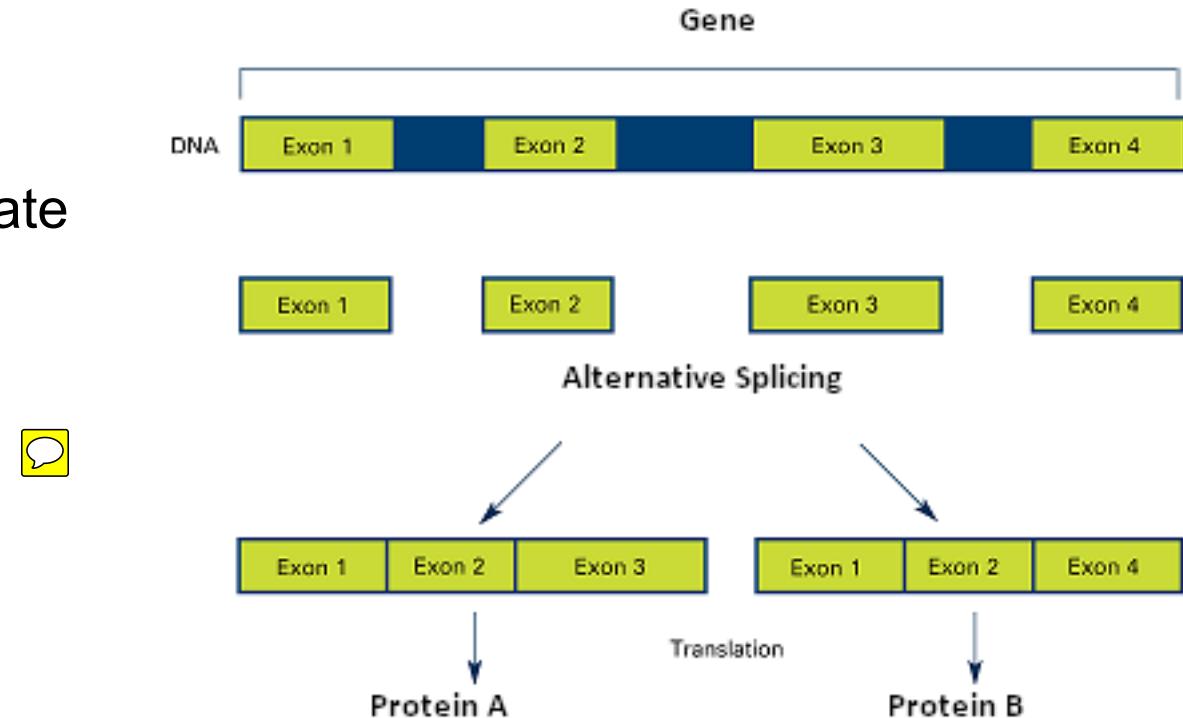
Exons and introns



- Exon: part of a gene that will encode a part of the final mature RNA
- Intron: any nucleotide sequence within a gene that is removed by RNA splicing during maturation of the final RNA product
- 1.1% of the genome is spanned by exons, whereas 24% is in introns, with 75% of the genome being intergenic DNA

Multiple transcripts / alternate splicing

- Exons can be alternatively arranged to create different proteins from the same gene



- Different types of cells often use different transcripts more/less frequently

The Human Genome

LEARNING OBJECTIVES



At the end of this lecture you should be able to:

- Define the terms used to describe the human genome
- Understand the structure of chromosomes
- Understand the coding regions of the human genome



MACQUARIE
University
SYDNEY · AUSTRALIA

Solving genetic problems

BIOL3120 – LECTURE 6



Where to find the problem sets?

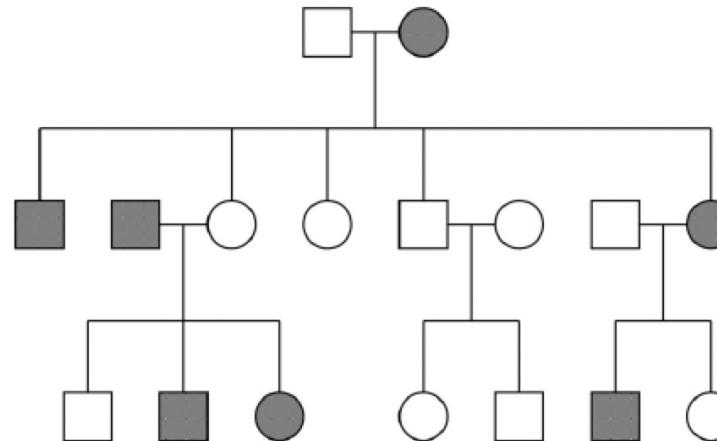


iLearn!

In the assessments tab

AT1

PROBLEM SET ASSESSMENT



1. What is the *most likely* mode of inheritance for the condition shown in the above pedigree? Why?
 - a) Autosomal dominant
 - b) Autosomal recessive
 - c) X-linked dominant
 - d) X-linked recessive

Special considerations



MACQUARIE
University
SYDNEY · AUSTRALIA

Student wellbeing / special considera... ▾

If you experience events or conditions that adversely affect your academic performance, please visit the [special consideration page](#).

Visit the [Student Support](#) webpage to see a range of services available to students at Macquarie University.

If you feel you are struggling with studies or have had disruptions in your personal life, you can speak to a trained professional at [Campus Wellbeing](#).

Your job as geneticists is to work out which tools you can use

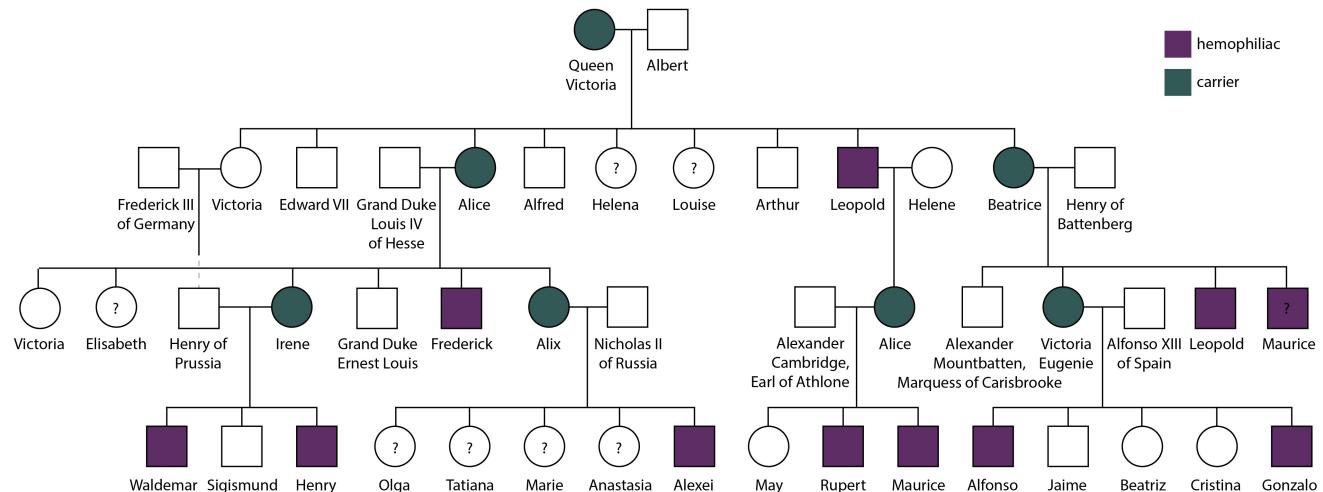
You already know all the tools to solve genetic problems

- Using family trees to infer inheritance patterns
- Hardy-Weinberg equilibrium
- Punnett squares to find expected phenotypes
- Simple probability



Using family trees to infer inheritance patterns

- Useful for inferring inheritance patterns
- Useful for inferring individual's genotype



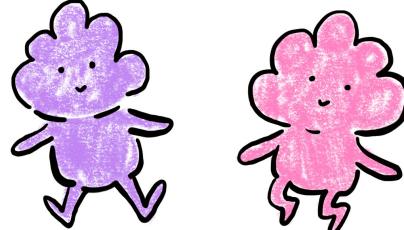
Hardy-Weinberg equilibrium

Hardy-Weinberg equilibrium

If there are only 2 alleles for a trait in a Population, then:

$$P + q = 1$$

frequency of dominant allele frequency of recessive allele



Purple is dominant to Pink

Hardy-Weinberg equilibrium

If there are only 2 alleles for a trait in a Population, then:

$$P^2 + 2Pq + q^2 = 1$$

frequency of homozygous dominant genotype frequency of heterozygous genotype frequency of homozygous recessive genotype

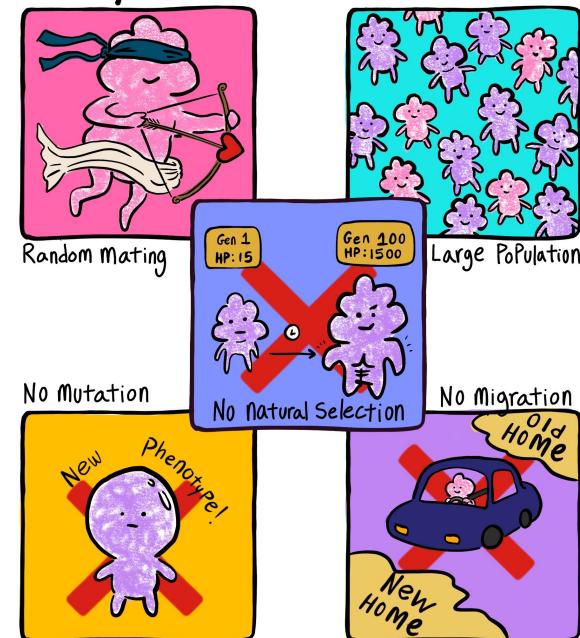


Purple is dominant to Pink

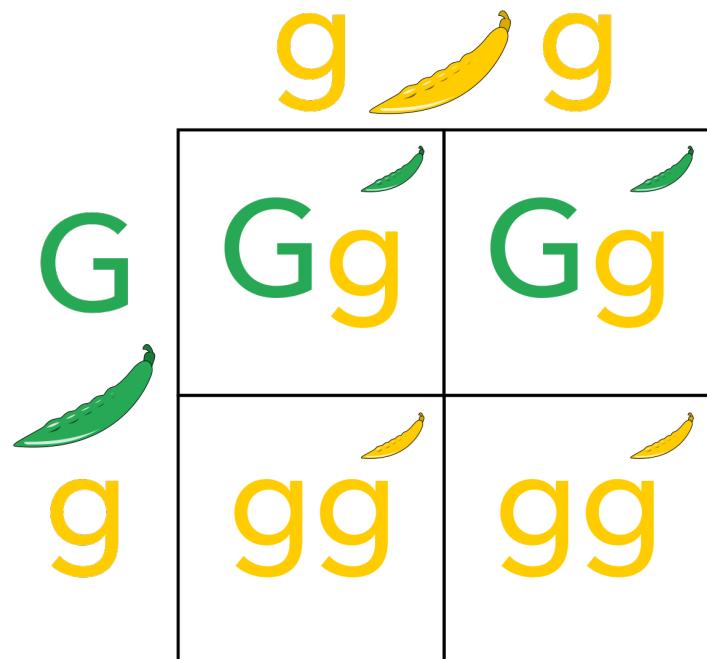
Hardy-Weinberg equilibrium

- Assumptions of Hardy-Weinberg equilibrium
- 1. Large population (avoids random chance having a big impact on allele frequencies).
- 2. No selection (ie no allele is necessarily more likely to get passed on than another)
- 3. Mating is random (ie any individual has an equal chance of mating with any other individual).
- 4. Mutation either does not occur or is in equilibrium.
- 5. Immigration and emigration do not occur.

Hardy-Weinberg Assumptions



Punnett squares to find expected phenotypes



Punnett Square
in Pea plant

Round Yellow (RrYy)

Round Green (Rryy)

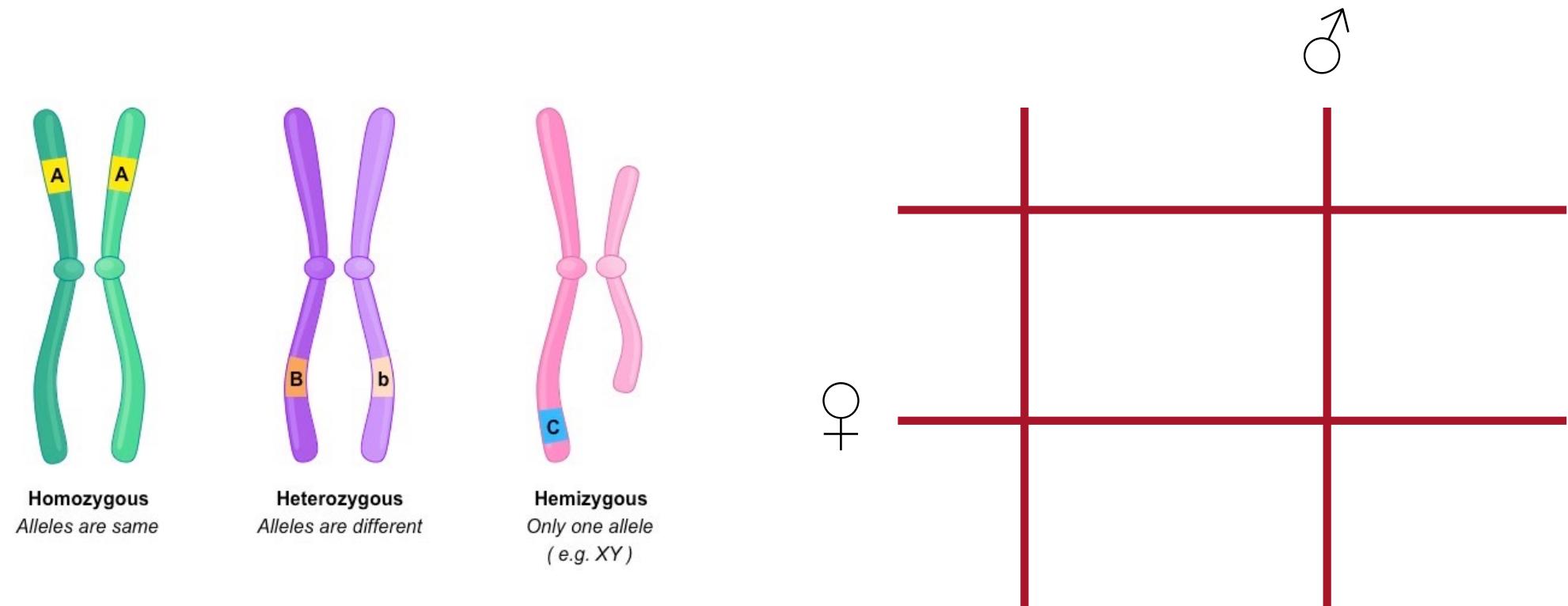
Wrinkled Yellow (rrYy)

Wrinkled Green (rryy)

RY	Ry	rY	ry	
RY	RYY ●	RRYy ●	RrYY ●	RrYy ●
Ry	RRYy ●	RRyy ●	RrYy ●	Rryy ●
rY	RrYY ●	RrYy ●	rrYy ●	rrYy ●
ry	RrYy ●	Rryy ●	rrYy ●	rryy ●

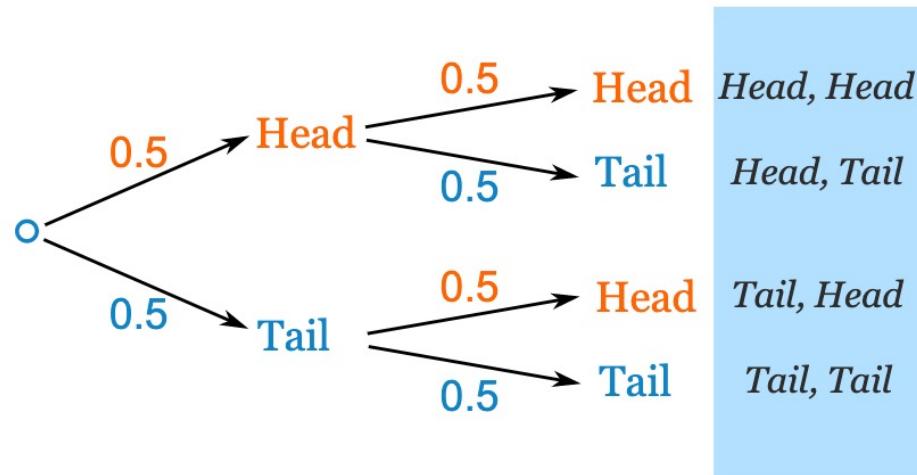
F2 Generation

Punnett squares to find expected phenotypes



Simple probability

- If you need to think about a probability problem, draw a probability tree



Solving genetic problems



LEARNING OBJECTIVES



At the end of this lecture you should be able to:

- List the tools available to solve genetic problems
- Think about which tools may be useful in certain circumstances
- Be prepared to work through genetic problems in week 3 and beyond



BIOL3120 –Human Genetics and Evolutionary Medicine

Heritability and Polygenic Inheritance



BIOL3120 –Heritability and Polygenic Inheritance

LEARNING OBJECTIVES



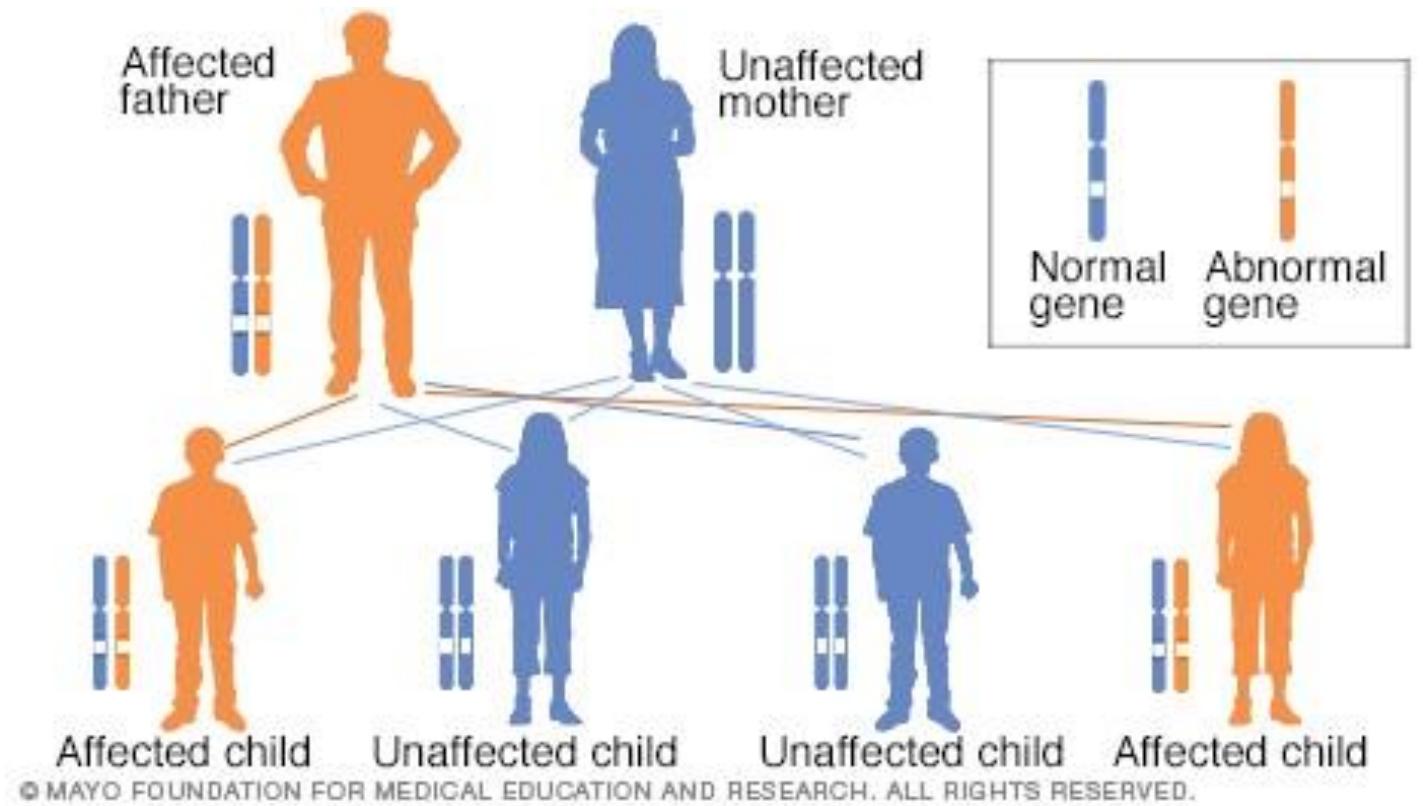
On successful completion of this lecture, you will be able to:

- Discuss factors that indicate a possible genetic basis for a condition
- Define heritability & understand it's limitations
- Discuss polygenic inheritance and polygenic risk scores

Terms

- Monogenic inheritance
 - Caused by variation in a single gene
 - Striking familial inheritance patterns
 - Sickle cell anaemia, cystic fibrosis, Huntington disease, Duchenne muscular dystrophy
- Polygenic inheritance
 - Caused by the combined action of more than one gene
 - Hypertension, coronary heart disease, diabetes
- Environmental factors
 - Complex diseases occur as a result of many genomic variants, paired with environmental influences

Monogenic inheritance



Polygenic inheritance



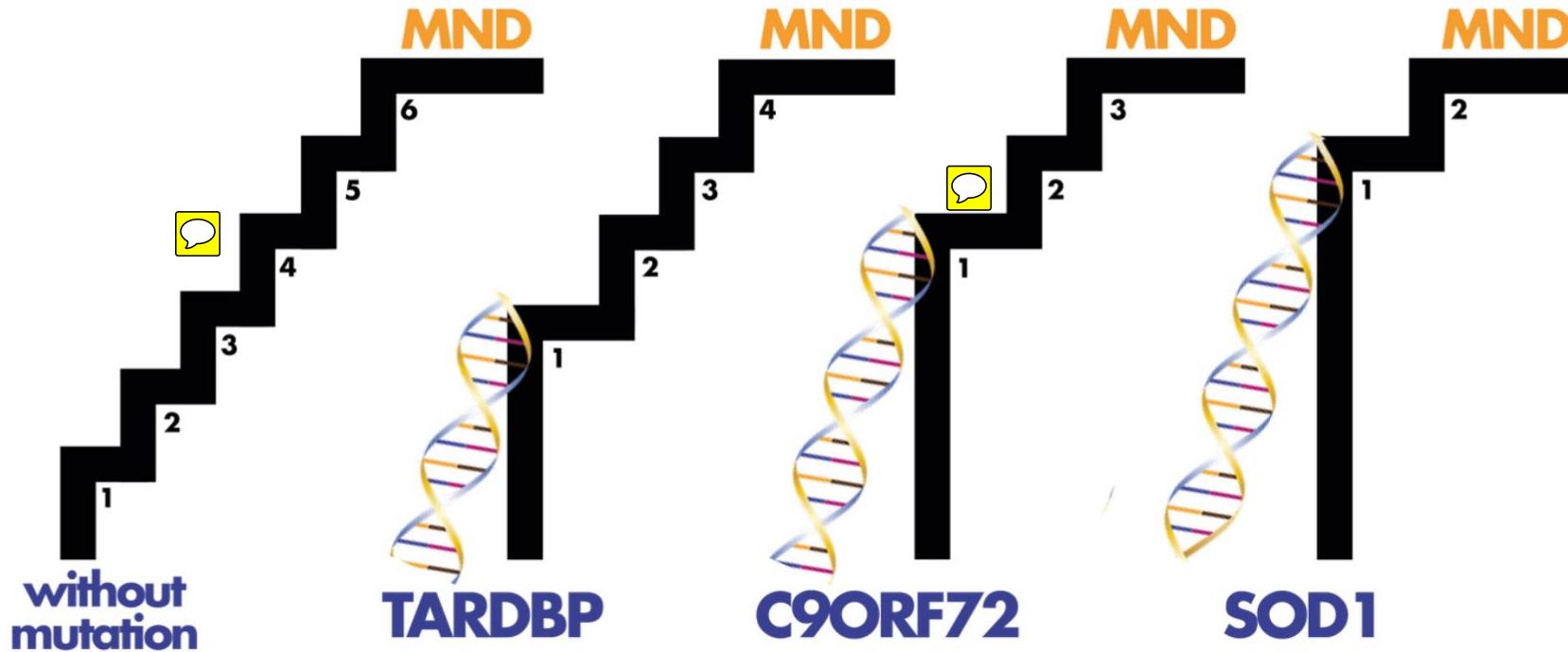
Human height.

There is great variation in human height between different individuals.

Jane Ades/National Human Genome Research Institute.

Complex disorders

Genetic and environmental risk factors



CAUSES AND DISEASE MECHANISMS / MND RESEARCH

Steps to understanding MND

© AUGUST 3, 2018

NICKJAMESCOLE

2 COMMENTS

THE OFFICIAL BLOG OF THE



Al-Chalbi *et al* Lancet Neurol; 13: 1108–13, 2014
Chio et al Neurology :e1-e8, 2018

How to tell how ‘genetic’ a disease is: Relative risk ratio

$$\lambda_r = \frac{\text{Prevalence of the disease in the relatives of an affected person}}{\text{Prevalence of the disease in the general population}}$$

TABLE 8-2 Risk Ratios λ_s for Siblings of Probands with Diseases with Familial Aggregation and Complex Inheritance

Disease	Relationship	λ_s
Schizophrenia	Siblings	12
Autism	Siblings	150
Manic-depressive (bipolar) disorder	Siblings	7
Type 1 diabetes mellitus	Siblings	35
Crohn disease	Siblings	25
Multiple sclerosis	Siblings	24

Value of 1 = no more likely to develop condition if you have an affected relative

Value higher than 1 = relative of affected person more likely to develop the condition 

How to tell how ‘genetic’ a disease is: Family history case-controlled studies

- Multiple sclerosis (MS) study example: 3.5% of first-degree relatives of people with MS also had MS themselves.
- Compared to 0.2% of first-degree relatives of people without MS
- $3.5 / 0.2 = 17.5$
- The odds of having a first degree relative with MS were 17.5x higher amongst MS patients than controls.
- Suggests genetic factors underlying

Problem:

- People who are related tend to have environmental factors in common.
- How do we tease apart genetic vs. environmental influences?



Heritability

- Define heritability & understand it's limitations

What is Heritability?

- A measure of how well differences in people's genes account for differences in their traits
 - Popularly referred to as 'nature versus nurture' debate
- Heritability is a statistical concept that describes how much of the variation in a given trait can be attributed to genetic variation
- Used in reference to the resemblance between parents and their offspring. In this context, high heritability implies a strong resemblance between parents and offspring with regard to a specific trait, while low heritability implies a low level of resemblance.
- An estimate of the heritability of a trait is specific to one population in one environment, and it can change over time as circumstances change.

What is Heritability?

- Heritability estimates range from zero to one.
- A heritability close to one indicates that almost all of the variability in a trait comes from genetic differences, with very little contribution from environmental factors.
 - Many disorders that are caused by mutations in single genes, such as phenylketonuria (PKU), have high heritability
- A heritability close to zero indicates that almost all of the variability in a trait among people is due to environmental factors, with very little influence from genetic differences.
 - Characteristics such as religion, language spoken, and political preference have a heritability of zero because they are not under genetic control.
- Most complex traits in people, such as multifactorial diseases, have a heritability somewhere in the middle, suggesting that their variability is due to a combination of genetic and environmental factors.

Heritability limitations

- Heritability does not indicate what proportion of a trait is determined by genes and what proportion is determined by environment.
 - a heritability of 0.7 does not mean that a trait is 70% caused by genetic factors; it means than 70% of the variability in the trait in a population is due to genetic differences among people.
- Heritability does not determine which genes/loci are involved in the trait/disease.
- Estimating Trait Heritability



विवरणीय

Heritability



Khan Academy

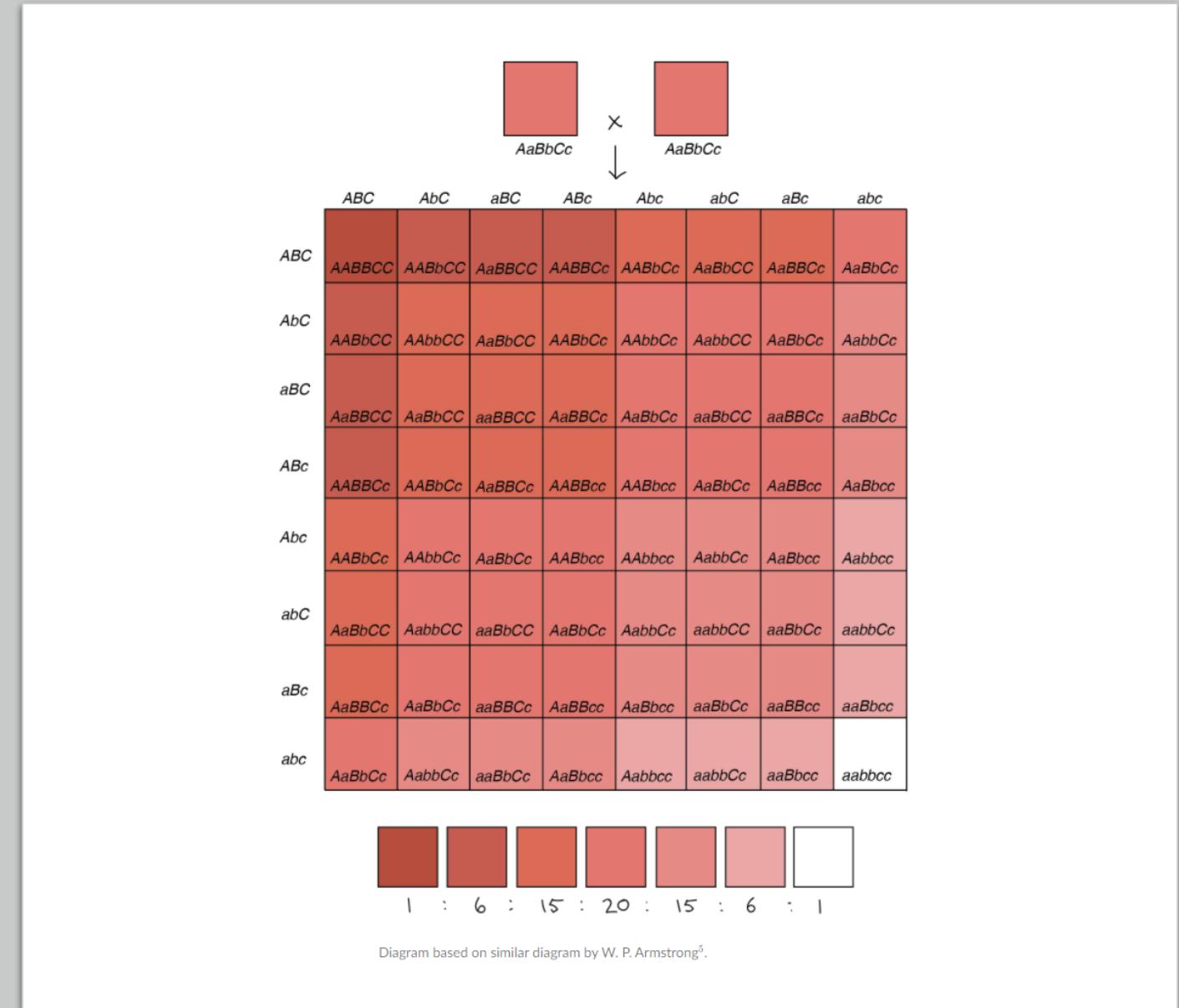


Polygenic Inheritance

- Discuss polygenic inheritance and polygenic risk scores

Polygenic inheritance

- Caused by the combined action of more than one gene
 - Height, skin colour
 - Hypertension, coronary heart disease, diabetes
- The inheritance of polygenic traits does not show the phenotypic ratios characteristic of Mendelian inheritance, though each of the genes contributing to the trait is inherited as described by Gregor Mendel.



Genetic susceptibility to complex disease

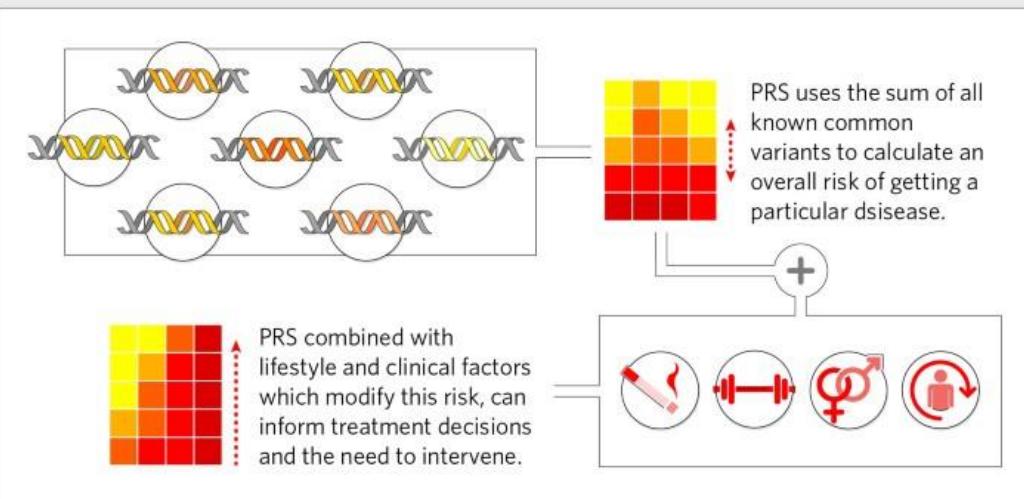
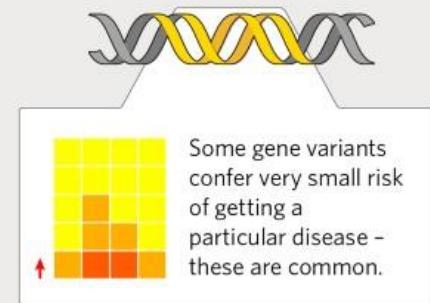
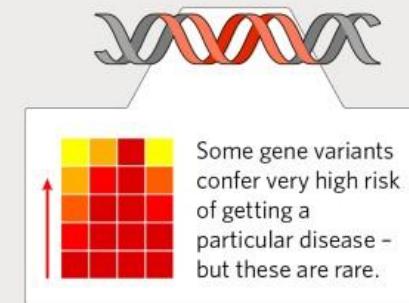
- **Complex disease:** Most medical problems such as heart disease, diabetes, obesity, Alzheimer's disease, asthma, Parkinson's disease, multiple sclerosis, osteoporosis, and sporadic MND, do not have a single genetic cause—they are likely associated with the effects of **multiple genes** in combination with **lifestyle and environmental factors**. These are called complex or multifactorial disorders.
- **Susceptibility (risk) alleles:** an allele, usually inherited, that increases the likelihood of developing a complex disease. The combination of multiple susceptibility alleles and environmental factors may be additive or synergistic, leading to disease.
 - On their own, they are neither necessary nor sufficient to cause disease

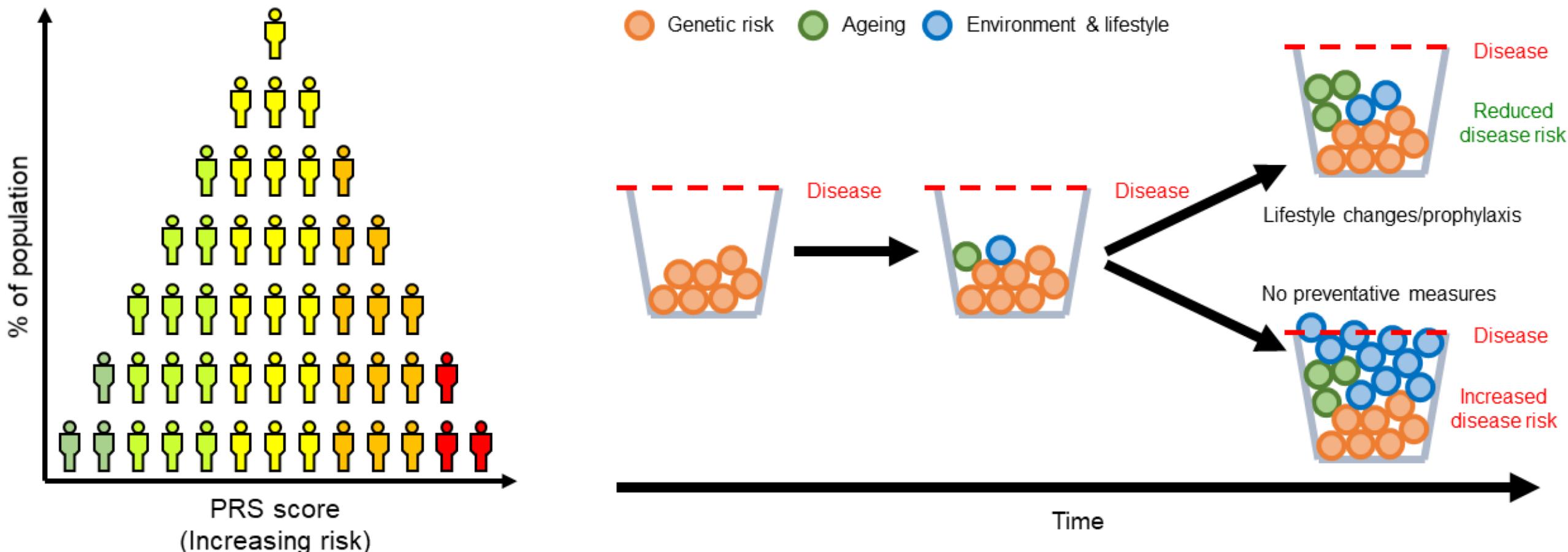
Polygenic Risk Score

- “a substitution of a single nucleotide that occurs at a specific position in the genome, where each variation is present to some appreciable degree within a population (e.g. > 1%)”
(Wikipedia)
- 4-5 million per person ~ every 1,000bp on average

CLINICAL APPLICATION OF PRS

A polygenic risk score (PRS) is calculated from many small genetic variants, and can often be modified by lifestyle factors.





Hall A, Bandres-Ciga S, Diez-Fairen M, Quinn JP, Billingsley KJ. Genetic Risk Profiling in Parkinson's Disease and Utilizing Genetics to Gain Insight into Disease-Related Biological Pathways. *International Journal of Molecular Sciences*. 2020; 21(19):7332. <https://doi.org/10.3390/ijms21197332>

BIOL3120 –Heritability and Polygenic Inheritance

LEARNING OBJECTIVES



On successful completion of this lecture, you will be able to:

- Discuss factors that indicate a possible genetic basis for a condition
- Define heritability & understand it's limitations
- Discuss polygenic inheritance and polygenic risk scores



MACQUARIE
University

BIOL3120 –Human Genetics and Evolutionary Medicine

Chromosomal Mutations

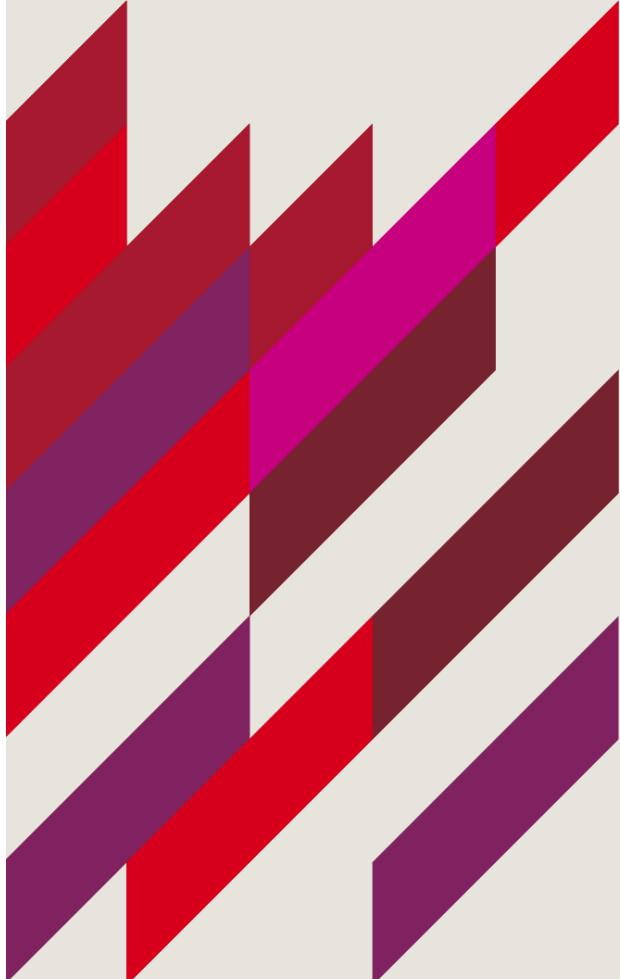




3	The Human Genome Modes of Inheritance and Population Genetics	Problem Set 1	Problem Set 1 (5%)	Explain the principles of evolutionary biology and their role in human health and disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources
4	Heritability and Polygenics Chromosomal Mutations	Problem Set 2	Problem Set 2 (5%)	Explain the principles of evolutionary biology and their role in human health and disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources
5	Nucleotide Mutations Human Genetic Diversity and Evolution	Problem Set 3	Problem Set 3 (5%)	Explain the principles of evolutionary biology and their role in human health and disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources

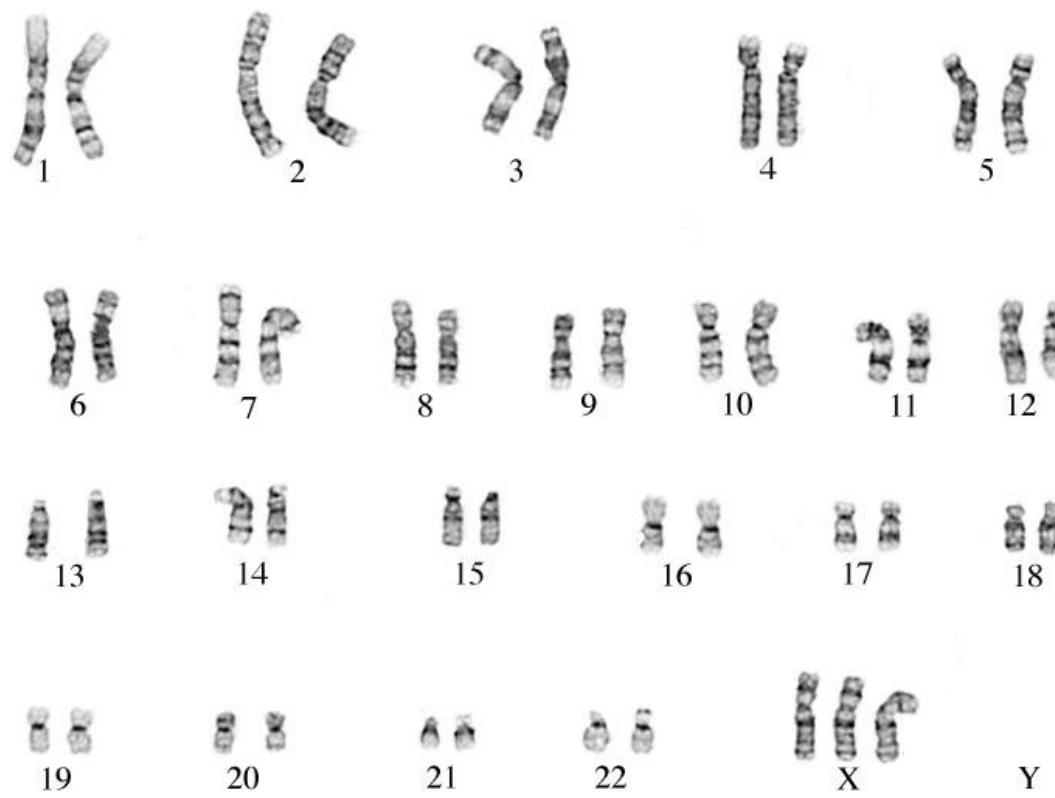
BIOL3120 –Chromosomal mutations

LEARNING OBJECTIVES



On successful completion of this lecture, you will be able to:

- Explain the different types of chromosomal mutations
- Use this knowledge to solve problems in human genetics relating to heritability, polygenic inheritance and chromosomal mutations



Chromosomal mutations overview

- Mosaicism
- Aneuploidy / other euploidies
- Uniparental disomy
- Translocations + Robertsonian translocations
- Changes within a chromosome



Chromosomal mutations

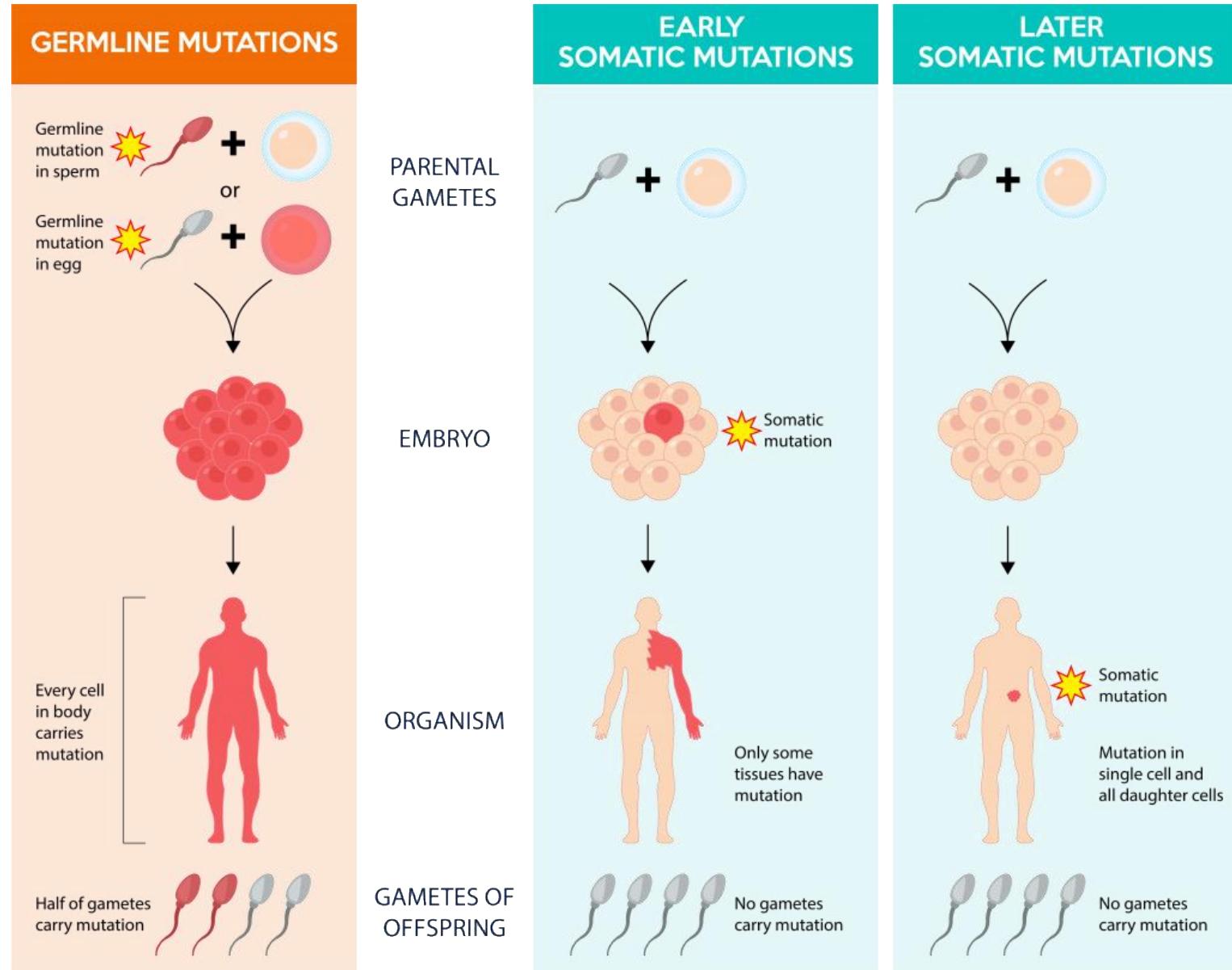
- Mosaicism
- Aneuploidy / other euploidies
- Uniparental disomy

Mosaicism

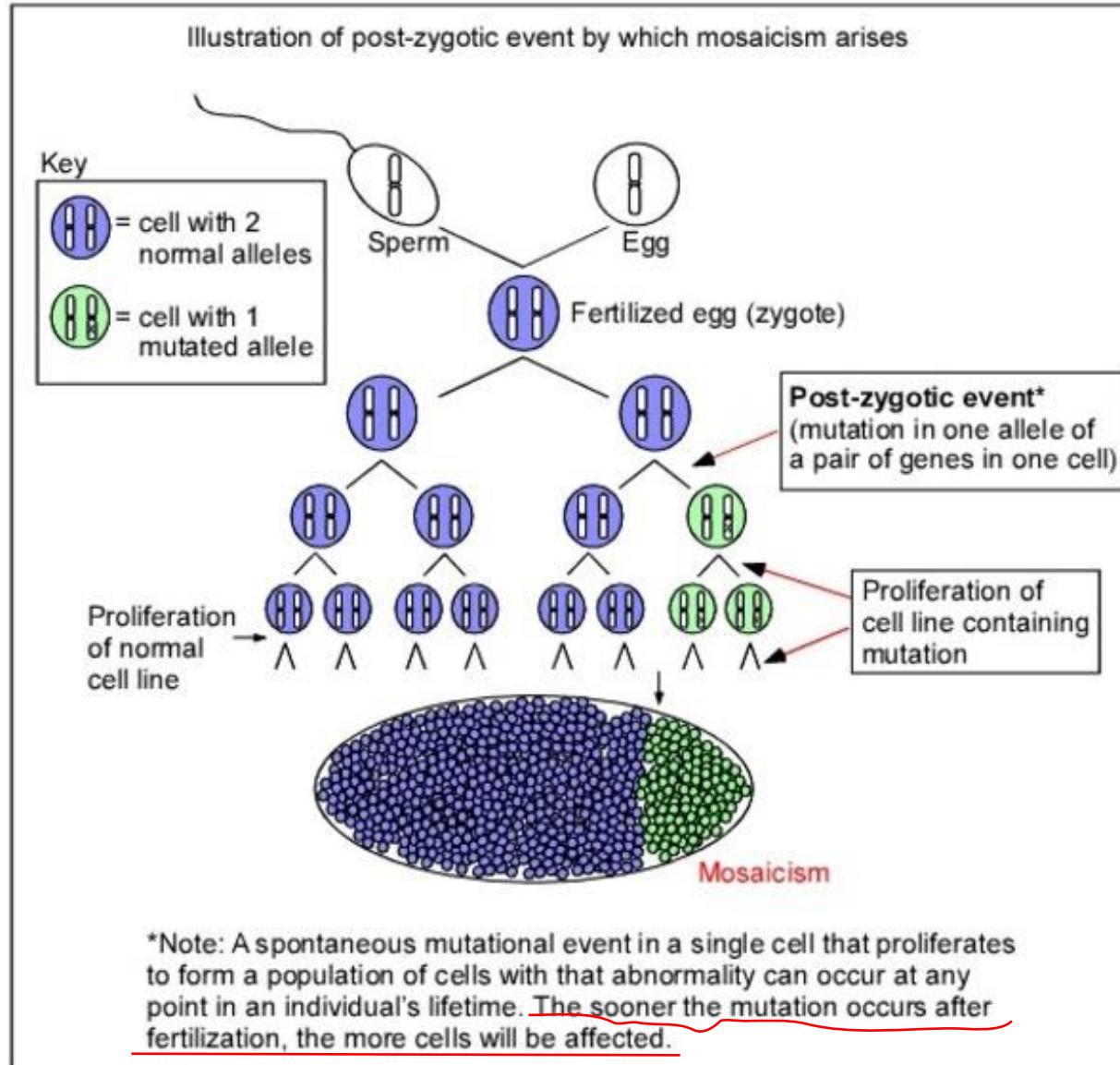


- Mosaicism = two sets of genomes in an organism
- Somatic mosaicism refers to the occurrence of two genetically distinct populations of cells within an individual, derived from a postzygotic mutation.
- In contrast to inherited mutations, somatic mosaic mutations may affect only a portion of the body and are not transmitted to progeny.
- These mutations affect varying genomic sizes ranging from single nucleotides to entire chromosomes and have been implicated in disease, most prominently cancer.

What cell is a mutation happening in, and when?



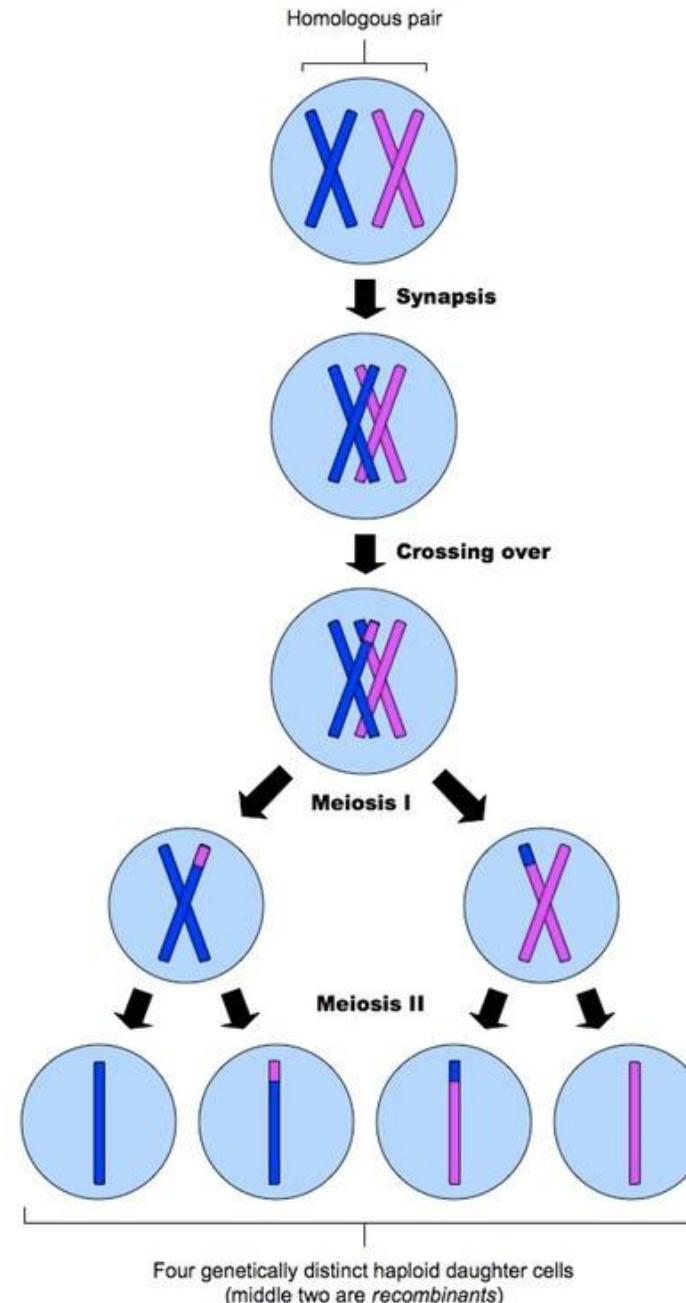
Mosaicism = two sets of genomes in an organism



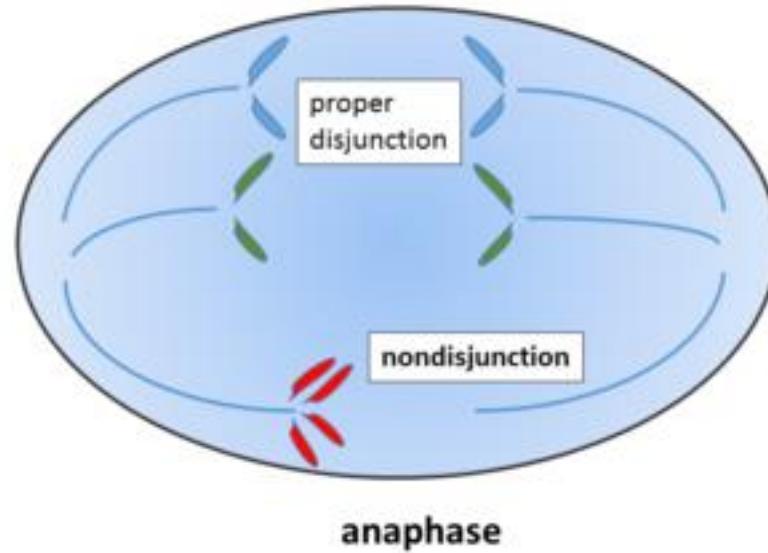
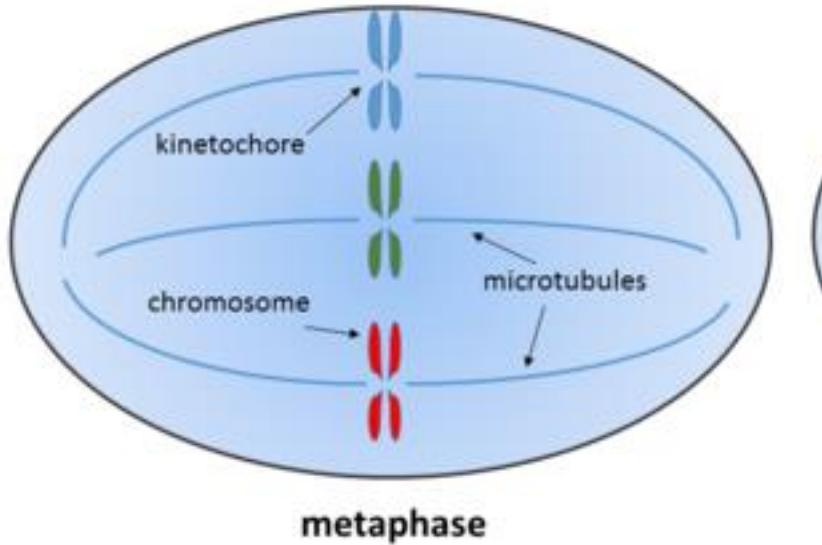
Aneuploidy

- Aneuploidy = abnormal number of chromosomes in a cell
- Aneuploidy usually caused by nondisjunction
 - Failure of homologous chromosomes to separate properly during cell division

Normal meiosis

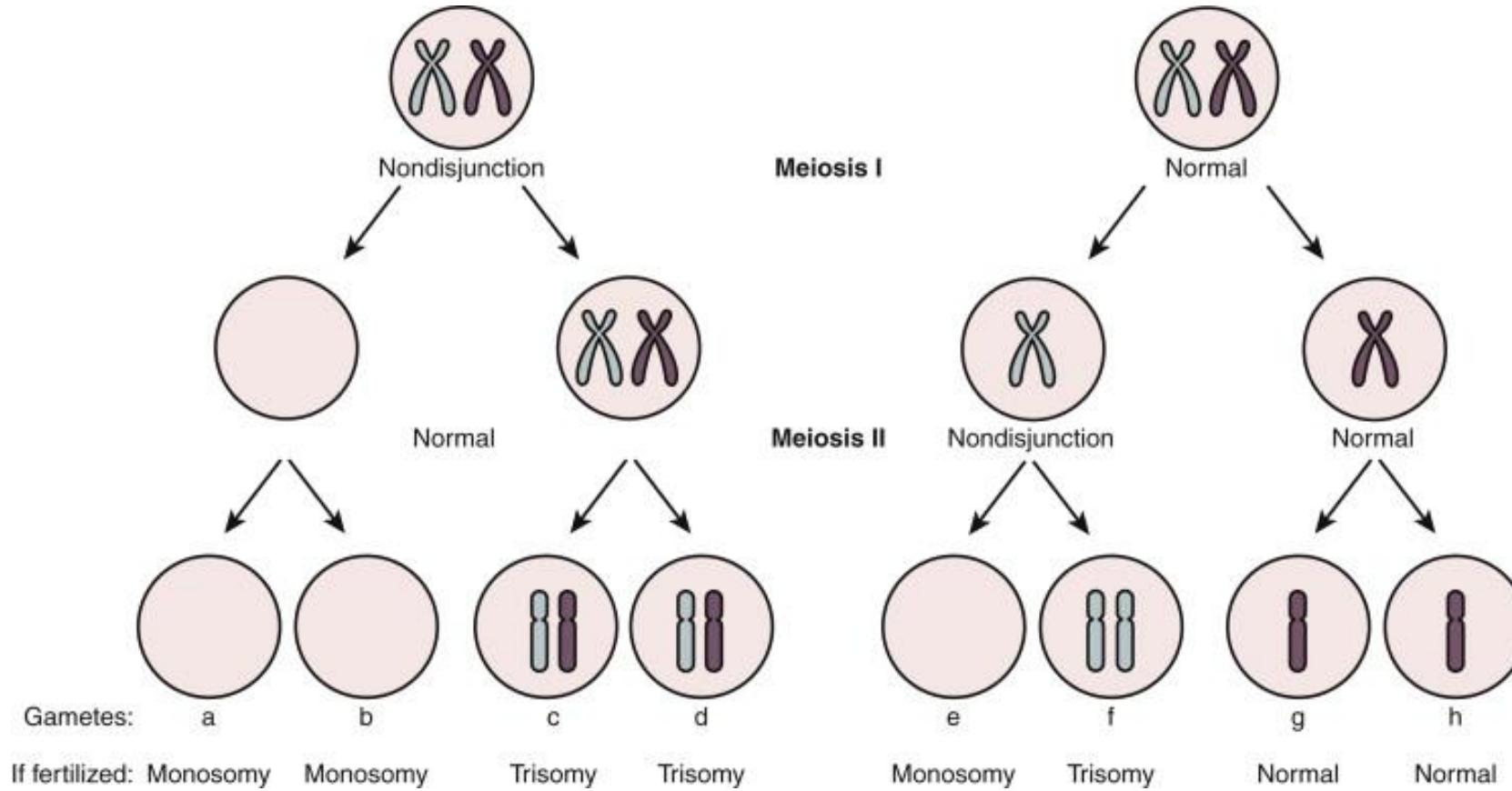


Nondisjunction



Can occur during mitosis, but only gets passed onto offspring if occurs during meiosis

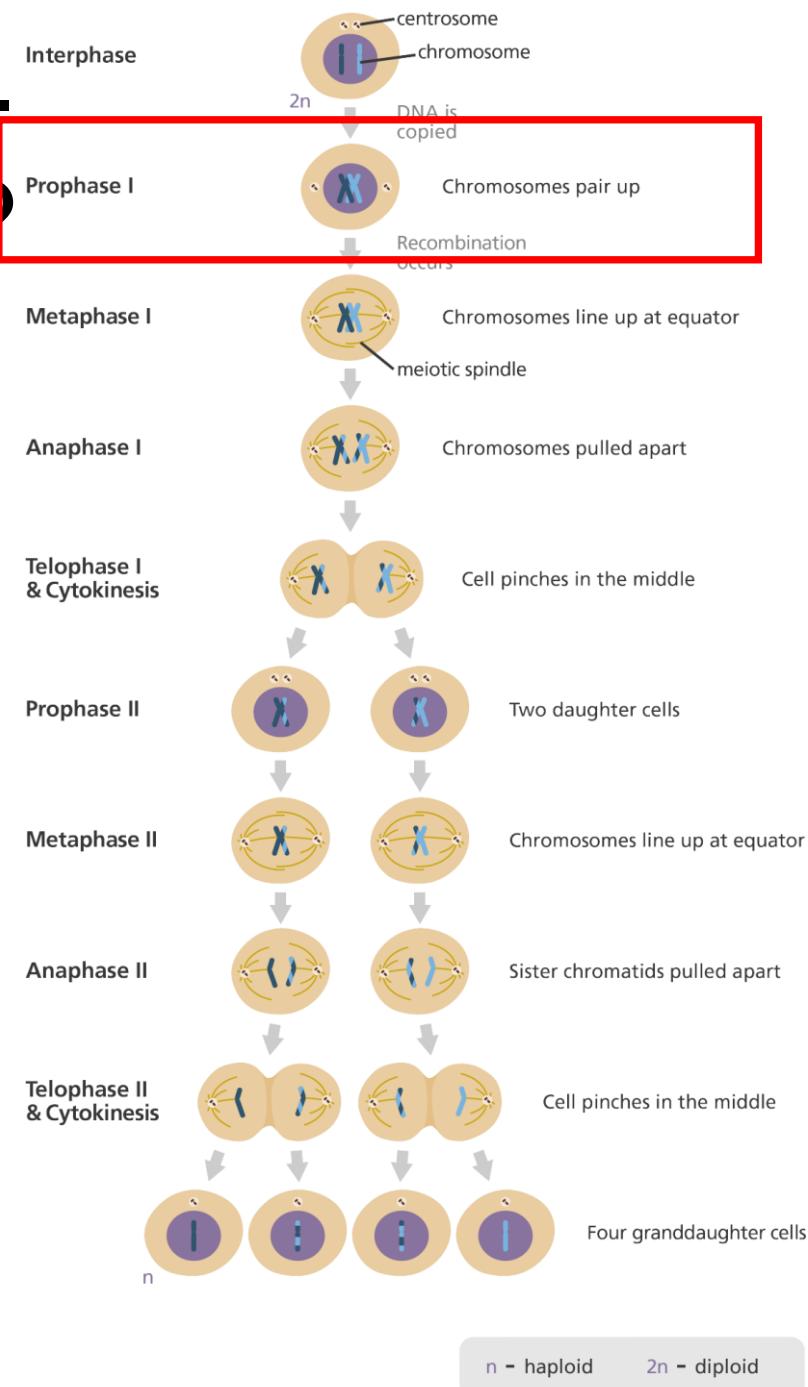
Aneuploidy usually caused by nondisjunction



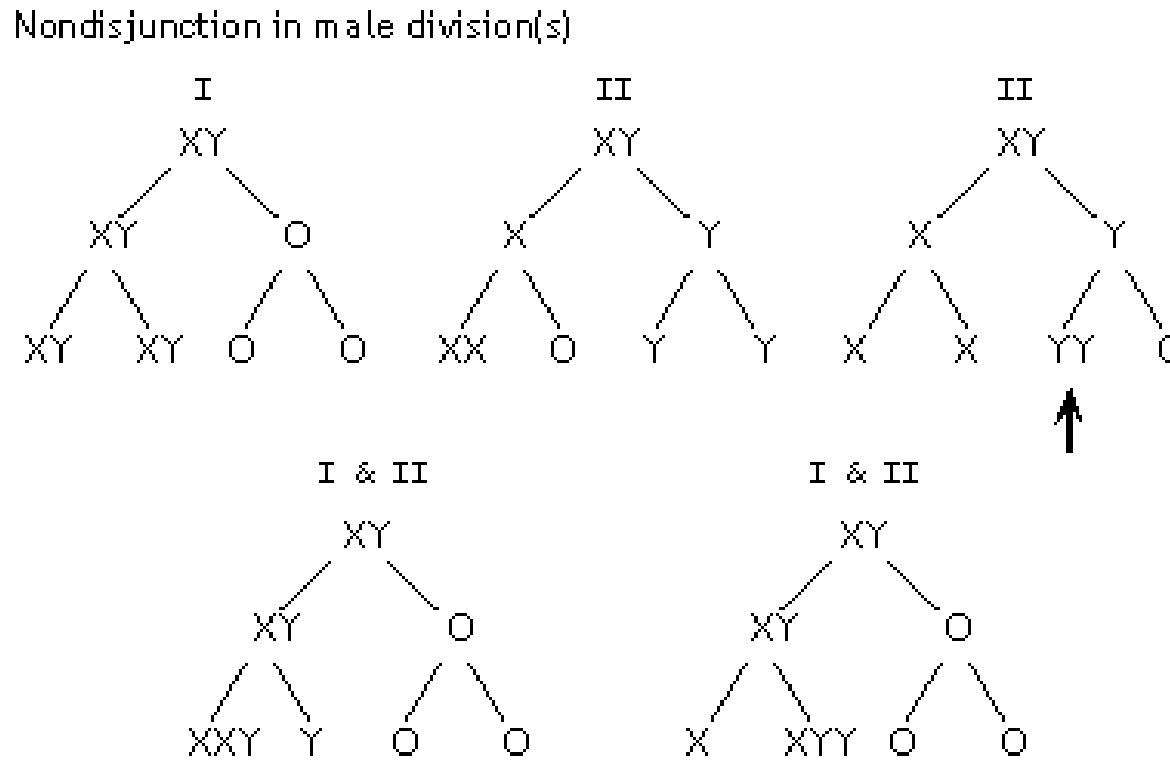
- Nondisjunction in meiosis 1 = pass on two different copies of chromosome
- Nondisjunction in meiosis 2 = pass on two same copies of chromosome

Why/how does non-disjunction happen?

- Happens more frequently in oocytes (egg cells)
 - Oocytes initially form before birth (at 3-4 months)
 - Arrested in prophase 1
 - Resume rest of meiosis after puberty as eggs are released
 - Cohesin which holds chromosomes together wears down over time?
 - Fewer crossover events in oocytes?
 - Spindle/centromere breakdown?



Male sex chromosome nondisjunction



- Also occurs in males:
 - 1 in 1,000 boys are XYY = nondisjunction had to occur in sperm

Aneuploidy = abnormal number of chromosomes in a cell

- Monosomy = 1 copy of a chromosome
 - Monosomy X (XO) = Turner syndrome
 - Short stature, delayed puberty, infertility, heart defects, learning disabilities
 - 1 in 2,000-2,500 live female births



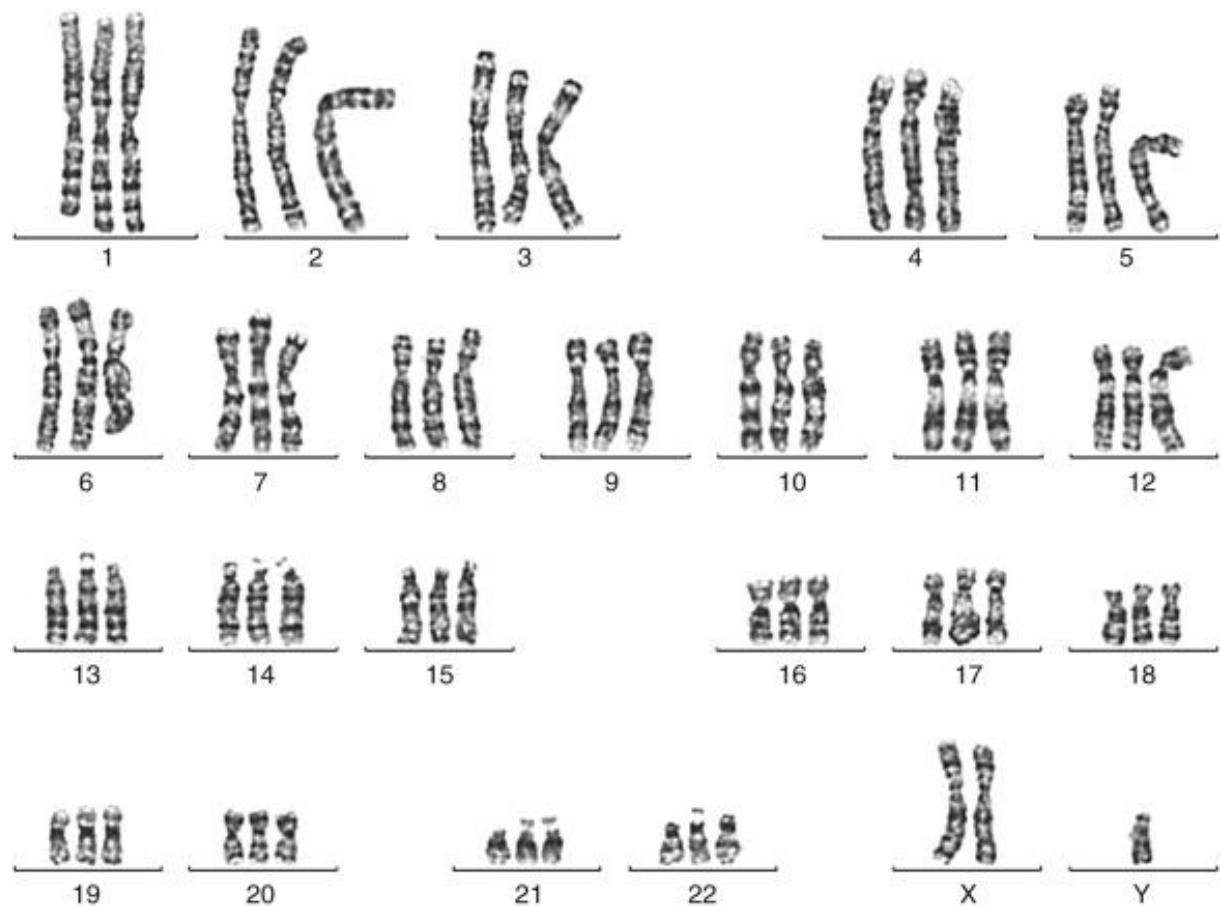
- Trisomy = 3 copies of a chromosome
 - Trisomy 21 = Down syndrome
 - Distinct facial appearance, intellectual disability, developmental delays
 - Maybe: thyroid/heart issues

- Others:
 - Trisomy X = Triple X syndrome 1 in 1,000 female births
 - XXY = Klinefelter syndrome = 1 in 500-1,000 males
 - XYY = XYY syndrome = 1 in 1,000 males
 - Trisomy 13 = Patau syndrome
 - Trisomy 18 = Edwards' Syndrome

Generally: deletion is more severe than duplication

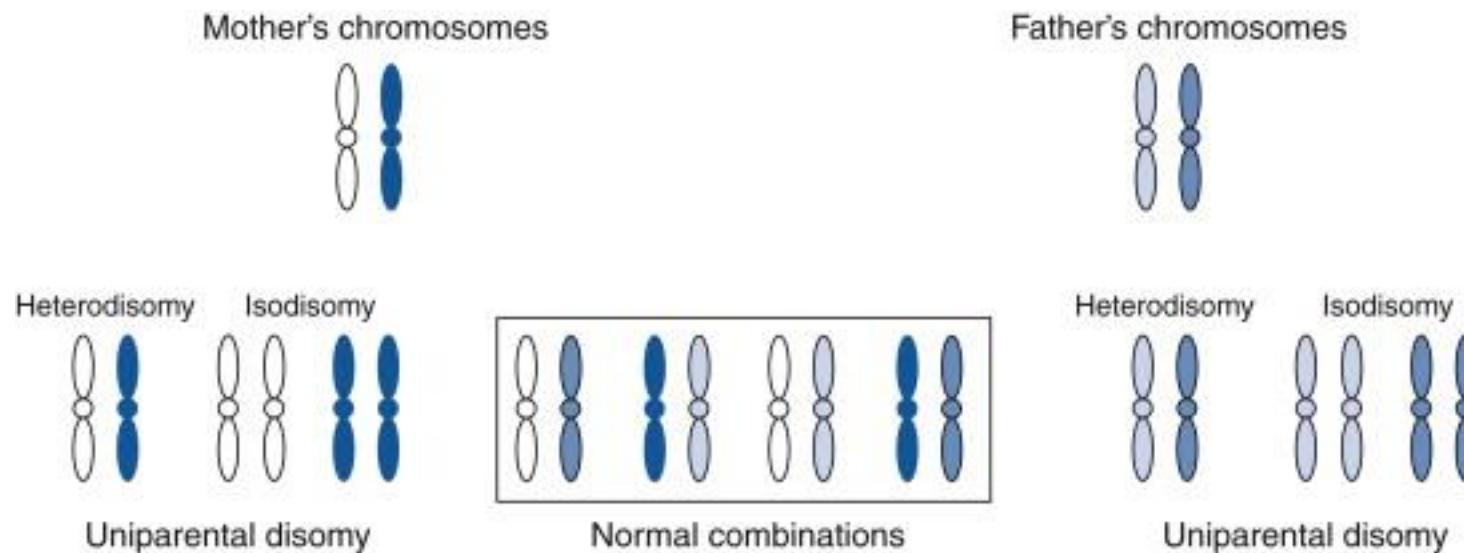
Triploidy

- 1-2% of all conceptions
- Usually early miscarriages but can survive to birth
- Euploid = an exact multiple of the haploid number of chromosomes
 - Monoploid
 - Diploid
 - Triploid
 - Tetraploid
 - etc



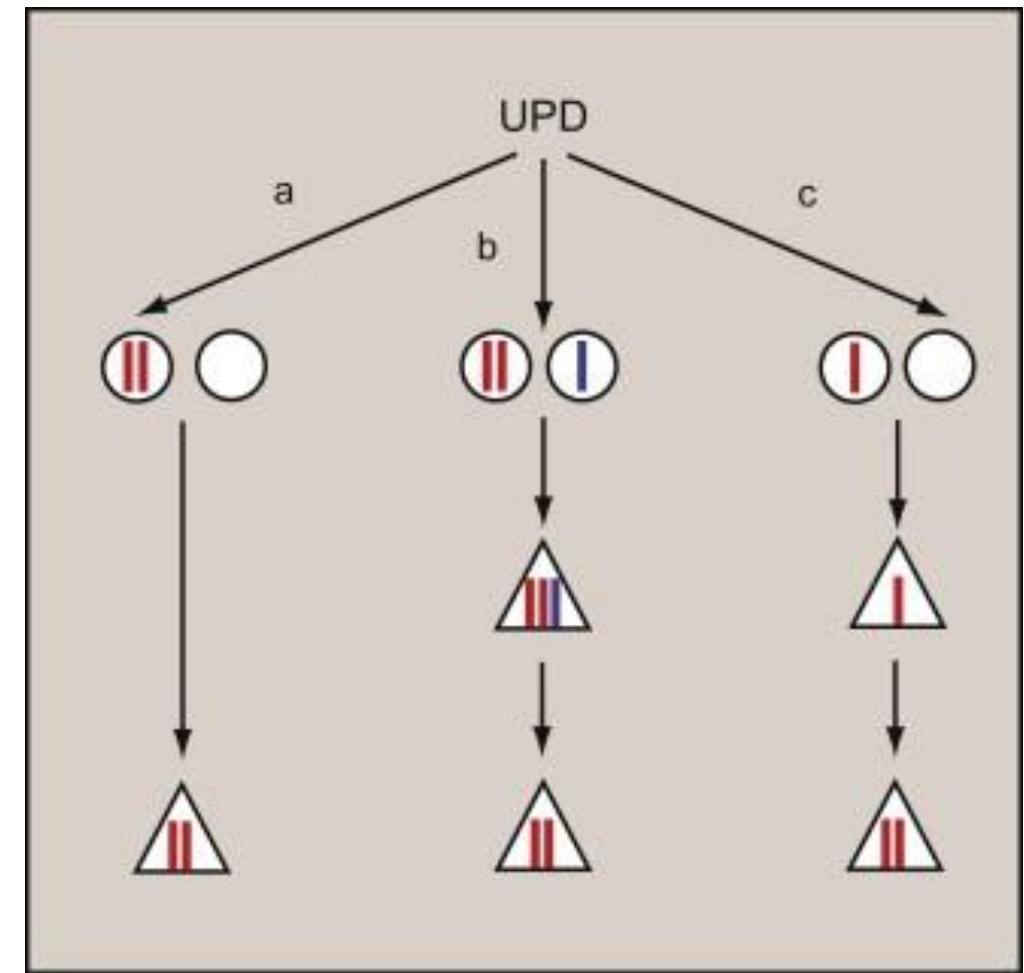
Uniparental disomy

- When two copies of a chromosome are inherited from the same parent
- **Heterodisomy** = inherited both of one parent's chromosomes
 - Error in meiosis I
- **Isodisomy** = inherited two identical copies of a chromosome from one parent
 - Error in meiosis II



Uniparental disomy – how?

- a) Errors in meiosis from both parents
(isodisomy or heterodisomy)
- b) Trisomy rescue
(isodisomy or heterodisomy)
- c) Duplication of single chromosome
(isodisomy only)



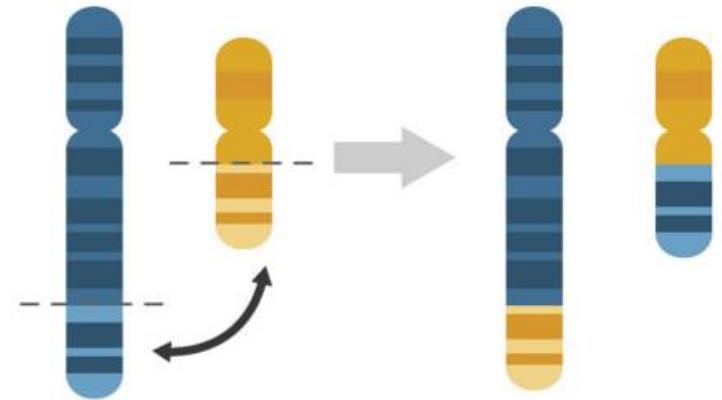


Abnormalities of Chromosome Structure

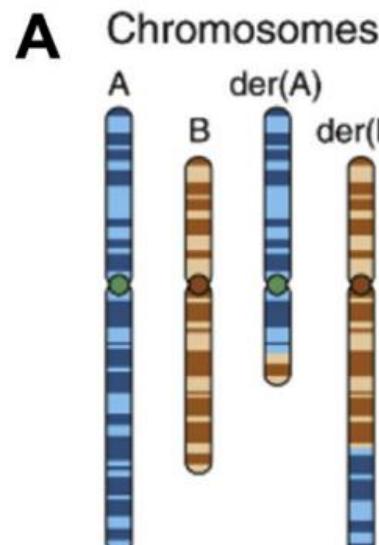
- Translocations + Robertsonian translocations
- Structural variants

Translocations

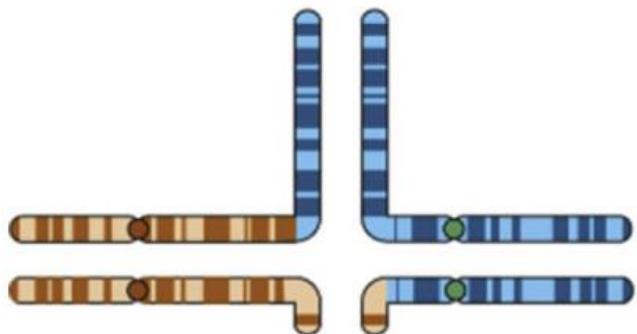
- Structural rearrangements of the chromosome
- Balanced
 - Normal complement of chromosomal material
- Unbalanced
 - Additional or missing material
- 1 in 375 newborns, most unaware until they try to have children
- Naming: **t(8;14)(q24;q32)** indicates a translocation between chromosomes 8 and 14
 - recombination points at 8q24 and 14q32
- Many specific translocations are seen more often, many linked with cancers



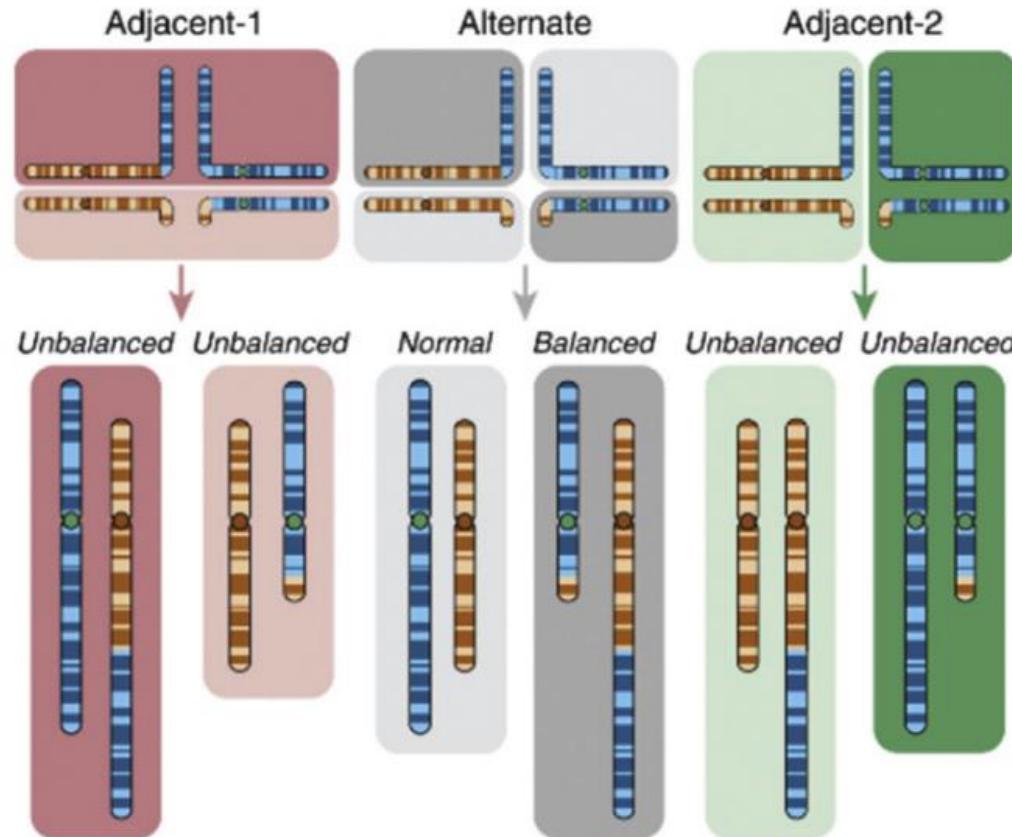
Mitosis with balanced translocations?



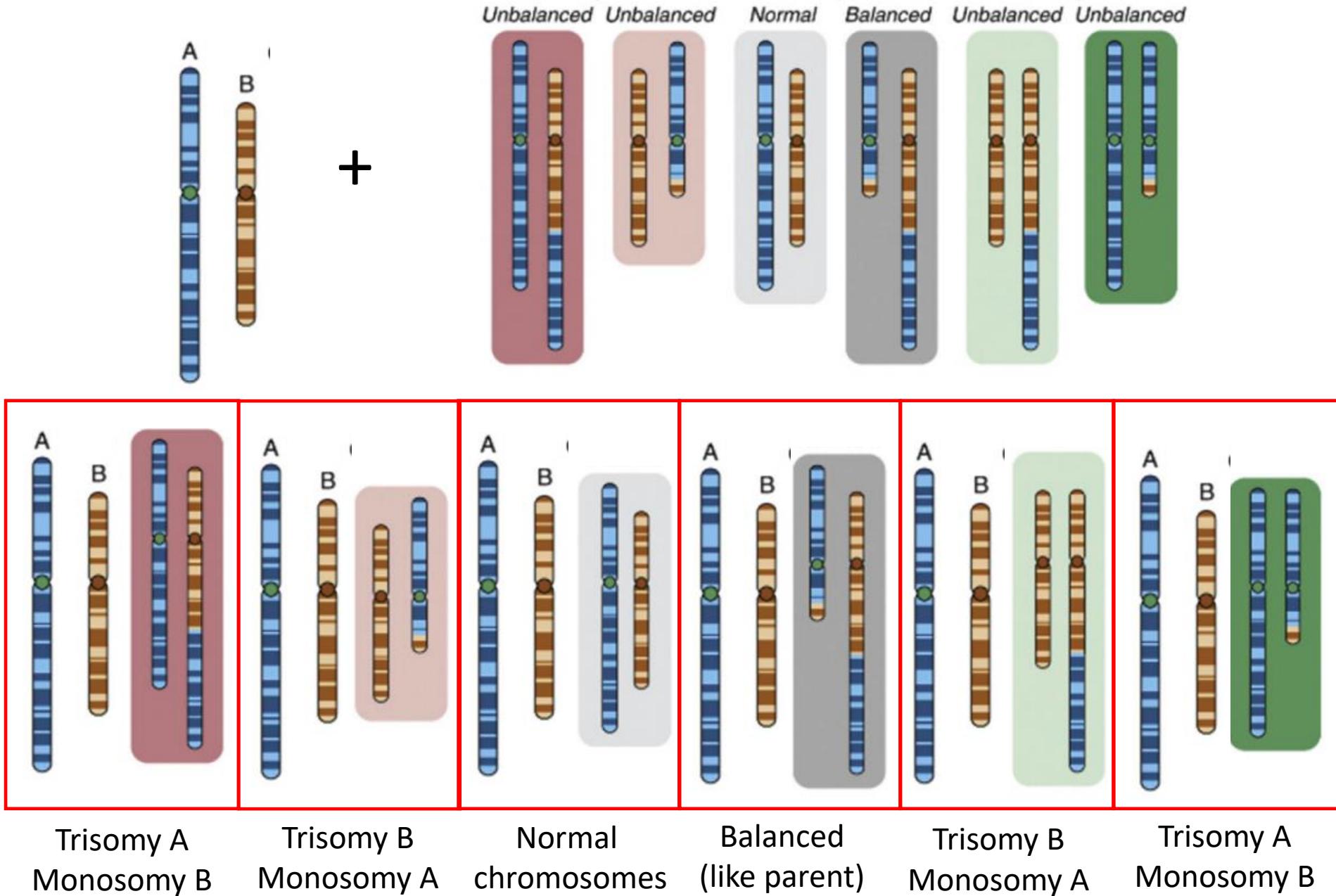
B Quadrivalent formation
in meiosis



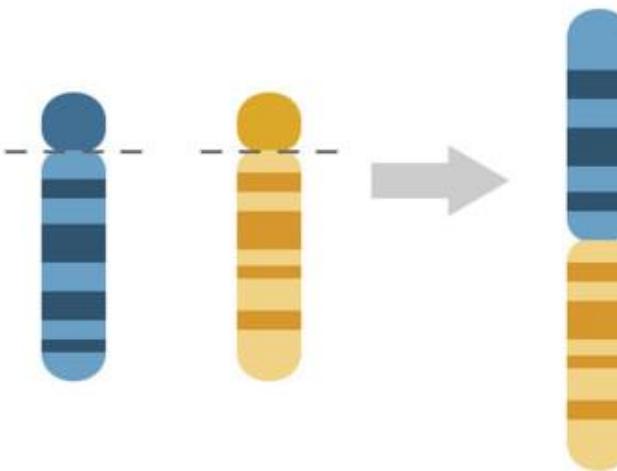
C Segregation and gametes



Fertilization with balanced translocations?



Robertsonian translocations



- Chromosomes 13, 14, 15, 21, 22 have only one functioning ‘arm’
- Robertsonian translocation = two of these chromosomes joining up into one chromosome, containing the long arm of both
- Balanced
- Individuals will have 45 chromosomes

Robertsonian meiosis/fertilization

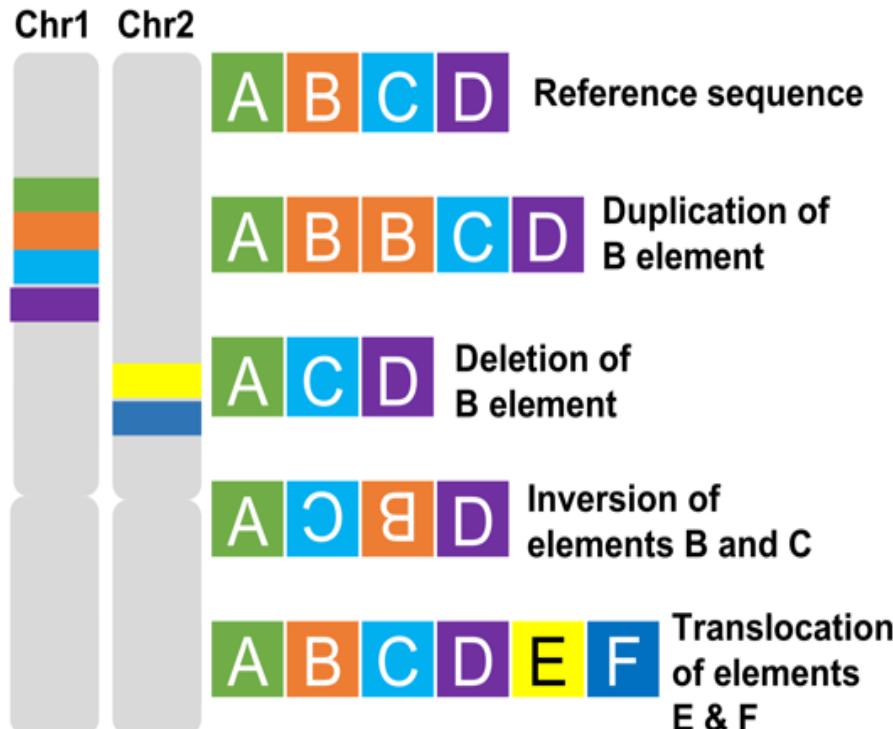
If one parent is a Robertsonian translocation carrier

Mother has 13;21, 14;21, 15;21 or 21;22	10-15% risk of a baby with translocation Down's.
Mother has 13;14, 13;15, 13;21 or 13;22	1% chance of having a baby with trisomy 13.
Mother has 14;15, 14;22 or 15;22	Almost certainly no risk of having a baby with a trisomy, but possible risk of miscarriage or UPD.
Father with any Robertsonian combination	Low risk, below 1%, of any child being affected.

Notes on translocations

- A small number (~5%) of down syndrome is caused by Robertsonian translocation.
- Some balanced translocations are causes of or associated with cancer types in **somatic** mutations (not germ line)
- Mechanism still under investigation, but happens more frequently at some known ‘breakage points’. Requires:
 - Double-stranded break
 - Non-homologous end joining

Structural variation (SV)



Jenn Fifita

What is it?

- Large genomic alterations affecting stretches of DNA >1000bp
- Affects chromosome structure
- Most prominent source of variation in the human genome
- Effects gene dosage and gene expression

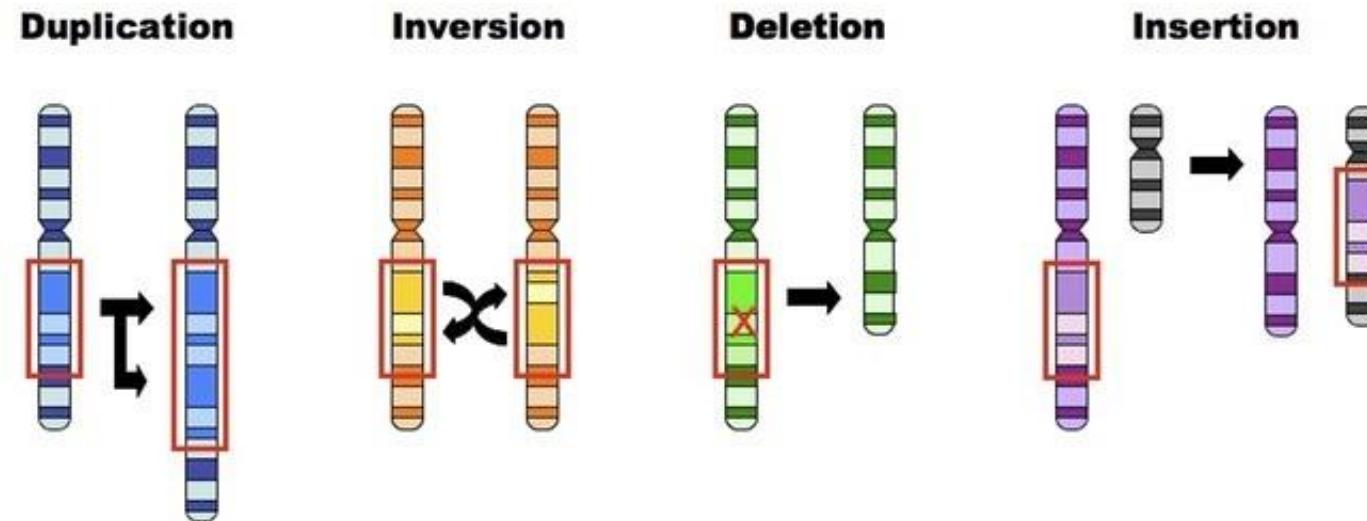
How to find it?

- Traditionally detected using laborious cytogenetic techniques
 - Multiplex ligation-dependent probe amplification (MLPA) or microarrays
- Recent development of bioinformatic tools for SV detection in NGS data now make high-throughput SV analysis possible

Relevance to ALS?

- Implicated in other neurodegenerative diseases
 - *incl. Parkinson's disease, Kennedy's disease, Spinocerebellar ataxias, Duchenne muscular dystrophy*
- ALS pathogenic *C9orf72* expansion is a similar phenomenon to copy number variation and is similar in size to SVs at up to 27,000bp long

Smaller/non-specific chromosome changes



- Insertion/Inversion similar to translocation
 - Maybe problems if gene interrupted, and in meiosis
- Duplications/deletions 5Mb or larger
 - <5Mb = microdeletion / microduplication
 - Can cause copy number variant (CNV)
 - Missing a gene copy more likely to cause problems than extra gene copy

Copy number variants (CNVs)

- In healthy population (Itsara et al., 2009):
 - 5-10% of people at least one del/dup larger than 500kb
 - 1-2% of people at least one del/dup larger than 1Mb
- What makes a CNV cause problems?
 - How many genes does it include? More = more likely to cause problem
 - Does it interrupt a gene?
- Extremely variable, even within same family
- Developmental delay / intellectual disability
 - 30% of unexplained cases had pathogenic duplication or deletion (2018 study)
- Many CNVs linked with Autism Spectrum Disorders (ASD), neurocognitive problems

Summary

- Mosaicism = mutation early in embryo development – person will have mutation in a certain % of their cells
- Aneuploidy = abnormal number of one chromosome set
 - Most trisomies/monosomies aren't survivable
- Uniparental disomy = both copies of a chromosome from one parent
- Translocations + Robertsonian translocations
- Changes within a chromosome = copy number variants
 - Deletions/duplications
- Next lecture: nucleotide mutations

BIOL3120 –Chromosomal mutations

LEARNING OBJECTIVES



On successful completion of this lecture, you will be able to:

- Explain the different types of chromosomal mutations
- Use this knowledge to solve problems in human genetics relating to heritability, polygenic inheritance and chromosomal mutations



MACQUARIE
University

BIOL3120 –Human Genetics and Evolutionary Medicine

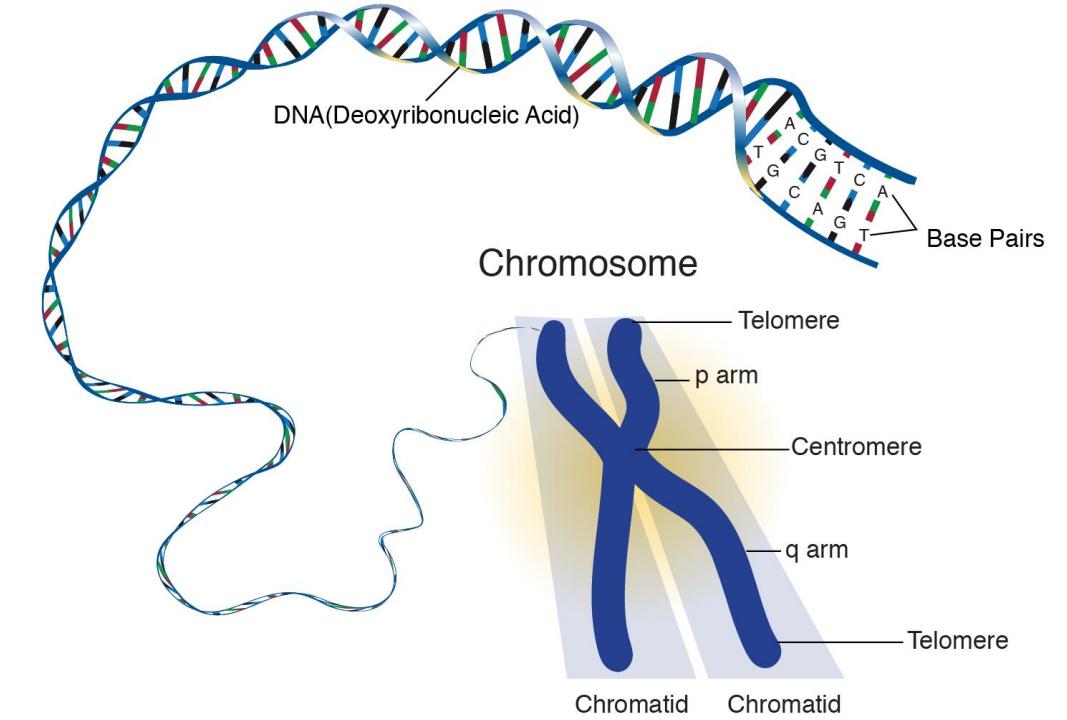
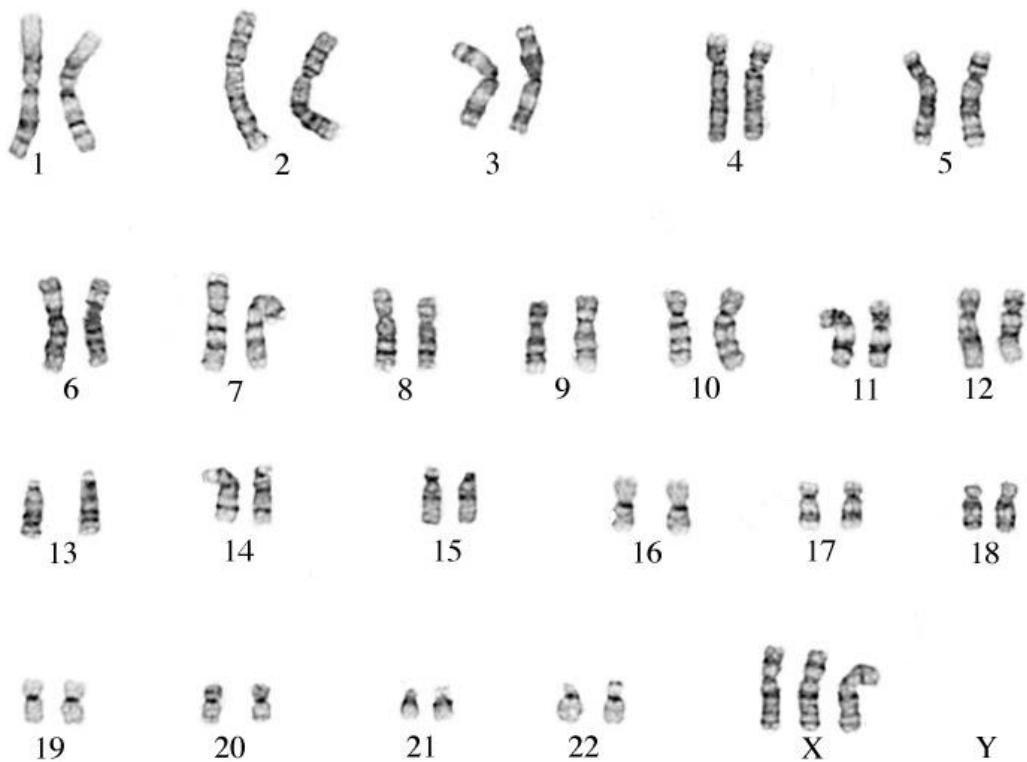
Nucleotide Mutations





4	Heritability and Polygenics Chromosomal Mutations	Problem Set 2	Problem Set 2 (5%)	Explain the principles of evolutionary biology and their role in human health and disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources
5	Nucleotide Mutations Human Genetic Diversity and Evolution	Problem Set 3	Problem Set 3 (5%)	Explain the principles of evolutionary biology and their role in human health and disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources
6	Genetic Testing Techniques GWAS	Problem Set 4	Problem Set 4 (5%)	Explain the importance of new techniques in human genetics for understanding human disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources
Recess				

Quantity (chromosomes) vs quality (nucleotides)



BIOL3120 –Nucleotide Mutations

LEARNING OBJECTIVES



On successful completion of this lecture, you will be able to:

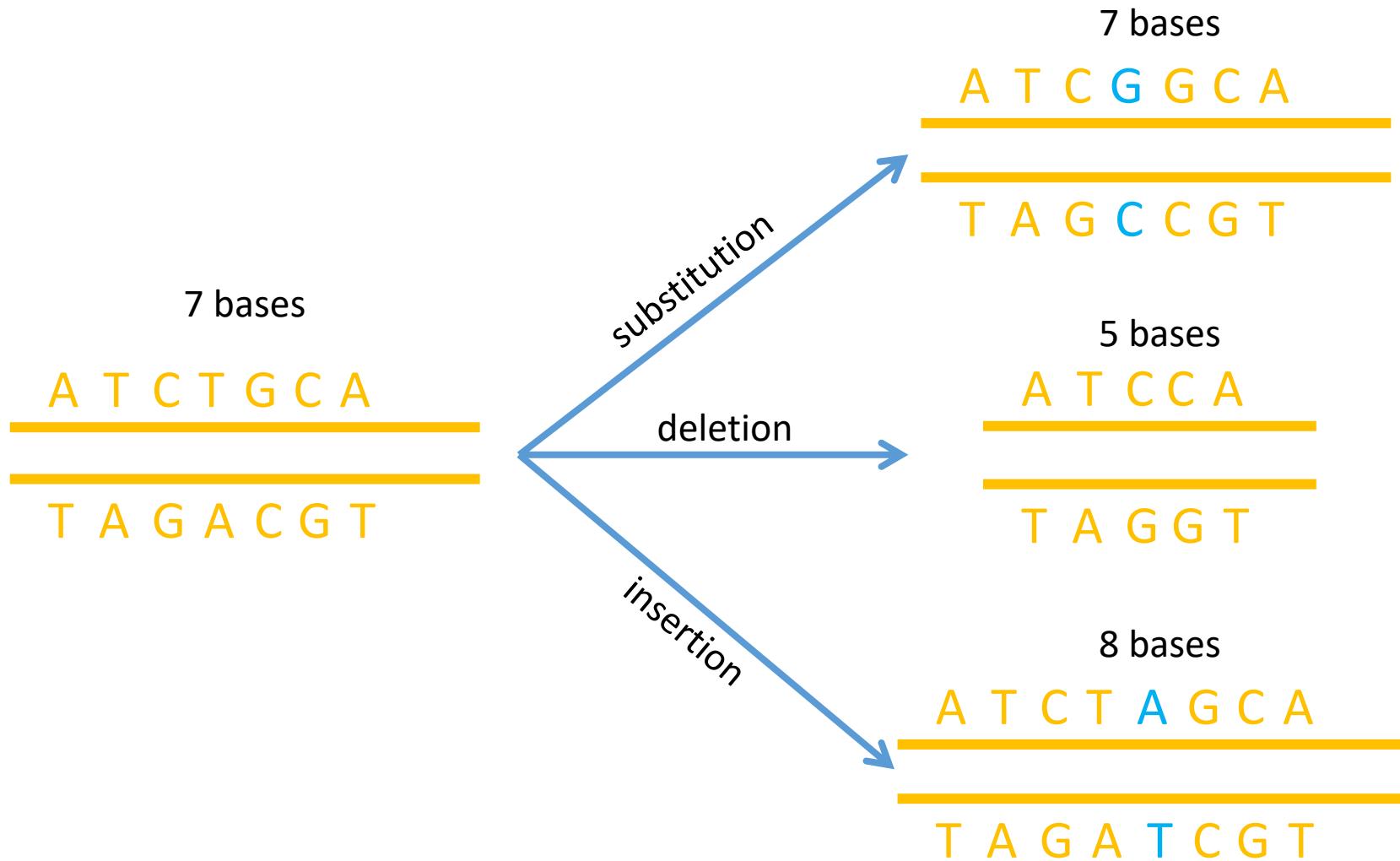
- Identify and name nucleotide mutations
- Interpret nucleotide variations
- Understand repeat expansions



Nucleotide Mutations

Nucleotide level variation

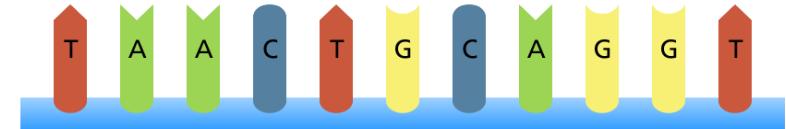
SINGLE NUCLEOTIDE AND INSERTION/DELETION VARIATION



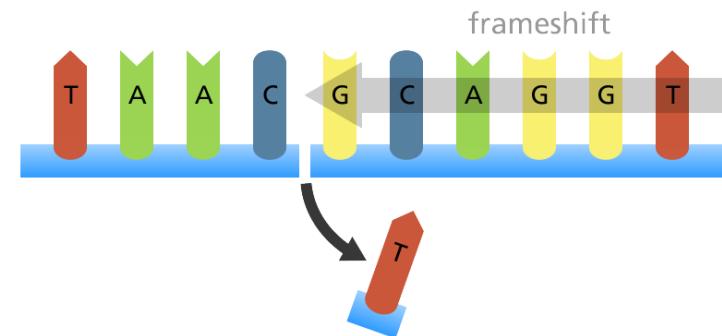
Deletions/Insertions

- Insertion or deletion of basepairs = Frameshift = all amino acids from there on affected
- Most likely a stop codon soon
 - 3/64 codons are stop codons = expect 1/23 codons to be a stop codon
- Very likely to impact function of protein
- Insertion/deletion of multiple of 3bp = no frameshift, extra or missing amino acids. May not impact function

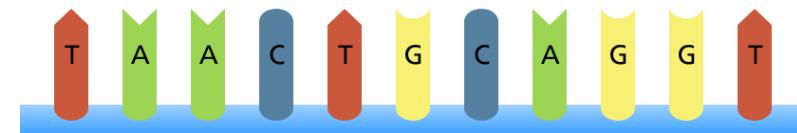
Original sequence



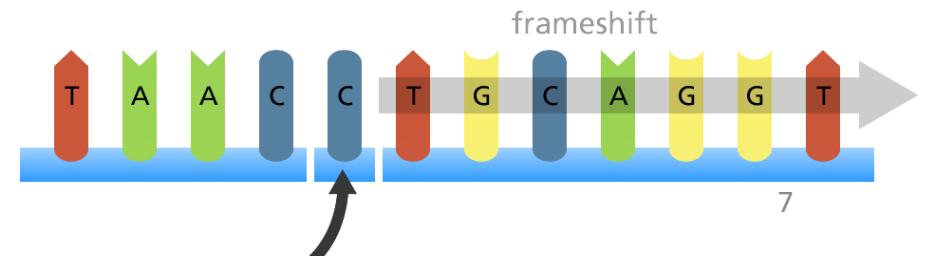
Deletion



Original sequence



Insertion



Frameshifts in coding region – what's changing?

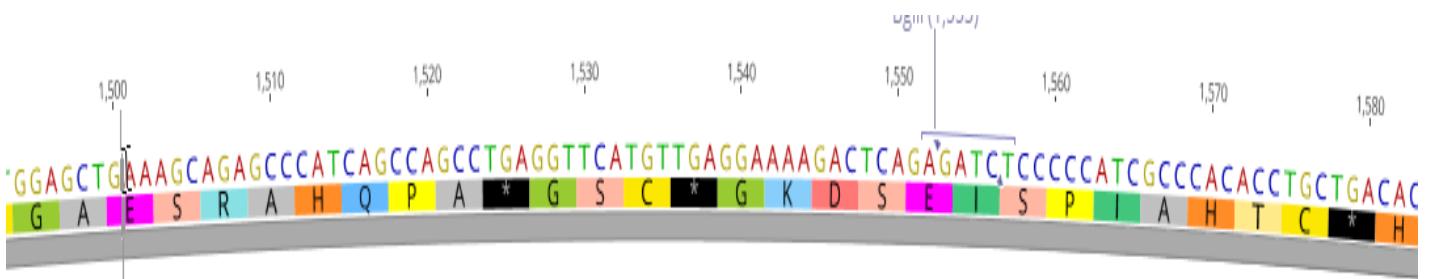
Codon Table
Second base in codon

Messenger RNA Codons	U	C	A	G		
U	UUU UUC UUA UUG	UCU UCC UCA UCG	UAU UAC UAA UAG	Tyr Stop Stop	UGU UGC UGA UGG	Cys Stop Trp
C	CUU CUC CUA CUG	CCU CCC CCA CCG	CAU CAC CAA CAG	His Gin	CGU CGC CGA CGG	Arg
A	AUU AUC AUA AUG	ACU ACC ACA ACG	AAU AAC AAA AAG	Asn Lys	AGU AGC AGA AGG	Ser Arg
G	GUU GUC GUA GUG	GCU GCC GCA GCG	GAU GAC GAA GAG	Asp Glu	GGU GGC GGA GGG	Gly

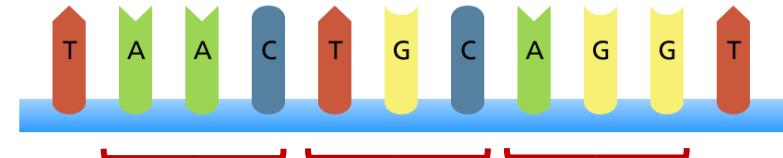
First base in codon *Third base in codon*

Deletions/Insertions

- Insertion or deletion of basepair = Frameshift = all amino acids from there on affected
- Most likely a stop codon soon
 - 3/64 codons are stop codons = expect 1/23 codons to be a stop codon
- Very likely to impact function of protein
- Insertion/deletion of multiple of 3bp = no frameshift, extra or missing amino acids. May not impact function



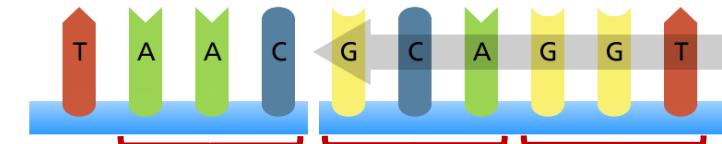
Original sequence



Deletion

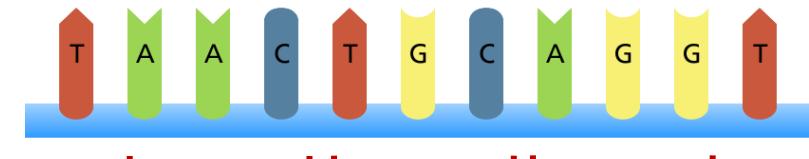
I C M

frameshift



I A G

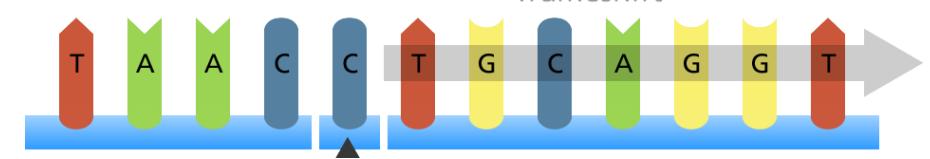
Original sequence



Insertion

I C M

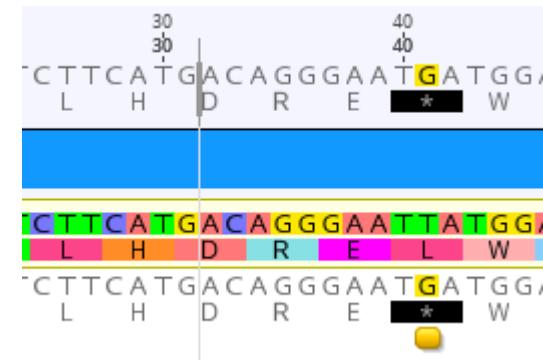
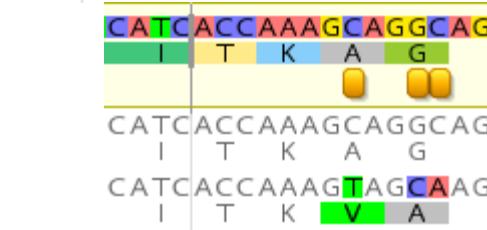
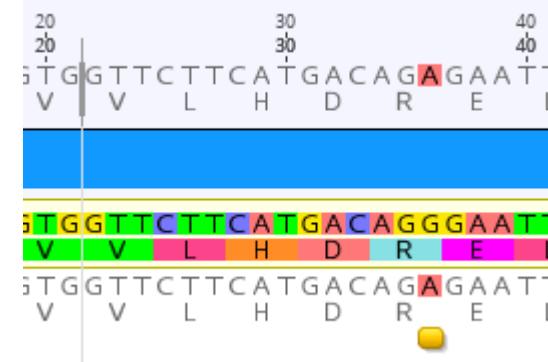
frameshift



I L Q

Substitutions/Point Mutations

- Silent or synonymous mutation
 - No change to amino acid
 - In some cases can cause disease
- Missense mutation
 - Changes to another amino acid
 - Many possibilities
- Nonsense mutation
 - Changes to stop codon
 - How much is the protein getting shortened?



Nomenclature for the description of sequence variations

- First, describe the reference sequence
 - DNA, RNA or Protein
- Then describe the location
 - State the basepair or amino acid where the mutation has occurred
- Then describe the kind of mutation
 - Was it a substitution, deletion, stop codon, frameshift, etc.

Indicate the reference sequence:

DNA		
	coding DNA	c.
	genomic DNA	g.
	mitochondrial DNA	m.
RNA		r.
Protein		p.

Code:

substitution (for bases)	>
range	-
more change in one allele	:
more transcripts / mosaicism	:
uncertain	0
allele	[]
deletion	del
duplication	dup
insertion	ins
inversion	inv
conversion	con
extension	ext
stop codon	X
frame shift	fsX
opposite strand	o
translocation	t

Type of variation/mutation:

Substitution	
c.123A>G	on cDNA, A in 123 is replaced by G
p.P252R	on protein, proline (P) replaced by arginine (R)
Deletion	
c.546delT	deletion of T in 546
c.586_591del	for six bases deleted
p.F508del	deletion of phenylalanine (F) in 508
Duplication	
c.546dupT	duplication of T in 546
c.586_591dup	duplication of the segment 586 to 591
p.G4_Q6dup	duplication of the segment from glycine (G) in 4 to glutamine (Q) in 6
Insertion	
c.546_547insT	insertion of T between 546 and 547
c.1086_1087insGCGTGA	insertion of GCGTGA
p.K2_L3insQS	insertion of glutamine serine between lysine (K) in 2 and leucine (L) in 3
Inversion	
c.546_2031inv	segment 546 to 2031 inverted
Frameshift	
p.R83SfsX15	arginine (R) is the first amino acid changed, it is in position 83, it makes serine (S) instead, the length of the shift frame is 15, including the stop codon (X)



Interpreting nucleotide variations

Biggest problem in clinical genetics now: Is this variant likely to cause a problem? ‘variant classification’

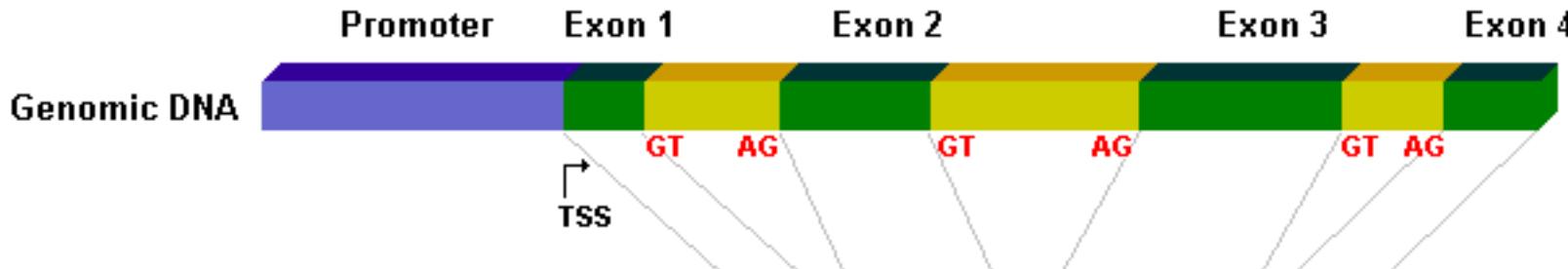
- In the past -> phenotype first
 - problem first, find the shared mutation
 - All research
- Now -> lots of sequence data
 - find variants first, try to determine likelihood of any of them causing the problem
 - Mutations may be seen first in the clinic

Variant classification

Questions to ask:

- Where in the gene is the change?
- If in an exon, what is the amino acid change? Bioinformatics predictions of the outcome on protein function
- Has this variant been reported in literature? Has a similar variant been reported (same residue?)
- Is this variant seen in healthy populations? What's the frequency?
- Is this residue conserved across other species' version of this protein?

Substitution – where?

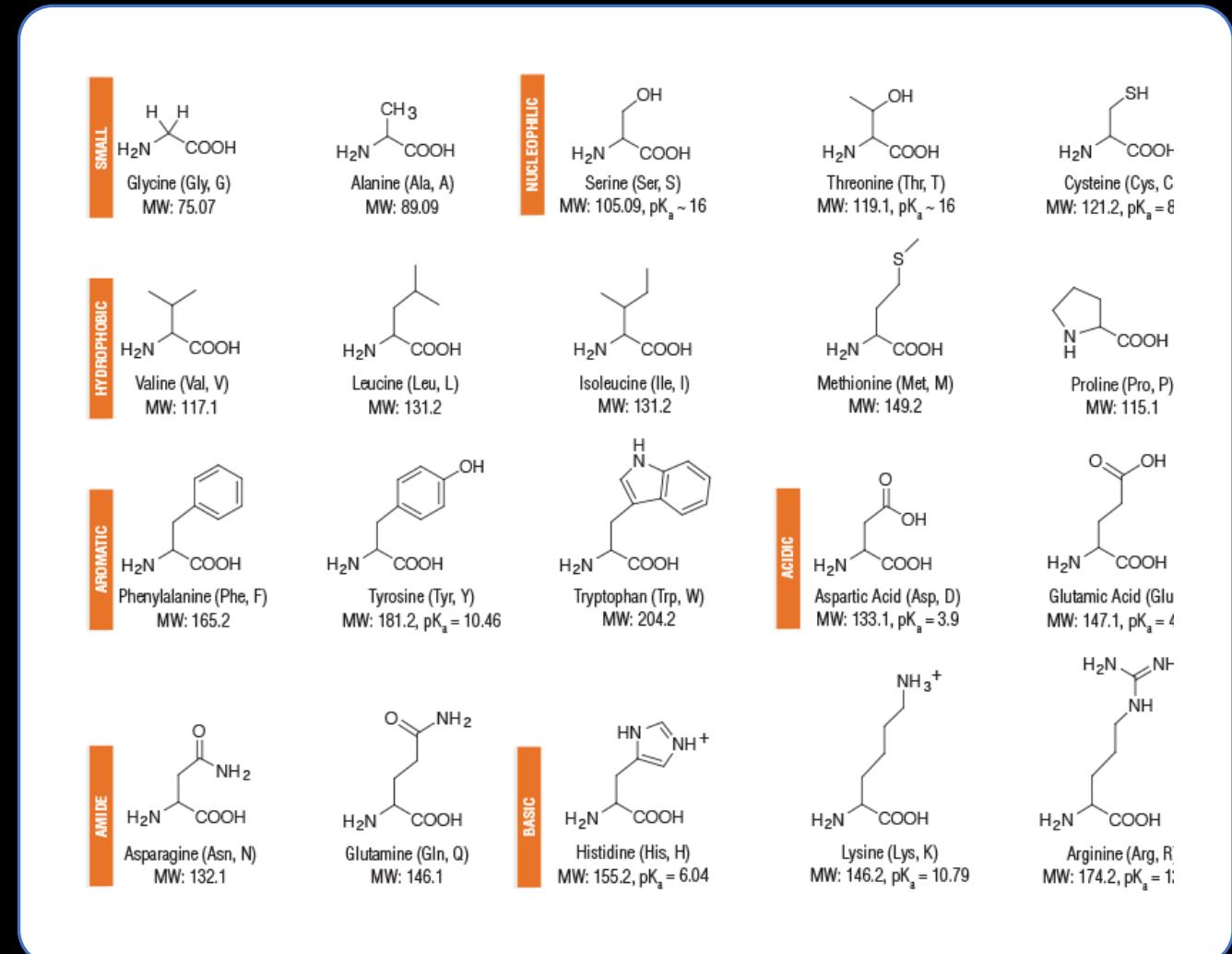


- Exon = can effect amino acid sequence
- Intron = within 8-10bp of exon may cause splicing issues
- Mutations within an exon near intron may affect splicing, even if it doesn't affect the amino acid (silent mutation)
- Promoter = maybe but far less likely
 - Might impact expression level 
- Further upstream, promoter?

Missense mutations: How similar are the amino acids?

Protein prediction algorithms:

- SIFT
- Polyphen



Variant seen in healthy populations?

- Databases can provide frequency of this variant in healthy population
- ExAC (Exome Aggregation Consortium)
 - <http://exac.broadinstitute.org/>
- gnomAD
 - <https://gnomad.broadinstitute.org/>
- Consider dominant vs recessive ‘allowed’ frequencies

Missense mutations: Is this residue conserved in other species' homologues of this protein?

Homo sapiens
Rhesus
Mouse
Rat
Cat
Dog
Bovine
Elephant
Chicken
X. tropicalis
Zebrafish

GNYNNQS-SNFGPMKGGNF~~GG~~-RSSG**P**YGGGGQYFAKPRNQGGY
GNYNNQS-SNFGPMKGGNF~~GG~~-RSSG**P**YGGGGQYFAKPRNQGGY
GNYNNQS-SNFGPMKGGNF~~GG~~-RSSG**P**YGGGGQYFAKPRNQGGY
GNYNNQS-SNFGPMKGGNF~~GG~~-RSSG**P**YGGGGQYFAKPRNQGGY
GNYNNQS-SNFGPMKGGNF~~GG~~-RSSG**P**YGGGGQYFAKPRNQGGY
GNYNNQS-SNFGPMKGGNF~~GG~~-RSSG**P**YGGGG-----GY
GNYNNQS-SNFGPMKGGNF~~GG~~-RSSG**P**YGGGGQYFAKPRNQGGY
GNYS**GQQQSNY**GP~~M~~KGG**SFGGRSSGS****P**YGGG**YGSGG**-----
GSYNNQS-SNFGPMKGGNF~~GG~~-RSSG**P**YGGG-----GY
GNYNNQS**SSSF**GP~~M~~KGG**NYGG**RNSG**P**YGG**SN**-----**A**
GNYS**QQ**-SN**Y**GP~~M~~KGN**F**GGG**RNSG****P**YGGG**YGGGSSG**-----

M9 core

Other things to consider:

- Functional studies?
- De novo (new mutation) or in parents?
- Segregate throughout family with disease?

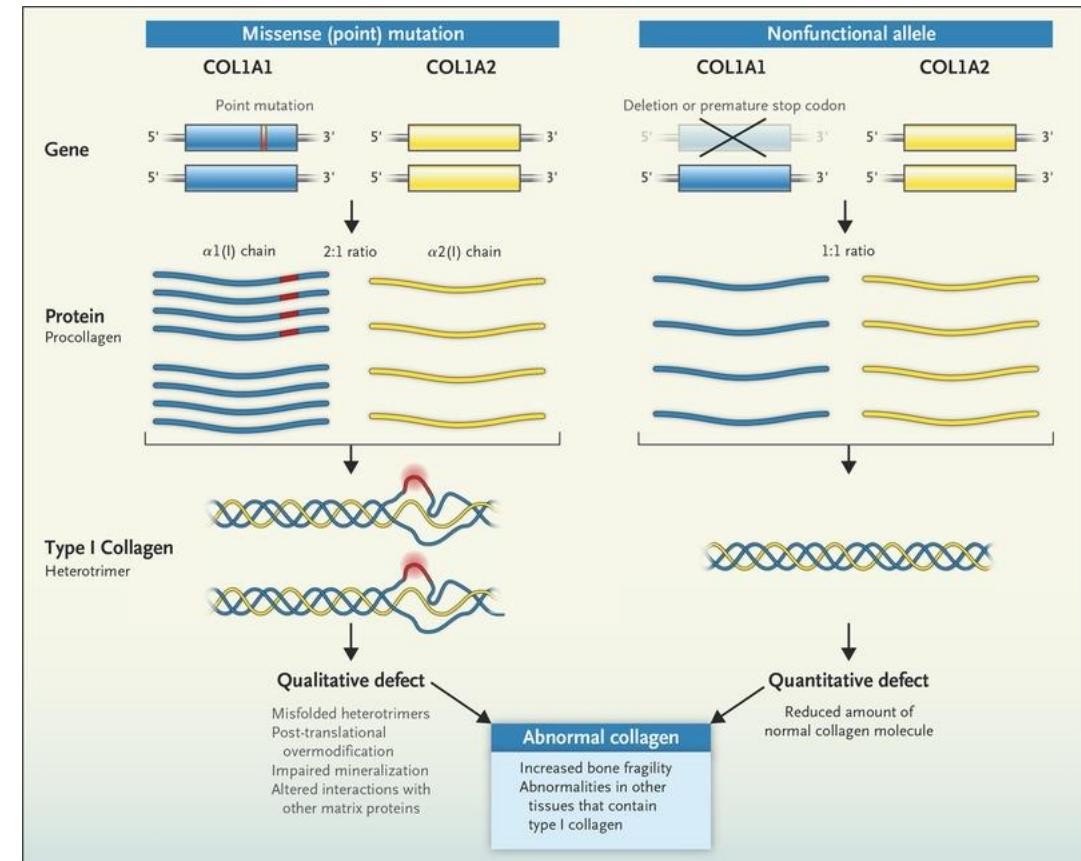
Why are diseases dominant or recessive?

Recessive: Having 1 normal allele gives normal phenotype

- 2 mutant alleles = **loss of function** = having no functioning protein

Dominant: 1 mutant allele causes condition

- **Gain of function** = new toxic function with mutant allele (majority of dominant conditions)
- **Haploinsufficiency** = the problem/condition is caused by not enough protein being made e.g. Marfan syndrome
- **Dominant negative** = the mutant protein product interferes with the function of the natural/normal protein product (eg osteogenesis imperfecta)





Repeat Expansions

Repeat expansions

- Some genes contain short repetitive sequences
- These repetitive sequence can expand and cause disease
 - >40 diseases, primarily affecting the nervous system
- Expanded trinucleotide repeat diseases were the first discovered and most frequent
 - Tetra-, penta-, hexa- and dodeca-nucleotide repeat expansions exist
- Can experience anticipation
 - Decreasing age of onset or increased severity of disease across generations

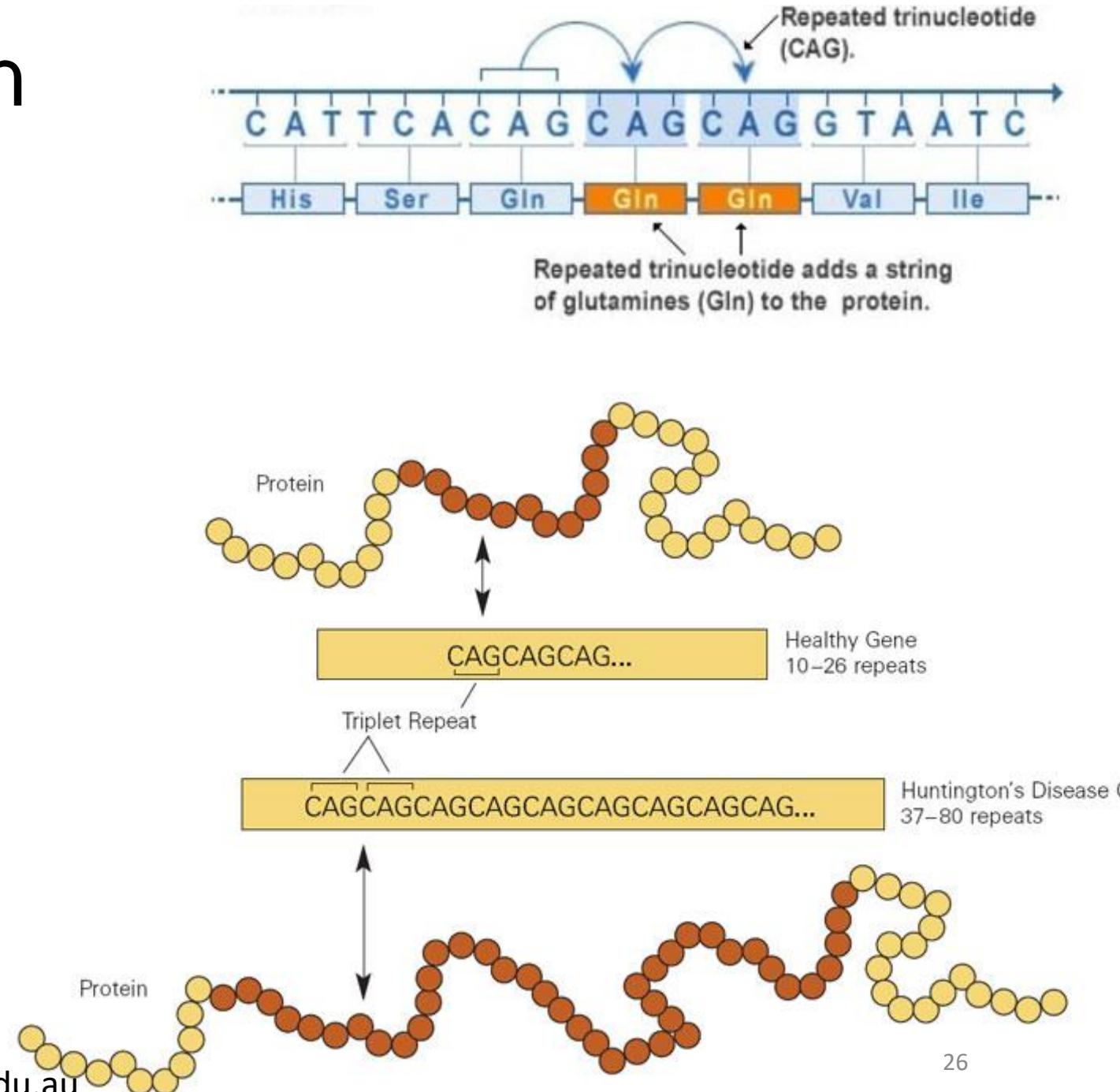
Repeat expansions

Huntington's Disease (HD)

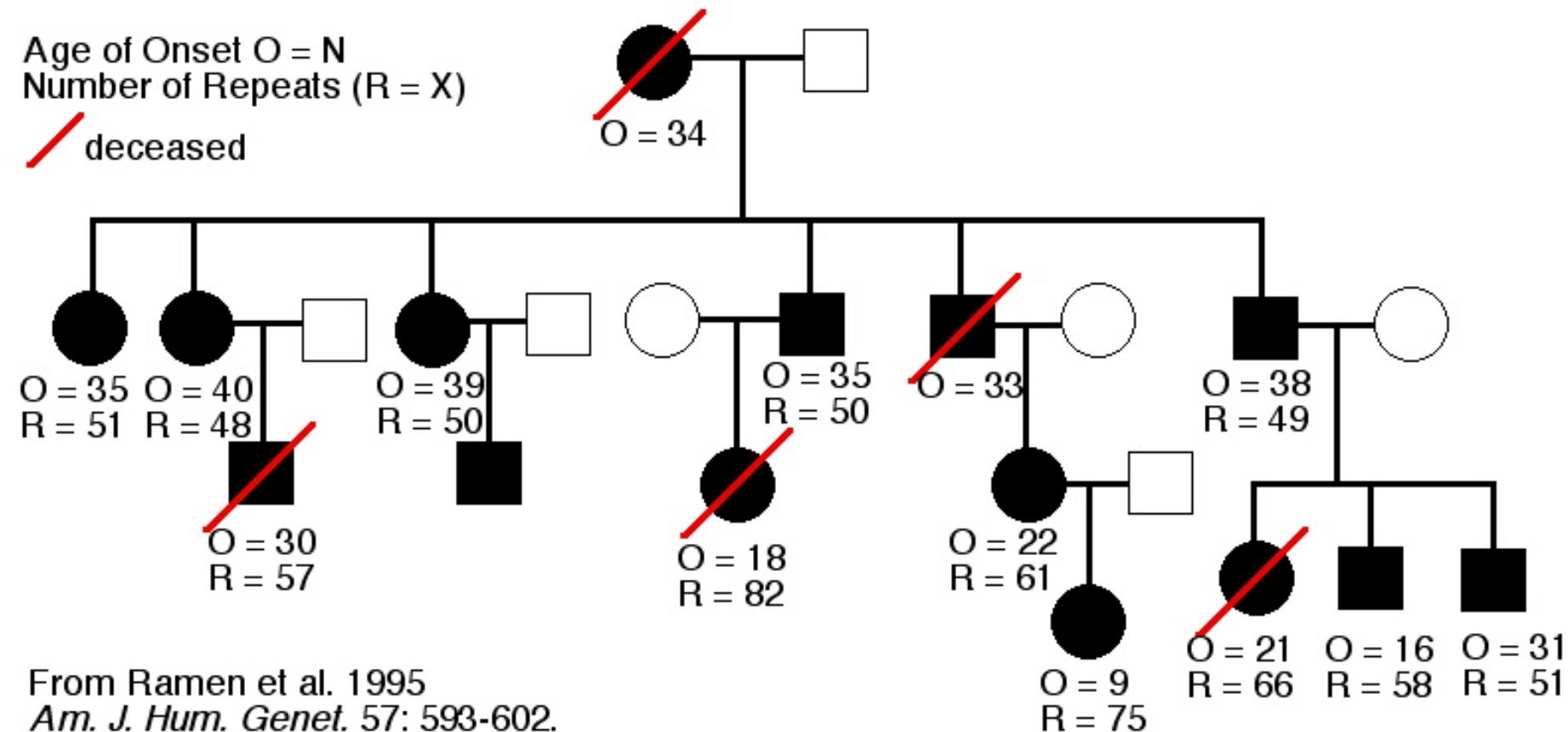
- Dominant inheritance
- HTT gene
- Onset 30s or 40s
- Progressive loss of motor abilities, psychological deterioration, cognitive decline
- Death 15~20 years after first symptoms
- Variable repeat of 3 nucleotides in 1st exon - CAG

Repeat expansions in HD

Number of CAG repeats*	What does it mean for you?
26 or less Normal range	You will not develop HD.
27 - 35 Intermediate range	You will not develop HD.
36 - 39 Increased risk range (reduced penetrance range)	You are likely to develop HD in your lifetime. However: <ul style="list-style-type: none"> • you might develop it at a late age • the condition might be less severe • you might not develop HD at all.
40 or more faulty gene HD range (full penetrance range)	You will almost certainly develop HD if you live a normal life expectancy.



Anticipation in Huntington's Disease (HD)



- Larger expansion = earlier onset

BIOL3120 –Nucleotide Mutations

LEARNING OBJECTIVES



On successful completion of this lecture, you will be able to:

- Identify and name nucleotide mutations
- Interpret nucleotide variations
- Understand repeat expansions

BIOL3120: Human genetics and evolutionary medicine

LECTURE 7: HUMAN GENETIC DIVERSITY AND EVOLUTION



Lecture 7: Human genetic diversity and evolution



On completion of this lecture you will be able to,

- Describe the origins and maintenance of human genetic variation
- Understand how pieces of the genome can be traced through populations
- Explain that genetic variation persists, within and between human populations and explain how variation can explain differences in phenotypes between populations, with examples



Human genetic variation

- Why are we so phenotypically different?
- Common vs. rare:
 - Common variants – minor allele frequency (polymorphism) > 1% in a population
 - Rare variants – <1% in a population
- Neutrality
 - The vast majority of genetic variants are likely neutral = no contribution to phenotypic variation
 - Some may reach significant frequencies, but this is chance 

The origin of genetic variation

- Evolution is based on genetic change in a population. Where does genetic change come from?
- The rate of evolution depends on the amount of genetic change. What maintains genetic variation?
- Classical view was that evolution acted very slowly and there is not a lot of variation in nature.
- Now more evidence that there is a lot of genetic variation upon which selection can act.
- The study of human genetic variation has evolutionary significance and medical applications. It may be important because some disease-causing alleles occur more often in people from specific geographic regions.



Genetic variation is greatest in Africa

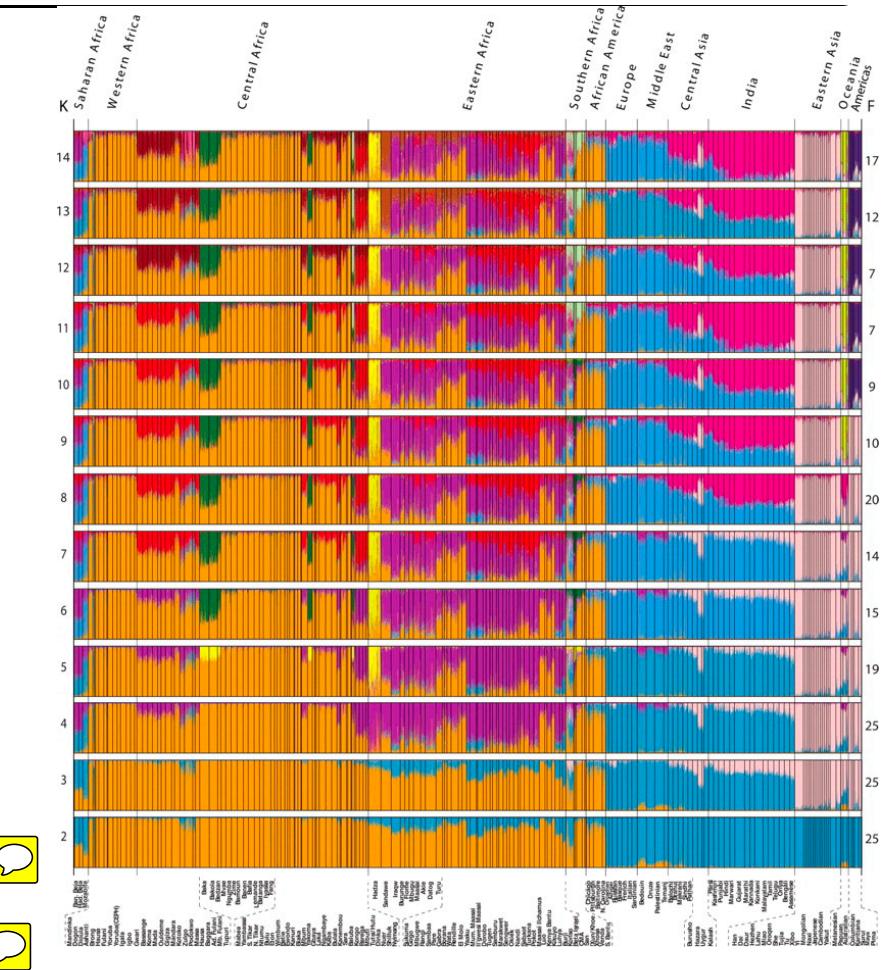


-
- Modern humans originated in Africa ~200,000 years ago and then spread across the rest of the globe within the past ~100,000 years
 - Thus, modern humans have existed continuously in Africa longer than in any other geographic region and have maintained relatively large effective population sizes, resulting in high levels of within-population genetic diversity
 - Africa contains more than 2000 distinct ethnolinguistic groups representing nearly one-third of the world's languages
 - Because of considerable environmental diversity, African populations show a range of linguistic, cultural, and phenotypic variation

Genetic variation is greatest in Africa

- Studied 121 African populations, four African American populations, and 60 non-African populations for patterns of variation at 1327 nuclear microsatellite and insertion/deletion markers
- Identified 14 ancestral population clusters in Africa that correlate with self-described ethnicity and shared cultural and/or linguistic properties
- Observed high levels of mixed ancestry in most populations, reflecting historical migration events across the continent.
- Provide evidence for shared ancestry among geographically diverse hunter-gatherer populations

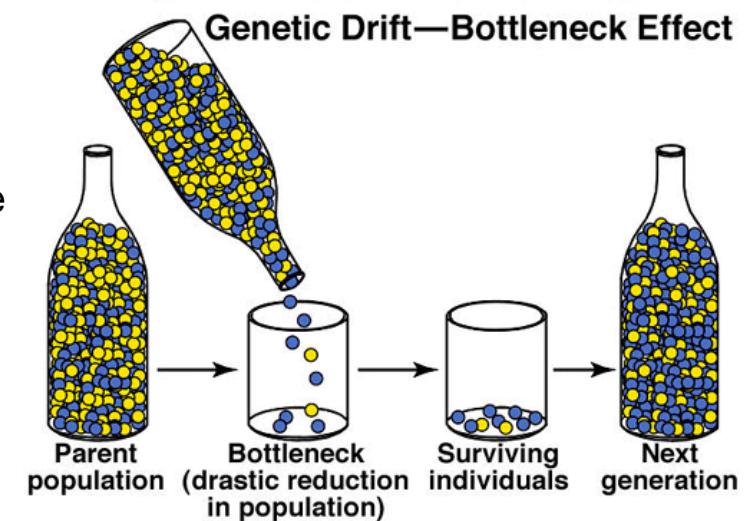
Tishkoff et al., 2009. The Genetic Structure and History of Africans and African Americans. *Science*, 324(5930) 1035-44.



Out of Africa

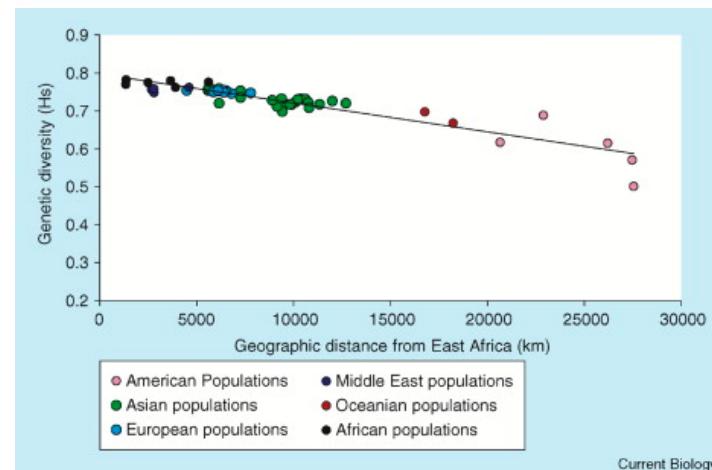
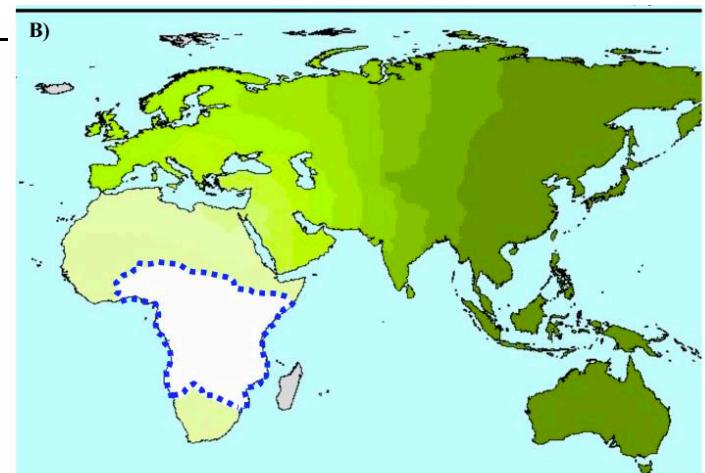
- Founder effects: when populations expand, local groups become isolated from each other and genetic differences begin to accumulate
- Each subsequent founder effect further reduces variation (specific populations in Africa are found to be the most diverse among all humans)

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



Ancient population bottlenecks on human diversity

- Dataset comprising 51 populations distributed worldwide that have been typed at 377 autosomal microsatellite loci
- Africa is the most genetically diverse place on Earth
- Geographic distance from East Africa along likely colonisation routes is an excellent predictor for genetic diversity of human populations
- History of colonisation of the world characterised by a very large number of small bottlenecks

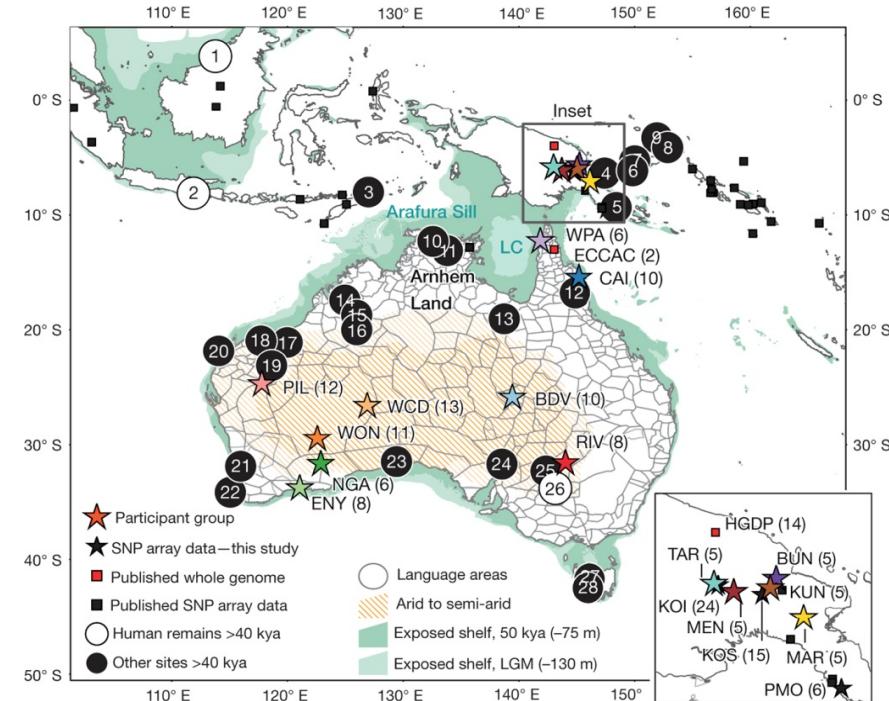


Prugnolle et al., 2005. Geography predicts neutral genetic diversity of human populations. Current Biology. 15(5), R159-E160.

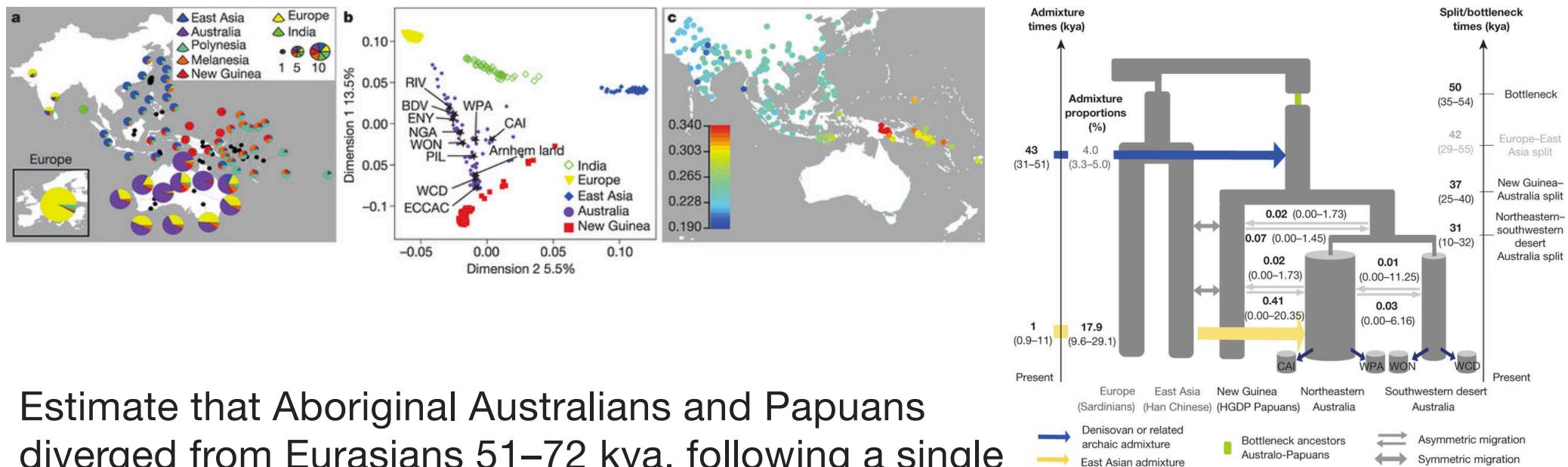
Current Biology

A genomic history of Aboriginal Australia

- During most of the last 100,000 years, Australia, Tasmania and New Guinea formed a single continent, Sahul, which was separated from Sunda (the continental landmass including mainland and western island Southeast Asia) by a series of deep oceanic troughs never exposed by changes in sea level
- Colonization of Sahul is thought to have required at least 8 sea crossings between islands, potentially constraining the occupation of Australia and New Guinea by earlier hominins.
- Aboriginal Australian and Papuan samples used in this study, as well as archaeological sites and human remains dated to ~40-60 kya in southern Sunda and Sahul



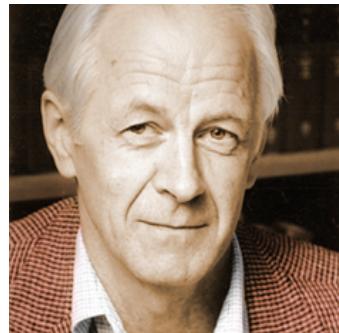
A genomic history of Aboriginal Australia



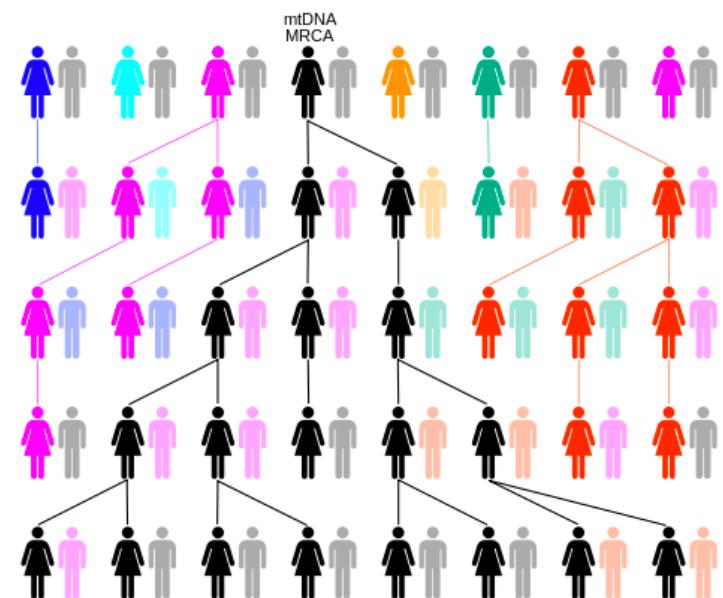
Estimate that Aboriginal Australians and Papuans diverged from Eurasians 51–72 kya, following a single out-of-Africa dispersal



Mitochondrial Eve

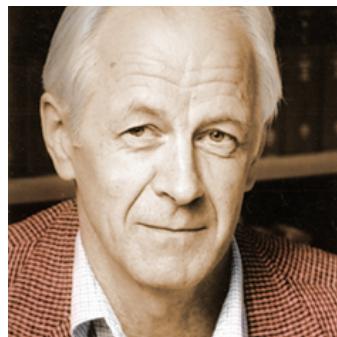


Allan Wilson

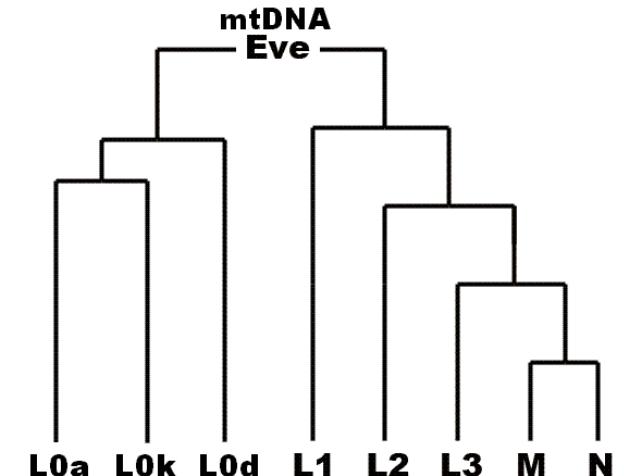


Mitochondrial Eve

- We can use mitochondrial DNA to reconstruct human evolutionary history through maternal lines
- Wilson et al (1987) collected mtDNA samples from people all over the world
- Using phylogenetics it is possible to find the ancestral sequence
- The most recent matrilineal common ancestor of all currently living humans, i.e., the most recent woman from whom all living humans descend in an unbroken line purely through their mothers

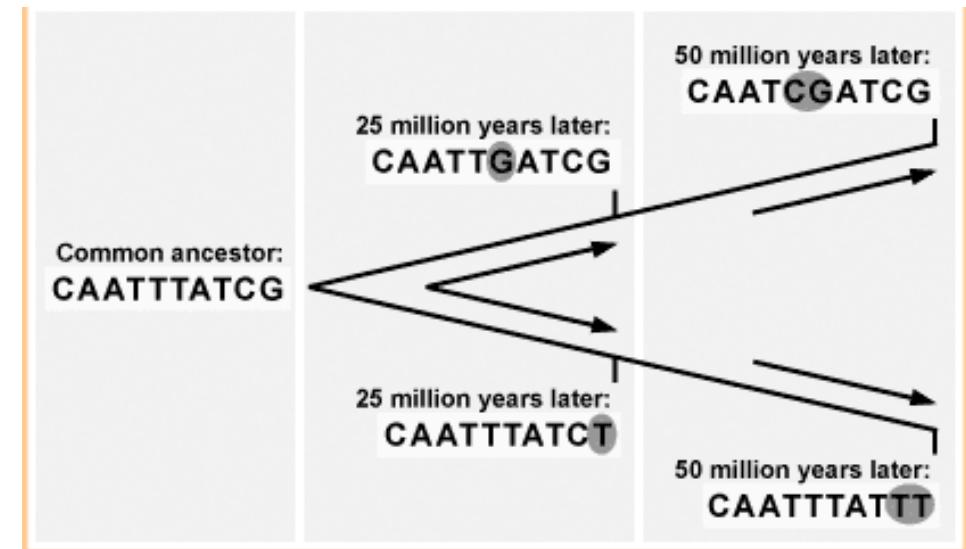


Allan Wilson



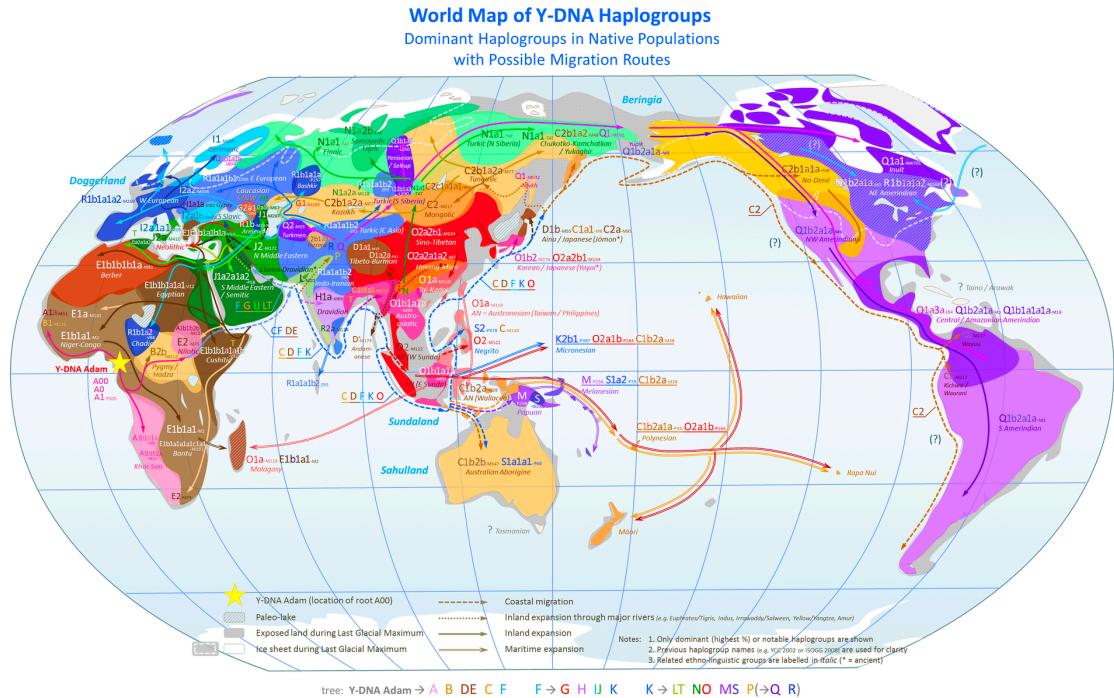
Mitochondrial Eve

- Along any particular line of descent, mitochondrial DNA accumulates mutations at the rate of approximately one every 3,500 years per nucleotide
- The molecular clock allows us to date divergence events
- 170,000 and 100,000 years ago



Y chromosome Adam

- Y-chromosomal most recent common ancestor
 - All currently living males are directly derived from the Y chromosome of this remote ancestor
 - Not permanently fixed to a single individual, but can advance over the course of human history as paternal lineages become extinct
 - Estimates of the age of the Y-MRCA range around 160,000 to 300,000 years ago
 - Likely origin is the north-western quadrant of the African continent



Scozzari et al., 2012. Molecular Dissection of the Basal Clades in the Human Y Chromosome Phylogenetic Tree. PLoS One, 7(11): e49170.

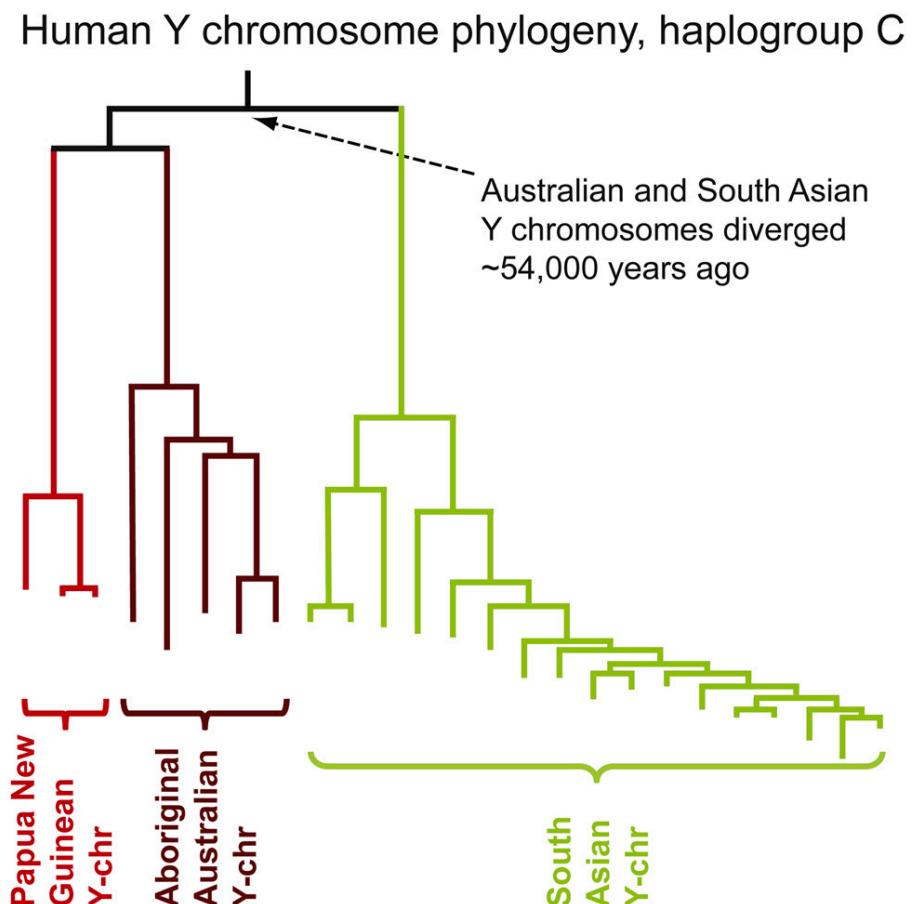
Deep Roots for Aboriginal Australian Y Chromosomes

Australia was one of the earliest regions outside Africa to be colonized by fully modern humans

Sequenced 13 Aboriginal Australian Y chromosomes to investigate their divergence times from Y chromosomes in other continents

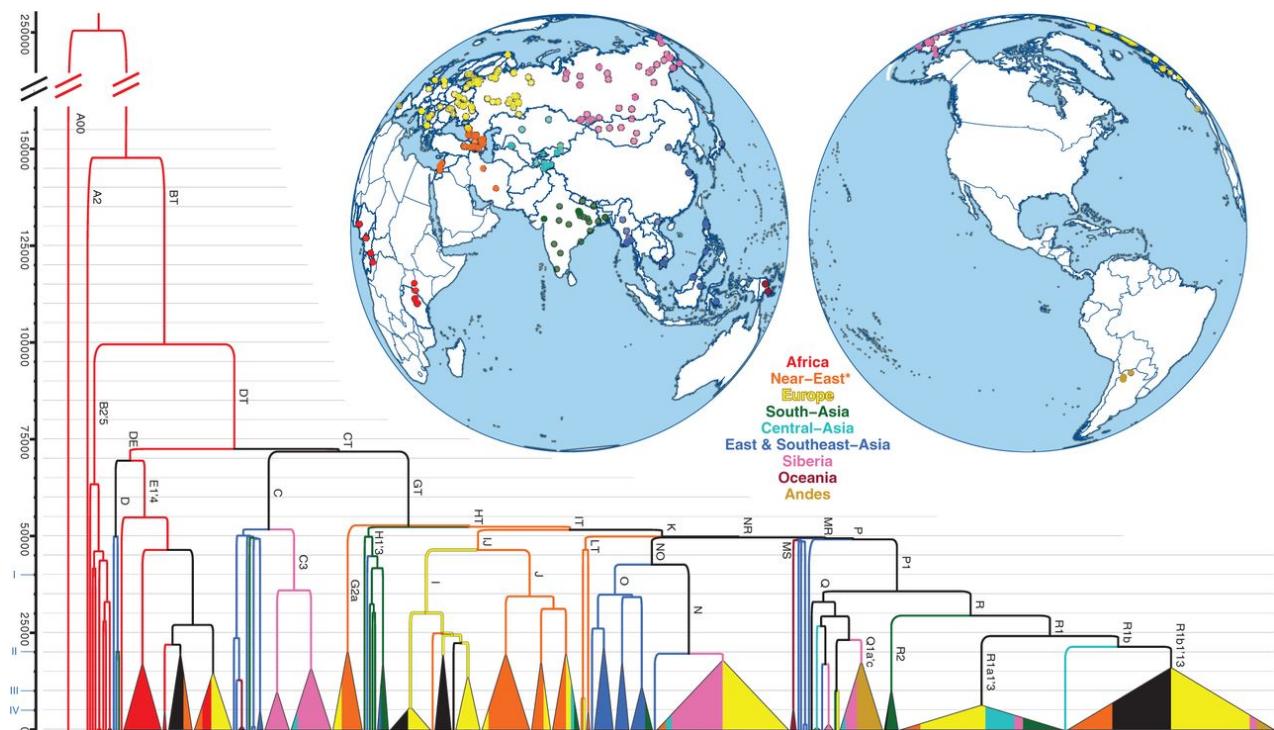
Divergence times dating back to ~50 kya, thus no evidence for recent gene flow from India into Australia

Bergstrom et al., 2016. Current Biology.



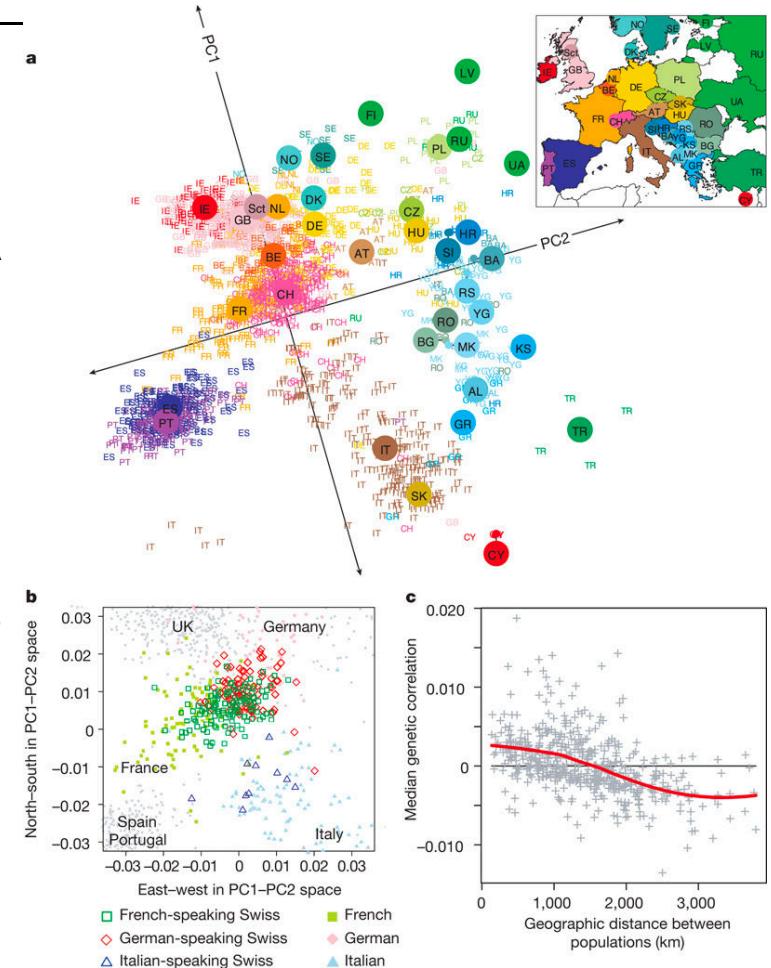
Human Y-chromosome variation

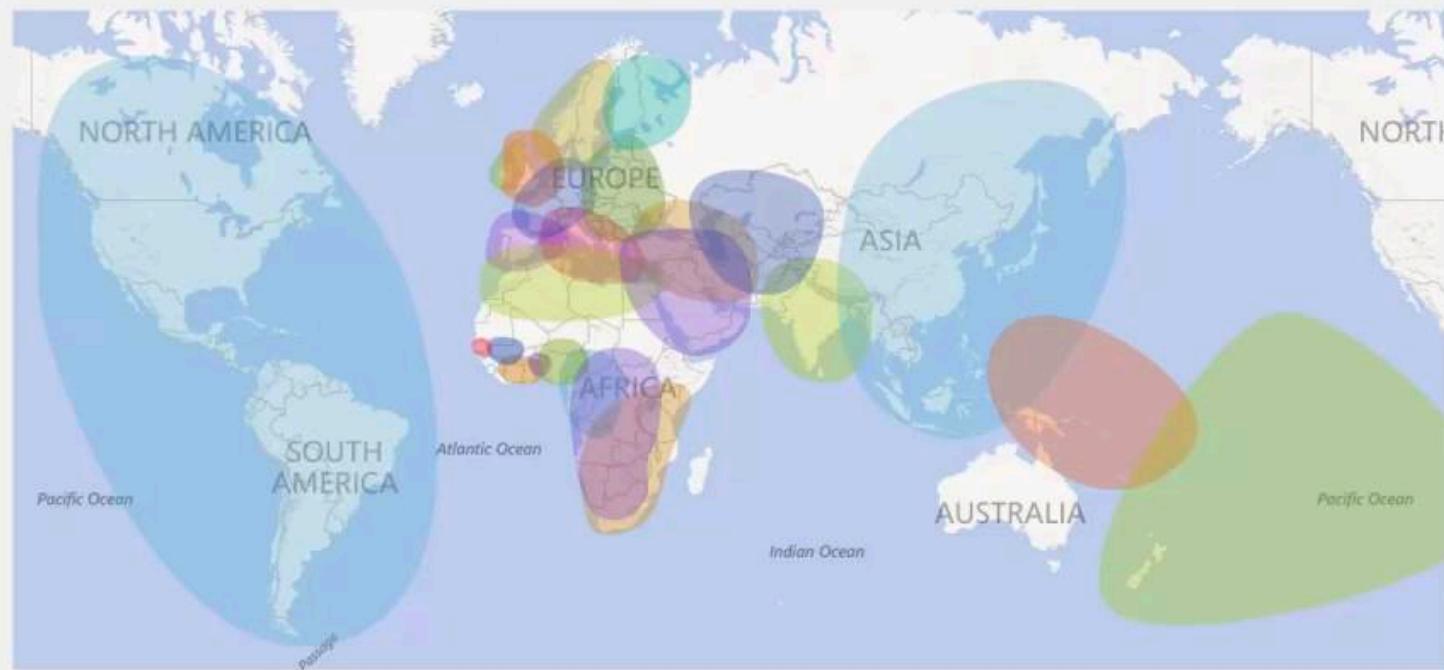
- The phylogenetic tree of 456 whole Y chromosome sequences and a map of sampling locations.
 - The data infer recent bottleneck in Y-chromosome lineages dating to the last 10k y
 - Hypothesise that recent bottlenecks are caused by cultural changes



Genes mirror geography within Europe

- Advances in high-throughput genotyping technology have markedly improved our understanding of global patterns of human genetic variation
- 3,000 European individuals genotyped at over half a million variable DNA sites in the human genome
- Despite low average levels of genetic differentiation among Europeans, found close correspondence between genetic and geographic distances
- A geographical map of Europe arises naturally as an efficient two-dimensional summary of genetic variation in Europeans
- The results emphasise that when mapping the genetic basis of a disease phenotype, associations can arise if genetic structure is not properly accounted for
- The results are relevant to the prospects of genetic ancestry testing; an individual's DNA can be used to infer their geographic origin with surprising accuracy



**AFRICA**

Africa North
Africa South-Central
Hunter-Gatherers
Africa Southeastern Bantu
Benin/Togo
Cameroon/Congo
Ivory Coast/Ghana
Mali
Nigeria
Senegal

AMERICA

Native American
ASIA
Asia Central
Asia East
Asia South

EUROPE

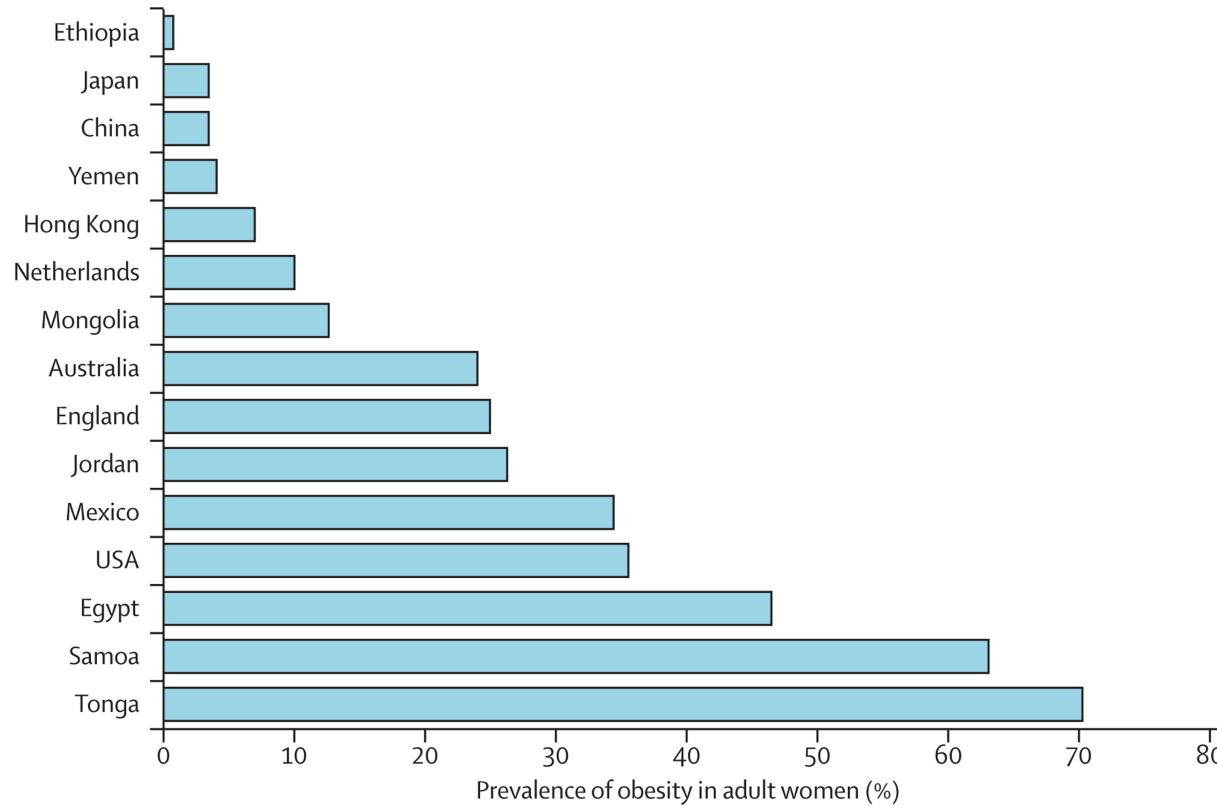
Europe East
Europe West
European Jewish
Finland/Northwest Russia
Great Britain
Iberian Peninsula
Ireland
Italy/Greece
Scandinavia

PACIFIC ISLANDER

Melanesia
Polynesia
WEST ASIA
Caucasus
Middle East

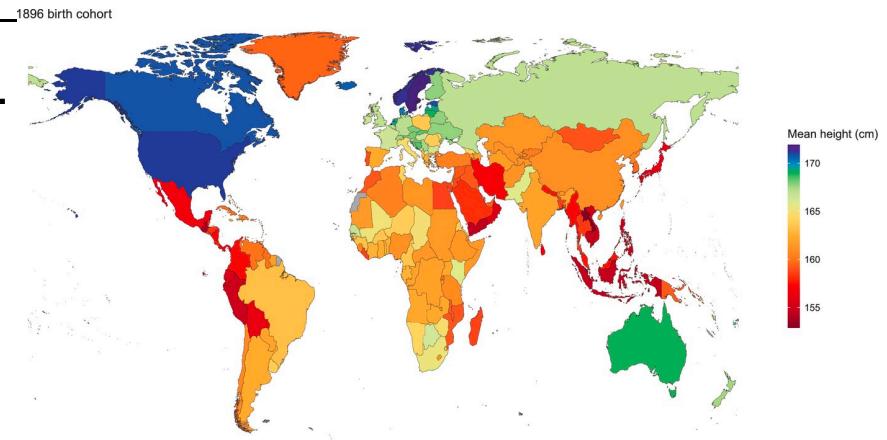
Variation in phenotype

e.g. obesity is highly variable



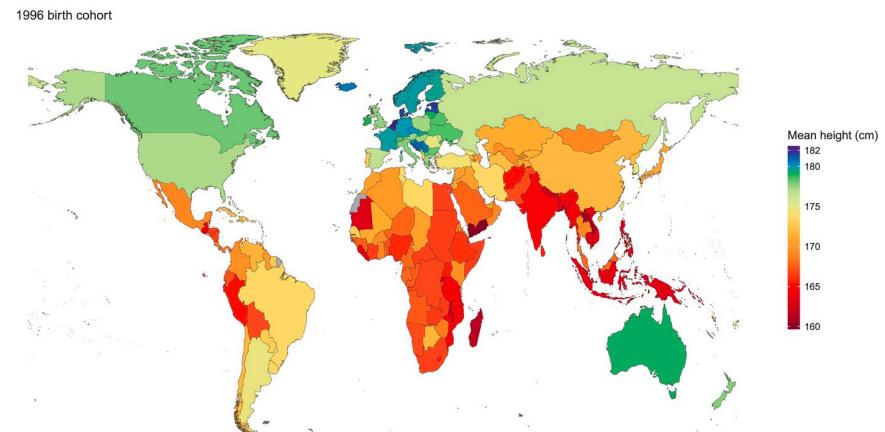
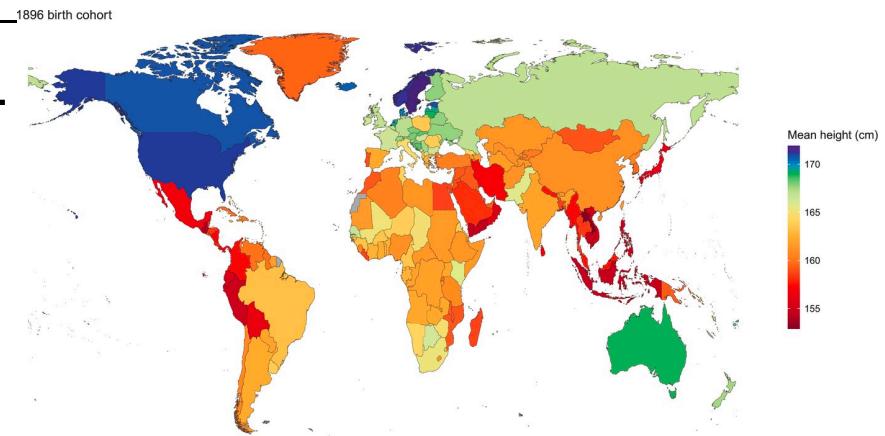
Adult human height

-
- People from different countries grow to different heights.



Adult human height

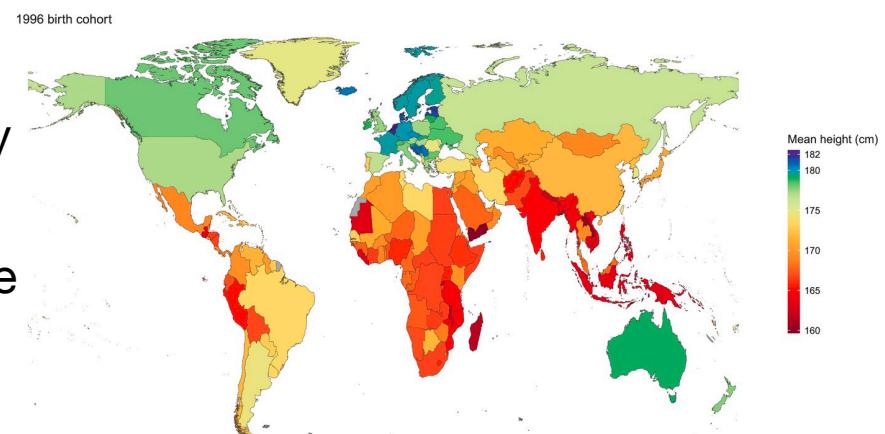
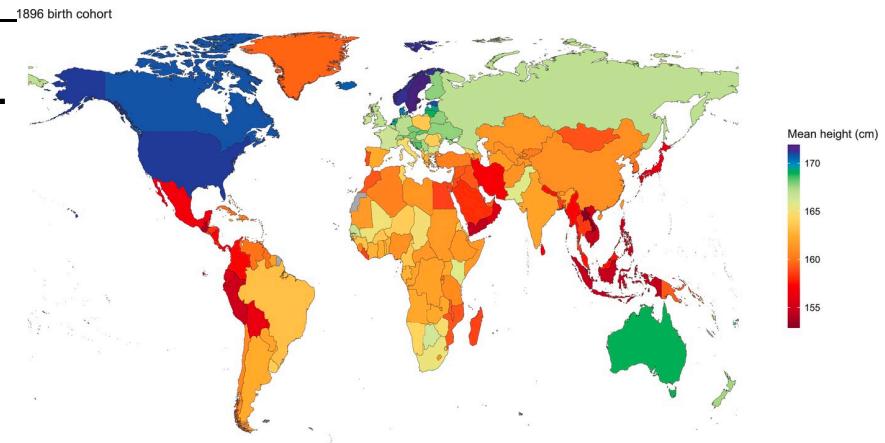
- People from different countries grow to different heights.



DOI: [10.7554/eLife.13410](https://doi.org/10.7554/eLife.13410)

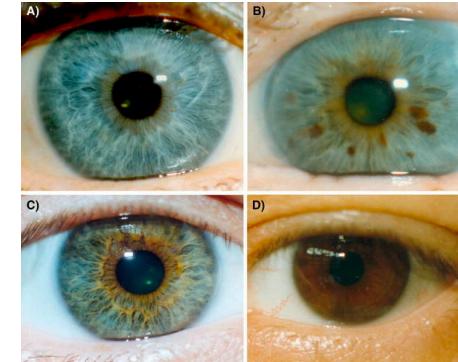
Adult human height

- People from different countries grow to different heights.
- This may be partly due to genetics, but also due to environmental causes.
- For example, children and adolescents who are malnourished, or who suffer from serious diseases, will generally be shorter as adults.
- This is important because taller people generally live longer, are less likely to suffer from heart disease and stroke, and taller women and their children are less likely to have complications during and after birth.
- The tallest people are men born in the Netherlands in the last quarter of 20th century, whose average heights surpassed 182.5 cm



Blue eye colour in humans may be caused by a perfectly associated founder mutation

- By linkage analysis of a large Danish family, finemapped the blue eye color locus to a 166 Kbp region within the *HERC2* gene
- One single haplotype from Denmark
- Data suggest a common founder mutation



Species	Eye color	DNA-Library	DNA sequence
Homo	Blue	hg18_dna	TTCATTTGAGCAT AAAGTG CAAGTTCTGCACGCTAT
Homo	Brown	hg18_dna	TTCATTTGAGCAT AAAATGT CAAGTTCTGCACGCTAT
Chimpanzee	Brown	panTro2_dna	TTCATTTGAGCAT AAAATGT CAAGTTCTGCACGCTAT
Rhesus monkey	Brown	rheMac2_dna	TTCATTTGAGCAT AAAATGT CAAGTTCTGCACGCTAT
Horse	Brown	equCab1_dna	TTCACTTGACGCT AAAATGT CAAGTGCTGCACAATGT
Cow	Brown	bosTau2_dna	TTCACTCTGCACGG AAAATGT CAAGTAC-ACACACTGT
Cat	Brown-yellow	felCat3_dna	TTCATTTGCATGT AAAATGT CAAGTACCAACACAATAC
Dog	Brown-yellow	canFam2_dna	TTCATTTGCATGT AAAATGT CAAGTGC-ACACAATAT
Rat	Brown	rn4_dna	TTCATTTGCCGTATT AAAATGT CAA
Mouse	Brown	Mm8_dna	TTCATTTGCCGTAT AAAATGT CAAATGCCATGCACTAT
Consensus sequence - blue eye		Ttca-ttg---- taaGtgtcaa-t-c---- c-tat	
Consensus sequence - brown eye		Ttca-ttg---- taaAtgtcaa-t-c---- c-tat	
Nkx-2.5 target site; match allele for blue eye color			TYAAGTG
CdxX-1 target site; match allele for brown eye color			YAKWAWW

Eiberg, H. et al (2008) *Human Genetics*

Lecture 7: Human genetic diversity and evolution



On completion of this lecture you will be able to,

- Describe the origins and maintenance of human genetic variation
- Understand how pieces of the genome can be traced through populations
- Explain that genetic variation persists, within and between human populations and explain how variation can explain differences in phenotypes between populations, with examples



MACQUARIE
University

BIOL3120 –Human Genetics and Evolutionary Medicine

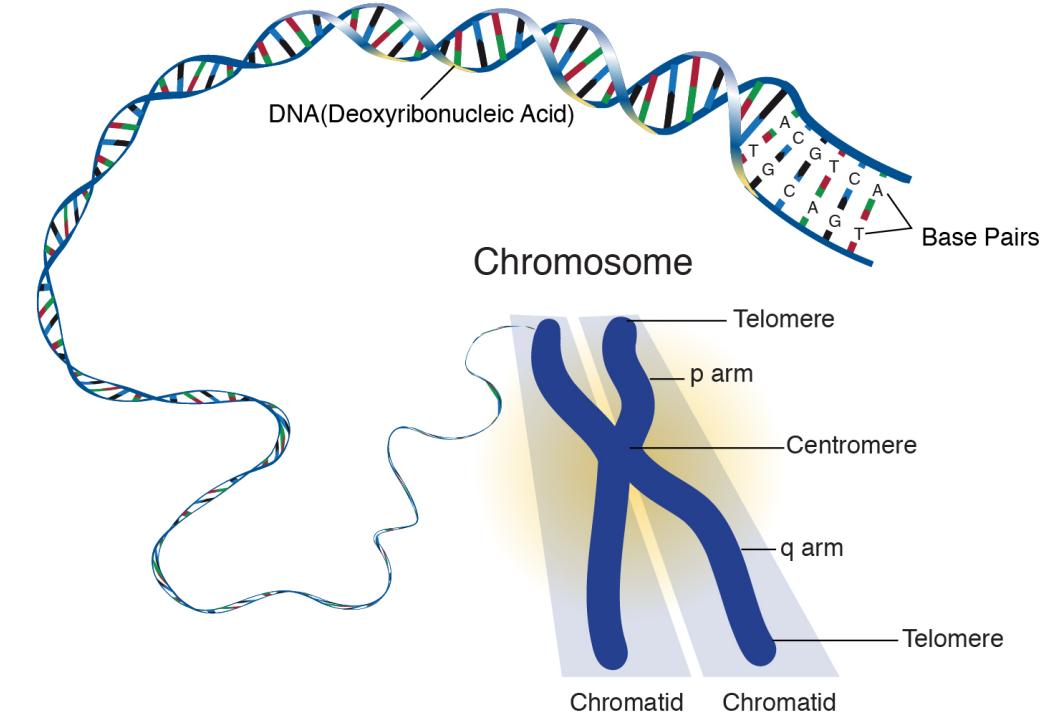
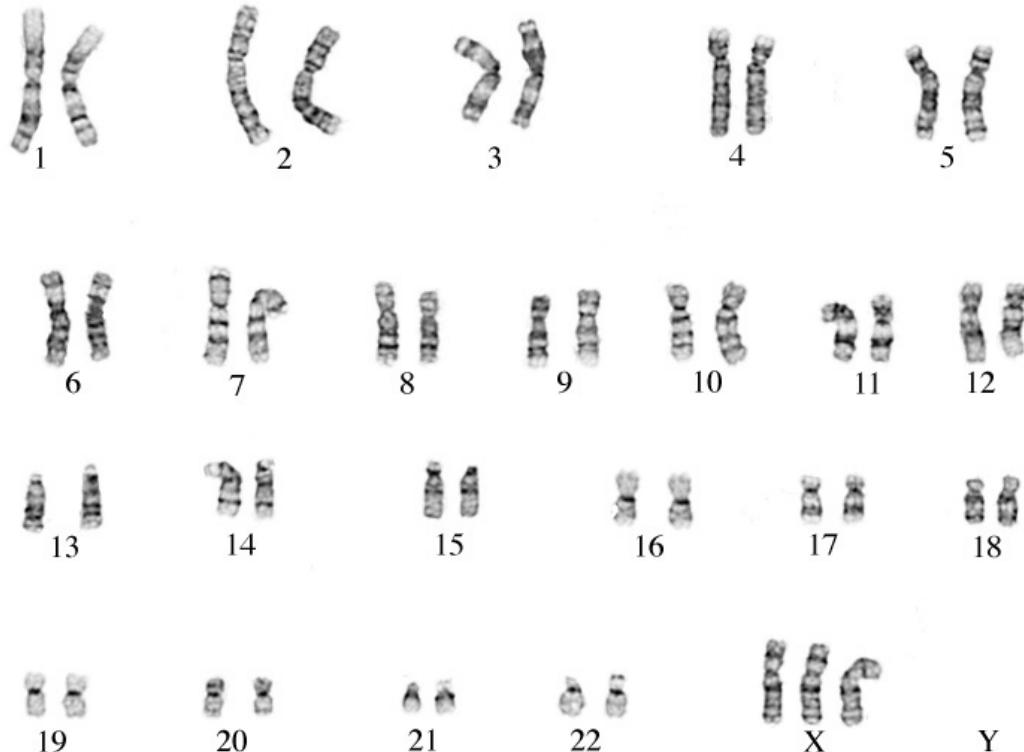
Genetic Testing Techniques





6	Genetic Testing Techniques GWAS	Problem Set 4	Problem Set 4 (5%)	Explain the importance of new techniques in human genetics for understanding human disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources
Recess				
Recess		Pracs for External Students only		
7	Treatment for Genetic Conditions Epigenetics and Imprinting	Problem Set 5	Problem Set 4 (5%) & Problem Set 5 (5%)	Explain the importance of new techniques in human genetics for understanding human disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources

Detection of chromosomal changes vs Detection of nucleotide changes in individuals



BIOL3120 –Genetic Testing Techniques

LEARNING OBJECTIVES



On successful completion of this lecture, you will be able to:

- Identify and understand the technologies that detect chromosomal mutations
- Identify and understand the technologies that detect nucleotide mutations

Overview

- Detection of chromosomal changes
 - Karotype
 - Non-Invasive Prenatal Testing (NIPT)
 - Microarrays
- Detection of nucleotide changes
 - Mutation panels
 - Genetic linking & mapping
 - Sequencing
 - Sanger sequencing
 - Next generation sequencing
 - Triplet repeat primed PCR
 - Multiplex Ligation-dependent Probe Amplification (MLPA)



Chromosomal Changes

Karyotype

- A karyotype is an individual's collection of chromosomes.
- The term also refers to a laboratory technique that produces an image of an individual's chromosomes.
- A picture of an individual's chromosomes
- The karyotype is used to look for abnormal numbers or structures of chromosomes.

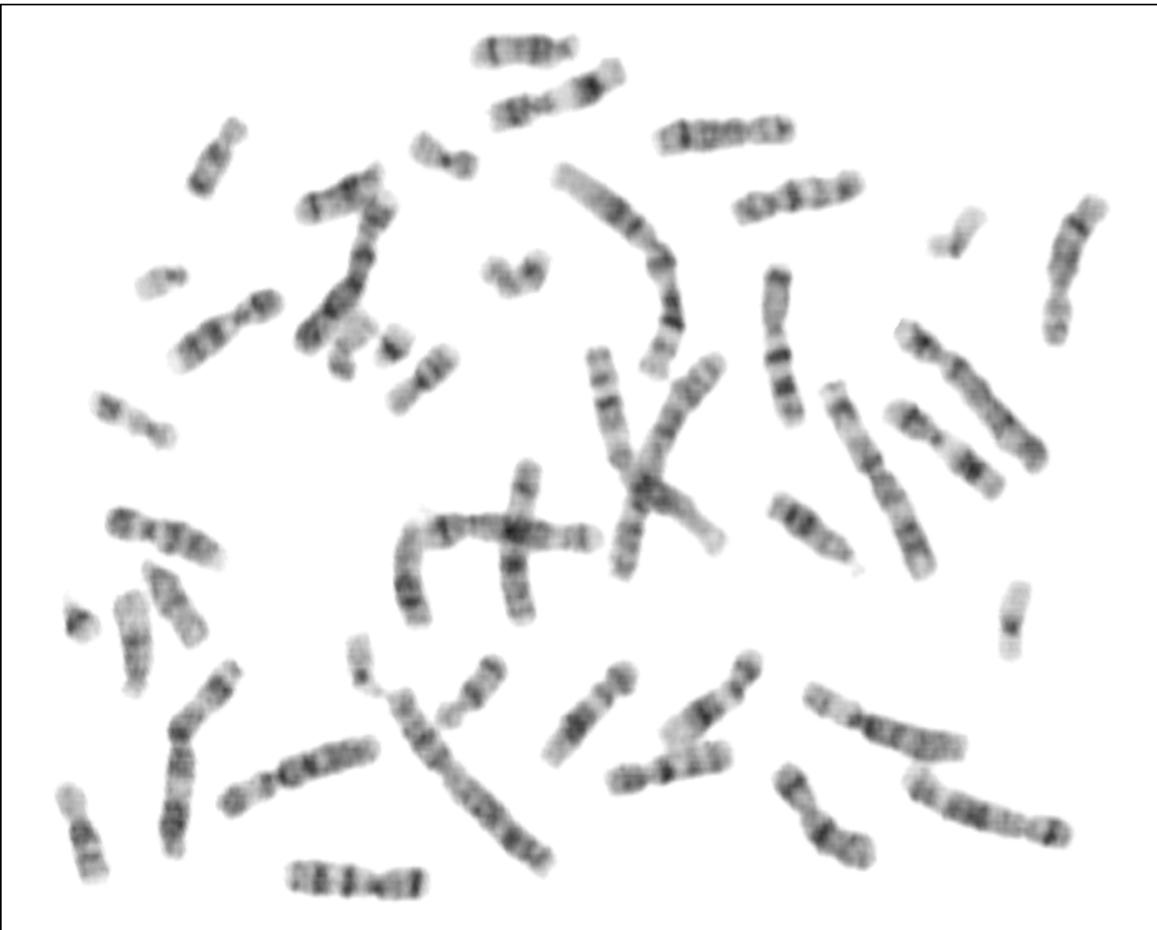
Basic Technique

- Using cells in tissue culture
- Pretreating cells in a hypotonic solution, which swells them and spreads the chromosomes
- Arresting mitosis in metaphase by a solution of colchicine
- Squashing the preparation on the slide forcing the chromosomes into a single plane
- Cutting up a photomicrograph and arranging the result into an indisputable karyogram.

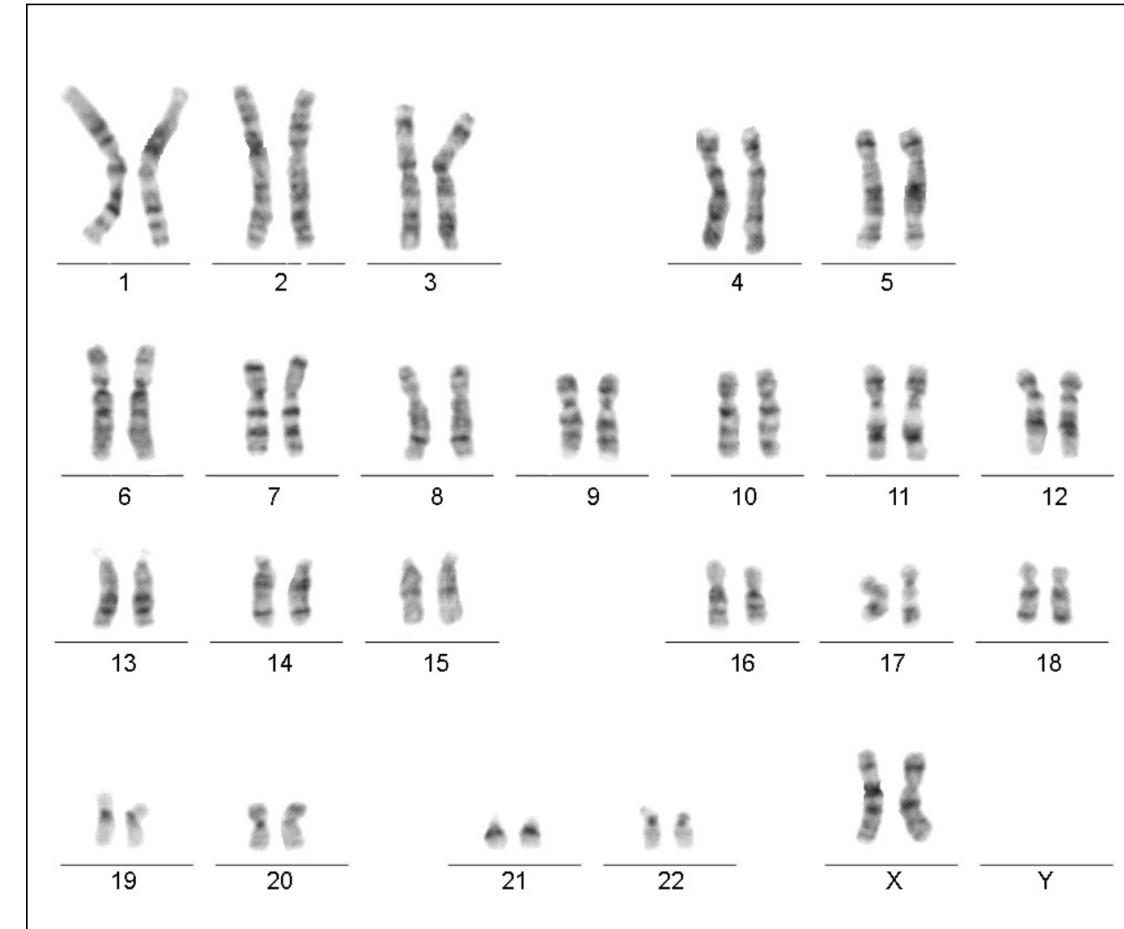
JOE HIN TJIO & ALBERT LEVAN, 1956

Karyotype

Metaphase Image

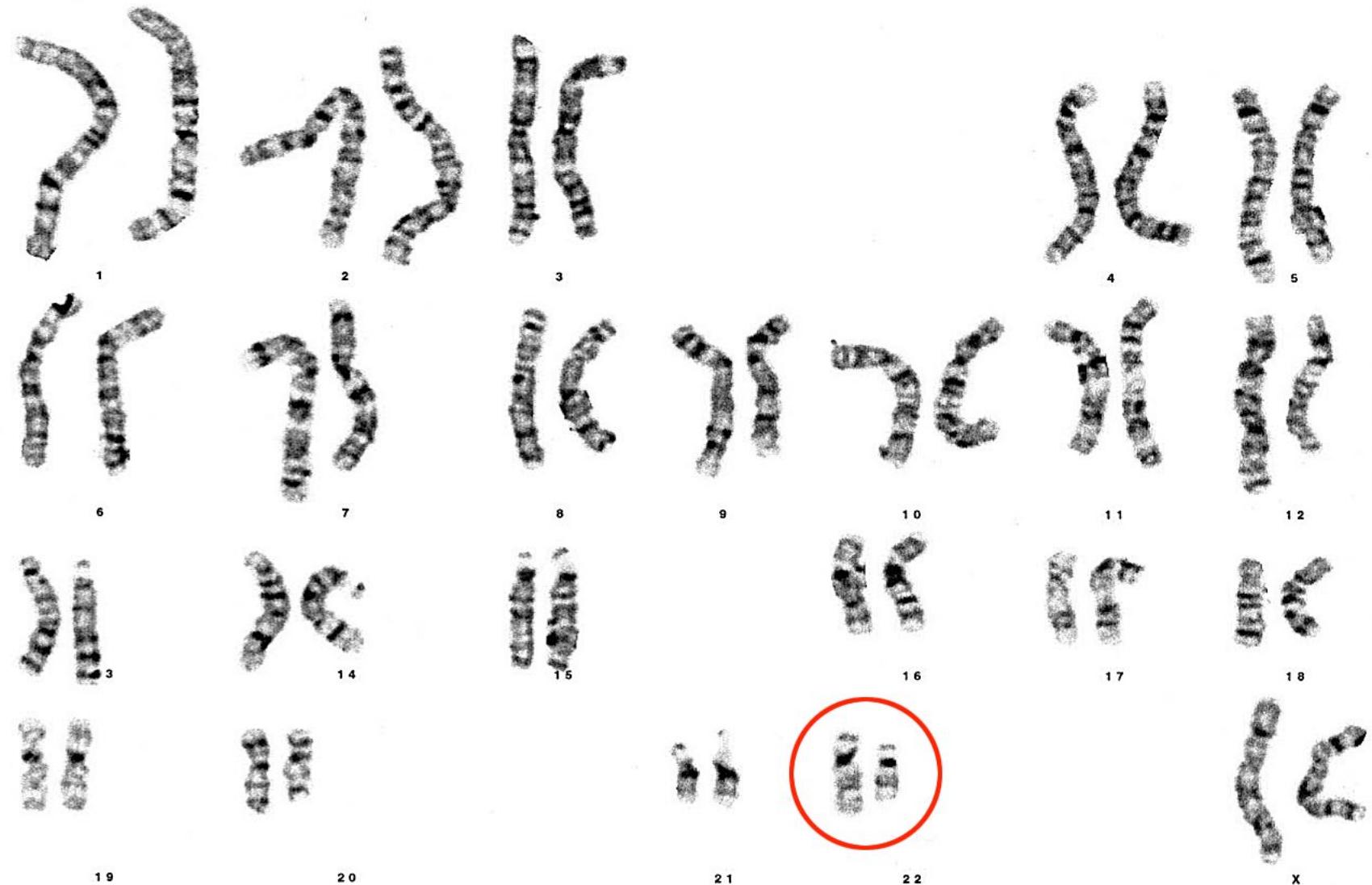


Karyotype Image

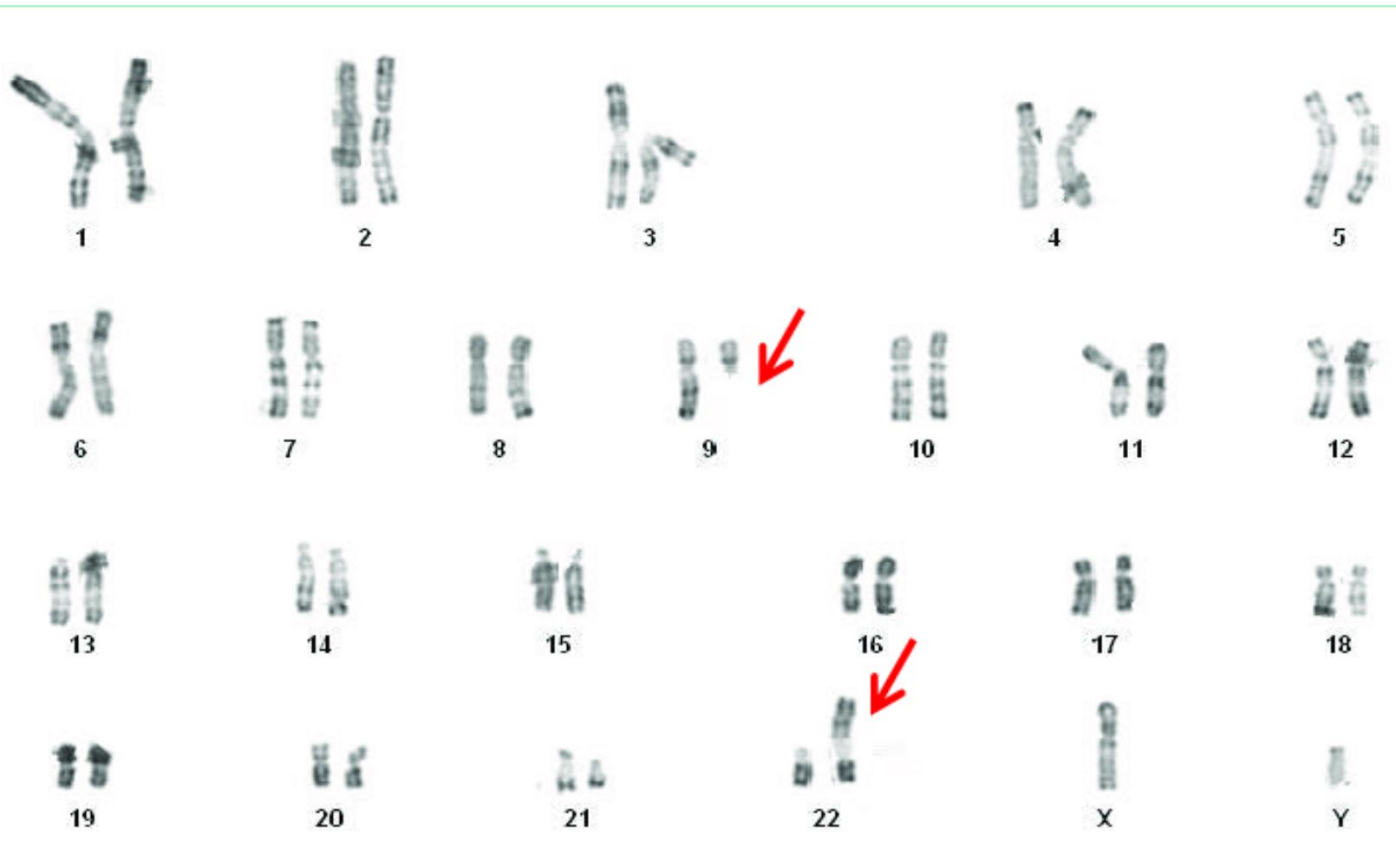


Karyotype

- ~5Mb and larger deletions/duplications detectable
- Can detect Translocations
- Mosaicism for detectable changes (count cells)

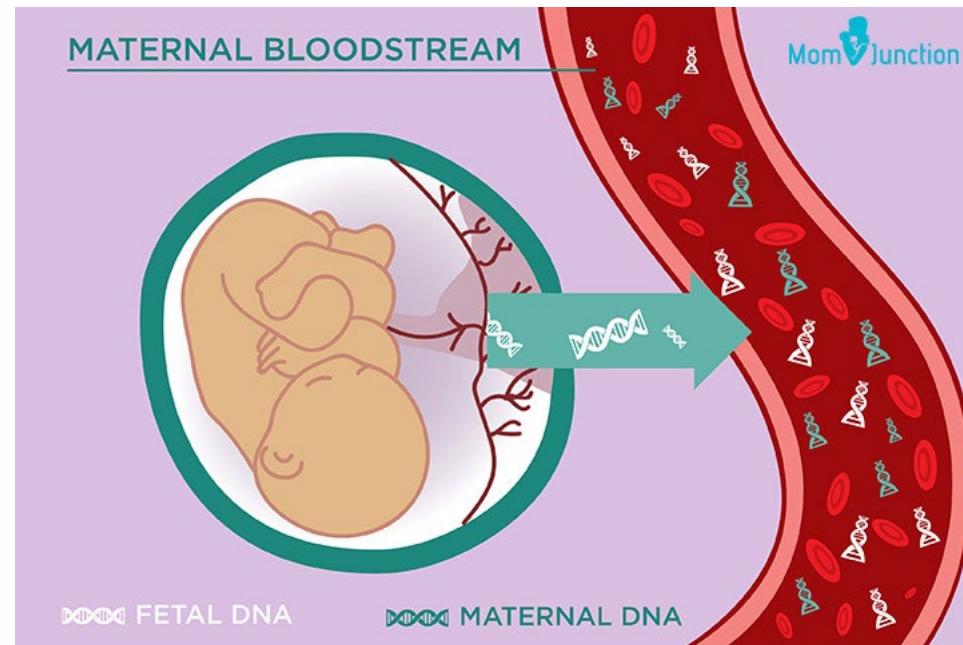


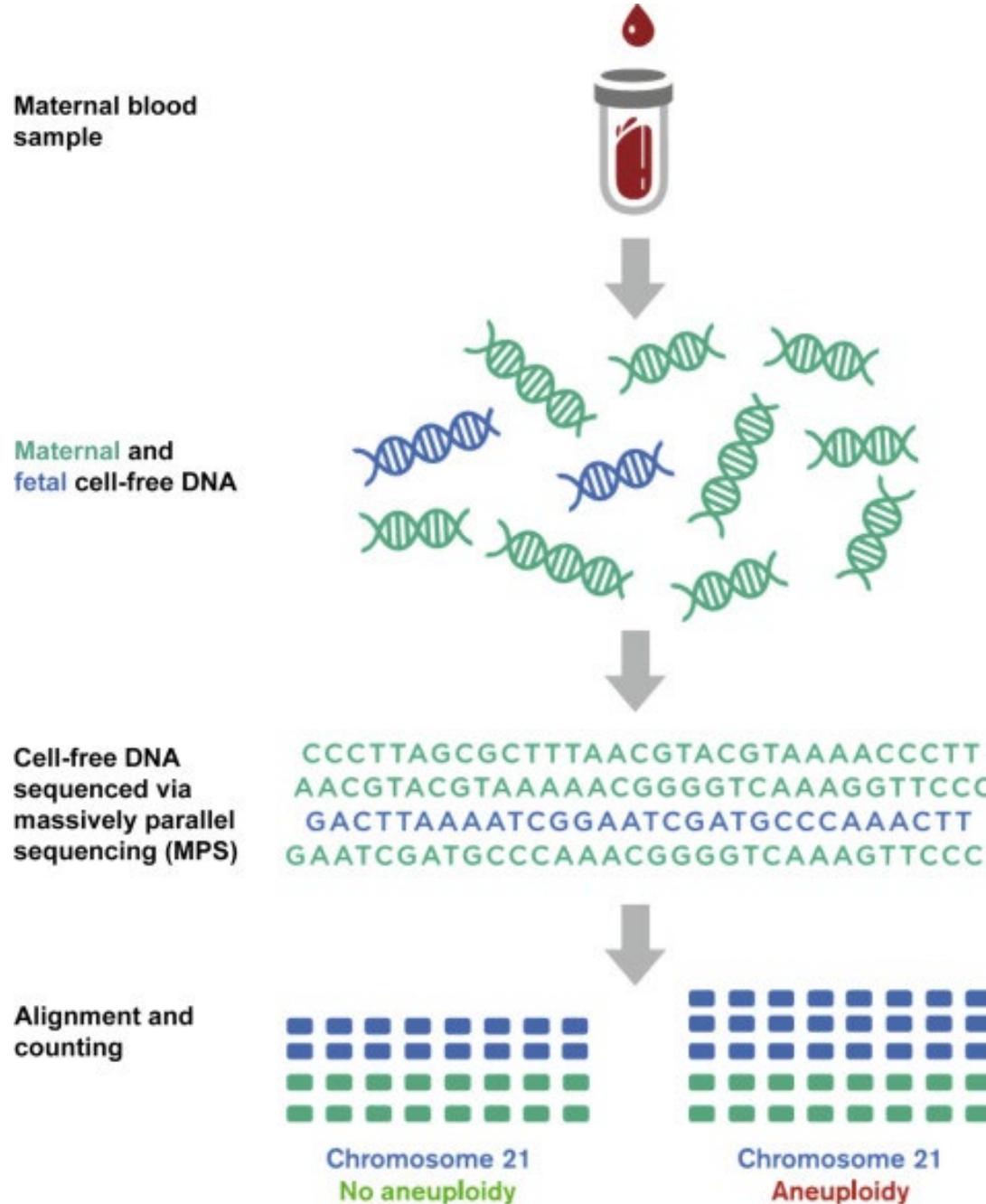
Karyotype



Non-Invasive Prenatal Testing (NIPT)

- Also called Non-invasive prenatal screening (NIPS)
- Fetal DNA present in mother's bloodstream cfDNA = cell-free DNA
- From 10 weeks





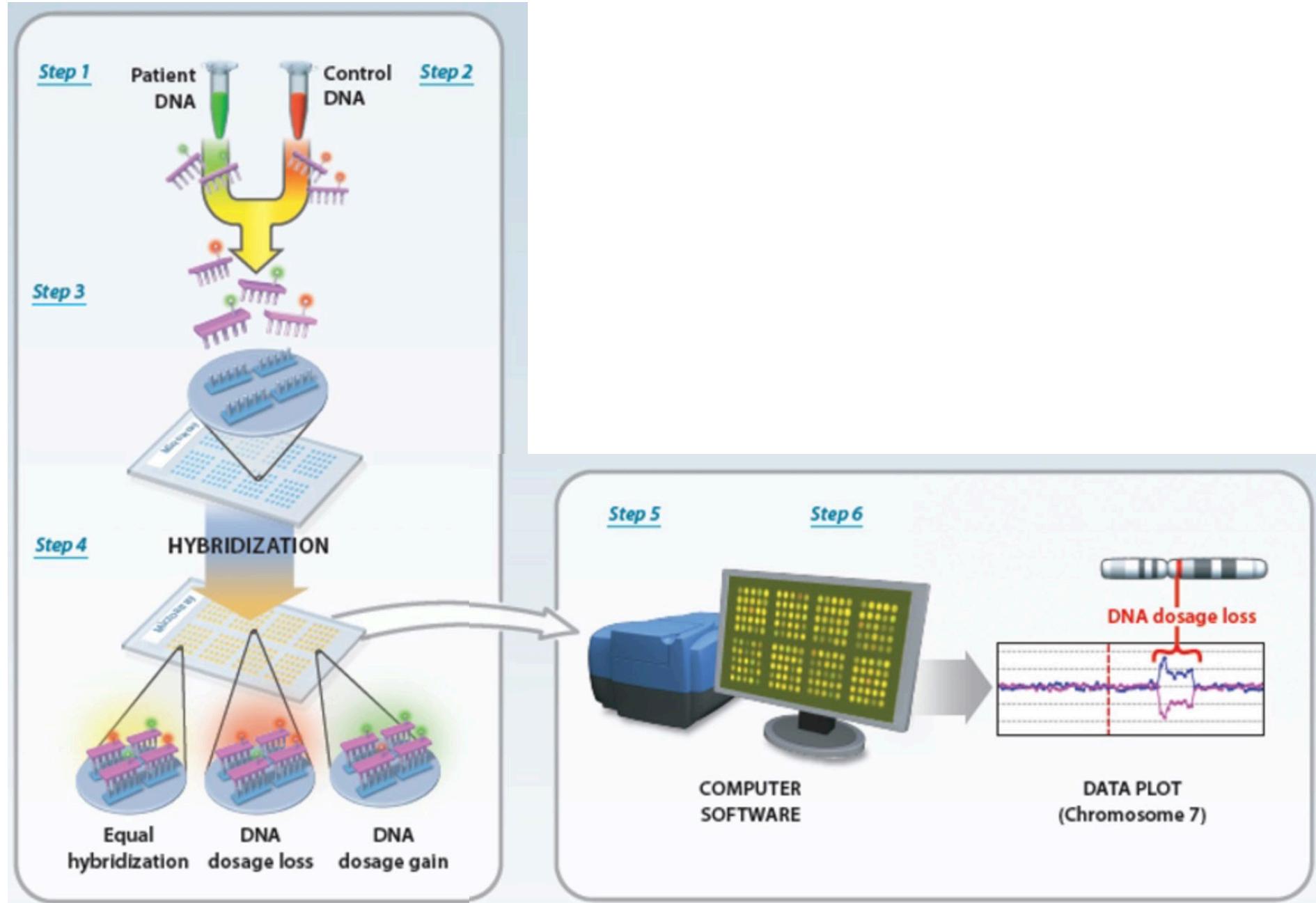
- Massively Parallel Sequencing = Next Generation Sequencing = High Throughput Sequencing
- ≠ Sangar sequencing
- To determine chromosomal aneuploidy, the most common method is to count all cfDNA fragments (both fetal and maternal).
 - If the percentage of cfDNA fragments from each chromosome is as expected, then the fetus has a decreased risk of having a chromosomal condition (negative test result).
 - If the percentage of cfDNA fragments from a particular chromosome is more than expected, then the fetus has an increased likelihood of having a trisomy condition (positive test result).
 - A positive screening result indicates that further testing (called diagnostic testing, because it is used to diagnose a disease) should be performed to confirm the result.

NIPT developing

- Some services now offering deletion/duplication
- Complete genome from cfDNA?
- Cancer detection

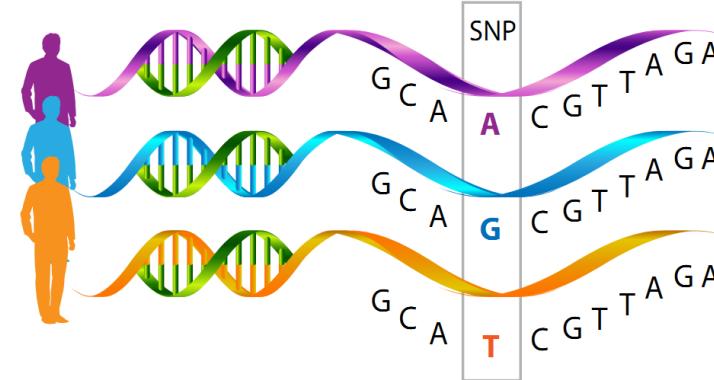
Microarray

- Detect the expression of thousands of genes at the same time
- DNA microarrays are microscope slides that are printed with thousands of tiny spots in defined positions, with each spot containing a known DNA sequence or gene
 - Referred to as gene chips or DNA chips
 - Each spot containing a known DNA sequence is referred to as a DNA probe
- cDNA is taken from an experimental sample and a reference sample and labelled with a fluorescent probe 
- The cDNA molecules bind the DNA probes on the chip
- Chip is then scanned to detect the fluorescence

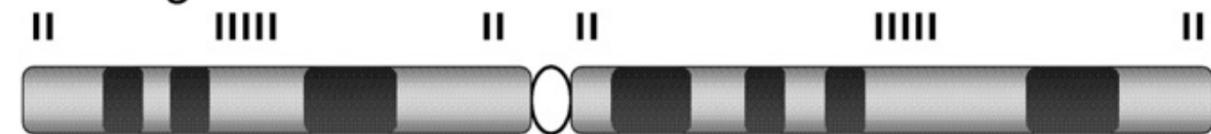


Microarray

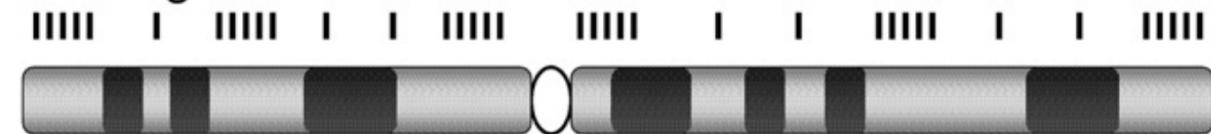
- Comparative genomic hybridisation
- Detects
 - Copy Number Variants over ~50kb
 - Homozygosity / uniparental disomy
- Can't detect
 - Balanced translocations
 - Inversions etc
 - Don't know where duplications are



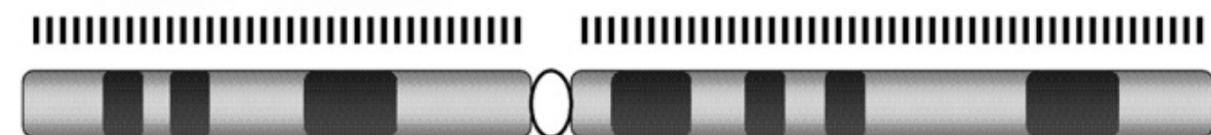
A Targeted



B Targeted with Backbone

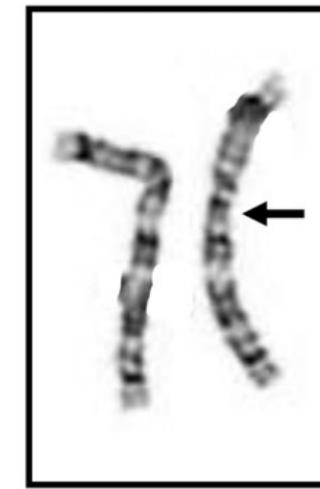
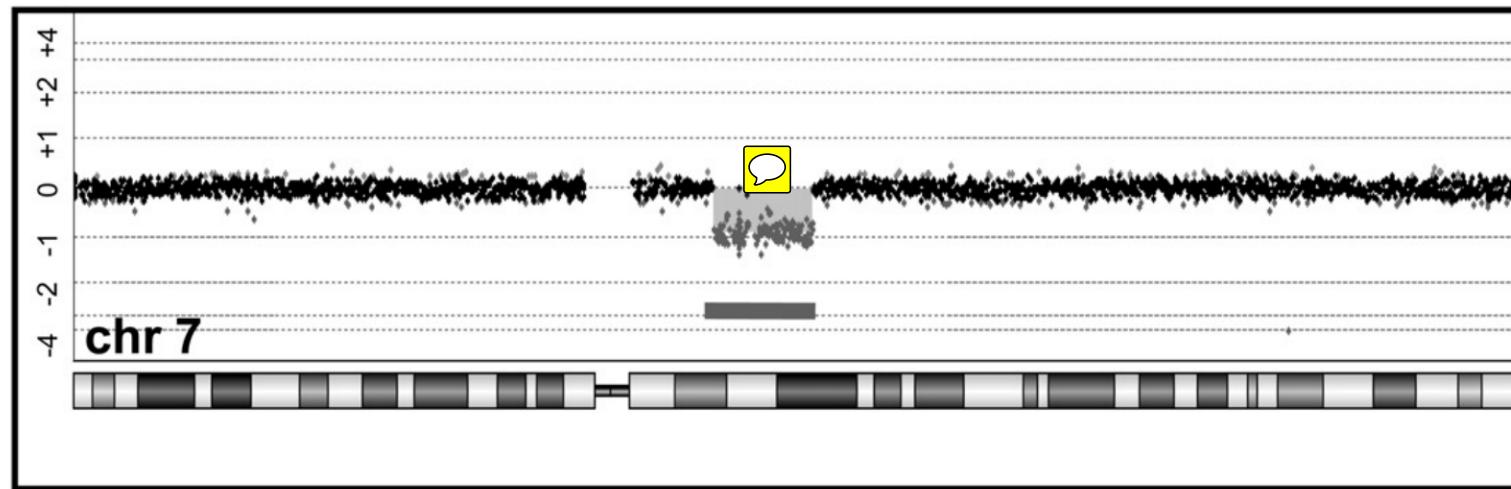


C Whole Genome

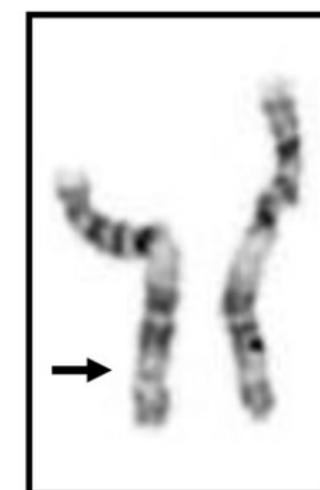
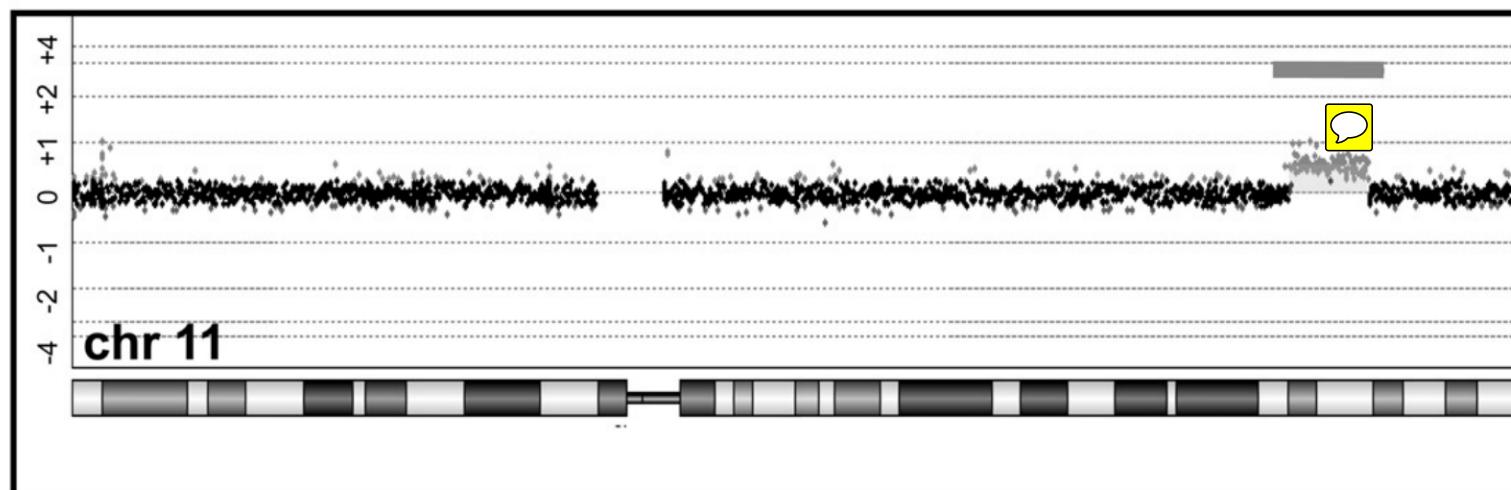


Microarray results

A



B



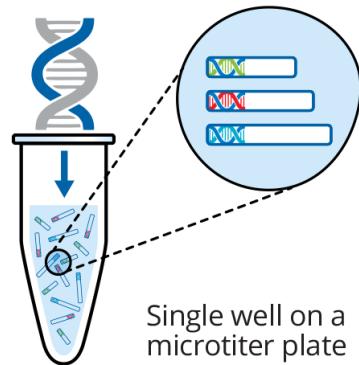


Nucleotide Mutations

Old School Methods

Targeted mutation panels

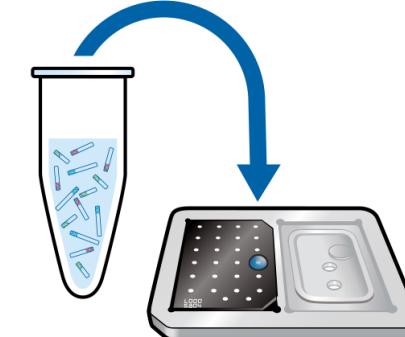
- Pre-set specific mutations to search for 
- Does not generate sequence – just tests for presence of mutations
- Quick and cheap



Endpoint PCR

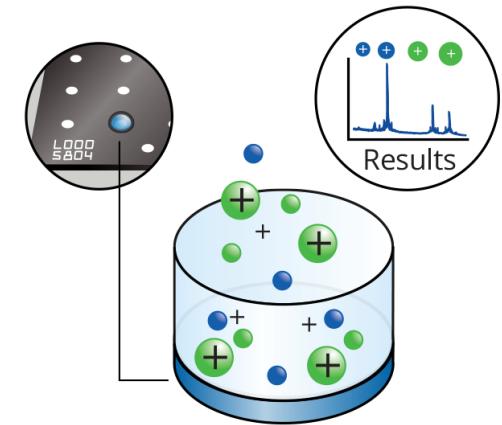
Amplify and extend up to 40 target-specific DNA fragments in a single reaction.

Sample Process Journey



Transfer Analyte

Transfer a small amount of sample to a single pad on the SpectroCHIP® Array.



Detection and Analysis

Multiple tests can be run on a single SpectroCHIP Array. Hundreds of mutations can be tested per sample.

* Use multiple reactions for >40 targets if required.

Targeted mutation panels

Cystic Fibrosis Test

Overview

Cystic fibrosis (CF) is an inherited condition affecting breathing and digestion. CF causes the build-up of thick mucus which traps bacteria, resulting in recurrent infections that damage the lungs. Thick mucus in the gut also makes digestion of food difficult. People with CF require daily physiotherapy to clear mucus from their lungs, frequent courses of antibiotics and need to take medicine to help with digestion. There is no cure for CF but better treatments are under research and development.

We screen for 175 cystic fibrosis transmembrane conductance regulator (CFTR) variants and 178 in diagnostic tests. [Download the full list of cystic fibrosis variants.](#)

Targeted mutation panels



Victorian Clinical Genetics Services
Murdoch Children's Research Institute
The Royal Children's Hospital
Flemington Road, Parkville VIC 3052
P +61 1300 11 8247 F +61 3 8341 6366
W vcgs.org.au

VCGS variant list for cystic fibrosis screening & diagnostic testing

Exon/intron number	Nucleotide change (NM_000492.3)	Protein change (NP_000483.3)	Diagnostic/screening
4	c.350G>A	p.Arg117His	Diagnostic assay only
Intron 9	c.1210-34TG[11_13]	polyTG tract	Diagnostic assay only
Intron 9	c.1210-12T[5_9]	polyT tract	Diagnostic assay only
Intron 2	c.165-3C>T	No protein name	Diagnostic and prepair® screening
Intron 2	c.165-1G>A	No protein name	Diagnostic and prepair® screening
3	c.166G>A	p.Glu56Lys	Diagnostic and prepair® screening
3	c.169T>G	p.Trp57Gly	Diagnostic and prepair® screening
3	c.170G>A	p.Trp57*	Diagnostic and prepair® screening
3	c.171G>A	p.Trp57*	Diagnostic and prepair® screening
3	c.175dupA	p.Arg59Lysfs*10	Diagnostic and prepair® screening
3	c.174_177delTAGA	p.Asp58Glufs*32	Diagnostic and prepair® screening
3	c.178G>T	p.Glu60*	Diagnostic and prepair® screening
3	c.200C>T	p.Pro67Leu	Diagnostic and prepair® screening
3	c.223C>T	p.Arg75*	Diagnostic and prepair® screening
3	c.233dupT	p.Trp79Leufs*32	Diagnostic and prepair® screening
2	c.254G>A	p.Glu85Glu	Diagnostic and prepair® screening

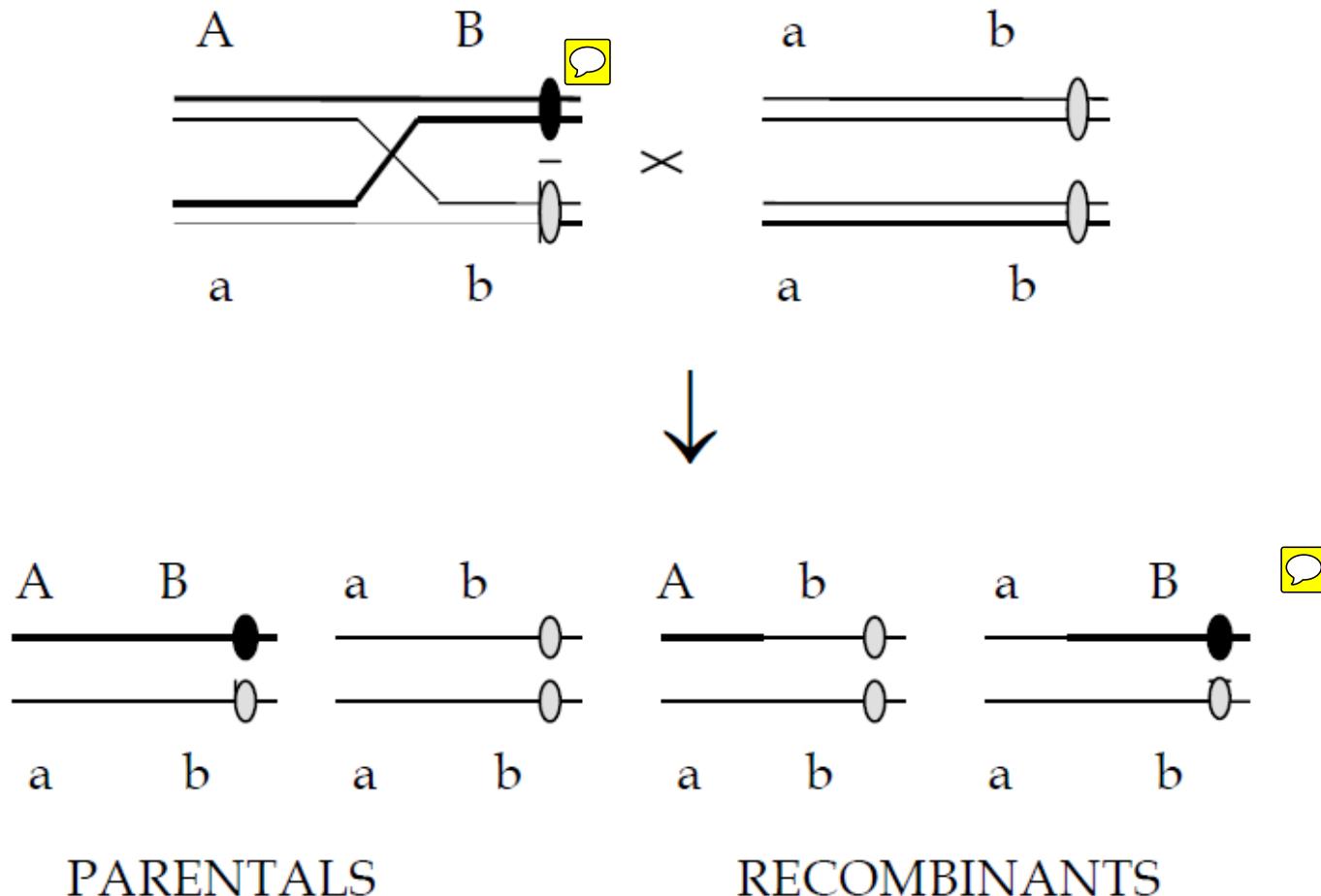
Genetic Linkage and Mapping



- Technique used to determine the locus of a disease linked mutation
- Basic principles of genetic linkage:
 - Linkage results from the location of genes on chromosomes
 - Gene loci on the same chromosome are inherited together
 - (sometimes)
 - Linkage is deduced from the progeny by observing the arrangement of alleles (haplotype)
 - Linkage phase: arrangement of alleles on each chromosome in a double heterozygote
- Individuals who have inherited a disease linked mutation are unlikely to be recombinant in that location

Genetic Linkage and Mapping

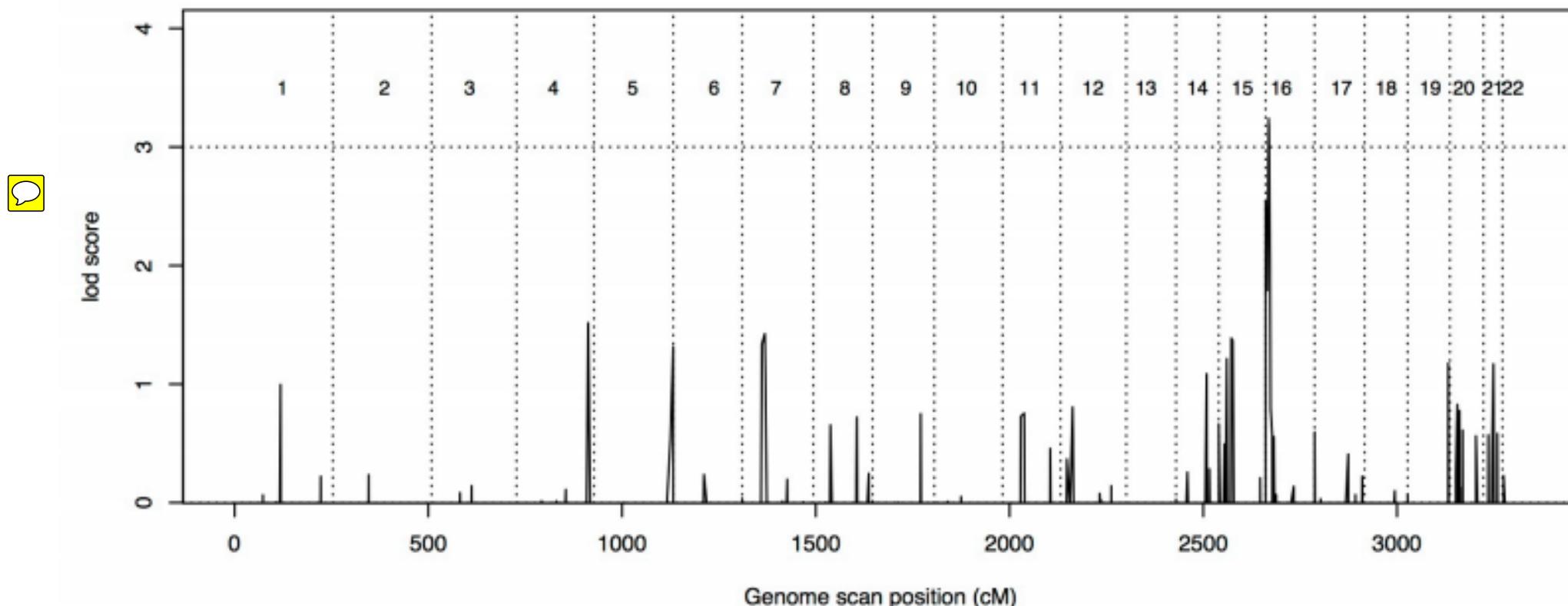
Basic testcross for linkage:



Genetic Linkage and Mapping

a

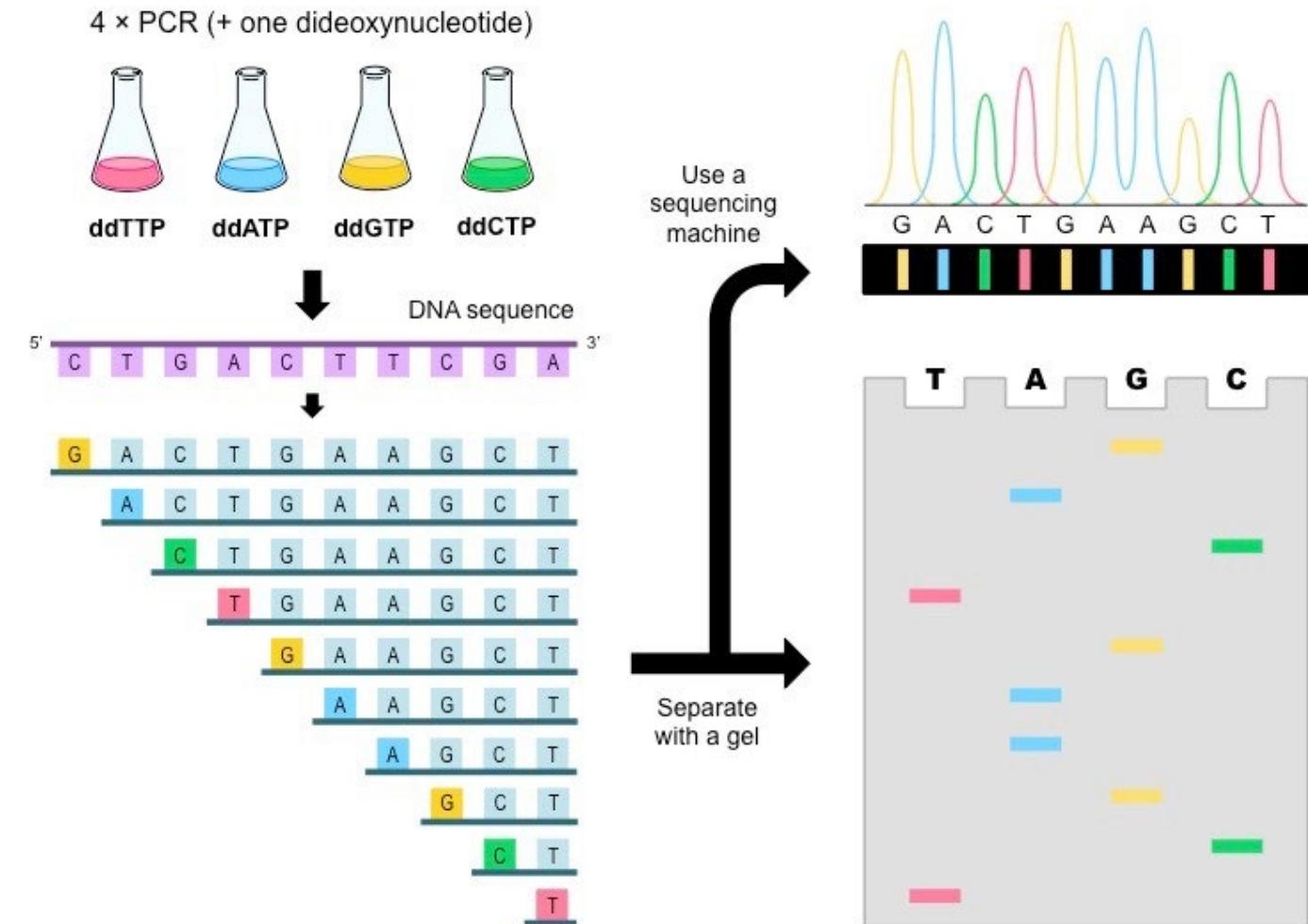
Genome-wide microsatellite 2-point linkage analysis



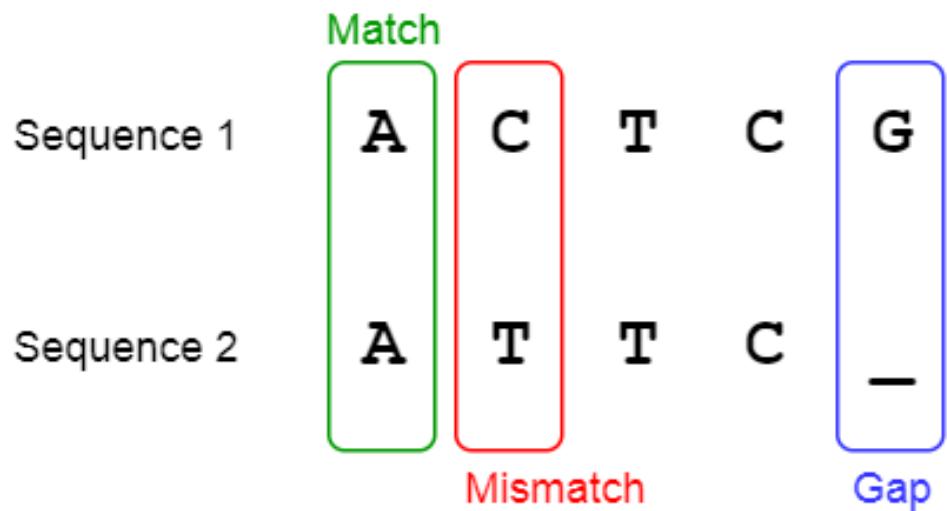
Sanger sequencing



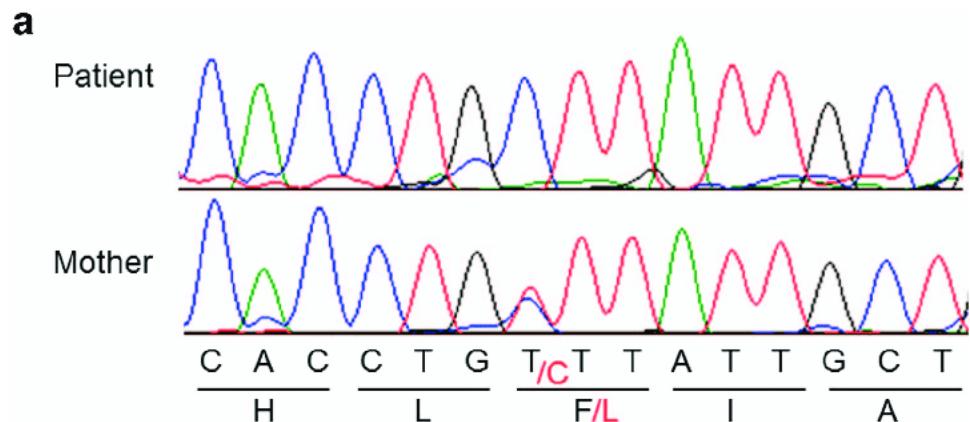
- Requires target region amplified by PCR – can be picky
- ~1kb limit per sequence – can build up overlapping reads to cover larger area



Sanger sequencing: compare to expected



- Heterozygous:



Sanger sequencing

- Large genes require a lot of work
- Titin gene (TTN)
 - 365,719bp total gene size (including introns)
 - Exons: 80,781bp across 365 exons
- DMD gene (Duchenne muscular dystrophy)
 - 2,220,390 bp including introns
 - ~14kb mRNA



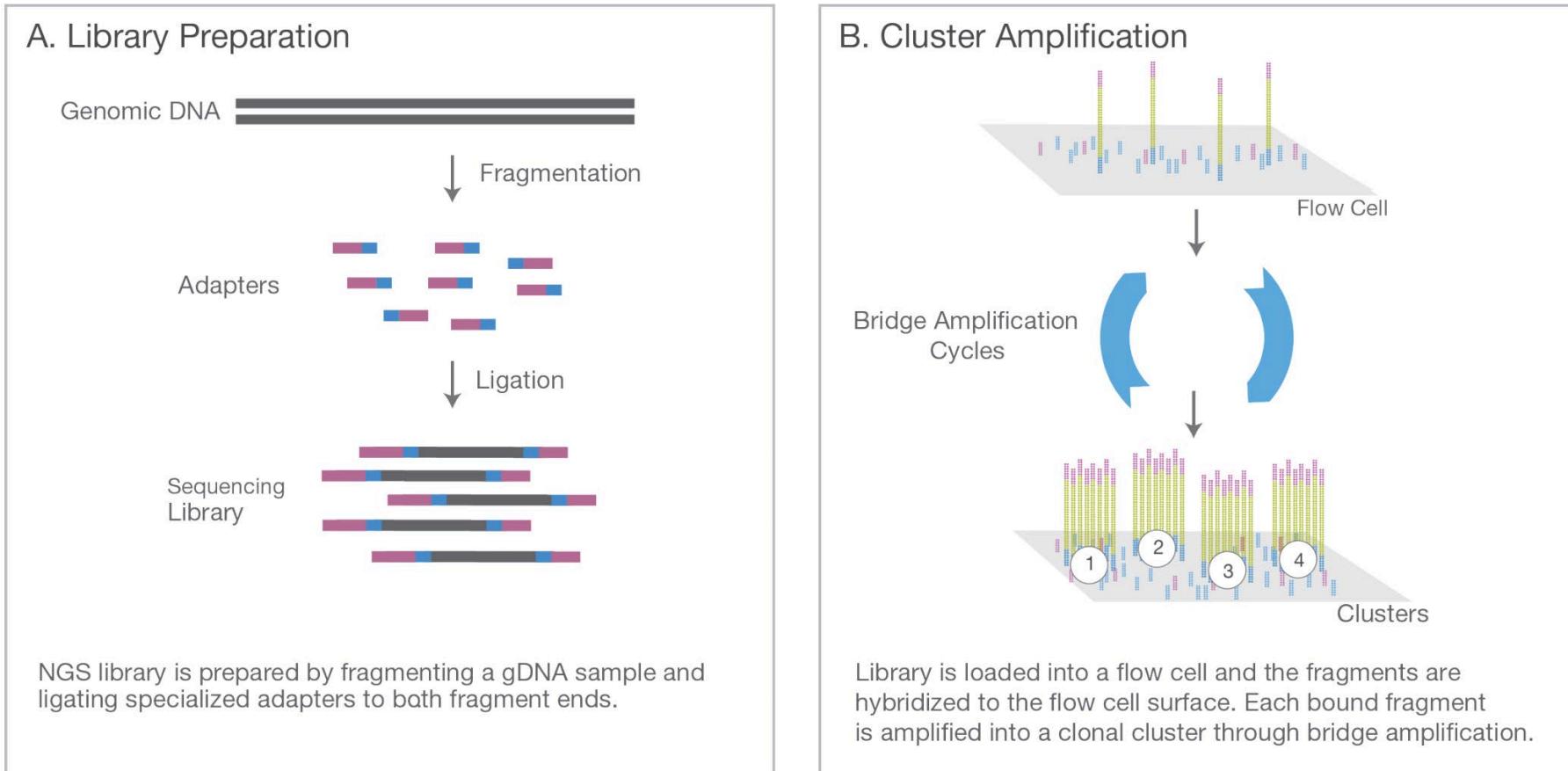
Nucleotide Mutations

Newer Technologies

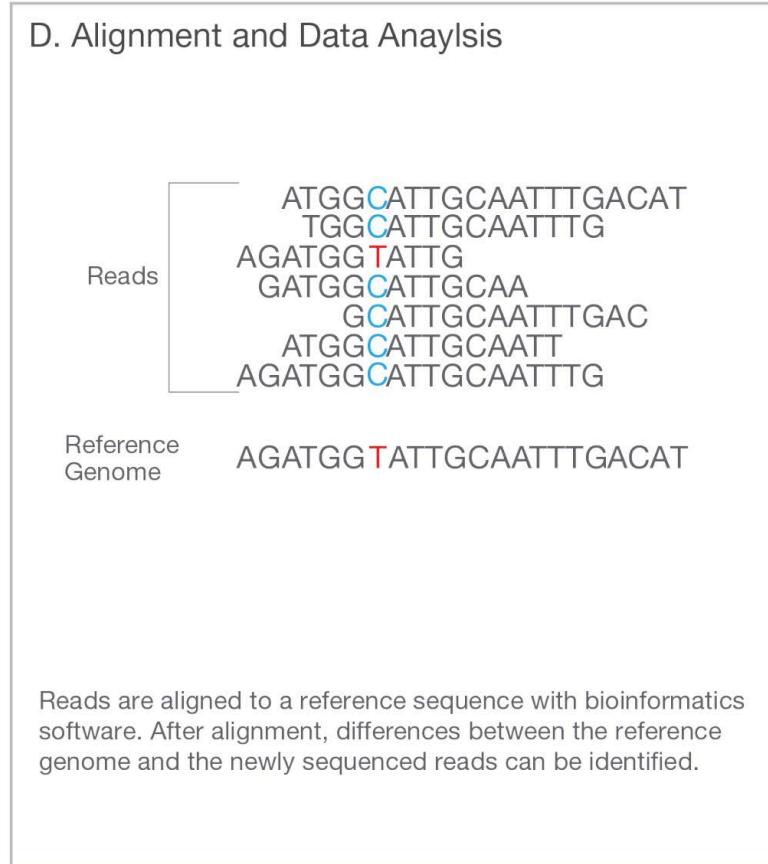
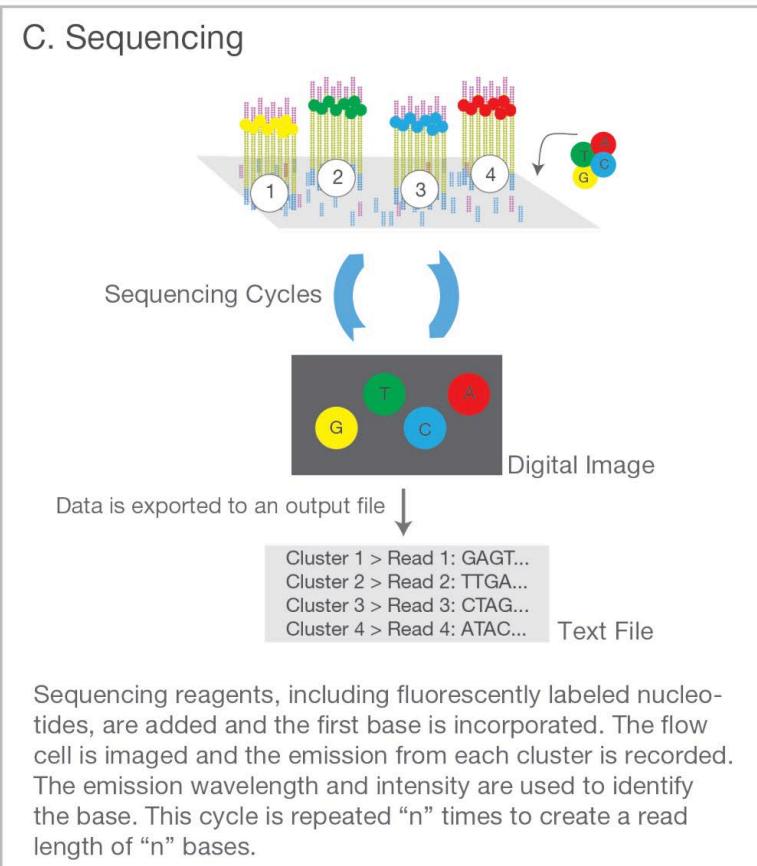
Next-generation sequencing (NGS)

- Sequence large regions of or entire genome
- Often cheaper than sequencing individual genes
 - Means we are more likely to test multiple genes at once (up to hundreds) 
 - Means we are more likely to find variants 

Next-generation sequencing: generation of reads



Next-generation sequencing: generation of reads

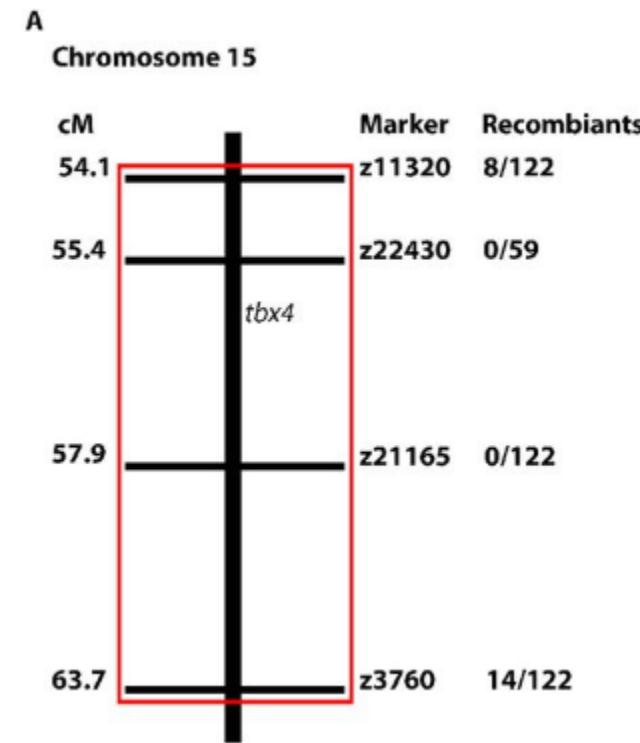


Next-generation sequencing: alignment of reads

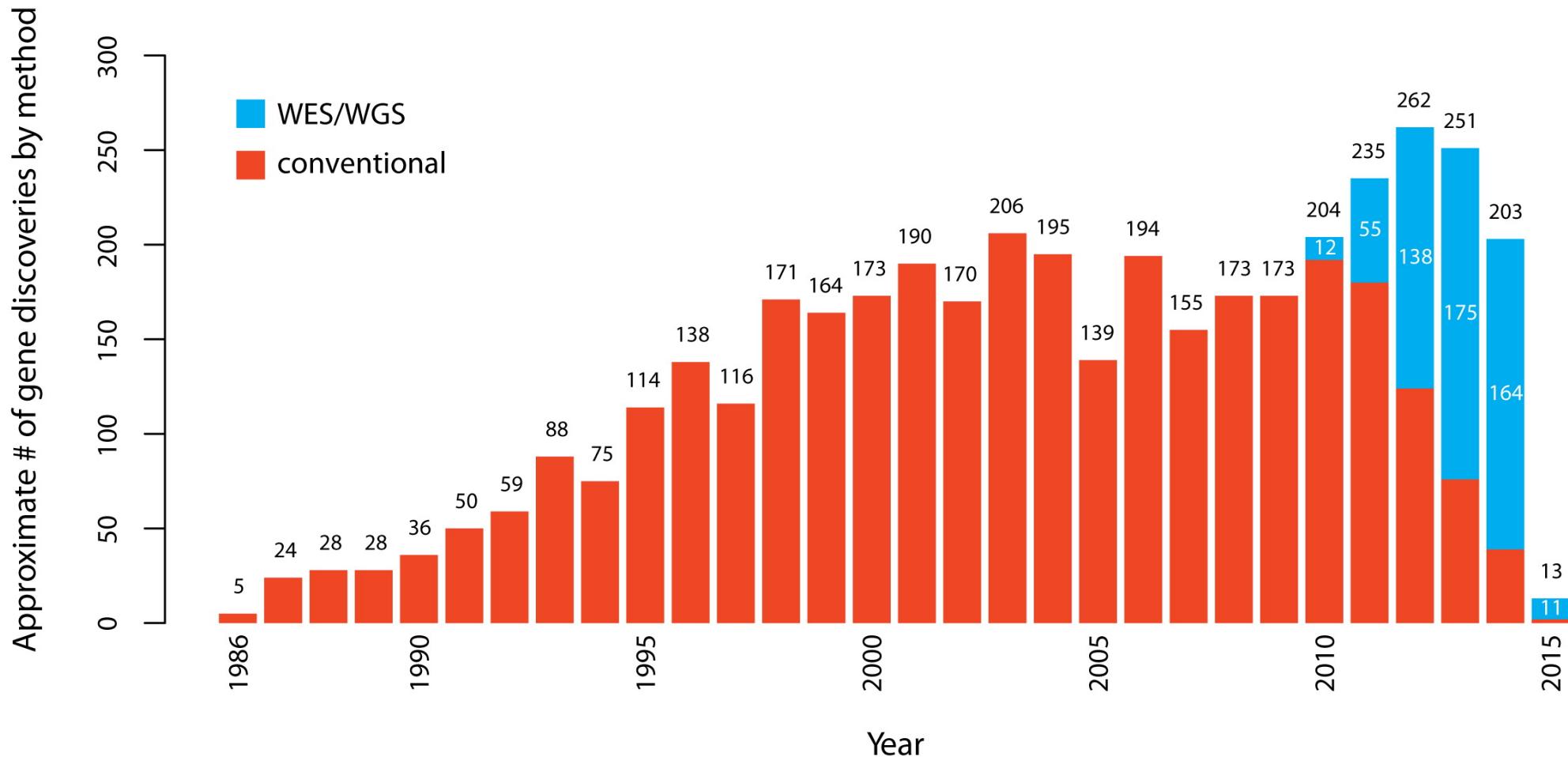


Next-generation sequencing

- Whole genome sequencing (WGS) = entire genome
- Whole exome sequencing (WES) = every known exon, usually with 10 base pairs either side
- Panel sequencing = selection of target genes (tens to hundreds)
- Targeted-enrichment next generation sequencing 



Since the introduction of WES and WGS in 2010, the pace of discovery of genes implicated in Mendelian phenotypes per year has increased substantially, and the proportion of discoveries made by WES or WGS (blue) versus conventional approaches (red) has steadily increased:

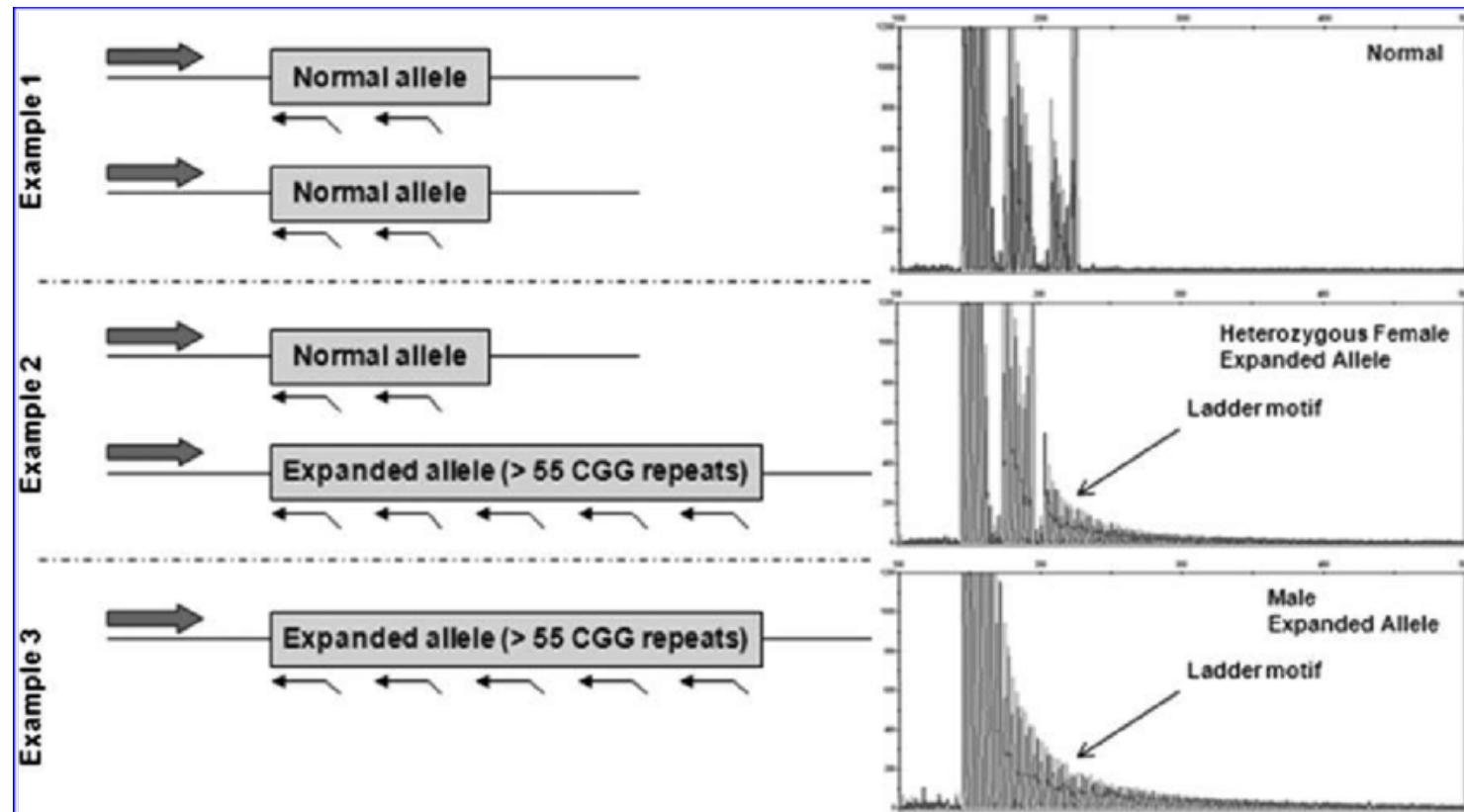


Next-generation sequencing

- Coverage/cost 
- Problems with deletions/duplications, and repeat sequences 

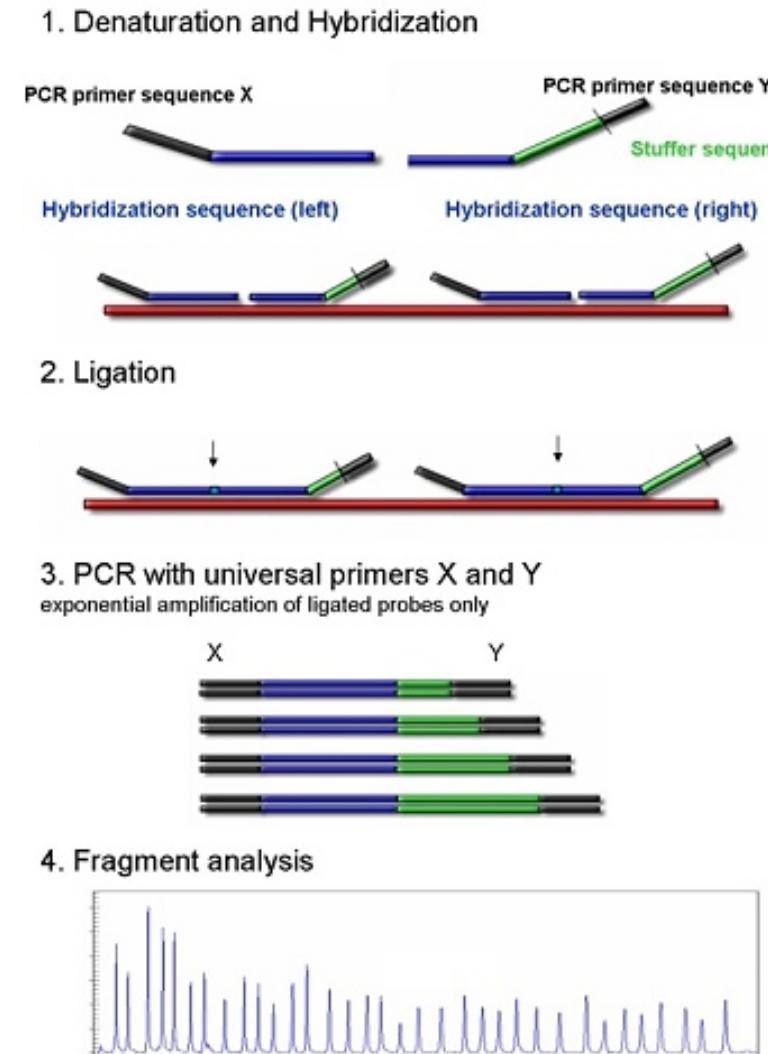
Triplet repeat primed PCR: how it works

- Detects expansion repeats
- Chimeric or triplet repeat primed PCR is defined as a PCR method that generates different sized amplicons due to multiple annealing sites on the template.



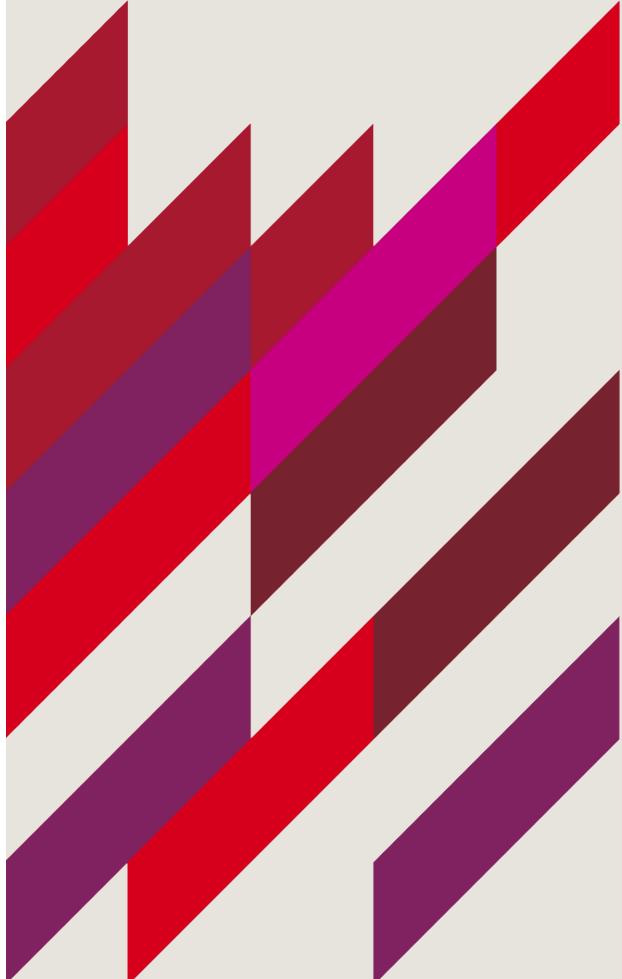
Multiplex ligation-dependent probe amplification (MLPA)

- Detects deletions/duplications in a target gene (often complements NGS) 
- Between next gen sequencing and microarray in deletion/duplication, sizewise
- Small probes throughout gene, amplification will only work if exact sequence is present
- Extra or missing regions show up as higher/lower amplification for that particular probe



BIOL3120 –Genetic Testing Techniques

LEARNING OBJECTIVES



On successful completion of this lecture, you will be able to:

- Identify and understand the technologies that detect chromosomal mutations
- Identify and understand the technologies that detect nucleotide mutations



MACQUARIE
University

BIOL3120 –Human Genetics and Evolutionary Medicine

Genome Wide Association Studies





6	Genetic Testing Techniques GWAS	Problem Set 4	Problem Set 4 (5%)	Explain the importance of new techniques in human genetics for understanding human disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources
Recess				
Recess		Pracs for External Students only		
7	Treatment for Genetic Conditions Epigenetics and Imprinting	Problem Set 5	Problem Set 4 (5%) & Problem Set 5 (5%)	Explain the importance of new techniques in human genetics for understanding human disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources

DNA sequence variation

- In humans, approximately 0.1–0.4% of nucleotides differ between any given pair of unrelated genomes
- The vast majority of sequence variation is comprised of single nucleotide polymorphisms (SNPs), which occur every 100–300 bases, and are mostly located within noncoding sequence
- A large number of inherited human diseases are caused by sequence variation in single genes
- Many complex diseases, including cancer, diabetes, and heart disease, are mediated, at least in part, by genetic factors
- The majority of rare diseases, such as those affecting only a small percentage of the population, result from hereditary or *de novo* genetic mutations
- Technological advances in high-throughput genotyping methods over the past two decades revolutionized the field of human genetics

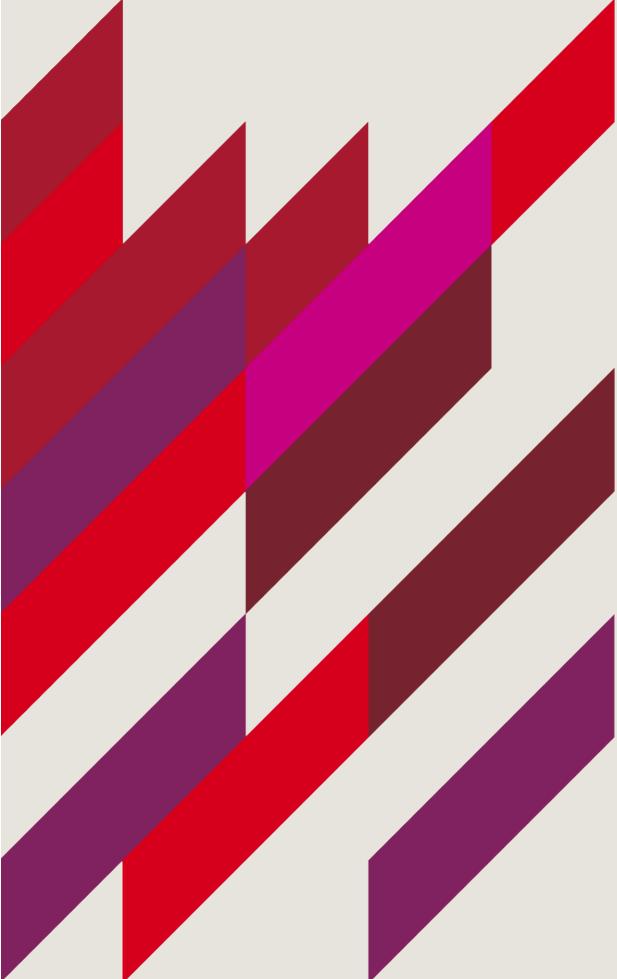
	SNP 1	SNP 2	SNP 3
Person 1:	acggtagctacaattttaaacgggaggaggatttattaaccagatgtg		
Person 2:	acggtatctacaattttaaacgggaggaggatttattaaccaatgtg		
Person 3:	acggtaactacaattttaaatgggaggaggatttattaaccagatgtg		
Person 4:	acggtaactacaattttaaatgggaggaggatttattaaccaatgtg		
Person 5:	acggtatctacaattttaaatgggaggaggatttattaaccaatgtg		
Person 6:	acggtatctacaattttaaatgggaggaggatttattaaccaatgtg		

Candidate Gene Approach

- Focuses on associations between genetic variation within pre-specified genes of interest, and disease
- Candidate genes are most often selected for study based on prior knowledge of the gene's biological functional impact on the trait or disease in question
- In contrast to genome-wide association studies (GWAS), which scan the entire genome for common genetic variation.

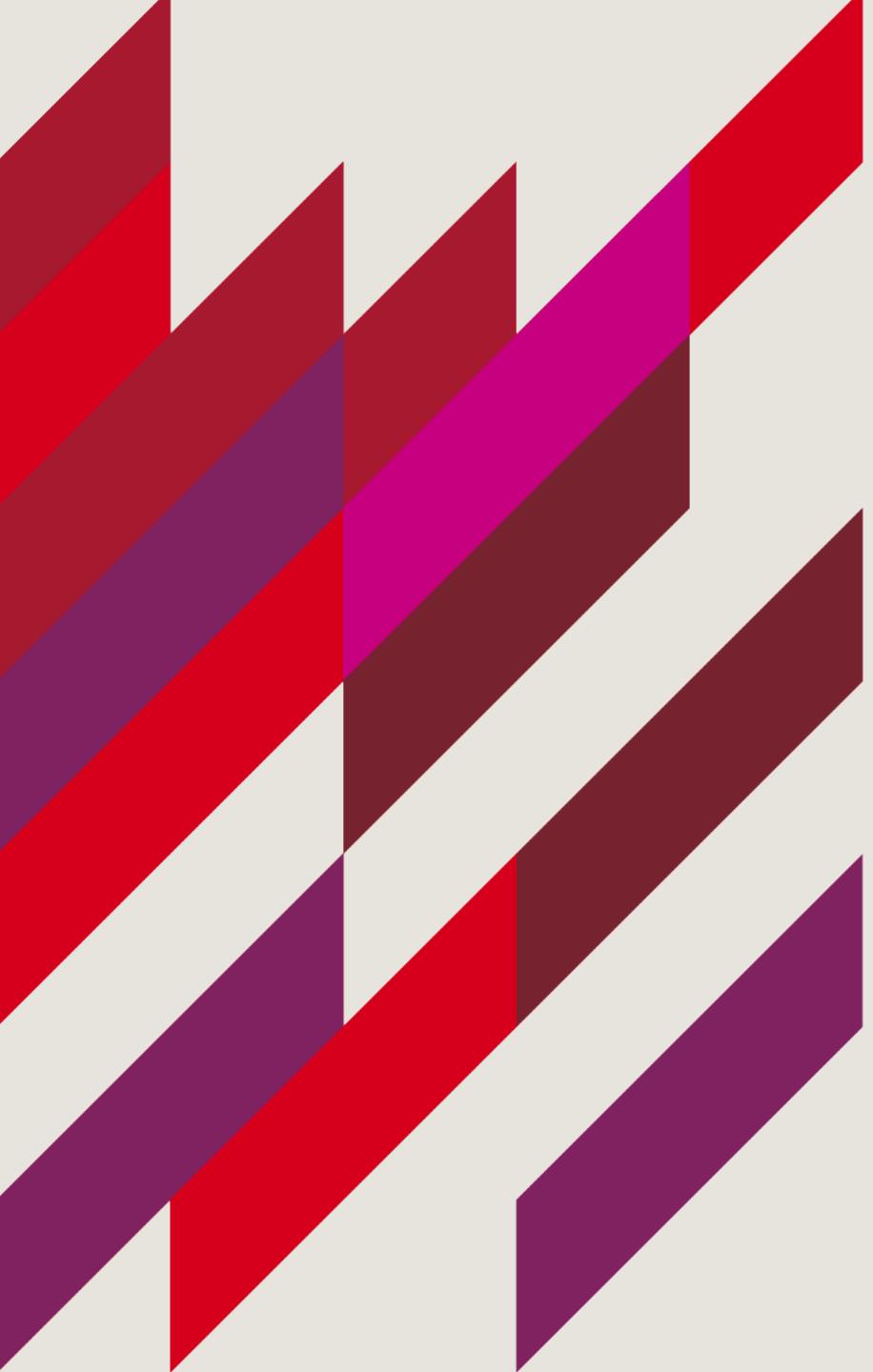
BIOL3120 –GWAS

LEARNING OBJECTIVES



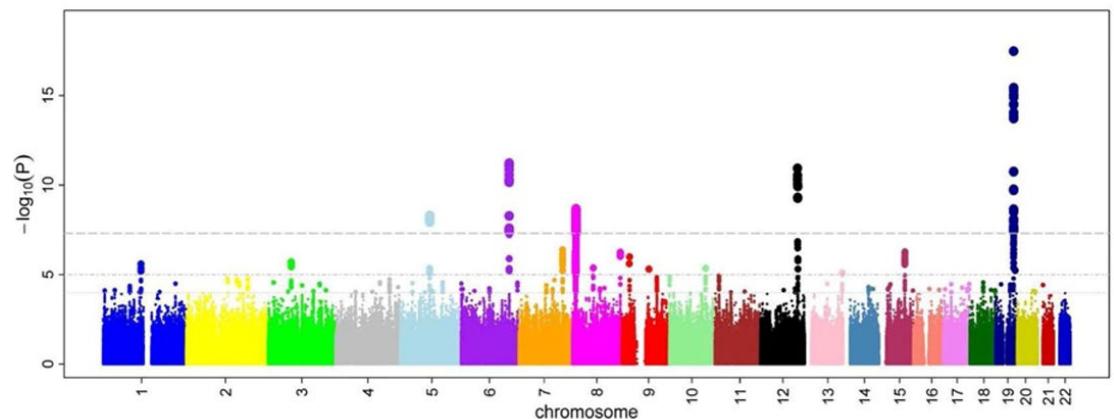
On successful completion of this lecture, you will be able to:

- Understand the difference between a candidate gene approach and GWAS
- Describe genome wide association studies
- Understand the limitations of GWAS



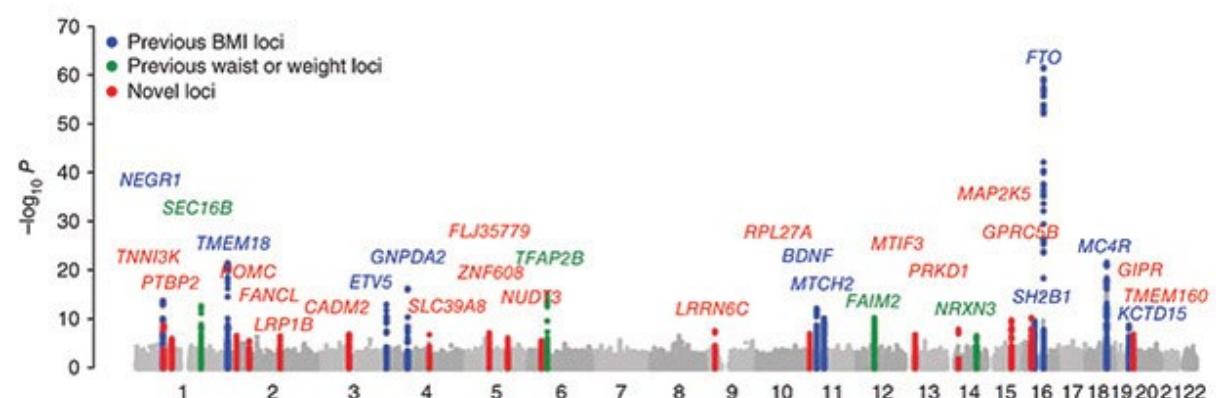
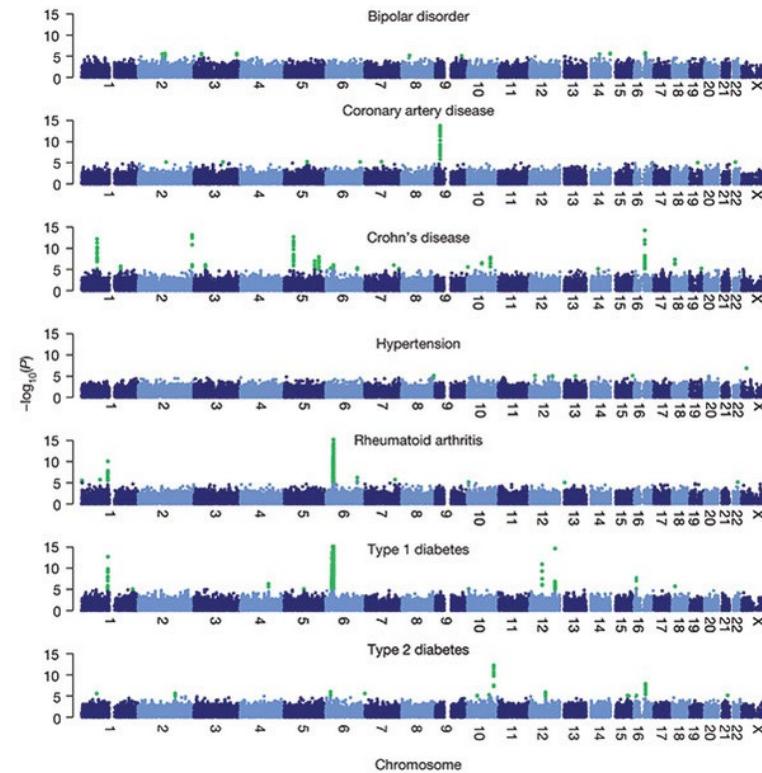
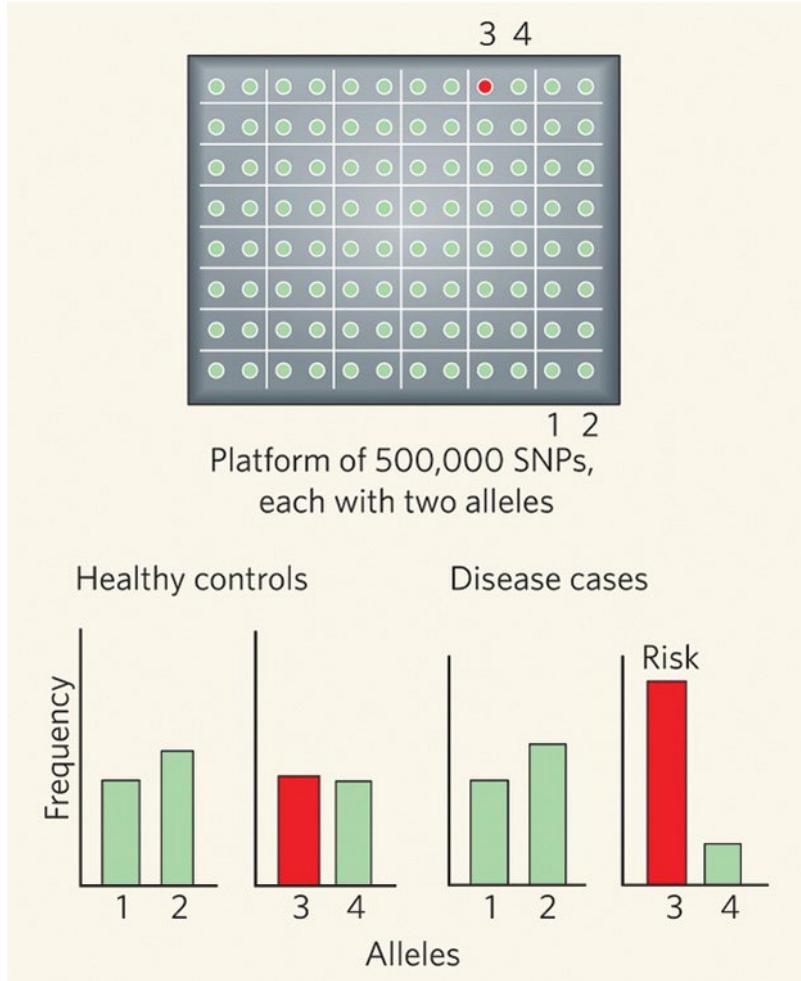
Genome Wide Association Studies

Genome-wide association studies (GWAS)



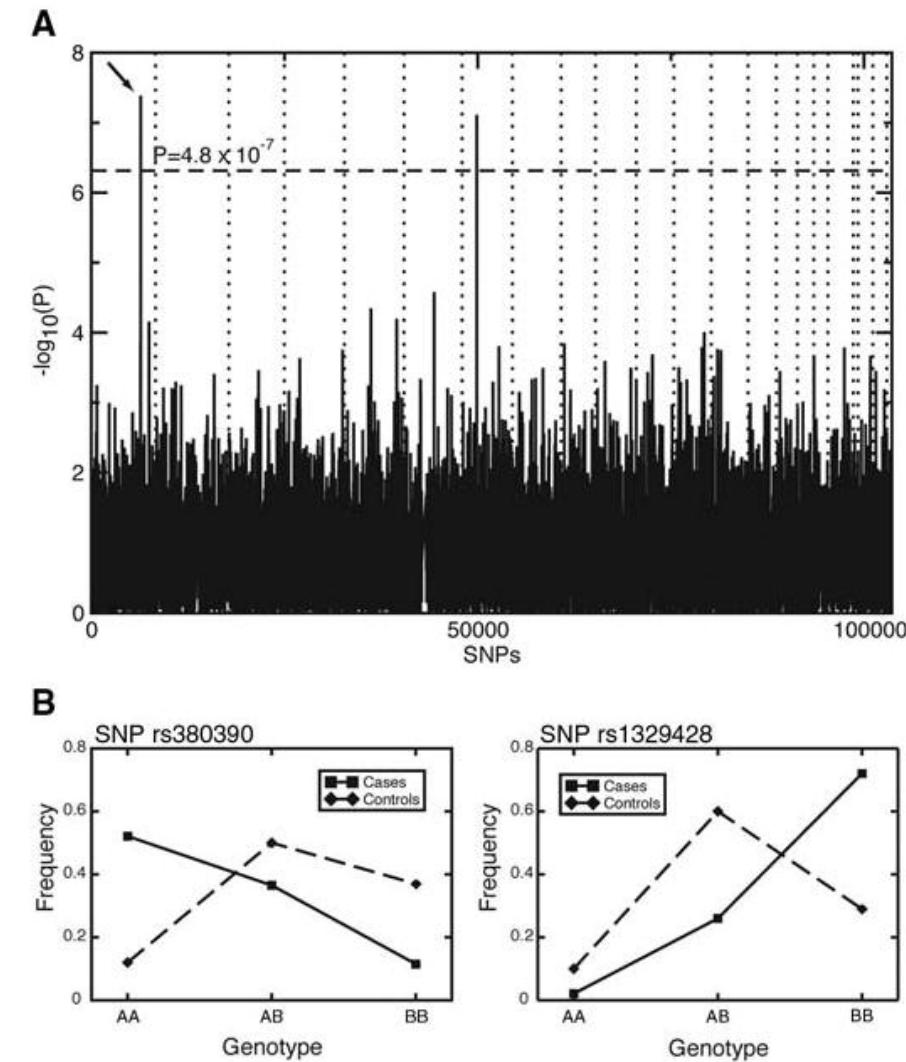
- A genome-wide set of genetic variants in different individuals to see if any variant is associated with a trait (disease)
- GWAS studies compare the DNA of participants having varying phenotypes for a particular trait or disease
- Each person gives a sample of DNA, from which millions of genetic variants are read using SNP arrays
- If one type of allele is more frequent in people with the disease, the variant is said to be *associated* with the disease
- The associated SNPs are then considered to mark a region of the human genome that may influence the risk of disease

Genome-wide association studies (GWAS)



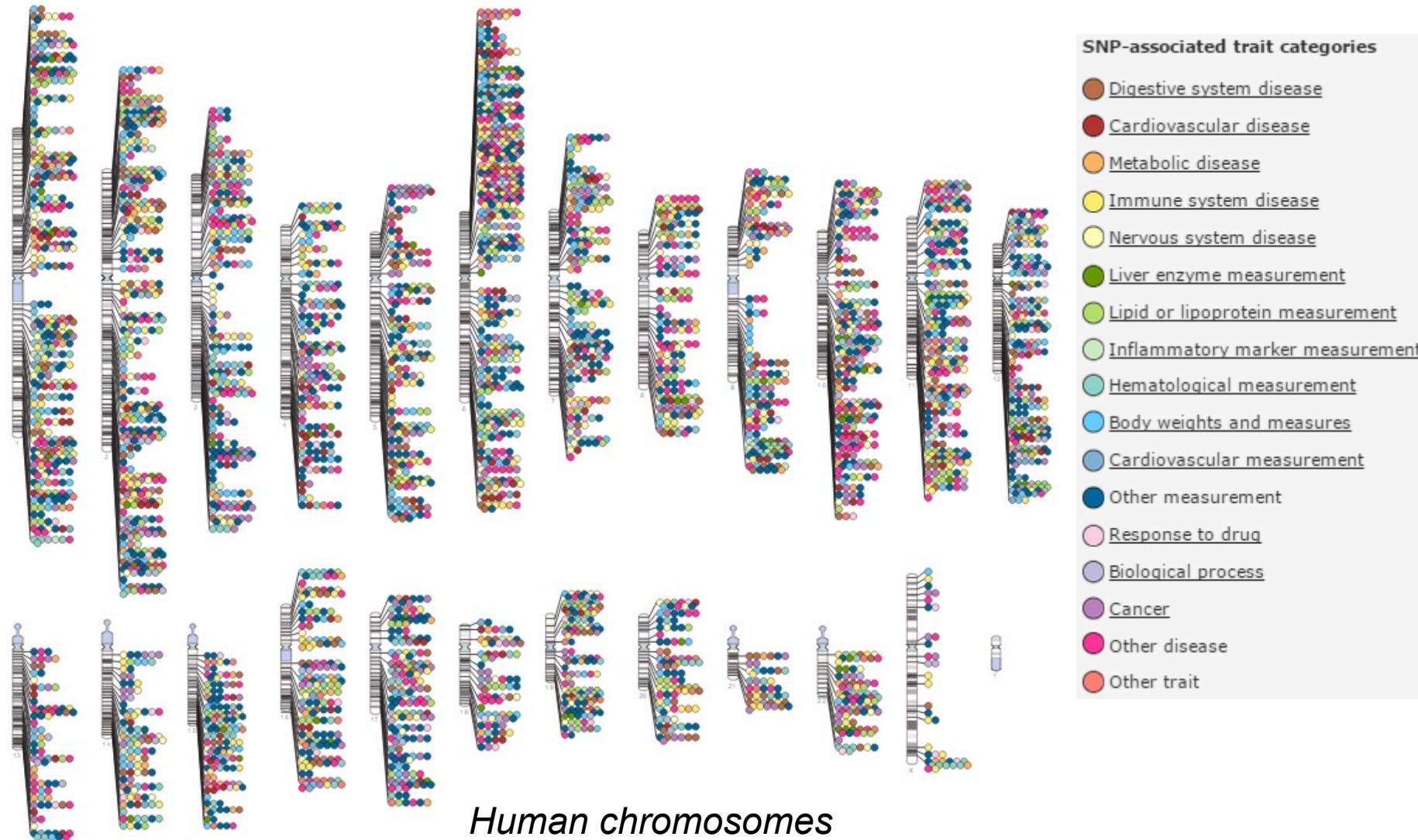
Genome-wide association studies (GWAS)

- The first successful GWAS was published in 2005
- It investigated patients with age-related macular degeneration and found two SNPs with significantly altered allele frequency compared to healthy controls
- Hundreds or thousands of individuals are tested in a typical GWA study, over 3,000 human GWA studies have examined over 1,800 diseases and traits, and thousands of SNP associations have been found



Klein et al., 2005. Complement Factor H Polymorphism in Age-Related Macular Degeneration. Science 308(5720).

Examples of GWAS Discoveries





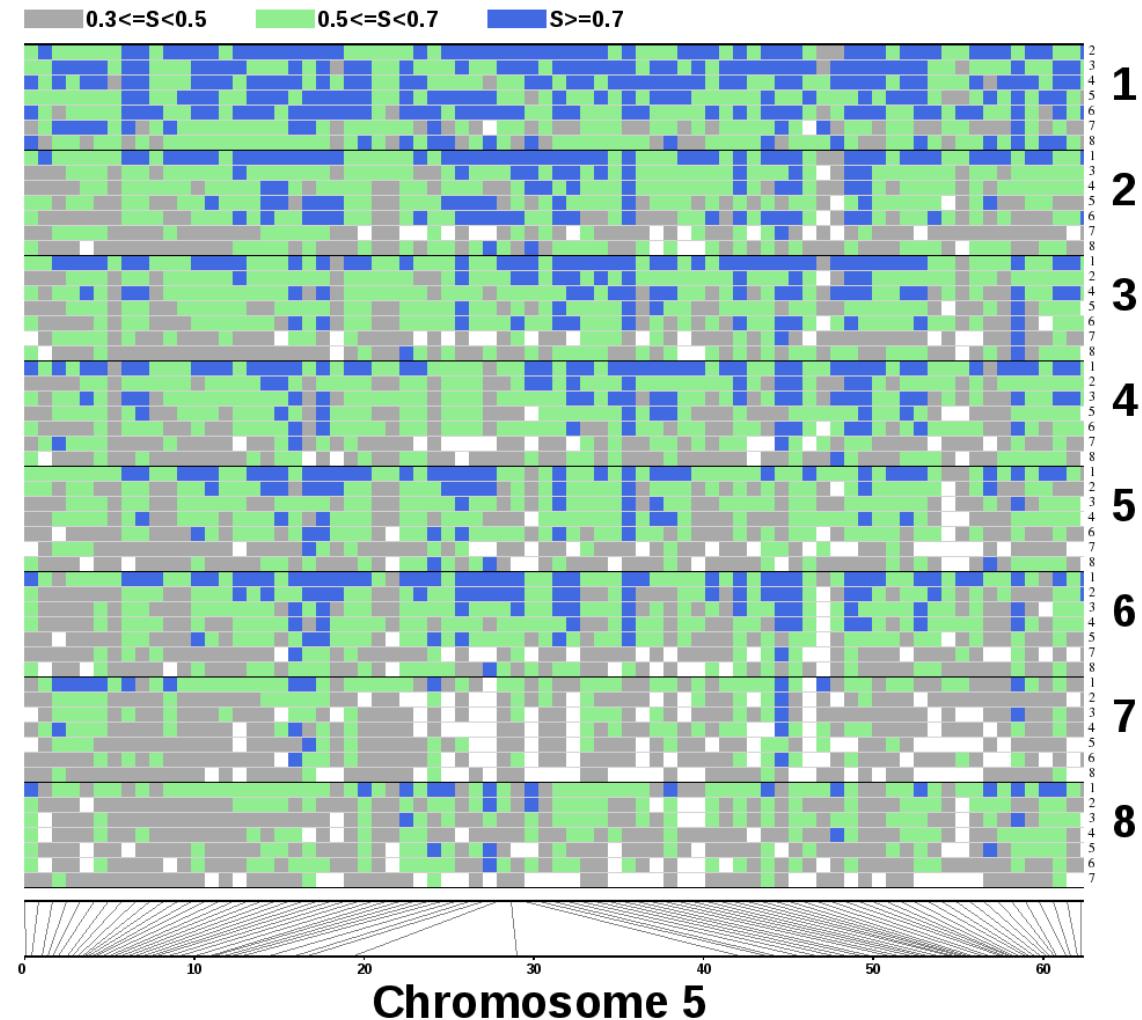
Success and Limitations of GWAS

Here at Macquarie

- Project MinE
- We plan to map the full DNA profiles of at least 15,000 people with ALS and 7,500 control subjects, and to perform comparative analyses on the resulting data.
- <https://www.projectmine.com/about/>

High-throughput genotyping

- Genome-wide association approaches have identified statistically significant evidence supporting relationships between complex human diseases and hundreds of common genetic variants in the human population
 - However, finding disease-associated alleles is only the first step on the path to identifying those variants that directly contribute to disease risk
 - A major challenge inherent in these studies is moving from identification of a genetic variant via association studies to determination of actual causal variants through functional genomics experimentation

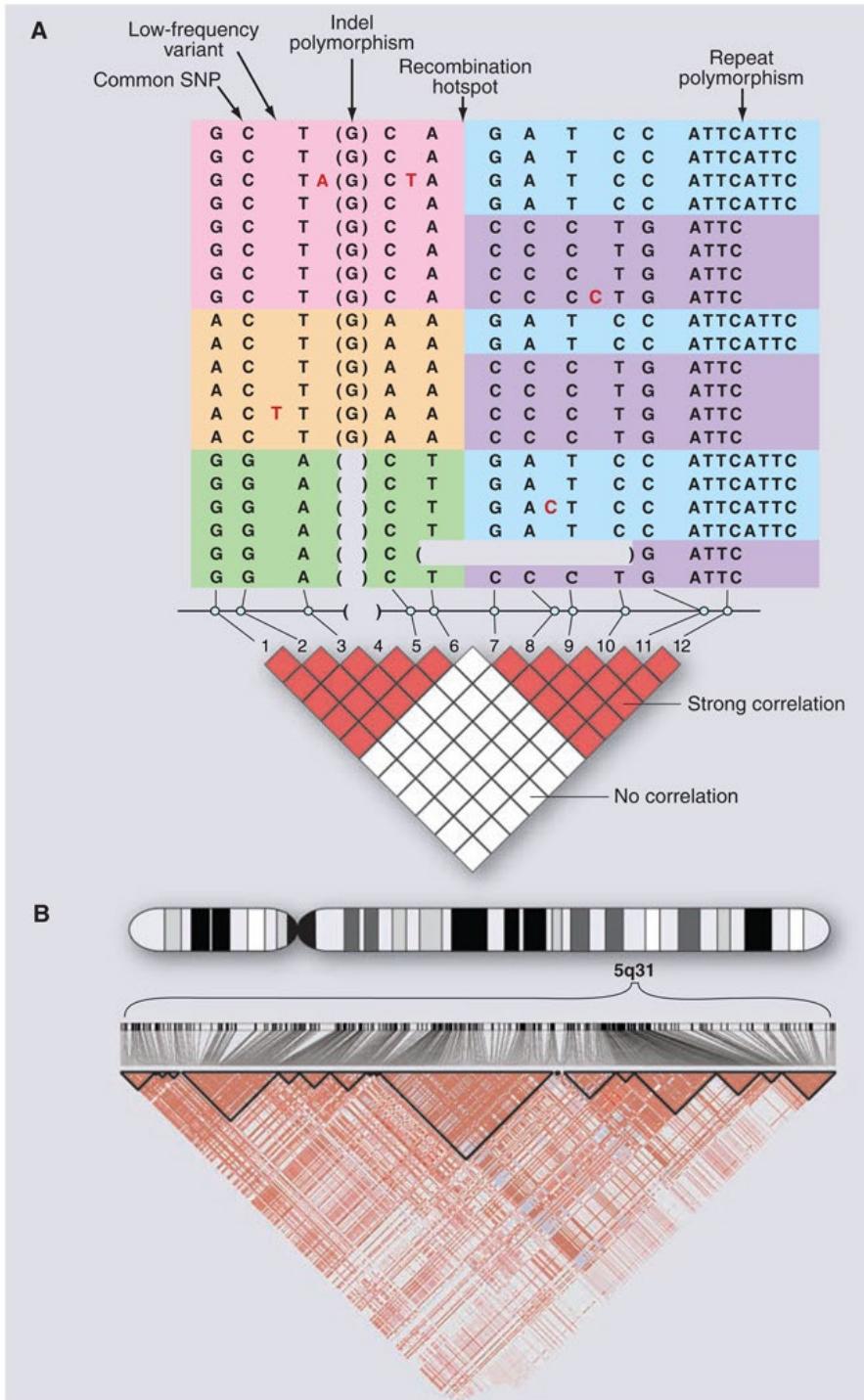


Limitations of GWAS

As the SNP catalog grows, a critical question looms:

Would GWASs require directly testing each of the ~10 million common variants for association to disease? That is, if only 5% of variants were tested, would 95% of associations be missed?

Or could a subset serve as reliable proxies for their neighbors?

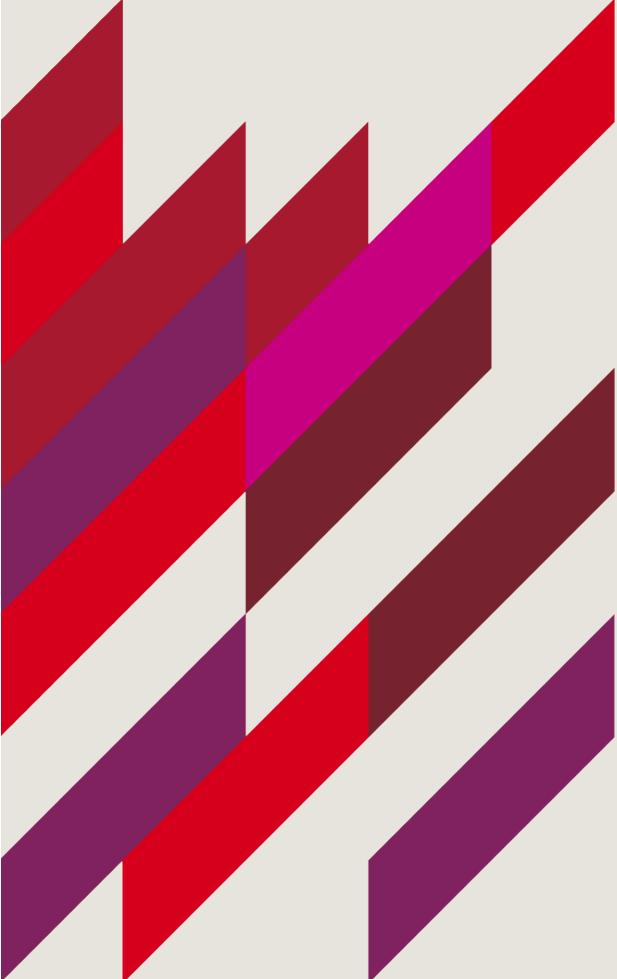


Limitations of GWAS

- Despite the success of GWAS in enhancing understanding of disease mechanisms, the variants identified by this approach represent only a fraction of the overall genetic contribution to common disease risk
- While many disease-associated variants have been identified through GWAS, they have mostly been common variants with moderate to high (i.e., >0.1) allele frequencies
- Assumption that common genetic variation plays a large role in explaining the heritable variation of common disease
- The question of whether common or rare variants underlie the majority of risk for common diseases continues to remain an open one

BIOL3120 –GWAS

LEARNING OBJECTIVES



On successful completion of this lecture, you will be able to:

- Understand the difference between a candidate gene approach and GWAS
- Describe genome wide association studies
- Understand the limitations of GWAS