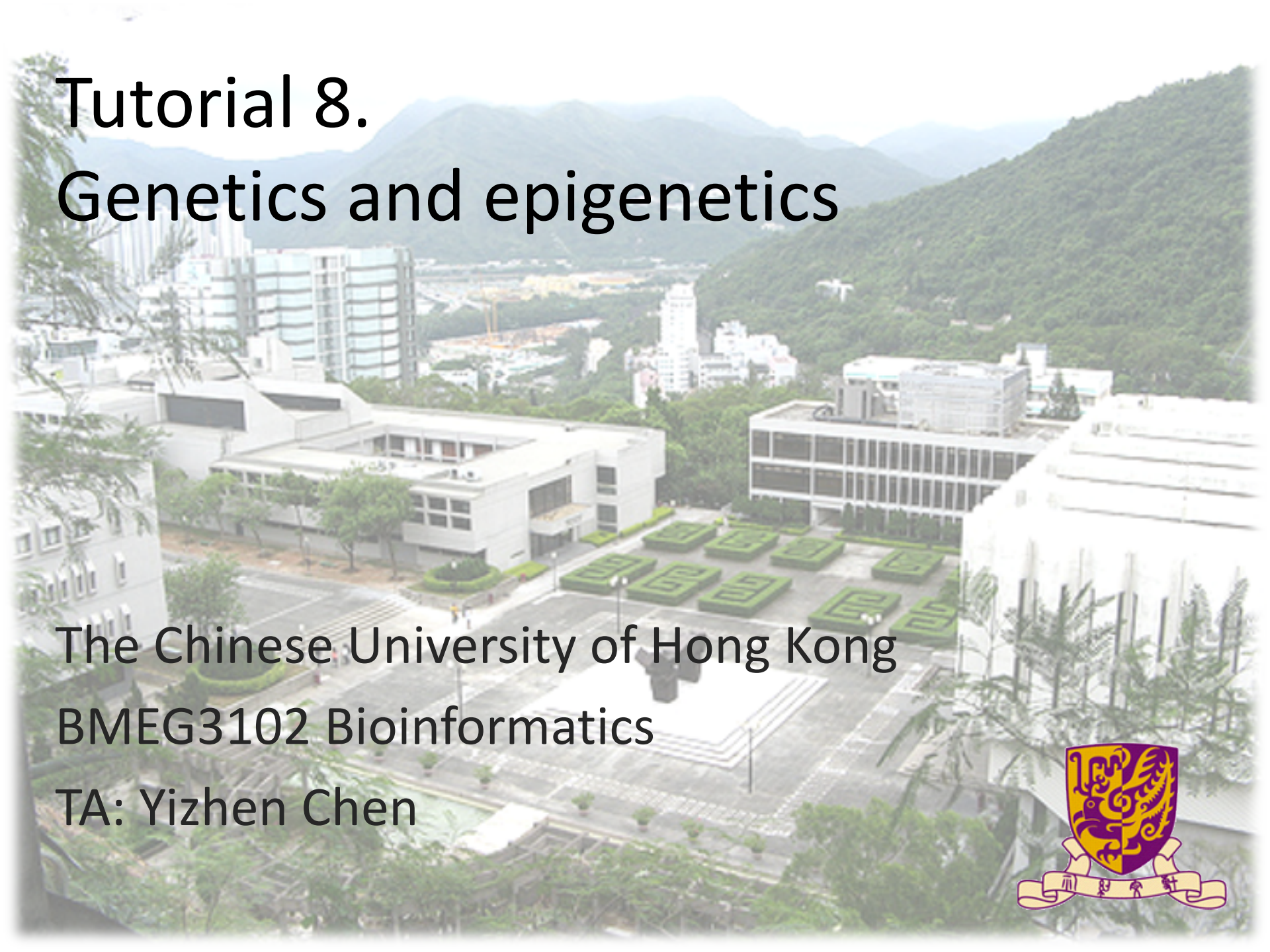
Tutorial 8.
Genetics and epigenetics



The Chinese University of Hong Kong

BMEG3102 Bioinformatics

TA: Yizhen Chen

Agenda

- Exercise from last time
- Genetics and epigenetics
  - Mendel's law
  - Case study: ABO blood type
  - Exercise on pedigree
  - Epigenetics
- High-throughput sequencing and applications
  - Case study: Prenatal diagnosis
  - Case study: Coronavirus discovery and detection

Exercise

Suppose in the genomic background, the percentages of A, C, G and T are 30%, 20%, 20% and 30%, respectively. What is the probability for each the following patterns to **appear on either strand** of a random DNA sequence of **5bp long**, assuming each position is **independent** of the others? Explain how these values are computed.

- AACTT
- ACNG, where N can be any of the four nucleotides
- ATA

i. AACTT

The probability for this pattern to appear on one strand is $(0.3)^4(0.2)^1 = 0.00162$.

Since whenever this pattern appears on one strand, it must not appear on the other strand at the same time, the total probability for it to appear on either strand is $2 \times 0.00162 = 0.00324$.

```
Strand 1: AACTT
Strand 2: TTGAA
```

ii. ACNG

On one strand, this pattern can appear at the first position or the second position, but not both at the same time. For each case, the probability for the pattern to appear is $(0.3)^1(0.2)^2 = 0.012$. Therefore, the probability for the pattern to appear on one strand is $2 \times 0.012 = 0.024$.

When the pattern appears at the first position of one strand, it can also appear on the first position of the other strand if the sequence is ACNGT. On the other hand, when the pattern appears at the second position of one strand, it cannot appear on the other strand at the same time.

Therefore, the total probability for it to appear on either strand is $2 \times 0.024 - (0.3)^2(0.2)^2 = 0.0444$.

```
            12345          12345
Strand 1:  ACNGT          CACNG
Strand 2:  TGNCA          GTGAC
```

iii. ATA

On one strand, this pattern can appear at the first, second or third position. If it appears at the second position, it cannot also appear at the first or third position at the same time. However, if it appears at the first position, it can also appear at the third position if the sequence is ATATA. Therefore, the probability for the pattern to appear on one strand is $3\times(0.3)^3 - (0.3)^5 = 0.07857$.

When the pattern appears at the first position on one strand, it can also appear on the other strand at the second position if the sequence is ATATN. Similarly, when the pattern appears at the third position on one strand, it can also appear on the other strand at the second position if the sequence is NTATA. The sequence ATATA is commonly shared by them.

Finally, when the pattern appears at the second position on one strand, it can also appear on the other strand at the first or third position if the sequence is NATAT or TATAN, respectively. The sequence TATAT is commonly shared by them.

Therefore, the total probability for the pattern to appear on either strand is $2\times0.07857 - (0.3)^4 - (0.3)^4 + (0.3)^5 - (0.3)^4 - (0.3)^4 + (0.3)^5 = 0.1296$.

Basic terms related to genetics

Term	Description
Traits	Some features of living things that we can see
Alleles	Each unique form of a single gene
Homozygous	Two same alleles
Heterozygous	Two different alleles
Dominant traits	Traits that are shown in heterozygous
Recessive traits	Traits that are not shown in heterozygous
Genotype	Type of alleles that are present
Phenotype	Type of traits that are present
Karyotype	The arrangement of the full chromosome set

Mendel's Laws

- (Law of Segregation) Every individual organism contains two alleles for each trait. During meiosis, the two alleles from a parent are separated so that each gamete contains only one of these alleles.
- (Law of Independent Assortment) Alleles for separate traits are passed independently of one another i.e., from parents to offspring.
- (Law of Dominance) Recessive alleles will be masked by dominant alleles.

Warm-up Exercise

- Assume that the loci of the phenotypes below are Mendelian and on the same autosome.

Dominant trait	Recessive trait
(A) Widow's peak	(a) Straight hair line
(B) Clockwise hair direction	(b) Counter-Clockwise hair direction
(C) Ability to roll tongue	(c) No ability to roll tongue

Note: the letters in brackets are genotypes

- If Amy has clockwise hair direction, what is (are) the possible genotype(s) for loci B? BB, Bb
- If Bob has straight hair line, counter-clockwise hair direction, and ability to roll tongue, what is (are) the possible haplotype(s) for the given loci?

abC, abc

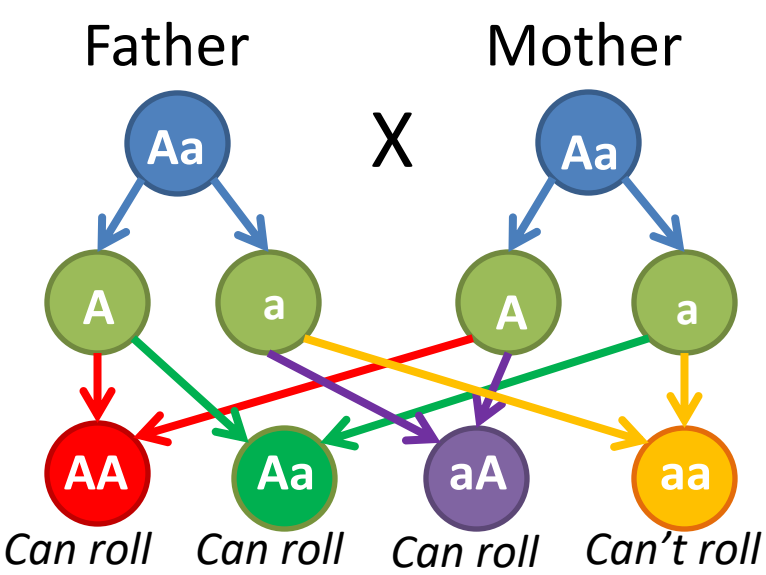
Genetics and Probability

- If both father and mother can roll tongue with heterozygous alleles, what is the probability that:
  - their child can roll tongue? $\frac{3}{4}$
  - their child has heterozygous alleles, if he/she can roll tongue? $\frac{2}{3}$
- Assuming equal chance of different alleles

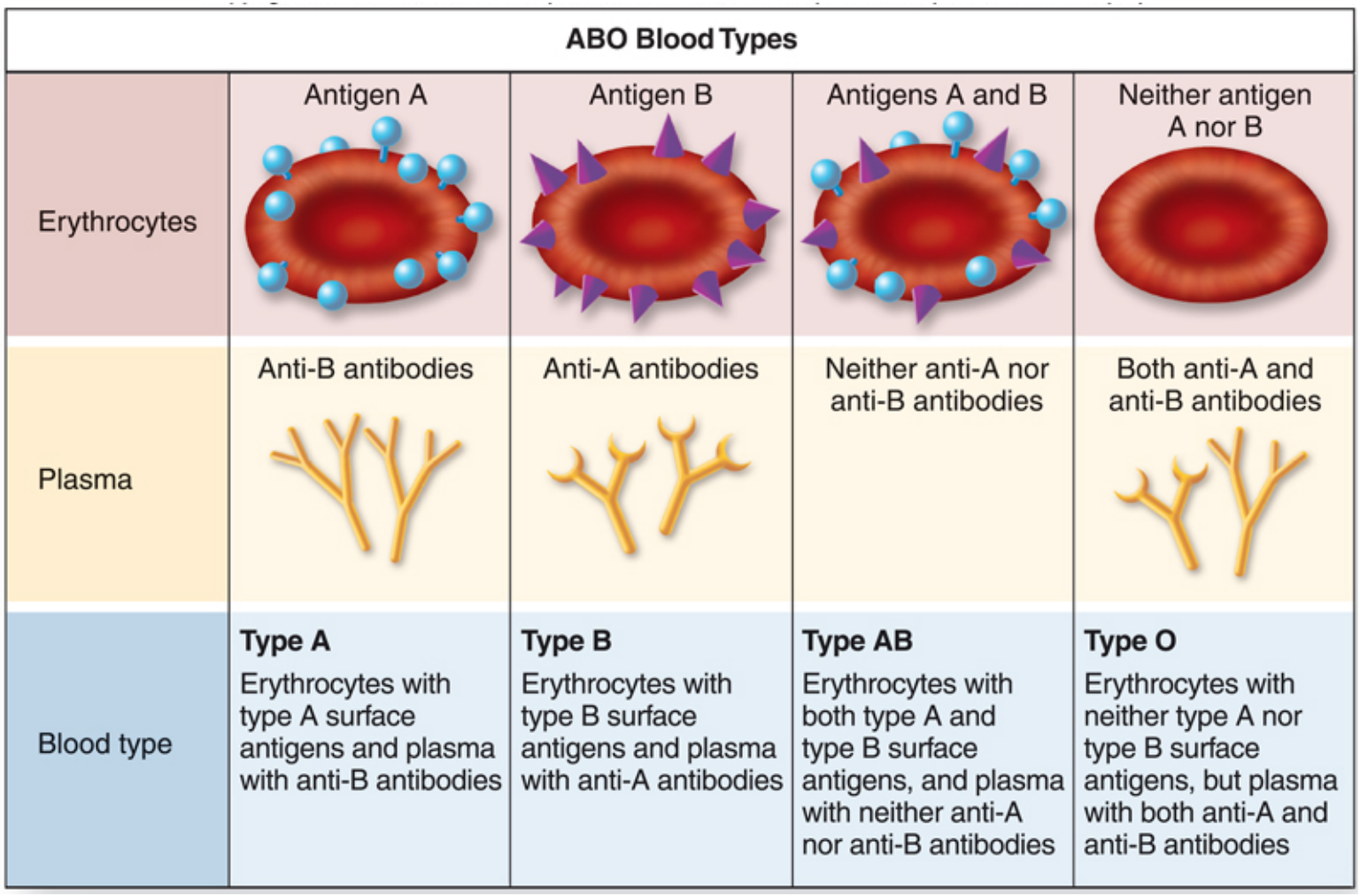
Punnet Square

Mother \ Father	X	A	a
	A	AA *Can roll*	Aa *Can roll*
	a	aA *Can roll*	aa *Can't roll*

Genetic Diagram



Case Study: ABO-blood type



ABO Blood Types				
Erythrocytes	Antigen A	Antigen B	Antigens A and B	Neither antigen A nor B
Plasma	Anti-B antibodies	Anti-A antibodies	Neither anti-A nor anti-B antibodies	Both anti-A and anti-B antibodies
Blood type	**Type A** Erythrocytes with type A surface antigens and plasma with anti-B antibodies	**Type B** Erythrocytes with type B surface antigens and plasma with anti-A antibodies	**Type AB** Erythrocytes with both type A and type B surface antigens, and plasma with neither anti-A nor anti-B antibodies	**Type O** Erythrocytes with neither type A nor type B surface antigens, but plasma with both anti-A and anti-B antibodies

Image source: http://www.biologycorner.com/anatomy/blood/images/bloodtypes.jpg

Case Study: ABO-blood type

Phenotype	Genotype
A	$I^A I^A$ or $I^A i$
B	$I^B I^B$ or $I^B i$
AB	$I^A I^B$
O	ii

- Exercise: if mother's blood type is AB, father's blood type is A, what is (are) the possible blood type(s) of their child?
- Answer: A, B or AB

Case Study: ABO-blood type

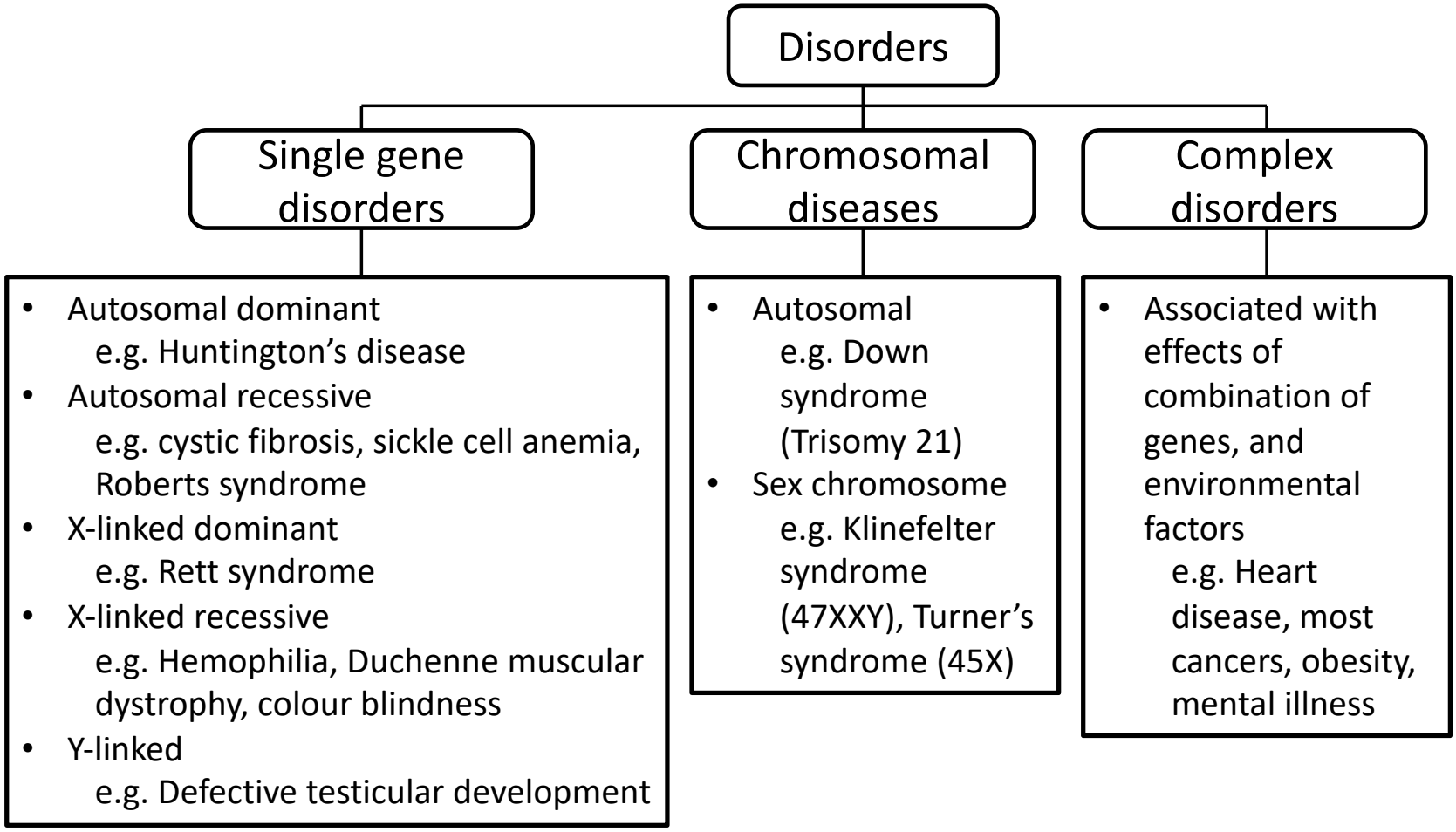
- ABO-blood type inheritance

Mother's Blood Type	Father's Blood Type: A	B	AB	O	Child's Blood type Must Be
A	A or O	A, B, AB, or O	A, B, or AB	A or O	
B	A, B, AB or O	B or O	A, B, or AB	B or O	
AB	A, B, or AB	A, B, or AB	A, B, or AB	A or B	
O	A or O	B or O	A or B	O	

Mother's Blood Type	Child's Blood Type: A	B	AB	O	Father's Blood Type Must Be
A	A, B, AB or O	B or AB	B or AB	A, B, or O	
B	A or AB	A, B, AB or O	A or AB	A, B, or O	
AB	A, B, AB or O	A, B, AB or O	A, B, or AB		
O	A or AB	B or AB		A, B, or O	

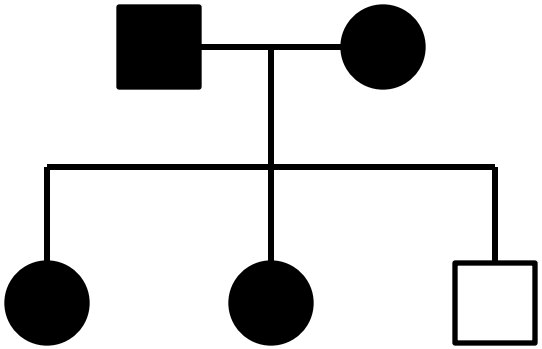
Image source: http://www.canadiancrc.com/paternity_determination_blood_type.aspx

Genetic disorder



Pedigree

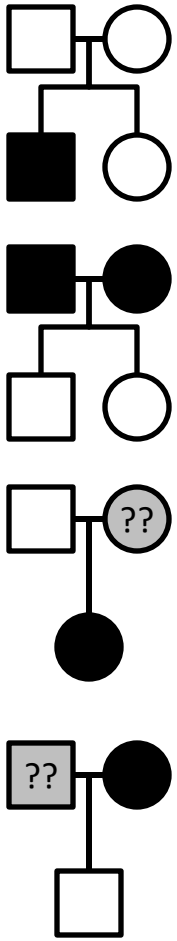
- In most cases, circle represents female, square represents male, filled shape represents affected with certain disorder.
- Example:



- For simplicity, in this tutorial, we suppose that
  - The disorder is caused by a single gene
  - The disorder follows Mendelian inheritance patterns.

Pedigree: useful rules

1. If both parents are unaffected but the child is affected, the disease trait must be recessive
2. If both parents are affected but the child is not, the disease trait must be dominant
3. For a recessive disease trait, if the father is unaffected but a daughter is affected, the disease gene must be autosomal
4. For a recessive disease trait, if the mother is affected but a son is unaffected, the disease gene must be autosomal

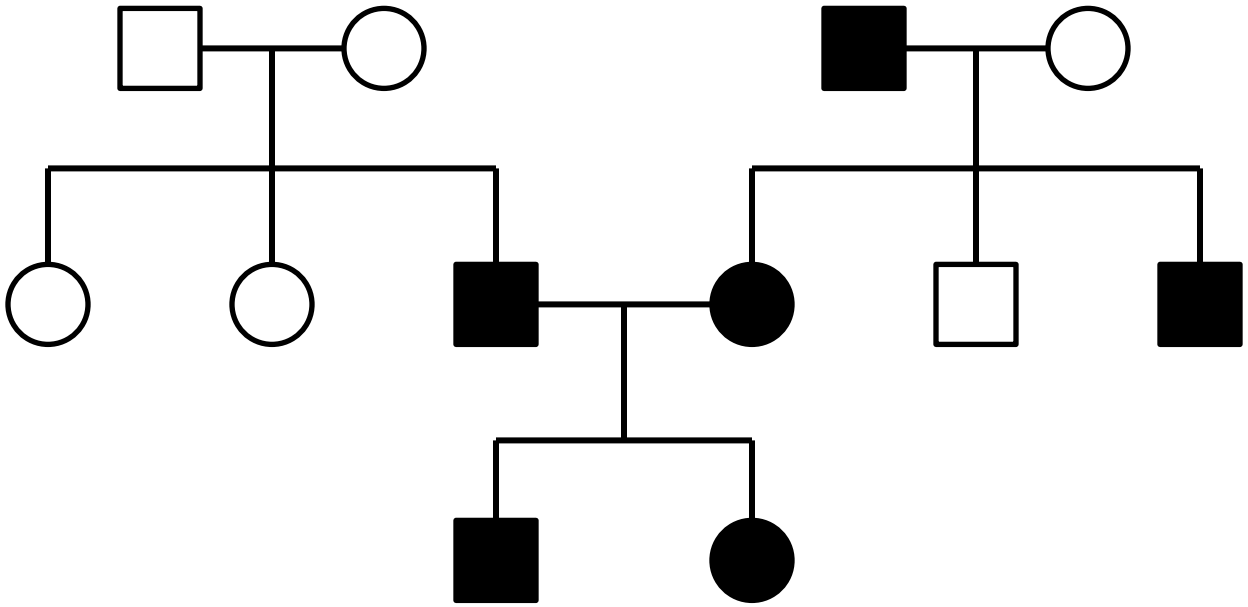


Pedigree Chart: Some Case Studies

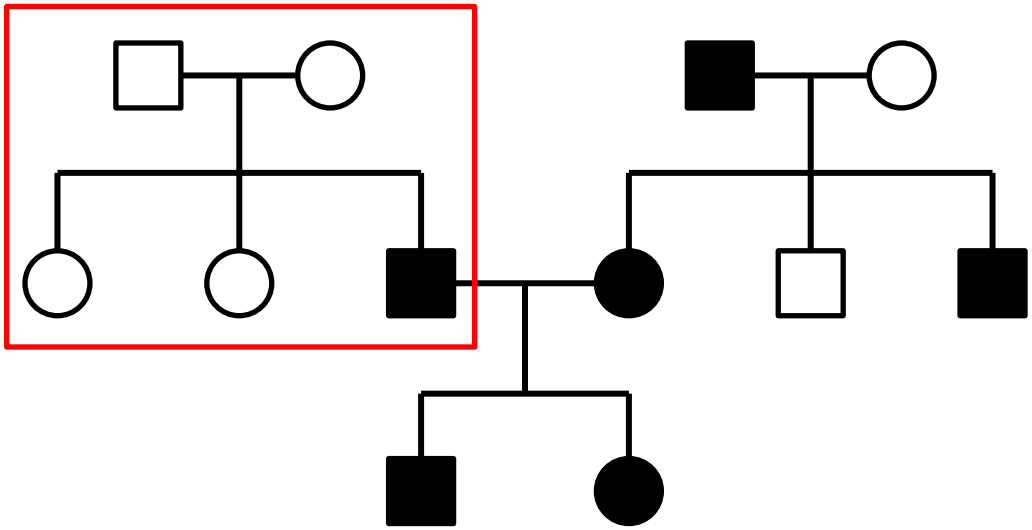
1. If affected ones appear in every generation, then it is likely to be dominant
2. If the fraction of affected males and the fraction of affected females are almost the same, then it is likely to be autosomal
3. If the fraction of affected males is much larger than that of affected females, then it is likely to be X-linked recessive
4. If the fraction of affected males is much smaller than that of affected females, then it is likely to be X-linked dominant
5. If all males are affected and all females are unaffected, then it is likely to be Y-linked

Exercise 1: Pedigree

- Given the pedigree chart, among dominant and recessive, also among autosomal, X-linked and Y-linked, discuss:
  - What are the possible cases of the disorder?
  - Which one is most likely?



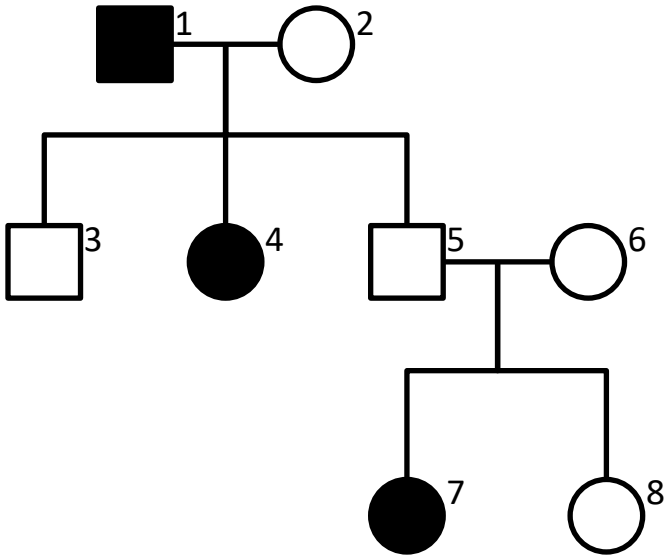
Answer 1: Pedigree



- It must be recessive (a boy gets affected from an unaffected couple).
- It can be autosomal, X-linked (females get affected).
- X-linked recessive is more likely than autosomal recessive.

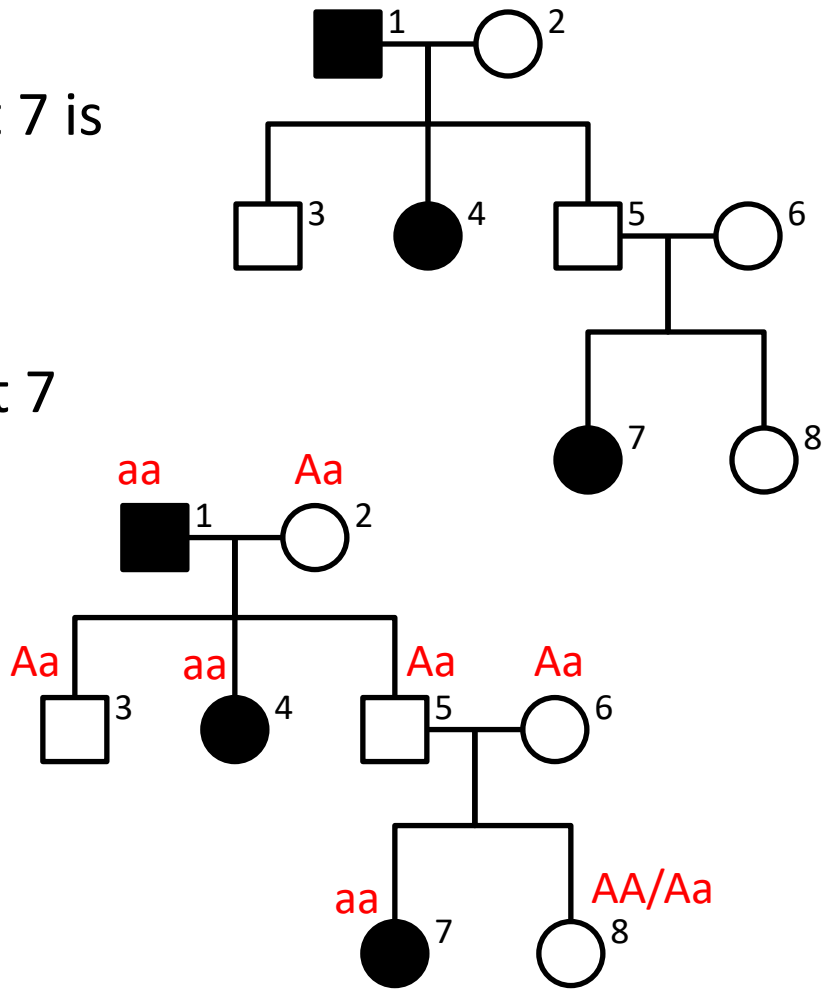
Exercise 2: Pedigree

- Given the pedigree chart, among dominant and recessive, also among autosomal, X-linked and Y-linked, discuss:
  - What are the possible cases of the disorder?
- In addition, infer the genotype of each person (suppose the locus related to the disorder only has two alleles A and a).



Answer 2: Pedigree

- Disease is recessive
  - 5 and 6 are not affected, but 7 is affected.
- Disease is on an autosome
  - 5 (father) is not affected, but 7 (daughter) is affected
- Inferred genotypes:
  - 1, 4, 7: `aa`
  - 2, 3, 5, 6: `Aa`
  - 8: `AA` or `Aa`



Epigenetics

- The study of inherited traits caused by mechanisms other than changes in the underlying DNA sequence.
- Epigenetic marks tell your genes to switch on or off
- Tissues have specific patterns of epigenetic modification
- Two types of epigenetic marks:
  - Chemical (e.g., Methylation)
  - Protein (e.g., Histones)
- Epigenetic marks can be inherited or accumulated.

HIGH-THROUGHPUT SEQUENCING AND APPLICATIONS

Genomics

- The study of an organism's complete set of genetic information.
- The genome includes both genes (coding) and non-coding DNA.
- 'Genome': the complete genetic information of an organism.

vs

Genetics

- The study of heredity
- The study of the function and composition of single genes.
- 'Gene': specific sequence of DNA that codes for a functional molecule.

Prenatal diagnosis

Diagnosis	Description	Danger to fetus
Ultrasonography	Use sound wave to obtain the image of physical features of fetus, like neck fold	No harm
Amniocentesis	Obtain sample of amniotic fluid and test for alpha-fetoprotein level	Risk of infection and spontaneous abortion
Chorionic Villus Sampling (CVS)	Collect fragments of placental tissue and perform sequencing or karyotyping experiments afterwards	Pregnancy loss 2 – 6%
Fetal genomic map from maternal blood	Obtain **maternal plasma**, separate fetal secrete and deduce **fetal genomic map** by parallel sequencing	No harm

Noninvasive prenatal genetic testing

- Diagnosis of fetal chromosomal aneuploidy.

PNAS

Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma

Rossa W. K. Chiu[a,b], K. C. Allen Chan[a,b], Yuan Gao[c,d], Virginia Y. M. Lau[a,b], Wenli Zheng[a,b], Tak Y. Leung[e], Chris H. F. Foo[f], Bin Xie[c], Nancy B. Y. Tsui[a,b], Fiona M. F. Lun[a,b], Benny C. Y. Zee[f], Tze K. Lau[e], Charles R. Cantor[g,1], and Y. M. Dennis Lo[a,b,1]

[a]Centre for Research into Circulating Fetal Nucleic Acids, Li Ka Shing Institute of Health Sciences, Departments of [b]Chemical Pathology and [e]Obstetrics and Gynaecology, and [f]Centre for Clinical Trials, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR, China; [c]Center for the Study of Biological Complexity and [d]Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23284; and [g]Sequenom, Inc., San Diego, CA 92121

Noninvasive prenatal genetic testing

- Construction of fetal genetic map.

RESEARCH ARTICLE | PRENATAL DIAGNOSIS

Maternal Plasma DNA Sequencing Reveals the Genome-Wide Genetic and Mutational Profile of the Fetus

Y. M. Dennis Lo[1,2,*], K. C. Allen Chan[1,2], Hao Sun[1,2], Eric Z. Chen[1,2], Peiyong Jiang[1,2], Fiona M. F. Lun[1,2], Yama W. Zheng[1,2], ...
+ See all authors and affiliations

Science Translational Medicine 08 Dec 2010:
Vol. 2, Issue 61, pp. 61ra91
DOI: 10.1126/scitranslmed.3001720

Noninvasive prenatal genetic testing

- Detection of fetal de novo mutation.

PNAS

Second generation noninvasive fetal genome analysis reveals de novo mutations, single-base parental inheritance, and preferred DNA ends

K. C. Allen Chan[a,b,1], Peiyong Jiang[a,b,1], Kun Sun[a,b,1], Yvonne K. Y. Cheng[c], Yu K. Tong[a,b], Suk Hang Cheng[a,b], Ada I. C. Wong[a,b], Irena Hudecova[a,b], Tak Y. Leung[c], Rossa W. K. Chiu[a,b,2], and Yuk Ming Dennis Lo[a,b,2]

[a]Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong; [b]Department of Chemical Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong; and [c]Department of Obstetrics and Gynaecology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong

Molecular Analytic Techniques

- Techniques for handling samples from CVS, amniocentesis, etc.

Technique	Description	Data
Karyotyping	Allow cells to grow, use microscope to detect the size, shape and number of chromosomes, so as to deduce anomalies like duplications, deletions and rearrangements	Karyotype
Fluorescent in situ Hybridization (FISH)	Detect and locate specific DNA sequence in interphase or metaphase	Fluorescence signals
Polymerase Chain Reaction (PCR)	Amplify specific DNA or RNA fragments, in order to identify specific gene mutation	Present or absence of target
Array Comparative Genomic Hybridization (Array-CGH)	Chromosomal microarray analysis, to detect copy number variation (CNV) at high resolution	Fluorescence signals
Next Generation Sequencing	To obtain many short reads from the amplified cDNA fragments	Many short reads

Check list

- Give one example of genetic disease which is autosomal dominant, autosomal recessive and X-linked recessive, respectively.
- Give one example of prenatal testing method and discuss its advantages and disadvantages.
- What is Polymerase Chain Reaction (PCR)?
- Why do people want to do genetic testing?
- What are the ethic issues related to genetic testing?