

TABLE 9.4 STRUCTURAL VARIATION IN SIZE AND ORGANIZATION OF HUMAN PROTEIN-CODING GENES

Human protein	Size of protein (no. of amino acids)	Size of gene (kb)	No. of exons	Coding DNA (%)	Average size of exon (bp)	Average size of intron (bp)
SRY	204	0.9	1	94	850	–
β-Globin	146	1.6	3	38	150	490
p16	156	7.4	3	17	406	3064
Serum albumin	609	18	14	12	137	1100
Type VII collagen	2928	31	118	29	77	190
p53	393	39	10	6.0	236	3076
Complement C3	1641	41	29	8.6	122	900
Apolipoprotein B	4563	45	29	31	487	1103
Phenylalanine hydroxylase	452	90	26	3	96	3500
Factor VIII	2351	186	26	3	375	7100
Huntingtin	3144	189	67	8.0	201	2361
RB1 retinoblastoma protein	928	198	27	2.4	179	6668
CFTR (cystic fibrosis transmembrane receptor)	1480	250	27	2.4	227	9100
Titin	34,350	283	363	40	315	466
Utrophin	3433	567	74	2.2	168	7464
Dystrophin	3685	2400	79	0.6	180	30,770

Where isoforms are evident, the given figures represent the largest isoforms. As genes get larger, exon size remains fairly constant but intron sizes can become very large. Internal exons tend to be fairly uniform in size, but the terminal exon or some exons near the 3' end can be many kilobases long; for example, exon 26 of the *APOB* gene is 7.5 kb long. Note the extraordinarily high exon content and comparatively small intron sizes in the genes encoding type VII collagen and titin. In addition to *SRY*, other single-exon protein-coding genes in the nuclear genome include retrogenes (see Table 9.8) and genes encoding other SOX proteins, interferons, histones, many G-protein-coupled receptors, heat shock proteins, many ribonucleases, and various neurotransmitter receptors and hormone receptors.