Many gene and domain families have convergent fates following independent whole-genome duplication events in *Arabidopsis*, *Oryza*, *Saccharomyces* and *Tetraodon*

Andrew H. Paterson, Brad A. Chapman, Jessica C. Kissinger, John E. Bowers, Frank A. Feltus and James C. Estill

Plant Genome Mapping Laboratory and Department of Genetics, University of Georgia, Athens, GA 30602 USA

Genome duplication is potentially a good source of new genes, but such genes take time to evolve. We have found a group of 'duplication-resistant' genes, which have undergone convergent restoration to singleton status following several independent genome duplications. Restoration of duplication-resistant genes to singleton status could be important to long-term survival of a polyploid lineage. Angiosperms show more frequent polyploidization and a higher degree of duplicate gene preservation than other paleopolyploids, making them well-suited to further study of duplication-resistant genes.

Introduction

How do some lineages survive repeated duplications of their entire genomes? Although genome duplication (GD) is potentially advantageous as a primary source of genes with new functions (called 'Ohnologs' after Susumu Ohno, who proposed this hypothesis [1,2], such genes take time to evolve. Most higher organisms continuously produce aberrant, unreduced gametes at low rates, but the rarity of GD shows that their evolutionary success is exceedingly unlikely. Rapid loss and restructuring of low-copy DNA [3–6], retrotransposon activation [7,8] and epigenetic gene silencing [9–11] following GD could provide raw material for evolutionary change; however, like other mutations, the majority of such new variations might be deleterious. For newly duplicated genes to persist long enough for adaptive evolution to occur, any major maladaptive consequences of genome duplication must first be resolved.

The angiosperms (flowering plants) are an outstanding model in which to elucidate the consequences of GD. Although the 'signal' remaining from ancient GD in many animal genomes is faint [12–14], all angiosperms seem to be paleopolyploids [15], many having survived multiple duplication events [16]. Collectively, these events provide naturally occurring replications that are useful for determining the properties shared by the rare cases in which polyploid gametes spawned successful lineages. These events also provide a foundation for comparing the two major branches of the angiosperms (*Arabidopsis* is from the eudicots and *Oryza* from the monocots), for which estimates of divergence times are converging on 140–180 million years ago (MYA) [17,18]. Population genetic theory predicts that the consequences of GD can be determined at least in part by effective population size (N_e) – angiosperms, with generally small N_e values, could be more appropriate models for exploring consequences of GD in large-bodied (small N_e) eukaryotes [19] than are large N_e microbes such as yeast.

Corresponding author: Paterson, A.H. (paterson@uga.edu)
Available online 18 September 2006.

Repertoire and abundance of protein functional domains

We determined the tendencies for individual protein functional (Pfam) domains to occur among duplicates or singleton genes (see the supplementary material online) resulting from independent GD events that occurred ~60 and ~70 MYA in the lineages of *Arabidopsis* (α) [15] and *Oryza* (ρ) [16], respectively. Analysis of Pfam domains as the experimental unit also permitted us to compare these angiosperm events with GDs in yeast [20] and *Tetraodon* [21]. Between 1334 and 1704 domain types were found in the four organisms, with duplicated genes containing an average of about one domain and singletons averaging 0.15–0.40 domains (see Table 1 in the supplementary material online).

The abundance of individual domain types was very closely correlated in the two angiosperms ($r = 0.95$, $p < 0.001$). Domain abundance in yeast and *Tetraodon* was also closely correlated with those of the angiosperms ($r = 0.49$–0.61, all $p < 0.001$) and one another ($r = 0.64$, $p < 0.01$).

Retention or loss of protein functional domain-containing genes has been convergent following independent duplications

For each domain type, the fraction of the total number of occurrences found in singleton genes (% singletons) was compared with the average fraction observed across all families, using the % singletons to calculate a normal deviate (z) score. In view of the large number of gene or domain families considered, only z scores exceeding 3.3, corresponding to $p < 0.001$, were considered significant (Table 1). Across all domain types, the % singletons

Table 1. Numbers of Pfam domains retained in singletons versus duplicated genes of *Arabidopsis, Oryza, Saccharomyces* and *Tetraodon*[a]

(a) Enriched in plant singletons										
Domain	**Name**	**Function**	**A1**	**A2**	**z**	**Significance**	**R1**	**R2**	**z**	**Significance**
PF07172	Glycine-rich	Nodulins, stress response	9	1	5.01	***	12	1	5.68	***
PF05678	VQ motif	Uncertain	8	14	1.46		15	4	5.52	***
PF01585	G-patch	RNA binding	7	0	4.81	***	4	0	3.51	***
PF05754	DUF834	Unknown (G-rich)	0	0	n.a.		7	0	4.64	***
PF00588	SpoU rRNA methylase	Ribose methylation; ribosome formation	6	0	4.46	***	5	0	3.93	***
PF05347	LYR	NADH-ubiquinone oxidoreductase complex I	6	0	4.46	***	3	0	3.04	
PF02037	SAP	DNA binding	6	0	4.46	***	3	2	1.85	
PF00856	SET	Protein–protein interaction	10	4	4.28	***	5	2	2.89	
PF00400	WD	GTP hydrolysis in transmembrane signaling	23	51	1.61		26	26	4.27	***
PF00515	TPR	Protein–protein interaction	17	17	3.70	***	14	8	4.27	***
PF04535	DUF588	Unknown	5	8	1.30		8	2	4.08	***
PF01040	UbiA prenyl-transferase	Transmembrane proteins	5	0	4.07	***	1	0	1.76	
PF01479	S4	RNA binding	5	0	4.07	***	1	0	1.76	
PF00533	BRCT	DNA repair	5	0	4.07	***	2	2	1.19	
PF04055	Radical SAM	Reductive SAM cleavage	5	0	4.07	***	2	2	1.19	
PF03242	LEA	Unknown	2	2	1.27		5	0	3.93	***
PF02996	Prefoldin subunit	Molecular chaperones	4	0	3.64	***	1	0	1.76	
PF02953	Tim10/DDP zinc finger	Mitochondrial protein import	4	0	3.64	***	1	1	0.84	
PF05498	RALF[b]	Root growth arrest	4	0	3.64	***	0	2	−0.81	
PF01266	FAD-dependent oxidoreductase	FAD-dependent oxidoreductase	4	0	3.64	***	0	0	n.a.	
PF00849	RNA pseudo-uridylate	RNA modification	4	0	3.64	***	4	0	3.51	***
PF03760	LEA group 1	Desiccation tolerance	1	2	0.42		4	0	3.51	***
PF00646	F-box	Protein–protein interaction in ubiquitination	15	18	3.03		19	21	3.38	***
(b) Enriched in plant duplicates										
Domain	**Name**	**Function**	**A1**	**A2**	***z***	**Significance**	**R1**	**R2**	***z***	**Significance**
PF00069	Protein kinase	Catalysis and ATP binding	4	297	−8.99	***	2	210	−7.98	***
PF00249	Myb-like	DNA binding	2	100	−5.08	***	4	78	−4.13	***
PF00847	AP2[b]	Transcriptional activation	1	69	−4.31	***	10	46	−1.16	
PF00560	LRR_1	Protein–protein interaction	8	89	−3.49	***	5	63	−3.29	
(c) Enriched in yeast duplicates										
Domain	**Name**	**Function**	**Y1**	**Y2**	***z***	**Significance**	***Arabidopsis***	**Significance**	**Rice**	**Significance**
PF00069	Protein kinase	Catalysis and ATP binding	0	45	−8.84	***	−8.99	***	−7.98	***
PF00153	Mitochondrial carrier	Energy transfer	0	30	−7.22	***	−0.56		−0.93	
PF00096	C2H2 zinc finger	Nucleic acid binding	2	25	−5.76	***	−0.93		−1.28	
PF00083	Sugar transporter	Small solute transport	1	13	−4.38	***	−2.33		−1.40	

Table 1 *(Continued)*

(c) Enriched in yeast duplicates										
Domain	**Name**	**Function**	**Y1**	**Y2**	***z***	**Significance**	**Arabidopsis**	**Significance**	**Rice**	**Significance**
PF00439	Bromodomain	Chromatin-associated: protein–protein interaction (?)	0	8	−3.72	***	−1.32		3.04	
PF00085	Thioredoxin	Protein folding	0	8	−3.72	***	−0.94		−0.47	
PF00012	HSP70	Protein–substrate binding	0	8	−3.72	***	−0.78		−1.14	
PF00071	Ras family	Regulating hematopoietic cells; signal transduction	0	8	−3.72	***	−3.2		−2.35	
PF00134	Cyclin N-terminal	Regulating cell division	2	10	−3.37	***	−1.69		−1.61	
(d) Enriched in *Tetraodon* duplicates										
Domain	**Name**	**Function**	**T1**	**T2**	***z***	**Significance**	**Yeast[e]**	***Arabidopsis*[e]**	**Rice[e]**	
PF00135	CO esterase	Acts on carboxylic esters	0	6	−4.98	***				
PF00324	AA permease	Transmembrane amino acid transport	0	6	−4.98	***				
PF00071	Ras family	Regulating hematopoietic cells; signal transduction	2	8	−4.83	***	−3.72	−3.20		
PF00063[c]	Myosin head	Motor activity	0	4	−4.07	***				
PF00104	Ligand binding	Transcription regulation	0	4	−4.07	***				
PF00209	Na^+ neurotransmitter	Neurotransmitter (re)cycling at synapses	0	4	−4.07	***				
PF00474	Na^+ symporter	Solute transport	0	4	−4.07	***				
PF00754	F5/8 type C	Cell adhesion, coagulation	0	4	−4.07	***				
PF00755	Carnitine o-acyltransferase	Acyltransferase	0	4	−4.07	***				
PF01413	C–terminal repeat	Collagen structural comp.	0	4	−4.07	***				
PF03137	OATP	Transmembrane transport	0	4	−4.07	***				
PF00002[d]	7tm-2	Secretin receptor (animal-specific)	2	6	−3.96	***				
PF00134	Cyclin N-terminal	Regulating cell division	2	6	−3.96	***	−3.37			
PF00520	Ion transporter	Transmembrane ion channels	2	5	−3.47	***				
PF00069	Protein kinase	Catalysis and ATP binding	14	12	−3.43	***	−8.84	−8.99	−7.98	
PF00022	Actin	Cytoskeleton	1	4	−3.42	***				
PF00266	Aminotransferase-5	Transaminase	1	4	−3.42	***				
PF00307	CH	Cytoskeleton, signal transduction	1	4	−3.42	***				
PF00992	Tn	Muscle contraction	1	4	−3.42	***				
PF03009	GDPD	Phosphodiesterase	1	4	−3.42	***				

[a]Abbreviations: A, *Arabidopsis*; R, rice (*O. sativa*); Y, yeast (*S. cerevisiae*); T, *Tetraodon*; A1, R1, Y1 and T1, singletons in each species; A2, R2, Y2 and T2, duplicated genes in each species. Triple asterisks indicate $p < 0.001$. n.a., not available.
[b]Plant-specific.
[c]All four genes containing Myosin head domains also contain PF00784, MyTH4 domains.
[d]All six of the duplicated genes containing secretin domains also contain PF02793, Hormone/ligand binding domains.
[e]Only the subset that are statistically significant are shown.

were closely correlated ($r = 0.59$, $p < 0.001$, $n = 1006$) in *Arabidopsis* and *Oryza*, with lesser but still highly significant correlations of the angiosperms to yeast ($r = 0.30$, $p < 0.001$ and $r = 0.39$, $p < 0.001$, respectively) and to *Tetraodon* (each $r = 0.29$, $p < 0.001$). By far the weakest relationship was between gene preservation rates for yeast and *Tetraodon* ($r = 0.17$), although the large number of data points still left this statistically significant.

In angiosperms we find non-random patterns of both duplicate gene retention and duplicate gene loss. Although

retention frequencies for the vast majority of duplicated genes were random, each of the two tails of the distribution was larger than would be expected by chance. For *Oryza*, 12 domains showed highly significant ($z > 3.3$, $p < 0.001$) enrichment in singleton genes and two showed enrichment in duplicates (Table 1). For *Arabidopsis*, 16 domains showed highly significant enrichment in singletons and four in duplicates. Only about one case of such enrichment in each taxon would be expected to occur by chance at $p < 0.001$.

With only one exception (PF05498), domains showing non-random distribution between singleton and duplicated genes in one angiosperm showed parallel distribution in association with independent duplication in the other angiosperm. A total of five and two domains showed highly significant enrichment in singleton and duplicated genes, respectively, in both *Oryza* and *Arabidopsis*, a correspondence exceedingly unlikely to occur by chance. Only three domains showed differential enrichment in singleton versus duplicated genes (see Table 2 in the supplementary material online).

Duplication-resistant genes

Although classical views highlight the potential long-term benefits of polyploidy as a source of duplicated genes that are free to acquire new functions [1,2], patterns of gene loss are as distinctive as those of gene retention in *Arabidopsis* and *Oryza*. That such patterns are closely related ($r = 0.59$) in independent duplications of about the same age in these taxa (thought to have diverged from a common ancestor 140–180 MYA) and are also related to those found in yeast ($r = 0.30$–0.39 and the animal *Tetraodon* ($r = 0.29$), suggests that gene loss could be part of an adaptive program. In particular, these findings identify domain-containing genes that might be to some degree maladaptive when duplicated, and for which restoration to singleton status might enhance the long-term fitness of a new polyploid lineage. The loss of duplicated copies of these genes in *Arabidopsis* seems to involve multiple spatially independent events distributed across the genome (see the supplementary material online).

Their relatively frequent polyploidization and high degree of duplicate gene preservation makes angiosperms well-suited to further study of duplication-resistant genes. Against the high background level of retained duplicates, angiosperm domains that occurred in only four singleton genes but no duplicates showed statistically significant bias ($p < 0.001$), whereas PF02519, which occurred in no singletons but in 26 *Arabidopsis* duplicates narrowly misses significance. By contrast, 80% of *Tetraodon* domains were in singletons (see Table 1 in the supplementary material online). Although duplication-resistant domains might have been found if we could have studied *Tetraodon* sooner after duplication, we could detect only one remaining duplication-resistant domain – PF0400, with 78 occurrences in singletons and three in duplicates – and this gene was also duplication-resistant in rice. In yeast, two-thirds of domains were in singletons (see Table 1 in the supplementary material online), and no significantly singleton-enriched domains were detected. However, against the generally low levels of duplicate gene retention, 22 and nine domains were enriched in duplicated *Tetraodon* and yeast genes, respectively (Table 1), compared with only four in plants. The most extreme duplicate-enriched domain was the same in *Tetraodon*, yeast and both angiosperms (PF0069, protein kinase).

Many domains that occur too few times to reach statistical significance might also be under selection for singleton versus duplicated status. Pooling of the *Arabidopsis* and *Oryza* data, which is statistically legitimate in that the duplication events are independent and show little nonlinear interaction, showed the distribution patterns of 17 additional singleton-enriched and six additional duplicate-enriched domains to be statistically significant ($p < 0.001$; see Table 3 in the supplementary material online).

Lineage-specific domain or gene family retention contributes to biological diversity

If GD is a primary source of new genes and functions, then 'deletion-resistant' genes, those most frequently retained in duplicate, should be members of large heterogeneous families with wide spectra of effects. This was generally true in angiosperms and yeast (Table 1). Indeed, the most duplicate-enriched domain in plants (PF0069, protein kinase) is also the most abundant domain in the rice genome. However, in *Tetraodon*, smaller domain families were mostly preferentially retained in duplicate. Moreover, despite the overall correlation in retention of domain-containing genes, some domain families showed heterogeneous retention patterns. For example, Myb-like domains (PF00249) are dramatically expanded and largely duplicated in the angiosperms but essentially randomly distributed between yeast singletons (six) and duplicates (four). *Tetraodon* has many domain types preferentially retained in duplicate, some related to animal-specific functions (motor and muscle function or the secretin receptor), that show random distribution in plants. Analysis of additional duplications, as well as comparison of divergent taxa affected by common duplications, could reveal further lineage-specific trends that might have contributed to biological diversity, such as the rapid growth and diversification of plant-specific AP2 gene families (Table 1).

New resolution of gene retention patterns

Gene classification based on shared protein functional domains resolves patterns that eluded detection based on broader gene ontology (GO)-based classification. Our conclusions mainly support *Arabidopsis*-specific studies [22,23] that suggest preferential retention of duplicated genes involved in signal transduction and transcription and loss of DNA repair genes. However, our analysis of specific protein domains reveals heterogeneity that is masked by limiting consideration to broad GO categories. For example among protein–protein interaction domains, an abundant one (LRR) that is almost invariably retained in duplicate masks less-abundant ones (SET, TPR; Table 1) that are duplication-resistant.

Although the presence of a known functional domain allows us to compare genes across divergent taxa for which orthology cannot yet be established, the lack of domains

leaves many genes unexplored (especially singletons). Additional genome sequences will provide greater power to join positional and sequence information to directly resolve orthologies in successive small steps across large evolutionary distances.

Domain combinations are diagnostic of deletion resistance

No pair of domains co-occurred in >2 angiosperm singleton genes, but many co-occurred in duplicated genes. Protein kinase domains (PF00069) co-occur with LRR domains (PF00560) in 53 *Arabidopsis* α duplicated genes and 43 rice duplicates; and with EF hand calcium signaling domains (PF00036) in 18 *Arabidopsis* and 14 rice duplicates. The Ras signaling domain (PF00071) and Arf GTP-binding domain (PF00025) co-occur 24 times in *Arabidopsis* and 10 in rice. In Tetraodon retained duplicates, PF00063 (myosin head) and PF00784 (MyTH4) domains always co-occurred, and PF00002 (7tm-2) and PF02793 (hormone or ligand binding) domains always co-occurred (Table 1). The exons of such domain-rich genes have recently been shown to evolve very conservatively [24]. Further, many such domain-rich genes, and deletion-resistant domain types in general, are involved in molecular complexes that tend to be dosage-sensitive [25], perhaps contributing further impetus to the preferential retention of both members of duplicated gene pairs.

Applications and implications

Angiosperms offer numerous advantages for gaining insight into the consequences of GD. Such insight is not only of evolutionary interest but could also contribute to better understanding of the consequences of pathophysiological ploidy, such as cancer-specific aneuploidies [26] and trisomy-associated diseases [27] as well as physiological ploidy in liver hepatocytes [28], placental trophoblast giant cells [29] and the platelet precursors, megakaryocytes [30,31].

In agriculture, 'synthetic' (man-made) polyploids provide potentially valuable diversity to the narrow gene pools of many crops [32]. Identification of genes for which restoration to singleton status improves fitness could open new doors into crop improvement. The 101 and 123 occurrences of protein domains enriched in *Arabidopsis* and *Oryza* singleton genes, respectively (Table 1), are readily traced back to their source genes, which in turn can be used as probes to investigate the degree to which our findings apply to additional taxa. For example, might orthologs of G-patch-containing proteins be confined to singletons in more recently formed polyploids, such as cotton, wheat, or canola, as they are in both *Arabidopsis* and *Oryza*? The answers to this question might depend on the antiquity of the duplication event being studied; given that *Arabidopsis* and *Oryza* are ancient polyploids, the analysis of additional genomes is needed to clarify the timetable of duplication-resistant gene loss.

Our findings also suggest a means to further investigate rapid responses to polyploidization such as epigenetic silencing and genomic restructuring. There presently exists little data to distinguish whether such rapid responses provide raw material for the beginnings of adaptation to genome duplication, or are symptomatic of imminent extinction. The extinction hypothesis seems more likely given that such a tiny fraction of duplication events result in successful lineages. If such rapid reactions were to be somehow directed, preferentially affecting gene or domain families that are to be eventually retained predominantly as singletons, then these mechanisms might be implicated as components of an adaptive program to resolve GD-associated genetic imbalances.

Acknowledgements

Aspects of this work were funded by the US National Science Foundation and USDA-CSREES National Research Initiative.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.tig.2006.09.003.

References

1 Ohno, S. (1970) *Evolution by gene duplication.* Springer
2 Stephens, S. (1951) Possible significance of duplications in evolution. *Adv. Genet.* 4, 247–265
3 Song, K. *et al.* (1995) Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proc. Natl. Acad. Sci. U. S. A.* 92, 7719–7723
4 Ozkan, H. *et al.* (2001) Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops-Triticum*) group. *Plant Cell* 13, 1735–1747
5 Shaked, H. *et al.* (2001) Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* 13, 1749–1759
6 Kashkush, K. *et al.* (2002) Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics* 160, 1651–1659
7 Kashkush, K. *et al.* (2003) Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat. Genet.* 33, 102–106
8 O'Neill, R.J.W. *et al.* (2002) Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature* 420, 106
9 Chen, Z.J. and Pikaard, C.S. (1997) Transcriptional analysis of nucleolar dominance in polyploid plants: biased expression/silencing of progenitor rRNA genes is developmentally regulated in *Brassica*. *Proc. Natl. Acad. Sci. U. S. A.* 94, 3442–3447
10 Comai, L. *et al.* (2000) Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. *Plant Cell* 12, 1551–1567
11 Lee, H.S. and Chen, Z.J. (2001) Protein-coding genes are epigenetically regulated in *Arabidopsis* polyploids. *Proc. Natl. Acad. Sci. U. S. A.* 98, 6753–6758
12 Bailey, J.A. *et al.* (2002) Recent segmental duplications in the human genome. *Science* 297, 1003–1007
13 Cheung, J. *et al.* (2003) Recent segmental and gene duplications in the mouse genome. *Genome Biol.* 4, R47
14 Tuzun, E. *et al.* (2004) Recent segmental duplications in the working draft assembly of the brown Norway rat. *Genome Res.* 14, 493–506
15 Bowers, J.E. *et al.* (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422, 433–438
16 Paterson, A.H. *et al.* (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl. Acad. Sci. U. S. A.* 101, 9903–9908
17 Sanderson, M.J. *et al.* (2004) Molecular evidence on plant divergence times. *Am. J. Bot.* 91, 1656–1665
18 Bell, C.D. *et al.* (2005) The age of the angiosperms: a molecular timescale without a clock. *Evolution Int. J. Org. Evolution* 59, 1245–1258
19 Lynch, M. and Conery, J.S. (2003) The origins of genome complexity. *Science* 302, 1401–1404

20 Wolfe, K.H. and Shields, D.C. (1997) Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* 387, 708–713
21 Jaillon, O. *et al.* (2004) Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431, 946–957
22 Maere, S. *et al.* (2005) Modeling gene and genome duplications in eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 102, 5454–5459
23 Blanc, G. and Wolfe, K.H. (2004) Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* 16, 1679–1691
24 Chapman, B.A. *et al.* (2006) Buffering crucial functions by paleologous duplicated genes may impart cyclicality to angiosperm genome duplication. *Proc. Natl. Acad. Sci. U. S. A.* 103, 2730–2735
25 Birchler, J.A. *et al.* (2005) Dosage balance in gene regulation: biological implications. *Trends Genet.* 21, 219–226
26 Duesberg, P. *et al.* (2000) Explaining the high mutation rates of cancer cells to drug and multidrug resistance by chromosome reassortments that are catalyzed by aneuploidy. *Proc. Natl. Acad. Sci. U. S. A.* 97, 14295–14300
27 Roubertoux, P.L. and Kerdelhue, B. (2006) Trisomy 21: from chromosomes to mental retardation. *Behav. Genet.* 36, 346–354
28 Guidotti, J.E. *et al.* (2003) Liver cell polyploidization: a pivotal role for binuclear hepatocytes. *J. Biol. Chem.* 278, 19095–19101
29 Zybina, T.G. and Zybina, E.V. (2005) Cell reproduction and genome multiplication in the proliferative and invasive trophoblast cell populations of mammalian placenta. *Cell Biol. Int.* 29, 1071–1083
30 Ravid, K. *et al.* (2002) Roads to polyploidy: the megakaryocyte example. *J. Cell. Physiol.* 190, 7–20
31 Olsen, L.E. *et al.* (2004) A chromosome 21 critical region does not cause specific Downs syndrome phenotypes. *Science* 306, 687–690
32 Zhang, P.Z. *et al.* (2005) Quantifying novel sequence variation and selective advantage in synthetic hexaploid wheats and their backcross-derived lines using SSR markers. *Mol. Breed.* 15, 1–10

0168-9525/$ – see front matter
doi:10.1016/j.tig.2006.09.003