

Supplementary tables

Supplementary Table 1: Raw and trimmed reads statistics for 2451S DNA sequencing.

Platform	Paired Insert Size	Average read length	Mill. raw reads (approx. genome coverage)	Average read length post-trim	Mill. reads post-trim (approx. genome coverage)
Roche 454	na	642.27 bp	5.92 (4.32x)	640.14 bp	5.88 (4.28x)
Illumina HiSeq	200 bp	100 bp	350.46 (39.83x)	87.0 bp	306.73 (30.32x)
Illumina HiSeq	300 bp	100 bp	138.38 (15.73x)	88.7 bp	131.11 (13.22x)
Illumina HiSeq	500 bp	100 bp	245.45 (27.89x)	88.8 bp	234.02 (23.61x)
Illumina HiSeq	3 kbp	80.02 bp	147.46 (13.41x)	78.66 bp	87.51 (7.82x)
Illumina HiSeq	8 kbp	84.85 bp	328.04 (31.63x)	80.77 bp	238.25 (21.87x)
Illumina HiSeq	20 kbp	82.65 bp	348.25 (32.71x)	76.77 bp	220.50 (19.24x)
Illumina HiSeq	40 kbp	76.77 bp	255.60 (22.30x)	76.25 bp	154.74 (13.41x)
Illumina MiSeq	~5 kbp	300bp	47.47 (16.18x)	127.4 bp	24.29 (3.52x)

Supplementary Table 2. Assembly statistics for 2451S reference genome assembly (BATG-0.5), after and before scaffolding, gap-filling and joining.

Scaffold Measure	Nuclear Scaffolds	Nuclear Contigs	Mitochondrial Scaffolds	Chloroplast Scaffold	Pre-scaffolding contigs
Number	89,487	118,931	26	1	235,463
Total size	866.8 Mbp	717.6 Mbp	580,788 bp	155,498 bp	663.1
Longest (bp)	884,900	209,591	184,534	-	97,536
Shortest (bp)	500	500	326	-	500
Number > 1K	40,751	68,601	25	1	141,677
Number > 10K	10,139	18,848	11	1	12,303
Number > 100K	2,520	208	1	1	0
Mean size (bp)	9,686	6,034	22,338	-	2,816
Median size (bp)	911	1,228	6,487	-	1,320
N50 length (bp)	104,030	25,305	60,627	-	5,698
L50 count (bp)	2,386	7,924	3	-	32,110
NG50 length (bp)	101,761	17,965	-	-	-
%A	27.22	32.87	27.49	30.79	32.87
%C	14.18	17.13	22.5	19.27	17.16
%G	14.19	17.14	22.29	18.56	17.14
%T	27.2	32.85	27.67	31.38	32.84
%N	17.21	0	0.04	0	0
CEGMA partial hits	238 (96%)	-	-	-	-
CEGMA complete hits	208 (84%)	-	-	-	-

Supplementary Table 3: Assembly statistics for five versions of 2451S genome assembly, after scaffolding, gap-filling and joining. (no scaffolding was performed on version 0.1, as only single-end reads were used). The final assembly was v.0.5.

Scaffold Measure	v0.1	v0.2	v0.3	v0.4	v0.5
Number of scaffolds	417,760	283,188	142,021	89,285	89,487
Total size of scaffolds (Mbp)	618.4	1,469.50	982.4	875.2	866.8
Longest scaffold (bp)	51,710	221,212	560,578	696,341	884,900
Shortest scaffold (bp)	100	344	500	500	500
Number of scaffolds > 1K	209,375	163,534	60,768	39,713	40,751
Number of scaffolds > 10K	1549	45,273	14,113	10,818	10,139
Number of scaffolds > 100K	0	16	2,202	2,484	2,520
Mean scaffold size (bp)	1,480	5,189	6,917	9,803	9,686
Median scaffold size (bp)	1,004	1,266	866	878	911
N50 scaffold length (bp)	2,412	14,228	68,494	98,766	104,030
L50 scaffold count (bp)	73,440	40,458	3,996	2,526	2,386
NG50 scaffold length (bp)	1,482	21,901	78,207	97,997	101,761
scaffold %A	32.6	27.16	24.07	26.54	27.22
scaffold %C	17.38	14.17	12.54	13.83	14.18
scaffold %G	17.15	14.16	12.54	13.84	14.19
scaffold %T	32.86	27.17	24.05	26.52	27.2
scaffold %N	0	17.34	26.81	19.27	17.21
CEGMA partial hits	220 (89%)	237 (96%)	242 (98 %)	241 (89%)	238 (96%)
CEGMA complete hits	127 (51%)	189 (76%)	214 (86 %)	220 (89%)	208 (84%)
Contig Measure	v0.1	v0.2	v0.3	v0.4	v0.5
Number of contigs	-	358,807	180,582	120,753	118,931
Total size of contigs (Mb)	-	1214.9	719.1	706.6	717.6
Longest contig	-	141,171	174,200	219,535	209,591
Shortest contigs	-	56	500	500	500
Number of contigs >1K	-	197,676	96,147	69,334	68,601
Number of contigs >10K	-	33,093	19,217	19,491	18,848
Number of contigs >100K	-	5	21	123	208
Mean contig size	-	3,386	3,982	5,852	6,034
Median contig size	-	1,156	1,084	1,246	1,228
N50 contig length	-	9,493	14,766	22,633	25,305
L50 contig count	-	35,755	12,859	8,788	7,924
NG50 contig length	-	13,224	9,819	15,556	17,965
contig %A	-	32.86	32.89	32.87	32.87
contig %C	-	17.14	17.13	17.13	17.13
contig %G	-	17.13	17.13	17.14	17.14
contig %T	-	32.86	32.85	32.85	32.85
contig %N	-	0.01	0	0	0

Supplementary Table 4: Summary of genic and repetitive element content of the ash genome reference assembly

Attribute		% of genome assembly
Gene Count	38,852	-
Total transcripts	50,743	-
Transcripts per gene	1.31	-
Transcript span (incl. intron) (bp)	5168.84	-
Transcript mean size cDNA (bp)	1,726	-
Transcript median size cDNA (bp)	1,522	-
Min cDNA (bp)	89	-
Max cDNA (bp)	16,883	-
Total (distinct) exons	236,602	-
Exons per transcript	6.07	-
Exon mean size * (bp)	302	-
CDS mean size * (bp)	222	-
Transcript mean size CDS (bp)	1,197	-
Transcript median size CDS (bp)	1,002	-
Min CDS (bp)	89	-
Max CDS (bp)	16,488	-
5UTR mean size * (bp)	174	-
3UTR mean size * (bp)	296	-
Class I (retrotransposons)**	224,209,197	25.87
LTR (bp)	214,958,285	24.8
LTR Copia (bp)	106,267,836	12.26
LTR Gypsy (bp)	100,315,432	11.57
LTR Retrovirus (bp)	9,370	0
LTR other (bp)	8,365,647	0.97
LINE (bp)	9,042,417	1.04
LINE L1 (bp)	6,610,333	0.76
LINE RTE (bp)	180,382	0.02
LINE other (bp)	2,251,702	0.26
SINE (bp)	208,495	0.02
Class II (DNA transposons) (bp)	32,907,896	3.8
Unclassified interspersed repeats (bp)	52,705,916	6.08
Tandem repeats (bp)	1,335,055	0.15
Total Repeats (bp)	311,158,064	35.9

* Based on distinct exons

**Categories of repeats are according to Wicker *et al.* A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8, 973–982 (2007).

Supplementary Table 5: Results of the BLAST search (E-value $1e-10$) against the combined custom ash repeat library for reads from the Illumina short insert libraries that fail to map back to the ash genome assembly.

Illumina Dataset	Number of reads	Number of reads with hits (%)
200 bp insert library, single reads	5,457,321	2,090,433 (38.31)
200 bp insert library, paired reads	9,261,078	1,644,689 (17.76)
300 bp insert library, single reads	3,210,763	1,335,378 (41.59)
300 bp insert library, paired reads	3,396,074	456,321 (13.44)
500 bp insert library, single reads	6,451,690	2,586,319 (40.09)
500 bp insert library, paired reads	4,751,260	593,683 (12.50)
All unmapped single reads	15,119,774	6,012,130 (39.76)
All unmapped paired reads	17,408,412	2,694,693 (15.48)
All unmapped reads	32,528,186	8,706,823 (26.77)

Supplementary Table 6: Statistics for RNA sequencing reads used in gene annotation.

Sample	Number of read pairs	Read length (bp)	Adapter / quality filtered reads (%)	rRNA reads removed (%)	Total reads removed (%)	Post-filtered read pairs
Selfed leaf	33,223,016	100	427,156 (1.29%)	323,725 (0.99%)	750,881 (2.26%)	32,472,135
Parent leaf	17,295,583	100	249,662 (1.44%)	291,272 (1.71%)	540,934 (3.13%)	16,754,649
Parent Cambium	32,544,085	100	467,632 (1.44%)	1,346,156 (4.20%)	1,813,788 (5.57%)	30,730,297
Parent Flower	37,354,933	100	488,566 (1.31%)	254,943 (0.69%)	743,509 (1.99%)	36,611,424
Parent Root	40,293,527	100	528,131 (1.31%)	556,007 (1.40%)	1,084,138 (2.69%)	39,209,389

Supplementary Table 7: RNA sequencing reads with hits (E-value $1e-10$) from the BLAST search against the combined custom ash repeat library

Sample	Number of trimmed reads*	Single reads with BLAST hit	Paired reads with BLAST hit**	Total number (%) of reads with BLAST hit
Selfed leaf	61,965,173	33,406	756,250	789,656 (1.27)
Parent leaf	32,194,435	18,708	403,422	422,130 (1.31)
Parent cambium	60,727,083	49,593	1,139,001	1,188,594 (1.96)
Parent flower	69,761,096	44,036	1,010,884	1,054,920 (1.51)
Parent root	75,293,284	39,945	954,733	994,678 (1.32)

* Reads trimmed with CLC; includes both paired reads and reads where one of the pair was discarded during quality trimming. See 'RNA Sequencing' in Methods for further details.

** Number of individual reads from intact pairs with hits. BLAST was performed on individual reads, therefore, either one or both of the reads in a pair may have a hit.

Supplementary Table 8: Summary of protein alignments used in gene annotation.

Dataset	Total Proteins	Source	Proteins Aligned to <i>F. excelsior</i> genome (%)	Total Protein Alignments to <i>F. excelsior</i> genome
<i>Amborella trichopoda</i> (v1.22)	27,313	Ensembl	8,500 (31.12%)	25,796
<i>Arabidopsis thaliana</i> (v167)	35,386	Phytozome	15,760 (44.54%)	48,376
<i>Fraxinus pennsylvanica</i> (r120313)	52,899	Hardwood Genomics	49,599 (93.76%)	95,017
<i>Mimulus guttatus</i> (v140)	28,282	Phytozome	19,083 (67.47%)	55,515
<i>Pinus taeda</i> (r130614)	33,595	NCBI	18,983 (56.51%)	88,046
<i>Poplar trichocarpa</i> (v210)	73,013	Phytozome	42,454 (58.15%)	129,186
<i>Solanum lycopersicum</i> (v225)	34,727	Phytozome	18,134 (52.22%)	54,941
<i>Solanum tuberosum</i> (v206)	51,472	Phytozome	26,619 (51.72%)	80,202
<i>Vitis vinifera</i> (v145)	26,346	Phytozome	14,774 (56.08%)	45,439

Supplementary Table 9: Paralog groups used for Ks analyses.

Number of genes in group	No. of groups in ash	No. groups in tomato	No. groups in mimulus	No. groups in grape	No. groups in Utricularia	No. groups in coffee	No. groups in olive (transcriptome)
2	3,657	2,547	2,569	1,859	2,042	1,734	2,175
3	1,350	838	790	598	616	546	675
4	747	438	358	259	244	260	331
5	440	236	199	158	117	170	148
6	273	145	126	89	67	88	92
7	148	106	87	83	35	61	50
8	121	82	70	38	34	42	38
9	69	57	50	47	12	29	26
10+	402	262	227	182	72	189	118
Total	7,207	4,711	4,476	3,313	3,239	3,119	3,653

Supplementary Table 10: Data used in the OrthoMCL analysis. Where necessary, mitochondrial and plastid genes were removed from downloaded protein sets.

Species	Number of protein sequences	Assembly/ Annotation version	File name/Dataset name	Source
<i>Arabidopsis thaliana</i>	27,2061	TAIR10	TAIR10_pep_20110103_representative_gene_model_updated.fasta	TAIR
<i>Amborella trichopoda</i>	26,846	v1.0	Atrichopoda_291_v1.0.protein_primaryTranscriptOnly.fa	Phytozome 10.3
<i>Coffea canephora</i>	25,574	v1.0	coffea_pep.faa	coffee-genome.org
<i>Fraxinus excelsior</i>	38,852	BATG-0.5/v2	Fraxinus_excelsior_38873_TGAC_v2.longestCDStranscript.gff3.pep.fa	This study
<i>Mimulus guttatus</i>	28,140	v2.0	Mguttatus_256_v2.0.protein_primaryTranscriptOnly.fa.gz	Phytozome 10.3
<i>Medicago truncatula</i>	50,894	Mt4.0/v1	Mtruncatula_285_Mt4.0v1.protein_primaryTranscriptOnly.fa.gz	Phytozome 10.3
<i>Pinus taeda</i>	34,059	v1.01/v2	ptaeda.v1.01.scaffolds.trimmed.all.maker.proteins.highq_whole.fasta	loblolly.ucdavis.edu
<i>Populus trichocarpa</i>	41,335	v3/v3.0	Ptrichocarpa_210_v3.0.protein_primaryTranscriptOnly.fa.gz	Phytozome 10.3
<i>Solanum lycopersicum</i>	34,727	SL2.40/ITAGv2.3	Slycopersicum_225_iTAGv2.3.protein_primaryTranscriptOnly.fa	Phytozome 10.3
<i>Utricularia gibba</i>	28,494	v1.0	Utricularia_gibba_PEP.fasta	Enrique Ibarra-Laclette*
<i>Vitis vinifera</i>	26,346	Genoscope.12X	Vvinifera_145_Genoscope.12X.protein_primaryTranscriptOnly.fa	Phytozome 10.3

*File obtained directly from the author of the genome sequence publication (Enrique Ibarra-Laclette et al (2013). Nature: doi:10.1038/nature12132).

Supplementary Table 11: Provenance source location, raw and trimmed read statistics for 38 trees in the *F. excelsior* European Diversity Panel. Latitude and longitude values were taken from the TREEBREEDEx database (<http://treebreedex.eu>), except for sample 20, where coordinates were provided by R. Sykes (Forest Research, UK), and samples 21 and 38, where coordinates were provided by EDK. All reads were 151 bp before trim (251 bp in Tree35).

Sample number	Provenance Source Location	Lat/Long ¹	Mill. raw reads (approx. genome coverage)	Average read length post-trim (bp)	Mill. reads post-trim (approx. genome coverage)	Mill. reads mapped to reference using CLC (% of trimmed reads)	Mill. reads mapped to reference using Bowtie (% of raw reads)	Average coverage of the non-N genome - Bowtie
1	La Romagne, France	49.69/4.29	44.75 (7.68x)	134.9	40.31 (6.18x)	32.26 (80.03)	38.56 (86.18)	8.11
2	Valle Pesio, Italy	44.33/7.67	78.65 (13.50x)	133.2	68.05 (10.30x)	54.86 (80.61)	63.92 (81.27)	13.44
3	Athis, France	49.02/4.61	54.77 (9.40x)	134.9	50.05 (7.67x)	40.87 (81.65)	48.17 (87.95)	10.13
4	Monti Lessini, Italy	45.67/11.17	58.50 (10.04x)	134.7	53.89 (8.25x)	44.14 (81.9)	52.19 (89.22)	10.98
5	Loch Tay, UK	56.57/-4.06	62.17 (10.67x)	134.9	57.06 (8.75x)	46.35 (81.23)	54.84 (88.22)	11.53
6	Hoge Bos, Belgium	50.83/2.95	57.66 (9.89x)	134.6	49.57 (7.58x)	40.01 (80.71)	48.04 (83.32)	10.10
7	Szczecinek, Poland	53.70/16.68	59.69 (10.24x)	135.7	52.13 (8.04x)	42.60 (81.73)	50.58 (84.76)	10.64
8	Wytham Woods, UK	51.77/-1.33	52.28 (8.97x)	135.5	46.74 (7.20x)	38.52 (82.40)	45.73 (87.5)	9.62
9	Dourdan, France	48.51/1.97	47.92 (8.22x)	135.6	38.80 (5.98x)	31.87 (82.15)	37.94 (79.18)	7.98
10	Kaisiadorys, Lithuania	54.89/24.37	70.61 (12.12x)	136.0	60.99 (9.43x)	49.28 (80.80)	58.14 (82.34)	12.23
11	Cadore, Italy	46.42/12.25	53.78 (9.23x)	136.3	50.38 (7.80x)	41.42 (82.21)	47.95 (89.17)	10.09
12	Settrington, UK	54.12/-0.71	71.41 (12.25x)	136.3	66.78 (10.34x)	54.50 (81.62)	63.16 (88.45)	13.28
13	Huttenheim, France	48.37/7.53	81.45 (13.98x)	130.9	71.94 (10.70x)	59.52 (82.74)	72.73 (89.29)	15.30
14	Rabštejn, Czech Rep.	49.56/17.15	60.21 (10.33x)	130.6	50.97 (7.56x)	41.97 (82.33)	51.96 (86.31)	10.93
15	Saint-Paul-de-Salers, France	45.12/2.53	61.41 (10.54x)	130.7	53.86 (8.00x)	44.10 (81.88)	54.17 (88.2)	11.39
16	Ravnholt, Denmark	55.26/10.58	72.69 (12.47x)	130.5	63.48 (9.41x)	51.93 (81.81)	64.42 (88.63)	13.55
17	Abetone, Italy	44.17/10.67	71.21 (12.22x)	130.9	62.24 (9.26x)	50.60 (81.30)	62.78 (88.17)	13.20
18	Włoszczowa,	50.51/19.01	86.04	130.6	74.01	60.76 (82.10)	75.01	15.78

	Poland		(14.76x)		(10.98x)		(87.19)	
19	Donadea, Rep. of Ireland	53.21/-6.45	60.52 (10.38x)	131.1	51.45 (7.66x)	42.44 (82.49)	52.87 (87.37)	11.12
20	Smolensk, Russia	54.20/32.00	89.96 (15.44x)	131.9	79.74 (11.95x)	65.49 (82.13)	80.28 (89.24)	16.88
21	Stjernebjerg, Langesø, Denmark	55.44/10.19	77.90 (13.37x)	131.4	61.26 (9.15x)	50.12 (81.82)	62.49 (80.23)	13.14
22	F. D. de Marly, France	48.89/2.04	65.46 (11.23x)	132.1	55.16 (8.28x)	45.30 (82.13)	55.38 (84.61)	11.65
23	Mircze, Poland	50.65/23.50	73.19 (12.56x)	131.8	63.78 (9.55x)	52.47 (82.27)	64.13 (87.63)	13.49
24	Le Hazey, France	49.16/1.28	59.65 (10.24x)	131.8	52.57 (7.87x)	43.49 (82.72)	52.94 (88.77)	11.14
25	Zeimelis, Lithuania	56.16/24.02	48.47 (8.32x)	134.0	42.04 (6.40x)	34.28 (81.53)	41.21 (85.04)	8.67
26	Monterolier, France	49.63/1.36	76.69 (13.16x)	134.1	67.32 (10.26x)	55.34 (82.20)	66.52 (86.75)	13.99
27	Enniskillen, UK	54.14/-7.28	62.66 (10.75x)	134.3	56.84 (8.67x)	46.63 (82.03)	55.37 (88.38)	11.65
28	Bois de Rosée, Belgium	50.23/4.72	72.88 (12.51x)	133.9	65.42 (9.95x)	53.74 (82.14)	64.5 (88.51)	13.57
29	Karlsruhe, Germany	48.98/8.30	65.36 (11.22x)	133.7	56.35 (8.56x)	46.00 (81.64)	55.52 (84.95)	11.68
30	Saint-Gatien, France	49.35/0.14	66.73 (11.45x)	134.4	58.89 (8.99x)	47.99 (81.49)	57.6 (86.33)	12.11
31	Aunais-en-Bazois, France	47.12/3.71	51.59 (8.85x)	132.6	36.96 (5.57x)	29.79 (80.59)	36.88 (71.5)	7.76
32	Haderslev Østerskov, Denmark	55.29/9.54	49.90 (8.56x)	133.7	43.01 (6.53x)	34.78 (80.87)	42.54 (85.26)	8.95
33	Bregentved, Denmark	55.34/11.96	59.79 (10.26x)	133.3	48.14 (7.29x)	39.13 (81.28)	47.89 (80.1)	10.07
34	Uncertain: Haguenau in France or Aisances in Belgium	n/a	64.44 (11.06x)	133.5	52.99 (8.04x)	42.82 (80.82)	51.81 (80.41)	10.90
35	Vassy, France	48.86/-0.70	54.51 (9.35x)	134.3	49.41 (7.54x)	40.07 (81.09)	48.02 (88.1)	10.10
36	Curraghchase, Rep. of Ireland	52.36/-8.53	59.89 (10.28x)	134.1	53.78 (8.20x)	43.93 (81.69)	52.92 (88.37)	11.13
37	Herzogtum Lauenburg, Germany	53.73/10.72	48.05 (8.24x)	134.0	42.49 (6.47x)	34.62 (81.49)	41.88 (87.17)	8.81
38	Tree35, Sorø, Denmark	55.39/11.59	43.75 (12.48x)	223.2	38.53 (9.77x)	30.07 (78.03)	--	--

Supplementary Table 12. Predicted effect of genome-wide SNPs called from the European Diversity Panel of *F. excelsior*

Type	Count
intergenic region	5,632,475
upstream gene variant (within 5kbp)	2,802,559
downstream gene variant (within 5kbp)	1,863,323
intron variant	1,295,581
missense variant	259,946
synonymous variant	251,249
3 prime UTR variant	201,107
5 prime UTR variant	106,353
splice region variant & intron variant	31,889
5 prime UTR premature start codon gain variant	19,151
missense variant & splice region variant	4,581
splice region variant & synonymous variant	4,335
splice region variant	3,795
stop gained	2,997
splice acceptor variant & intron variant	1,475
splice donor variant & intron variant	1,467
start lost	564
stop lost	530
stop retained variant	394
stop gained & splice region variant	109
stop lost & splice region variant	93
initiator codon variant	71
splice region variant & stop retained variant	31
start lost & splice region variant	24
splice region variant & initiator codon variant	5
initiator codon variant & non-canonical start codon	1
stop lost & splice region variant & splice region variant	1
total	12,484,106

Supplementary Table 13. Summary results of the 47 randomly selected SNPs for validation. Pass is scored as TRUE when over 30 trees amplified the marker. Results for individual trees are given in Supplementary Data 7.

Contig	Position	Ref	Alt	Quality	Pass	Polymorphic	Effect
Contig2569	40924	A	G	445.108	TRUE	TRUE	intergenic_region
Contig536	70861	G	A	1196.83	TRUE	TRUE	downstream_gene_variant
Contig4437	46511	T	C	97.1063	TRUE	FALSE	downstream_gene_variant
Contig70	255978	G	T	1786.15	TRUE	TRUE	downstream_gene_variant
Contig832	137377	A	T	377.912	TRUE	TRUE	3_prime_UTR_variant
Contig9889	58942	A	G	167.919	TRUE	FALSE	downstream_gene_variant
Contig5214	26904	T	C	292.191	TRUE	TRUE	downstream_gene_variant
Contig2395	38845	C	A	52.208	TRUE	FALSE	intergenic_region
Contig41052	270	A	G	125.041	TRUE	FALSE	intergenic_region
Contig1810	55914	A	T	3363.72	TRUE	TRUE	intergenic_region
Contig4244	27613	C	A	182.729	TRUE	TRUE	intergenic_region
Contig225	139884	G	A	998.159	TRUE	TRUE	intergenic_region
Contig8712	17141	G	T	346.033	TRUE	FALSE	intergenic_region
Contig112	188152	T	C	5750.62	FALSE	FALSE	missense_variant
Contig3188	19191	A	G	1316.64	TRUE	TRUE	missense_variant
Contig1169	11529	G	A	773.611	TRUE	FALSE	missense_variant
Contig3708	31929	A	G	1435.39	TRUE	TRUE	missense_variant&splice_region_variant
Contig52	257647	G	A	7174.26	TRUE	TRUE	missense_variant
Contig1833	8348	A	G	3785.39	TRUE	TRUE	missense_variant
Contig308	258788	C	T	1266.31	TRUE	TRUE	missense_variant
Contig3441	105934	A	G	1181.45	TRUE	TRUE	missense_variant
Contig136	29781	A	G	181.218	TRUE	TRUE	missense_variant
Contig53446	16720	A	G	2645.79	TRUE	TRUE	missense_variant
Contig4527	62938	C	T	6644.51	TRUE	TRUE	missense_variant
Contig3065	4053	G	A	2001.58	TRUE	TRUE	missense_variant
Contig207	5712	T	A	1359.45	TRUE	TRUE	intergenic_region
Contig5217	51255	G	A	2114.53	TRUE	TRUE	synonymous_variant
Contig1749	87622	T	G	1400.95	FALSE	TRUE	splice_region_variant&intron_variant
Contig56153	40340	C	A	624.346	TRUE	TRUE	stop_gained
Contig281	61829	C	T	7604.66	TRUE	TRUE	5_prime_UTR_variant
Contig497	67247	G	A	1304.16	TRUE	TRUE	stop_gained
Contig11696	8651	A	T	6009.17	TRUE	TRUE	downstream_gene_variant
Contig510	177482	A	T	334.743	TRUE	TRUE	missense_variant
Contig1510	30461	G	A	2928.75	TRUE	TRUE	stop_gained
Contig1861	159968	C	T	3986.3	TRUE	TRUE	stop_gained
Contig958	36608	G	T	588.347	TRUE	TRUE	upstream_gene_variant
Contig645	56724	G	A	1995.45	TRUE	TRUE	stop_gained
Contig1915	132716	G	A	961.81	TRUE	TRUE	downstream_gene_variant
Contig2758	30618	G	A	266.378	TRUE	TRUE	intron_variant
Contig4762	11030	A	C	1720.11	TRUE	TRUE	intron_variant
Contig4102	6779	G	A	1969.01	TRUE	TRUE	upstream_gene_variant

Contig68220	248	A	G	1267.1	FALSE	TRUE	intergenic_region
Contig5232	32074	A	G	125	TRUE	TRUE	upstream_gene_variant
Contig4644	3806	G	A	120.196	TRUE	FALSE	upstream_gene_variant
Contig1427	94985	T	C	205.472	TRUE	FALSE	upstream_gene_variant
Contig1962	3396	C	T	105.907	TRUE	FALSE	upstream_gene_variant
Contig3003	65637	G	A	5516.64	TRUE	TRUE	upstream_gene_variant

Supplementary Table 14. Timeline of the project

Date	Milestone
24/10/12	Ash dieback is discovered for the first time in a natural woodland in the UK
29/10/12	UK Environment Secretary bans the import of ash trees
02/11/12	UK Prime Minister's emergency committee "Cobra" meet to discuss ash dieback response
17/12/12	NERC Grant NE/K01112X/1 "Ash tree genomics: an urgent need" funded
24/01/13	Sample taken from reference tree 2451S at Earth Trust
28/01/13	DNA extracted from reference tree
05/02/13	Genome sequencing commences at Eurofins on 2451S
20/03/13	454 Sequence Data for 2451S received at QMUL
01/03/13	BBSRC/Defra Grant BBS/E/J/000CA523 "An open consortium for molecular understanding of ash dieback disease" (Nornex) funded
22/04/13	BATG-0.1-Newbler released online at www.ashgenome.org
01/05/13	Illumina data for 200bp, 300bp and 500bp paired-end and 8kb mate pair libraries of 2451S received at QMUL from Eurofins
01/05/13	Sampling of leaves from Danish Scored Panel for RNA extraction for association transcriptomics
24/05/13	Tissue samples for transcriptome sequencing collected from 2451S and its parent tree in the Cotswolds
10/06/13	Transcriptomic RNA from 2451S and parental tree sent to QMUL Genome Centre
11/06/13	BATG-0.2-CLCbio released online at www.ashgenome.org
03/07/13	Illumina data for 40kb mate pair libraries of 2451S received at QMUL from Eurofins
31/07/13	Illumina data for 3kb and 20kb mate pair libraries of 2451S received at QMUL from Eurofins
23/09/13	BATG-0.3-CLCbioSSPACE released online at www.ashgenome.org
27/10/13	Transcriptome sequence data received from QMUL Genome Centre
11/11/13	BATG-0.4-CLCbioSSPACE released online at www.ashgenome.org
13/03/14	Samples taken for European Diversity Panel from Earth Trust "Realising Ash's Potential" Provenance trial
01/05/14	Sampling of leaves from Danish Test Panel for RNA extraction for association transcriptomics
29/05/14	European Diversity DNA samples sent from QMUL to TGAC for Illumina Sequencing
29/05/14	2451S DNA sample sent from QMUL to TGAC for Nextera long mate pair MiSeq sequencing
24/06/14	Annotation Version 2 - BATG-0.4 released online at www.ashgenome.org
26/06/14	Transcriptome data sent from QMUL to TGAC for annotation of genome
22/08/14	European Diversity Panel Illumina Read Data received at QMUL from TGAC
01/09/14	Samples collected from Danish trees for metabolomic profiling
10/02/15	Nextera long-mate pair Illumina MiSeq data received at QMUL from TGAC
26/02/15	Annotation Version 3 - BATG-0.4 released online at www.ashgenome.org
18/05/15	Samples collected from Earth Trust for British Screening Panel

09/09/15	DNA samples sent to TGAC from QMUL for KASP analysis
29/10/15	BATG-0.5 released online at www.ashgenome.org
21/11/15	Annotation Version 4 Gene Models sent to York University for re-analysis of Gene Expression Markers
14/12/15	KASP results received at QMUL from TGAC
21/01/16	qRT-PCR analyses of GEMs complete for British Screening Panel
08/02/16	Annotation Version 4 - BATG-0.5 released online at www.ashgenome.org
15/02/16	Validation of SSR marker panels complete
23/02/16	Manuscript submitted to Nature
10/06/16	Revised manuscript submitted to Nature
18/08/16	Re-revised manuscript submitted to Nature
05/10/16	Further revision submitted to Nature