

**Table 1: Distribution of (TG/CA)<sub>n</sub> repeats in large gene families**

Functional class and gene families	Chromosomal Distribution <sup>a</sup>	Average Gene Length (kb) <sup>b</sup>	Genes in the family	Proportion of Genes with (TG/CA) <sub>n</sub> repeats	Number of Genes with (TG/CA) <sub>n</sub> repeats in three categories			Average number of repeats per gene in three categories		
					Type I	Type II	Type III	Type I	Type II	Type III
<b>Cell cycle class<sup>c</sup>(1)</b>										
Histone proteins family	Dispersed	3.55	76	7.9	4	4	0	1.7	0.7	0
<b>Immune and related functions class (9)</b>										
Interleukins	Dispersed	14.67	43	32.6	9	6	2	1.3	0.9	0.1
Serine (or cysteine) proteinase inhibitor family	Dispersed	17.87	32	50	12	9	2	1.4	0.8	0.2
Tumor necrosis factor (ligand) superfamily	Dispersed	23.49	38	63.2	21	14	1	1.9	0.9	0.1
CD antigens	Dispersed	26.96	54	46.3	20	11	3	1.9	1.3	0.3
Immunoglobulin heavy chains	Intrachromosomal	0.38	162	0.6	1	0	0	1	0	0
Immunoglobulin kappa chains	Intrachromosomal	0.55	73	5.5	1	3	0	0.3	3	0
Immunoglobulin lambda chains	Intrachromosomal	0.35	88	0	0	0	0	0	0	0
Interleukin receptors family	Dispersed	29.77	32	59.4	17	10	0	3.4	0.9	0
T cell receptor beta chains	84 Intrachromosomal, 9 Dispersed	0.42	94	9.6	5	3	1	0.8	0.3	0.1
<b>Information class (6)</b>										
Homeo box	Dispersed	5.48	40	25	6	6	1	1.2	0.9	0.2
Eukaryotic translation initiation factor	Dispersed	36.54	33	45.5	13	10	3	2.1	0.9	0.2
Zinc finger protein family	Dispersed	30.33	200	42.5	63	44	6	2.2	1.1	0.1
DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptides	Dispersed	43.44	32	62.5	17	8	2	1.5	0.6	0.1
Ribosomal protein genes	Dispersed	4.94	96	6.3	6	0	0	1	0	0
Mitochondrial ribosomal protein genes	Dispersed	16.92	74	23	14	12	1	1.8	0.9	0.1
<b>Metabolism class (2)</b>										
Cytochrome P450 superfamily	Dispersed	31.34	45	46.7	15	11	2	1.9	1	0.1
Proteasome subunit genes	Dispersed	25.64	40	32.5	11	8	0	1.5	0.8	0
<b>Signaling and Communication class (8)</b>										
G protein-coupled receptor family	Dispersed	24.71	98	33.7	26	19	6	2.3	1.5	0.2
Tripartite motif-containing family	Dispersed	29.29	40	60	19	12	2	1.6	0.8	0.1
Solute carrier family	Dispersed	59.19	223	62.8	134	87	22	2.9	1.5	0.2
RAS oncogene family	Dispersed	39.92	60	65	38	17	4	1.8	0.7	0.1
ATP-binding cassette transporters gene family	Dispersed	73.85	44	68.2	29	24	4	3.6	2.5	0.3
Guanine nucleotide binding protein (G protein) polypeptide genes	Dispersed	58.83	32	59.4	18	12	5	3.3	1.9	0.4
Potassium voltage-gated channel genes	Dispersed	104.95	38	57.9	17	16	6	8.6	4.4	0.5
Protein phosphatase subunit genes	Dispersed	65.62	59	57.6	27	22	7	2.9	1.8	0.3
<b>Structure and motility class (1)</b>										
Collagen family	Dispersed	132.83	37	86.5	29	23	10	5.7	2.3	0.4

<sup>a</sup>: Chromosomal distribution of the members of gene families. 'Dispersed' indicates that members are distributed on different chromosomes.

<sup>b</sup>: average gene length (in kb) for each gene family.

<sup>c</sup>:Numbers in parentheses show the number of large sized gene families in each functional class.