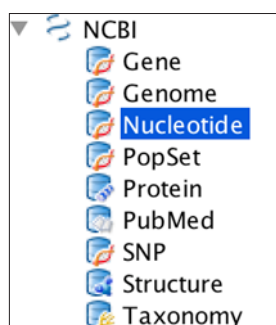


# Populationsgenetik hos to hvalarter

I 'Regn med biologi' side 98, figur 94 vises et fylogenetisk stamtræ over grønlandshval og sydlig rethval. Dette supplerende materiale viser hvorledes man kan finde sekvenser, fremstille et stamtræ som det viste i figuren, og hvordan man kan regne på genetiske forskelle i populationer ud fra sådant et stamtræ.

## Find sekvenser

For at undersøge om der er forskel i genetisk variation hos de to hvalarter, kan man anvende programmet 'Geneious' til at søge på NCBI's server. Herefter anvendes programmet MEGA til at fremstille et stamtræ og til at kvantificere forskellene i de to nucleotiddatasæt (grønlandshval og sydlig rethval). Begynd med at åbne Geneious og markér 'Nucleotide' på NCBI's søgefunktion – som, vist på figur 1.



Figur 1. Søgefunktionen i Geneious. Der skal søges på nucleotider.

Der vil i populationer kun forekomme små forskelle i DNA, og det gælder derfor om at søge på et område med en hurtig substitutionsrate og ingen bevarende selektion – kontrolregionen i mtDNA er på den baggrund velegnet til disse undersøgelser. Man skal derfor søge på hvalarten grønlandshval (*Balaena mysticetus*) og på 'control region'. I figur 2 vises efter at have klikket på 'More options' i Geneious' søgefelt og så valgt '+' for flere søgninger i søgefeltet: Søg på '*Balaena mysticetus*' i det øverste søgefelt og på 'control region' i det nederste søgefelt.



Figur 2. Søgning på grønlandshval og kontrolregionen i Geneious.

I denne undersøgelse skal der bruges data til et stamtræ med de to hvalarter, og der skal bruges data fra hver hvalart til kvantitativ analyse. Fremstil derfor en mappe ved at højreklikke på 'Local'

i stikstrukturen til højre og der vælges 'New folder'; navngiv den 'Hvaler'. På samme måde opstilles to undermapper der henholdsvis hedder 'Sydlig Rethval' og 'Grønlandshval'. I søgningen kommer der nu nogle sekvenser på 453 baser frem blandt mange andre. Sortér i sekvenserne ved at klikke på 'Sequence length' i resultatvinduet som vist med rød ring i figur 3. Markér alle sekvenserne fra og med AF355204 til og med AF355271 (68 i alt) og træk dem over i mappen 'Hvaler'.

Name	Description	Modified	Organism	Sequence ...	Accession	GID	Topology	Molecule Type	Taxonomy	
AF355258	Balaena mysticetus haplotype CCC mitochondri...	09 May 2002	Balaena mys...	453	38.6%	AF355258.1	15420497	linear	DNA	Eukaryota; ...
AF355259	Balaena mysticetus haplotype DDD mitochondri...	09 May 2002	Balaena mys...	453	38.9%	AF355259.1	15420498	linear	DNA	Eukaryota; ...

Figur 3. Man kan sortere de fundne data ved at klikke på overskriften over de enkelte kategorier.

Disse 68 filer trækkes også over i mappen 'Grønlandshval'. Man kan kopiere filerne fra en mappe til en anden ved at holde <Ctrl> eller <CMD> nede, mens man trækker filerne fra den ene mappe til den anden.

Det samme gentages for sydlig rethval:

Foretag en søgning som før, hvor 'Nucleotide' markeres, og der søges efter '*Eubalaena australis*' og 'control region'.

Filerne med længden 499-500 baser (AF395044-AF395053), i alt 9 sekvenser, overføres til mappen 'Hvaler' og til mappen 'Sydlig rethval'.

Fordi Geneious' funktioner er begrænsede, skal data færdigbehandles i programmet 'MEGA'. Der er derfor to muligheder for at arbejde videre. Man kan lade Geneious lave alignmentet og herefter eksportere alignede data til MEGA, eller man kan eksportere de fundne datafiler over i MEGA som '\*.fasta'-filer og så foretage alt det yderligere arbejde, herunder alignment, i MEGA. Begge dele gennemgås i denne tekst fordi man på den måde kan lære noget om, hvordan data kan udveksles mellem programmer. Det er dog klart nemmest til dette formål at anvende den førstnævnte fremgangsmåde.

## Fremgangsmåde 1: Alignment i Geneious

Dobbeltklik på mappen 'Hvaler' og markér alle filerne ved at trykke på <Ctrl>+A (pc) eller <CMD>+A (Mac). Vælg herefter 'Align/Assemble' → 'Pairwise/Multiple Align' – vent til der kommer en alignment-fil frem i mappen. Den hedder 'Nucleotide alignment'. Klik på den fil så den markeres.

Klik derefter i menulinjen: 'Files' → 'Export' → 'Selected Documents' → 'MEGA alignment (\*.meg)'. Herefter kommer en dialogboks frem hvor man kan navngive filen. Kald den 'Hvaler'. Under 'Tags' kan man skrive en note, men der behøver ikke at stå noget så bare lad feltet stå tomt. Nederste felt i dialogboksen henviser til hvor man vil gemme filen.

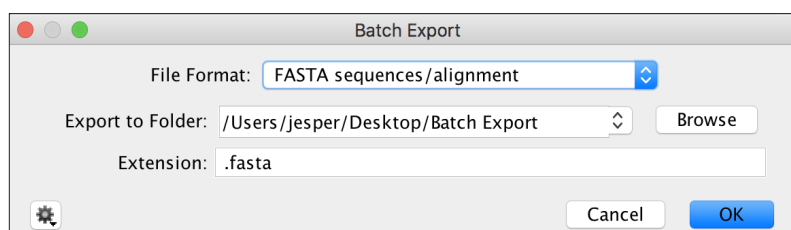
Gem fx på skrivebordet og flyt evt. senere filen ind i en anden mappe. Man advares om at der kan være et datatab ved at gå fra '\*.geneious'-formatet til mega-, men tryk bare på 'Proceed'. Fremstil på samme måde alignment af alle filerne i hver af mapperne 'Grønlandshval' og 'Sydlig Rethval' og eksportér resultaterne som hhv. grønlandshval.meg og sydligrethval.meg (uden mellemrum i filnavnet – filtypenavnet sætter programmet automatisk).

## Fremgangsmåde 2: Alignment i MEGA

Filerne fundet i Geneious, eksporteres som \*.fasta-filer hvor hver enkel sekvens placeres i en enkel \*.fasta-fil.

Disse samles til én \*.fasta-fil, og man laver et alignment i MEGA.

Markér alle datafiler i mappen 'Hvaler'. Klik herefter på 'File' → 'Export' → 'Batch Export'. Geneious foreslår at filerne eksporteres som Geneious-format, men vælg i stedet 'FASTA sequences/alignment' som vist på figur 4. Eksportér filerne til skrivebordet.



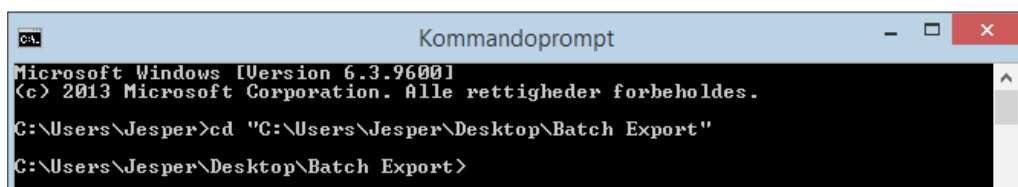
Figur 4. Eksport af data som \*.fasta-filer.

I næste dialogboks trykkes 'OK' uden flueben i 'Wrap sequence line', og 'Sequence' skal være i 'Upper case'. Der dukker nu en mappe op på skrivebordet (hvis det er der, I har gemt den) som hedder 'Batch Export'. I den mappe ligger nu 69 fasta-filer som skal kombineres til én \*.fasta-fil, hvor sekvenserne står under hinanden, og som kan importeres i MEGA. Der er forskel i proceduren, alt efter om man bruger en pc eller en Mac.

## Samling af information i en fil

### Pc

Klik på starttasten (den med de fire felter som et flag) og skriv 'CMD'. Der kommer et vindue op med et kommandoprompt. Et kommandoprompt fungerer på den måde at man kan skrive kommandoer og udføre dem ved at trykke på <enter>. De kommandoer der skal foretages, skal ske i mappen 'Batch Export' på skrivebordet. Derfor skal der skiftes bibliotek til mappen 'Batch Export'. Det gøres som vist i figur 5, ved at skrive 'cd' (change directory), trykke på mellemrumstasten og trække mappen 'Batch Export' ind i kommandopromptet.



Figur 5. Kommandopromptvinduet i en pc.

Efter tryk på <enter> kan man som vist i figur 5, tredje linje, se at stien før '>' peger ind i 'Batch Export', og de kommandoer man laver herefter, foregår alene i mappen 'Batch Export' på skrivebordet. Da \*.fasta-filer er tekstfiler, kan de omdøbes til tekstfiler vha. følgende kommando:

'rename \*.fasta \*.txt' <enter>

Alle filerne beholder deres navne, men filtypen ændres. Herefter kopieres alle filernes indhold sammen i en fil som kaldes 'hvaler.txt'. Det gøres vha. følgende kommando:

```
'copy *.txt hvaler.txt' <enter>
```

Kommandoen kan læses som: Kopiér indholdet af alle filer der ender med 'txt', ind i filen 'hvaler.txt'. Dette gøres i alfabetisk rækkefølge efter filnavnet.

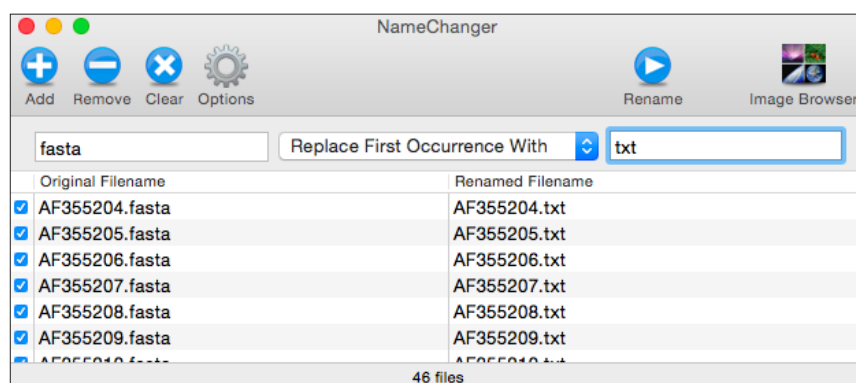
```
'rename *.txt *.fasta' <enter>
```

Filen 'hvaler.fasta' trækkes ud på skrivebordet, og mappen 'Batch Export' slettes. Proceduren gentages for filerne hhv. i mappen 'Grønlandshval' og 'Sydlig Rethval'. Filerne der flyttes ud på skrivebordet, kaldes hhv. 'grønlandshval.fasta' og 'sydligrethval.fasta'.

## Mac

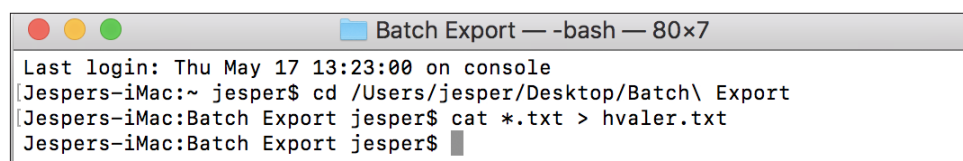
Man skal bruge to programmer: programmet 'NameChanger' (hentes i AppStore) og programmet 'Terminal' som computeren har indbygget. Gå til 'Lanchpad' og start NameChanger og Terminal.

Dobbeltklik på mappen 'Batch Export', markér alle filerne og træk dem over i NameChangers venstre side hvor der står 'Original Filename', som vist i figur 6. I boksen over 'Original Filename' skrives 'fasta', og i boksen over 'Renamed Filename' skrives 'txt'. Herved omdøbes \*.fasta-filer til tekst-filer når man trykker på 'Rename'.



Figur 6. Omdøbning af filer med NameChanger (Mac).

Start Terminal. Som vist i figur 7, er Terminal et kommandoprogram hvor man i tekst kan indsætte forskellige kommandoer. I Terminalvinduet skriver man 'cd ' (cd efterfulgt af et mellemrum) og trækker mappen 'Batch Export' ind i Terminalvinduet. Tryk <enter>. Skriv herefter 'cat \*.txt > hvaler.txt' (læg mærke til at der er mellemrum på begge sider af '>'). Afslut med at trykke <enter>.



Figur 7. Terminalvinduet på en Mac.

I NameChanger trykkes på 'Clear', og i mappen 'Batch Export' findes filen 'hvaler.txt' som trækkes over i NamChangers venstre side, hvor der i venstre boks så skrives 'txt' og i højre boks skrives 'fasta' – tryk 'Rename', og filen er omdøbt til en \*.fasta-fil. I mappen 'Batch Export' vælges filen 'hvaler.fasta' som trækkes ud på skrivebordet. Mappen 'Batch Export' slettes. Proceduren gentages for filerne i hhv. mappen 'Grønlandshval' og 'Sydlig Rethval'. Filerne der flyttes ud på skrivebordet, kaldes hhv. 'grønlandshval.fasta' og 'sydligrethval.fasta'.

## Fra \*.fasta til en alignet fil i MEGA

Man har når man er nået hertil, tre filer liggende på skrivebordet: enten hvaler.meg, grønlandshval.meg og sydligrethval.meg eller hvaler.fasta, grønlandshval.fasta og sydligrethval.fasta – alt efter om man har fulgt fremgangsmåde 1 eller fremgangsmåde 2. Har man \*.fasta-filer liggende på skrivebordet, skal de nu alignes:

Åben filen 'hvaler.fasta' i MEGA ved at vælge 'File' → 'Open a File/Session'. En dialogboks spørger herefter om man vil analysere eller aligne data. Vælg 'Align'. Programmet 'Alignment Explorer' åbner og ser ud som figur 8. Ved at klikke på 'W' i menuen laves et alignment ved hjælp af ClustalW-algoritmen. Vælg 'Align DNA' og svar 'OK' til de forslag der derefter kommer op – alignmentet er færdigt på få øjeblikke. Vælg herefter 'Data' → 'Export Alignment' → 'MEGA format'. Skriv OK til 'Title' og 'No' til 'Protein-coding nucleotide sequence data'. Filen kommer ud på skrivebordet. Proceduren gentages for hhv. grønlandshval.fasta og sydligrethval.fasta.

Derved har man uanset fremgangsmåde, filerne hvaler.meg, grønlandshval.meg og sydligrethval.meg liggende på skrivebordet. Man kan også vælge at lave et alignment ud fra Muscle-algoritmen. Så skal man klikke på ikonet ved siden af W, som forestiller en bøjet arm.

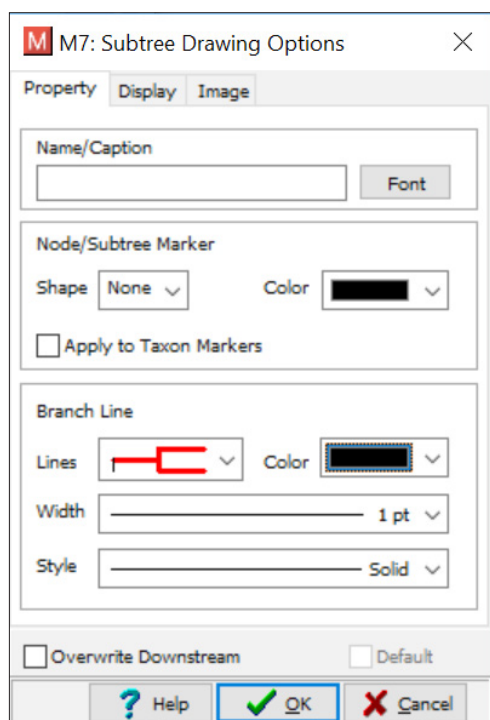
Species/Abbrv	Group Name	*	*	*	*	*	*	*	*					
1. AF355204_Balaena_mysticetus_haplotype_A_mitochondrial_control_region_		G	A	A	A	A	T	A	T	T	G	T	A	
2. AF355205_Balaena_mysticetus_haplotype_B_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
3. AF355206_Balaena_mysticetus_haplotype_C_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
4. AF355207_Balaena_mysticetus_haplotype_D_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
5. AF355208_Balaena_mysticetus_haplotype_E_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
6. AF355209_Balaena_mysticetus_haplotype_F_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
7. AF355210_Balaena_mysticetus_haplotype_G_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
8. AF355211_Balaena_mysticetus_haplotype_H_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
9. AF355212_Balaena_mysticetus_haplotype_I_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
10. AF355213_Balaena_mysticetus_haplotype_J_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
11. AF355214_Balaena_mysticetus_haplotype_K_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
12. AF355215_Balaena_mysticetus_haplotype_L_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
13. AF355216_Balaena_mysticetus_haplotype_M_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
14. AF355217_Balaena_mysticetus_haplotype_N_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
15. AF355218_Balaena_mysticetus_haplotype_O_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T
16. AF355219_Balaena_mysticetus_haplotype_P_mitochondrial_control_region_		G	A	A	A	A	T	A	T	A	T	T	G	T

Figur 8. Alignment Explorer i MEGA.

## Fremstilling af stamtræ i MEGA

For di Geneious i gratisversionen ikke kan farve dele af stamtræer, kan man lave stamtræet i MEGA som vist på figur 94 side 98 i 'Regn med biologi'. Åben filen 'hvaler.meg'. Vælg 'Phylogeny' → 'Construct/Test UPGMA tree'. Sæt 'Model/Method' til 'Jukes-Cantor-model'. Klik 'Compute'. Stamtræet der kommer frem, indeholder en masse information om grenlængde og navn på sekvenserne og andet. Klik på ikonerne i venstre side så ingen af ikonerne er markerede. Derved fremstår stamtræet som tegnet – uden yderligere information end selve stamtræet.

Find den knude som samler alle individerne af sydlig rethval. Klik på knuden og vælg ikon nummer 7 i venstre side som viser en farvet del af stamtræet med en firkantet parentes omkring. Ved klik på dette ikon åbnes en boks der hedder 'Subtree Drawing Options' og som er vist i figur 9. I fanebladet 'Property' sættes 'Branch Line' til at være en anden farve end sort. I 'Regn med biologi' er valgt blå. I fanebladet 'Display' fjernes flueben i 'Display Bracket'. Klik på 'OK'. Stamtræets størrelse kan reguleres. Ikonet øverst der viser en hammer (se rød ring figur 10), kan sætte afstanden mellem grenene, grenlængden og bredden af stamtræet. Et brugbart stamtræ fremkommer hvis man sætter 'Taxon Separation' til 10 pixels, 'Branch Length' til 100 og 'Tree Width' til 400.



Figur 9. Mulighed for at markere dele af et stamtræ som noget andet end det øvrige stamtræ.

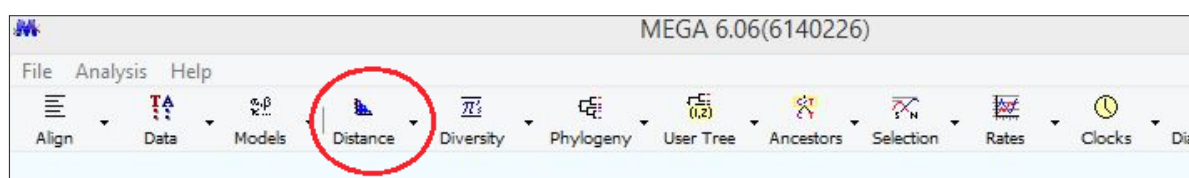


Figur 10. Redigering af stamtræets størrelsesforhold.

## Kvantitativ analyse

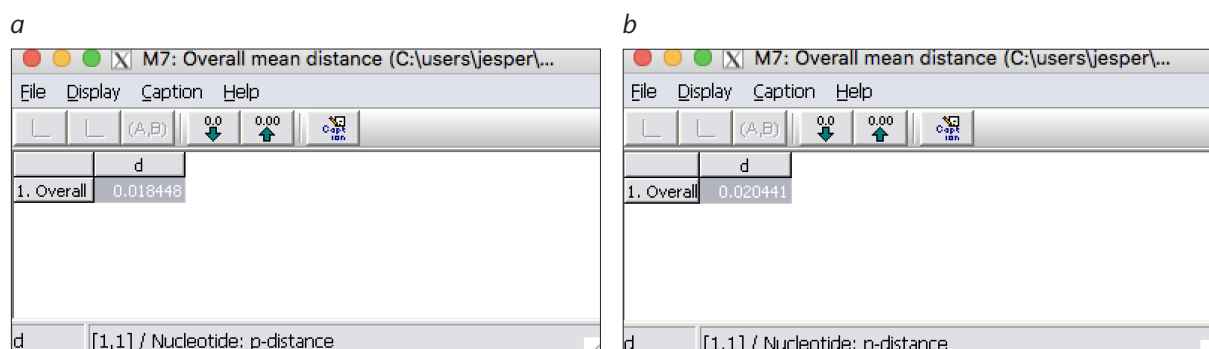
En rigtig god funktion i MEGA er at programmet kan lave afstandsmatrixer. Det skal nu undersøges hvad den gennemsnitlige forskel på sekvenserne er for henholdsvis grønlandshval og sydlig rethval. Begynd med at åbne filen 'grønlandshval.meg' i MEGA.

Vælg som vist på figur 11, herefter 'Distance' → 'Compute Overall Mean Distance'. Vælg Model/Method til at være '*p*-distance'.



Figur 11. Menulinjen med mulighed for at fremstille afstandsmatrixer.

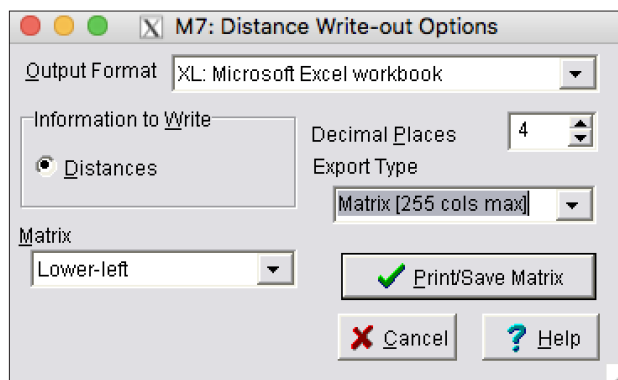
Der fremkommer herefter en beregning af den gennemsnitlige forskel i sekvenserne inden for de fundne arter af grønlandshval. Det ses af figur 12a at forskellen er 0,0184 for grønlandshval. Gøres det samme for sydlig rethval fås som vist i figur 12b, en middelværdi på 0,0204, der altså er højere end for grønlandshvalen. Med de data der her er til rådighed, kan man derfor underbygge at jagten på grønlandshval har medført et tab af genetisk variation, idet variationen er større blandt de 9 rethvaler end blandt de 68 grønlandshvaler.



Figur 12. Gennemsnitlig genetisk forskel på bestande (*p*-afstand). a. Grønlandshval, b. Sydlig rethval.

## Evt. viderebehandling i Excel

Ønsker man at overføre sine data til Excel for at foretage yderligere beregninger, kan man vælge 'Distance' → 'Compute Pairwise Distances'. Svar 'OK' til de dialogbokse der fremkommer. Vælg derefter 'Compute', og programmet danner en afstandsmatrix med *p*-afstande. Hvis man klikker på 'XL' i menuen for oven, kan man få overført matrixen til Excel og regne videre dér. Hvis man vil eksportere en matrix og ikke bare en kolonne med parvise sammenligninger, skal 'Export Type' som vist i figur 14', ændres til 'Matrix' (der er en begrænsning på 255 kolonner), og man kan også fastlægge antallet af decimaler (vælg 4) – jo lavere *p*-værdier, jo flere decimaler er nødvendige. Ved tryk på 'Print/Save Matrix' starter Microsoft Excel (Windows), eller man kan gemme en Excel-fil som man navngiver (Mac).



Figur 14. Eksport af afstandsmatrix til Excel fra MEGA.