Tracing the Evolution of Streptophyte Algae and their Mitochondrial Genome

Supplementary Material

Monique Turmel, Christian Otis, and Claude Lemieux

Institut de biologie intégrative et des systèmes, Département de biochimie, de microbiologie et de bio-informatique, Université Laval, Québec (QC) Canada

Corresponding Author: Monique Turmel
Institut de Biologie Intégrative et des Systèmes
1030 ave de la medicine, Pavillon Charles-Eugène Marchand
Université Laval, Québec (QC) Canada G1V 0A6;
phone: 418-656-2131 ext. 7623; Fax: 418-656-7176;
email: monique.turmel@bcm.ulaval.ca

Table S1

Comparison of evolutionary models using cross validation and the mitochondrial data set of 9,013 positions

Models Compared	**Likelihood Difference (±SD)**
GTR+Γ4 vs LG+Γ4	357.49 ± 31.62
CAT+Γ4 vs LG+Γ4	495.77 ± 55.21
CATGTR+Γ4 vs LG+Γ4	709.3 ± 55.86

Fig. S1. – Gene maps of the two klebsormidialean mitochondrial genomes sequenced in this study. Coding sequences (filled boxes) on the outside of the map are transcribed in a clockwise direction. Introns are represented by open boxes. tRNA genes are indicated by the one-letter amino acid code followed by the anticodon in parentheses (Me, elongator methionine: Mf, initiator methionine). The *Klebsorminium* genome carry two non-identical copies of *trnI*(gau); the copy located between *trnG*(gcc) and *trnE*(uuc) is more similar in sequence to its orthologues in other streptophyte mtDNAs.

Fig. S2. – Gene maps of the two zygnematalean mitochondrial genomes sequenced in this study. Coding sequences (filled boxes) on the outside of the map are transcribed in a clockwise direction. Introns are represented by open boxes. tRNA genes are indicated by the one-letter amino acid code followed by the anticodon in parentheses (Me, elongator methionine: Mf, initiator methionine).

Fig. S3. – Shared mitochondrial gene clusters between the moss *Physcomitrella* and the charophyceans *Chara/Nitella*, *Roya*, *Closterium* and *Chaetosphaeridium*. The gene order data sets used for these four pair-wise genome comparisons are presented. Each data set includes all genes/pseudogenes that are shared by *Physcomitrella* and the charophycean analyzed. Conserved clusters are denoted by boxes of different colors. The + and - signs indicate gene polarities.

Fig. S4. – Scenario of mitochondrial genome rearrangements during charophycean evolution and the transition to land plants. This is the best tree we inferred from a data set of 55 gene/pseudogene positions using GRAPPA and the algorithm of Caprara. The tree features the minimal number of reversals (109). Branch lengths indicate numbers of reversals.

Fig. S5. **–** Correspondence analysis of amino acid usage in the data set of 9,013 positions. Species names are abbreviated as in table 1.

Fig. S6. – Saturation levels of the original amino acid data set of 9,013 positions (A) and of the trimmed data set of 7,210 positions (B). The level of saturation was estimated for each data set by computing the slope of the regression line of patristic distances versus observed distances (solid line). The dashed line represents the saturation plot predicted for a data set with no substitution saturation (slope = 1). The patristic distances were inferred from branch lengths of the best tree inferred under the CATGTR+Γ4 model.

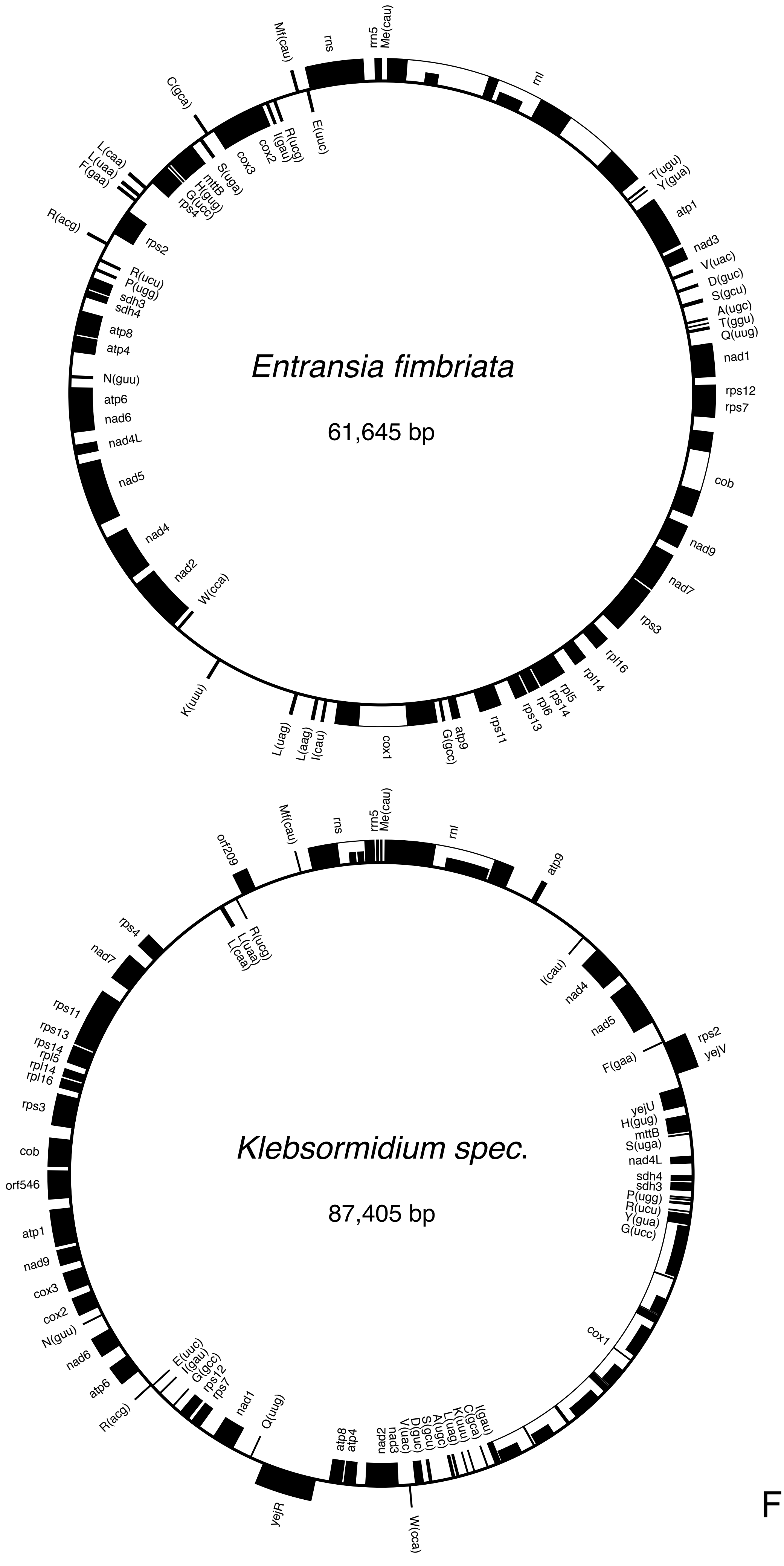

Entransia fimbriata
61,645 bp
Mf(cau)
rns
rrn5
Me(cau)
rnl
T(ugu)
Y(gua)
atp1
nad3
V(uac)
D(guc)
S(gcu)
A(ugc)
T(ggu)
Q(uug)
nad1
rps12
rps7
cob
nad9
nad7
rps3
rpl16
rpl14
rpl5
rps14
rpl6
rps13
rps11
atp9
G(gcc)
cox1
I(cau)
L(aag)
L(uag)
K(uuu)
W(cca)
nad2
nad4
nad5
nad4L
nad6
atp6
N(guu)
atp4
atp8
sdh4
sdh3
P(ugg)
R(ucu)
rps2
R(acg)
F(gaa)
L(uaa)
L(caa)
rps4
G(ucc)
H(gug)
mttB
S(uga)
cox3
C(gca)
cox2
I(gau)
R(ucg)
E(uuc)
Klebsormidium spec.
87,405 bp
Mf(cau)
rns
rrn5
Me(cau)
rnl
atp9
I(cau)
nad4
nad5
rps2
yejV
F(gaa)
yejU
H(gug)
mttB
S(uga)
nad4L
sdh4
sdh3
P(ugg)
R(ucu)
Y(gua)
G(ucc)
cox1
I(gau)
C(gca)
K(uuu)
L(uag)
A(ugc)
S(gcu)
D(guc)
V(uac)
nad3
nad2
W(cca)
atp4
atp8
yejR
Q(uug)
nad1
rps7
rps12
G(gcc)
I(gau)
E(uuc)
R(acg)
atp6
nad6
N(guu)
cox2
cox3
nad9
atp1
orf546
cob
rps3
rpl16
rpl14
rpl5
rps14
rps13
rps11
nad7
rps4
orf209
R(ucg)
L(uaa)
L(caa)

Figure S1

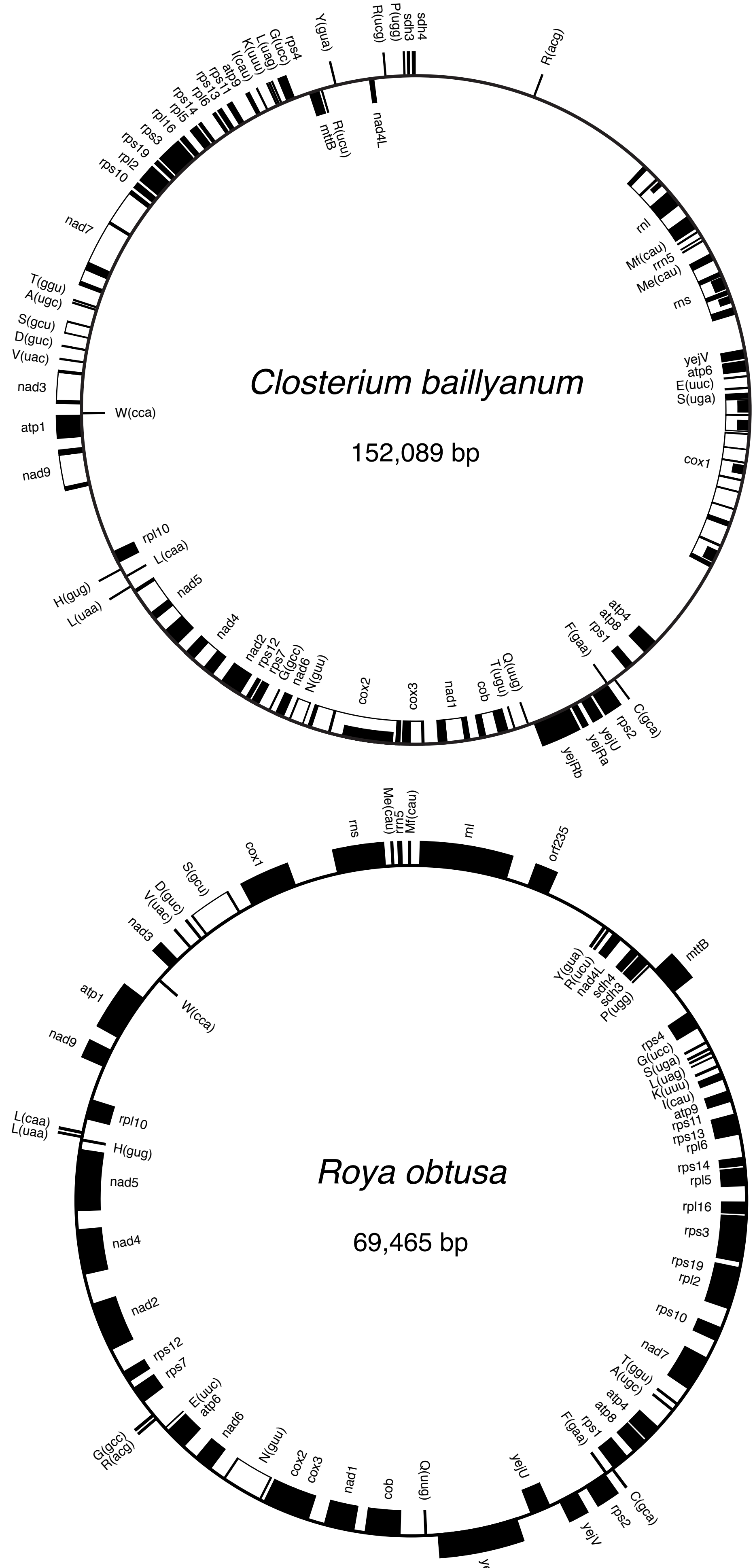
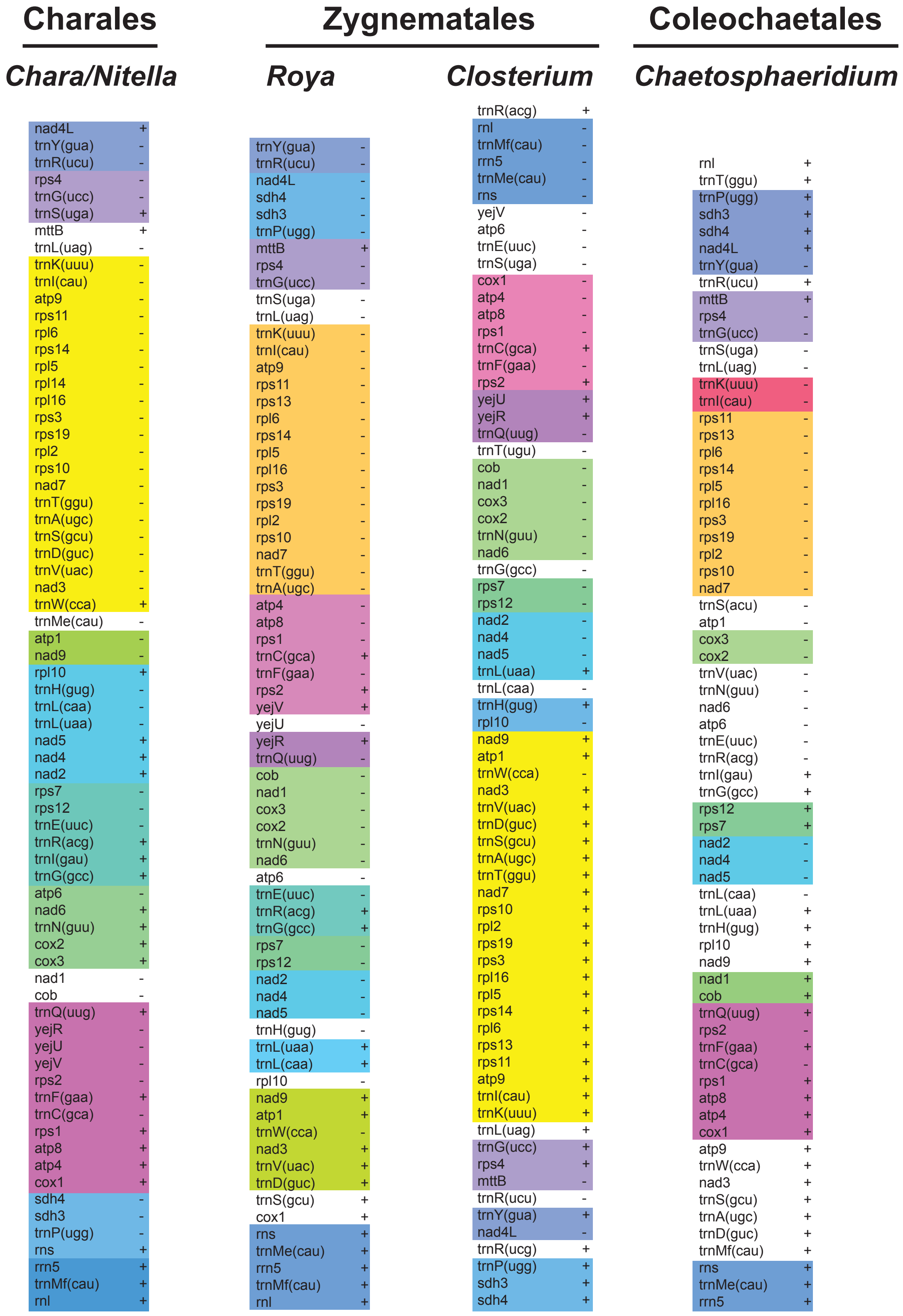

Closterium baillyanum
152,089 bp
sdh4
sdh3
P(ugg)
R(ucg)
Y(gua)
rps4
G(ucc)
L(uag)
K(uuu)
I(cau)
atp9
rps11
rps13
rpl6
rps14
rpl5
rpl16
rps3
rps19
rpl2
rps10
R(ucu)
mttB
nad4L
R(acg)
rnl
Mf(cau)
rrn5
Me(cau)
rns
yejV
atp6
E(uuc)
S(uga)
cox1
nad7
T(ggu)
A(ugc)
S(gcu)
D(guc)
V(uac)
nad3
W(cca)
atp1
nad9
rpl10
L(caa)
H(gug)
L(uaa)
nad5
nad4
nad2
rps12
rps7
G(gcc)
nad6
N(guu)
cox2
cox3
nad1
cob
T(ugu)
Q(uug)
F(gaa)
rps1
atp8
atp4
C(gca)
rps2
yejU
yejRa
yejRb
Roya obtusa
69,465 bp
rns
Me(cau)
rrn5
Mf(cau)
rnl
orf235
cox1
S(gcu)
D(guc)
V(uac)
nad3
W(cca)
atp1
nad9
Y(gua)
R(ucu)
nad4L
sdh4
sdh3
P(ugg)
mttB
rps4
G(ucc)
S(uga)
L(uag)
K(uuu)
I(cau)
atp9
rps11
rps13
rpl6
rps14
rpl5
rpl16
rps3
rps19
rpl2
rps10
nad7
T(ggu)
A(ugc)
atp4
atp8
rps1
F(gaa)
C(gca)
rps2
yejV
yejU
yejR
L(caa)
L(uaa)
rpl10
H(gug)
nad5
nad4
nad2
rps12
rps7
E(uuc)
atp6
G(gcc)
R(acg)
nad6
N(guu)
cox2
cox3
nad1
cob
Q(uug)

Figure S2



Charales

Chara/Nitella

nad4L +
trnY(gua) -
trnR(ucu) -
rps4 -
trnG(ucc) -
trnS(uga) +
mttB +
trnL(uag) -
trnK(uuu) -
trnI(cau) -
atp9 -
rps11 -
rpl6 -
rps14 -
rpl5 -
rpl14 -
rpl16 -
rps3 -
rps19 -
rpl2 -
rps10 -
nad7 -
trnT(ggu) -
trnA(ugc) -
trnS(gcu) -
trnD(guc) -
trnV(uac) -
nad3 -
trnW(cca) +
trnMe(cau) -
atp1 -
nad9 -
rpl10 +
trnH(gug) -
trnL(caa) -
trnL(uaa) -
nad5 +
nad4 +
nad2 +
rps7 -
rps12 -
trnE(uuc) -
trnR(acg) +
trnI(gau) +
trnG(gcc) +
atp6 -
nad6 +
trnN(guu) +
cox2 +
cox3 +
nad1 -
cob -
trnQ(uug) +
yejR -
yejU -
yejV -
rps2 -
trnF(gaa) +
trnC(gca) -
rps1 +
atp8 +
atp4 +
cox1 +
sdh4 -
sdh3 -
trnP(ugg) -
rns +
rrn5 +
trnMf(cau) +
rnl +

Zygnematales

Roya

trnY(gua) -
trnR(ucu) -
nad4L -
sdh4 -
sdh3 -
trnP(ugg) -
mttB +
rps4 -
trnG(ucc) -
trnS(uga) -
trnL(uag) -
trnK(uuu) -
trnI(cau) -
atp9 -
rps11 -
rps13 -
rpl6 -
rps14 -
rpl5 -
rpl16 -
rps3 -
rps19 -
rpl2 -
rps10 -
nad7 -
trnT(ggu) -
trnA(ugc) -
atp4 -
atp8 -
rps1 -
trnC(gca) +
trnF(gaa) -
rps2 +
yejV +
yejU -
yejR +
trnQ(uug) -
cob -
nad1 -
cox3 -
cox2 -
trnN(guu) -
nad6 -
atp6 -
trnE(uuc) -
trnR(acg) +
trnG(gcc) +
rps7 -
rps12 -
nad2 -
nad4 -
nad5 -
trnH(gug) -
trnL(uaa) +
trnL(caa) +
rpl10 -
nad9 +
atp1 +
trnW(cca) -
nad3 +
trnV(uac) +
trnD(guc) +
trnS(gcu) +
cox1 +
rns +
trnMe(cau) +
rrn5 +
trnMf(cau) +
rnl +

Closterium

trnR(acg) +
rnl -
trnMf(cau) -
rrn5 -
trnMe(cau) -
rns -
yejV -
atp6 -
trnE(uuc) -
trnS(uga) -
cox1 -
atp4 -
atp8 -
rps1 -
trnC(gca) +
trnF(gaa) -
rps2 +
yejU +
yejR +
trnQ(uug) -
trnT(ugu) -
cob -
nad1 -
cox3 -
cox2 -
trnN(guu) -
nad6 -
trnG(gcc) -
rps7 -
rps12 -
nad2 -
nad4 -
nad5 -
trnL(uaa) +
trnL(caa) -
trnH(gug) +
rpl10 -
nad9 +
atp1 +
trnW(cca) -
nad3 +
trnV(uac) +
trnD(guc) +
trnS(gcu) +
trnA(ugc) +
trnT(ggu) +
nad7 +
rps10 +
rpl2 +
rps19 +
rps3 +
rpl16 +
rpl5 +
rps14 +
rpl6 +
rps13 +
rps11 +
atp9 +
trnI(cau) +
trnK(uuu) +
trnL(uag) +
trnG(ucc) +
rps4 +
mttB -
trnR(ucu) -
trnY(gua) +
nad4L -
trnR(ucg) +
trnP(ugg) +
sdh3 +
sdh4 +

Coleochaetales

Chaetosphaeridium

rnl +
trnT(ggu) +
trnP(ugg) +
sdh3 +
sdh4 +
nad4L +
trnY(gua) -
trnR(ucu) +
mttB +
rps4 -
trnG(ucc) -
trnS(uga) -
trnL(uag) -
trnK(uuu) -
trnI(cau) -
rps11 -
rps13 -
rpl6 -
rps14 -
rpl5 -
rpl16 -
rps3 -
rps19 -
rpl2 -
rps10 -
nad7 -
trnS(acu) -
atp1 -
cox3 -
cox2 -
trnV(uac) -
trnN(guu) -
nad6 -
atp6 -
trnE(uuc) -
trnR(acg) -
trnI(gau) +
trnG(gcc) +
rps12 +
rps7 +
nad2 -
nad4 -
nad5 -
trnL(caa) -
trnL(uaa) +
trnH(gug) +
rpl10 +
nad9 +
nad1 +
cob +
trnQ(uug) +
rps2 -
trnF(gaa) +
trnC(gca) -
rps1 +
atp8 +
atp4 +
cox1 +
atp9 +
trnW(cca) +
nad3 +
trnS(gcu) +
trnA(ugc) +
trnD(guc) +
trnMf(cau) +
rns +
trnMe(cau) +
rrn5 +

Figure S3

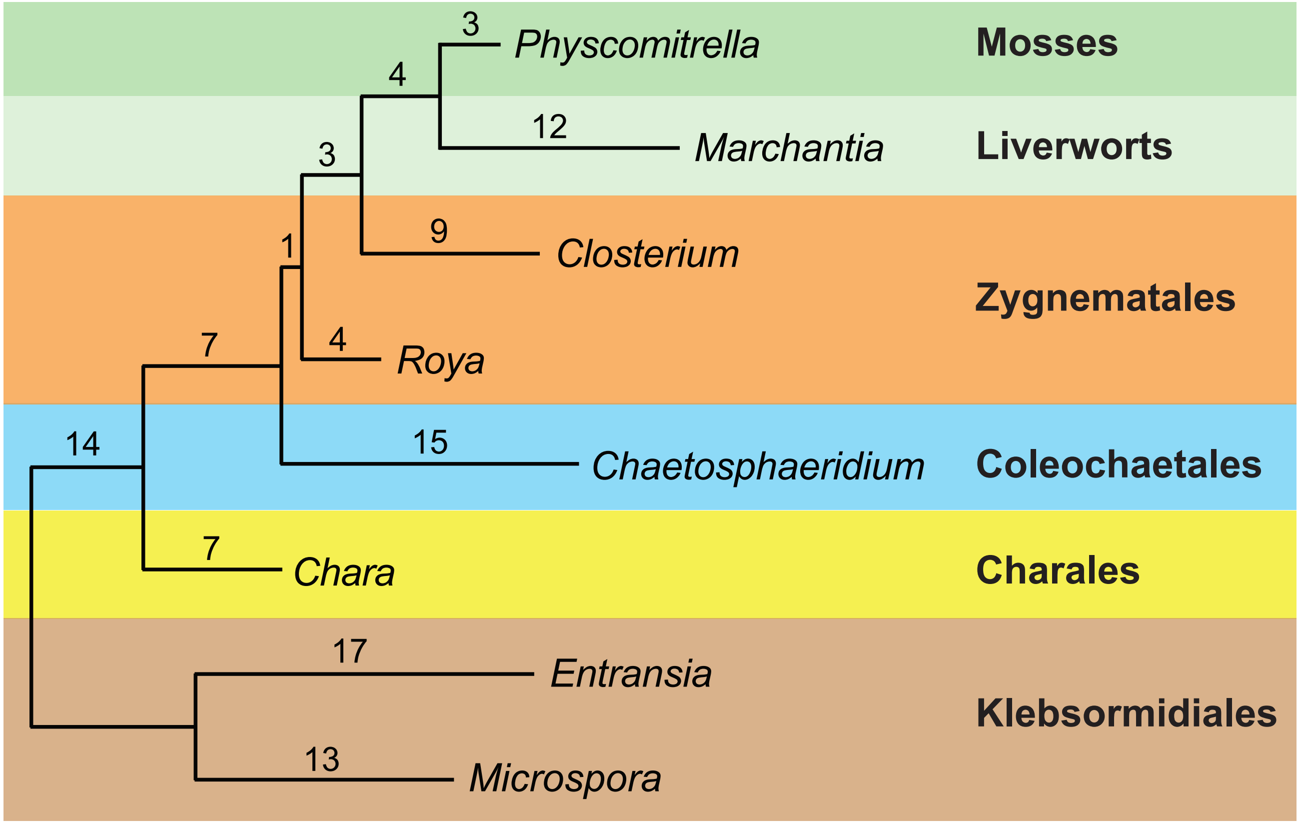
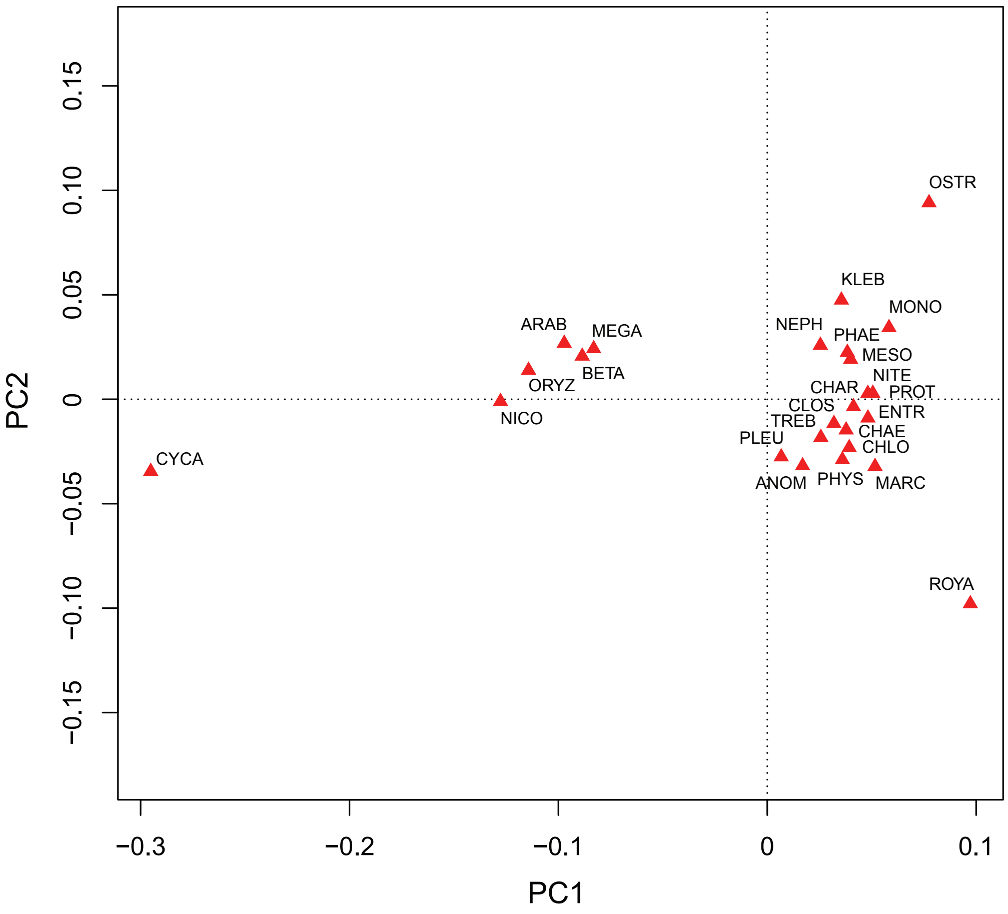

3
Physcomitrella
Mosses
4
12
Marchantia
Liverworts
3
9
Closterium
1
Zygnematales
7
4
Roya
14
15
Chaetosphaeridium
Coleochaetales
7
Chara
Charales
17
Entransia
Klebsormidiales
13
Microspora

Figure S4



PC2
0.15
0.10
0.05
0
−0.05
−0.10
−0.15
PC1
−0.3
−0.2
−0.1
0
0.1
OSTR
KLEB
MONO
NEPH
PHAE
MESO
NITE
CHAR
PROT
CLOS
ENTR
TREB
CHAE
PLEU
CHLO
ANOM
PHYS
MARC
ROYA
ARAB
MEGA
BETA
ORYZ
NICO
CYCA

Figure S5

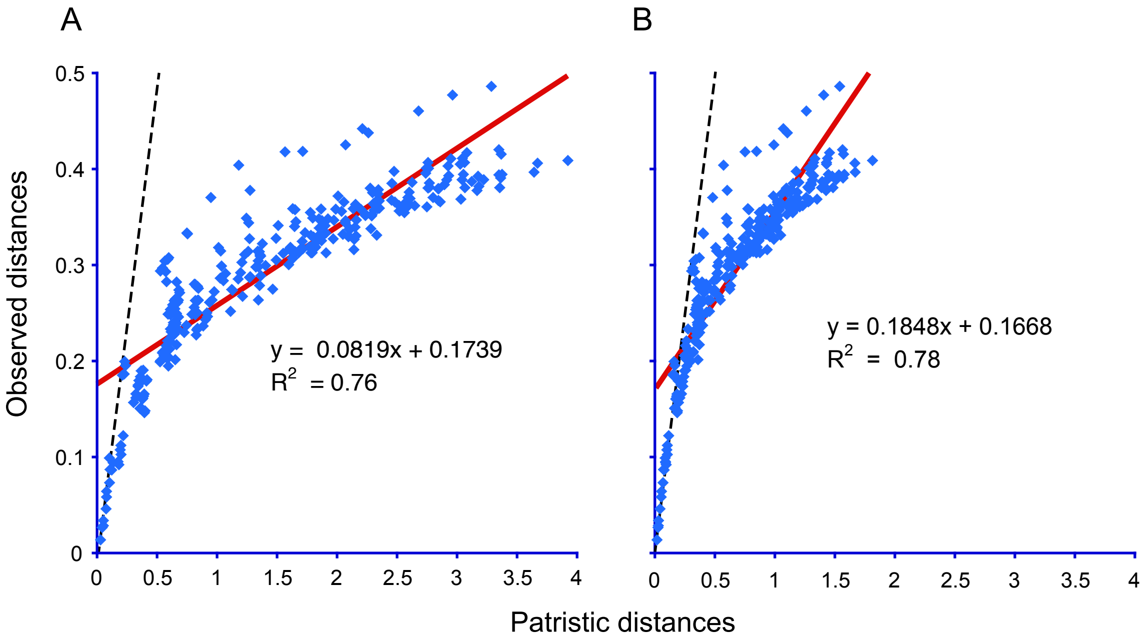
A
B
0.5
0.4
0.3
0.2
0.1
0
Observed distances
0
0.5
1
1.5
2
2.5
3
3.5
4
Patristic distances
y = 0.0819x + 0.1739
R2 = 0.76
y = 0.1848x + 0.1668
R2 = 0.78

Figure S6