Supplementary Materials: Analysis of the Mitochondrial Genome in *Hypomyces aurantius* Reveals a Novel Twintron Complex in Fungi

Youjin Deng, Qihui Zhang, Ray Ming, Longji Lin, Xiangzhi Lin, Yiying Lin, Xiao Li, Baogui Xie and Zhiqiang Wen

1 2 2	CLUSTAL O(1.2.1) multiple sequence alignment							
4567	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	ATGACAACTTTAACTAGAAGTAATTTTCAAGATCATCCTTTCCATTTAGTGTCACCTTCT ATGACAACTTTAACTAGAAGTAATTTTCAAGATCATCCTTTCCATTTAGTGTCACCTTCA ******************************						
8 9 10	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	CCTTGACCTTTATACACAAGTATATCTTTATTTACTTTAACAGTAAATGGAGCATTATCT CCGTGACCTTTATACACAAGTATATCTTTATTTAATTTA						
12 13 14	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	ATGCATCTATTCAGTAATAGCTATATATTATTCTTTATAGCTCTAGGTACAGTAGTAGCA ATGCATCTTTTCAATAATAGCTATATATTCTTCTTTTATAGCTTTAGCTACAGTAATTTCA ******** **** ***********************						
16 17 18	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	TCAATGTCTTTATGATTCAGAGATATTATCTCTGAAGGT ACATATTTAGGTAACCATACA TCAATGGCTTTATGATTCAGAGACATAGTATCAGAAGGT ACATTTTTAGGTAATCACACAT ****** *************** ** * ** ****** ****						
20 21 22	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	TTATCCGTTCAAAAAGGATTAAATTTAGGAGTTATATTATTATAGTATCTGAAGCTTTA CTTGCTGTTCAAAAAGGATTAAATTTAGGTGTAATATTATTTAT						
23 24 25 26	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	TTCTTCTTAGCTATATTCTGAGCATTCTTTCATAGTGCTTTAACACCTACAGTTGAATTA TTCTTCTTAGCTATATTCTGAGCATTCTTCCATAGTGCTTTAACACCTACAGTTGAATTA *****************************						
27 28 29 30	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	GGTGCTCAATGACCACCTATGGGTATAGAACCTATAAATCCTTTCGAATTACCTTTATTA GGTGCTCAATGACCACCTATGGGTATAGAACCTATAAATCCTTTTGAATTACCTTTATTA *************************						
31 32 33 34	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	AACACAGTTATATTGTTATCTAGTGGTGCTACAATTACTTATGCTCACCATTCTTTAATT AACACAGTTATATTATTATCTAGTGGTGCTACAATTACTTTTGCTCATCATAGTTTAATA ****************************						
36 37 38	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	AAAGGTGAAAGATCAGGAGCTTTATATGGTACTATTTCGACAGTAGTATTAGCTTTAGTA AAAGGTGAAAGATCAGGAGCATTATATGGAACTATTTTCACAGTAGTATTAGCTTTAATA *******************************						
40 41 42	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	TTTACTTTATTTCAAGGAGTAGAATATAGCGTTTCTTCATTTACAATAAGTGACGGTGTA TTTACTTTATTCCAAGGAGTAGAATATAGCGTTTCTTCATTCA						
43 44 45 46	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	TTTGGAACATGTTTCTTCTTCGGAACAGGTTTTCATGGCT CCATGTTATCATAGGTACA TTTGGAACATGTTTCTTCTTCGGAACAGGTTTCCACGGAT TCCACGTTATTATAGGTACA						
48 49 50	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	ATCTTCTTATCTGTAGGATTATGAAGAATAATGGCATATCATTTAACAGATCATCACCAT GTTTTCTTATCTGTTGCATTGTGAAGAATAATGGCATACCATTTAACAGATCATCATCAT * ********** * *** ******************						
52 53	Hypomyces_aurantius_cox3 Hypocrea_jecorina_cox3	CTTGGTTATGAAGCTGGAATATTATACTGACATTTTGTAGATGTTGTTTGATTATTCTTA CTTGGTTACGAAGCTGGAATATTATACTGACATTTTGTAGATGTTGTTTGATTATTCTTA						

Figure S1. *Cox3* gene sequence alignment between *H. aurantius* and *H. jecorina*. Intron-less *cox3* gene sequence of *H. jecorina* was used to compare against that of *H. aurantius* to confirm locations of introns of cox3 gene in *H. aurantius*. Results showed that the second intron located in position of 640 bp and was 2,877 bp in size. No exon sequences was found in the large intron. Consensus sequences were marked by asterisks. Arrows represent probable insertion sites. The alignment was performed using the program CLUSTAL O [27].

Int. J. Mol. Sci. 2016, 17, 1049; doi:10.3390/ijms17071049

1 2 3	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	TCAATTGTTATATTTTTAACAATGTAATTAATCGTTCTTCTTCTCTA TTAGTTATTCTAACAATGTTAATTATTATTCTCACTTAAAAACTAA * * ** ********* * * ***
4 5 7	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	GTAACAAAAGATTTTATTCTACTATGTCTGTTGAGCGTAAAATACTATCTCC: GCAATAAAAGATTCTATTCTACTATACCCGAAGAGCGTAATAACCTTTCTCCC * ** ******** ************** * * ******
8 9 10 11	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	TAACAGGATTTGCGGATGCTGAATCAAGTTTTAGTCTAAAGGTATCTAAAAA TAACAGGATTTGCCGACGCAGAGTCAAGTTTTAGTCTAAAAGTATCTAAAAA **************
12 13 14 15	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	CTAGATCAGGATGACATATTATTCCTGAATTTAGGATAGAGTTACATAATAGA CTAAATCAGGTGACACGTAATACCTGAATTTAGAATAGAGTTACACAATAGA
16 17 18 19	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	ATCTATTAAGAGAAATACATGCCTTTTTTGGGGTAGGTATTATCACAGAGAT ACTTATTAAGAGAAATACACCCCTTTTTTGGGGTAGGTCATATTAGTGAGTA * *****************
20 21 22 23	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	AAAATAAGGTAGTTTATAGTGTACAATCGCGTCGAGATTTAACTAATGCAAT TAAATAAGGTAGTTTACAGTGTTCCAATCCTCTCGTGATTTACGTAATGTAAT ******************
24 25 26 27	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	ATTTTGATCAATATCCTTTAATAACTCAAAAAAAGGCGGATTACTTTTAATT ATTTTGACAAATATCCTTTAATAACTCAAAAAAATGCTGATTACCTTTTATT ******* ************************
28 29 30 31	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	CTGTTGAGATGTTATGTTTAAATGCGCAATCAGATATTAGGGGAGTTCAAGA CTATTGAGCTATTAAATTTAAAGATACATTCAGACATTAACGGTTTTCAAGG ** ***** * *** ****** ** ****** ** *****
32 33 34 35	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	GTCTAAAAGCAGCAATGAATAAAGGATTGTCTGATAAATTAAATTCTGAGTT GACTTAAAGCAGCAATGAACTGAGGATTGTCAGACCAAATTAATGAATG
36 37 38 39	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	ATGTTATCCCCGCCGAACGTCCTACAATTAATTTTGAAAGAATACCTGATCC2 GTGTTATTTCAGCTATACGTCCTGCAGTAAATTTTTGAAGGTATACCCCAATCC2 ****** * ** ******* ** ********* * *****
40 41 42 43	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	TAGCAGGTTTCGTGGACGGAGAGGGGTTGTTTCTATATTAACACAAAAAAAA
44 45 46 47	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	ATATAACTGGTTATCAGGTTATCATGACATTCTCTATATCTCAACATGTAAG ATTTAACAGGATATCAGGTTATTATGTCCTTCTCTATATGTCAACATGTAAA ** **** ** *********** *** * ********
48 49 50 51	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	TTTTATTAAGAAAGTTTATAGATTACTTAAATTGTGGTAATATAGAAAAAGT TTTTATTAAATAAGTTCACAGAGATTACTTAGGTTGTGGTAAAAAAGAGAAAAAGT ********* * ***** * *****************
53 54 55	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	GACCTGATGGAGTTAAATTTATTGTATATAAATTTAGTGATATTAGAGACAA GACCTGATCAGGTAACATTTGTTACCTATAAATTTAGTGATATTCGTGATAA ******** ** * **** ** *************
57 58 59	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	CTTTTTTCCAAAATTACCCTTTACATGGAATGAAGTCTAAGGATTTTAGAGA CTTTTTTCCAAAACTATCCTTTACACGGGGTGGAAATCTAAGGATTACAGAGA ************** ** ******* ** *****
61 62 63	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	AAATAGCTAAACTTATGGAAAACAAATCACATTTAACTCTTGACGGTTTCAAC AGATAGTTAAACTTATGGAAAATAAATCACATTTAACTCTTGACGGTCTTAA * **** ***************** **********
65 66 67	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	AATCTTTAAAATCTGGTATGAATAGAAGAAGAATGTAGAATGCGCGATTAAT AATCATTAAAATCTGGAATGAATAGAGGGTAGAATATACGAT **** *********** ********* ********
69 70 71 72	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	CGATTATGATGTGGAGGCCCTTTTAATTATTAATATTAATTA
73 74 75 76	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	TTAATAGCAAATATTGGAACCTTTATAACGTGGAGGCCCTTTTATATGATTA: TAATATTGGAACCTTTATAACGTGGAGGCCCTTTTATATGATTA: **********************************
77 78 79 80	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	ТАААТАТТААТАТТААТААТТАААТАGААААСАААGTTAAATGCATAGAACT ТАААТАТТААТАТТААТААТТАААТАGAAAACAAAGTTAATTGCATGAAACT **********************************
81 82 83 84	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	-TAAGAGACAATATGCAGCCAAATATTATTAGAACATTAAAATAATTTCTAA TATTGAGATAATATGCAGCCAAACATTATTAGAATATTAAATTAATT
85 86 87	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	GTTCAACGACTAAAAAGATAGATTTCCTATCTTGCAAATTTTGTTGTATAAT GTTCAACGACTAAAAAGATAGATCTTCTATCTTGCAAATTTTGCTGTATAGT?
88 89 90 91	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	TATTATACAAGTGACATAGTCTGAACAATAATGAAAATTATTGAAGGATTGAT TATTATATGGATGATATAGTCTGAACAATAATTAAAATTATTGAAGGATTTAA ******* *** *********************
92 93 94 95	Chaetomium_thermophilum_varthermophilum_cox3-i2 Hypomyces_aurantius_cox3-i2_intron_with_orf330	TGCGCATTAAATCCGTTAACAAATTTTTTTTTTGATCCGTTAACAAAAAAAAAA

Figure S2. Alignment between cox3-i2 of *Chaetomium thermophilum* var. thermophilum and upstream intron in cox3-i2 of Hypomyces aurantius. Among 1339 base pairs compared, 1069 were identify (79.8%). Consensus sequences were marked by asterisks. Gaps were indicated by dash lines. Comparison were performed by CLUSTAL O 1.2.1.

Int. J. Mol. Sci. 2016, 17, 1049; doi:10.3390/ijms17071049

```
H. aurantius cox3-i2 downstream intron
                                             TAACATATGGAGCCCTTTTTACA---
                                                                     --TATATATTGTTTAATACAAAAAAATCTTATG
    F. solani cox3-i2
                                             H. aurantius cox3-i2 downstream intron
    F. solani cox3-i2
    H. aurantius cox3-i2 downstream intron F. solani cox3-i2
                                             TACCTGAATCTAAAGATAATTATTCTATAGATAAACAGTTTATAGAATGATTAGTAGGAT
    H. aurantius cox3-i2 downstream intron
                                             14
15
     F. solani cox3-i2
                                             TTACAGATGCAGAGGGTAATTTTAACCTTAAACTTACAGATCTAAAAGATAATACGTTTA
    H. aurantius cox3-i2 downstream intron
F. solani cox3-i2
                                             AATATGTTCAATTCCACATTCCAGATAGGTCTTCATGAAGATGACGTTCATGTGTTAGAAT
AATATGTTCAATATACATATCCAAATATCTCTTCATGAAGACGACATTGAAGATATAGAAT
**********
    H. aurantius cox3-i2 downstream intron
                                             ATATTATGAATACATTAAAATGTGGACATATTTCTAGATCAAAGGATAGAGTTAATTATT
    F. solani cox3-i2
                                             24
25
26
27
28
    H. aurantius cox3-i2 downstream intron
                                             TTGTTAATGATTTGAATTCTTTACTTTATGTAATAATTCCTATATTTAATTATGTTAATC
    F. solani cox3-i2
                                             TTGTTAATGATTTAAACTCTTTTGTTATACATAATAATTCCTATATTTAACTACGTTAATC
                                             29
30
    H. aurantius cox3-i2 downstream intron
    F. solani cox3-i2
32
33
34
35
    H. aurantius cox3-i2 downstream intron
                                             ATAATAAAATTATCTGATACAAATAAATTAGAAATTATAAAACTTCAAAAAGAAATGC
     F. solani cox3-i2
                                             AAAATAAAAAATTATCAGATGCTAAGAAGTTAGAAATAATAAAACTCCCAAAAGGAAATGC
    H. aurantius cox3-i2 downstream intron
                                             AAAATATGTCAGGTAAATGAATACCTAATTCTATAAATGATAAGATAATAATAACTAAGT
    F. solani cox3-i2
                                             40
    H. aurantius cox3-i2 downstream intron F. solani cox3-i2
41
42
                                             TTTGATTAGCAGGATTTATTGACGGGGAAGCCACTTTTTCAACTAATAAGTATATACCCA
                                             43
44
45
    H. aurantius cox3-i2 downstream intron
                                             GATTTAAATTAGAAAATCATATAAAAGAGTTGGAATTATATAAAAGAGATAAGAGATTTCT
46
47
48
                                             F. solani cox3-i2
    H. aurantius cox3-i2 downstream intron F. solani cox3-i2
                                             49
50
53
54
    H. aurantius cox3-i2 downstream intron
                                             TAGAAATAAATAAAATCCTGGAACTAAAAGAAAACTTAATTCCTTTAATGTCTCATGATA
                                             F. solani cox3-i2
    H. aurantius cox3-i2 downstream intron
                                             ATAGTGTTATATAAAAACGTTAAAAATCTAAAGATTTCTTATTATGATTAAGACTAGTCG
    F. solani cox3-i2
                                             GTAATGTTATACTAAGAACTTTAAAACATAAAGATTTCTTACTATGATTAAAATTAGTAG
                                             61
62
    H. aurantius cox3-i2 downstream intron
    F. solani cox3-i2
63
64
65
66
67
    H. aurantius cox3-i2 downstream intron
                                             AATTGCATATGAATAAATATAGATTAACTACTAATAGTAATTACTTATCAATAAGAAGAAC
                                             F. solani cox3-i2
69
70
    H. aurantius cox3-i2 downstream intron
                                             F. solani cox3-i2
                                             TTATTTCTATGGTAGAAATAGATAATCTAATGTCTAAGCTTTATTTGACAGATAGTCCTT
71
72
73
74
75
76
77
78
79
80
81
82
                                            ACGAAATAAAAACAAAATATAAGATATTATAGAAATACTGCTAAATTAGTAAGTGAAGCAA
ATGAAATAAAAGATAATAATAAGATTTTATAGAAATACTGATAAACTAGTAAGTGAAAGCAA
********
    H. aurantius cox3-i2 downstream intron
    F. solani cox3-i2
    H. aurantius cox3-i2 downstream intron
                                             CTAAAATTGTAGCAATGAAAGATAATCATGTGAAAATATATGATAGTATATCTGAATGTG
     F. solani cox3-i2
                                             CAAAAATTATAGCAATTAAAGATAACCAAAGTAAAATGTATAACAGTATATCTGAATGTG
    H. aurantius cox3-i2 downstream intron
                                             CAAAAGACCTTAATATTTCTAGAAAATGTATTAAAGGATGTATAGATTCAGGAAAATCAT
    F. solani cox3-i2
                                             CAAAAGATATTAGTATATCTCGTAAATATATCAAAGAATGTTTAATTTCAGGTAAATTTT
                                                  *****
                                                  Hypomyces_aurantius_cox3-i2_intron_with_orf327
Fusarium_solani_cox3-i2
                                                  Hypomyces_aurantius_cox3-i2_intron_with_orf327
Fusarium_solani_cox3-i2
                                                  AGTAAAAGAATGGCATAAAACTCTACTTATTAATTATCATTAGTTGATAAGCAC
AGTAAAAGAATGGCACAAAACTCTACTTATTAATTATCATTAATTTGATAAGAAC
            Hypomyces_aurantius_cox3-i2_intron_with_orf327
Fusarium_solani_cox3-i2
            Hypomyces_aurantius_cox3-i2_intron_with_orf327
Fusarium_solani_cox3-i2
                                                  AGATGACATAGTCTAAACAAGAATGAAAATTATTGATAATCTCTTTTTATTG
                                                  AGATTACATAGTCTAAGCAATAACGAGAGTTATTGATAATATATTTG---
```

Figure S3. Alignment between cox3-i2 of *Fusarium solani* and downstream intron in cox3-i2 of Hypomyces aurantius. Among 1538 base pairs compared, 1234 were identify (80.2%). Consensus sequences were marked by asterisks. Gaps were indicated by dash lines. Comparison were performed by CLUSTAL O 1.2.1.



0.2

Figure S4. Phylogenetic analysis for translated protein sequences of six ORFs/HEGs. HEG in upstream intron of cox3-i2 showed highest similarity with ORF330 in cox3-i2 of *C. themophilum* var. thermophilum and ORF338 in cox3-i2 of A. ruber, and HEG in downstream intron matched well with ORF426 in cox3-i2 of *F. solani* and ORF436 in cox3-i3 of *F. culmorum*. ORF309 in cox3-i1 (group IB) of *H. aurantius* was used as an outgroup. Comparing with the ORF309, the relationship between two HEGs in cox3-i2 of *H. aurantius* was much closer.

Table S1. Conserved protein-coding genes and ribosomal RNA subunits in *H. aurantius* mitochondrial genome.

Gene	Start	End	Strand	Gene Size, bp	CDSSize, bp	Intron Number	Average Intron Length, bp	Percent of Intronic Sequences, %	Stop Codon
rns	1	1505	+	1505				0	
cox3	3516	8330	+	4815	810	2	2003	83.2	TAA
nad6	8565	9242	+	678	678			0	TAA
rnl	11,589	22,088	+	10,500		4	1993	80	
rps3	17,762	19,186	+	1425	1425			0	TAG
nad2	30,234	31,901	+	1668	1668			0	TAA
nad3	33,265	33,678	+	414	414			0	TAA
atp9	33,856	34,050	+	195	195			0	TAG
cox2	35,317	39,361	+	4045	750	1	1795	81.5	TAA
nad4L	41,027	41,296	+	270	270			0	TAA
nad5 ª	41,296	44,055	+	2760	2760			0	TAA
cob	44,271	49,297	+	5027	1170	3	1286	76.7	TAA
cox1	49,947	60,590	+	10,644	1587	5	1811	85.1	TAA
nad1	61,752	62,852	+	1101	1101			0	TAA
nad4	63,999	65,456	+	1458	1458			0	TAA
atp8	66,581	66,727	+	147	147			0	TAA
atp6	66,860	70,630	+	3771	783	2	711	79.2	TAA

^a *Nad*5 gene uses the last A of nad4L termination codon as first nt, blank space means no present or observed, or unable to calculate.

Amino acid	Codon	Frequency									
Phe	UUU	571	Ser	UCU	426	Tyr	UAU	568	Cys	UGU	87
	UUC	256		UCC	48		UAC	165		UGC	19
Leu	UUA	1158		UCA	269	TER	UAA	34	Trp	UGA	130
	UUG	104		UCG	27		UAG	10		UGG	11
	CUU	201	Pro	CCU	254	His	CAU	200	Arg	CGU	93
	CUC	8		CCC	22		CAC	69		CGC	4
	CUA	139		CCA	91	Gln	CAA	264		CGA	3
	CUG	29		CCG	16		CAG	35		CGG	1
Ile	AUU	501	Thr	ACU	350	Asn	AAU	836	Ser	AGU	387
	AUC	117		ACC	24		AAC	216		AGC	73
	AUA	791		ACA	339	Lys	AAA	955	Arg	AGA	372
Met	AUG	277		ACG	20		AAG	153		AGG	18
Val	GUU	307	Ala	GCU	328	Asp	GAU	448	Gly	GGU	380
	GUC	32		GCC	72		GAC	73		GGC	24
	GUA	367		GCA	186	Glu	GAA	537		GGA	279
	GUG	68		GCG	31		GAG	136		GGG	65

Table S2. The frequency of utilization of the 64 standard codons.

14,074 codons in Average of genes (used Filamentous fungi Mito).