



Hvor kommer du fra?

Hvordan kan vi bruge data fra projektet i undervisningssammenhæng?

Slutkonference

Aulaen på Aarhus Universitet, d. 31 marts 2014

Frank Grønlund Jørgensen

Ph.d. i biologi fra AU med fokus på molekylær evolution
Lektor på Tørring Gymnasium i biologi og fysik

Email: fg@toerring-gym.dk



Fem ideer til undervisning

Bestem genotype- & allelfrekvenser	Eleverne kan bestemme genotype- og allelfrekvenser i hele populationen eller i mindre subpopulationer (region, klasse mm)
Test for Hardy-Weinberg proportioner	For hver enkelt SNP-loci er det muligt at teste om en population udviser HW-proportioner. Det samme gælder for subpopulationer.
F_{ST} & UPGMA	F_{ST} kan bruges til at sammenligne forventede genotypfrekvenser i forskellige populationer. Ud fra en F_{ST} afstandsmatrice kan man lave et UPGMA-træ over subpopulationer.
Retsgenetiske undersøgelser	Ud fra de enkelte elevs genotype profil i udvalgt SNP-loci kan man undersøge simple retsgenetiske spørgsmål.
Hvad skyldes egenskaben?	Ved hjælp af et simpelt associationskortlægningsstudie kan eleverne bruge egne data til at undersøge om genotypen i forskellige SNPs er korreleret med en bestemt egenskab.



Hvilke typer data får vi?

- En SNP (Single Nucleotide Polymorphism) er en bestemt position i genomet vi ved varierer når man sammenligner forskellige mennesker. 6 SNPs er markeret med rødt herunder.
- Op mod en million variable steder (SNPs) i genomet genotypes ved hjælp af en DNA-chip



... **T**TGAG ... AG**T**T ... AAT**G**ACA ...
 ... **A**ACTC ... T**C**AA ... TT**A**CTGT ...

GT**C**T ... T**A**CG**T**AC ... ATAT**T**CA**T** ...
CAG**A** ... AT**G**C**A**TG ... TATA**A**GT**A** ...

... **T**TGAG ... AG**T**G ... AAT**G**ACA ...
 ... **A**ACTC ... T**C**AC ... TT**A**CTGT ...

GT**C**A ... T**A**CG**T**AC ... ATAT**T**CA**T** ...
CAG**T** ... AT**G**C**A**TG ... TATA**A**GT**A** ...

TT **TG** **GG** **TA** **TT** **TT**



... **T**TGAG ... AG**T**T ... AAT**G**ACA ...
 ... **A**ACTC ... T**C**AA ... TT**A**CTGT ...

GT**C**G ... T**A**CG**G**AC ... ATAT**T**CA**T** ...
CAG**C** ... AT**G**C**C**TG ... TATA**A**GT**A** ...

... **T**GGAG ... AG**T**T ... AAT**G**ACA ...
 ... **A**CCTC ... T**C**AA ... TT**A**CTGT ...

GT**C**C ... T**A**CG**G**AC ... ATAT**C**CA**T** ...
CAG**G** ... AT**G**C**C**TG ... TATA**G**GT**A** ...

TG **TT** **GG** **GC** **GG** **TC**



... **T**GGAG ... AG**T**T ... AAT**G**ACA ...
 ... **A**CCTC ... T**C**AA ... TT**A**CTGT ...

GT**C**A ... T**A**CG**T**AC ... ATAT**T**CA**T** ...
CAG**T** ... AT**G**C**A**TG ... TATA**A**GT**A** ...

... **T**GGAG ... AG**T**T ... AAT**G**ACA ...
 ... **A**CCTC ... T**C**AA ... TT**A**CTGT ...

GT**C**G ... T**A**CG**T**AC ... ATAT**T**CA**T** ...
CAG**C** ... AT**G**C**A**TG ... TATA**A**GT**A** ...

GG **TT** **GG** **AG** **TT** **TT**

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014

Hvordan ser data ud i virkeligheden?

Rådata kommer som en simpel tekst fil med følgende informationer for hver enkelt person

# rsid	chromosome	position	genotype
rs4477212	1	82154	AA
rs3094315	1	752566	AG
rs3131972	1	752721	AG
rs12124819	1	776546	AG
rs6681049	1	800007	CC
rs4970383	1	838555	CC



... **T**AGAG ... AG**T**A ... AAT**A**ACA ...
 ... **A**TCTC ... T**C**A**T** ... TT**A**TGT ...

GT**C**A ... T**A**CG**C**AC ... ATAT**C**CA**T** ...
CAG**T** ... AT**G**C**G**TG ... TATA**G**GT**A** ...

... **T**AGAG ... AG**T**G ... AAT**G**ACA ...
 ... **A**TCTC ... T**C**AC ... TT**A**CTGT ...

GT**C**G ... T**A**CG**C**AC ... ATAT**C**CA**T** ...
CAG**C** ... AT**G**C**G**TG ... TATA**G**GT**A** ...

AA **AG** **AG** **AG** **CC** **CC**

↑
rs4477212

↑
rs3131972

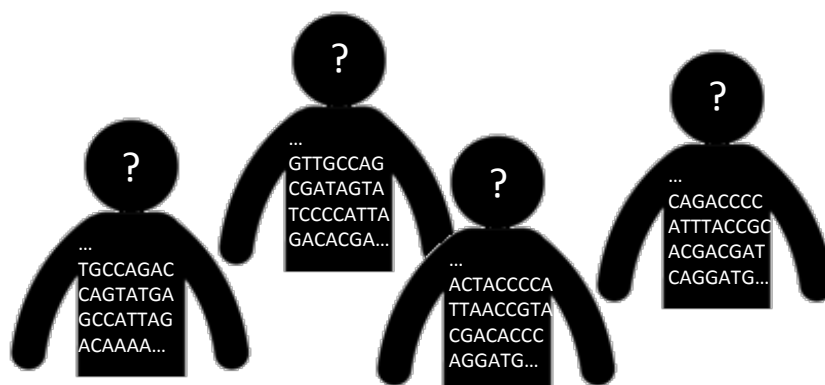
↑
rs4970383

- Ca. 600.000 rækker for hver person
- Vi får ca. 800 personers rådata

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014

Hvilke typer data får I?

- Studiet er anonymiseret så I kan ikke få fri adgang til jeres personlige data!
- Udvalgte dele af data anonymiseres og frigives til brug i undervisningen.
- Udvalgt data og tilhørende vejledninger kan hentes på hvorkommerdufra.dk efterhånden som de bliver færdigudviklet.



Hvor kommer du fra - Frank Grønlund Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014



Bestem genotype- og allelfrekvenser

- Genotypefrekvenser og allelfrekvenser (hyppigheder) er grundlæggende for mange populationsgenetiske undersøgelser



Genotypefrekvenser

Hvor kommer du fra - Frank Grønlund Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014



Bestem genotype- og allelfrekvenser

- Genotypefrekvenser og allelfrekvenser (hyppigheder) er grundlæggende for mange populationsgenetiske undersøgelser



Genotypefrekvenser

$$f(AA) = 5 / 10 = 0,5$$

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Bestem genotype- og allelfrekvenser

- Genotypefrekvenser og allelfrekvenser (hyppigheder) er grundlæggende for mange populationsgenetiske undersøgelser



Genotypefrekvenser

$$f(AA) = 5 / 10 = 0,5$$

$$f(TT) = 2 / 10 = 0,2$$

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Bestem genotype- og allelfrekvenser



Hvad bliver $f(AT)$?

- A. 0,2
- B. 0,3
- C. 0,4
- D. 0,5
- E. 1

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Bestem genotype- og allelfrekvenser

• Genotypefrekvenser og allelfrekvenser (hyppigheder) er grundlæggende for mange populationsgenetiske undersøgelser



Genotypefrekvenser

$$f(AA) = 5 / 10 = 0,5$$

$$f(TT) = 2 / 10 = 0,2$$

$$f(AT) = 3 / 10 = 0,3$$



I et enkelt locus skelner vi ikke mellem AT og TA

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Bestem genotype- og allelfrekvenser

• Genotypefrekvenser og allelfrekvenser (hyppigheder) er grundlæggende for mange populationsgenetiske undersøgelser



Genotypefrekvenser

$$f(AA) = 5 / 10 = 0,5$$

$$f(TT) = 2 / 10 = 0,2$$

$$f(AT) = 3 / 10 = 0,3$$

$$\text{I alt: } f(AA) + f(TT) + f(AT) = 1$$

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Bestem genotype- og allelfrekvenser



Genotypefrekvenser

$$f(AA) = 5 / 10 = 0,5$$

$$f(TT) = 2 / 10 = 0,2$$

$$f(AT) = 3 / 10 = 0,3$$

Hvad bliver allelfrekvensen $f(A)$?

- A. 0
- B. 0,50
- C. 0,55
- D. 0,65
- E. 0,75
- F. 1

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Bestem genotype- og allelfrekvenser

- Genotypfrekvenser og allelfrekvenser (hyppigheder) er grundlæggende for mange populationsgenetiske undersøgelser



Genotypfrekvenser

$$f(AA) = 5 / 10 = 0,5$$

$$f(TT) = 2 / 10 = 0,2$$

$$f(AT) = 3 / 10 = 0,3$$

$$\text{I alt: } f(AA) + f(TT) + f(AT) = 1$$

Allelfrekvenser

- Optæl allelantal og divider med populationens samlede antal alleller

$$f(A) = 13 / 20 = 0,65$$

$$f(T) = 7 / 20 = 0,35$$

$$\text{I alt: } f(A) + f(T) = 0,65 + 0,35 = 1$$

I forbindelse med undersøgelser af Hardy-Weinberg proportioner kaldes de to allelfrekvenser normalt for p og q (lige meget hvilke type bi-allel SNP vi arbejder med).

Hvor kommer du fra - Frank Grønlund Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014



Bestem genotype- og allelfrekvenser

- Man kan bestemme og sammenligne genotype- og allelfrekvenser fra forskellige subpopulationer.

Subpopulation 1



Genotypfrekvenser

$$f(AA) = 5 / 10 = 0,5$$

$$f(TT) = 2 / 10 = 0,2$$

$$f(AT) = 3 / 10 = 0,3$$

Allelfrekvenser

$$f(A) = 13 / 20 = 0,65$$

$$f(T) = 7 / 20 = 0,35$$

Subpopulation 2



Genotypfrekvenser

$$f(AA) = 3 / 10 = 0,3$$

$$f(TT) = 2 / 10 = 0,2$$

$$f(AT) = 5 / 10 = 0,5$$

Allelfrekvenser

$$f(A) = 11 / 20 = 0,55$$

$$f(T) = 9 / 20 = 0,45$$

Hvor kommer du fra - Frank Grønlund Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014



Bestem genotype- og allelfrekvenser

- Eller man kan se på de samlede genotype- og allelfrekvenser i den totale population

Subpopulation 1



Genotypefrekvenser

$$f(AA) = 5 / 10 = 0,5$$

$$f(TT) = 2 / 10 = 0,2$$

$$f(AT) = 3 / 10 = 0,3$$

Allelfrekvenser

$$f(A) = 13 / 20 = 0,65$$

$$f(T) = 7 / 20 = 0,35$$

Subpopulation 2



Genotypefrekvenser

$$f(AA) = 3 / 10 = 0,3$$

$$f(TT) = 2 / 10 = 0,2$$

$$f(AT) = 5 / 10 = 0,5$$

Allelfrekvenser

$$f(A) = 11 / 20 = 0,55$$

$$f(T) = 9 / 20 = 0,45$$

Totalpopulation



Genotypefrekvenser

$$f(AA) = 8 / 20 = 0,4$$

$$f(TT) = 4 / 20 = 0,2$$

$$f(AT) = 8 / 20 = 0,4$$

Allelfrekvenser

$$f(A) = 24 / 40 = 0,60$$

$$f(T) = 16 / 40 = 0,40$$

Hvor kommer du fra · Frank Grønland Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Hardy-Weinberg proportioner

- Hvis man har en uendelig stor lukket population uden mutation og selektion opstår der efter en generation med tilfældig parring en bestemt sammenhæng mellem allelfrekvenserne og genotypefrekvenserne i populationen
- Denne sammenhæng kaldes for Hardy-Weinberg proportioner.
- I et SNP-locus med to alleler (1,2) er der tre mulige genotyper (11, 12, 22)
- Vi kalder de to allelfrekvenser $f(1)$ og $f(2)$ i det undersøgte SNP-locus for henholdsvis p og q
- $p + q = 1$

Hvor kommer du fra · Frank Grønland Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Hvilke af følgende sammenhænge beskriver HW-proportioner?

- A. $p^2 + q^2 = 1$
- B. $p^2 + q^2 = 2pq$
- C. $p^2 + pq + q^2 = 1$
- D. $p^2 + 2pq + q^2 = 1$
- E. $p + q + 2pq = 0$

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Hardy-Weinberg proportioner

• Hvis man har en uendelig stor lukket population uden mutation og selektion opstår der efter en generation med tilfældig parring en bestemt sammenhæng mellem allelfrekvenserne og genotyperefrekvenserne i populationen

Forventede genotyperefrekvenser ved HW-proportioner

$$f(11) = p^2$$

$$f(12) = 2pq$$

$$f(22) = q^2$$

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

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Test for Hardy-Weinberg proportioner

• Eksempel baseret på en tilfældig SNP

Subpopulation 1

Østjylland (n = 117)



Genotypfrekvenser

$$f(AA) = 65 / 117 = 0,56$$

$$f(AT) = 46 / 117 = 0,39$$

$$f(TT) = 6 / 117 = 0,05$$

Allelfrekvenser

$$p = f(A) = 176 / 234 = 0,75$$

$$q = f(T) = 58 / 234 = 0,25$$

Subpopulation 2

København (n = 59)



Genotypfrekvenser

$$f(AA) = 40 / 59 = 0,68$$

$$f(AT) = 16 / 59 = 0,27$$

$$f(TT) = 3 / 59 = 0,05$$

Allelfrekvenser

$$p = f(A) = 96 / 118 = 0,81$$

$$q = f(T) = 22 / 118 = 0,19$$

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Test for Hardy-Weinberg proportioner

• χ^2 -test kan bruges til at teste nulhypotesen: H_0 = Populationen udviser HW-proportioner

Subpopulation 1

Østjylland (n = 117)



Genotypfrekvenser

$$f(AA) = 65 / 117 = 0,56$$

$$f(AT) = 46 / 117 = 0,39$$

$$f(TT) = 6 / 117 = 0,05$$

Allelfrekvenser

$$p = f(A) = 176 / 234 = 0,75$$

$$q = f(T) = 58 / 234 = 0,25$$

Forventet antal af hver genotype ved HW-proportioner

$$AA: p^2 \cdot N = 0,75^2 \cdot 117 = 65,8$$

$$AT: 2pq \cdot N = 2 \cdot 0,75 \cdot 0,25 \cdot 117 = 43,9$$

$$TT: q^2 \cdot N = 0,25^2 \cdot 117 = 7,3$$

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Test for Hardy-Weinberg proportioner

• χ^2 -test kan bruges til at teste nulhypotesen: H_0 = Populationen udviser HW-proportioner

Subpopulation 1 Østjylland (n = 117)



Genotypfrekvenser

$$f(AA) = 65 / 117 = 0,56$$

$$f(AT) = 46 / 117 = 0,39$$

$$f(TT) = 6 / 117 = 0,05$$

Allelfrekvenser

$$p = f(A) = 176 / 234 = 0,75$$

$$q = f(T) = 58 / 234 = 0,25$$

χ^2 -test

$$\chi^2 \text{ - teststørrelsen} = \sum_{i=1}^k \frac{(\text{observerede} - \text{forventede})^2}{\text{forventede}}$$

Der er 1 frihedsgrad i en χ^2 -test for HW-proportioner.

Forventet antal af hver genotype ved HW-proportioner

$$AA: p^2 \cdot N = 0,75^2 \cdot 117 = 65,8$$

$$AT: 2pq \cdot N = 2 \cdot 0,75 \cdot 0,25 \cdot 117 = 43,9$$

$$TT: q^2 \cdot N = 0,25^2 \cdot 117 = 7,3$$

χ^2 -teststørrelsen

$$\chi^2 \text{ - teststørrelsen} = \frac{(65 - 65,8)^2}{65,8} + \frac{(46 - 43,9)^2}{43,9} + \frac{(6 - 7,3)^2}{7,3} = 0,347$$

Teststørrelsen skal vurderes i en χ^2 -fordeling med 1 frihedsgrad hvilket giver p-værdi = 0,56.

Altså afviger subpopulation 1 (Østjylland) **ikke signifikant** fra HW-proportioner.

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Test for Hardy-Weinberg proportioner

• Er der HW-proportioner når man betragter den totale stikprøve af danskere?

Alle elever på tværs af landet Alle syv geografiske regioner (n = 499)



Genotypefordeling

$$AA: 304$$

$$AT: 157$$

$$TT: 38$$

$$\text{I alt: } 304 + 157 + 38 = 499$$

Allelfrekvenser

$$p = f(A) = 765 / 998 = 0,77$$

$$q = f(T) = 233 / 998 = 0,23$$

Forventet antal under HW-proportioner

$$AA: 0,77^2 \cdot 499 = 295,9$$

$$AT: 2 \cdot 0,77 \cdot 0,23 \cdot 499 = 176,7$$

$$TT: 0,23^2 \cdot 499 = 26,4$$

$$\chi^2 \text{ - teststørrelsen} = \frac{(304 - 295,9)^2}{295,9} + \frac{(157 - 176,7)^2}{176,7} + \frac{(38 - 26,4)^2}{26,4} = 7,300$$

$$p\text{-værdi} = 0,007$$

Den totale population afviger **signifikant** fra HW-proportioner på et 5 % signifikansniveau.

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Årsager til afvigelse fra HW-proportioner

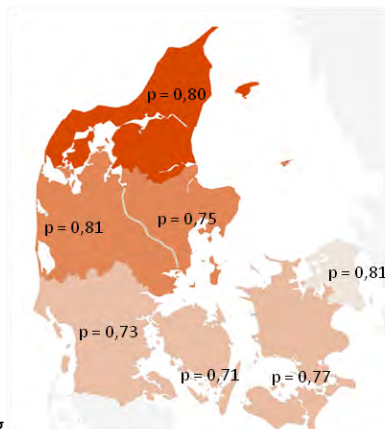
Den totale population afviger **signifikant** fra HW-proportioner ($p = 0,007$)

Tilfældigheder

- Med en p-værdi på 0,007 forventer vi at komme til at forkaste nulhypotesen (HW-proportioner) selvom den er sand 7/1000 gange.
- Vi undersøger ca. 600.000 SNPs så vi forventer at se ekstreme resultater ind i mellem ved et tilfælde. Derfor korrigerer man normalt p-værdien når man laver mange tests!

Regionale forskelle

- Korter viser at allelfrekvenserne p (og q) varierer en smule fra region til region.
- Måske afspejler dette reel genetiske differentiering mellem landsdelene, men det er i så fald meget lidt og vi kan ikke sige det med sikkerhed uden yderligere undersøgelser.



Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Regionale forskelle – kan de måles?

Regionale forskelle

- Hvis afvigelsen fra HW-proportioner afspejler reel genetisk differentiering kan man bruge parvise F_{ST} mål som en slags afstandsmål mellem populationerne.

Først estimeres H_S for et par af subpopulationer som et vægtet gennemsnit af de forventede antal heterozygoter under antagelse af HW-proportioner.

$$H_S(1,2) = (N_1 \cdot 2p_1q_1 + N_2 \cdot 2p_2q_2) / (N_1 + N_2) = (95 \cdot 0,3502 + 67 \cdot 0,3218) / (95 + 67) = 0,3384$$

Derefter estimeres H_T for samme par. Dette gøres ved at lægge alle observationer sammen, estimere p_{1+2} og q_{1+2} for den samlede population (1+2) og udregne $2p_{1+2}q_{1+2}$.

$$H_T(1,2) = 2p_{total}q_{total} = 2 \cdot 0,783951 \cdot 0,216049 = 0,3387$$

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

$$F_{ST}(1,2) = \frac{0,3387 - 0,3384}{0,3387} = 0,00088$$

Parvis F_{ST}	1	2	3	4	5	6
2	0,00088					
3	0,00063	0,00280				
4	0,00441	0,00999	0,00167			
5	0,00109	0,00002	0,00306	0,01108		
6	0,00274	0,00668	0,00071	0,00035	0,00715	
7	0,00207	0,00036	0,00422	0,01484	0,00023	0,00904

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014

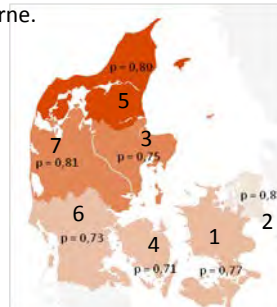


Regionale forskelle – kan de måles?

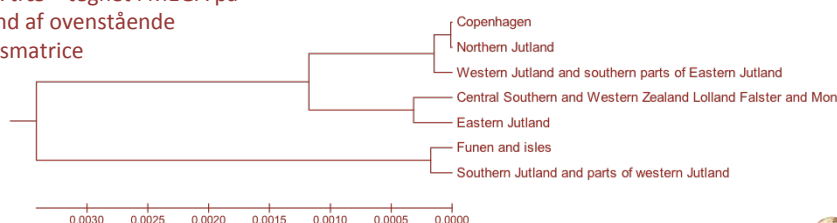
Regionale forskelle

- Hvis afvigelsen fra HW-proportioner afspejler reel genetisk differentiering kan parvise F_{ST} værdier bruges som en slags afstandsmål mellem populationerne.

Parvis F_{ST}	1	2	3	4	5	6
2	0,00088					
3	0,00063	0,00280				
4	0,00441	0,00999	0,00167			
5	0,00109	0,00002	0,00306	0,01108		
6	0,00274	0,00668	0,00071	0,00035	0,00715	
7	0,00207	0,00036	0,00422	0,01484	0,00023	0,00904



UPGMA træ – tegnet i MEGA på baggrund af ovenstående afstandsmatrice



Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Retsgenetiske undersøgelser

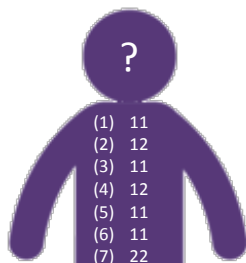
- Normalt benyttes STR-baserede metoder, men SNP-baserede bliver mere almindeligt

	STR-baseret	SNP-baseret
Variation mellem mennesker per locus	Høj	Lav
Type polymorfi	Længe	Baseforskelle
Antal alleler per locus	Mange	Typisk kun 2
Antal mulige genotyper per locus	Mange	Typisk kun 3
Mutationsrate	Høj (10^{-2} - 10^{-4})	Lav (10^{-7} - 10^{-8})
Antal loci der normalt benyttes	10-15 (i DK 10 fra forskellige kromosomer + amelogenin gen)	50-100 uafhængige (ikke koblede) SNP.
Mulige undersøgelsesmetode	Kan laves med PCR og gelelektroforese	Typisk PCR efterfulgt af DNA-sekventering eller genotyping med fx DNA-chip.

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014

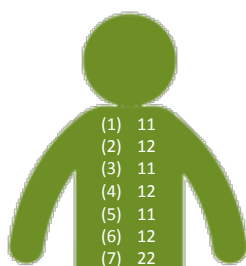


Retsgenetiske undersøgelser



Blodspor fundet på gerningssted

1	2	3	4	5	6	7
11	12	11	12	11	11	22



Kan vi frikende nogen?

- A. Den lilla
- B. Den grønne
- C. Ingen af dem
- D. Dem begge

Hvor kommer du fra - Frank Grønlund Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014



Retsgenetiske undersøgelser

• Udvalgte bi-allele SNPs er undersøgt fra jeres DNA. Allel 1 er altid den mest almindelige allel i populationen, 2 altid den mindst almindelige i det pågældende locus.

Locus	1	2	3	4	5	6	7	8	9	10
Genotype (11)	509	373	534	247	598	424	214	716	475	326
Genotype (12)	182	290	170	332	118	249	363	0	218	318
Genotype (22)	23	56	15	141	5	48	143	0	29	77
I alt genotypet	714	719	719	720	721	721	720	716	722	721

Hvilket af ovenstående 10 loci er mindst brugbar til at identificere en person?

- A. Locus 1
- B. Locus 2
- C. Locus 3
- D. Locus 4
- E. Locus 5
- F. Locus 6
- G. Locus 7
- H. Locus 8
- I. Locus 9
- J. Locus 10

Hvor kommer du fra - Frank Grønlund Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014



Retsgenetiske undersøgelser

- Udvalgte bi-allele SNPs er undersøgt fra jeres DNA. Allel 1 er altid den mest almindelige allel i populationen, 2 altid den mindst almindelige i det pågældende locus.

Locus	1	2	3	4	5	6	7	8	9	10
Genotype (11)	509	373	534	247	598	424	214	716	475	326
Genotype (12)	182	290	170	332	118	249	363	0	218	318
Genotype (22)	23	56	15	141	5	48	143	0	29	77
I alt genotypet	714	719	719	720	721	721	720	716	722	721

- Vi kan udregne allelfrekvenserne $p = f(1)$ og $q = f(2)$ på sædvanlig vis for hvert locus
- Hvis der er HW-proportioner i populationen kan vi udregne de forventede genotypfrekvenser

HW-proportioner	1	2	3	4	5	6	7	8	9	10
$f(11) = p^2$	0,7062	0,5190	0,7412	0,3290	0,8303	0,5787	0,3017	1,0000	0,6543	0,4525
$f(12) = 2pq$	0,2683	0,4028	0,2395	0,4892	0,1618	0,3640	0,4951	0,0000	0,3092	0,4404
$f(22) = q^2$	0,0255	0,0782	0,0193	0,1818	0,0079	0,0572	0,2031	0,0000	0,0365	0,1071

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Retsgenetiske undersøgelser

	Locus									
HW-proportioner	1	2	3	4	5	6	7	8	9	10
$f(11) = p^2$	0,7062	0,5190	0,7412	0,3290	0,8303	0,5787	0,3017	1,0000	0,6543	0,4525
$f(12) = 2pq$	0,2683	0,4028	0,2395	0,4892	0,1618	0,3640	0,4951	0,0000	0,3092	0,4404
$f(22) = q^2$	0,0255	0,0782	0,0193	0,1818	0,0079	0,0572	0,2031	0,0000	0,0365	0,1071

- Hvad er sandsynligheden for at finde en person, med nedenstående genotype, i populationen ved et tilfælde?

	Locus									
DNA-profil	1	2	3	4	5	6	7	8	9	10
	11	12	11	12	11	11	22	11	11	22

Eksempel på beregning for de første to loci

Locus 1: $p^2 = 0,7062$

Locus 2: $2pq = 0,4028$

Samlet sandsynlighed, antagelse af uafhængighed: $\Pr(11(1), 12(2)) = 0,7062 \cdot 0,4028 = 0,2845$

Tilføjer vi locus 8 fås: $\Pr(11(1), 12(2), 11(8)) = 0,7062 \cdot 0,4028 \cdot 1 = 0,2845$

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Retsgenetiske undersøgelser

HW-proportioner	Locus									
	1	2	3	4	5	6	7	8	9	10
$f(11) = p^2$	0,7062	0,5190	0,7412	0,3290	0,8303	0,5787	0,3017	1,0000	0,6543	0,4525
$f(12) = 2pq$	0,2683	0,4028	0,2395	0,4892	0,1618	0,3640	0,4951	0,0000	0,3092	0,4404
$f(22) = q^2$	0,0255	0,0782	0,0193	0,1818	0,0079	0,0572	0,2031	0,0000	0,0365	0,1071

DNA-profil	Locus									
	1	2	3	4	5	6	7	8	9	10
DNA-profil	11	12	11	12	11	11	22	11	11	22
Sandsynlighed	0,7062	0,4028	0,7412	0,4892	0,8303	0,5787	0,2031	1,0000	0,6543	0,1071

Hvor kommer du fra - Frank Grønland Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014



Retsgenetiske undersøgelser

DNA-profil	Locus									
	1	2	3	4	5	6	7	8	9	10
DNA-profil	11	12	11	12	11	11	22	11	11	22
Sandsynlighed	0,7062	0,4028	0,7412	0,4892	0,8303	0,5787	0,2031	1,0000	0,6543	0,1071

Hvilke af ovenstående 10 loci ser ud til at bidrage mest til at identificere personen?

- A. Locus 1
- B. Locus 2
- C. Locus 3
- D. Locus 4
- E. Locus 5
- F. Locus 6
- G. Locus 7
- H. Locus 8
- I. Locus 9
- J. Locus 10

Hvor kommer du fra - Frank Grønland Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014



Retsgenetiske undersøgelser

	Locus									
HW-proportioner	1	2	3	4	5	6	7	8	9	10
f(11) = p ²	0,7062	0,5190	0,7412	0,3290	0,8303	0,5787	0,3017	1,0000	0,6543	0,4525
f(12) = 2pq	0,2683	0,4028	0,2395	0,4892	0,1618	0,3640	0,4951	0,0000	0,3092	0,4404
f(22) = q ²	0,0255	0,0782	0,0193	0,1818	0,0079	0,0572	0,2031	0,0000	0,0365	0,1071

	Locus									
DNA-profil	1	2	3	4	5	6	7	8	9	10
DNA-profil	11	12	11	12	11	11	22	11	11	22
Pr (loci)	0,7062	0,4028	0,7412	0,4892	0,8303	0,5787	0,2031	1,0000	0,6543	0,1071
Pr (akkumuleret)	0,7062	0,2844	0,2108	0,1031	0,0856	0,0496	0,0101	0,0101	0,0066	0,0007

7 / 10.000

Hvor kommer du fra - Frank Grønlund Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014



Retsgenetiske undersøgelser

	1	2	3	4	5	6	7
DNA-profil	11	12	11	12	11	11	22
Pr (loci)	0,7062	0,4028	0,7412	0,4892	0,8303	0,5787	0,2031
Pr (akkumuleret)	0,7062	0,2844	0,2108	0,1031	0,0856	0,0496	0,0101

Blodspor fundet på gerningssted

1	2	3	4	5	6	7
11	12	11	12	11	11	22



(1) 11
(2) 12
(3) 11
(4) 12
(5) 11
(6) 11
(7) 22



Kan vi frikende den blå eller grønne person?
Har vi fat i den rigtige person?



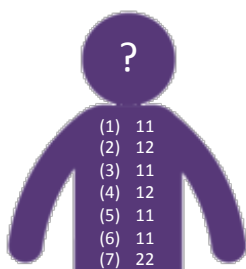
(1) 11
(2) 12
(3) 11
(4) 12
(5) 11
(6) 12
(7) 22

Hvor kommer du fra - Frank Grønlund Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014



Har vi fat i den rigtige person?

	1	2	3	4	5	6	7
DNA-profil	11	12	11	12	11	11	22
Pr (loci)	0,7062	0,4028	0,7412	0,4892	0,8303	0,5787	0,2031
Pr (akkumuleret)	0,7062	0,2844	0,2108	0,1031	0,0856	0,0496	0,0101



1 / 100

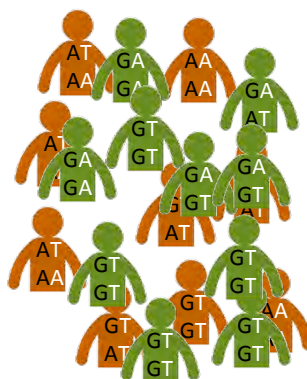


Hvor kommer du fra - Frank Grønland Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014



Hvad skyldes egenskaben?

- DNA chip data kan bruges til at undersøge genetikken bag forskellige egenskaber
- Fx kan vi undersøge arvegangen for grøn / orange hudfarve og se om det korrelerer med en af de to forskellige SNPs vist (en sort SNP og en hvid SNP)
- Princippet er simpelt – Vi inddeler populationen efter hvorvidt de er grønne og orange og ser om der er et mønster i de to gruppers genotyper.

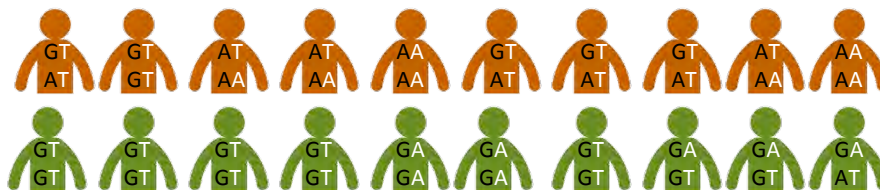


Hvor kommer du fra - Frank Grønland Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014



Hvad skyldes egenskaben?

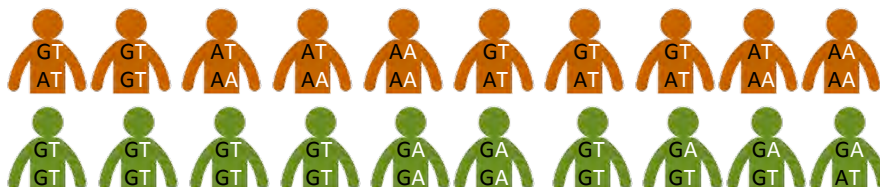
- DNA chip data kan bruges til at undersøge genetikken bag forskellige egenskaber
- Fx kan vi undersøge arvegangen for grøn / orange hudfarve og se om det korrelerer med en af de to forskellige SNPs vist (en sort SNP og en hvid SNP)
- Princippet er simpelt – Vi inddeler populationen efter hvorvidt de er grønne og orange og ser om der er et mønster i de to gruppers genotyper.



Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Hvad skyldes egenskaben?



Hvilke af følgende sætninger tror du er mest korrekt?

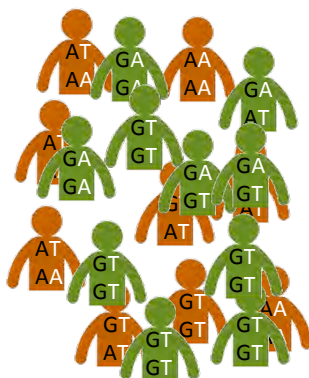
- At have genotypen TT I SNP-2 (hvid) er positivt korreleret med grøn hud
- At have genotypen GG I SNP-1 (sort) er giver altid grøn hud
- Ingen af de undersøgte SNPs ser ud til at være positivt korreleret med grøn hud
- At have genotypen GG I SNP-1 (sort) ser ud til at være positivt korreleret med grøn hud.

Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



Hvad skyldes egenskaben?

- Mange egenskaber er ikke enten/eller egenskaber
- Ikke mange egenskaber styres 100% af en SNP
- Ofte samspil mellem flere SNPs (gener)
- Ofte samspil med miljøet
- Vi bliver hele tiden klogere men mange ting er meget komplekse.

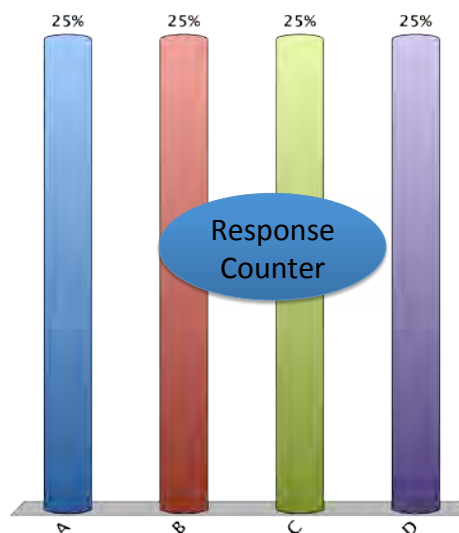


Hvor kommer du fra · Frank Grønlund Jørgensen · Tørring Gymnasium & Centre for Biocultural History · Aarhus Universitet · Marts 2014



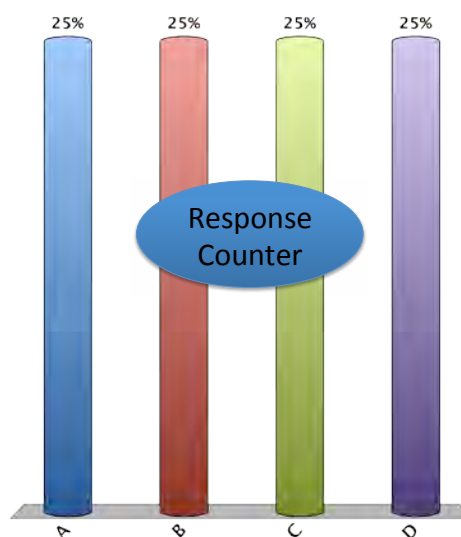
Hvad er din mening om at studiet er anonymt?

- Det har jeg det bedst med
- Jeg ville gerne have kendt mit eget data
- Det har jeg ikke nogen mening om



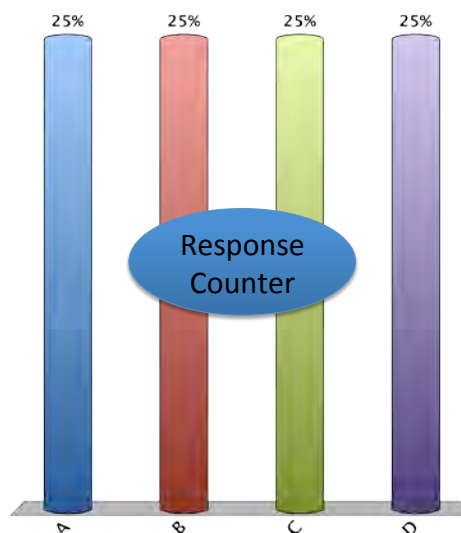
Hvad synes du om det faglige niveau i projektet?

- A. Det har været passende
- B. Det har været for højt
- C. Det har været for lavt



Hvad er dit overordnede indtryk af projektet?

- A. Det har været interessant og lærerigt
- B. Det har ikke givet mig noget fagligt eller personligt
- C. Det har jeg ingen mening om



Tusind tak fordi I deltog i projektet

Følg med i projektet på hvorkommerdufra.dk



SPØRGSMÅL?



Hvor kommer du fra - Frank Grønlund Jørgensen - Tørring Gymnasium & Centre for Biocultural History - Aarhus Universitet - Marts 2014

