Supplementary Tables for

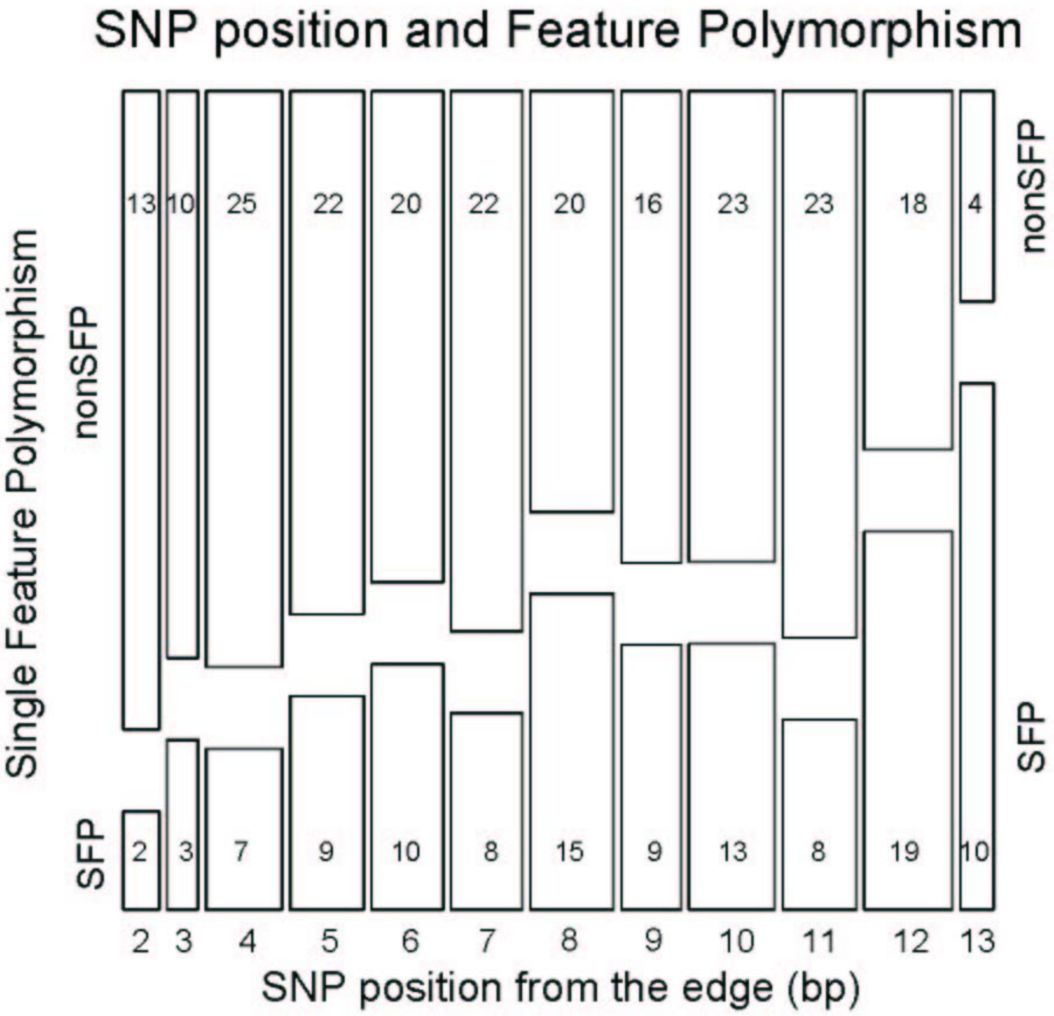
Large Scale Identification of Single Feature Polymorphisms in Complex Genomes.
Justin O. Borevitz, David Liang, David Plouffe, Hur-Song Chang, Tong Zhu, Detlef Weigel, Charles C. Berry, Elizabeth Winzeler, Joanne Chory

Supplement Table 1

General		number of features
Total PM, MM, and controls	534X534=	285156
PM excluding controls	100%	131822
Unique match	83%	109825
High - stringency unique match	79%	103860
Overlapping features (22,23,24, or 25 bp)		
No overlap > 22 bp	61%	80986
Two features overlap	7%	8606
Three features overlap	1%	1835
Four features overlap	1%	1313
Five features overlap	0%	12
Six features overlap	0%	3
Unique positions	**70%**	**92924**
Chromosome		
1	15%	20317
2	25%	33241
3	5%	5984
4	21%	27172
5	5%	6210
Genes		number of genes
5 or more features		6859

Supplement Table 1. Description of features on the AtGenome1 array. Each 25 bp PM feature was aligned to the five *Arabidopsis thaliana* pseudo-chromosomes (1/7/02 release) using BLAST. Repetitive features and those that were not perfect matches were removed. We further increased the stringency by lowering the BLAST threshold so that features aligning imperfectly to multiple positions were also removed (see Methods). The exact chromosome location of each feature in this unique set was recorded. Since features were sometimes designed to cover only a small portion of a gene, they often overlapped. When features overlapped by 22, 23, 24 or 25 bp, we grouped them together so that just one test was performed on the group. Many groups consisted of 2 features, but some contained as many as 6. PM, perfect match; MM, mismatch

Supplement Figure 1



Supplement Figure 1 Description

We aligned the 340 Cereon candidate polymorphisms with the corresponding 25bp feature sequence, to determine the relationship between the SNP position and the ability of that feature to detect a polymorphism. ClustalW was used followed by manual sequence alignment. 11 features contained polymorphisms at more than 1 location and are not included in this analysis. Polymorphisms are considered symmetrical, and were scored as the distance from the nearest edge of the 25bp feature. Thus position 2 and 24 or 3 and 23, etc., are equivalent. Supplement Figure 1 shows the proportion of SFPs detected at different SNP positions. When a polymorphism is at the central base (position 13), it is more often detected as a Single Feature Polymorphism. Furthermore when a polymorphism is near the edge of the 25bp feature (position 2 or 3) it is less often identified as a SFP. Polymorphisms at intermediate distances have a similar chance of being identified. The trend between SNP position and SFP detection is significant (Chi-squared Test for Trend in Proportions, p-value = 0.0004), however other factors besides position may also be important. In addition some Cereon candidate polymorphisms may be sequencing errors. These would not show a relationship with SNP position and would therefore weaken the apparent association of SFP detection with central location in the 25 bp region.

Supplement Table 2 (part1)

PM only

SAM threshold
5% FDR

		GeneChip		
		SFPs	nonSFPs	
		3806	89118	
Sequence	817	121	696	Sensitivity
Polymorphic	340	117	223	34%
Non-polymorphic	477	4	473	

False Discovery rate: 3%
Test for independence of all factors:
Chisq = 177.34, df = 1, p-value = 1.845e-40

SAM threshold
18% FDR

		GeneChip		
		SFPs	nonSFPs	
		10627	82297	
Sequence	817	223	594	Sensitivity
Polymorphic	340	195	145	57%
Non-polymorphic	477	28	449	

False Discovery rate: 13%
Test for independence of all factors:
Chisq = 265.13, df = 1, p-value = 1.309e-59

PM - MM

SAM threshold
5% FDR

		GeneChip		
		SFPs	nonSFPs	
		3806	89118	
Sequence	817	105	712	Sensitivity
Polymorphic	340	96	244	28%
Non-polymorphic	477	9	468	

False Discovery rate: 9%
Test for independence of all factors:
Chisq = 123.04, df = 1, p-value = 1.364e-28

SAM threshold
18% FDR

		GeneChip		
		SFPs	nonSFPs	
		10627	82297	
Sequence	817	178	639	Sensitivity
Polymorphic	340	142	198	42%
Non-polymorphic	477	36	441	

False Discovery rate: 20%
Test for independence of all factors:
Chisq = 136.39, df = 1, p-value = 1.636e-31

Supplement Table 2 (part2)

MM only

SAM threshold
5% FDR

		GeneChip SFPs	GeneChip nonSFPs	
		3806	89118	
Sequence	817	64	753	Sensitivity
Polymorphic	340	57	283	17%
Non-polymorphic	477	7	470	

False Discovery rate: 11%
Test for independence of all factors:
Chisq = 64.34, df = 1, p-value = 1.048e-15

SAM threshold
18% FDR

		GeneChip SFPs	GeneChip nonSFPs	
		10627	82297	
Sequence	817	145	672	Sensitivity
Polymorphic	340	100	240	29%
Non-polymorphic	477	45	432	

False Discovery rate: 31%
Test for independence of all factors:
Chisq = 54.27, df = 1, p-value = 1.746e-13

PM and MM

SAM threshold
5% FDR

		GeneChip SFPs	GeneChip nonSFPs	
		3806	89118	
Sequence	817	109	708	Sensitivity
Polymorphic	340	104	236	31%
Non-polymorphic	477	5	472	

False Discovery rate: 5%
Test for independence of all factors:
Chisq = 149.82, df = 1, p-value = 1.894e-34

SAM threshold
18% FDR

		GeneChip SFPs	GeneChip nonSFPs	
		10627	82297	
Sequence	817	199	618	Sensitivity
Polymorphic	340	168	172	49%
Non-polymorphic	477	31	446	

False Discovery rate: 16%
Test for independence of all factors:
Chisq = 198.4, df = 1, p-value = 4.656e-45

Supplement Table 2. Perfect match and mismatch model comparisons. False Discovery Rates and Sensitivities are shown for PM only, PM-MM, MM only, and PM and MM models. Mismatch signal is subtracted from the perfect match signal in the PM-MM model. The mismatch feature is used as a second observation in the PM and MM model (see below). Chi square test statistic (Chisq).

Supplement Table 2 Description

The mismatch (MM) feature on GeneChips was designed as a control for background hybridization (Lipshutz et al. 1999; Lockhart et al. 1996). Under this assumption, the intensity from the MM feature should be subtracted from the perfect match (PM) feature to reveal the true signal of a 25 bp oligonucleotide. We tested three models using the MM feature, none of which performed better the PM alone. Difference values (PM - MM) were used to compute false discovery and sensitivity rates by comparison with sequence data. For a direct comparison with PM only scores we choose thresholds where 3806 or 10627 SFPs could be identified. After log transformation and spatial correction the MM signal was subtracted from the PM signal. Spatial correction also slightly improved the PM-MM results. The PM-MM difference values were effective at predicting markers however not as effective as PM alone (part1). PM and MM hybridization intensities are highly correlated (r=0.8). We further searched the *Arabidopsis* genome using all 25 bp MM sequences and found that a MM feature aligned uniquely to the same place as the corresponding PM feature 98.8% of the time, indicating that the MM signal was indeed quite specific and that it may not be a good control for background hybridization. Furthermore SFP analysis using MM data only was also effective at discovering markers, albeit with lower accuracy. The specificity of the MM feature suggested its use as a second observation of the PM feature. When a feature is a bona fide marker, both the PM and MM features may have a significantly lower intensity. PM together with MM data was quite effective at predicting markers, but including the MM as a second observation again was not an improvement over PM alone. Therefore we see no practical use of the MM feature in our studies.

Supplement Table 3

Mean t-statistic	Original data	Permuted data	Difference	FDR
0.2	165	120	45	73%
0.4	147	64	83	43%
0.6	137	23	114	17%
0.8	123	11	112	9%
1.0	**105**	**5**	**100**	**5%**
1.2	82	3	79	4%

Supplement Table 3. Potential deletions in 2000 clusters. A threshold was set for potential deletions by calculating the number of potential deletions in the real data and comparing it to the average number in the permuted data. The difference between the original data and permutated data and the false discovery rates were also determined. Mean t-statistic is the average of all the t-statistics in the cluster.

Supplement Table 4. Potentially deleted genes. 105 deletions were detected that cover 111 genes. The probe set(s), chromosome, location, size, number of features covered, and gene (At and BAC name) and description are shown. Deletions shown in grey were confirmed by Cereon shotgun Ler sequence.

Deletion number	Affy probe set1	Affy probe set2	chrom	chrom position start	width	number features	Annotation 1.7.02
1	18310_at		CHR1	4057090	415	5	At1g12010 F12F1.12 putative amino-cyclopropane-carboxylic acid oxidase (ACC oxidase) Strong similarity to amino-cyclopropane-carboxylic acid oxidase gb|L27664 from Brassica napus. ESTs gb|Z48548 and gb|Z48549 come from this gene; supported by cDNA: gi_15450652_gb_AY052694.1_
2	16182_at		CHR1	8363359	2121	15	At1g23610 F28C11.24 F5O8.17 putative OBP32pep protein
	13375_at		CHR1				At1g23620 F28C11.21 F5O8.18 pseudogene, hypothetical protein
3	19026_at		CHR1	9480330	251	7	At1g27285 F17L21.7 putative polyprotein similar to GB:AAC02666
4	18161_at		CHR1	11064863	238	8	At1g31030 F17F8.5 putative reverse transcriptase similar to a multitude of putative reverse transcriptase sequences
5	19691_at		CHR1	11087815	351	4	At1g31070 F17F8.1 UDP-N-acetylglucosamine pyrophosphorylase-like protein similar to sperm associated antigen 2, GB:4507759 from [Homo sapiens], similar to ESTs gb|N95844, gb|AI162709, gb|N65104, dbj|AU031622, and gb|T45941

6	16439_at		CHR1	11311477	269	10	At1g31580 F27M3.20 F27M3.20 ORF1, putative similar to ORF1 GI:457716 from (Arabidopsis thaliana); supported by cDNA: gi_16649160_gb_AY059950.1_
7	20025_at		CHR1	12457181	171	8	At1g34200 F23M19.12 F12G12.25 unknown protein ; supported by cDNA: gi_13605868_gb_AF367333.1_AF367333
8	12386_at		CHR1	16532954	135	7	At1g44970 F27F5.6 peroxidase, putative similar to peroxidase GI:993004 from [Mercurialis annua];supported by full-length cDNA: Ceres:26375.
9	18927_at		CHR1	18970876	732	13	At1g52100 F5F19.16 jasmonate inducible protein, putative similar to jasmonate inducible protein GI:9279645 from [Arabidopsis thaliana]; contains non-consensus AG donor splice site at exon3
10	16834_at	15944_at	CHR1	19734201	1043	13	At1g53950 T18A20.25 F15I1.3 ubiquitin, putative similar to ubiquitin GI:159038 from [Euplotes eurystomus]
11	16384_at		CHR1	20285376	359	7	At1g55430 T5A14.18 hypothetical protein similar to putative CHP-rich zinc finger protein GB:CAB80684 GI:7268575 from [Arabidopsis thaliana]
12	16073_f_at		CHR1	20821348	231	6	At1g56650 F25P12.92 anthocyanin2, putative similar to anthocyanin2 (An2) GI:7673088 from [Petunia integrifolia]; supported by cDNA: gi_3941507_gb_AF062908.1_AF062908

13	18166_at	CHR1	20905035	276	8	At1g57560 T8L23.3 DNA-binding protein, putative similar to DNA-binding protein GI:19058 from [Hordeum vulgare]; supported by full-length cDNA: Ceres:250386.
14	17787_at	CHR1	21753373	173	5	At1g60020 T2K10.7 polyprotein, putative similar to polyprotein GI:2865437 from [Arabidopsis arenosa]
15	18163_at	CHR1	21883140	437	7	At1g60310 T13D8.19 pseudogene, putative non-LTR retroelement reverse transcriptase
16	19796_at	CHR1	21907131	254	4	At1g60400 T13D8.27 hypothetical protein similar to heat shock transcription factor HSF30-like protein GI:9759202 from [Arabidopsis thaliana]
17	15032_at	CHR1	22239091	1093	6	At1g61250 F11P17.4 secretory carrier membrane protein identical to secretory carrier membrane protein GI:7109228 from [Arabidopsis thaliana];supported by full-length cDNA: Ceres:27197.
18	12363_at	CHR1	22602677	283	7	At1g62090 F19K23.4 hypothetical protein similar to putative receptor-like protein kinase GI:3461841 from [Arabidopsis thaliana]
19	13840_at	CHR1	24813600	2192	15	At1g67240 F1N21.6 mutator-like transposase, putative similar to mutator-like transposase GI:5306250 from [Arabidopsis thaliana]
	19198_at	CHR1				At1g67250 F1N21.7 unknown protein ;supported by full-length cDNA: Ceres:27632.

20	14096_at	CHR1	28573805	194	5	At1g76960 F22K20.6 unknown protein ; supported by cDNA: gi_13926184_gb_AF370567.1_AF370567
21	12194_at	CHR2	151414	168	7	At2g01310 F10A8.19 unknown protein
22	19048_at	CHR2	337068	266	5	At2g01780 T8O11.5 putative S-locus glycoprotein
23	12640_s_at	CHR2	497977	161	5	At2g02060 F14H20.13 F5O4.17 unknown protein
24	16722_at	CHR2	1079050	679	12	At2g03560 F19B11.1 F-box protein family, AtFBX7 contains F-box domain PF:00646
25	15916_at	CHR2	1525190	139	4	At2g04380 T1O3.21 unknown protein
26	14125_at	CHR2	2900217	184	6	At2g06990 T4E14.10 unknown protein ; supported by cDNA: gi_16024935_gb_AY050658.1_
27	14792_at	CHR2	4365859	128	4	At2g11140 T13H18.4 putative retroelement pol polyprotein
28	16714_at	CHR2	4366037	2729	9	At2g11150 T13H18.5 putative replication protein A1
29	20304_at	CHR2	4426468	5527	19	At2g11240 T13H18.14 putative non-LTR retroelement reverse transcriptase
	20672_at	CHR2				At2g11260 T13H18.16 unknown protein
30	20182_at	CHR2	5636422	105	4	At2g13650 T10F5.19 putative vanadate resistance protein
31	20192_s_at	CHR2	5641067	230	6	At2g13660 T10F5.25 F13J11.1 unknown protein
32	14927_at	CHR2	5673036	407	7	At2g13720 F13J11.7 unknown protein
33	17763_at	CHR2	5805349	102	4	At2g13940 F9B22.5 putative retroelement pol polyprotein contains a Zn-finger CCHC type domain (prosite:QDOC50158) and a DDE integrase signature motif

34	16272_at		CHR2	6006070	290	9	At2g14290 T1O16.12 hypothetical protein predicted by genscan and genefinder
35	19239_at		CHR2	6371867	187	6	At2g14930 T26I20.9 putative retroelement pol polyprotein
36	19264_at		CHR2	6671461	496	16	At2g15410 F26H6.7 putative retroelement pol polyprotein
37	17829_at		CHR2	6727680	246	6	At2g15540 F9O13.9 putative non-LTR retroelement reverse transcriptase
38	19813_at		CHR2	6939935	440	13	At2g16110 F7H1.13 putative non-LTR retrolelement reverse transcriptase
39	15233_at		CHR2	7009545	1096	7	At2g16300 F16F14.20 hypothetical protein predicted by genscan
40	13190_s_at	17486_at	CHR2	7513871	575	10	At2g17420 F5J6.18 putative thioredoxin reductase
41	20394_at		CHR2	7599860	218	7	At2g17590 MJB20.15 unknown protein
42	20318_at		CHR2	7600101	207	7	At2g17600 MJB20.16 unknown protein
43	19777_s_at		CHR2	7603398	3424	17	At2g17610 T19E12.5 putative non-LTR retroelement reverse transcriptase
44	17327_at		CHR2	7730596	315	7	At2g17910 T13L16.7 putative non-LTR retroelement reverse transcriptase
45	16144_s_at		CHR2	7856383	105	5	At2g18170 F8D23.5 MAP kinase (ATMPK7) identical to GB:Q39027; MAP-mitogen activated protein kinase; supported by cDNA: gi_457405_dbj_D21843.1_ATHATMPK7
46	13888_s_at		CHR2	8293352	256	10	At2g19230 F27F23.3 putative receptor-like protein kinase
47	16777_at		CHR2	8924757	327	6	At2g20850 F5H14.18 putative LRR receptor protein kinase ; supported by cDNA: gi_15810274_gb_AY056176.1_

48	19865_g_at	CHR2	9056723	113	8	At2g21270 F3K23.3 putative ubiquitin fusion-degradation protein ;supported by full-length cDNA: Ceres:34470.
49	20104_at	CHR2	9282818	1737	6	At2g21900 F7D8.22 putative WRKY-type DNA binding protein
50	20483_at	CHR2	9980905	134	5	At2g23570 F26B6.22 putative acetone-cyanohydrin lyase
51	18416_at	CHR2	10214336	350	8	At2g24150 F27D4.6 hypothetical protein predicted by genefinder; supported by cDNA: gi_13877832_gb_AF370179.1_AF370179
52	20426_at	CHR2	10442884	163	5	At2g24680 F25P17.2 unknown protein
53	13331_at	CHR2	11040466	200	5	At2g26030 T19L18.16 hypothetical protein predicted by genscan
54	17693_at	CHR2	11389579	1023	13	At2g26820 F12C20.14 similar to avrRpt2-induced protein 1
55	15286_s_at	CHR2	11815859	483	4	At2g27870 T1E2.21 F15K20.3 putative reverse transcriptase
56	18144_at	CHR2	13171713	328	8	At2g31080 T16B12.11 putative non-LTR retroelement reverse transcriptase
57	20001_at	CHR2	14027694	513	4	At2g33220 F25I18.4 unknown protein predicted by genefinder and 5' EST; supported by full-length cDNA: Ceres: 954.
58	17458_at	CHR2	15542346	450	9	At2g37110 T2N18.13 unknown protein ;supported by full-length cDNA: Ceres:123128.
	19972_at	CHR2				At2g37120 T2N18.12 unknown protein ; supported by full-length cDNA: Ceres: 1517.
59	17760_at	CHR2	18599894	160	8	At2g45230 F4L23.26 putative non-LTR retroelement reverse transcriptase

60	17334_at		CHR3	9011584	4213	21	At3g24675 MSD24.5 pseudogene, reverse transcriptase similar to reverse transcriptase GB:S65812 from [Arabidopsis thaliana]
	18364_at		CHR3				At3g24680 MSD24.2 zinc-finger protein, putative almost identical to zinc-finger protein GB:S65811 from [Arabidopsis thaliana]
61	12237_at		CHR3	18896817	1428	7	At3g50810 F18B3.90 putative protein
62	17833_at		CHR4	823853	411	6	At4g01910 T7B11.17 putative CHP-rich zinc finger protein similar to several A. thaliana CHP-rich proteins
63	16829_at		CHR4	828481	183	7	At4g01920 T7B11.18 putative CHP-rich zinc finger protein similar to A. thaliana CHP-rich proteins
64	14573_at		CHR4	1040562	985	5	At4g02360 T14P8.17 putative protein
	13648_at		CHR4				At4g02370 T14P8.18 coded for by A. thaliana cDNA R64991 ;supported by full-length cDNA: Ceres:19543.
65	18762_at	18763_g_at	CHR4	1096573	1196	4	At4g02490 T10P11.21 T14P8.10 hypothetical protein similar to reverse transcriptases (Pfam: rvt.hmm, score: 53.76)
66	16240_at		CHR4	1467128	169	6	At4g03340 F4C21.27 putative glycosylation enzyme similar to B. taurus core2-GLCNAC-transferase, GenBank accession number U41320
67	19104_at		CHR4	1541677	181	4	At4g03470 F9H3.10 putative protein
68	19072_at		CHR4	1541897	6789	23	At4g03480 F9H3.11 hypothetical protein similar to A. thaliana hypothetical protein F21B7.8, GenBank accession number AC002560

69	15239_at	CHR4	1548706	576	5	At4g03490 F9H3.12 hypothetical protein similar to A. thaliana hypothetical protein F21B7.8, GenBank accession number AC002560
70	19324_at	CHR4	1555288	1573	9	At4g03500 F9H3.13 hypothetical protein
	17552_s_at	CHR4				At4g03510 F9H3.14 RMA1 RING zinc finger protein identical to RMA1 gi:3164222; supported by cDNA: gi_3164221_dbj_AB008518.1_AB008518
71	14272_at	CHR4	1742879	393	5	At4g03800 T5H22.5 hypothetical protein
72	17635_at	CHR4	1851621	160	4	At3g31440 T22C2.4 pseudogene, putative helicase
73	17834_at	CHR4	1923067	678	8	At4g04000 T24H24.17 putative reverse transcriptase
74	13871_at	CHR4	2145670	139	7	At4g04390 T19B17.11 hypothetical protein
75			2146011	151	6	
76	17371_at	CHR4	2239307	264	4	At4g04500 T26N6.11 putative receptor-like protein kinase
77	17813_at	CHR4	3558830	195	7	At4g07770 F5K24.12 pseudogene
78	19350_at	CHR4	5089939	6953	20	At4g09690 F17A8.40 T25P22.5 putative protein hypothetical protein T15B16.6 - Arabidopsis thaliana,PIR2:T01999
	17333_at	CHR4				At4g09710 F17A8.60 RNA-directed DNA polymerase -like protein RNA-directed DNA polymerase (EC 2.7.7.49)- Arabidopsis thaliana retrotransposon Ta11-1,PIR2:S65812

79	17200_at	CHR4	5124099	512	6	At4g09780 F17A8.130 putative protein hypothetical protein F9D12.8 - Arabidopsis thaliana,PID:g3319350
80	17731_at	CHR4	5599291	194	4	At4g10780 T12H20.8 putative disease resistance protein similar to RPS2, GenBank accession number U14158 functional catalog ID=11.05.01
81			5599562	272	8	
82	16869_at	CHR4	6241576	817	16	At4g12180 T4C9.20 putative reverse transcriptase other putative reverse transcriptase -Arabidopsis thaliana
83	20429_at	CHR4	7262374	200	4	At4g14400 FCAALL.190 hypothetical protein ; supported by cDNA: gi_15810232_gb_AY056155.1_
84	13356_at	CHR4	7292534	131	7	At4g14470 FCAALL.217 reverse transcriptase like protein
85			7292671	230	5	
86	20498_at	CHR4	7862991	183	4	At4g15570 FCAALL.334 SEN1 like protein
87	19739_at	CHR4	7866999	231	8	At4g15590 FCAALL.338 reverse transcriptase like protein
88	16655_at	CHR4	9138091	148	4	At4g18410 F28J12.70 MuDR transposable element - like protein mudrA protein - maize transposon MuDR, PIR2:S59141
89	12561_at	CHR4	9425785	170	7	At4g19120 T18B16.90 putative protein several hypothetical proteins - Arabidopsis thaliana; supported by cDNA: gi_15320409_dbj_AB039927.1_AB039927
90	19089_at	CHR4	9602140	155	5	At4g19510 F24J7.70 TMV resistance protein N-like TMV resistance protein N, Nicotiana glutinosa, PIR2:A54810

91	15311_at	CHR4	9799030	177	6	At4g19990 F18F4.90 putative protein hypothetical protein, Arabidopsis thaliana, PATCHX:E326848
92	12341_s_at	CHR4	9842567	227	5	At4g20110 F1C12.2 F18F4.210 vacuolar sorting receptor-like protein BP-80 vacuolar sorting receptor, Pisum sativum, PATCHX:G1737222
93	18821_at	CHR4	9879008	2096	7	At4g20210 F1C12.130 cadinene synthase like protein (+)-delta-cadinene synthase isozyme XC14, Gossypiumarboreum, PIR2:S68366
94	18597_at	CHR4	10138326	159	5	At4g20860 T13K14.20 berberine bridge enzyme - like protein ; supported by cDNA: gi_14194126_gb_AF367269.1_AF367269
95	20223_at	CHR4	11117491	171	6	At4g23250 F21P8.140 protein kinase - like protein receptor kinase 1, Brassica rapa
96	13025_at	CHR4	12041772	213	6	At4g25640 L73G19.20 putative protein NorM, Vibrio parahaemolyticus, gb:AB010463
97	15330_at	CHR4	12585516	228	7	At4g27190 T24A18.140 putative protein NBS/LRR disease resistance protein - Arabidopsis thaliana, PID:g3309619
98	14755_at	CHR4	12598659	339	11	At4g27220 M4I22.30 putative protein disease resistance protein RPS2, Arabidopsis thaliana, PIR2:A54809
99	12658_at	CHR4	13021071	382	7	At4g28420 F20O9.100 tyrosine transaminase-like protein tyrosine transaminase (EC 2.6.1.5)-rat, PIR1:XNRTY

100	20450_at	CHR5	8148310	709	10	At5g24150 K12G2.2 MLE8.1 squalene monooxygenase
101	19173_at	CHR5	9097640	567	8	At5g26120 T1N24.13 arabinosidase -like protein arabinosidase asdI, Bacteroides ovatus, EMBL:BO15178
102	14876_at	CHR5	9106663	636	9	At5g26140 T1N24.16 putative protein various predicted lysine decarboxylase proteins, Arabidopsis thaliana
103	19395_at	CHR5	9439456	876	4	At5g26940 F2P16.200 putative protein various predicted bacterial DNA polymerase subunits;supported by full-length cDNA: Ceres:32265.
	13997_at	CHR5				At5g26950 F2P16.17 putative protein various predicted proteins, Arabidopsis thaliana
104	18516_at	CHR5	9492135	181	7	At5g27100 T21B4.10 ion channel -like protein ligand gated channel-like protein, Brassica napus, EMBL:AF109392
105	18585_at	CHR5	21251785	233	6	At5g53130 MFH8.6 cyclic nucleotide-regulated ion channel (emb|CAA76178.1) ; supported by cDNA: gi_13877752_gb_AF370139.1_AF370139

References

Lipshutz, R.J., S.P. Fodor, T.R. Gingeras, and D.J. Lockhart. 1999. High density synthetic oligonucleotide arrays. *Nat Genet* **21:** 20-24.

Lockhart, D.J., H. Dong, M.C. Byrne, M.T. Follettie, M.V. Gallo, M.S. Chee, M. Mittmann, C. Wang, M. Kobayashi, H. Horton, and E.L. Brown. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* **14:** 1675-1680.