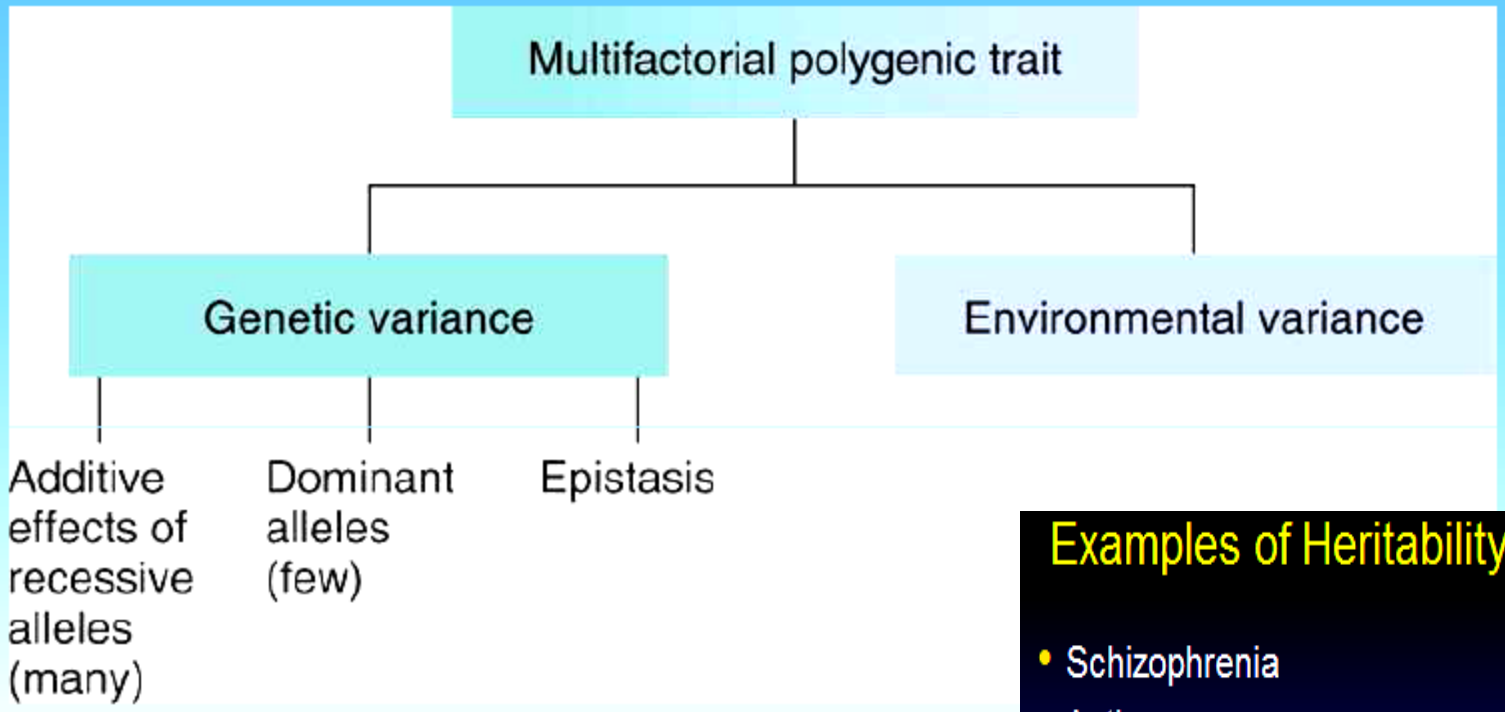
MULTIFACTORIAL DISEASES

MG L-10

July 7th 2014

Heritability (H)

Estimates the proportion of the phenotypic variation in a population due to genetic differences



- Heritability is estimated from the proportion of people sharing a trait compared to the proportion predicted genetically to share the trait
- May vary between populations and time period

Examples of Heritability Estimates

• Schizophrenia	85
• Asthma	80
• Pyloric stenosis	75
• Ischaemic heart disease	65
• Essential hypertension	60
• Spina bifida	60
• Diabetes mellitus	40

Estimates of Heritability of Some Disorders

Disorder	Frequency (%)	Heritability
Schizophrenia	1	85
Asthma	4	80
Cleft Lip = Cleft palate	0.1	76
pylonc stenosis	0.3	75
Ankylosingspondylitis	0,2	70
Club foot .	0.1	68
Coronaryartery dlsease	3	65
Hypertension {essential)	5	62
Congenital dislocction of the hip	0.1	60
Anencephaly and spina pifida	0.1	60
Peptic Ulcer	4	37
Congenital Heart Disease	0.5	35

Concordance

- **Concordance -** the percentage of pairs in which both twins express the trait
- Used to determine heritability
- Has limitations, assumes both type of twins share similar environments
- MZ twins often share more similar environments

Concordance in MZ and DZ Twins

	Concordance Values (%)	
Trait	MZ Twins	DZ Twins
Blood types	100	66
Eye color	99	28
Mental retardation	97	37
Hair color	89	22
Down syndrome	89	7
Handedness (left or right)	79	77
Epilepsy	72	15
Diabetes	65	18
Tuberculosis	56	22
Cleft lip	42	5

Twin studies provide an insight into the interaction of genotypes and environment

How to identify Quantitative Trait Loci (QTL)

Linkage and Association Studies

"Linkage Disequilibrium" – alleles are inherited together (rather than genes)

- LD only ranges a short distance ~ 10,000 bases
- Because alleles are so close they are always inherited together (no crossing over)

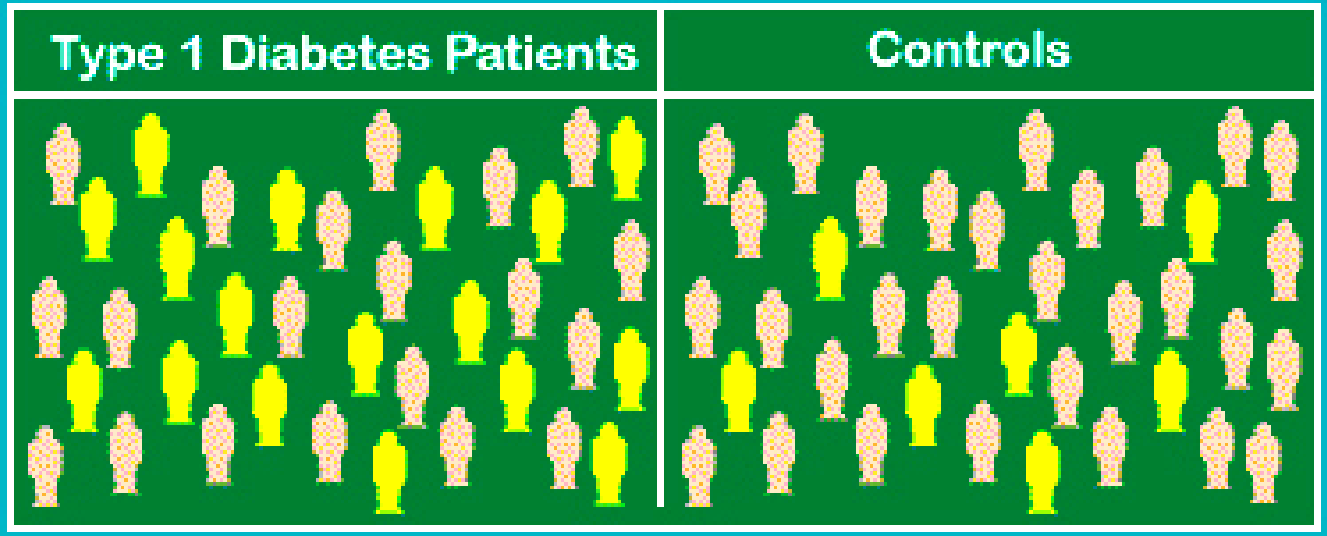
- Association comparing alleles
- Linkage usually done in families, association usually done case vs. control

Association Studies

- Studies which compare a group of interest (cases) to a control group for the presence of a gene or SNP.
- Controls are matched to cases for characteristics that may confound results: age, ethnicity, gender, environment.
- If the SNP is present more often in cases than controls, it is associated with the trait and implies that the SNP may be near a gene impacting the trait.

Association studies in diabetes type 1

Association Studies



Genotype	Type 1	Controls	Total
HLA DR4	17	7	24
NON-HLA DR4	20	30	50
	37	37	

$X^2_{.05} = 5.377$

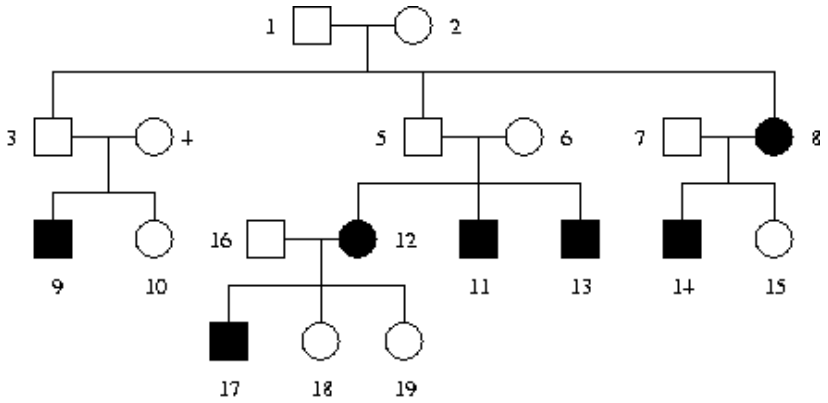
$p < 0.025$

= HLA DR4

= non-HLA DR4

Genetic linkage and linkage analysis

- Two loci are **linked** if they appear close by in the same chromosome.
- The task of linkage analysis is to find markers that are linked to the hypothetical disease locus
- Complex diseases in focus $\rightarrow$ usually need to search for one gene at a time
- Requires mathematical modelling of meiosis
  - One of the two main approaches in gene mapping.
  - Uses pedigree data



Correlation

- Correlation coefficient
  - ✓The fraction of genes shared by two relatives
- Identical twins have 100% of their genes in common (correlation coefficient = 1.0)

Relationship to Proband	Proportion of Alleles in Common with Proband
Monozygotic (MZ) twins	1
Dizygotic (DZ) twins	1/2
First-degree relative	1/2
Second-degree relative	1/4
Third-degree relative	1/8

Type of Information Used in Genome-Wide Association Studies

Marker Type	Definition
SNP	A single nucleotide polymorphism is a site in the genome that is a different DNA base in >1% of a population.
CNV	A copy number variant is a tandemly repeated DNA sequence, such as CGTA CGTA CGTA
Gene expression	The pattern of genes that are overexpressed and/or overexpressed in people with a particular trait or disease.
	Epigenetic signature

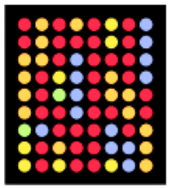
Genome-wide association studies seek SNPs that are shared with much greater frequency among individuals with the same trait than among others

People with disorder

People without disorder

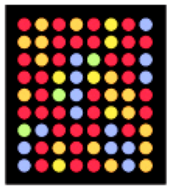
Patient DNA

Non-Patient DNA



Disease-specific SNPs

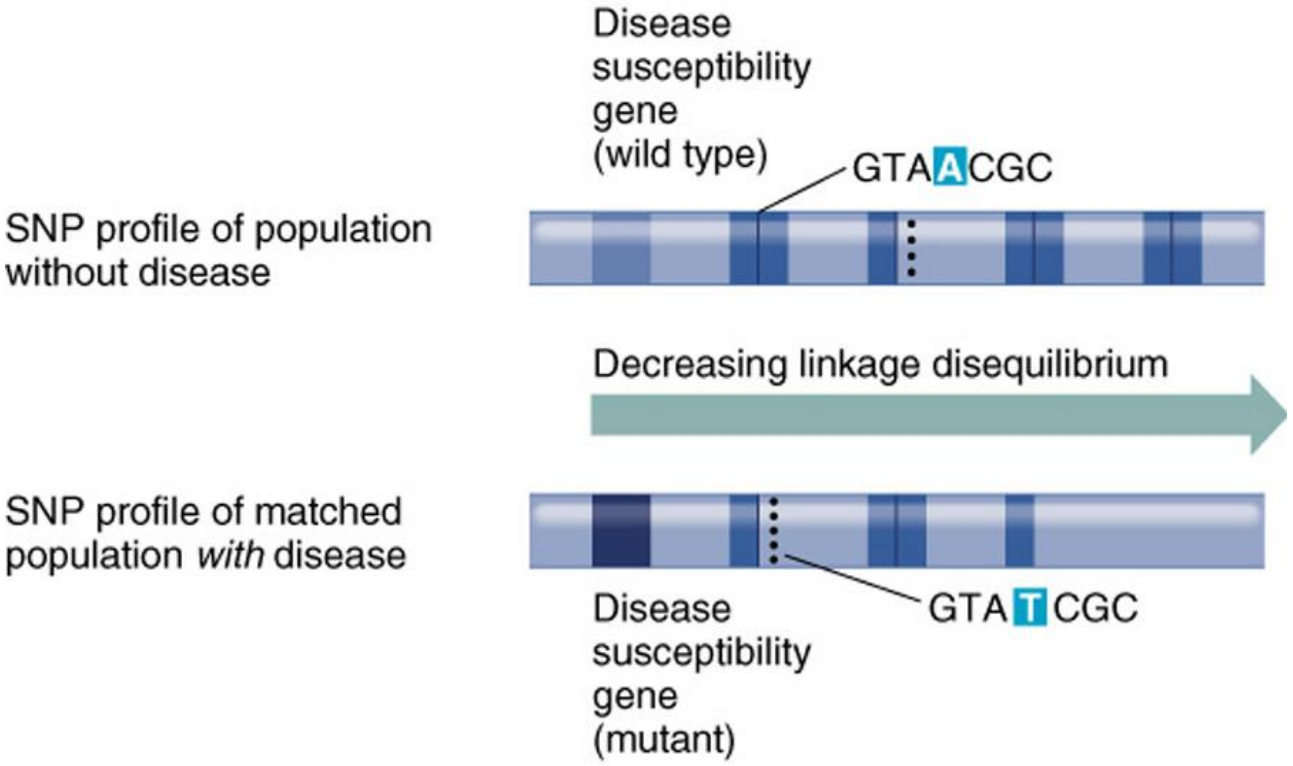
Compare differences to discover SNPs associated with disease



Nondisease SNPs

SNP (single nucleotide polymorphism)

Nucleotide site with more than one allele is a polymorphism.



- **On average between two random individuals, there is one SNP every 1000 bases => 3 million differences!**

Conclusions

- Multifactorial disorders are more common than single gene and chromosomal disorders
- They are caused by the interaction of many genes with environmental factors
- Optimum preventive measures rely on avoidance of the bad environmental factors since avoidance of inheriting the bad genes is at present not possible.
- These measures can be explained through counseling such as preconception and chronic noncommunicable diseases counseling.

Population Genetics

Population

- Population - is an interbreeding group of the same species within a given geographical area
- Subpopulation - any of the breeding groups within a population among which migration is restricted
- Local population - subpopulation within which most individuals find their mates
- Gene pool - the collection of all alleles in the members of the population
- Gene Flow - alleles can move between populations when individuals migrate and mate

Phenotypic Evolution: Process

Population Differentiation

What forces are responsible for population differentiation and how do they affect genetic diversity?

- Mutation ↑ genetic diversity
- Selection ↑↓ genetic diversity
- Genetic drift ↓ genetic diversity
- Migration ↑↓ genetic diversity
- Non-random mating ↓ genetic diversity

Heterozygote Superiority

- Heterozygote superiority = fitness (measurement of viability and fertility) of heterozygote is greater than that of both homozygotes
- When there is heterozygote superiority, neither allele can be eliminated by selection
- In sickle cell anemia, allele for mutant hemoglobin is maintained in high frequencies in regions of endemic malaria because heterozygotes are more resistant to his disease

Population Genetics

- Gene pool = the complete set of genetic information in all individuals within a population
- Genotype frequency = proportion of individuals in a population with a specific genotype may differ from one population to another
- Allele frequency = proportion of any specific allele in a population, are estimated from genotype frequencies

Allele Frequencies

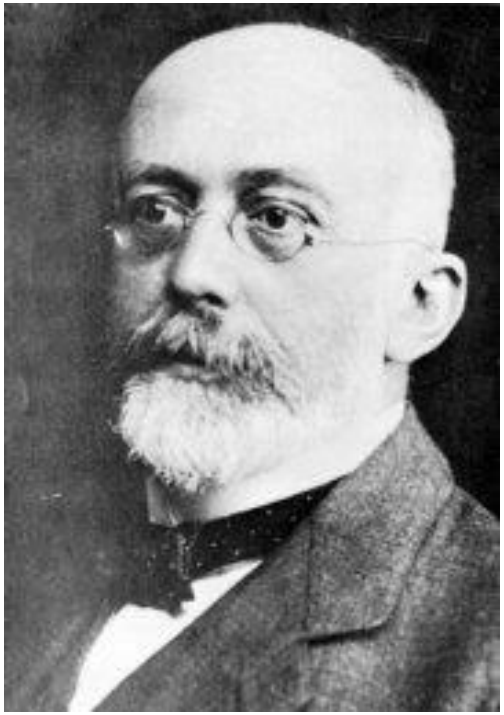
$$\text{Allele frequency} = \frac{\text{No of particular allele}}{\text{total No of alleles in the population}}$$

- Both chromosomes should be count of each individual
- Allele frequencies affect the genotype frequencies (frequency of each type of homozygote and heterozygotes) in the population.

Frequency of PKU in different Populations

Population	Frequency of PKU
Chinese	1/16,000
Irish, Scottish, Yemenite Jews	1/5,000
Japanese	1/119,000
Swedes	1/30,000
Turks	1/2,600
United States Caucasians	1/10,000

Hardy-Weinberg Theory



Hardy-Weinberg Principle Depends Upon the Following Assumptions

1. There is no selection
2. There is no mutation
3. There is no migration
4. There are no chance events
5. 5. Individuals choose their mates at random

Using the Hardy-Weinberg Law in Human Genetics

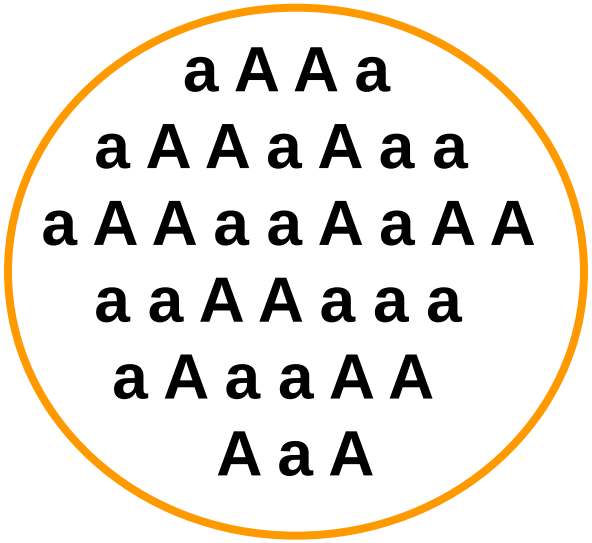
- The Hardy-Weinberg Law can be used to
  - Estimate frequencies of autosomal dominant and recessive alleles in a population
  - Detect when allele frequencies are shifting in a population (evolutionary change)
  - Measure the frequency of heterozygous carriers of deleterious recessive alleles in a population

Assumptions:

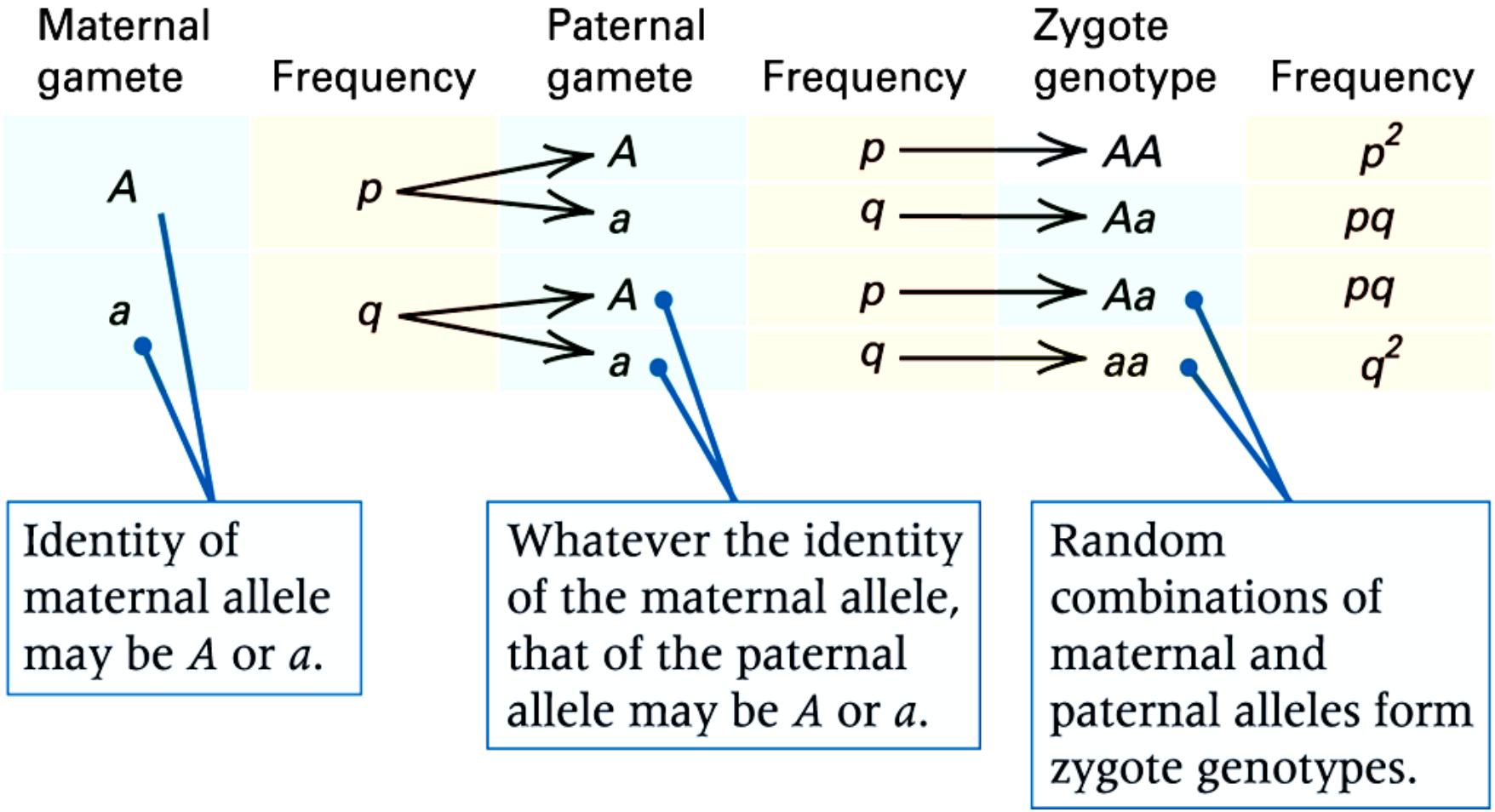
1) Diploid, autosomal locus with 2 alleles: **A** and **a**
2) Simple life cycle:

PARENTS (DIPLIOD) → GAMETES (HAPLOID) → ZYGOTES (DIPLOID)

These parents produce a large gamete pool (Gene Pool) containing alleles **A** and **a**.



Allele frequencies when mating is random



One locus, 2 Allele Model

In **A diploid organism, there are two alleles for each locus. Therefore there are three possible genotypes:**

Genotype A_1A_1 A_1A_2 A_2A_2

Given:

Frequency of allele A_1 = p
Frequency of allele A_2 = 1 - p = q

Then:

Genotype A_1A_1 A_1A_2 A_2A_2

Frequency p^2 $2pq$ q^2

A population that maintains such frequencies is said to be at Hardy-Weinberg Equilibrium

Hardy-Weinberg Equilibrium

$p + q = 1$

All of the allele frequencies together equals 1 or the whole collection of alleles

p allele frequency of one allele
q allele frequency of a second allele

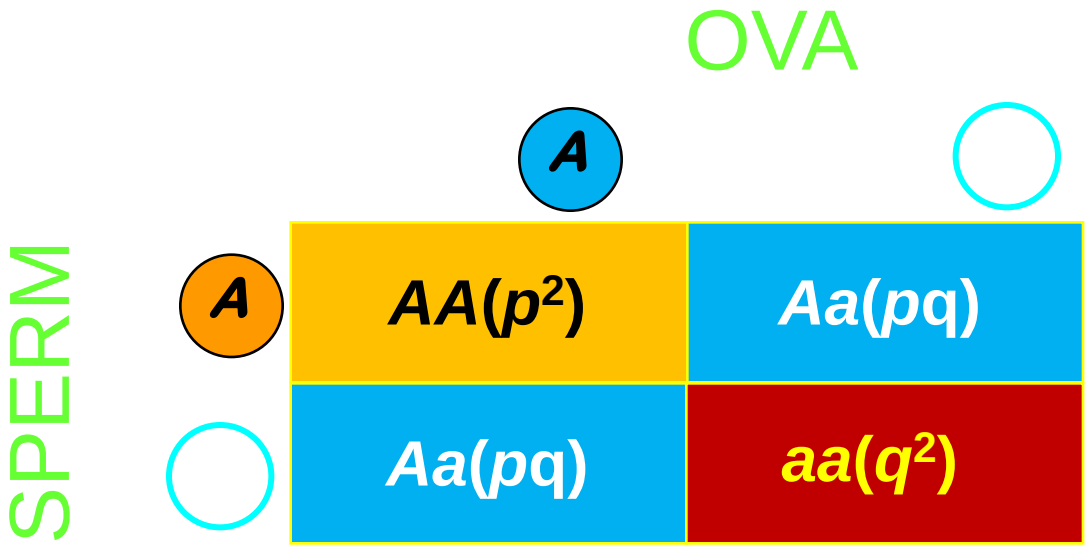
$p^2 + 2pq + q^2 = 1$

All of the genotype frequencies together equals 1

p^2 and q^2 genotype frequencies for each homozygote

2pq genotype frequency for heterozygotes

EXAMPLES OF HARDY WEINBERG



p^2 = homozygous dominant

2pq = heterozygous

q^2 = homozygous recessive

THE HARDY WEINBERG PRINCIPLE

Step 1

- Calculating the gene frequencies from the genotype frequencies
- Easily done for codominant alleles (each genotype has a different phenotype)

HARDY-WEINBERG PROBLEM EXAMPLE 1:

- **Given: In a population of 747 individuals (1494 alleles),**

- **Problem:**
  - **Find the allele frequencies for A and a.**
  - **Find the genotypic frequencies of AA, Aa, and aa.**

Example : The MN blood group

Sample Population	Phenotypes	Type M	Type MN		Type N
	Genotypes	M^mM^m	M^mM^n		M^nM^n
747	Numbers	233	385		129
	Contribution to gene pool	2 M^m alleles per person	1 M^m allele per person	1 M^n allele per person	2 M^n alleles per person

MN blood group in Iceland

Total $\mathbf{M^m}$ alleles = (2 x 233) + (1 x 385) = 851

Total $\mathbf{M^n}$ alleles = (2 x 129) + (1 x 385) = 643

Total of **both** alleles =1494

= 2 x 747

(humans are diploid organisms)

Frequency of the $\mathbf{M^m}$ allele = 851/1494 = 0.57

or 57%

Frequency of the $\mathbf{M^n}$ allele = 643/1494 = 0.43

or 43%

In General for a diallellic gene A and a (or A^x and A^y)

If the frequency of the **A** allele = **p**

and the frequency of the **a** allele = **q**

Then **p+q** **=** **1**

Step 2

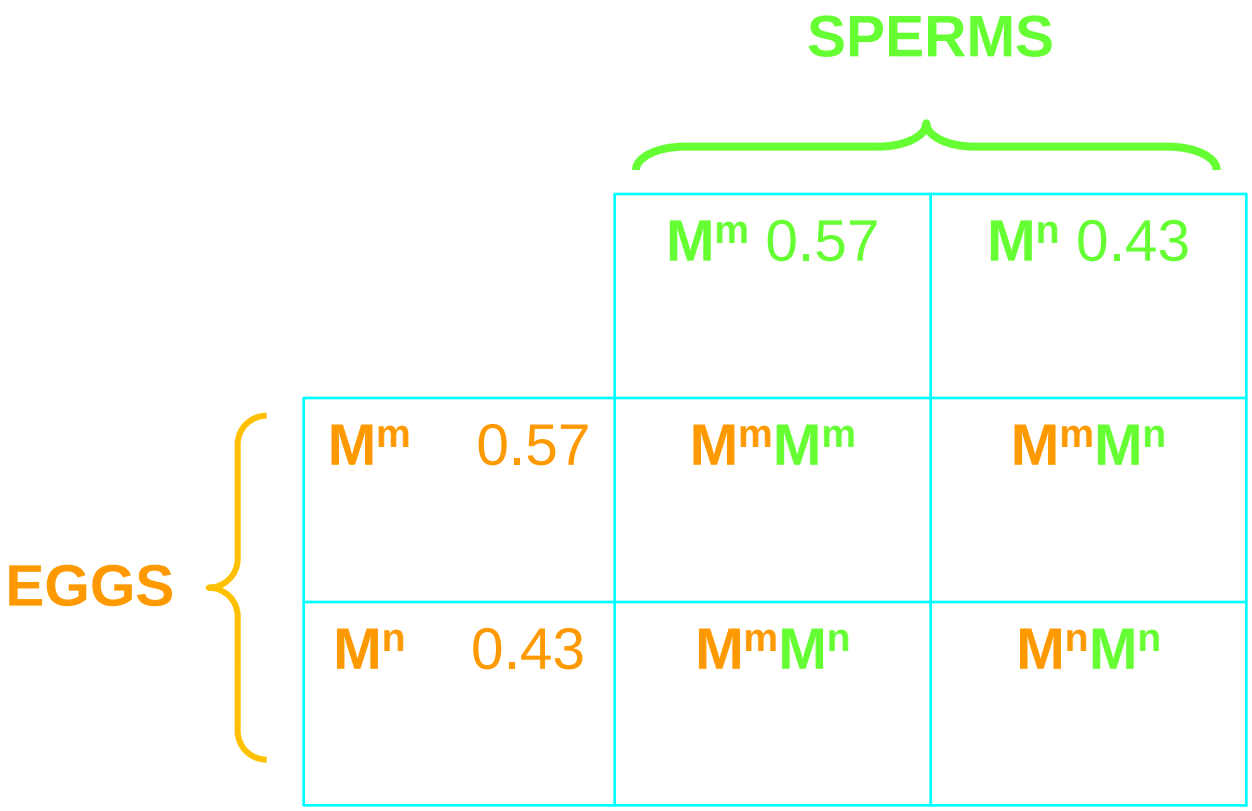
- **Using the calculated gene frequency to predict the EXPECTED genotypic frequencies in the NEXT generation**

OR

- **to verify that the PRESENT population is in genetic equilibrium**

Assuming all the individuals mate <u>randomly</u>

NOTE the **gene frequencies** are the **gamete frequencies** too



Close enough for us to assume genetic equilibrium

Genotypes	Expected frequencies	Observed frequencies
M^mM^m	0.32	233 ÷ 747 = 0.31
M^mM^n	0.50	385 ÷ 747 = 0.52
M^nM^n	0.18	129 ÷ 747 = 0.17

Important

- **Need to remember the following:**

p^2 = homozygous dominant

$2pq$ = heterozygous

q^2 = homozygous recessive

Describing genetic structure

- Genotype frequencies
- Allele frequencies

rr = white

Rr = pink

RR = red

Describing genetic structure

- **genotype frequencies**
- **allele frequencies**

200 white

500 pink

300 red

genotype frequencies:

200/1000 = 0.2 rr

500/1000 = 0.5 Rr

300/1000 = 0.3 RR

Describing genetic structure

- **genotype frequencies**
- **allele frequencies**

200 rr = 400 r

500 Rr = 500 r
= 500 R

300 RR = 600 R

allele frequencies:

900/2000 = 0.45 r

1100/2000 = 0.55 R

total = 2000 alleles

Keep In Mind

- *The frequency of recessive alleles in a population cannot be measured directly*

Calculating the Frequency of **Autosomal Dominant and** Recessive Alleles

- Count the frequency of individuals in the population with the recessive phenotype, which is also the homozygous recessive genotype (*aa*)
  - The frequency of genotype *aa* = q^2
  - The frequency of the *a* allele is $\sqrt{q^2} = q$
  - The frequency of the dominant allele (*A*) is calculated $p = 1 - q$

Calculating the Frequency of Alleles for X-Linked Traits

- For X-linked traits, females (XX) carry 2/3 of the alleles and males (XY) carry 1/3 of the alleles
- The number of males with the mutant phenotype equals the allele frequency for the recessive trait

  - Frequency of an X-linked trait in males is q
  - Frequency of the trait in females is q^2

Risk Calculations in X-linked Traits

Females: $p^2 + 2pq + q^2 = 1$ — All of the women in the population

Males: $p + q = 1$ — All of the men in the population

Hemophilia is X-linked and occurs in 1 in 10,000 males

p= 1/10,000 = .0001

q= .9999

Carrier females = 2pq
= .0002 — 1 in 5000 are carriers

Affected females = $p^2 = (.0001)^2$
= .00000001 — 1 in 100 million women will have hemophilia

Calculating the Frequency of Multiple Alleles

- In ABO blood types, six different genotypes are possible (*AA, AO, BB, BO, AB, OO*)
  - Allele frequencies: $p\ (A) + q\ (B) + r\ (O) = 1$
  - Genotype frequencies: $(p + q + r)^2 = 1$

- Expanded Hardy-Weinberg equation:
  - $p^2\ (AA) + 2pq\ (AB) + 2pr\ (AO) + q^2\ (BB) + 2qr\ (BO) + r^2\ (OO) = 1$

Factors that effect the genetics of populations:

Intrinsic factors

- Segregation
- Recombination
- Transposition
- Mutation

Extrinsic factors

- Population size
- Patterns of mating
- Geographic distribution
- Migration
- Natural selection

Hardy-Weinberg Theorem

- Population gene and genotypic frequencies don’t change over generations if is at or near equilibrium.

- Population in equilibrium means that the populations isn’t under evolutionary forces (Assumptions for Equilibrium*)