XI. Appendix D

Mutations in human pathology

The following is a listing of the types of mutations that have been linked to human disease. An example for each kind of mutation and its associated disease is given in Table XI-1.

A. Localised mutations

Localised mutations are pathologic changes affecting single bases or short stretches of genomic sequence.

A.1. Base substitutions

Among the localised mutations, base substitutions are the most confined: they affect a single nucleotide.

A.1.1. Silent mutations

Pathogenicity through affecting splicing or regulatory elements.

When silent mutations occur in coding DNA, they result in a synonymous substitution resulting in no net change of the gene product. Although one could argue that certain exchanges would confer susceptibility to a particular condition, most do not have an observable effect. However, through aberrant splicing – for example, by activating a cryptic splice site[1381] (see XI.A.4.2) or by affecting regulatory sequences such as exonic splicing enhancers[721] – some of them are, indeed, pathogenic[1382].

A.1.2. Missense mutations

Non-synonymous amino acid exchange. Pathogenicity by alteration of the gene product's characteristics.

Missense mutations are base substitutions leading to a non-synonymous amino acid exchange. This group of mutations can be subdivided into conservative and non-conservative substitutions, depending on the degree of similarity between the amino acids that have been exchanged. Non-conservative substitutions have been implied in a large number of human pathologies[1383]. Although less frequently involved in pathogenesis, conservative substitutions have also been implicated in human disease[204].

A.1.3. Nonsense mutations

Non-synonymous exchange leading to a termination codon. Pathogenicity through null allelism or dominant negative effects.

Nonsense mutations are a special group of non-synonymous alterations replacing a codon specifying an amino acid with a termination codon. They either trigger NMD[1384], resulting in the degradation of the faulty message and hence representing null alleles, or they lead to truncated proteins, which may represent null alleles or exert a dominant negative effect. A plethora of different examples of pathogenic nonsense mutations is described in the literature[1383].

A.2. Insertions and deletions

Addition or removal of one or more bases. Pathogenicity by altering the gene product's characteristics (in-frame) or by frameshift mostly leading to a premature termination codon and subsequent null allelism or dominant negative effects (out of frame).

In-frame insertions[1385] or deletions[209] mostly alter gene function by upsetting the protein's conformation through the addition or removal of one or a few amino acids. However, often insertions and deletions lead to frameshift mutations, causing a shift in the translational reading frame. Such mutations change the amino acid sequence following the frameshift and often incorporate a stop codon, leading to premature termination of translation[58]. The outcome is similar to that discussed under nonsense mutations (see previous section). Special examples of insertions/deletions are expansions of triplet repeats (see next section), DNA transposition (see XI.B.1) and unequal cross-over (see XI.B.3).

A.3. Expansions of triplet repeats

In-frame insertions. Pathogenicity through unstable repeat sizes when expanded beyond a critical threshold.

A particular group of in-frame insertions are the expansions of intragenic triplet repeats. Such expansions have been shown to cause a variety of diseases and are found in the coding sequence, as well as in the UTRs. The coding repeats are moderately expanded $(CAG)_n$ triplets seldomly exceeding 100 repetitions[1386]. The resulting poly-glutamine tracts are believed to cause protein oligomerisation[1387]. The non-coding repeats, in contrast, are

Table XI-1 | **Mutational mechanisms underlying human pathology**

Type of mutation	Mutation	Gene/Locus/Genotype	Effect	Pathology (OMIM)	Ref.
Local mutations					
Single-base substitution	Silent c.852T>C	*MAPT*	p.L284L, abolition of an exon splicing silencer	FTDP-17 (600274)	1382
	Conservative missense c.419C>T	*MECP2*	p.A140V, alteration of the wedge-shaped structure of the MBD	NS-XLMR	204
	Non-conservative missense c.410A>G	*MECP2*	p.E137G, located in the MBD	NS-XLMR	206
	Nonsense c.352G>T	*ZNF674*	p.E118X, truncated protein	NS-XLMR	73
	Nonsense c.3645T>G	*FBN1*	p.Y1215X, exon skipping, restoration of the ORF	Marfan syndrome (154700)	1388
In-frame deletion	c.1161_1400del	*MECP2*	Loss of 80 C-terminal AA	NS-XLMR	209
In-frame insertion	c.3670_3678dupATCCAATCC	*CACNA1A*	p.1133_1135dupPNS, decrease in current density, shift of voltage threshold of activation, slower activation	Late-onset EA2 (108500)	1385
Insertion	c.3898_3899dupAG	*PQBP1*	Frameshift, truncated protein, aberrant subcellular localisation	SHS S-XLMR	58

Table XI-1 | **Mutational mechanisms underlying human pathology**

Type of mutation	Mutation	Gene/Locus/Genotype	Effect	Pathology (OMIM)	Ref.
Deletion	c.3898_3899delAG	*PQBP1*	Frameshift, truncated protein, aberrant subcellular localisation	S-XLMR	58
Triplet repeat expansion	Coding c.366(CAG)$_{37\text{-}120}$	*HD*	Poly-Q tract in the Huntingtin protein, formation of inclusion bodies	HD (143100)	975
	Non-coding c.*224(CTG)$_{50\text{-}4000}$	*DMPK*	CUG repeats in the 3'UTR	Myotonic dystrophy (160900)	1389
Defective splice acceptor site	g.IVS3-2A>G	*FTSJ1*	Skipping of the downstream exon, frameshift, truncated protein	NS-XLMR	1390
	c.1003-2A>G	*ACSL4*	Use of an intronic cryptic splice site, retention of 28 bp intronic sequence, frameshift, truncated protein	NS-XLMR	93
Defective splice donor site	c.2642A>G	*ATP7A*	Skipping of the upstream exon, use of an exonic cryptic splice site, 220 bp exonic deletion, NMD	OHS (304150)	1391
	g.IVS20+1G>A	*OPA1*	Retention of 25 bp intronic sequence, frameshift, truncated protein	ADOA (165500)	1392
	g.IVS66+2T>C g.IVS66+5G>T	*Dystrophin*	Skipping of the upstream exon, frameshift, truncated protein	DMD with severe MR (310200)	1393
Activation of a cryptic splice site	g.IVS7+26G>A	*E1α PDH*	Generation of an SC35 binding motif, activation of a downstream cryptic splice site, in-frame retention of 45 bp of intronic sequence, unstable protein	Lactic acidosis & MR (300502)	1394
Defective branch site	g.IVSP-19A>C	*L1CAM*	In-frame 69 bp insertion, 23 additional AA, skipping of the downstream exon, frameshift, truncated protein	HSAS (307000)	349

Table XI-1 | **Mutational mechanisms underlying human pathology**

Type of mutation	Mutation	Gene/Locus/Genotype	Effect	Pathology (OMIM)	Ref.
5'UTR	c.-17A>G	*F9*	Abolition of a C/EBP TF binding site	Haemophilia B Leyden (306900)	1395
3'UTR	AATAAA → AATAAG	*HBA2*	Abolition of the canonical AAUAAA poly(A) signal, read-through transcripts, reduced accumulation of *HBA2* mRNA	α-thalassemia (141850)	1396
Imprinting	Maternal heterodisomy	15q11 – q13	Absence of paternal contribution to 15q11-q13	PWS (176270)	1397
Abberant methylation at promoter	Methylated promoter	*CDX1*	Absence or reduction of *CDX1* mRNA expression	CRC (600746)	1398
Mitochondrial	m.11778G>A	*ND4*	p.R340H, reduced functionality	LHON (535000)	1399
Global mutations					
DNA transposition	Intronic *Alu* insertion	*NF1*	Skipping of the downstream exon, frameshift, truncated protein	NF1 (162200)	1400
	Exonic *LINE-1* insertion	*F8*	Insertion of novel amino acids and frameshift, truncated protein	Haemophilia A (306700)	1401
Altered chromosomal environment	t(4;11)(q22;p13) inv(11)(p13p13) BPs ≥ 85 kb distal of the 3' end of *PAX6*	*PAX6*	Inappropriate chromatin environment for normal *PAX6* expression on the rearranged chromosome 11p13	Aniridia (106200)	1402
Contiguous gene syndrome	~2 Mb interstitial deletion	*DAX1* and *IL1RAPL1* on Xp21.2-p21.3	Deletion of *DAX1* and the 3' end of *IL1RAPL1*	Adrenal hypoplasia & MR (300200)	1403

Table XI-1 | **Mutational mechanisms underlying human pathology**

Type of mutation	Mutation	Gene/Locus/Genotype	Effect	Pathology (OMIM)	Ref.
	1.55 Mb deletion	several genes on 7q11.23	Upsetting relative gene dosage	WBS (194050)	1404
Unequal cross-over	~1.5 Mb reciprocal duplication	*PMP-22* on 17p11.2	Upsetting relative gene dosage	CMT1A (118220)	1405
	Fusion of the 5' end of *OPN1LW* to the 3' end of *OPN1MW* and *vice versa*	*OPN1LW* and *OPN1MW*	Formation of fusion genes	Red & green colour blindness (303800, 303900)	1406
	Fusion of the 5' end of a functional gene with the 3' end of its pseudogene	*GBA*	Introduction of c.1448T>C, p.L444P, c.1483G>C, p.A456P, c.1497G>C and p.V460V from the pseudogene	Gaucher disease (230800, 230900, 231000)	1407
Dosage effect	c.265C>T	*IGF1R*	p.R89X[§], truncated protein, haploinsufficiency	Primary microcephaly, mild MR, intra-uterine and post-natal growth retardation (147370)	1408
	0.4 – 0.8 Mb duplication	Several genes on Xq28, including *MECP2*	Increased MeCP2 dosage	Severe MR	1409
	c.1385G>A	*RNASEL*	p.R462Q, ~3-fold reduced *in vitro* RNase activity	HNPCC1 (120435)	1410
Inversion	Unequal cross-over between an intronic and a telomeric copy of sequence A	*F8*	Inversion of exons 1 – 22, loss-of-function	Haemophilia A (306700)	680
	inv(11)(p13q13) BP > 75 kb distal of the 3' end of *PAX6*	*PAX6*	Position effect disrupting regulatory elements, loss-of-function, haploinsufficiency	Aniridia (106200)	681

Table XI-1 | **Mutational mechanisms underlying human pathology**

Type of mutation	Mutation	Gene/Locus/Genotype	Effect	Pathology (OMIM)	Ref.
Reciprocal translocation	t(X;7)(p22.3;p15) t(X;6)(p22.3;q14)	*CDKL5*	Disruption of *CDKL5*, absence of mRNA	ISSX (308350)	474
	t(4;11)(q21;q23) t(9;11)(p22;q23)	*ALL-1*	Gene fusions between *ALL-1* and *AF-4* [t(4;11)] or *AF-9* [t(9;11)], gain-of-function of the chimeric proteins	Acute leukemia (159555)	688
Uniparental disomy	Maternal disomy 14	46,XX,UPD14$_{mat}$	Reduced heterozygosity and possibly imprinted gene effects	Intra-uterine growth retardation, early onset of puberty, short stature, small hands (NA)	1411
	Paternal disomy 14	46,XX,UPD14$_{pat}$	Reduced heterozygosity and possibly imprinted gene effects	Abdominal muscular defects, skeletal anomalies, characteristic facies (608149)	1412
Autosomal trisomy	Additional chromosome 13	47,XX,+13 or 47,XY,+13	Likely upsetting relative gene dosage	Patau syndrome (NA)	1413
	Additional chromosome 18	47,XX,+18 or 47,XY,+18	Likely upsetting relative gene dosage	Edwards syndrome (NA)	1414
	Additional chromosome 21	47,XX,+21 or 47,XY,+21	Likely upsetting relative gene dosage	Down syndrome (190685)	686
Sex chromosome aneuploidy	Supernumery X and/or Y chromosomes	47,XXX 47,XXY 47,XYY	Mostly mild to moderate developmental problems; 47,XXY boys are infertile	Klinefelter syndrome [47,XXY] (NA)	1415
	Monosomy X in girls	45,X	A variety of abnormal physical features, short stature, infertility	Turner syndrome (NA)	1416

Table XI-1 | **Mutational mechanisms underlying human pathology**

Type of mutation	Mutation	Gene/Locus/Genotype	Effect	Pathology (OMIM)	Ref.
Polyploidy	Triploidy	69,XXX	Mostly spontaneous abortion Premature birth, life expectancy usually < 24 hrs.	General dysmaturity, a plethora of pathological features (NA)	1417
	Tetraploidy	92,XXXX	Mostly spontaneous abortion Life expectancy days to months	Facial dysmorphism, severe growth delay, developmental delay (NA)	1418

§ Although it is obvious from their supplementary Fig. 2 that amino acid 89 is mutated, Abuzzahab *et al.*[1419], the authors who identified this mutation, erroneously refer to it as p.R59X. Unfortunately other authors[1408] as well as OMIM have adopted this mistake.

massively expanded triplets, with up to several thousand repetitions, thought to interfere with transcription or translation[1420]. Both types of repeats are stable in mitosis and meiosis below a certain threshold length, but become unstable once that critical length is surpassed. Through expansion, these so-called pre-mutations eventually develop into full mutations when transmitted from one generation to the next. With increasing repeat size, the condition becomes more severe in successive generations[1421], a genetic phenomenon called anticipation[1422]. The average size change depends on the length of the repeat[1423] and on the gender of the transmitting parent[1424]. It is thought that instability of triplet repeats involves DNA polymerase slippage[1425], slipped strand mispairing[1426], or unequal sister chromatid exchange[1427] (see XI.B.3), but the exact mechanisms underlying the inheritance of expanded triplet repeats are unclear.

A.4. Mutations affecting messenger RNA and transcription

Several mutations affect regulatory elements critical to the processing of mRNAs, such as poly-adenylation signals, splice sites or splicing enhancers.

A.4.1. Mutations of the poly-adenylation signal

Abolition of the canonical poly(A) signal. Pathogenicity through read-through transcripts and destabilised mRNAs.

In mammalian cells, the enzyme Poly(A) polymerase adds a poly(A) tail of around 200 residues to the 3' end of most mRNA molecules[1428]. Such a poly(A) tail stabilises at least some of the eukaryotic mRNAs[1429]. Addition of the poly(A) tail critically depends on the canonical AAUAAA poly-adenylation signal[1430]. Mutations affecting this signal have been implied in human disease[1396].

A.4.2. Mutations affecting RNA splicing

(i) Mutation of a splice site. Pathogenicity through use of a cryptic splice site, retention of intronic or deletion of exonic sequence, or exon skipping. (ii) Mutation of a regulatory element. Pathogenicity arising from generation of a cryptic splice site or defective branching.

Several pathogenic mutations are known which affect splicing in a variety of ways. Failure of splicing can occur when a splice site is mutated, resulting in the retention of intronic sequence in the mature mRNA[1392]. Alternatively, the splicing machinery has been found to accidentally use a cryptic splice site, thereby circumventing the faulty splice accep-

tor site, but incorporating intronic sequence when the illegitimate splice site is located within an intron[93], or deleting coding sequence in case of an exonic cryptic site[1391].

Another outcome of a mutated splice acceptor site is skipping of the downstream exon[1390]. When a mutation affects the splice donor site, this results in skipping of the upstream exon[1393]. Some nonsense mutations have also been reported to induce exon skipping[1388].

Sometimes, mutations can cause abnormal RNA splicing by activation of cryptic splice sites: a sequence which normally has no effect on RNA splicing, but which is reminiscent of a splice signal, is changed by the mutation so that it is falsely interpreted by the splicing machinery as a splice site[1394]. As with the use of cryptic splice sites described earlier, the location of the mutated sequence is important in the effect of the mutation. This implies that some silent mutations in coding DNA may not be neutral but instead are pathogenic because they activate an exonic cryptic splice site[1431]. Finally, pathogenic mutations affecting the branch site have also been reported[349].

In all cases, the ultimate effect of mutations upsetting RNA splicing depends on the nature of the resulting insertion or deletion, as has been described under XI.A.2.

A.4.3. Mutations in the untranslated regions

Affecting regulatory elements. Pathogenicity by compromising transcript stability or reducing binding affinity to interacting factors.

Mutations in the UTRs of mRNAs are believed to exert their effect in either of two ways: they compromise the stability of the RNA message, for example by aberrant poly-adenylation (see XI.A.4.1), or they reduce binding affinity to interacting factors, such as those involved in the translational apparatus[1395].

A.4.4. Mutations in regulatory sequences

Affecting sequences involved in regulating gene transcription. Pathogenicity through faulty gene regulation.

All types of mutations described in this overview can affect gene expression by acting on regulatory sequences. Depending on the type of regulatory sequence involved and on the nature of the mutation, such alterations can lead to up- or downregulation of transcription[1383]. In several cases, distant regulatory units have been discovered through investigation of chromosomal BPs.

A.5. Epigenetic mutations

Epigenetic mutations are a class of mutations influencing the phenotype without altering the genotype. Although changes are inherited, they do not represent a change in genetic information.

A.5.1. Imprinting

Upsetting parental origin of imprinted locus. Pathogenicity arises from same parent heterodisomy.

The maternal and paternal genomes that are combined in a diploid individual do not function interchangeably. Their different mode of action is, for example, visible in genomic imprinting, a parental-specific transcriptional silencing mediated by DNA methylation[1432]. XCI, described in detail in Appendix B, is a special form of genomic imprinting in which one of both X chromosomes is inactivated in females[27,28]. Only a limited number of autosomal genes are subject to imprinting[1249]. When the WT situation is upset, this has been shown to lead to disease[1397]. In other words, some human conditions only manifest when inherited from one parent.

A.5.2. Methylation at the promoter

Changing the methylation pattern at the promotor. Pathogenicity due to aberrant transcription.

As in imprinting, DNA methylation at the promoter is a way to transcriptionally regulate genes[1433]. Aberrant transcription due to altered methylation patterns at certain promoters leads to disease[1398].

A.6. Mitochondrial mutations

Mutation in the mitochondrial genome. Pathogenicity because of mitochondrial dysfunction.

The 16.6 kb human mitochondrial genome[1434] is characterised by (i) a matrilineal inheritance[1435], (ii) frequent heteroplasmy[1436] and (iii) a high mutation rate[1437]. Numerous mutations in the mitochondrial genome have been reported to cause human disease[1438]. Because cells typically contain thousands of mtDNA molecules, the severity of the phenotype often varies with the proportion of affected mtDNA molecules. It has also been observed that mutations evolve within individuals and that the number of mtDNA variants increases with age[1439].

B. Global mutations

In contrast to localised mutations, global mutations are those changes affecting several loci at once or involving large portions of the genome.

B.1. DNA transposition

Repetitive units spreading through the genome by retrotransposing. Pathogenicity due to malignant insertion.

The human genome is littered with large numbers of repetitive sequences. The *Alu* repeat is the most abundant sequence in the genome and occurs, on average, once every 4 kb[21]. It is thought that the *Alu* repeat represents a processed 7SL RNA pseudogene[1440]. LINE-1, another type of repetitive sequence, occurs, on average, once every 50 kb in the human genome[21]. Both these repetitive units are known to spread through the genome by retrotransposing[1441]. Although not very frequent, malignant insertions of both *Alu*[1400] and LINE-1[1401] elements have been implicated in pathology.

B.2. Alteration of the chromosomal environment

Position effect. Pathogenicity because of aberrant transcription.

Several human disease phenotypes have been attributed to position effects, i.e. the impairment of transcription because of an unusual chromosomal location. Cases have been reported in which chromosomal rearrangements cause a characteristic phenotype without affecting the causative gene, implying that some crucial regulatory sequences are no longer functional[1402]. This effect could be one of purely physical isolation of the regulatory elements. Alternatively, the regulatory sequences may have been translocated in a transcriptionally silenced chromatin domain.

B.3. Unequal cross-over and unequal sister chromatid exchange

Recombination between non-allelic sequences. Pathogenicity arises through formation of fusion genes, introduction of mutations from a pseudogene in its functional counterpart and dosage effects.

Sometimes recombination occurs between sequences that are non-allelic but highly homologous such as repetitive elements (see XI.A.3) or genes and their pseudogenes. This process is called unequal cross-over when it happens between non-sister chromatids of a pair of homologues. It is termed unequal sister chromatid exchange when taking place be-

tween sister chromatids. These mechanisms lead to an insertion on one chromatid and a corresponding deletion on the other chromatid, and have been shown to exert their effect on human health in a variety of ways. Formation of fusion genes[1406], introduction of deleterious mutations from a pseudogene into its functional counterpart[1407], and dosage effects[1405] have all been observed. Affecting the relative abundance of a gene product, gene dosage effects can also become apparent through loss of function of one copy of a gene (i.e. haploinsufficiency)[1408], gene duplication[1409] or mutations partly abolishing gene function[1410]. Each of these cases has been described in human pathology.

B.4. Contiguous gene syndrome

Patients carrying deletions removing a string of adjacent genes may show combined features of several monogenic disorders. Such a condition is referred to as a contiguous gene syndrome.

B.4.1. X-chromosomal

Removal of several adjacent X-chromosomal genes. Pathogenicity occurs most prominently in males due to gene deletion.

Males carrying X-chromosomal deletions often are characterised by well-defined contiguous gene syndromes, showing superimposed features of several Mendelian X-linked conditions[1403]. The severity of the phenotype is roughly concordant with the extent of the deletion and the number of genes deleted.

B.4.2. Autosomal

Removal of several adjacent autosomal genes. Pathogenicity mainly due to dosage effects.

The situation in autosomal contiguous gene syndromes is less clear-cut than that in X-chromosomal deletions, as only dosage-sensitive genes contribute to the phenotype. There exists a remarkable phenotypic similarity among patients carrying such microdeletions: developmental abnormalities, MR, growth retardation and facial dysmorphism[1404]. It is thought that this is a reflection of the high number of genes active in brain function and during embryonic development.

B.5. Chromosomal aberrations

Disease-associated chromosomal rearrangements are described in detail under I.B.1.3.3.

B.6. Numerical chromosomal abnormalities

A special class of global mutations are changes involving whole chromosomes or even entire genomes. An improper parental origin of chromosomes and deviations from the correct number of chromosomes belong to this category.

B.6.1. Uniparental disomy

Two copies of a chromosome with the same parental origin. Pathogenicity likely due to reduced heterozygosity and possibly imprinted gene effects.

Uniparental disomy is believed to be the result of a survival strategy of non-viable trisomic zygotes: by eliminating one copy of the trisomic chromosome, they restore the normal number of chromosomes. As described under XI.A.5.1, the parental origin of genetic material is an important factor in human disease, so it should not be surprising that uniparental disomies are involved in human pathology[1442]. As would be expected, both forms of human uniparental diploidy (hydatidiform mole, i.e. paternal diploidy and ovarian teratoma, i.e. maternal diploidy) are lethal during early development.

B.6.2. Aneuploidy

Incorrect number of chromosomes. Aneuploidy of the sex chromosomes results in minor problems; autosomal aneuploidy is embryonically or perinatally lethal, with the exception of trisomy 21.

As described under I.B.1.3.3.3, zygotic aneuploidy, an incorrect number of chromosomes, results from fertilisation involving a gamete carrying a Robertsonian translocation. Aneuploidy of the autosomes is almost always lethal: all monosomies are embryonic lethal and only trisomies 13 (Patau syndrome), 18 (Edwards syndrome) and 21 (Down syndrome) may survive to term[1443]. Assuming adequate medical care, trisomy 21 individuals may reach ~50 years[1444]. Aneuploidy of the sex chromosomes sharply contrasts with autosomal aneuploidy: 47,XXX; 47,XXY (Klinefelter syndrome)[1445,1446]; and 47,XYY individuals suffer from minor problems[1415] and although the vast majority of 45,X (Turner syndrome)[1447,1448] fetuses abort spontaneously[1449], surviving individuals are of normal intelligence, showing relatively minor physical characteristics and infertility[1416].

Partial monosomies and trisomies can arise when a gamete carrying a reciprocal translocation is involved in fertilisation (see I.B.1.3.3.3). The associated phenotype entirely depends on the extent and the nature of the aneuploidy.

B.6.3. Polyploidy

Supernumerical set of chromosomes. Almost all polyploid conceptions abort spontaneously or die perinatally.

Polyploidy, a supernumerical set of chromosomes, results (i) when an egg is fertilised by more than one sperm, (ii) due to fertilisation involving at least one abnormal diploid gamete or (iii) from DNA duplication without cell division of the zygote. With an estimated 1 – 2% and ~3% of conceptions, tri- and tetraploidy, respectively, are remarkably frequent in humans[1450,1451]. However, most triploid fetuses[1417] and almost all tetraploid conceptions[1418] abort spontaneously or die perinatally.