CONSANGUINITY

*

*

I- DEFINITION

- A subject is in a situation of consanguinity if, for a given locus, (s)he has two identical alleles, per copy of *one and the same ancestor gene*.
- The coefficient of consanguinity (Cc or F) is the probability that the two allele genes that an individual has at a locus are identical *by descendance*.

--> this assumes that a common ancestor (A) is shared by the parents, F and M, of the individual, I studied.

II- COEFFICIENT OF CONSANGUINITY OF AN INDIVIDUAL

II-1. FORMULATION

- First, look for the common ancestor (or common ancestors) on the genealogical tree.
- Then, calculate the probabilities; for example:

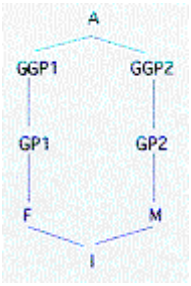

Figure 1

- A possesses alleles a1 and a2. It transmits to the great-grandparents GGP:
    - either identical alleles (a1 and a1 or: a2 and a2) --> proba: 1/2
    - or different alleles (a1 and a2 or: a2 et a1),

but... if A itself exhibits consanguinity (with a coefficient of consanguinity F_A), then a1 and a2 have a probability F_A of being identical, and A transmits a1 and a2 with a proba 1/2, i.e. F_A x 1/2

Overall, A transmits the identity with a proba: 1/2 + 1/2 F_A, or: 1/2 (1 + F_A) Note: F_A can be equal to zero.

- Each generation i has a proba 1/2 of transmitting this allele to i+1; therefore a proba $(1/2)^n$ after n generations; or $(1/2)^p$ to go from GGP1 to I and $(1/2)^m$ to go from GGP2 to I, if *f and m are the number of links linking the father and mother respectively to the common ancestor* (here p = m = 3)

- Therefore: $FI = (1/2)^{p+m+1}(1+F_A)$ ...
- And, if there are several common ancestors (not consanguinous with each other), a sum Σ, addition of the various consanguinities, gives us the:

II-2. GENERAL EQUATION: $FI = \Sigma(1/2)^{p+m+1}(1+F_{Ai})$

- Note: F_A is negligible in man and at the level of the individual, but may not be in Drosophila, particularly in an entire population.

- Genealogy studies are indispensable for determing consanguinity; see: **Genealogy and Coefficient of Consanguinity, Exercices**

III- CONSANGUINITY OF A POPULATION

The mean coefficient of consanguinity is equal to the mean of the various individual coefficients weighted by the frequencies of the various types of crosses between related individuals.

To evaluate this, an inventory is compiled for the individuals of the different types of crossings between related individuals, and they are classified on the basis of the value of Fx.

EQUATION $\alpha = \Sigma F_i f_i$ where f_i is the frequency of the subjects with consanguinity F_i.

- Example: a population in which 6% are consanguin, and among these: 2.5% have a value of F = 1/8; 2% a value of F = 1/16, and 1.5% a value of F = 1/32; what is the consanguinity of this population?

- Answer: α= (2.5 X 1/8) + (2 X 1/16) + (1.5 X 1/32) = 0.484%

IV- SELF-FERTILIZATION

This means that each genotype is fertilized exclusively by itself (a situation that is possible in maize (corn), but not in Drosophila, or in Man).

In a population of plants, underHW in Go, which is then put in a situation of self-fertilization:

	AA		Aa		aa
Go	0.25		0.50		0.25
self-fertilization	AA	AA	Aa	aa	aa
	X 1	X1/4	X1/2	X1/4	X1
G1	0.25	0.125	0.25	0.125	0.25
	AA		Aa	aa	etc...

What is the frequency H_n of the heterozygotes in generation n?

$H_n = 1/2\ H_n\text{-}1 \dashrightarrow H_n = (1/2)^n H_o$; tends towards zero.
$D_n = D_n\text{-}1 + 1/4\ H_n\text{-}1$
$R_n = R_n\text{-}1 + 1/4\ H_n\text{-}1$

--> at the equilibrium of self-fertilization: $D_{eq} = D_o + 1/2\ H_o$; $H_{eq} = 0$; $R_{eq} = R_o + 1/2\ H_o$

Being under HW in G_o, $D_o = p^2$; $H_o = 2pq$; $R_o = q^2$

--> $D_{eq} = p^2 + 1/2\ 2pq = p^2 + pq = p\ (p + q) = p$; similarly for R_{eq} -->

- Genotype frequencies at equilibrium:

$D_{eq} = p$
$H_{eq} = 0$
$R_{eq} = q$

V- GENERALIZATION

Genotype	General case	panmixia	self-fertilization
	$0 <= F <= 1$	F=0	F=1
	allozygotism + autozygotism		
AA	$p^2(1-F) + pF$	p^2	p
Aa	2pq(1-F)	2pq	0
aa	$q^2(1-F) + qF$	q^2	q

Thus, any population (and this includes a consanguin population) will behave as though:

- one fraction (1-F) was developing in panmixia

- one fraction F was developing in self-fertilization
F being the mean coefficient of consanguinity of the population. F=0 in panmixia, F=1 in self-fertilization

- The genotype frequencies at equilibrium will be:

$F(AA)eq = p^2(1 - F) + pF = p^2 - p^2F + pF = p^2 + Fp (1 - p) = p^2 + Fpq$; similarly for F(aa); thus:

V-1. GENOTYPE FREQUENCIES AT EQUILIBRIUM
EQUATION:

$F(AA)eq = p^2 + Fpq$
$F(Aa)eq = 2pq(1 - F)$
$F(aa)eq = q^2 + Fpq$

EQUATION
The risk that a consanguin subject will be homozygotic for the allele a is: $F(aa) = q^2 + Fpq$
- Another demonstration of the relations $F(AA) = p^2 + Fpq$, $F(Aa) = 2pq(1 - F)$, et $F(aa) = q^2 + Fpq$ is given in: **Genetic Constitution of Consanguin Populations**

- Are the allele frequencies altered?

$F(A) = D + H/2 = p^2 + Fpq + 2pq(1 - F)/2 = p^2 + Fpq + pq - Fpq = p^2 + pq = p(p + q) = p$ --> invariable; therefore:

V-2. PROPERTIES OF CONSANGUINITY
Consanguinity:

- modifies the genotype frequencies. We can see an increase in the frequency of homozygotes and a reduction in that of heterozygotes.
- does not modify the allele frequencies

VI- HUMAN POPULATION

It is usual within the human population for there to be several common ancestors (e.g. below: AM and AF male and female ancestors, parents of GP 1 and 2). In practice, the equation is simplified to:

EQUATION $CcI = \Sigma(1/2)^{p+m+1}$

Example: Parents first cousins: p=2; m=2; Σ is the sum of 2 terms, since there are 2 possibilities of having identical alleles: via AM and via AF (i.e. 2 common ancestors);

so,

Answer: $Fi = (1/2)^{2+2+1} + (1/2)^{2+2+1} = 1/16$

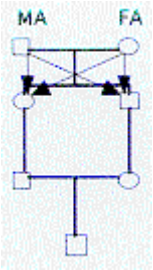

Figure 2

VII-1. GENETIC COUNSELLING

For a deleterious mutant recessive autosomal allele (rare by definition) with a frequency q, the risk that a consanguin child will be homozygotic for this allele is: ***q x Cc*** *whereas it is* q^2 *for the children of non-consanguin parents.*

- Note: the exact equation q^2+ pqCc is replaced by the approximation: q x Cc. This is applicable to human genetics (genetic counselling) if/because q is very small.

Exercice 1: Parents first cousins; for a mutant gene with recessive autosomal transmission with frequency q = 1/100 (example of phenylketonuria, one of the most common recessive autosomal diseases): what is the risk in the general population? what is the risk that I could be affected?
Answer:

- for the general population: $q^2 = 1/10\ 000$
- for I, it is, according to the equation: q x Cc = 1/100 x 1/16 = 1/1 600
- Note: the risk of a recessive autosomal disease in the individual I, relative to the general population, is increased by the factor: q x Cc / q^2 = Cc / q here = 6.25 (or, if we use the accurate equation (q^2+ pqCc) / q^2 = 7.19)

Exercice 2: same exercice, but for a frequency q = 1/10 000 of the mutant gene.br
Answer:

- for the general population: $q^2 = 1/100\ 000\ 000$
- for I, it is, according to the equation: q x Cc = 1/10 000 x 1/16 = 1/160 000
- the risk for individual I relative to the general population, is increased by a factor Cc / q here = 625 or, if we use the accurate equation, q^2+ pqCc / q^2 = 626 (the rarer the allele, the better the approximation q x Cc).

<

VIII- RARE ALLELE - COMMON ALLELE

VIII-1. EXERCICE

- consider a gene A of which the recessive allele a has a freqency F(a) = q = 0.5,
- and a gene B of which the recessive allele b has a freqency F(b) = q = 0.0001, a fairly usual frequency for a morbid allele,

calculate the frequency/risk of being recessive homozygote(s) for each of these two genes
1. according to Hardy-Weinberg (HW)
2. for a consanguin child whose parents are first-cousins
3. compare

Answer:

1. according to HW:

$F(aa) = q^2 = (0.5)^2 = 0.25$

$F(bb) = q^2 = (0.0001)^2 = (10^{-4})^2 = 10^{-8}$

2. for a consanguin child whose parents are first-cousins:

- $\Sigma(1/2)^{p+m+1} = (1/2)^5 + (1/2)^5 = (1/2)^4 = 0.0625$
- $F(aa) = q^2 + Fpq = (0.5)^2 + 0.0625 \text{ X } 0.5 \text{ X } 0.5 = 0.2656$
- $F(bb) = q^2 + Fpq = (10^{-4})^2 + 0.0625 \text{ X } 1 \text{ X } 10^{-4} = (1 + 625)10^{-8} = 626 \text{ X } 10^{-8}$

3. comparison: the increase in the frequency (risk) of homozygotism due to the consanguinity will be

F(consang.)/F(under HW), i.e.:

1. for the common allele: 0.26256 / 0.25 = 1.06, slight increase
2. for the rare morbid allele: $626 \text{ X } 10^{-8}/10^{-8} = 626$!!!
3. **VIII-2. PRACTICAL CONSEQUENCES**
   In other words, the increase is: $(q^2 + Fpq)/q^2 = 1 + Fp/q$

- when p = q: --> = 1 + F ≅ 1
- when q is rare, p ≅ 1 --> = 1 + F/q ≅ F/q, with Fmax = 0.25 (incest), F is often of the order of 1 to 5 X 10^{-2} and q = 10^{-3} or 10^{-4}, therefore a risk increased by a factor of 10 to 10^3, which is generally the case for recessive diseases. Isolates, with a high level of consanguinity, allow unusual diseases to emerge.

IX- CONSANGUINITY - HETEROZYGOTISM - ISOGENETIC LINE

In the human species, the percentage of heterozygotic loci, calculated from enzymatic polymorphism, has a value H = 0.067. We can take it that there are 30000 structural genes, and in consequence 2010 genes in the heterozygotic state in the human genome (30000 x 0.067 = 2010).
If an individual results from an uncle-niece cross:
this individual will be more "homogenous" than his or her parents, because of the increased consanguinity,
the percentage of his/her heterozygotic genes falls from 2010 to 1759 genes (2010 x 7/8) since Fi = 1/8 (1/8 of genes are identical as a result of consanguinity).

Consequences: If regular consanguin crosses are made (for example brother/sister crosses in mice), at each generation:
--> Fi tends towards a value of 1,
--> the individuals will become totally homozygotic.

Within each familly, all the individuals will be identical in the genetic sense of the word.
exactly the same genome
exactly the same genes.

This leads to the concept of the isogenetic line.

X- MULTIALLELE SYSTEM

The genotype frequencies at equilibrium for every A_iA_i homozygote and every A_iA_j: heterozygote will be:

$$F(A_iA_i) = p_i^2(1 - F) + p_iF$$
$$F(A_iA_j) = 2p_ip_j(1 - F)$$

X-1. EXERCICE: CONSANGUINITY FOR A LOCUS AND THREE ALLELES

In a population of a diploid species with separate sexes and separate generations, we are dealing with a triallele autosomal locus (three possible allele states: A1, A2 and A3).

A sample of 400 individuals is examined. The numbers of the various genotypes are as follows:

A1A1	A1A2	A1A3	A2A2	A2A3	A3A3
32	36	60	57	90	125

4. Estimate the allele frequencies.
5. Can we assume that we have the proportions of panmixia in the sample?
6. Knowing that there is no selection, no mutation, no migration, no drift (large population), can consanguinity account for the difference? What are the theoretical proportions of the various genotypes, knowing that the mean coefficient of consanguinity is F?
7. Estimate the value of F from the proportions of the sample.

Answer:

8. The allele frequencies are estimated by counting the alleles. Thus, for allele A1 for example:

   frequency of A1 = ((2 x 32) + 36 + 60) / (2 x 400) = 0.20 = p
   similarly: frequency of A2= 0.30 = q; frequency of A3 = 0.50 = r

9. The proportions of panmixia are in fact those indicated by the Hardy-Weinberg law.

   a) Theoretical frequencies of the genotypes according to the Hardy-Weinberg law

A1A1:	$p^2 = 0.20^2$	A1A2:	2 pq = 2 x 0.20 x 0.30
A2A2:	$q^2 = 0.30^2$	A1A3:	2 pr = 2 x 0.20 x

			0.50
A3A3:	$r^2 = 0.50^2$	A2A3:	2 qr = 2 x 0.30 x 0.50

with p, q, r: the respective frequencies of the alleles A1, A2 and A3.

b) Theoretical numbers

A1A1:	0.202 x 400 =16	A1A2:	48
A2A2:	36	A1A3:	80
A3A3:	100	A2A3:	100

c) Comparison of the theoretical numbers and the actual numbers found by a chi-squared test of conformity
χ^2 = (32 - 16)2/16 + ... + (125 - 120)2/100 = 50
dd1 = 6 - 2 - 1 = 3; to a 5% threshold, [[chi]]2 calculated is greater than the value given in the table (7.815) and is therefore highly significant. Consequently the proportions of the genotypes do not comply with those of the Hardy-Weinberg law, the hypothesis of panmixia can be rejected.

10. Consanguinity has the effect of increasing the frequency of homozygotes and of reducing that of heterozygotes, relative to the proportions given by the Hardy-Weinberg law. This is indeed what we find. Consequently, consanguinity can account for the significant differences between the previous theroretical and actual numbers.

    For the whole population, the theoretical genotype frequencies are:

A1A1: (1- F) p^2 + Fp</TD< TD>	A1A2: 2 pq (1 - F)
A2A2: (1 - F) q^2 + Fq</TD< TD>	A1A3: 2 pr (1 - F)
A3A3: (1 - F) r^2 + Fr</TD< TD>	A2A3: 2 qr (1 - F)

11.

12. The frequency of A1A2 can be used simply to calculate F:

frequency of A1A2 = 2 pq (1 - F)
36/400 = 2 x 0.20 x 0.30 (1 - F) therefore 1 - F = 0.75
F = 0.25

Contributor(s)

Written 06-2002 Robert Kalmes, Jean-Loup Huret

Citation

This paper should be referenced as such :

Kalmes R, Huret JL . Consanguinity. Atlas Genet Cytogenet Oncol Haematol. June 2002 .
URL : http://AtlasGeneticsOncology.org/Educ/ConsangID30039ES.html