

# **Evolutionsbiologi**

**Svar til tidligere eksamensopgaver**

Hans R. Siegismund

Biologisk Institut  
Københavns Universitet

# Evolutionsbiologi

## Skriftlig eksamen, den 17. juni 2015

## Med hjælpemidler (Lærebøger, noter, computer)

For hver opgave oplyses dens vægtning.

*Dette opgavesæt indeholder **9** sider inklusive forsiden  
Der må **IKKE** benyttes blyant.  
Opgaverne skal afleveres i den rækkefølge, de er stillet i.*

## Opgave 1 (20 %)

En forsker starter to eksperimentelle populationer af kaniner: Én med individer indsamlet fra en stor skotsk population af vilde kaniner og én indsamlet på en lille ø ud for Skotland, hvor der findes en population, der er blevet sat ud for 50 generationer siden og som har haft en lille populationsstørrelse lige siden. Dens effektive populationstørrelse ( $N_e$ ) har været på 25 gennem alle 50 generationer



I begge populationer er gennemsnitsvægten – som er en kvantitativ karakter – 3 kg, og den fænotypiske varians for vægten,  $V_P$ , er 0,5 i begge populationer. Der selekteres ved kunstig selektion for en øget middelværdi. Dette gøres ved at begrænse avlen til særligt tunge individer med en gennemsnitsvægt på 3,6 kg i begge eksperimentelle populationer. I eksperimentet med kaniner fra den store population havde afkommets middelværdi en vægt på 3,480, medens afkommets middelværdi i eksperimentet med kaniner fra øen var på 3,175 kg.

- a) Hvad er heritabiliteten i de to eksperimentelle populationer?

$$R = h^2 S \Rightarrow h^2 = R/S$$

$$\text{Stor population: } h^2 = 0,480/0,6 = 0,800$$

$$\text{Lille population: } h^2 = 0,175/0,6 = 0,292$$

- b) Hvor stor er den additive genetiske varians,  $V_A$ , i de to populationer?

$$h^2 = V_A/V_P \Rightarrow V_A = V_P \times h^2$$

$$\text{Stor population: } V_A = 0,5 \times 0,800 = 0,400$$

$$\text{Lille population: } V_A = 0,5 \times 0,292 = 0,146$$

- c) Hvorfor er der forskel på heritabiliteten i de to eksperimentelle populationer? Hvordan kan det forklares?

I populationen på øen har populationen tabt gendiversitet pga genetisk drift i de 50 generationer. Det gælder også den additive genetiske varians,  $V_A$ .

Gendiversiteten (og additiv genetisk varians) aftager efter formlen

$$H_t = H_0(1 - 1/(2Ne))^t,$$

så den brøkdel af  $V_A$ , der er tilbage efter 50 generationer er givet ved

$$\begin{aligned} H_t/H_0 &= (1 - 1/(2Ne))^t \\ &= (1 - 1/(2 \times 25))^{50} \\ &= 0,364. \end{aligned}$$

$V_A(\text{lille}) = V_A(\text{stor}) \times 0,364 = 0,4 \times 0,364 = 0,146$  (under antagelse af fravær af genetisk drift i den store population).

- d) I et andet forsøg selekteres der ved kunstig selektion for en øget middelværdi på en anden kvantitativ karakter i den eksperimentelle population, der stammer fra øen. Karakteren

har meget høj varians, men lige meget hvor meget der selekteres, ændres fordelingen af den kvantitative karakter i population sig ikke. Middelværdien og variansen ændres ikke fra generation til generation. Hvad fortæller dette os om heritabilitet og den additive genetiske varians af denne karakter (Kort)

Heritabiliteten og dermed den additive genetiske varians  $V_A$  er nul.

- e) Der selekteres for en anden karakter—og denne gang øges middelværdien af den kvantitative karakter fra 1,5 til 2,4 i løbet af 10 generationer. En tredje karakters middelværdi reduceres samtidig fra 3,2 til 2,1, uden at der har været nogen direkte selektion på denne karakter. Hvordan kan det bedst forklares.

Den indirekte selektion kan skyldes pleiotropi, hvor nogle af de gener, der er under selektion, også har (negativ) indflydelse på den tredje karakter. Koblede gener i LD med en negativ effekt på den tredje karakter kan også spille en rolle.

### Opgave 2 (25%)

En klassisk genetisk polymorfisme i mennesket er evnen til at smage phenylthiocarbamide (PTC). Nogle kan smage PTC, mens andre ikke kan. Hvis du kan smage det, ville du ønske du ikke kunne, når du bliver budt det. Det smager meget bittert. Evnen til at smage PTC er bestemt af et enkelt autosomalt gen med to alleler. Alle ikke-smagere er recessivt homozygote for ikke-smager allelen. Antag, at smager og ikke-smager bestemmes af et autosomalt gen med to alleler.

I en given population, hvor genotyperne optræder i Hardy-Weinberg proportioner, er frekvensen af ikke-smagere 4%.

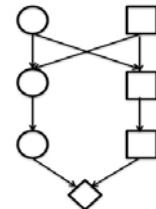
- a) Hvad er frekvensen af ikke-smager allelen? Vis dine udregninger.

$$q = \sqrt{0,04} = 0,2$$

- b) Hvilken andel af alle ægteskaber i populationen er smager  $\times$  ikke-smager (vi antager at ægteskaber er et tilfældigt møde mellem populationens genotyper).

$$2 \times q^2 \times (1 - q^2) = 2 \times 0,04 \times 0,96 = 0,0768$$

Et studie undersøgte fordelingen af smagere og ikke-smagere blandt børn fra fætter-kusine ægteskaber (vist på figur) i populationen.



- c) I hvilken retning vil indavlen påvirke fordelingen af smagere og ikke-smagere i denne gruppe?

Indavl fører til flere homozygoter for begge alleler og færre heterozygoter. Der vil altså komme flere homozygote for ikke-smager allelen dvs. flere med ikke-smager fænotypen blandt afkom af fætter-kusine ægteskaber.

- d) Udregn andelen af ikke-smagere i denne gruppe (dvs. i børn fra fætter-kusine ægteskaber).

$$p(\text{ikke-smager}) = q^2 + Fpq \\ F = 1/16$$

$$p(\text{ikke-smager}) = 0,2^2 + 1/16 \times 0,2 \times 0,8 = 0,05$$

Ronald Fisher – én af grundlæggerne af populationsgenetikken – tænkte i 1930erne over om vores nærmeste slægtning – chimpansen – havde den samme polymorfi. Han besøgte Edinburgh Zoo og tilbød chimpanserne PTC. Konklusionen var meget klar. Han blev spyttet på af den første chimpanse, der blev budt på PTC. Andre chimpanser reagerede ikke. Slutningen var åbenlys: chimpanser og mennesker har begge en polymorfi, hvor de er variable for at kunne smage eller ikke-smage PTC. Han undrede sig. Han mente, at sandsynligheden for at en sådan polymorfi i et locus kan findes hos både chimpanser og mennesket er meget lille på grund af den forholdsvis lange tid siden, hvor de spaltede fra hinanden (5-6 millioner år, men det eksakte tal kendte Fisher ikke). Derfor foreslog han, at polymorfien måtte opretholdes gennem overdominant selektion.



- e) Hos mennesket er den effektive populationsstørrelse ca. 10.000 medens den for chimpansen er ca. 25.000. Benyt dette til at estimere, hvad alderen (i år) er for den sidste fælles forfader for neutrale alleler i et locus hos mennesket og chimpansen. Vi ved, at tiden til most recent common ancestor er givet ved  $t_{MRCA} = 4 \times N_e$  generationer. Antag, at generationstiden er 25 år hos mennesket og 20 år hos chimpansen. Er det sandsynligt, at polymorfien i både menneske og chimpanse vil kunne opretholdes fra deres fælles stamform?

**Vi ved, at den sidste fælles stammor/far til gener er  $4 \times N_e$  generationer tilbage. Det er  $4 \times 10.000$  generationer for mennesket. Dette svarer til én million år med en generationstid på 25 år. Hos chimpansen er det  $4 \times 25.000$  generationer. Dette svarer til to millioner år med en generationstid på 20 år. De to arter spaltede fra hinanden for 5-6 millioner år siden. Derfor er det meget usandsynligt at finde neutrale polymorfier, der er opretholdt hos både chimpansen og mennesket. Fishers intuition var god nok.**

- f) Som nævnt, så er den effektive populationsstørrelse hos mennesket ca. 10.000. Vi er flere end 7 milliarder mennesker. Der er en kæmpe forskel på den aktuelle, ( $N$ ) og den effektive populationsstørrelse,  $N_e$ . Hvilke faktorer reducerer  $N_e$  i forhold til  $N$ ? Hvilken faktor anser du for at være den vigtigste hos mennesket?

**Afvigelser fra ligelig kønsratio. Variation i antal afkom. Swingninger i populationsstørrelsen. Det er nok sidstnævnte, der er vigtigst. Mennesket har været igennem flere flaskehalse og levet i meget små populationsstørrelser da vi splittedes fra den fælles stamform.**

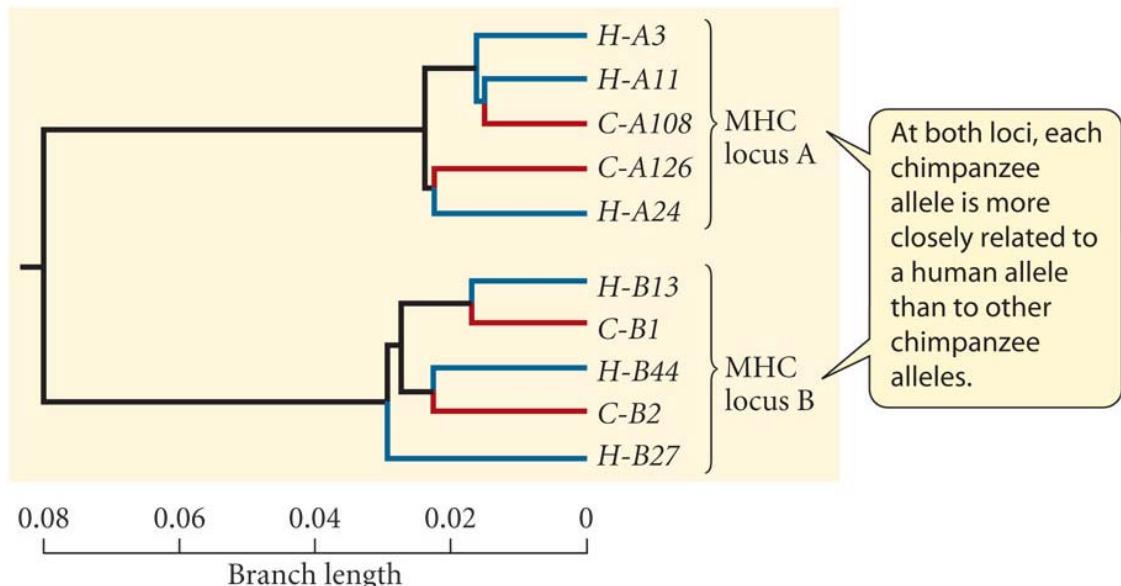
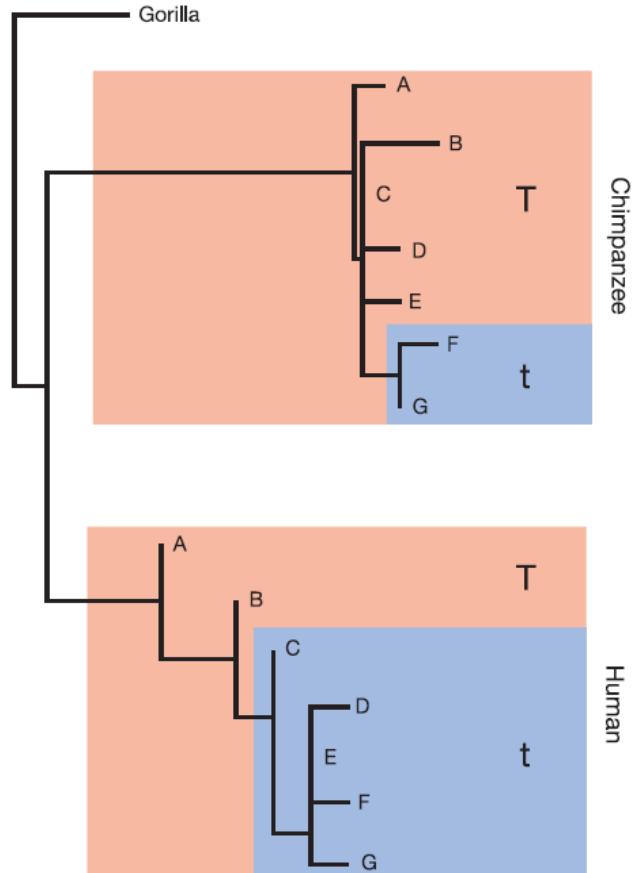
Evnen til at smage PTC vides nu at være kontrolleret af genet TAS2R38. Det viser sig at både smager ( $T$ : taster) og ikke-smager ( $t$ : non-taster) fænotypen kodes af flere forskellige haplotyper. Fylogenien af genet har gorillaen som udgruppe. Bemærk, at A-G dækker over forskellige haplotyper hos mennesket og chimpansen.

- a) Kan man ud fra fylogenien udtales sig om hvorvidt ikke-smager allelen er oprindelig og passer det med Fishers hypotese?

Den er opstået uafhængigt to gange: én gang hos mennesket og én gang hos chimpansen. Det er konvergent evolution. Fishers forklaring holdt ikke.

- b) Betragt et locus, hvor der er balancerende selektion hos chimpansen, mennesket og deres fælles stamform, hvor nogle alleler er ældre end arterne. Hvordan ser et fylogenetisk træ ud for sådan et gen sammenlignet med et artstræ? (Husk: der er flere alleler end de to arter.)

Man vil forvente art nogle af chimpansens alleler er nærmere beslægtet med menneskealleler (og omvendt.) Vi forventer ikke, at de danner to monofyletiske træer, ét for hver art. Se f.eks to MHC loci for mennesket og chimpansen



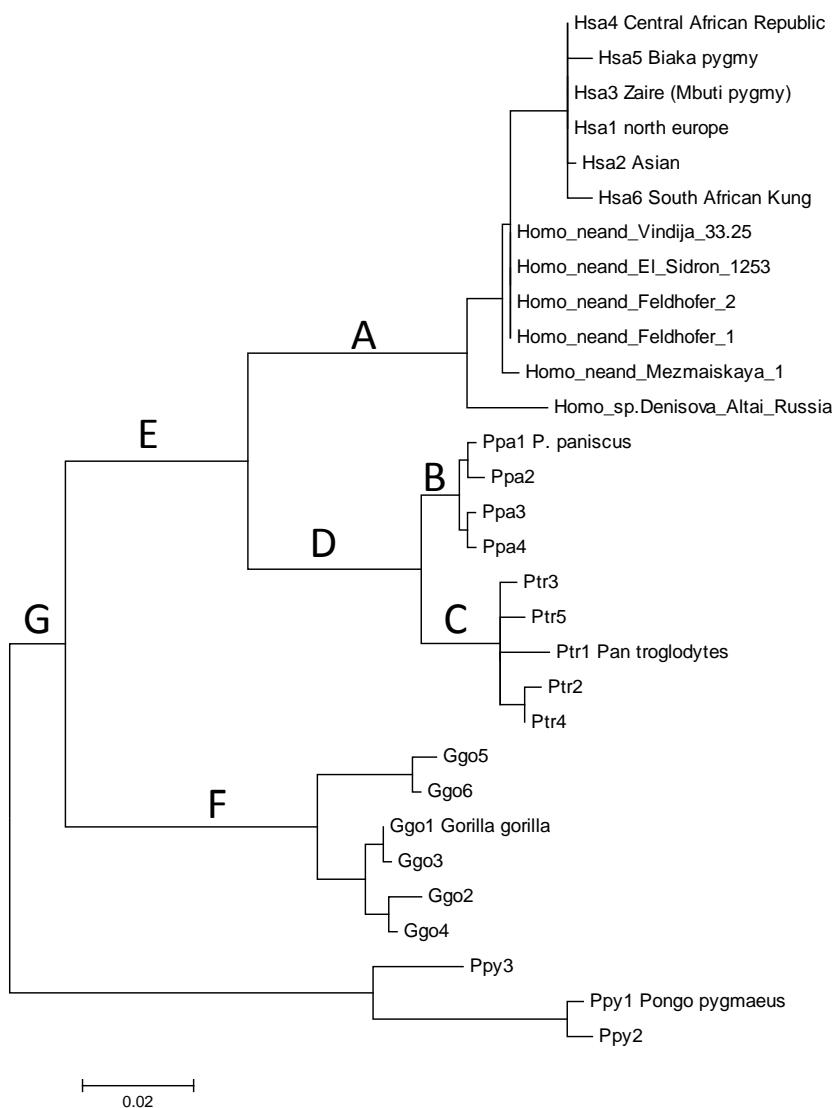
**EVOLUTION 3e, Figure 12.32**

© 2013 Sinauer Associates, Inc.

### Opgave 3 (25 %)

I denne opgave analyserer vi slægtskabet af mennesket og dets nære slægtninge. Vi benytter programmet MEGA og tager som data en række mitokondrielle DNA sekvenser af cytochrome oxydase II (COII) fra både mennesker og menneskeaber. Desuden inkluderes data fra Neanderthalere og fra Denisova-fundet, som er nogle af menneskets nærmeste slægtninge. Data er indeholdt i filen “apes-COII.fas”, som er analyseret i en øvelse i kurset.

Som analysemetode benytter vi Maximum Likelihood, hvor der skal specificeres en substitutionsmodel. Vi vælger Kimura 2-parameter modellen, når vi åbner vinduet “Analysis preferences” i programmet. Vi inkluderer information fra alle tre sites i alle codons. Fylogenien er vist i figuren nedenfor.



Forkortelserne for navnene er:

Hsa: *Homo sapiens* mennesket

Homo\_neand: Neanderthalere

Homo\_sp.Denisova: Denisova-fundet

Ppa: *Pan paniscus* (Bonobo)

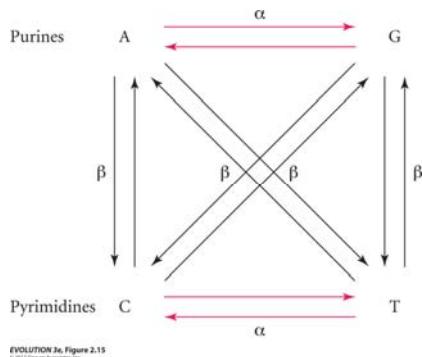
Ptr: *Pan troglodytes* (Almindelig chimpanse)

Ggo: *Gorilla gorilla* (Gorilla)

Ppy: *Pongo pygmaeus* Orangutang

- a) Hvad tages der hensyn til i Kimura 2-parameter modellen?

**Der tages hensyn til, at transitioner og transversioner kan foregå med forskellig rate.**



I MEGA kan vi i "Sequence Data Explorer" under Statistics vælge "Nucleotide pair frequencies". Her får vi oplyst at der observeret 52 transitioner og 6 transversioner i datasættet.

- b) Hvad er transitions/transversions ratioen?

**Den er meget høj (8,7). Der sker langt mere transitioner end transversioner**

- c) Hvis transitioner og transversioner var lige sandsynlige, hvor mange transitioner ville du så forvente at observere for hver observeret transversion?

**Der er dobbelt så mange mulige transversionsmåder som for transitioner. Derfor forventer vi, at der ville observeres en transitions/transversionsrate på 0,5.**

- d) Neighbor joining træer er ofte hurtigere at beregne end parsimony træer og maximum likelihood træer. Hvorfor?

**Ved NJ beregnes kun et træ ud fra en afstandsmatrix, men der ved parsimony og ML beregnes mange træer og der derefter søges efter det korteste eller det der har den højeste likelihood.**

- e) Er mennesket et monofyletisk taxon?

**Ja, alle efterkommere efter den fælles stamform er indeholdt i dette taxon.**

- f) Hvordan vil du karakterisere Neanderthalerne som taxon.

**Det er et parafyletisk taxon. Den indeholder nogle, men ikke alle efterkommere efter stamformen. Neanderthaleren fra Mezmaiskaya er en udgruppe til at alle Neanderthalere samt til mennesket.**

- g) Hvem er nærmest beslægtet med chimpanse og Bonobo: Denisova, Neanderthaler eller mennesket?

**De er alle tre lige tæt beslægtet med dem.**

- h) Denisova, Neanderthaler og mennesket udgør et taxon. Hvem er deres nærmeste slægtning, den almindelige chimpanse eller bonoboen?

**De er begge to lige beslægtede med mennesket**

Tabellen nedenfor angiver hvilke baser der findes i seks positioner i alignmenten over cytochrome oxidase i de fem studerede arter.

	Position					
	42	75	99	300	345	464
Menneske,						
Neanderthaler	C	T	T	T	A	A
og Denisova						
Chimpanse	C	C	C	C	G	A
Bonobo	C	C	C	C	A	G
Gorilla	T	T	C	C	A	A
Orangutang	T	C	A	C	A	A

Antag at baserne i orangutangen er de oprindelige og at andre baser er afledte herfra.

- i) Benyt parsimoni metoden til at angive på hvilke grene (A-G) der er sket hvilken mutation for de seks positioner.

42: E: T->C

75: (A: C->T og F: C->T)  
eller (G:C->T og D: T->C)

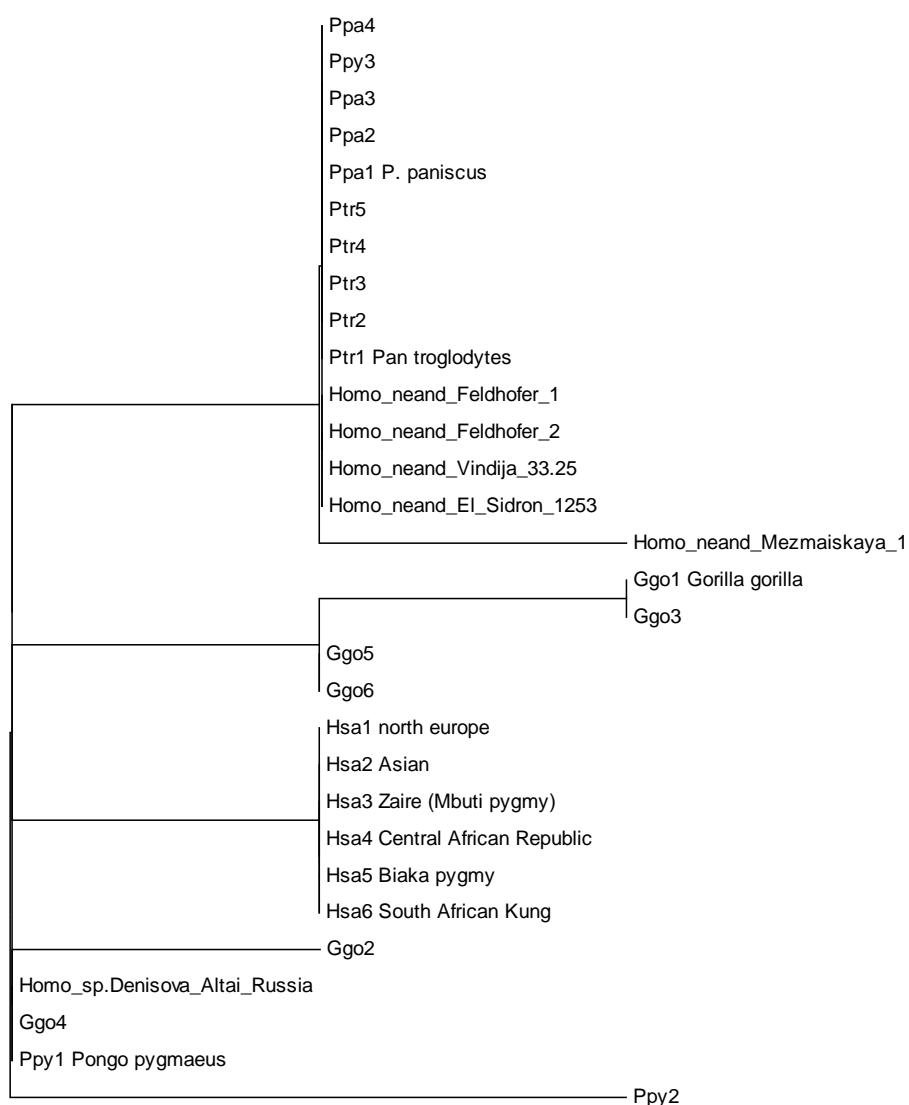
99: G: A->C og A:C->T

300: A: C-> T

345: C: A->G

464: B: A->G

I figuren her har vi beregnet den samme fylogeni med den samme alignment og samme indstillinger som før bortset fra, at vi kun har inkluderet data fra position 2 i alle codons.



- j) Fylogenien ser noget anderledes ud, når vi kun ser på position 2. Giver den mening? Hvad er forklaringen på dette?

Fylogenien giver ikke megen mening. Der er ikke meget der tyder på, at den afspejler noget biologisk realistisk billede af disse arters slægtskab. Forklaringen er, at position 2 af codons i dette gen ikke er særlig variabel, og at der simpelthen ikke er information nok til stede for at estimere en formuftig fylogeni ud fra dette.

#### Opgave 4 (15 %)

En evolutionsbiolog simulerede en population med en størrelse på 200 individer i fire forskellige scenarier (I-IV) med forskellige typer naturlig selektion. Scenarierne var givet ved de følgende respektive fitness-værdier for de tre genotyper i fire autosomale loci:

Genotype			
Locus	AA	AB	BB
I	1,2	1	1,2
II	1	1	1
III	0,8	1	0,8
IV	1,1	1	0,9

Han simulerede 10 populationer for hvert scenarie i 50 generationer.

- a) Hvilken form for naturlig selektion er der tale om i de fire scenarier?  
 I: underdominans  
 II: neutral evolution  
 III: overdominans  
 IV: retningsselektion (for allele A)
- b) Hvor mange ligevægte (stabile og ustabile) findes der for hvert af de fire scenarier?  
 I: 2 stabile (0, 1), 1 ustabil (0,5)  
 II: ingen, allelfrekvensens ligevægt er neutral  
 III: 2 ustabile (0, 1), 1 stabil (0,5)  
 IV: 1 ustabil (0), 1 stabil (1) (allelfrekvensen er for A)
- c) Hvilket locus (I-IV) er illustreret i de følgende fire scenarier (A-D)? (begrundes kort)  
 I: C (Når allelfrekvensen skubbes bort fra 0,5, fortsætter den enten mod 1 eller 0. Når frekvensen er over 0,6 eller under 0,4 bevæger den sig ikke tilbage og krydser 0,5.)  
 II: B (Her divergerer allelfrekvenserne tilfældigt. Det kan gå både op og ned. Allelfrekvenserne spredes over et stadigt bredere interval med tiden.)  
 III: D (Overdominant selektion med en ligevægtsfrekvens på 0,5 holder allelfrekvensen tæt på 0,5.)

IV: A (retningsselektion fikserer alle populationer i  $p = 1$ .)

Bonuspoint: med selektionskoefficienter på 0,1 eller 0,2, som er tilfældet her, og  $N = 200$ ,  
gælder der, at

$$4Ns > 1$$

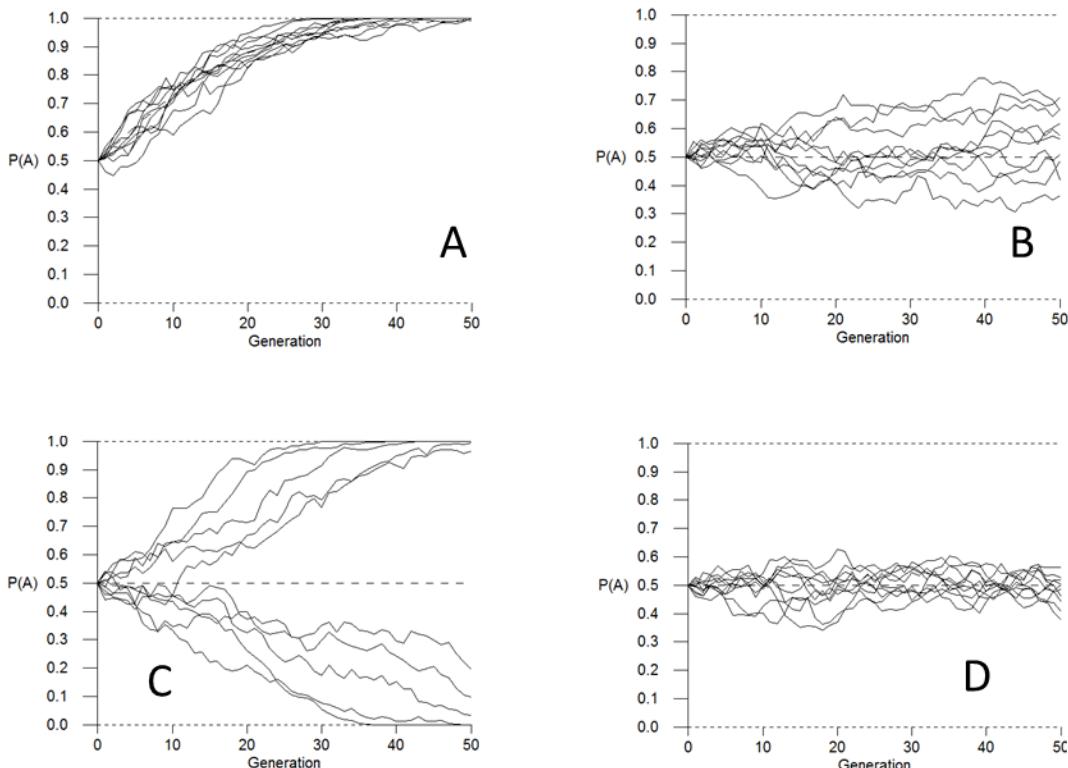
her

$$4 \times 200 \times 0,1 = 80 \gg 1$$

og

$$4 \times 200 \times 0,2 = 160 \gg 1$$

Dvs, at det er naturlig selektion, der afgør udfaldet. Genetisk drift spiller ingen stor rolle.



### Opgave 5 (10 %)

Under sidste istid har en sommerfugleart i Europa fået begrænset sit udbredelsesområde og er blevet trængt tilbage til to refugier, et i Spanien og et på Balkan. Her har populationsstørrelserne været stærkt reducerede. Efter istiden har de spredt sig mod nord og mødes igen.

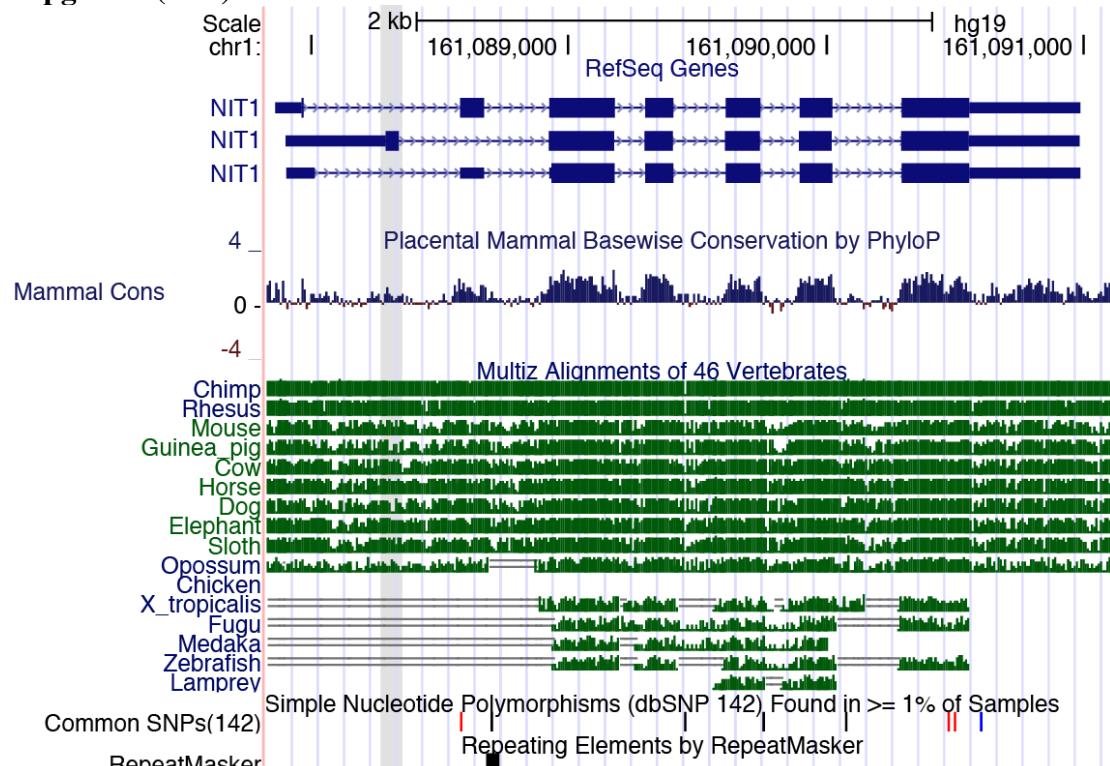
- Skitser **kort** hvad forskellen i udfald kan være, hvis hybriderne har enten *mindre* fitness end de parentale former eller har *samme* fitness som de parentale former.  
Hvis hybriderne har mindre fitness, vil der blive selekteret mod hybridisering, dvs for øget reproduktiv isolation (reinforcement). Den præzygotiske isolation øges. Hvis den reproduktive isolation ikke ender med at blive 100%, kan der ske gen flow for neutrale alleler mellem arterne.  
Hvis hybriderne har samme fitness, vil der ske introgression imellem de to populationer,

og de vil i sidste ende kunne smelte sammen.

Ovenstående scenarie over hybridisering mellem populationer af dyr kan se anderledes ud, når man ser på planter.

- b) Skitser **kort** hvad der kan ske i denne situation.  
**Der kan dannes nye arter gennem polyploidisering.**
- c) Hvor hurtig er denne arts dannelse?  
**Den kan ske over 1-2 generationer.**

#### Opgave 6 (5 %)



Ovenfor vises en region af det humane kromosom 1 i UCSC genom browseren. I denne region transkriberes genet NIT1. Kig på de forskellige annoteringer og svar på nedenstående spørgsmål.

- a) Hvor mange forskellige transkriptions start-sites findes der for dette gen ifølge RefSeq annoteringen?  
**2**
- b) Hvor mange forskellige polyadenylering sites findes der for dette gen ifølge RefSeq annoteringen?  
**1**
- c) Hvor mange forskellige proteiner koder de viste RefSeq mRNA for?  
**3**
- d) Hvilke er de mindst konserverede exons?  
**1 + 2: De findes kun hos pattedyr.**
- e) Hvilke er de mest konserverede exons?  
**5 + 6: De findes også hos lampretten.**

# **Evolutionsbiologi**

Skriftlig eksamen, den 22. juni 2016

**Med alle tilladte hjælpemidler  
(Lærebog, noter, computer)**

For hver opgave oplyses dens vægtning.

*Dette opgavesæt indeholder 9 sider inklusive forsiden  
Opgaverne skal afleveres i den rækkefølge, de er stillet i.*

Jørgen Bundgaard    Jeppe Vinther    Søren Rosendahl    Jesper Givskov Sørensen  
                            Rasmus Heller                         Hans R. Siegismund

## Opgave 1 (15 %)

Den almindelige kjove (*Stercorarius parasiticus*) findes i to faser, en lys og en mørk. (De er illustreret på billederne til højre.) I et studie over nedarvningen af disse to faser observerede man de følgende fænotypeudspaltninger i en række krydsninger af forskellige forælderkombinationer på Fair Isle øen ved Skotland.



Forældre-kombination	Afkom	
	Lys	Mørk
Lys × Lys	29	0
Lys × Mørk	52	86
Mørk × Mørk	25	240
Total	106	326

- a) Hvorledes nedarves de to faser?

Den lyse fase må være recessiv homozygot (LL), medens den mørke fase er enten heterozygot (LM) eller homozygot (MM) for den mørke allele.

- b) Bestem frekvensen af allelen for den lyse fase. (Tag udgangspunkt i udspaltningen af det totale antal afkom.)

$$q = \sqrt{106/432} = 0,495$$

Vi antager, at der er Hardy-Weinberg proportioner i populationen.

En molekulær genetisk analyse har vist, at den mørke fase skyldes en aminosyresubstitution i MC1R (melanocortin-1 receptor) genet. Allelen for den mørke fase er den afledte. En sekvensanalyse af haplotyper, der bærer variationen for mørk og lys fase, viste at nukleotiddiversiteten ( $\pi$ ) var henholdsvis 0,021 og 0,094 (i procent). Diversiteten er signifikant lavere for haplotyper, der bærer variationen for den mørke fase.

- c) Forklar hvorfor der observeres lavere diversitet for haplotyper, der bærer variationen for den mørke fase?

Den mørke fase må være blevet selekteret for, siden den opstod. Gennem et selektivt sweep er variationen, der er i LD (linkage disequilibrium) med den, også øget i frekvens og har fået reduceret diversiteten.

- d) Hvor høj forventes nukleotiddiversiteten langs kromosomet at være som funktion af afstanden fra den aminosyre, der koder for den mørke fase?

Vi forventer at nukleotiddiversiteten vil stige med større afstand, da rekombinationen nedbryder LD til den selekterede aminosyre.

### Opgave 2 (25 %)

Fie er forsker og arbejder med at fremavle kalkuner med hurtigere vækst. Den egenskab Fie selekterer på er, hvor hurtigt kalkunerne opnår en slagtevægt på 6 kg. I gennemsnit tager det almindelige kalkuner 12 uger at opnå en vægt på 6 kg. Heritabiliteten i snæver forstand,  $h^2$ , for dette træk er 0,4. Fie vil etablere en avlsstamme for hurtigere vækst. Hun udvælger en gruppe kalkuner, som i gennemsnit vejer 6 kg efter kun 11 uger, og deres afkom udgør første afkomsgeneration.

- a) Hvor stort er selektionsdifferentialet?

$$S = 11 \text{ uger} - 12 \text{ uger} = -1 \text{ uge}$$

- b) Hvor mange uger vil det i gennemsnit tage for den første afkomsgeneration, at opnå en vægt på 6 kg?

Respons:  $R = S \times h^2 = -1 \times 0,4 = -0,4$

Middelværdi af afkomsgenerationen:  $12 - 0,4 = 11,6$

Efter 8 generationers selektion lykkes det Fie at nå sit mål, nemlig at have kalkuner, der når en slagtevægt på 6 kg efter 10 uger. Responsen i den 8. generation er mindre end responsen i den første generation med selektion. Hun har benyttet det samme selektionsdifferentiale i hver generation.

- c) Hvorfor aftager responsen i løbet af et selektionsexperiment?

Under kunstig (og naturlig) selektion aftager den additive genetiske varians  $V_A$ . Derfor falder heritabiliteten og den samlede respons bliver således mindre end 8 gange responsen i første generation.

I mange selektionsforsøg på kvantitative karakterer når man et plateau, hvor man ikke længere ser en respons, lige meget hvor stærkt man selekterer på en karakter.

- d) Betyder det, at der ikke længere er genetisk variation der påvirker karakteren i populationen?

Nej, det betyder kun, at den additive genetisk varians,  $V_A$ , må være 0. Der kan stadig være variation for karakteren.

Fie har brugt en del SNP-markører i sit avlsarbejde. Når hun sammenligner nukleotiddiversiteten i udgangspopulationen med nukleotiddiversiteten efter 8 generationer, opdager hun at diversiteten er faldet til det halve.

- e) Ud fra det slutter hun, at den effektive populationstørrelse i avlspopulationen har været ca. 6 under selektionsexperimentet. Hvordan kommer hun til det resultat?

$$H_t = H_0(1 - 1/(2N_e))^t$$

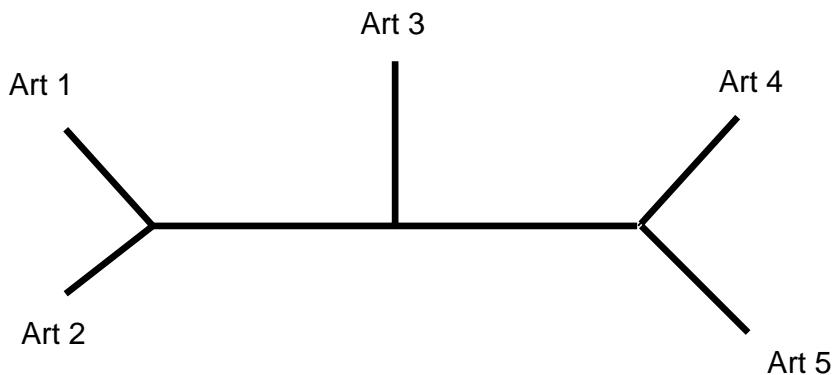
$$H_t/H_0 = 0,5 \approx (1 - 1/(2 \times 6))^8$$

- f) Fie mener, at det er vigtigt at bevare så meget af den genetiske variation som muligt i avlsstammen. Hvilke(n) faktor(er) vil påvirke tabet af genetiske variation i avlsstammen og hvordan kan Fie modvirke dette?

Det er primært genetisk drift der vil påvirke den genetiske variation i avlsstammen. Tabet af variation reduceres bedst ved at have så stor population som muligt i hver generation (min. genetisk drift)+ lige mange haner som høner (max  $N_e$ ) + lige meget afkom pr høne (max  $N_e$ ).

### Opgave 3 (20%)

Betrægt den følgende fylogeni:



- a) Hvad er parsimoniscoren udregnet på dette træ for følgende sekvenser:

	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
Art 1	T	C	T	C	A	G	A	C	A	G	T	G	C	G	G	A
Art 2	A	C	T	C	A	G	T	G	G	A	T	G	G	A	C	C
Art 3	A	T	C	A	G	A	C	G	A	T	G	G	A	C	G	C
Art 4	A	T	T	C	A	G	T	G	G	C	A	G	T	G	G	C
Art 5	A	T	T	C	A	G	T	G	G	C	A	G	T	G	G	C

7

- b) Hvilke monofyletiske grupper findes i dette træ, hvis roden er placeret på den gren, der leder til art 5? (Medregn ikke grupper, der kun består af én art eller af alle 5 arter)?

(1, 2), (1, 2, 3), (1, 2, 3, 4).

- c) Hvor mange transitioner og transversioner adskiller sekvenserne fra art 1 og art 2?

2 transversioner, 1 transition

- d) Hvis rodens i stedet var placeret på den gren som leder til art 3, hvilke monofyletiske grupper ville da kunne findes i træet?

(1, 2), (4, 5), (1, 2, 4, 5).

- e) Tyder substitutionsmønstret på, at dette gen er et almindeligt proteinkodende gen?  
(Hjælp: se på antallet af substitutioner i position 1, 2 og 3.)

Nej, der er ingen substitutioner på position 3 og flere på position 2 end på position 1. Det er det stik modsatte af, hvad vi forventer, hvor position 3 har den højeste rate efterfulgt af position 1 og så position 2.

- f) Hvilke to parametre er de vigtigste i den neutrale teori?

Den neutrale mutationsrate,  $u_0$ , og populationsstørrelsen,  $N$ .

- g) Hvilke to væsentlige forudsigelser indeholdes i den neutrale teori?

Den forudsiger et molekylært ur, hvor substitutionsraten er lig med den neutrale mutationsrate. Desuden forudsiger den graden af diversitet i en population som funktion af 4 gange produktet af den neutrale mutationsrate og populationsstørrelsen.

#### Opgave 4 (20%)

Husmus (*Mus domesticus*) forekommer i mange kromosomale racer i Europa. I en undersøgelse af to parapatrisk udbredte racer i det nordlige Italien lavedes de følgende krydsninger mellem forskellige dyr. Den kromosomale genotype af dyr fra den ene race betegnes som AA, medens den kromosomale genotype for den anden race betegnes som BB, dvs. at de er homozygote for den samme kromosomale genotype. I forskellige krydsninger mellem individer samt F<sub>1</sub> (AB) fra de to racer observerede man de følgende kuldstørrelser:



Krydsning	Kuldstørrelse
AA × AA	6,8 ± 0,8
BB × BB	6,6 ± 0,7

AB × BB	2,3 ± 0,6
AA × AB	1,9 ± 0,7

- a) Kuldstørrelsen kan betragtes som en vigtig fitness-komponent. Hvordan er fitness af de tre kromosomale genotyper?

Fitness af heterozygoter er mindre end fitness for de to homozygoter, der cirka er ens.

- b) Hvilken form for naturlig selektion er der tale om?

Underdominant selektion.

- c) Hvor mange ligevægte findes der for sådan et system? Hvor mange er stabile, hvor mange er ustabile?

3, 2 stabile i 0 og 1, samt en indre ustabil ligevægt.

- d) Hvis et B-kromosomsæt introduceres i en meget stor population, der er homozygot AA, hvad vil der så mest sandsynligt ske med det nye B-kromosomsæt?

Det ville blive elimineret.

- e) Under hvilke forhold vil man kunne forestille sig, at B-kromosomsættet bliver introduceret i en AA population og bliver fikseret?

Hvis populationsstørrelsen bliver meget lille, kan genetisk drift øge hyppigheden af B og i sidste ende – sammen med naturlig selektion – fiksere den i populationen.

- f) Ovenstående naturlige selektion sammen med den parapatriske udbredelse af kromosomracer kan indgå som komponent i arts dannelse. Beskriv hvordan dannelsen af nye arter på denne baggrund kan foregå.

Parapatrisk arts dannelse: Når hybrider mellem kromosomracerne har nedsat fitness vil der ske selektion for øget isolation mellem dem (reinforcement).

- g) Hos planter indgår en anden form for kromosommutation ofte i arts dannelsen. Beskriv denne **kort**.

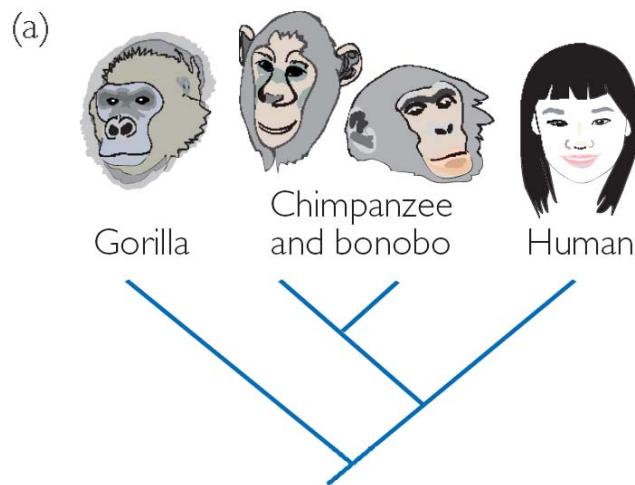
Polyploidisering

- h) Arts dannelse, som nævnt under g) forvolder komplikationer i fylogenetiske analyser. Hvordan?

De fleste fylogenetiske analyser antager forgreninger og ikke retikulær evolution.

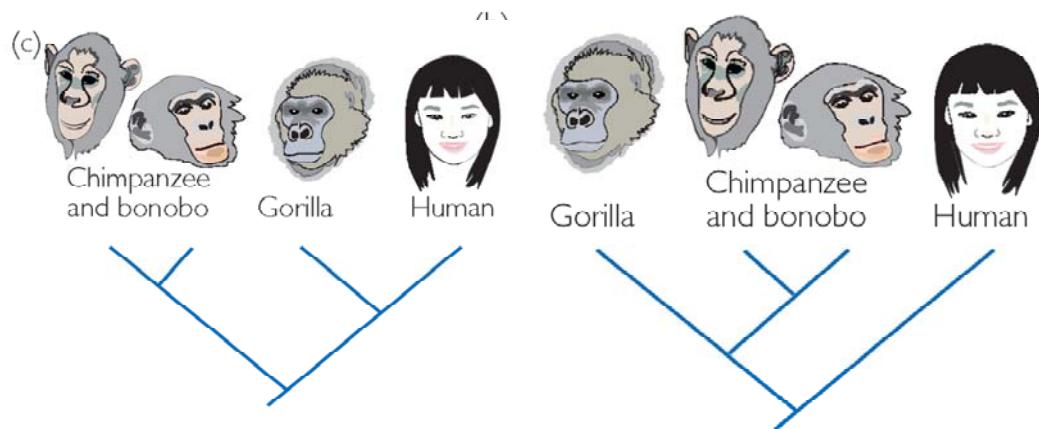
### Opgave 5 (20%)

Når man laver fylogenier baseret på hele genomer hos mennesket, chimpansen og gorillaen finder man at kun i 70% af genomet er gentræer identiske med artstræet, som er illustreret nedenfor.



I 15 % af genomet er gorillaen en søstergruppe til mennesket og i 15% af genomet er mennesket en udgruppe til de øvrige.

- a) Tegn disse træer

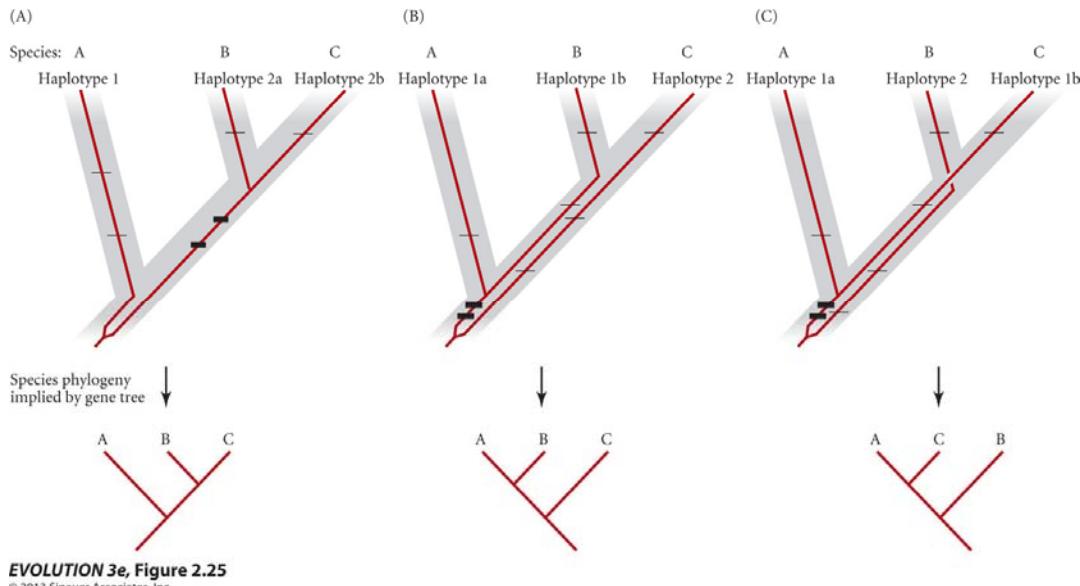


- b) Hvem er menneskets nærmeste slægtning i artstræet, chimpansen eller bonoboen?

Det er de begge.

- c) Hvor er årsagen til at hele 30% af de nævnte arters genom har gentræer, der afviger fra artstræet?

Incomplete lineage sorting. Når de fælles stamfædre har været polymorfe for forskellige haplotyper, kan det resultere i at gentræet bliver forskelligt fra artstræet.



- d) Hvilke forhold hos den fælles forbader til chimpansen (samt bonoboen) og mennesket kan have haft indflydelse på den afvigelse mellem gentræer og artstræer?

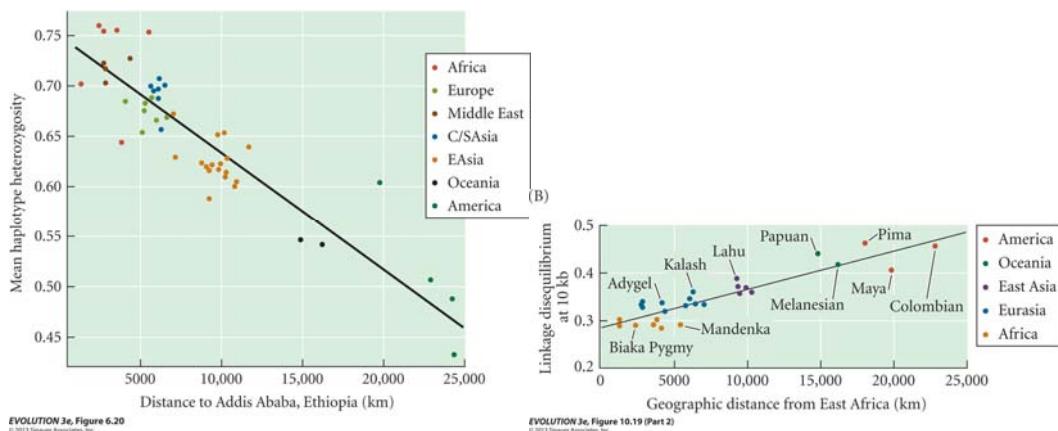
Hvis den fælles stamform for chimpanse (samt bonobo) og mennesket har haft en stor effektiv populationsstørrelse er sandsynligheden for incomplete lineage sorting stor. Det samme gælder hvis tiden fra splittet mellem den fælles stamform for gorillaen og menneske/chimpanse linjen har været kort

- e) Når vi taler om, at 30% af det humane genom har gentræer, der afviger fra artstræet, betyder det så, at 30% af de humane kodende gener har en anden fylogeni end arten har?

Nej, det er 30% af genomet i stykker af varierende længde. Det er ikke de enkelte gener, der følger artstræet eller ej.

- f) Det moderne menneske forlod Afrika for ca 50.000 år siden. Hvordan har koloniseringen af Asien, Europa, Australien og sidst Amerika efterladt spor i genomet hos populationer i disse kontinenter? (Hjælp: Hvilke mål ændrer sig som funktion af afstanden fra Afrika?)

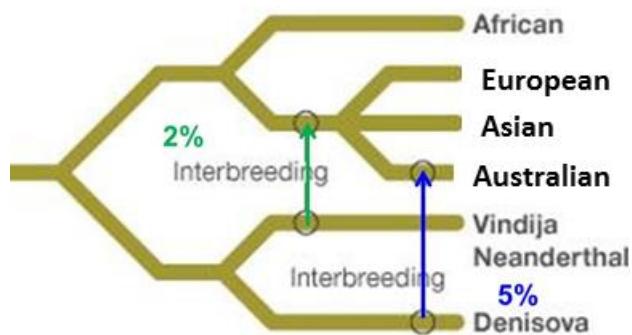
Som følge af gentagne founder effects falder gendiversiteten og koblingsuligevægten (LD, linkage disequilibrium) stiger som funktion af afstanden fra Afrika.



- g) På vejen bort fra Afrika mødte det moderne menneske to andre hominider (Neandertaler og Denisova). Hvordan har disse møder efterladt sig spor i genomet hos populationer der i dag findes i Asien, Europa og Australien?

Der er først sket en hybridisering med Neandertalere, så alle ikke-afrikanske mennesker indeholder i gennemsnit 2% Neandertalgenom i deres genom. De mennesker, der endte i Australien (og omliggende øer) mødte desuden Denisova, med hvilken de hybridiserede og modtog 5% Denisova genom.

### Human family tree



Source: Nature

# **Evolutionsbiologi**

Skriftlig eksamen, den 21. juni 2017

**Med alle tilladte hjælpemidler  
(Lærebog, noter, computer)**

For hver opgave oplyses dens vægtning.

*Dette opgavesæt indeholder **9** sider inklusive forsiden  
Opgaverne skal afleveres i den rækkefølge, de er stillet i.*

Jørgen Bundgaard      Jeppe Vinther      Jesper Givskov Sørensen  
Rasmus Heller            Hans R. Siegismund

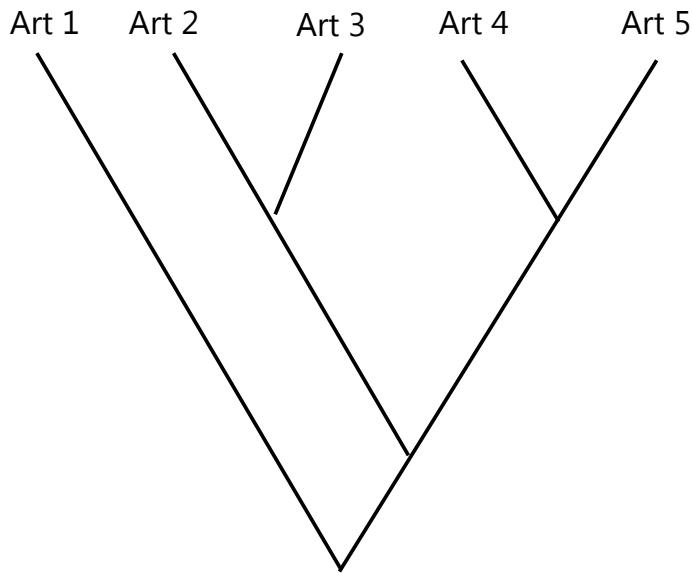
### Opgave 1 (25 %)

Betræt den følgende alignment af 40 basepar fra sekvensen af et ikke-kodende gen fra fem beslægtede arter. En prik i en arts sekvens betyder, at basen er identisk med art 1's sekvens.

Art_1	actaacaaac	cgacaatgct	actgaaaatag	agctctcaga
Art_2	ct.....g.t	....t...c..	.....	....gc....
Art_3	c.....g.t	....t...c..	.....	....gc....
Art_4	c...g..g.t	.....	....g...a	....c.gc.
Art_5	c...g..g.t	.....	....g...a	....c.gc.

Art 1 kan betragtes som udgruppe til de øvrige.

- a) Benyt parsimony princippet til at finde det mest parsimoniske træ for de fem arter.



- b) Hvad er parsimoniscoren udregnet på dette træ?

13

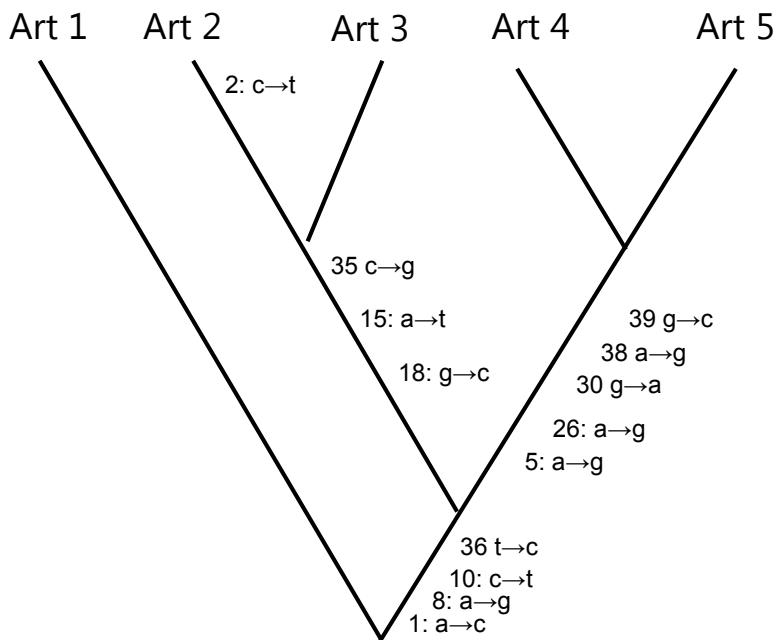
- c) Hvor mange monofyletiske grupper findes i dette træ (Medregn ikke grupper, der kun består af én art eller af alle arter)? Skriv grupperne op ved hjælp af parenteser.

3: (2,3), (4,5), ((2,3),((4,5))

- d) Hvordan betegnes en gruppe, der består af arterne 2, 4 og 5?

Parafyletisk

- e) Angiv i det fylogenetiske træ, hvilke substitutioner der er opstået hvorhenne.



- f) Hvordan er substitutionsraten i et gen relateret til den neutrale mutationsrate af genet?

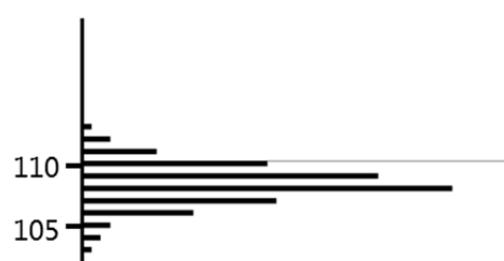
**Substitutionsraten er lig med den neutrale mutationsrate for et givet gen.**

- g) Har den effektive populationsstørrelse  $N_e$  indflydelse på substitutionsraten af neutral variation?

**Nej, den er uafhængig af populationsstørrelsen.**

### Opgave 2 (25 %)

Johannes Schmidt studerede morfologisk variation hos ålevabben (*Zoarces viviparus*) i forskellige populationer i Danmark. Ålevabben er en levendefødende fisk. Det benyttede Schmidt sig af til at estimere heritabiliteten for forskellige karakterer. Blandt andet talte han antallet af hvirvler. I en population fra bunden af Mariager Fjord fandt han den følgende fordeling over antal hvirvler



Middelværdien viste sig at være 108 i denne population. Ved at undersøge sammenhængen mellem hvirvler hos hun-fisk og antallet af hvirvler hos afkom, kunne Schmidt estimere narrow sense heritabiliteten for denne karakter til at være  $h^2 = 0,8$ .

- a) Hvordan kan man estimere narrow sense heritabiliteten  $h^2$  når man har målt fænotypiske karakterer hos forældre og afkom?

I kurset Evolutionsbiologi er nævnt de følgende to måder: Heritabiliteten kan beregnes ved regressioner mellem beslægtede individer. Det kan vises, at heritabiliteten er lig med regressionskoefficienten  $b$  mellem afkommets værdi og forældrenes middelværdi for den observerede karakter. Laves en regression mellem afkom og den ene forælder, kan det vises, at heritabiliteten er lig med to gange regressionskoefficienten  $b$  mellem afkommets og forælderens værdi for den observerede karakter.

- b) Hvilken komponent af den fænotypiske varians måler man med narrow sense heritabiliteten  $h^2$ ?

Det er den brøkdel den additive genetiske varians,  $V_A$ , udgør af den totale fænotypiske varians,  $V_P$ , det vil sige  $V_A/V_P$ .

Antag, at man udvælger fisk med et højere antal hvirvler, for eksempel alle med et højere antal end 110. Denne gruppe af fisk har en gennemsnitsværdi for antallet af hvirvler på 111. Disse forplanter sig indbyrdes.

- c) Hvad vil afkommets gennemsnitlige antal hvirvler være?

Selektionsdifferentialet er  $S = 3$

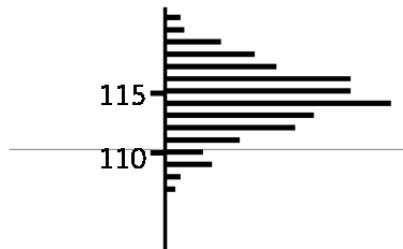
Responsen er givet ved

$$R = Sh^2 = 3 \times 0,8 = 2,4$$

Afkommets middelværdi bliver så

$$110,4 (= 108 + 2,4)$$

Schmidt undersøgte også en anden population ude ved mundingen af Mariager Fjord. Her fandt han den følgende fordeling af hvirvler.



Her var det gennemsnitlige antal hvirvler 114. Denne population havde samme heritabilitet for denne karakter som populationen i bunden af Mariager Fjord, nemlig 0,8.

- d) Hvordan ville du undersøge, om forskellen i antallet af hvirvler mellem populationerne fra bunden og mundingen af Mariager Fjord er genetisk betinget?

Der kan laves et common garden eksperiment, hvor man lader dyr fra hver population i vokse op i deres eget og i den anden populations miljø. Hvis det er genetisk betinget ville man forvente, at middelværdien af en bestemt population er den samme i begge de testede miljøer. Dette blev også gjort af Johannes Schmidt. Forskellen var genetisk betinget.

Antag, at forskellen imellem hvirvelantallet hos populationer i den indre del af fjorden og munden er genetisk betinget.

- e) Hvad fortæller det så om mængden af migration mellem de to populationer?

Migrationen må være begrænset. Ellers ville de to populationer blandes op og få de samme værdier for antal hvirvler.

Johannes Schmidt kortlagde også ålens vandring. Han fandt ud af, at alle europæiske ål svømmer til Sargasso-havet, når de skal yngle. Her gyder de sæd og æg frit i vandmassen. Man må antage, at det sker ved tilfældig parring uanset, hvor ålene kommer fra i Europa.



- f) Når man benytter genetiske markører til at beskrive populationsstrukturen af de europæiske ålepopulationer, hvilken værdi får man så, når man kvantificerer den med  $F_{ST}$ ?

Vi forventer at  $F_{ST} = 0$ , da der er tilfældig parring mellem alle europæiske ål.

Betrægt neutrale genetiske markører i en samling af populationer inden for en art. Populationerne antages alle at have den samme størrelse,  $N_e$ , og at udveksle gener mellem hinanden.

- a) Hvilke faktorer spiller en rolle for størrelsen af differentieringen mellem populationerne, målt som  $F_{ST}$ ? (Vi bruger her Wright's  $\phi$ -model som tilnærmede/antagelse.)

Det er den effektive populationsstørrelse,  $N_e$ , og migrationsraten  $m$  mellem populationerne. Genetisk drift differentierer med raten  $1/(2N_e)$ , medens migrationen homogeniserer med raten  $m$ . Under ligevægt mellem genetisk drift og migration, kan  $F_{ST}$  findes som

$$F_{ST} = \frac{1}{1 + 4N_e m}$$

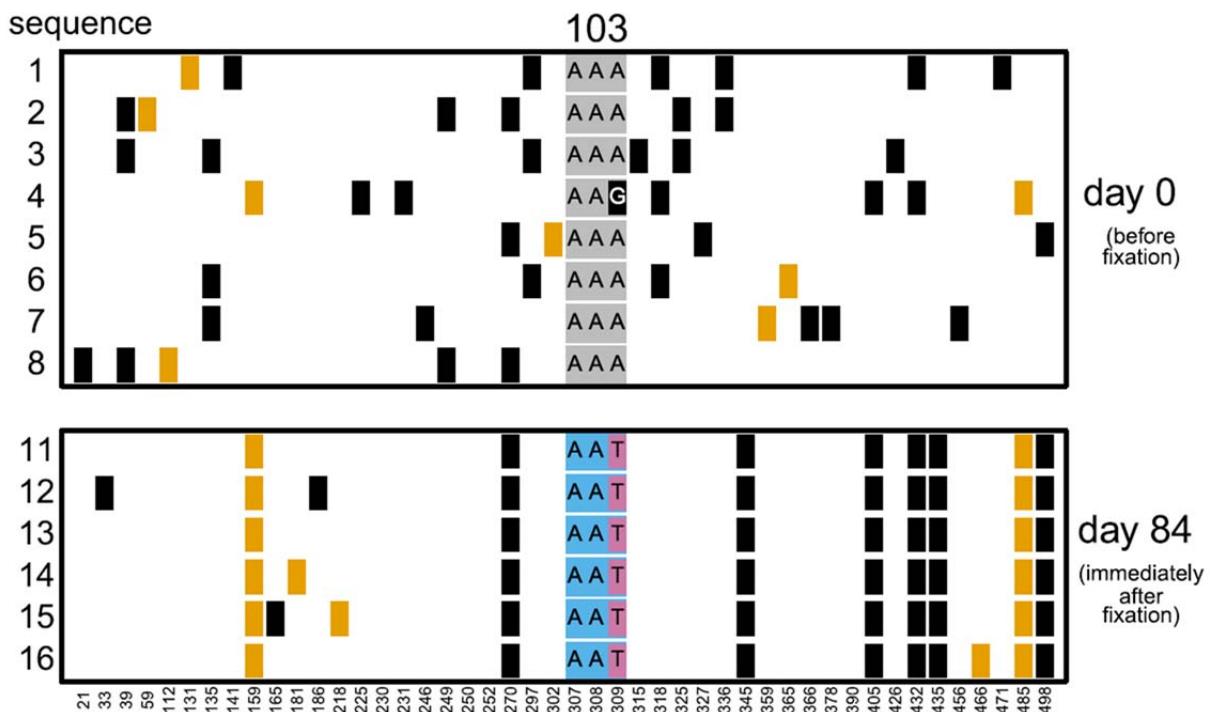
Det er således produktet  $N_e m$ , der er afgørende for differentieringen.

### Opgave 3 (30%)

HIV virus kan nu til dags behandles med en række forskellige medikamenter, der er i stand til at holde virus populationen i en patient nede på et niveau, hvor patienten ikke bliver syg. Det sker ofte, at HIV virus udvikler resistens overfor et bestemt medikament. Den har en meget høj mutationsrate. Nedenstående figur illustrerer evolutionen i virus populationen i en enkelt patient over tid. Hver række i figuren



viser sekvensen af polymorfe sites i en enkelt virus sekvens af patienten. (Positionerne findes nederst i figuren.) De øverste 8 viser sekvenser, der blev taget fra patientens blod før medikamentet blev givet til patienten, medens de nederste 6 sekvenser viser sekvenser på dag 84 efter behandlingen med medikamentet blev påbegyndt. Desuden vises codon 103 (baseposition 307-309). AAA og AAG koder for lysin, medens AAT koder for asparagin. Asparagin i codon 103 medfører resistens overfor det nye medikament. Sorte firkanter angiver polymorfe synomime sites, medens orange firkanter angiver ikke-synomime sites.



- a) Hvor mange synomime og ikke synomime polymorfe sites findes i den øverste figur (sekvens 1 til 8).

Synomime: 24 (Husk at tælle position 309 med)

Ikke-synomime: 8

- b) Hvilken form for naturlig selektion synes at have foregået før dag 0, når man fokuserer på forekomsten af synomime og ikke-synomime sites på dette tidspunkt?

Negativ selektion, der fjerner skadelige mutationer.

- c) Hvilken evolutionær proces illustreres med figuren? (Når man sammenligner dag 0 og dag 84.)

Selective sweep.

- d) Beskriv **kort** hvad der sker under denne proces.

Der er sket en mutation i codon 103 (en transversion af A til T), som så er blevet fikseret i viruspopulationen. Naturlig selektion har fikseret den resistente mutation.

- e) Hvordan ændres nukleotiddiversiteten sig under denne proces?

Den aftager på grund af hitchhiking, da andre sites koblet til det selekterede locus også bliver fikseret. Kan ses på figuren. Den nederste figur indeholder næsten ingen polymorfe sites.

- f) Hvordan påvirkes LD (linkage disequilibrium, koblingsulige vægt) under denne proces?

LD stiger.

- g) Hvorfor findes der polymorfe sites i den nederste figur? Hvor kommer de fra?

Det er formentlig nye mutationer, der er blevet introduceret i populationen.

- h) Hvad er antallet af synonome og ikke synonome polymorfe sites findes i den nederste figur (sekvens 11 til 16)?

Synonyme: 3

Ikke-synonyme: 3

- i) Hvorfor er forholdet nu anderledes end det, som fundet i spørgsmål a?

Der er forholdsvis flere ikke-synonome polymorfier, som negativ naturlig selektion ikke har fjernet endnu.

#### Opgave 4 (15%)

Betræt de følgende arter indenfor korsblomstfamilien (*Brassicaceae*). De tilhører alle slægten *Brassica*. Tre af arterne er diploide (deres kromosomtal er nævnt i parentes efter deres navn)

*Brassica nigra* (sort sennep) ( $2n = 16$ )

*Brassica oleracea* (havekål) ( $2n = 18$ )

*Brassica rapa* ( $2n = 20$ )



Slægten indeholder desuden (ud over en del andre arter) de følgende arter (deres kromosomtal er nævnt i parentes efter deres navn)

*Brassica oleracea* (havekål)

*Brassica carinata* ( $2n = 34$ )

*Brassica juncea* (sareptasennep) ( $2n = 36$ )

*Brassica napus* (raps) ( $2n = 38$ )

- a) Hvilken form for artsdannelse er involveret i dannelsen af de sidstnævnte tre arter?

Allopolyploid artsdannelse.

- b) Hvilken form for artsdannelse er der tale om i de sidste tre arter, når man ser på det i et geografisk perspektiv?

Sympatisk artsdannelse

- c) Hvad anser man for at være den mest almindelige artsdannelse hos dyr, når man ser på det i et geografisk perspektiv?

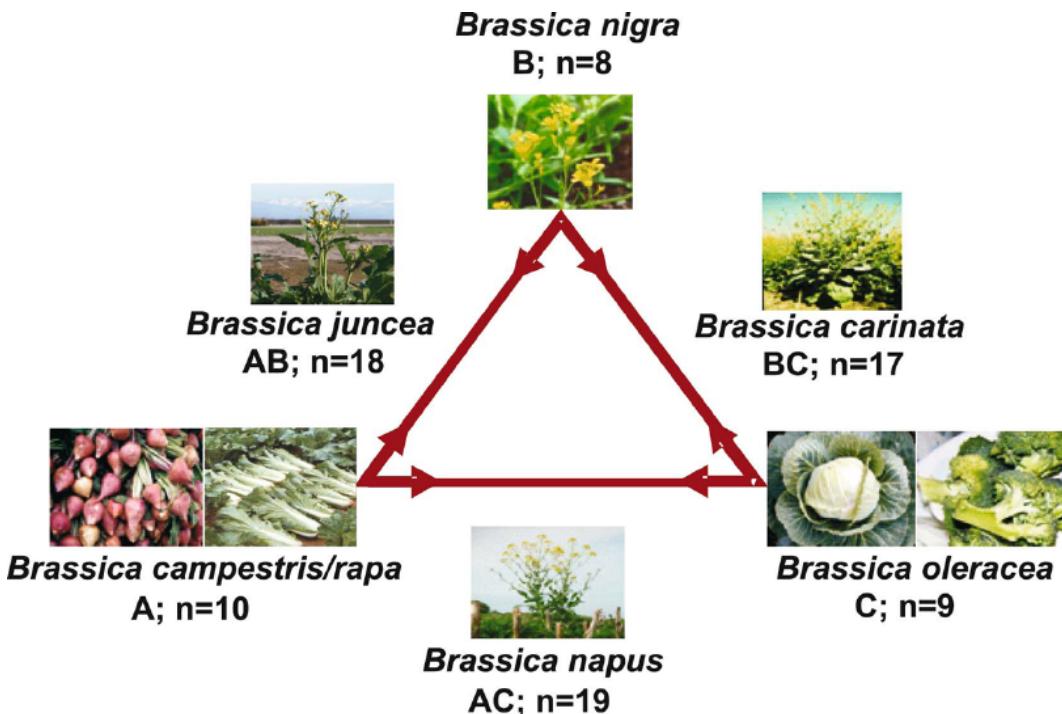
Allopatrik artsdannelse

- d) Hvordan er de tre sidstnævnte *Brassica* arter relateret til de tre førstnævnte i slægten?

*B. carinata* er opstået som en allopolyploid mellem *Brassica nigra* og *Brassica oleracea*.

*B. juncea* er opstået som en allopolyploid mellem *Brassica nigra* og *Brassica rapa*.

*B. napus* er opstået som en allopolyploid mellem *Brassica rapa* og *Brassica oleracea*.



## Opgave 5 (5%)

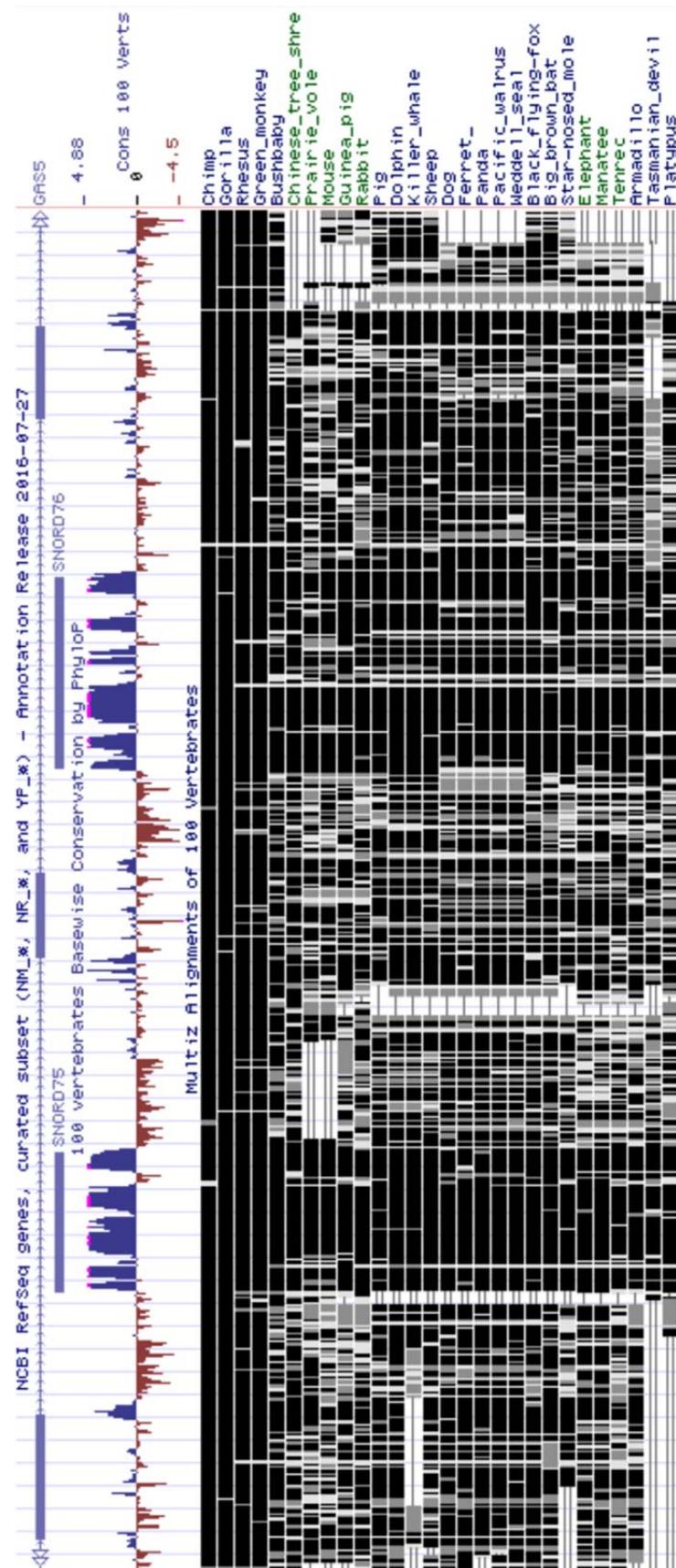
Dette UCSC screenshot viser en del (578 bp) af det humane genom, samt konserveringen til en række andre organismer. I området finder vi en del af GAS5 genet, samt to ikke-protein kodende gener. Annoteringen af generne er vist i NCBI RefSeq Gen annoteringstræk øverst i figuren.

- a) Kommenter på konserveringen af de forskellige regioner af de viste gener.

Figuren viser tre exons samt to introns samt en del af et intron i GAS5 genet. Ud fra rækken med "Cons 100 Verts" samt fra den taxonomiske oversigt over konserveringen, ses at disse exons er mindre konserverede end de to SNORD gener, der befinner sig i introns. Det ses endvidere, at den flankerende sekvens af GAS5 exons er mere konserveret, sikkert fordi der er selektion på de sekvenser, der kontrollerer splicing.

- b) Tror du GAS5 er et vigtigt protein kodende gen?  
Hvorfor/hvorfor ikke?

GAS5 viser ikke en konservering, som er typisk for et proteinkodende gen. Ud fra rækken med "Cons 100 Verts" ses, at exons ikke er konserverede. Det kan derfor ikke være et vigtigt proteinkodende gen. (GAS5 er faktisk et ikke-protein kodende gen med snoRNA gener placeret i introns [inside-out gen]).



## **Evolutionsbiologi**

Skriftlig eksamen, den 20. juni 2018

**Med alle tilladte hjælpemidler  
(Lærebog, noter, computer)**

For hver opgave oplyses dens vægtning.

*Dette opgavesæt indeholder **10** sider inklusive forsiden  
Opgaverne skal afleveres i den rækkefølge, de er stillet i.*

Censorer: Tove Hedegaard Jørgensen      Jesper Givskov Sørensen  
Eksaminatorer:      Rasmus Heller      Hans R. Siegismund

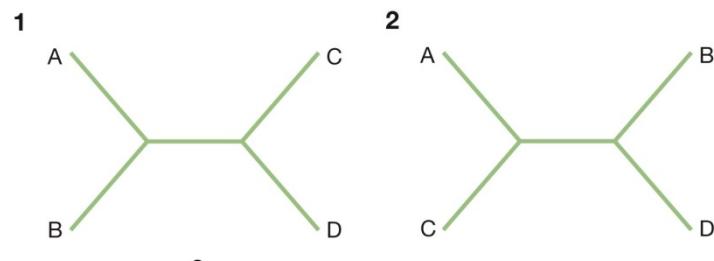
### Opgave 1 (14 %)

En undersøgelse af DNA-sekvenser fra fire arter resulterede i den følgende tabel over afstande imellem dem:

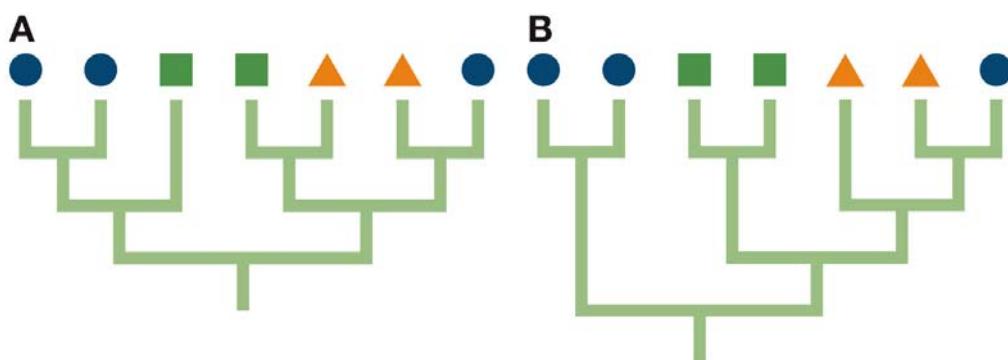
	Art A	Art B	Art C	Art D
Art A	0.18	0.07	0.17	
Art B		0.16	0.05	
Art C				0.18
Art D				

- a) Hvilket af de afbildede træer er det mest sandsynlige for den givne afstandstabel? Begrundes **kort**.  
 (Man behøver ikke regne for at give svaret.)

Det må være træ nummer 2. Vi ser fra tabellen, at parrene (A, C) og (B, D) har de laveste afstande, og at afstandene mellem de to grupper stort set er den samme.



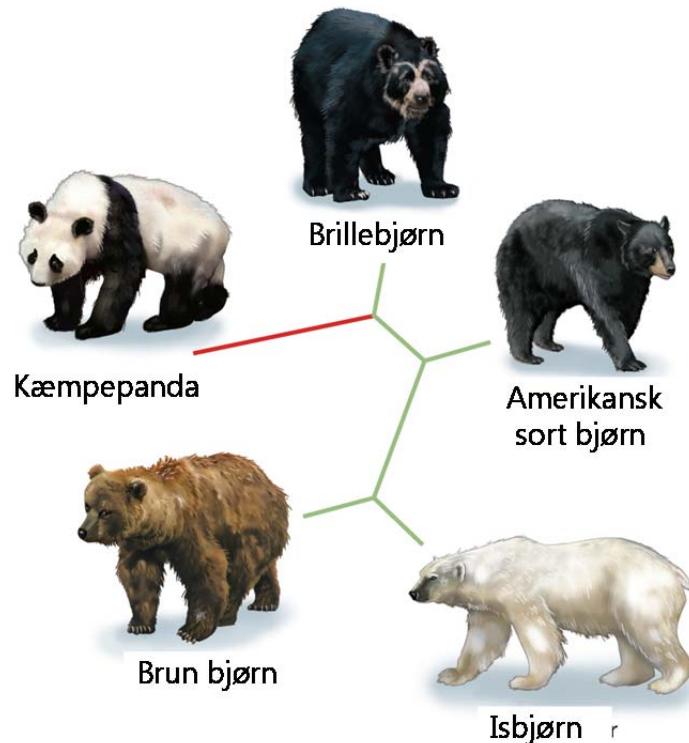
Betræt de følgende to træer, hvor firkanten antages at være den oprindelige karakter:



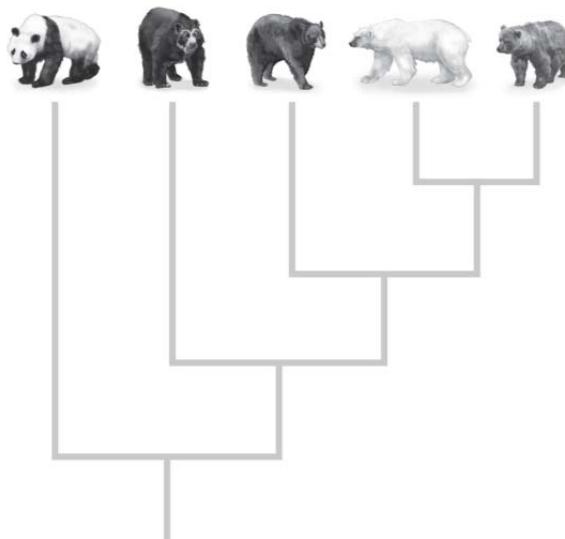
- b) Hvilket af de to træer giver den mest parsimoniske forklaring på fordelingen af karaktererne i de to træer? (Angiv antal karakterændringer for hvert af dem.)

Det må være træ B. Det kræver tre ændringer, medens det andet træ kræver fire ændringer.

Betragt det følgende træ. Det illustrerer en fylogeni uden rod af flere arter af bjørne baseret på DNA-sekvenser. Kæmpepandaen kan benyttes som udgruppe til de øvrige.



- c) Benyt udgruppe metoden til at tegne en fylogeni med rod over de fem arter af bjørne.



- d) Hvor mange monofyletiske grupper er der indeholdt i træet (medtag alle grupper med over 1 art)?

4

- e) Hvordan betegnes en gruppe, der udelukkende indeholder den brune bjørn, den amerikanske sorte bjørn samt deres fælles forfader?

Den er parafyletisk.

## Opgave 2 (22 %)

Hos den nordamerikanske sorte bjørn (*Ursus americanus*) findes en mutant i melanocortin 1 receptor genet (mc1r). En mutation fra A til G i codon 298 resulterer i en aminosyresubstitution af Tyrosin til Cystein. Individer, der er homozygote GG får fænotypen hvid (se billedet). Vi har derfor tre genotyper, som resulterer i to fænotyper

Genotype	AA	AG	GG
Fænotype	Sort	Sort	Hvid

Antag, at vi har observeret  $N_W$  hvide og  $N_B$  sorte bjørne i en population med summen  $N_T$ .



- a) Hvordan ville du beregne frekvenserne for de to alleler ud fra fænotyperne? Angiv de formler, man kan bruge til at beregne hhv. frekvensen af allele A og allele G.

$$q = \sqrt{N_W/N_T}$$

$$p = 1 - q$$

- b) Hvilke forudsætninger blev gjort for at kunne beregne frekvenserne?

Der blev antaget, at vi havde Hardy-Weinberg proportioner.

I et molekylært studie sekventerede man mc1r genet, og bestemte på denne måde frekvensen af A og G i codon 298 af mc1r genet. På tre små øer ud for Canadas vestkyst fandt man den følgende genotypefordeling

Population	Genotype			$\chi^2$	$F$
	AA	AG	GG		
Gribbell	7	6	10	??	??
Princess Royal	26	17	10	4.59	0.29
Roderick	9	1	2	6.70	0.75

- c) Beregn frekvenserne for allele A og G for de tre populationer.
- d) Test, om der er afvigelse fra Hardy-Weinberg proportioner for Gribbell populationen.
- e) Beskriv afvigelser fra Hardy-Weinberg proportioner ved hjælp af indavlskoefficienten  $F$ . [ $F = (H_e - H_o)/H_e$ , hvor  $H_e$  og  $H_o$  er henholdsvis forventet og observeret hyppighed af heterozygoter]. Er der overskud af heterozygoter eller overskud af homozygoter? Det skal beregnes for Gribbell populationen, for de to øvrige er tallene angivet.

c, d, e

AA	AG	GG	N	p	q	Exp(AA)	Exp(AG)	Exp(GG)	F	$\chi^2$	P
7	6	10	23	0.435	0.565	4.35	11.30	7.35	0.469	5.06	0.02443
26	17	10	53	0.651	0.349	22.46	24.08	6.46	0.294	4.59	0.03223
9	1	2	12	0.792	0.208	7.52	3.96	0.52	0.747	6.70	0.00963

Alle tre populationer afviger fra Hardy-Weinberg proportioner med et overskud af homozygoter.

Der blev fremsat to forskellige forslag til at forklare afvigelsen fra Hardy-Weinberg proportioner. Det kunne være positiv assortativ parring eller indavl.

- f) Er det muligt ud fra fordelingen af genotypefrekvenserne i et enkelt locus (her mc1r) at afgøre om det er positiv assortativ parring eller indavl, der er skyld i afvigelsen fra Hardy-Weinberg proportioner? Begrund dit svar **kort**.

Nej, begge resulterer i et overskud af homozygoter.

Der blev i det samme studie undersøgt variationen i 10 autosomale mikrosatelliter i de undersøgte tre populationer. Her fandt man ingen afvigelse fra Hardy-Weinberg proportioner.

- g) Benyt denne oplysning til at afgøre om afvigelsen fra Hardy-Weinberg proportioner skyldes indavl eller positiv assortativ parring. Giv en **kort** begrundelse.

Det må være assortativ parring. Indavl påvirker alle autosomale loci ens og ville derfor være blevet observeret hos de undersøgte mikrosatellit loci. Assortativ parring påvirker kun de loci, der koder for de fænotyper, hvor der foregår assortativ parring.

En sammenligning af allelfrekvenserne i de tre ø-populationer viste, at de var signifikant forskellige fra hinanden. En undersøgelse af flere fastlandspopulationer viste, at frekvensen af den lyse morf var meget lav i dem alle sammen, og at der ikke var nogen forskel mellem populationerne.

- h) Giv en forklaring på, hvorledes populationerne på øerne er divergeret fra hinanden og fra fastlandspopulationen.

Differentieringen skyldes formentligt genetik drift. Der har været founder events, da øerne blev koloniseret og siden hen har genetisk drift medvirket til en yderligere differentiering i de små populationer. Migrationen mellem øerne er betydelig mindre end migrationen på fastlandet.

### Opgave 3 (14 %)

Betrægt et autosomalt locus med to alleler, der har indflydelse på bærerens fitness, som angivet i tabellen:

Genotype	$A_1A_1$	$A_1A_2$	$A_2A_2$
Fitness	0.6	1	0.4

- a) Hvilken form for selektion foregår? Hvad er selektionskoefficenterne for de to homozygoter?

Overdominant selektion.  $A_1A_1: s = 0.4 \quad A_2A_2: t = 0.6$

Betrægt en population af dyr med tilfældig parring, hvor vi har zygoter med Hardy-Weinberg proportioner og fitness som angivet ovenfor. Vi antager at populationen er meget stor og derfor ikke udsættes for genetisk drift.

- b) Hvor mange ligevægte findes i dette system? Hvor mange stabile, hvor mange ustabile? Angiv frekvensen for alle mulige ligevægte.

2 ustabile:  $p = 0$  og  $p = 1$

1 stabil:  $p = t/(s + t) = 0.6/(0.4 + 0.6) = 0.6$

Betrægt nu en population af en selvbestøvende plante med fitness som angivet ovenfor. De forplanter sig udelukkende ved selvbestøvning. Vi betragter det samme locus. Antag, at vi starter med en 1/3 af hver genotype. Der ses igen bort fra genetisk drift.

- c) Hvordan vil evolutionen i en sådan plante foregå?

Populationen fikseres for  $A_1A_1$ : Heterozygoten har en selektiv fordel, men i hver generation vil den spalte ud med hver homozygot og vil efterhånden forsvinde. Genotype  $A_1A_1$  har højere fitness end  $A_2A_2$  og vil over tiden blive fikseret.

Betrægt nu en population af en apomiktisk plante (som mælkebøtter f.eks.). I en apomiktisk population er alt afkom af en plante en klon af moderens genotype. Antag, at vi starter med en 1/3 af hver genotype, og at fitness er som angivet ovenfor. Der ses igen bort fra genetisk drift.



- d) Hvordan vil evolutionen i en sådan plante foregå?

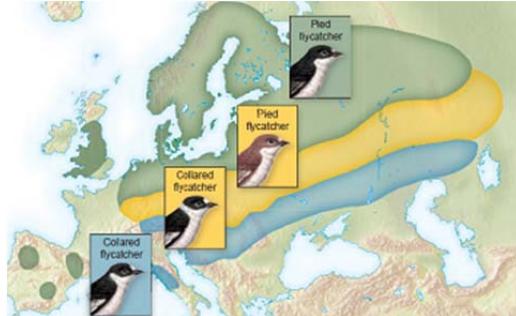
Heterozygoten  $A_1A_2$  vinder. Den har den højeste fitness og bliver fikseret i populationen.

- e) I hvilken af de tre nævnte tilfælde har populationen den højeste fitness, når der er opnået en ligevægt?

I tilfælde for den apomiktiske plante, der bliver fikseret for heterozygoten med fitness 1.

#### Opgave 4 (14%)

Figuren viser udbredelsen af to arter af fliesnappere: broget fliesnapper (*Ficedula hypoleuca*, pied flycatcher) og hvidhalset fliesnapper (*Ficedula albicollis*, collared flycatcher). I områder, hvor begge arter forekommer, sker det sjældent at de hybridiserer. Det viser sig, at hybridhunner altid er sterile, medens hybridhanner kan få afkom. De får dog færre afkom relativt til de to forældrearter gennem hele deres liv. De opfylder Haldane's regel.

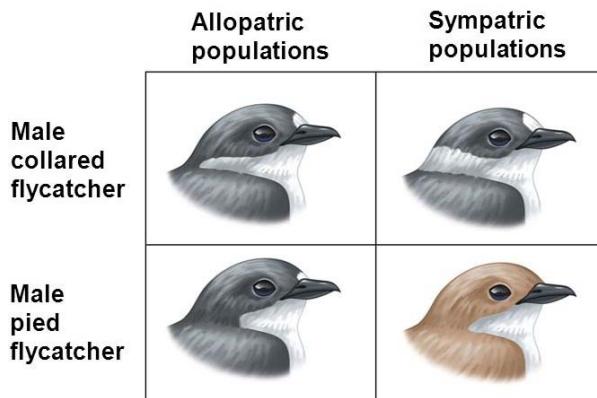


- a) Hvilket køn er det heterogametiske køn?

**Det er hunnen.**

Figuren viser hanner fra de to arter, hvor de henholdsvis forekommer allopatriske og sympatriske.

- b) Diskuter **kort**, hvilke evolutionsmekanismer, der kan forklare hvorfor de to arter ligner hinanden i allopatriske populationer, men ser mere forskellige ud i sympatriske populationer.



Der er sket et reproduktivt character displacement, så arterne i sympati minder mindre om hinanden end i allopatriske områder.

- c) Dobzhansky delte reproduktive isolationsmekanismer op i to kategorier. Hvilken form er der tale om i henholdsvis 1) sterilitet hos hybridhunner og nedsat fertilitet hos hanner samt 2) det emne, der diskutes i spørgsmål b?

**Postzygotiske og præzygotiske isolationsmekanismer.**

- d) Hvilken af de to mekanismer øges gennem naturlig selektion i sympatriske områder? Begrundes **kort**.

Det er præzygotiske isolationsmekanismer, som forhindrer dannelsen af hybrider med nedsat fitness. Naturlig selektion virker ikke på en måde, hvor den reducerer fitness ved at øge den postzygotiske isolation.

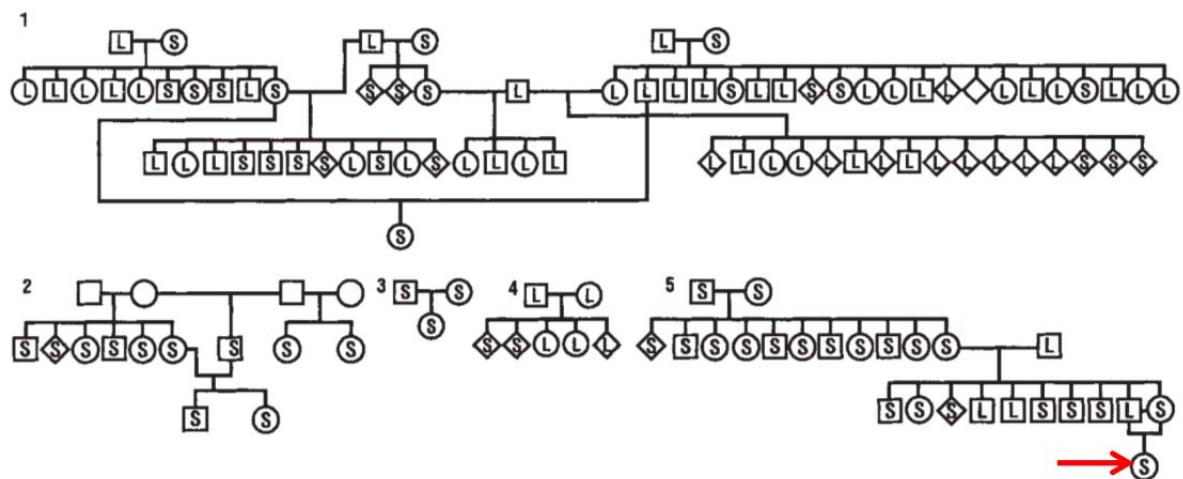
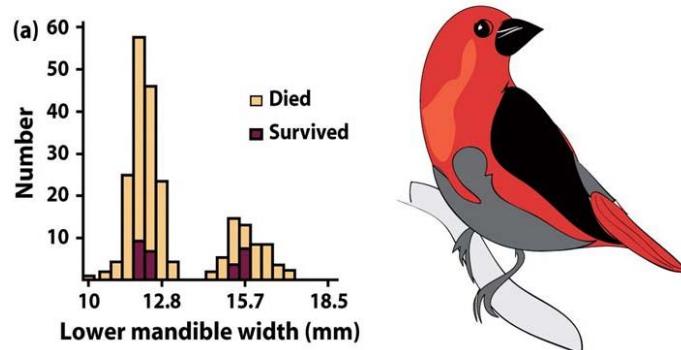
- e) Hvilken form for artsdannelse anses for at være den mest udbredte hos dyr?

## Allopatrisk artsdannelse.

### Opgave 5 (14%)

Den afrikanske finke (*Pyrenestes ostrinus*) er blevet studeret af Thomas Bates Smith. Fugle i denne art findes i to grupper med forskellig næbstørrelse; en gruppe med store næb og en gruppe med små næb. Figuren til højre viser en fordeling af næbstørrelsen i en naturlig population i Cameroun (som størrelse for næbstørrelsen er vist breden af næbbet).

Bates Smith analyserede udspaltningen af afkom fra forskellige parringer i en population. Der foregår tilfældig parring med hensyn til næbstørrelse. Den følgende figur viser stamtavler for forskellige krydsninger. I stamtavlen får hunner en cirkel,  $\circ$ , medens hanner får en firkant,  $\square$ . Er kønnet ukendt, angives det som  $\diamond$ . Individer med stor næbstørrelse er angivet med L (Large), medens individer med små næb er angivet med S (Small)



- a) Benyt stamtavlerne til at udlede den genetiske baggrund for fænotyperne S og L. Angiv hvilke genotyper, der bestemmer de to fænotyper og angiv dominansforhold. Begrundes *kort*.

Udspaltingerne peger på et enkelt autosomalt locus med to alleler, S og L, hvor L dominerer over S. Gentyperne LL og LS resulterer i fænotypen stort næb, medens gentypen SS resulterer i fænotypen lille næb. S krydset med S får altid S afkom, medens de øvrige krydsninger spalter ud.

- b) Angiv genotyperne for forældrene i stamtavle nr. 4.

## LS og LS

- c) Hvad er indavlskoefficienten af individet, der er angivet med en pil, hvis vi antager at olde- og bedsteforældrene er ubeslægtede?

Det er afkom af to helsøkende, så indavlskoefficienten er  $(1/2)^3 + (1/2)^3 = (1/2)^2$

Finken lever hovedsageligt af frø, hvor størrelsen af de frø, der spises, er korreleret med næbstørrelsen. Finker med små næb spiser små bløde frø, medens finker med store næb spiser store hårde frø. Fordelingen over næbstørrelsen vises for unge fugle, hvor de mørke sjøbler er dem, der overlever til voksenstadiet, medens de lyse sjøbler viser fugle, som ikke overlever til voksenstadiet.

- d) Hvilken form for naturlig selektion virker, når man betragter begge størrelsesklasser af finken?

Disruptiv selektion, hvor de to grupper har forskellige fødenicher.

- e) Hvilken form for naturlig selektion virker, når man betragter de to størrelsesklasser af finken hver for sig?

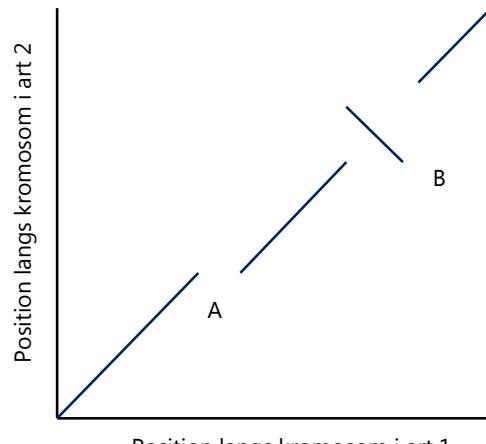
Stabiliserende selektion, hvor individer tæt på middelværdien af en given størrelsesklasse har højere overlevelse end individer længere væk fra middelværdien.

## Opgave 6 (14%)

I figuren til højre sammenligner man det samme kromosom fra to nært beslægtede arter i et såkaldt "syntetic dot plot". Der er åbenlyst sket ændringer i to områder af kromosomet, som er angivet med A og B.

- a) Hvilke evolutionære hændelser er sket i de to områder af kromosomet? I A kan der være sket to forskellige hændelser.

I A er der sket enten en insertion eller deletion. I B er der sket en inversion.



- b) Hvordan ville du undersøge hvilken af de to hændelser, der er sket i A? (Vink: inddrag yderligere en art.)

Ved at sammenligne med en fjerne beslægtet art vil man kunne se, om den i A ligner art 1 eller art 2. Hvis den ligner art 1, må der være sket en deletion i art 2. Hvis den ligner art 2, må der være sket en insertion i art 1.

- c) Hvordan varierer rekombinationsraten langs kromosomer hos eukaryote organismer, som for eksempel mennesket?

Rekombinationsraten er meget ujævnt fordelt, med høje rekombinationsrater i rekombinations hot spots.

- d) Hvordan påvirkes LD (linkage disequilibrium, koblingsuligevægt) i et område, hvor der er høj rekombinationsrate i hver ende af området og mindre rekombinationsrate ind imellem?

Der vil være højt LD i områder imellem de to rekombinations hot spots. (En haplotypeblok)

Betræt et område som nævnt ovenfor (i spørgsmål d), hvor man har fundet to haplotyper i en population, haplotype I og haplotype II. Haplotype I forefindes med en frekvens på 40% i populationen, medens haplotype II forefindes med en frekvens på 60% i populationen. De to haplotyper bærer den følgende variation i locus C og D. I haplotype I har locus C allelen  $C_1$  medens den i locus D har allelen  $D_1$ . Tilsvarende forefindes der i haplotype II allelerne  $C_2$  og  $D_2$ .

- e) Hvad er frekvensen af de to alleler i hvert locus,  $p(C_1)$ ,  $p(C_2)$ ,  $p(D_1)$ ,  $p(D_2)$ ?

I begge loci må allel 1 have frekvensen 0.4 og allel 2 må have frekvensen 0.6.

- f) Er der koblingsuligevægt (LD) mellem de to loci? (Begrundes **kort**). Beregn dernæst koblingsuligevægten  $D$ .

Ja,  $C_1$  er altid sammen med  $D_1$  og  $C_2$  er altid sammen med  $D_2$ .

$$\begin{aligned}f_{C_1D_1} &= p(C_1)p(D_1) + D \\D &= f_{C_1D_1} - p(C_1)p(D_1) \\&= 0.4 - 0.4 \times 0.4 \\&= 0.24\end{aligned}$$

- g) Hvad sker der, når der er retningsselektion i locus D, hvor fitness for de tre genotyper er givet ved

Genotype	$D_1D_1$	$D_1D_2$	$D_2D_2$
Fitness	1.2	1.1	1

Der skal ses bort fra genetisk drift.

Allel  $D_1$  fikseres i populationen.

- h) Hvordan påvirkes variationen i locus C?

Allel  $C_1$  vil på grund af hitchhiking øges og vil også blive fikseret, medmindre der sker rekombination mellem C og D før det sker.

## **Evolutionsbiologi**

Skriftlig eksamen, den 19. juni 2019

**Alle hjælpemidler tilladt  
(Lærebog, noter, computer)**

For hver opgave oplyses dens vægtning.

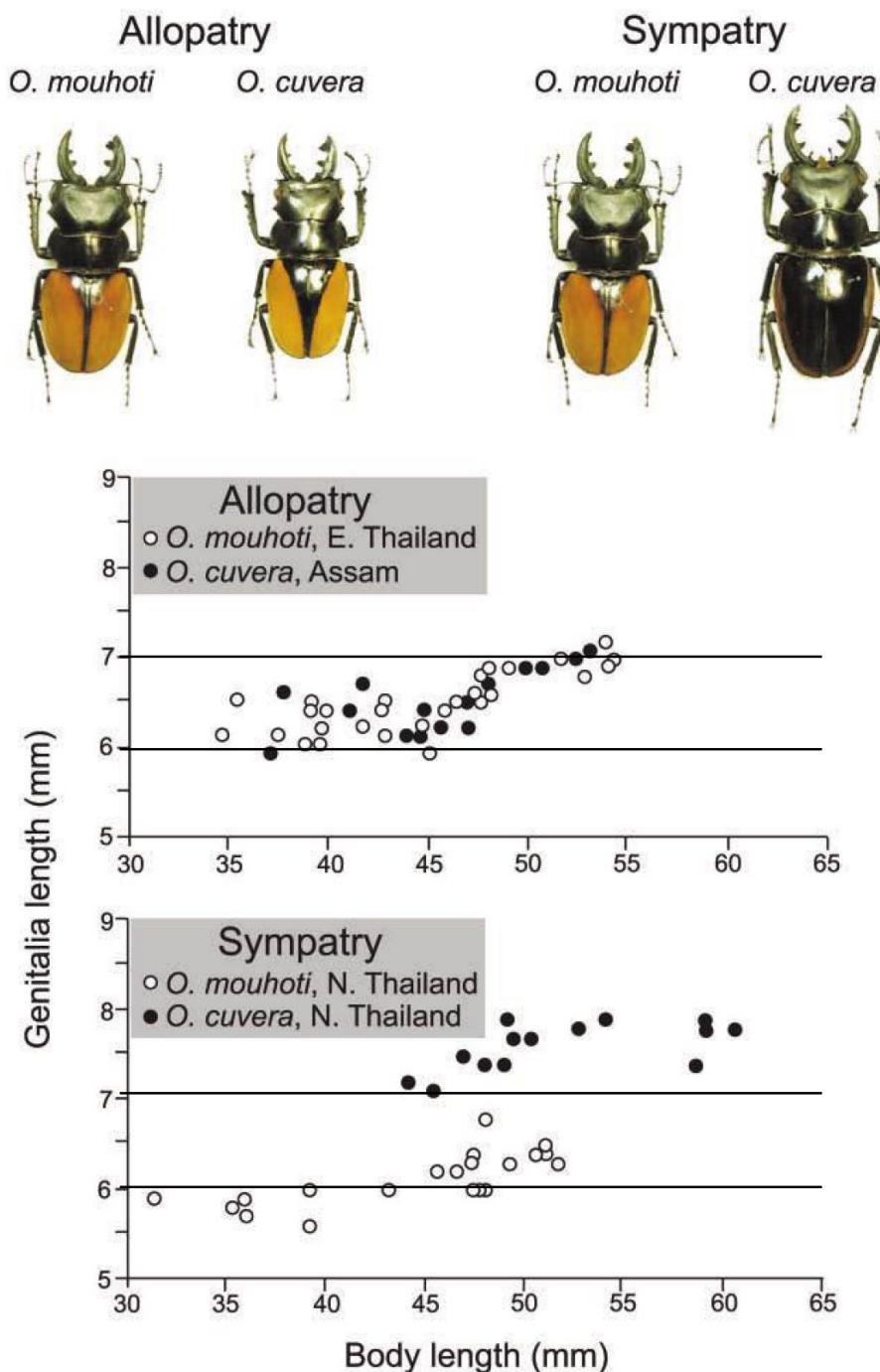
*Dette opgavesæt indeholder **9** sider inklusive forsiden  
Opgaverne skal afleveres i den rækkefølge, de er stillet i.*

Censorer: Tove Hedegaard Jørgensen  
Eksaminatorer: Rasmus Heller

Jesper Givskov Sørensen  
Hans R. Siegismund

### Opgave 1 (20 %)

Kawano (2003)<sup>1</sup> har undersøgt morfologisk variation hos to billearter i slægten *Odontolabis*, *O. mouhoti* og *O. cuvera*, i Sydøstasien. Billerne tilhører familien hjortebiller (*Lucanidae*), hvor også eghjorten (*Lucanus cervus*) findes.



<sup>1</sup> Kawano, K. 2003 *Annals of the Entomological Society of America* **96**: 503–511.

Figuren viser de målinger, som Kawano foretog på hanner af de to arter fra populationer, der var henholdsvis allopatriske og sympatriske mellem de to arter. Desuden er der vist billeder af hanner fra allopatrisk og sympatisk udbredelse.

- Sammenlign udseendet og de morfologiske målinger af de to arter i allopatri og sympatri.

**Udeende:** I allopatri ligner de to arter hinanden, medens de ser meget forskellige ud i sympatri. Specielt *O. cuvera* har ændret sig.

**Morfologi:** I allopatri er der stort set et fuldstændigt overlap mellem de to arter med hensyn til de to målte karakterer. I sympatri ses, at *O. cuvera* er blevet længere og at dens genitalielængde er blevet højere. Tilsvarende er *O. mouhoti* blevet kortere og dens genitalielængde er blevet lavere.

- Hvilket fænomen er illustreret i figuren?

“Reproductive character displacement”

- Giv et bud på, hvordan fitness af hybrider mellem de to arter i sympatri er relativt til fitness af afkom, hvor begge forældre tilhører samme art.

Hybrider mellem de to arter har nedsat fitness. Hvis ikke det var tilfældet ville der ikke blive selekteret for “Reproductive character displacement”

- Beskriv hvilke evolutionære processer, der har resulteret i de observerede forskelle mellem populationer i allopatri og sympatri.

De to arter har formentligt mødt hinanden i en sekundær kontaktzone. Her har de hybridiseret, hvorefter der sket et “Reproductive character displacement”, hvor naturlig selektion har øget den morfologiske forskel mellem dem for at forhindre dannelsen af hybrider med nedsat fitness. Den præzygotiske isolation øges.

- Hvilke typer inddeltes arts dannelse i, når der fokuseres på udbredelsesmæssige aspekter?

Allopatrisk, parapatrisk og sympatisk.

## Opgave 2 (30 %)

Chen et al. (2017)<sup>2</sup> har rapporteret de følgende genotypefordelinger for en SNP [C1orf62 (rs7551421)] for to grupper af mennesker i Europa, finner og ikke-finske personer:

Population	Genotype			Sum
	AA	AB	BB	
Europæisk (finsk)	864	1630	812	3306

<sup>2</sup> Chen, B, Cole, JW & Grond-Ginsbach, C (2017) *Front. Genet.* **8**:167. doi: 10.3389/fgene.2017.00167

Europæisk (ikke-finsk)	5564	15393	12372	33329
------------------------	------	-------	-------	-------

- a. Beregn frekvenserne for allel A og B for de to populationer.

Finsk:  $p(A) = (2 \times 864 + 1630)/(2 \times 3306)$   
 $= 0,508$

$p(B) = (2 \times 812 + 1630)/(2 \times 3306)$   
 $= 0,492$

Ikke-finsk:  $p(A) = (2 \times 5564 + 15393)/(2 \times 33329)$   
 $= 0,398$

 $p(B) = (2 \times 12372 + 15393)/(2 \times 33329)$   
 $= 0,602$

- b. Stemmer genotypefordelingerne overens med Hardy-Weinberg-fordelingen? (Angiv de forventede genotyper og lav en  $\chi^2$ -test.)

Population	Genotype			$\chi^2$ -test
	AA	AB	BB	
Europæisk (finsk)	852,70	1652,59	800,70	0,62
Europæisk (ikke-finsk)	5275,91	15959,18	12083,91	43,39

- c. Hvordan afviger de observerede fra de forventede genotypefordelinger?

Der er et signifikant overskud af homozygoter i den ikke-finske population.

- d. Hvad er den mest sandsynlige forklaring på afvigelsen fra Hardy-Weinberg proportioner?

Der er tale om en Wahlund effekt. Chen et al. (2017) har slået alle ikke-finske europæiske populationer sammen, selv om der formentlig har været forskelle i allelfrekvenser mellem dem.

Finnerne har en forskellig allelfrekvens fra andre europæere i den ovennævnte SNP. Dette synes at være generelt fænomen hos finnerne. Desuden har de også variation i arvelige sygdomme, som gør dem unikke. For eksempel, sygdommen diastrophic dysplasi, som er en recessiv arvelig lidelse, forekommer hos én ud af 15000 i Finland, medens den forekommer hos én ud af 500000 i USA.

- e. Bestem frekvensen for den recessive allel og angiv hvilken antagelse der blev taget for at beregne den.

Frekvensen af den recessive allel kan beregnes under antagelse af Hardy-Weinberg proportioner.

$$q_{Finland} = \sqrt{\frac{1}{15000}} = 0,0082$$

$$q_{USA} = \sqrt{\frac{1}{5000000}} = 0,0014$$

- f. Hvordan kan man forklare, at finnerne er så genetisk unikke?

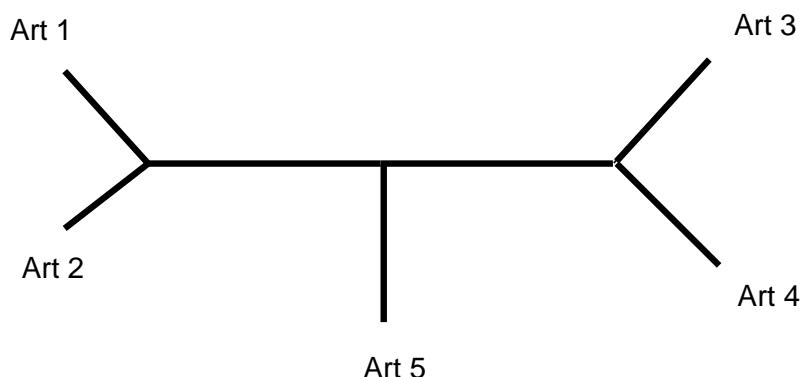
Finnerne har været igennem en founder-effekt, da de koloniserede Finland og har været relativt isoleret fra andre sidenhen. (De bosatte sig i Finland for cirka 2000 år siden.)

- g. Hvordan kan det være, at skadelige recessive alleleler har fået så en høj frekvens i den finske befolkning?

Genetisk drift virker stærkere i små populationer og kan “overdøve” naturlig selektion, når populationens størrelse bliver tilstrækkelig lille.

### Opgave 3 (20%)

Den følgende fylogeni skal analyseres.



- a. Beregn parsimoniscoren for dette træ baseret på de følgende sekvenser:

	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
Art 1	T	C	C	A	G	A	C	C	A	G	G	T	G	G	A	G
Art 2	T	C	C	A	G	A	C	C	A	G	G	T	G	G	A	G
Art 3	T	C	T	A	A	A	C	A	G	A	T	G	G	C	A	T
Art 4	G	C	T	A	A	A	C	A	G	A	T	G	G	C	A	T
Art 5	T	C	T	A	C	A	G	A	T	G	G	G	A	T	G	G

- b. Antag, at art 5 er udgruppe til de øvrige. Hvor mange monofyletiske grupper findes i dette træ hvis rodens placering på den gren, der leder til art 5? (Medregn ikke grupper, der kun består af én art eller af alle 5 arter)?  
 (1, 2), (1, 2, 3), (1, 2, 3, 4).
- c. Hvor mange transitioner og transversioner adskiller sekvenserne fra art 1 og art 4?

**1 transversion, 5 transitioner**

- d. Hvis rodens placering var på den gren som leder til art 3, hvilke monofyletiske grupper ville da kunne findes i træet?  
 (1, 2), (1, 2, 5), (1, 2, 5, 4).
- e. Tyder substitutionsmønstret på, at dette gen er et almindeligt proteinkodende gen? (Hjælp: antag at 1, 2 og 3 svarer til kodonpositioner.)

Ja, der flest substitutioner på position 3 (fire), færre på position 1 (to) og ingen på position 2. Det er, hvad vi forventer for proteinkodende gener, hvor position 3 har den højeste rate efterfulgt af position 1 og så position 2.

#### Opgave 4 (10 %)

En evolutionsbiolog har lavet fire simulationer, hvor populationens størrelse har været sat til 100 i alle tilfælde. I tre af tilfældene indgik der forskellig form for naturlig selektion. Frekvensen af allele A var i alle tilfælde 0,1 i generation 0. Den anden allele er B. Simulationerne varede 500 generationer, og der indgik 50 populationer i hver af de fire simulationer. Resultaterne af simulationerne fremgår af figuren.

- a. Angiv for Figur I til IV om der foregår selektion og i givet fald, hvilken form for selektion, der foregår. Angiv fitnessværdierne for de tre genotyper  $W_{AA}$ ,  $W_{AB}$  og  $W_{BB}$  relativt til hinanden. (F.eks.  $W_{AA} > W_{AB} > W_{BB}$ ) Begrundes kort.

I: Overdominant selektion.  $W_{AA} < W_{AB}$  og  $W_{BB} < W_{AB}$ . Allelfrekvenserne holder sig tæt omkring 0,1, men i de fleste populationer fikseres allele B.

II: Ingen selektion.  $W_{AA} = W_{AB} = W_{BB}$ . Allelfrekvenserne bevæger sig tilfældigt op og ned. Desuden fikseres 5 populationer for allele A, hvilket er lig med forventningen, da  $5/50 = 0,1$ , hvilket var allelfrekvensen til at starte med.

III: Retningsselektion.  $W_{AA} \geq W_{AB} \geq W_{BB}$  med højst ét lighedstegn. Langt de fleste populationer fikseres for allele A.

IV: Underdominant selektion.  $W_{AA} > W_{AB}$  og  $W_{BB} > W_{AB}$ . Vi ser, at når en population får en frekvens af A på over 0,5 bliver den altid fikseret for allele A. Ændringerne i allelfrekvenser er ikke så tilfældige, som man forventer med genetisk drift.

- b. Tre af simulationerne har et sammenfaldende ligevægtspunkt. Hvad er allelfrekvensen af det?

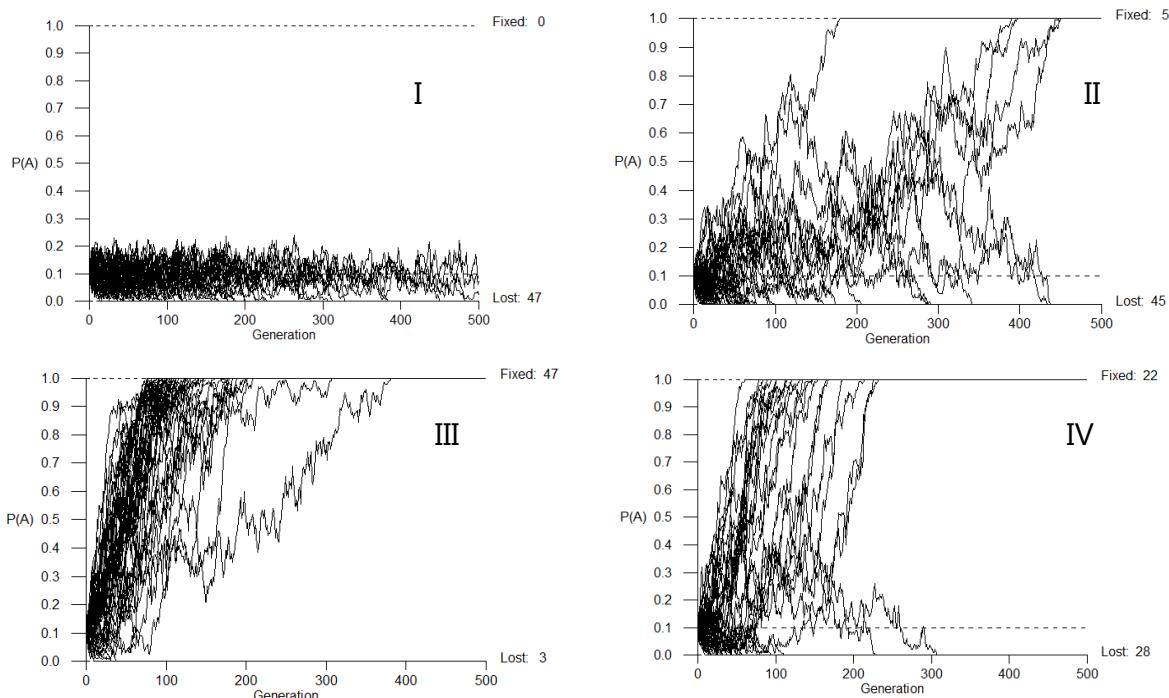
0.1. I tilfældet for ingen selektion (II) vil ligevægten vil være 0,1, som er frekvensen i generation 0. (For en uendelig stor population) Derfor må I og IV have det samme ligevægtspunkt. For øvrigt, så viser det benyttede simulationsprogram (PopG) ligevægte som stiplede linjer for tilfældet, hvor der ikke virker genetisk drift.

c. Hvilken form for ligevægt er der tale om i de tre tilfælde?

I: stabil ligevægt

II: neutral ligevægt

IV: ustabil ligevægt



### Opgave 5 (10 %)

Kelley (2008)<sup>3</sup> har lavet et selektionsforsøg på blomsterstørrelsen hos planten åben abebłomst (*Mimulus guttatus*) fra maskeblomstfamilien (Scrophulariaceae), se billedet forneden. Udgangspopulationen havde en gennemsnitsbredde af blomsten på 18.9 mm. De 20% individer med de bredeste blomster blev udvalgt som forældre til næste generation. Disse havde en gennemsnitsstørrelse på 22,9 mm. Narrow sense heritabiliteten i denne population blev estimeret til  $h_N^2 = 0,3$ .

<sup>3</sup> Kelley J.K. (2008) *Genetica* **132**:187–198



- a. Hvad bliver gennemsnitsstørrelsen af blomsterbredde i afkomsgenerationen?

$$\text{Selektionsdifferentialet: } S = 22,9 - 18,9 = 4$$

$$\text{Responsen: } R = h^2 S = 0,3 \times 4 = 1,2$$

$$\text{Gennemsnitsstørrelsen af blomsterbredde i afkomsgenerationen} = 18,9 + 1,2 = 20,1.$$

Selektionsexperimentet blev fortsat i 5 generationer. Samtidigt blev der lavet et parallelt forsøg, hvor man selekterede for mindre blomster. Da man stoppede eksperimentet, målte man endvidere planternes biomasse (tørvægt af plantedelen over jorden i mg). "Kontrol" i den følgende tabel er en fortsættelse af udgangspopulationen med tilfældig parring over 5 generationer. Bemærk, der blev udelukkende selekteret på blomsternes størrelse.

Selektionsexperiment			
Karakter	Små	Kontrol	Store
Blomsterbredde (mm)	13,0	18,9	26,2
Biomasse (mg)	28,0	44,8	81,9

- b. Hvorfor ændrede biomasse sig i selektionsexperimentet for mindre og større blomsterbredde?

Pleiotropi: de loci, der har bidraget til blomsternes størrelse har også haft indflydelse på plantens størrelse som sådan.

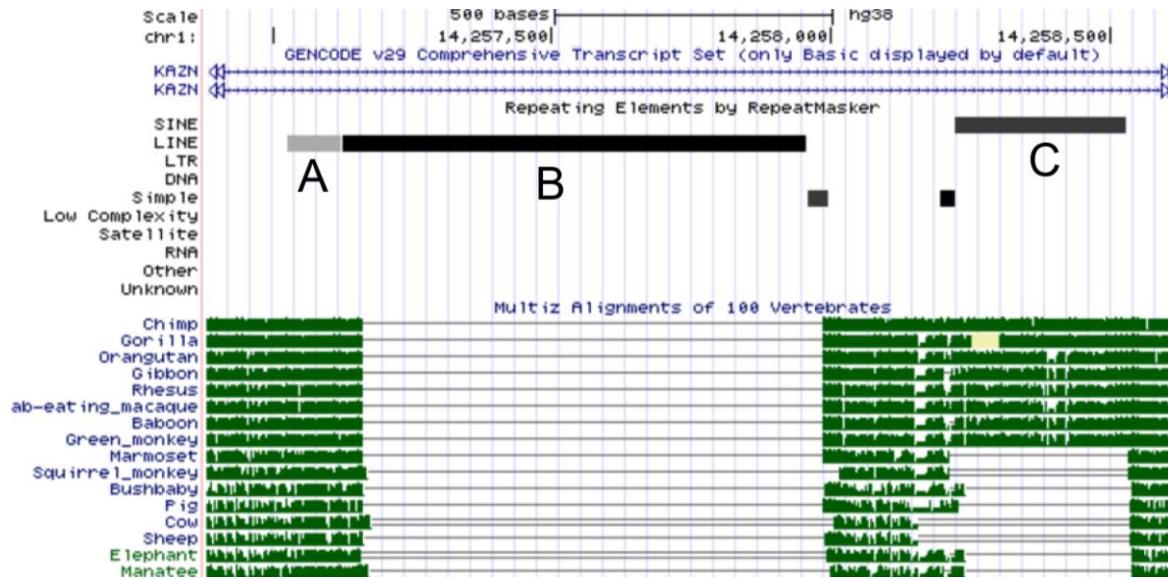
Når man laver selektionsexperimenter, der varer i meget lang tid, resulterer det ofte i et selektionsplateau, hvor der ikke sker nogen ændring.

- c. Skyldes dette, at der ikke forefindes genetisk variation for den karakter, der er blevet selekteret på?

Nej, der kan stadig være genetisk variation. Responsen er 0. Derfor må den additive genetiske varians  $V_A$  være 0.  $R = 0 = h^2 S = V_A/V_P S \Rightarrow V_A = 0$ .

### Opgave 6 (10 %)

På UCSC genom browser figuren vises en genomisk region på det humane kromosom 1, nærmere bestemt en del af et intron i KAZN genet. I dette intron findes i det humane genom nogle transposons.



- a. Hvad er transposons og hvorfor findes de i det humane genom?

Transposons er selviske genetiske elementer, der kan kopiere sig selv til nye positioner i humane genomet vha. reverse transskription.

- b. Hvad er forskellen på LINE og SINE transposons?

LINE er længere end SINE fordi de koder for deres egen reverse transskriptase, mens SINES nasser på LINEs transskriptase for at kunne kopiere sig til en ny position.

- c. Kan du ud fra figuren sige noget om i hvilken forfader de 3 transposons (A, B, C) er blevet en del af genomet?

A: Forfader før Manatee, B: Menneske/urfader efter chimpanse, C: Forfader efter marmoset, men før Green Monkey.

## **Evolutionsbiologi**

Skriftlig eksamen, den 19. juni 2020

**Alle hjælpemidler tilladt  
(Lærebog, noter, computer)**

For hver opgave oplyses dens vægtning.

Dette opgavesæt indeholder **11** sider inklusive forsiden  
*Opgaverne skal afleveres i den rækkefølge, de er stillet i.*

Censorer:            Tove Hedegaard Jørgensen            Jesper Givskov Sørensen  
Eksaminatorer:      Rasmus Heller                        Hans R. Siegismund

### Opgave 1 (10 %)

Mennesker har immigreret til Grønland i flere omgange, men de fleste immigrationer uddøde. Den sidste immigration skete for cirka tusind år siden. Menneskene har levet af en ernæring der er ekstrem protein- og fedtholdig kost, som har medført selektion for at kunne fordøje denne kost. Blandt andet er der tegn på positiv selektion i carnitine palmitoyl-transferase 1A (*CPT1A*) genet, der er involveret i stofskiftet af fedtsyrer. De oprindelige grønlændere er blevet fikseret for basen T i en SNP. Europæere, herunder danskere, er fikseret for basen C. I en undersøgelse over den genetiske variation af mennesker fra alle større bygder og byer på Grønland fandt man den følgende fordeling af genotyper

TT	TC	CC	Sum
1800	991	214	3005

- a) Bestem frekvenserne af de to alleler.

$$p(T) = (2 \times 1800 + 991)/(2 \times 3005) = 0.764$$

$$p(C) = (2 \times 214 + 991)/(2 \times 3005) = 0.236$$

- b) Er der Hardy-Weinberg proportioner?

Nej

Obs 1800	991	214
Exp 1753.52	1083.97	167.52

$$\chi^2 = 22.10 \quad P = 2.584004e-06$$

- c) Beregn indavlskoefficienten  $F$ . Hvordan afviger de observerede genotyper fra de forventede ifølge Hardy-Weinberg proportioner?

$$F = (1083.97 - 991)/1083.97 = 0.086$$

Der er overskud af homozygoter i forhold til Hardy-Weinberg proportioner

- d) Hvad kan årsagen til en eventuel afvigelse være?

Stikprøven er fra hele Grønland, hvor der tilsyneladende er populationsstuktur. Vi får dermed en Wahlund effekt.

- e) Med hvor stor en del er den grønlandske befolkning blevet op blandet med europæiske gener? Tag udgangspunkt i *CPT1A* genet. Giv et bud på, hvordan denne opblanding har fundet sted.

De oprindelige grønlændere har været fikseret for basen T, europæere er fikseret for C. Nu er allelfrekvensen 0.764. Derfor må den europæiske del udgøre 0.236.

### Opgave 2 (20 %)

En forsker analyserer 40 basepar fra en region og håber ved hjælp af disse data at estimere en fylogeni for fire arter. Hun finder følgende alignment:

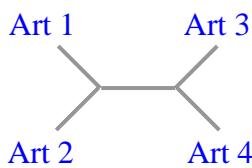
Art 1: acgttgcagt acattgcagt tcattgcagt acattgcagt  
 Art 2: ..... .... a..... ....  
 Art 3: ..a..... ....c.... a..... ....t.  
 Art 4: ..a.....a. .... a..... ....t.

En prik i en sekvens viser, at basen er identisk med basen i sekvensen for art 1

- a) Hvor mange forskellige topologier er der for fylogeni for fire arter med og uden rod?

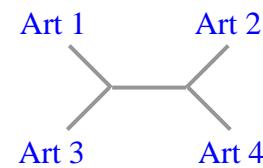
**15 og 3**

- b) Hvis art 1 og art 2 er søstertaxa, hvor mange mutationer må der som minimum være sket i disse sekvenser.

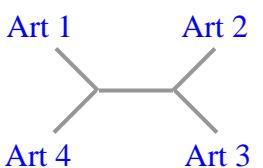


**5 mutationer**

- c) Tæl antallet af mutationer på hvert af de (andre) mulige træer (uden rod) under antagelse af parsimoni.



**7 mutationer**



**7 mutationer**

- d) Hvilket træ (uden rod) er maksimum parsimoni træet?

**Det under b)**

- e) Antag at art 1 og art 2 er søster taxa og at art 4 er udgruppe (outgroup). Bortset fra grupper, der inkluderer enten 1 eller 4 arter, hvilke monofyletiske grupper findes der da i denne fylogeni?

**(1,2), (1,2,3)**

- f) Antag at art 1 og art 2 er søster taxa og at art 4 er udgruppe (outgroup). Hvordan ville du så betegne en gruppe, der består af art 1 og art 3?

**parafyletisk**

- g) Brug parsimonikriteriet til at bestemme hvilket nukleotid (base) der var i den sidste fælles stamfader (most recent common ancestor) for disse fire sekvenser i det første,

tredje, og 21. position i den viste alignment.

1: a

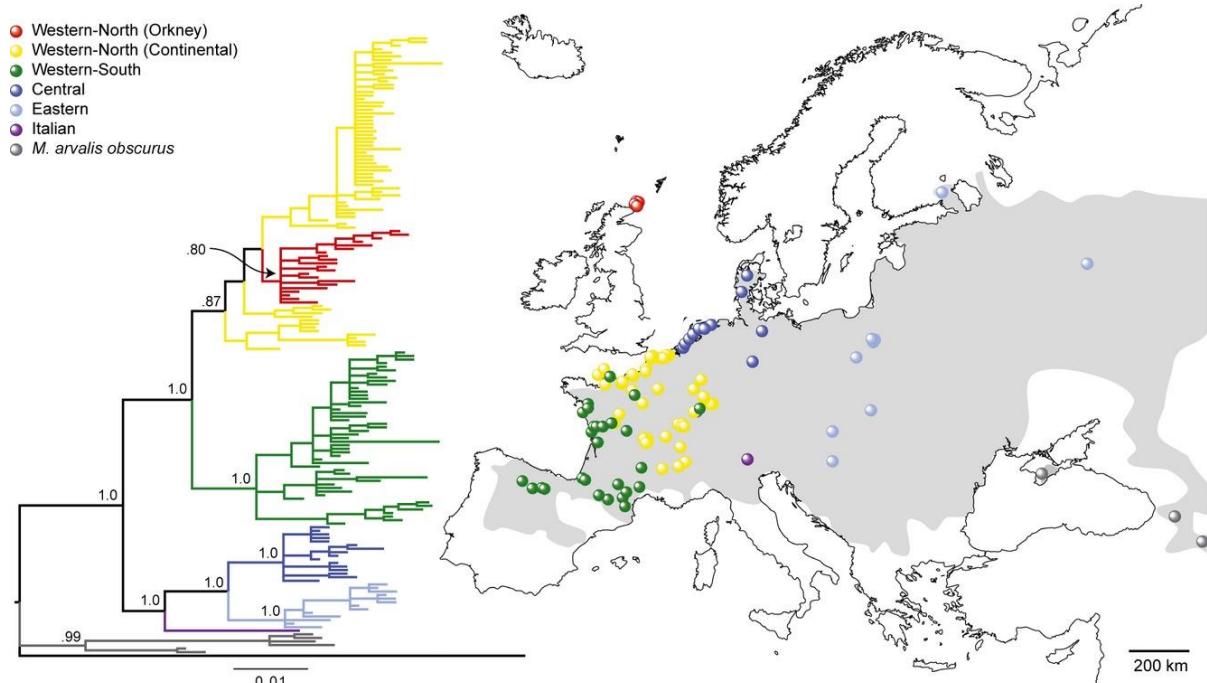
3: a

21: a

### Opgave 3 (20%)

Udbredelsen af Sydmarkmusen *Microtus arvalis* fremgår af det grå område i den følgende figur. Det er overraskende at finde arten på Orkney ("Western-North (Orkney)"), når den ellers ikke findes i Storbritannien.

For at forstå denne disjunkte udbredelse blev DNA fra 10-26 individer i en række populationer indsamlet (på de med farver angivne lokaliteter i figuren). Nogle få individer per lokalitet blev sekventeret for *cytb*-genet og disse sekvenser blev brugt til en fylogenetisk analyse. I træet er grenlængden et udtryk for genetisk afstand og alle grene med høj statistisk sikkerhed har et tal associeret med sig.



Alle indsamlede individer i 10 udvalgte populationer blev genotypt i en række neutrale loci (mikrosatellitter) og disse data blev brugt til at beregne en forventet heterozygoti ( $H_e \pm$  standardfejl den følgende tabel).

Population	Land	$H_e (\pm SE)$
1	Belgien	0,72 (0,15)
2	Frankrig	0,75 (0,17)
3	Frankrig	0,73 (0,25)
4	Tyskland	0,80 (0,09)
5	Frankrig	0,70 (0,23)

6	Tyskland	0,74 (0,11)
7	Holland	0,75 (0,23)
8	Orkney (UK)	0,23 (0,26)
9	Orkney (UK)	0,21 (0,26)
10	Orkney (UK)	0,19 (0,29)

- a) I populationsgenetikken beregner vi ofte to typer af heterozygoti: observeret heterozygoti ( $H_o$ ) og forventet heterozygoti ( $H_e$ ). Forklar forskellen mellem de to typer med dine egne ord.

$H_o$  er den observerede frekvens af heterozygoter i et givet locus.

$H_e$  beregnes fra allelfrekvenser som den forventede frekvens af heterozygoter under antagelse af Hardy-Weinberg proportioner.

- b) Brug træet til at give et bud på en koloniseringsvej for *M. arvalis* til Orkney.

Orkney udgør monofyletisk gruppe som er en del af en større monofyletisk gruppe (gule grene i træet), der kun inkluderer 'western north – continental'. Altså solid støtte til en hypotese om at musen har koloniseret Orkney fra kontinentet (Frankrig/Belgien/Holland)

- c) Brug træet og  $H_e$  til at diskutere hvilke evolutionære processer, der har haft betydning for niveauet af den genetiske variation i populationerne på Orkney.

Founder effekt– derfor lav genetisk variation ( $H_e$  er lav).

- d) Brug træet og/eller  $H_e$  til at diskutere, hvorvidt koloniseringen af Orkney er sket for nyligt eller længere tid siden.

Den røde gruppe i træet er monofyletisk og findes udelukkende på Orkney. Der må været gået nogen tid for at få mutationer til at danne de observerede sekvenser.

#### Opgave 4 (20 %)

Grevy's zebra (*Equus grevyi*), afbildet i figuren til højre, er én af tre eksisterende zebraarter. De to andre er steppezebra (*Equus quagga*), der har en større udbredelse (fra Etiopien i nord til Sydafrika i syd) og bjergzebraen (*Equus zebra*), der kun findes i Namibia og Sydafrika. Grevy's zebra findes i Etiopien og Kenya. Den har været igennem en drastisk tilbagegang og er i kategorien "truet" ifølge IUCNs rødliste. I 1970erne fandtes der ca. 16.000, men i dag er der kun cirka 2500 tilbage.



I 2009 publicerede Cordingley et al. en artikel med titlen "Is the endangered Grevy's zebra threatened by hybridization?" i *Animal Conservation* 12, 505–513. De havde i en sympatrisk bestand af steppezebra og Grevy's zebra observeret hybridiseringer mellem de to arter. Som titlen ganske rigtigt antyder kan hybridisering være en trussel mod sjældne arter,

hvis de i høj grad udsættes for hybridisering med en art, der er mere almindelig. De analyserede mitokondrie og Y-kromosom DNA-sekvenser.

Det fylogenetiske træ over mitokondriesekvenser fra det undersøgte område i Kenya viser slægtskabet mellem Grevy's zebra, steppezebra samt hybriderne. Hybriderne er fremhævet med **fed**. I alt blev fundet 10 hybrider mellem de to arter. Hanlige hybrider fik også analyseret deres Y-kromosomer. Det viste sig, at de alle bar på en haplotype, der grupperede dem sammen med Grevy's zebra.

- a) Giv et bud på hvilket køn forældrene til hybriderne havde.

Alle hybrider har mitokondriesekvenser, der grupperer dem med steppezebraer. Deres mor må have været en steppezebra. Samtidigt har alle hannerne modtaget et Y-kromosom fra Grevy's zebra. Det tyder på, at i det mindste alle hanlige hybrider har haft en Grevy's zebra som far. Det er nærliggende at tro, at fædrene har været Grevy's zebra.

- b) Baseret på det fylogenetiske træ bedes du besvare: Hvilken art har haft den største effektive populationsstørrelse for hunner?

Der observeres stort set ingen variation i mitokondriesekvenser hos Grevy's zebra i modsætning til steppezebraen. Steppezebraen må have haft en betydelig højere effektiv populationsstørrelse for hunner.

- c) Alle hybrider blev observeret i flokke af steppezebraer. Giv dit bud på, hvilken vej genflowet mellem de to arter har.

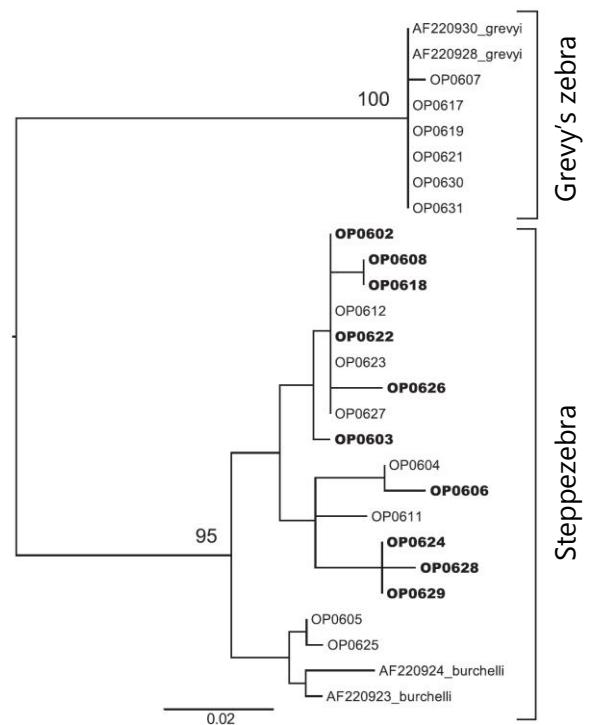
Genflowet ser ud til at gå fra Grevy's zebra og ind i steppezebra.

- d) Vurder, hvorvidt hybridisering er en trussel mod Grevy's zebra. Giv dit bud på, hvordan man bør håndtere de to lokale bestande ud fra et conservation synspunkt.

Hybridisering ser ikke ud til at være en trussel mod Grevy's zebra. Det ser mere ud til at være tilfældet for den undersøgte population af steppezebra.

- e) Cordingley et al. benyttede udelukkende mitokondrie og Y-kromosomal variation. Er det tilstrækkeligt til også at analyse tilbagekrydsninger? Hvilke genetiske data ville du inkludere for at undersøge dette? (Besvares **kort**)

Man burde have undersøgt autosomale loci (flere). mt og Y-kromosom fortæller kun



noget om, hvor forældrenes mt og Y-kromosom kommer fra.

Den følgende tabel viser kromosomantal (2n) i hestefamilien

Art	Antal
<i>Equus przewalski</i> (vildhest, forvildet)	66
<i>E. caballus</i> (tamhest)	64
<i>E. asinus</i> (æsel)	62
<i>E. onager</i> (halvæsel)	56
<i>E. kiang</i> (halvæsel)	52
<i>E. grevyi</i> (Grevy's zebra)	46
<i>E. quagga</i> (Steppezebra)	44
<i>E. zebra</i> (bjergzebra)	32

- f) Hvordan er fertiliteten af en krydsning mellem hest og æsel? Hvilken form for isolationsmekanisme er der tale om? Tyder denne isolationsmekanisme på, at arterne er dannet sympatrisk eller allopatrisk? (Begrundes kort)

Hest og æsel har forskelligt antal kromosomer. Hybrider, muldyr og mulæsel, er sterile. Det er postzygotisk isolationsmekanisme. Artsdannelsen er formentlig sket allopatrisk.

- g) Hvad forventer du, a fertiliteten for en krydsning mellem en Grevy's zebra og steppezebra er?

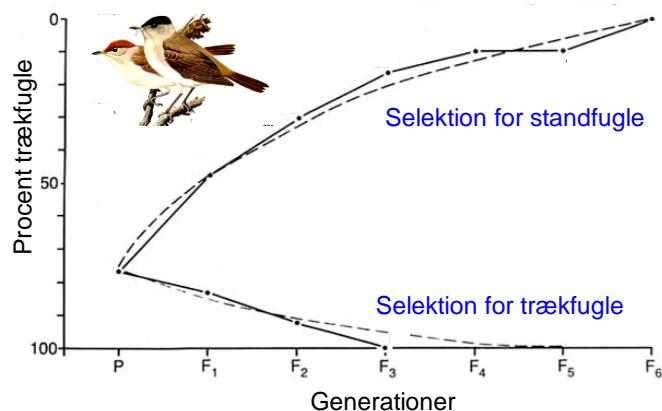
Da de har forskelligt antal kromosomer, forventer vi ligeledes at de har en nedsat fertilitet.

- h) Hvis der sker hybridisering mellem arter, hvor hybrider har nedsat fertilitet, hvilken isolationsmekanisme vil så over tid blive øget gennem naturlig selektion?

Præzygotisk isolationsmekanisme.

### Opgave 5 (10%)

Munken (*Sylvia atricapilla*) har populationer i Tyskland, hvor nogle individer er standfugle og andre individer er trækfugle. Man har lavet forsøg i en af disse populationer, hvor man har selekteret for større og mindre "trækuro". Dette er et mål for hvorvidt en fugl ender som standfugl eller som trækfugl. Den population, man undersøgte, havde cirka 75%, der var trækfugle, og 25%, der var standfugle. Man lavede to forsøg: Et, hvor man selekterede for stigende trækuro og et, hvor man selekterede for lav trækuro. De to forsøg er illustreret i



figuren. (Se bort fra de stiplede kurver. Data for de forskellige generationer er angivet med prikker og er forbundet med linjer.)

- a) Hvilken form for selektion er der tale om?

**Retningsselektion for henholdsvis øget og reduceret trækuro.**

- b) Hvilke andre former for selektion har vi på kvantitative karakterer?

**Stabiliserende selektion og disruptiv selektion**

- c) Vurderer du at den additive varians for trækuro er høj eller lav baseret på de to forsøg?

Den må være forholdsvis høj. Ellers ville man ikke kunne komme så langt i et selektionsforsøg. (Lærebogen viser en figur over regression af afkom på gennemsnit af forældres trækuro, der viser at heritabiliteten er 0.45.)

- d) I forsøgene med munken ser vi, at selektion for trækfugle endte med 100% trækfugle efter tre generationers selektion. Er der så stadig genetisk variation for karakteren i populationen? Begrundes **kort**.

**Ja, det er bare den additive genetiske varians,  $V_A$ , der er 0.**

### Opgave 6 (20%)

- a) Hvilke data kan anvendes til fylogeni

- 1) fossiler
- 2) morfologi
- 3) gensekvenser
- 4) alle de nævnte typer af data**

- b) Genetisk drift medfører

- 1) stigende indavl
- 2) tab af genetisk variation
- 3) differentiering mellem populationer
- 4) alle de nævnte ting**

- c) Hvad er indavlskoefficienten F et udtryk for?

- 1) sandsynligheden for at et individ er homozygot
- 2) sandsynligheden for at et individ bærer to gener, der er identiske gennem arv**
- 3) sandsynligheden for at et individ bærer to gener, der er ens
- 4) alle de nævnte muligheder

- d) Hvad er en substitution?

- 1) et andet udtryk for mutation
- 2) når en purin erstattes med en pyrimidin eller omvendt
- 3) når en allele erstattes fuldstændig af en anden i en population eller art over evolutionær tid**
- 4) når en plesiomorf karakter bliver til en apomorf karakter

e) For en karakter med en høj heritabilitet:

- 1) er forskelle mellem populationer genetisk betingede
- 2) er miljøvariansen  $V_E$  meget stor
- 3) er en stor del af populationens fænotypiske varians genetisk betinget**
- 4) gælder alle tre udsagn

f) En adaptation:

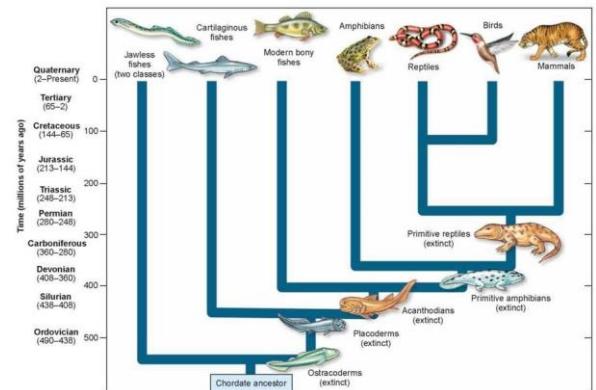
- 1) skyldes negativ selektion
- 2) er en egenskab, der øger et individs overlevelse eller reproduktion**
- 3) opträder, når genetisk drift får en population ned igennem en fitness-dal
- 4) er resultatet af genetisk hitchhiking

g) Figuren viser en fylogeni over chordata. Hvem er ifølge denne fylogeni tættest beslægtet med amfibier? (Der skal kun kigges på nulevende dyr.)

- 1) benfisk
- 2) reptiler
- 3) reptiler, fugle og pattedyr**
- 4) benfisk, reptiler, fugle og pattedyr

h) En nukleotid substitution i et protein-kodende gen som forårsager en ændring af proteinsekvensen kaldes:

- 1) Translokation
- 2) Synonym substitution
- 3) Ikke-synonym substitution**
- 4) Tilbagemutation
- 5) En duplikation



i) Hvilke(n) af følgende processer bidrager med ny variation?

- 1) Rekombination**
- 2) Genetisk drift
- 3) Mutation**
- 4) Selektion
- 5) Indavl

j) Udviklingen af samme karaktertilstand hos forskellige uafhængige evolutionære linjer kaldes

- 1) dominans
- 2) konvergens**
- 3) homologi
- 4) founder effect
- 5) naturlig selektion

k) Mennesket og dets nærmeste uddøde (fossile) slægtninge kaldes homininer. Hvilken blandt homininerne kunne man med rette kalde for den første "værktøjs-mand"?

- 1) *Homo erectus***
- 2) *Homo sapiens***
- 3) *Homo habilis***

4) *Homo neanderthalensis*

l) Hvilken er størst

- 1) Broad sense heritabilitet
- 2) Narrow sense heritabilitet

m) I humane populationer vil der som funktion af deres afstand fra Afrika gælde for gendiversitet og linkage disequilibrium (LD), at

- 1) gendiversitet stiger og linkage disequilibrium (LD) stiger
- 2) gendiversitet stiger og linkage disequilibrium (LD) aftager
- 3) gendiversitet aftager og linkage disequilibrium (LD) stiger
- 4) gendiversitet aftager og linkage disequilibrium (LD) aftager

n) Hvilken proces indgår *ikke* i horizontal gen transfer hos prokaryoter?

- 1) conjugation
- 2) transduktion
- 3) inversion
- 4) transformation

o) De proteinkodende gener i det menneskelige genom udgør den følgende procentdel

- 1) 15%
- 2) 1.5%
- 3) 25%
- 4) 50%

p) Eukaryote genomer er ofte opdelt i haplotype blokke. For disse gælder, at linkage disequilibrium (LD) er

- 1) høj indenfor en haplotype blok, høj imellem haplotype blokke
- 2) høj indenfor en haplotype blok, lav imellem haplotype blokke
- 3) lav indenfor en haplotype blok, lav imellem haplotype blokke
- 4) lav indenfor en haplotype blok, høj imellem haplotype blokke

q) I en pattedyrpopulation er den effektive populationsstørrelse  $N_e$ . Der er det samme antal hunner og hanner i populationen. Hvad er den effektive populationsstørrelse for mitokondrier?

- 1)  $N_e$
- 2)  $2N_e$
- 3)  $N_e/2$
- 4)  $N_e/4$

## **Evolutionsbiologi**

Skriftlig eksamen, den 23. juni 2021

**Alle hjælpemidler tilladt  
(Lærebog, noter, computer)**

For hver opgave oplyses dens vægtning.

Dette opgavesæt indeholder 13 sider inklusive forsiden  
*Opgaverne skal afleveres i den rækkefølge, de er stillet i.*

Censorer:

Tove Hedegaard Jørgensen

Jesper Givskov Sørensen

Eksaminatorer:

Rasmus Heller

Hans R. Siegismund

**Eksamensvejledning**

Eksamenssættet udleveres som et word-dokument. I skal skrive jeres besvarelse ind i Word-dokumentet (benyt gerne en anden farve end i det udleverede dokument) og uploadet det på Digital Eksamen.

I forventes at lave eksamensbesvarelsen selvstændigt, det vil sige, at I sidder alene i et rum og ikke kommunikerer med andre.

### Opgave 1 (20 %)

Den vilde hvede *Avena barbata* er udbredt omkring Middelhavet. Arten er selvkompatibel, det vil sige, at den kan selvbestøve. Den er blevet indført til Nordamerika, hvor den har spredt sig som invasiv art, specielt i Californien. Her har den fortrængt en del hjemmehørende græsser. Marshal og Allard har undersøgt genetisk variation i en række enzymloci. I et locus fandt de den følgende genetiske fordeling:



Genotype	11	12	22	Sum
Antal	31	6	50	87

- a) Bestem frekvenserne af de to alleler.

$$p = (2 \times 31 + 6)/(2 \times 87) = 0,391$$
$$q = (2 \times 50 + 6)/(2 \times 87) = 0,609$$

- b) Er der Hardy-Weinberg proportioner? (Benyt  $\chi^2$ -test.)

Nej

Obs 31	6	50
Exp 13,29	41,43	32,29

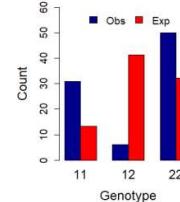
$$\chi^2 = 63,62 P = 1,51\text{e-}15$$

- c) Beregn indavlskoefficienten  $F$ . Hvordan afviger de observerede genotyper fra de forventede ifølge Hardy-Weinberg proportioner?

$$F = (41,43 - 6)/41,43 = 0,855$$

Der er overskud af homozygoter i forhold til Hardy-Weinberg proportioner

- d) Hvad kan årsagen til en eventuel afvigelse være?



Afgelsen skyldes, atarten som nævnt er delvist selvbestøvende. Selvbestøvning medfører et overskud af homozygoter i forhold til Hardy-Weinberg proportioner.

*Avena barbata* har et kromosomsæt på  $2n = 28$ . Dens to nærmeste slægtninge er *Avena hirtula* og *Avena wiestii*, der begge har kromosomsæt på henholdsvis  $2n = 12$  og  $2n = 16$ .

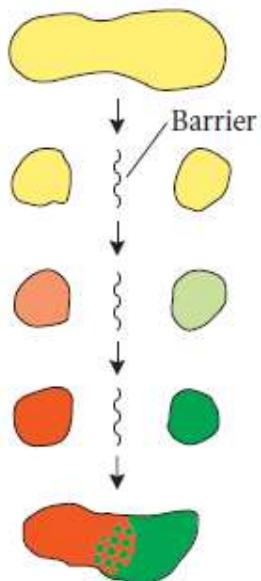
- e) Hvordan er de tre arter relateret til hinanden. Mere specifikt, giv et bud på, hvordan *Avena barbata* er opstået som art.

*Avena barbata* er opstået ved polyploid arts dannelse mellem de to diploide arter *Avena hirtula* og *Avena wiestii*.

- f) Hvilken arts dannelse anses for den mest udbredte hos dyr? Forklaries **kort**.

### Allopatrisk artsdannelse

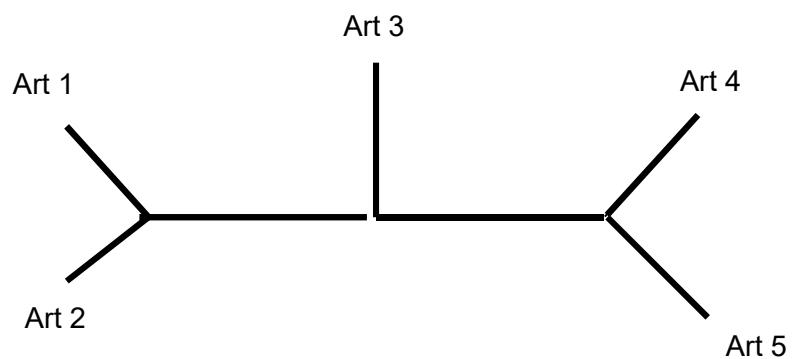
Allopatric speciation  
by vicariance



Barrier removed  
or new species  
disperse over it,  
re-establishing  
sympathy

### Opgave 2 (20 %)

Betrægt den følgende fylogeni:



- a) Hvilke monofyletiske grupper findes i dette træ, hvis roden er placeret på den gren, der leder til art 5? (Medregn ikke grupper, der kun består af én art eller af alle 5 arter)?

(1, 2), (1, 2, 3), (1, 2, 3, 4).

- b) Baseret på parsimoniprincippet skal antallet af mutationer udregnes på dette træ for følgende sekvenser:

Art 1 TCTCAGACCAGATGGCGGATGGCG  
 Art 2 ACTCAGACCTGATGGCGGATGGCA  
 Art 3 ACTCAGACCAGATGGCAGATGGCA  
 Art 4 ACCCAGACCTGATGGCAGATGGCG  
 Art 5 ACCCAGACCAGATGGCAGATGGCA

7

- c) Hvor mange transitioner og transversioner adskiller sekvenserne fra art 1 og art 2?

2 transversioner, 1 transition

- d) Baseret på parsimoniprincippet udregn antallet af transitioner og transversioner i de fem ovenstående sekvenser baseret på fylogenien afbildet foroven, hvis man placerer roden på den gren, der leder til art 5.

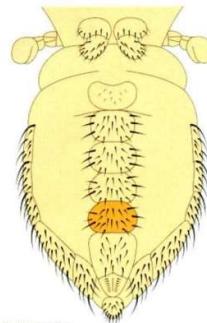
3 transversioner, 4 transitioner

- e) Hvis roden i stedet var placeret på den gren som leder til art 3, hvilke monofyletiske grupper ville da kunne findes i træet?

(1, 2), (4, 5), (1, 2, 4, 5).

### Opgave 3 (20 %)

Fra en meget stor population af *Drosophila melanogaster* udtagtes en stikprøve, der bruges som første forældregeneration i et selektionsforsøg for antal børster på en abdominalplade. Gennemsnittet af børster i stikprøven er 20. Man udvælger de 20 fluer (10 hanner og 10 hunner), der har det højeste antal børster, i gennemsnit 28, og benytter disse som forældre (Generation 0).



- a) Hvor stort er selektionsdifferentialt?

$$S = 28 - 20 = 8.$$

- b) Afkommet (generation 1) har i gennemsnit 24 børster. Benyt denne oplysning til at beregne responset på selektionen.

$$R = 24 - 20 = 4.$$

- c) Hvor stor er heritabiliteten for denne karakter?

$$R = h^2 S. \rightarrow h^2 = R/S = 4/8 = 0,5$$

- d) Man fortsætter forsøget igennem 10 generationer. Ud fra responset i generation 1, hvad er så den maksimalt mulige middelværdi for antal børster, som populationen forventes at have når forsøget slutter? (Besvares **kort**)

Den samlede respons er  $10 \times 4 = 40$ . Derfor forventes middelværdien at ændre sig fra 20 til 60.

- e) Hvilken antagelse er dette skøn baseret på? (Besvares **kort**)

Det forudsætter at heritabiliteten er konstant igennem de 10 generationer.

- f) Det oprindelige eksperiment fortsættes igennem mange generationer. Hvad forventes heritabiliteten at blive til sidst?

$h^2 = 0$ , da responsen til sidst er 0. Dvs, at der ikke er additiv genetisk varians tilbage.

#### Opgave 4 (20 %)

Den almindelige eland antilope (*Taurotragus oryx*) er den største af alle antiloper og findes vidt udbredt på savannerne i det østlige og sydlige Afrika. I en undersøgelse af variation i det mitokondrielle d-loop af Lorenzen *et al.* (2010) indsamledes der bl.a. 46 prøver fra Tanzania. Her er kun angivet variationen i to basepositioner, der sidder tæt sammen (med to baser ind imellem sig)

$$f_{TT} = 22/46 = 0,48$$

$$f_{TC} = 0/46 = 0$$

$$f_{CT} = 20/46 = 0,43$$

$$f_{CC} = 4/46 = 0,09$$



Her repræsenter  $f_{XY}$  variationen i den første position (X) og den anden position (Y) i sekvensen.

- a) Beregn frekvenserne af de to baser i de to positioner. Lad  $p_T$  og  $p_C$  være frekvenserne i position 1 og lad  $q_T$  og  $q_C$  være frekvenserne i position 2.

$$p_T = (22 + 0)/46 = 0,48$$

$$p_C = (20 + 4)/46 = 0,52$$

$$q_T = (22 + 20)/46 = 0,91$$

$$q_C = (0 + 4)/46 = 0,09$$

- b) Beregn koblingsuligevægten  $D$  mellem de to positioner.

$$f_{TT} = p_T \times q_T + D$$

↓

$$D = f_{TT} - p_T \times q_T$$

$$= 0,4783 - 0,4783 \times 0,9130$$

$$= 0,0416.$$

Koblingsulige vægtens størrelse er bestemt af frekvenserne af de to alleler i de to loci (her basepositioner). Det kan vises at den maksimalt kan være  $p_T \times q_C$ .

- c) Er den fundne værdi for  $D$  ekstrem?

$$\begin{aligned} p_T \times q_C &= 0,4783 \times 0,0870 \\ &= 0,0416 \end{aligned}$$

Den fundne værdi er lig med den maksimalt mulige. Den må siges at være ekstrem.

- d) Hvad kan forklaringen være?

Årsagen er manglende rekombination i det mitokondrielle genom.

- e) Der er kun fundet tre haplotyper. Hvor mange mutationer skal der som minimum til for at forklare det?

To mutationer. Vi kender ikke de oprindelige sekvenser, men det kunne se ud som følgende:

$TT \rightarrow CT \rightarrow CC$ .

- f) Hvilke evolutionære processer kan i fremtiden være med til at danne den fjerde haplotype,  $f_{TC}$ ?

Den fjerde haplotype kunne opstå enten ved rekombination eller ved en tredje mutation.

### Opgave 5 (20%)

- a) Antag, at en founder population har en allelfrekvens, der afviger fra den population, som den kommer fra. Hvilken af de følgende effekter vil sandsynligvis føre til genetisk homogenisering mellem den oprindelige population og founder populationen?

- 1) genetisk drift i founder population
- 2) tilfældig parring i founder populationen
- 3) ingen mutationer i begge populationer
- 4) migration mellem den oprindelige population og founder populationen
- 5) Intet af ovenstående

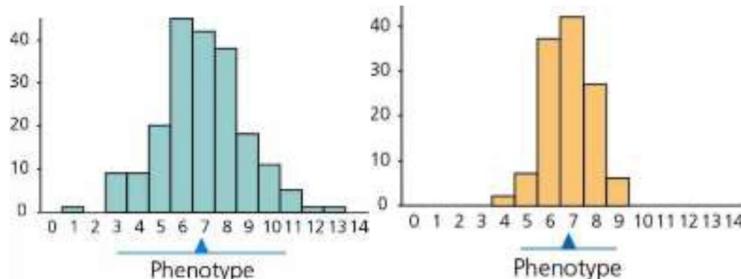
- b) Når genotyper i et locus er uafhængige af genotyper i et andet locus, så er de

- 1) i koblingsulige vægt (linkage equilibrium)
- 2) i koblingsulige vægt (linkage disequilibrium)
- 3) ikke koblede
- 4) Intet af ovenstående.

c) Hvilken af de følgende evolutionære processer kan producere koblingsulige vægt (linkage disequilibrium)?

- 1) genetisk drift
- 2) naturlig selektion
- 3) blanding af to forskellige populationer
- 4) alle de ovennævnte processer

d) Betragt nedenstående figurer. Figur A viser en kvantitativ karakter i en population på tidspunktet 0, og figur B repræsenterer det samme karakter nogle generationer senere. Hvilken form for selektion har da virket?

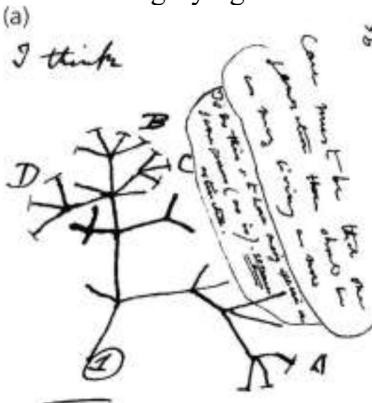


Figur A

Figur B

- 1) Stabiliserende selektion
- 2) Disruptiv selektion
- 3) Retningsselektion
- 4) Frekvensafhængig selektion

e) Den følgende figur repræsenterer et tidligt fylogenetisk træ som er tegnet af



- 1) Mendel
- 2) Wallace
- 3) Lamarck
- 4) Darwin

f) En ny afledt karakter kaldes også en

- 1) apomorfi
- 2) plesiomorfi
- 3) synapomorfi

4) klade

g) En gruppe, der består af en fælles forfader og nogle af, men ikke alle, dens efterkommere kaldes en

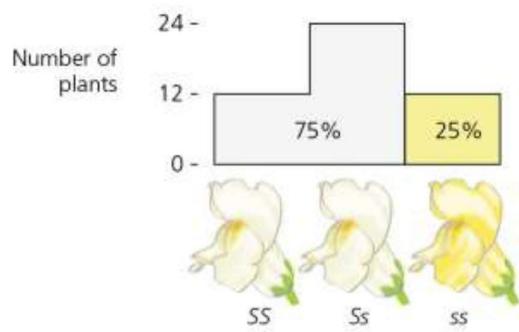
- 1) monofyletisk gruppe
- 2) parafyletisk gruppe**
- 3) polyfyletisk gruppe
- 4) polytomi gruppe

h) Mere end 45 procent af det menneskelige genom er sammensat af

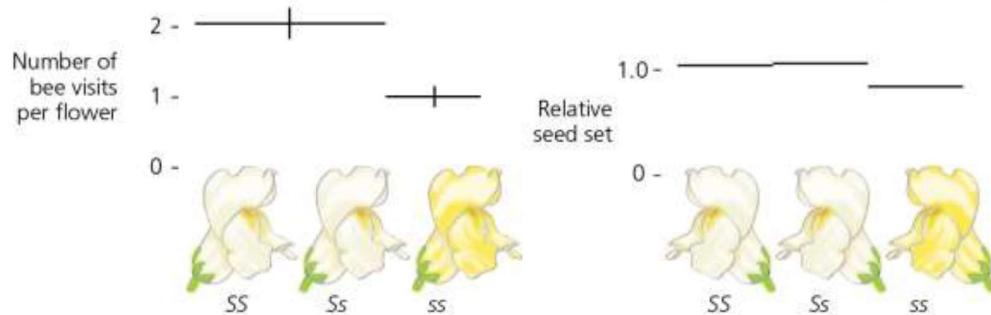
- 1) repeterede sekvenser**
- 2) exons
- 3) introns
- 4) heterochromatin
- 5) mikrosatellitter

i) Den følgende figur viser forsøg med løvemunden, hvor gul blomsterfarve er recessiv overfor hvid blomsterfarve. Hvad er det mest sandsynlige resultat?

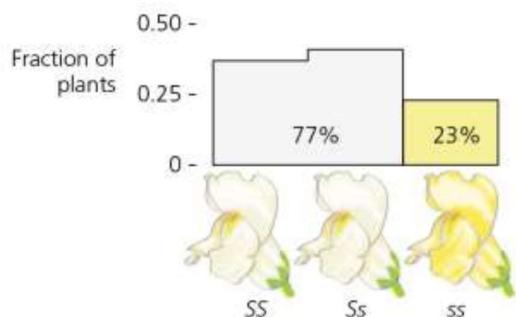
(a) Composition of parental population



(b) Differences in reproductive success through male function (left) and female function (right)

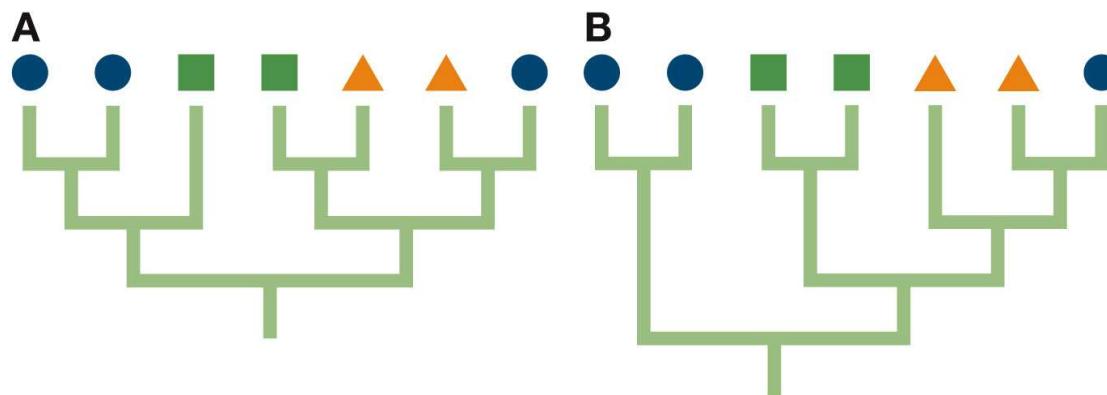


(c) Composition of offspring population



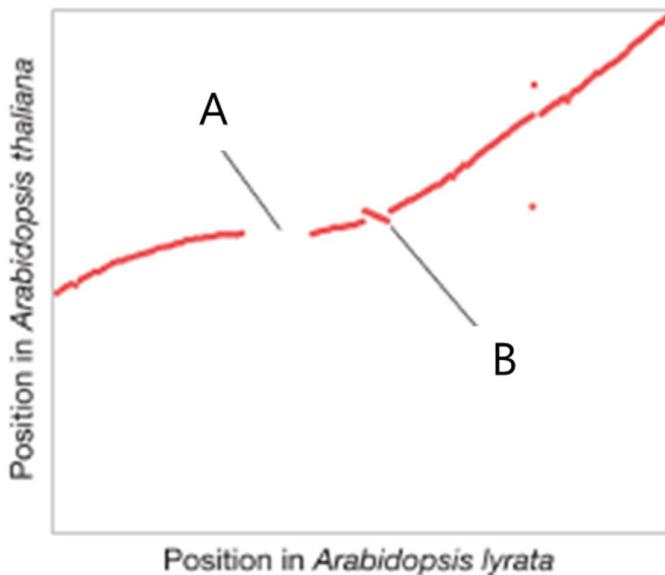
- 1) En ændring på 2% er for lille til at være signifikant, så procentdelen af gule og hvide blomster forbliver nogenlunde lige i fremtidige generationer.
  - 2) Gule og hvide blomster svinger i frekvenser afhængigt af miljøet.
  - 3) Hvide blomster vil fortsætte med at stige i hyppighed sammenlignet med gule blomster.
  - 4) Fordi det gule træk er recessivt, vil dets hyppighed forblive omkring 0,25.
- j) Hvilket af de følgende udtryk repræsenterer den fænotypiske varians  $V_P$  i en kvantitativ karakter?
- 1)  $V_G + V_A$
  - 2)  $V_A + V_E$
  - 3)  $V_G + V_E$
  - 4)  $V_G + V_M$
  - 5)  $V_M + V_E$

.k) Hvilket af de følgende to fylogenetiske træer er det mest parsimoniske?



- 1) A  
2) B
- l) Hvordan adskiller eukaryote gener sig fra generne fra bakterier og archaea?
- 1) Gener hos bakterier og archaea kodes af RNA, ikke DNA.
  - 2) **Eukaryote gener indeholder ofte introns.**
  - 3) Eukaryote gener er meget mindre.
  - 4) Eukaryote gener mangler cis-regulerende elementer.
  - 5) Intet af det ovenstående; eukaryote gener adskiller sig ikke fra generne fra bakterier og archaea.
- m) Virus genomer er meget små, på grund af
- 1) codon bias.
  - 2) at vira har først udviklet sig for nyligt.
  - 3) at små genomer er en tilpasning, der muliggør hurtigere kopiering.
  - 4) at mange vira anvender deres værtscelles molekulære maskineri.
  - 5) **Både 3 og 4**
- n) Hvilken af de følgende hypoteser kan forklare, hvorfor forskellige gener kan understøtte forskellige fylogenier?
- 1) Variationen af mutationsrate på tværs af genomet
  - 2) **Incomplete lineage sorting**
  - 3) Variation af rekombinationsrate på tværs af genomet
  - 4) Variation i effektiv populationsstørrelse blandt forskellige arter
  - 5) Alternativ splejsning
- o) Konjugation
- 1) skyldes optagelse af frie DNA fragmenter
  - 2) **er et eksempel på horizontal gentransfer**
  - 3) involverer bakteriofager
  - 4) foregår mellem bakterier og bakteriofager

p) Figuren nedeunder sammenligner kromosomer fra to arter af *Arabidopsis* i et såkaldt "syntenic dot plot"



De evolutionære hændelser, der er sket i A og B er

- 1) A: Inversion, B: inversion
- 2) A: Inversion, B: (deletion eller insertion)
- 3) A: (Deletion eller insertion), B: inversion**
- 4) A: (Deletion eller insertion), B: (Deletion eller insertion)

q) I et locus findes to alleler (A og B). Under antagelse af der ikke findes nogen stabiliserende kræfter, vil A-allelen i sidste ende blive

- 1) fikseret.
- 2) enten tabt eller fundet.
- 3) bevaret.
- 4) enten fikseret eller tabt.**

r) Den additive genetiske varians af strandkrabbenes klos længde er 0,50, mens miljøvariansen er 0,30. Hvad er heritabiliteten for denne karakter? (Der er ingen andre komponenter, der bidrager til den fænotypiske varians.)

- 1) 1.667
- 2) 0,600
- 3) 0,625**
- 4) 0,200
- 5) 0,500



s) Hvis der observeres to arter af cichlider i den samme sø, siges de at være \_\_\_\_\_. Hvis to fuglearter forekommer på separate kontinenter, siges de at være \_\_\_\_\_.

- 1) sympatriske; allopatriske**
- 2) parapatriske; parapatriske

- 3) parapatriske; allopatriske
  - 4) allopatriske; sympatriske
  - 5) allopatriske; parapatriske
- t) En afledt karakter som deles af to evolutionære linjer kaldes også en
- 1) apomorfi
  - 2) plesiomorfi
  - 3) synapomorfi
  - 4) klade

## **Evolutionsbiologi**

Skriftlig eksamen, den 25. juni 2022

**Alle hjælpemidler tilladt**

For hver opgave oplyses dens vægtning.

Dette opgavesæt indeholder xx sider inklusive forsiden  
*Opgaverne skal afleveres i den rækkefølge, de er stillet i.*

Censorer: Tove Hedegaard Jørgensen Jesper Givskov Sørensen  
Eksaminatorer: Rasmus Heller Hans R. Siegismund

## **Eksamensvejledning**

Eksamenssættet udleveres som et Word-dokument. I skal skrive jeres besvarelse ind i dokumentet (benyt gerne en anden farve end i det udleverede dokument) og uploadet det på Digital Eksamen.

### Opgave 1 (30 %)

En ung uerfaren biolog<sup>1</sup> har indsamlet data for den humane SNP rs16891982 for europæere (EUR) og afrikanere (AFR) fra flere forskellige lande. Han fandt den følgende genotypetfordeling

	CC	CG	GG	Sum
AFR O	617	41	3	661
AFR E	614.84	45.33	0.84	
EUR O	4	54	445	503
EUR E	1.91	58.18	442.91	

- a) Beregn allelfrekvensen af C i de to grupper.

$$p(C, AFR) = 0.964$$

$$p(C, EUR) = 0.062$$

- b) Afviger genotypefrekvenserne fra Hardy-Weinberg proportioner i de to stikprøver? (Benyt  $\chi^2$  test og karakteriser en afvigelse ved hjælp af indavlskoefficienten  $F$ .)

$$AFR: \chi^2 = 6.03; \quad P = 0.014; \quad F = 0.096$$

$$EUR: \chi^2 = 2.60; \quad P = 0.107; \quad F = 0.072$$

Der er en signifikant afvigelse i den afrikanske population, og der er også et overskud af homozygoter i den europæiske population, hvilket dog ikke er signifikant.

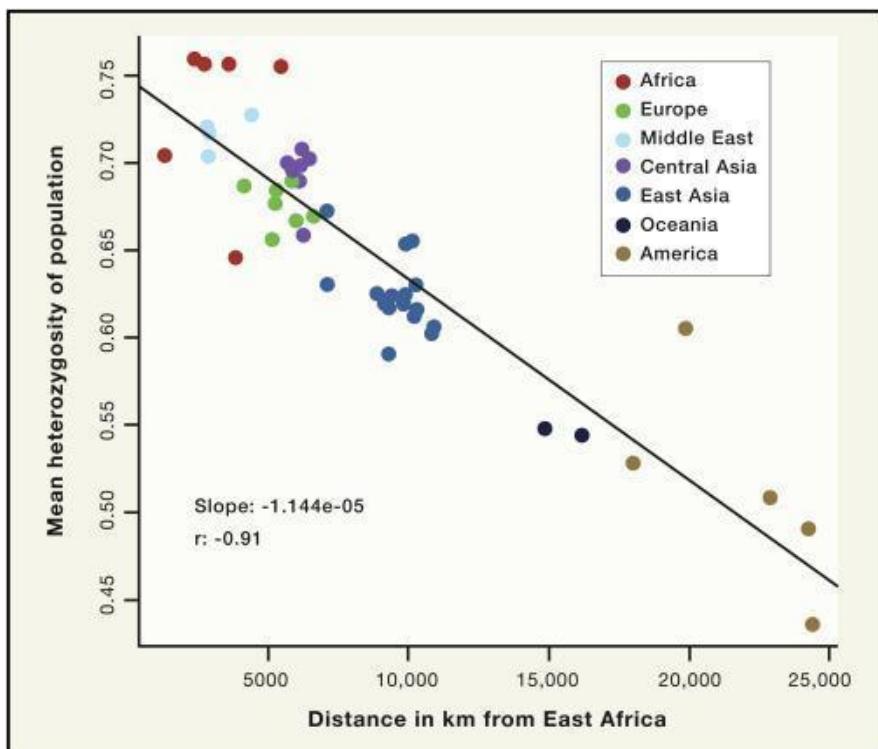
- c) Hvad kan have forårsaget en mulig afvigelse fra Hardy-Weinberg proportioner?

Begge stikprøver er en blanding af flere populationer (Wahlund effekt), som forårsager overskud af homozygoter.

Når man mäter genetisk diversitet hos forskellige menneskepopulationer får man den følgende sammenhæng mellem diversitet (Mean heterozygosity of population) i en population og dens afstand fra Østafrika

---

<sup>1</sup> Han havde ikke taget kurset i evolutionsbiologi endnu.



- d) Diversiteten aftager lineært som funktion af populationens afstand fra Østafrika. Hvad kan forklaringen for denne sammenhæng være?

Der har været gentagne founder effects, som har reduceret den genetiske diversitet. Effekten er størst for amerikanske populationer.

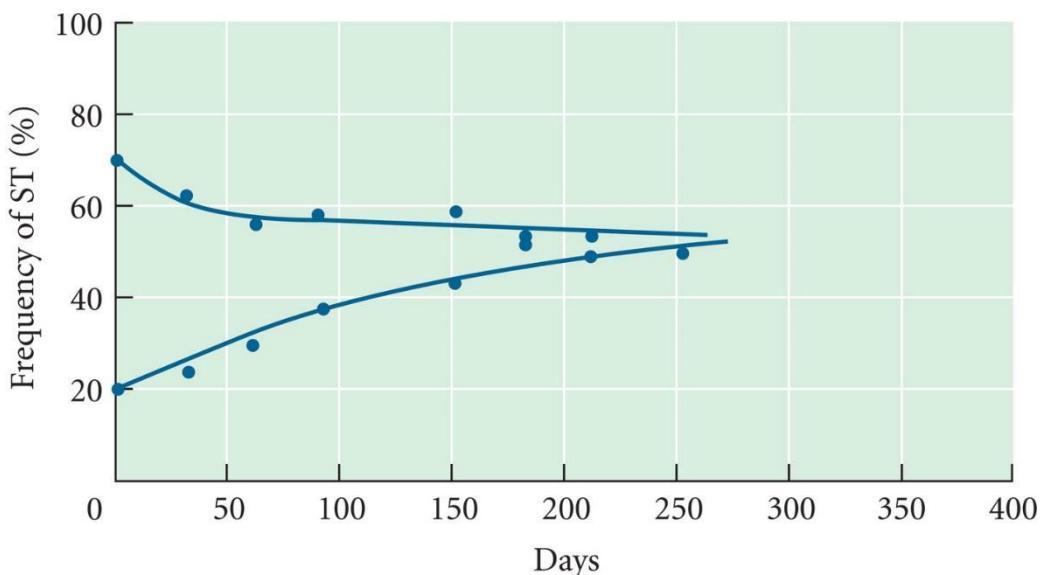
- e) I menneskepopulationer med meget lav genetisk diversitet, observeres det ofte at i gener, hvor der forekommer skadelige alleler, har disse alleler en forholdsvis høj hyppighed. Hvad skyldes dette?

Genetisk drift har været årsag til den forhøjede allelfrekvens. Når populationsstørrelsen bliver tilstrækkelig lille, kan genetisk drift “overdøve” naturlig selektion.

## Opgave 2 (30 %)

Dobzhansky lavede flere forsøg med kromosompolymorfier hos *Drosophila pseudoobscura*. Han benyttede to forskellige inversioner i sine forsøg, kaldet ST og AR. Den følgende figur viser frekvensen af kromosominversionen ST i to forsøg, der startede med forskellige frekvenser af ST inversionen til dag 0 i forsøget (ca. 0,20 og 0,70). Observationerne er angivet med prikker, medens de fuldt optrukne kurver er tilnærmelser.





- a) Beskriv forløbet af ST i de to forsøg.

I forsøget, der starter ved 0,7, falder frekvensen af ST jævnt til ca. 0,55. I forsøget, der starter ved 0,2, stiger frekvensen af ST jævnt til ca. 0,55.

- b) Hvad er forklaringen på forløbet i de to forsøg?

Det må være overdominant selektion (Eller stabiliserende selektion), med en stabil indre ligevægt.

- c) Nævn KORT forskellige typer af naturlig selektion i et enkelt locus samt hvorvidt der opretholdes genetisk variation i stabile ligevægte?

Overdominant selektion (dette tilfælde, variation opretholdes), retningsselektion (ingen variation), underdominant selektion (ingen variation), negativ frekvensafhængig selektion (variation opretholdes) og positiv frekvensafhængig selektion (ingen variation).

Lad fitnessværdierne af genotyperne ST/ST, ST/AR AR/AR være  $W_{ST/ST}$ ,  $W_{ST/AR}$  og  $W_{AR/AR}$ .

- d) Hvordan er de tre fitnessværdier relateret til hinanden i ovenstående forsøg?

$$W_{ST/ST} < W_{ST/AR}$$

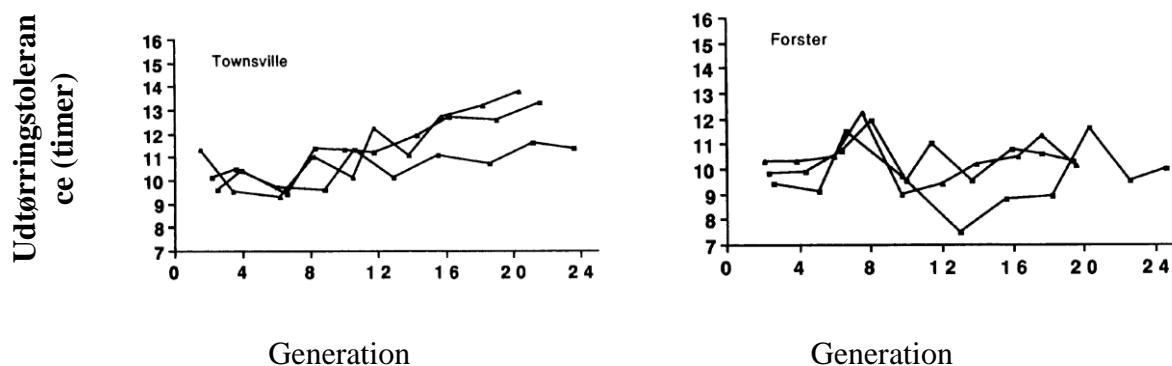
$$W_{AR/AR} < W_{ST/AR}$$

- e) Ligevægten ser ud til at være tæt på 0,5. Benyt dette til at udtale dig om hvor forskellige fitnessværdierne for de to homozygoter er.

Da ligevægtsfrekvensen er tæt på 0,5, må fitnessværdierne af de to homozygoter være næsten ens.

### Opgave 3 (vægter 20%)

Resultaterne fra en undersøgelse af udtørringstolerance i seks populationer af den tropiske bananflue *Drosophila serrata* er vist i nedenstående figur. Her er tre populationer fra artens centrale udbredelsesområde nær Townsville og tre populationer fra artens sydlige udbredelse på Australiens østkyst blev udsat for selektion for øget udtørringstolerance i op til 25 generationer.



- a) Beskriv helt kort resultaterne i figuren og diskutér hvad der kan forklare forskellen mellem populationer fra Townsville og populationer fra Forster.

Vi observerer konsistent respons til selektion i Townsville populationerne, mens der ikke er et tilsvarende respons i Forster. Dette må skyldes manglende additiv genetisk variation i Forster, som følge af denne populations lille populationsstørrelse, og formodentlig stærk drift.

De gennemsnitlige heritabiliteter ( $\pm$  standard fejl) for udtørringstolerance blev i forsøget estimeret til  $0,21 \pm 0,007$  (Townsville) og  $0,007 \pm 0,010$  (Forster).

- b) Beregn responsen til én generations selektion i en population fra Townsville og en population fra Forster, under antagelse af at populationernes gennemsnitlige udtørringstolerance er 10 timer, og de individer, der udvælges til at grundlægge næste generation, har en gennemsnits udtørringstolerance på 13 timer.

$R = h^2 \times S$ . Selektionsdifferentialet er  $S = 13 - 10$  timer = 3 timer. Responsen i Townsville er  $3 \text{ timer} \times 0,21 = 0,63$  timer højere tolerance. Responsen i Forster er  $3 \times 0,007 = 0,021$  timer.

- c) Beskriv de teoretiske forventninger for udfaldet af selektion for større udtørringstolerance, hvis den fortsætter i mange generationer endnu.

Under retningsselektion vil additiv genetisk varians blive brugt op, efterhånden som alleler der påvirker trækket fikses eller tabes. Derfor vil heritabiliteten og selektionsresponsen falde og der vil opnås et selektionsplateau.

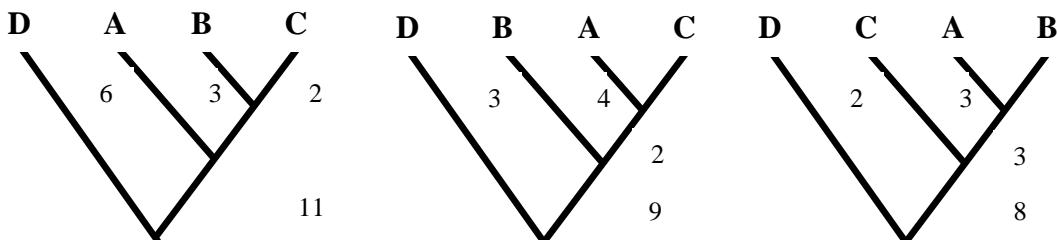
#### Opgave 4 (10%)

En biolog har undersøgt fire arter for tilstedeværelse af oprindelig (0) eller afledt (1) tilstand for 6 karakterer. Art D er en udgruppe til de tre andre arter.

	Karakter					
	1	2	3	4	5	6
Art A	1	1	1	1	1	1
Art B	0	1	0	1	0	1
Art C	1	0	1	0	0	0
Art D	0	0	0	0	0	0

- a) Tegn de mulige stamtræer for disse arter (Husk på, at D er en udgruppe.)
- b) Tæl antallet af ændringer i karaktertilstanden i de grene, hvor der er sket ændringer.

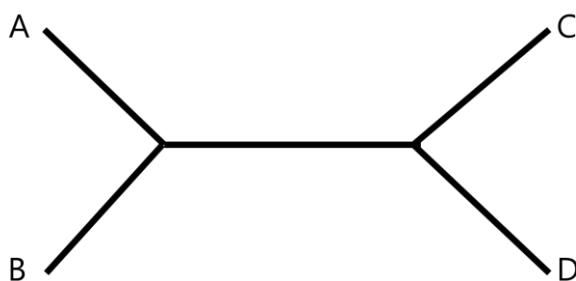
**a + b**



- c) Hvilket er det mest parsimoniske træ?

**Det til højre.**

- d) Hvordan ser et urodet træ ud for det mest parsimoniske træ?

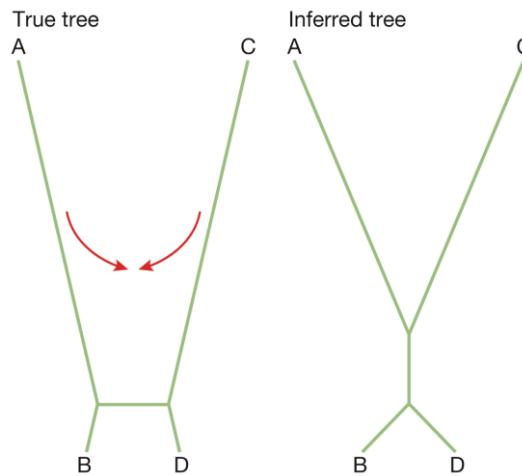


- e) Nævn nogle alternativer til parsimoni når man skal finde det bedst mulige træ baseret på karakterer i forskellige arter.

**Maximum likelihood, afstandsbaseret, Bayesiansk**

- f) En af metoderne kan vises at være ikke-konsistent. Dette betyder, at jo mere data man har, jo mere sandsynligt er det at få et forkert træ. Hvilken metode er det, og under hvilke betingelser sker det?

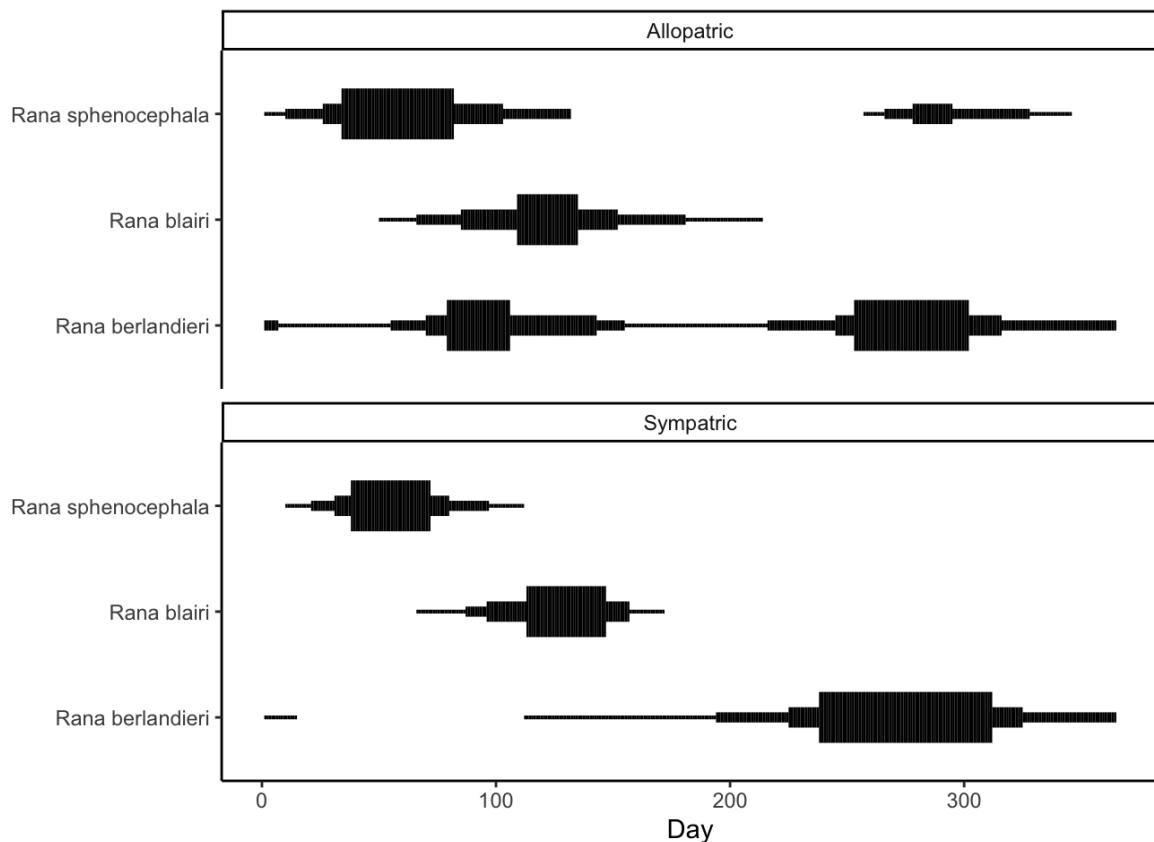
Parsimony: hvis evolutionære ændringer har forskellig rate i forskellige dele af træet, kan grene med høje rater ofte beregnes til at have en fælles oprindelse. (Long branch attraction)



### Opgave 5 (10%)



Tre arter frøer i slægten *Rana* har udbredelsesområder, der hovedsageligt er allopatriske, men deres udbredelse overlapper i Texas og Oklahoma. Den følgende figur viser deres yngle- sæson i løbet af året for de tre arter. (Jo tykkere stregen er, jo højere er yngleaktiviteten. "Day" er dag i løbet af året.)



- a) Beskriv KORT ligheder og forskelle i ynglesæsonen mellem allopatriske og sympatriske populationer.

I de allopatriske populationer har to af arterne to ynglesæsoner i løbet af året. Ynglesæsonen overlapper tidsmæssigt for de tre arter. I sympatriske områder har alle arter kun en ynglesæson, og der er stort set ikke overlap i ynglesæsonen mellem de tre arter.

- b) Forklar, hvad der kan have forårsaget forskellen i ynglesæsonen i allopatriske og sympatriske populationer.

Forskellen i ynglesæsonen mellem allopatriske og sympatriske populationer skyldes reproduktivt character displacement, der forhindrer parring mellem de forskellige arter.

- c) Kan du give et bud på, om hybrider mellem de tre arter har lige så høj fitness, lavere fitness eller højere fitness som krydsninger inden for arter.

Hybriderne mellem arterne har haft lavere fitness. Ellers ville der ikke være blevet selekteret for character displacement, men de tre taxa ville smelte sammen.