

Gerard C.P. Schaafsma and Mauno Vihinen

5.1 Introduction

Bruton agammaglobulinemia tyrosine kinase (*BTK*) variations lead to X-linked agammaglobulinemia (XLA, MIM# 300300), a hereditary primary immunodeficiency [26, 29]. XLA is caused by a block in B cell differentiation resulting in severely decreased numbers of B lymphocytes and an almost complete lack of plasma cells and very low or missing immunoglobulin levels of all isotypes. The patients have increased susceptibility to mainly bacterial infections because of virtually absent humoral immune responses. The frequency of XLA has been estimated to be 1:200,000 live births. The disease is considered to have full penetrance. Female carriers are healthy but display nonrandom X-chromosome inactivation in their B cells. Only a few female patients have been identified.

The *BTK* gene (LRG_128, reference sequence used U78027.1) contains 19 exons (Fig. 5.1) and codes for a protein of 77 kDa. Exon 1 is outside the coding region. *BTK* is expressed in all hematopoietic lineages except for T lymphocytes and plasma cells [23]. *BTK* belongs to the Tec family of related cytoplasmic protein tyrosine kinases (PTKs) formed by *BMX* (*BMX* non-receptor tyrosine kinase), *ITK* (*IL2*-inducible T cell kinase), *TEC* (*tec* protein tyrosine kinase), and *TXK* (*TXK* tyrosine kinase). Except for *TXK*, they have the same domain organization, from the N-terminus pleckstrin homology (PH) domain, Tec homology (TH) domain, Src homology 3 (SH3) domain, SH2 domain, and catalytic tyrosine kinase (TK) domain.

The three-dimensional structure has been determined for the PH domain and the first half of the TH domain [7], the SH3 domain [4], the SH2 domain [6], and the kinase domain [12]. For the full-length *BTK*, there is a low-resolution structure in an extended conformation [13]. *BTK* interacts with several partners [15].

G.C.P. Schaafsma • M. Vihinen (✉)
Department of Experimental Medical Science, Lund University,
Lund SE-22184, Sweden
e-mail: mauno.vihinen@med.lu.se

Variations in *BTK* account for about 80 % of agammaglobulinemia cases. Several other genes can lead to a failure of B cell development and agammaglobulinemia [1]. These genes encode components of the pre-B cell receptor or proteins that are activated by cross-linking of the pre-B cell receptor. Defects in these genes lead to a block in B cell differentiation at the pro-B to pre-B cell transition. Other forms of agammaglobulinemia appear with growth hormone deficiency or as autosomal recessive diseases. Some autosomal recessive agammaglobulinemias have been identified involving pre-B cell receptor (pre-BCR) or BCR component genes μ -heavy chain (*IGHM*), λ 5/14.1 (*IGLL1*, immunoglobulin lambda-like polypeptide 1), $Ig\alpha$ (*CD79A*), and $Ig\beta$ (*CD79B*). Variations in the B cell linker protein (*BLNK*), which is essential for $Ig\mu$ signal transduction, and *PIK3R1* (phosphoinositide-3-kinase, regulatory subunit 1 (alpha)) for phosphoinositide 3-kinase regulator are downstream of BCR.

5.1.1 BTKbase

BTKbase is the first immunodeficiency variation database (IDbase) founded in 1994 [32]. Subsequently more than 130 immunodeficiency variation databases (IDbases) have been released [19]. BTKbase contains public variation entries for 1362 patients from 1198 unrelated families (total number of variants in these unrelated families is 1209) showing 742 unique molecular events.

BTKbase aims at collecting all published variations. Data are either directly submitted or derived from more than 100 publications. The database format has been previously published [31, 34]. The data are presented as individual entries, each carrying a unique patient identification number (PIN) and accession number, systematic names according to the Human Genome Variation Society (HGVS) variation nomenclature, a short verbal description of the variation, submission information (submission and update dates, version numbers, and submitter details), literature citations, and annotation in detail at DNA, RNA, and protein levels. In addition, the most important clinical parameters and laboratory findings are included, provided they are available.

IDbases, including BTKbase, follow a number of standards including the use of HUGO Gene Nomenclature Committee (HGNC) gene names (www.genenames.org), HGVS variation nomenclature [3], and IDRefSeqs (reference sequences for primary immunodeficiency genes and proteins). Currently, IDbases are in the process of changing to Locus Reference Genomic (LRG) reference sequences, which are already available for some 100 immunodeficiency genes (www.lrg-sequence.org). BTKbase follows the recommendations for locus-specific variation databases (LSDBs) [33] and their curation [2].

BTKbase is freely available at <http://structure.bmc.lu.se/idbase/BTKbase/>. The website contains information related to XLA and *BTK*. The bioinformatics pages include several tables for statistics of *BTK* variations. The variation distributions are shown along sequences in illustrative ways. The submission page provides variation checking facilities and electronic submission services. The variation browser allows

visual means for browsing variations along the protein sequence. The reference information for variation publications and related protein structures are included in their own sections.

5.2 Analysis of BTK Variations

XLA arises as a block in B cell development. BTKbase contains information in many entries for the immunological status of patients. These properties have been extensively discussed in a previous publication [28]. The majority of the reported patients have significantly reduced numbers of B cells and Ig levels. A large portion of patients with X-linked diseases have de novo variations.

5.2.1 Variation Statistics

Extensive statistical analyses of variations at the three molecular levels, DNA, RNA, and protein, were performed. Since data per unique families are considered the most representative regarding, e.g., mutational effects and prevalence, the discussion about variation statistics mainly relates to these.

Variations appear throughout the BTK domains as well as in exons and introns (Fig. 5.1, Table 5.1); however, the distribution is not even. Some exons contain more variations than expected. The PH and SH2 domains contain approximately the expected number of variations, whereas there are less than expected in the TH and SH3 domains and more than expected in the kinase domain (Table 5.2). The TH domain has two structural elements [31, 35], an N-terminal BTK motif and a C-terminal proline-rich region which contains two proline-rich regions capable of intra- and intermolecular interactions [4, 17]. The reason for under-representativeness of the TH domain may be that it likely has a partially intrinsically disordered structure in the C-terminal half of the domain, and therefore, variations do not have a major effect. On the other hand, XLA-causing variants do appear in the Zn²⁺-binding BTK motif.

We have recently investigated the putative effects of all possible amino acid substitutions due to single nucleotide changes in the BTK TK domain [27]. Altogether 67 % of the 1495 substitutions were predicted to be harmful. Although this number

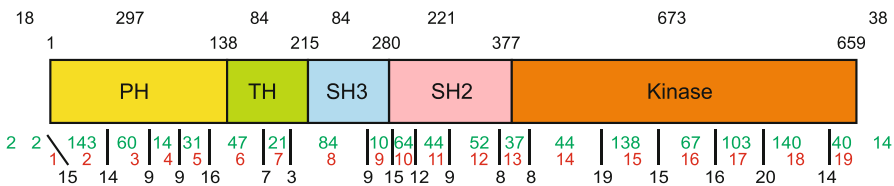


Fig. 5.1 Distribution of all variations to *BTK* gene regions and BTK protein domains. Variations in exons are indicated by green numbers in exons which are numbered in red. Variations in introns are in black below the domain chart. Domain borders are above the chart and numbers of variations in the domains above them

Table 5.1 Distribution of variation and variation types in BTK domains for all cases, independent families, and unique variations

Domain Type	Upstream		PH (414)		TH (231)		SH3 (194)		SH2 (292)		TK (846)		Others		Total		% of total											
	All	Fam	All	Uniq	All	Uniq	All	Uniq	All	Uniq	All	Uniq	All	Uniq	All	Uniq	All	Uniq	Fam	Uniq								
Missense	0	0	0	0	119	100	48	11	10	10	0	0	0	0	328	290	147	0	0	0	570	493	247	41.5	40.8	33.3		
Nonsense	0	0	0	0	44	40	19	16	14	9	46	39	11	23	22	12	97	85	43	0	0	0	226	200	94	16.4	16.5	12.7
Deletion; in-frame	1	1	1	1	11	6	6	2	2	2	0	0	0	5	5	4	21	18	11	0	0	0	40	32	24	2.9	2.6	3.2
Deletion; frameshift	0	0	0	0	3	2	1	0	0	0	0	0	0	1	1	1	2	1	1	4	3	3	10	7	6	0.7	0.6	0.8
Gross deletion	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	14	14	14	14	14	1.0	1.2	1.9
Insertion; in-frame	0	0	0	0	3	3	2	0	0	0	0	0	0	0	3	3	3	0	0	0	6	6	6	5	0.4	0.5	0.7	
Insertion; frameshift	0	0	0	0	3	2	1	0	0	0	1	1	0	0	0	0	0	0	0	0	4	3	1	0.3	0.2	0.1		
Indel	0	0	0	0	3	3	3	3	3	3	2	2	2	1	1	1	6	6	5	0	0	0	15	15	14	1.1	1.2	1.9
Intron	11	10	6	27	23	14	9	8	5	6	5	6	5	3	29	28	20	65	59	35	6	6	153	139	89	11.1	11.5	12.0
Intron; in-frame	0	0	0	0	3	3	2	0	0	0	1	1	1	1	6	4	3	0	0	1	1	1	11	9	7	0.8	0.7	0.9
Intron; out-of-frame	0	0	0	0	7	6	5	1	1	1	1	1	1	1	6	3	3	15	12	10	2	2	32	25	22	2.3	2.1	3.0
Frameshift	0	0	0	0	63	56	41	37	31	26	23	22	18	37	34	30	86	81	66	2	1	1	248	225	182	18.0	18.6	24.5
Multiple variant	0	0	0	0	7	6	4	5	5	5	2	2	2	0	0	0	6	4	2	0	0	0	20	17	13	1.5	1.4	1.8
Complex	2	1	1	3	2	2	0	0	0	0	2	2	2	1	1	1	2	2	2	1	1	1	11	9	9	0.8	0.7	1.2
Unknown	4	4	1	1	1	1	0	0	0	0	0	0	0	0	0	0	2	2	2	8	8	8	15	15	15	1.1	1.2	2.0
Total	18	16	12	297	253	149	84	74	61	61	84	75	40	221	192	117	633	563	327	38	36	36	1375	1209	742	100	100	100

Between brackets length of the domain in terms of cDNA nucleotides, all analyzed alleles, fam unrelated families, uniq unique molecular events in DNA sequence

Table 5.2 Spectrum of variants in the structural BTK domains

Domain	Length	Length/total length	Normalized expected	Observed	χ^2	P^a
PH	414	0.209	242	253	0.47384	0.49122
TH	231	0.117	135	74	27.69467	<10 ⁻⁶ ***
SH3	194	0.098	114	75	13.07900	0.00030***
SH2	292	0.148	171	192	2.60845	0.10630
TK	846	0.428	495	563	9.31070	0.00228**
Total	1977	1	1157	2257		

^aSignificance levels: ** $p < 0.01$ *** $p < 0.001$

seems very high, it is considered to be realistic because the kinase domain contains so many conserved regions and has several functions. The situation is likely very different in the SH3 and TH domains.

The variants are classified in Table 5.1 based on their effects on DNA or RNA level. The largest group of the variants is amino acid substitution causing missense variations (41 % of independent families). The SH3 domain is the only one where amino acid substitutions do not occur. Although SH3 domains are abundant in the human proteome, no disease-causing amino acid substitutions have been reported in any of them.

Nonsense variations account for 17 % of all variations, frameshift variations 19 %, and intronic variations 14 %. The proportions of deletions (4 %) and insertions (0.7 %) are very low and different from those reported in previous publications [10, 28] where proportions of 20 % (deletions) and 7 % (insertions) were given. These differences are due to the way the variants were counted, e.g., a variant with a DNA name “deletion” and an RNA name “frameshift” has been considered here as a frameshift variation. In the future we will avoid this kind of issues by adopting variation naming according to the Variation Ontology [30].

The distribution of variation types is very similar compared to the other IDbases [19]. The ratio of missense/nonsense variations, 2.5, is slightly higher in BTKbase compared to IDbases (1.5). Multiple variants in *BTK* have been identified in 17 families, complex variations in 9 families, and miscellaneous cases in 15 families.

There are altogether 341 unique amino acid substitutions. The theoretical maximum is 4151: thus, until now we have 8.2 % of the total variation; however, just a fraction of them is harmful and thus identifiable from XLA patients. In the case of nonsense variations, a larger portion has been seen in patients. There are 94 (28 %) of all the possible ($n = 297$) variants in the BTKbase. According to χ^2 statistics, there is highly significant overrepresentation ($p < 0.0001$).

When we are looking at the changes at amino acid level, it is apparent that arginine, as previously indicated, harbors the largest number of variants (Table 5.3). However, the most common outcome at protein level is protein truncation due to incorporation of a stop codon to the coding region. Altogether 29.5 % of single nucleotide changes lead to protein truncation.

Table 5.3 Amino acid substitutions indicated in percentages

	Hydrophobic							Hydrophilic							Special				Total				
	A	F	I	L	M	V	W	Y	D	E	H	K	R	N	Q	S	T	C		G	P	X	
A	0					1		0.6	0.4							0	0		0	0.8	0	2.8	
F		0	0	0.4		0.1		0.3								1		0.1			0.1	2.1	
I		0.1	0	0	0.1	0					0	0	1	0	0.1	0.6						0	2
L		1.3	0.3	0	0	0.3	0.1			0		0.6		0.1	0.6					3.8	0.6	7.6	
M			1.1	0.1	0	0.4					0.4	0.1					2				0	4.2	
V	0.4	0.7	0	0	0	0.1		0.4	0.1										0.3	0	0.1	2.2	
W				0.1									0.8			0.3		0.3	0		4.1	5.6	
Y		0						0	0.7		0.6			0.6		0.8		1.3			4.8	8.7	
D	0					0.4	0.1	0	0.1	0.1				0.3					0.3		0	1.4	
E	0					0		0.7	0		0.3				0				0.4		1.8	3.2	
H				0			0.1	0.1		0		0.4	0	0.1						0.3	0	1.1	
K			0.1		0				1.1		0	0.4	0.3	0	0	0					1.4	3.4	
R			0	0.3	0		6.3			4.1	0.6	0.1		4.5	1	0.3	2.5	2	1.1	8.3	31		
N			0				0.1	0		0	0.1		0		0	0					0.1	0.4	
Q			0	0					0	0.3	0	0.1		0						0.1	6.2	6.7	
S	0	0.7	0.1	0.1			0	0.4					0.1	0		0	0	0	0	0	0.7	1	3.2
T	0.1		0.4		0							0	0	0		0	0			1.1	0	1.7	
C		0.7					0.3	1.7					0.4			0.3		0	0.4		0.7	4.5	
G	0.1					0.1	0.1	1.1	2.1				1.5		0		0	0	0.1		0.3	5.6	
P	0.3			0.6		0					0		0.3		0	0.7	0.7			0	0	2.5	
Total	1	3.5	2.1	1.7	0.1	2.5	6.9	2.8	3.6	3.9	5	1.4	5	2.1	4.8	4.8	3.5	4.2	3.5	8	29.5	100	

Arginine is by far the most frequently substituted amino acid (31 %). This has not only been observed in BTK before [28] but also in variant datasets extracted from dbSNP [21], and this overrepresentation of arginine is known to be due to the high mutability of the codons containing CpG dinucleotides. Arginine is coded by six codons, four of which have a CpG dinucleotide in the first and second codon position [18]. The overrepresentation of arginine as the most frequently substituted amino acid also leads to the enrichment of tryptophan as the residue other amino acids are substituted to; arginine was replaced by tryptophan in 6.3 % of all amino acid substitutions (Table 5.3).

Proline is the amino acid to which most amino acids have been substituted to (8 %) closely followed by tryptophan (6.9 %), histidine (5 %), arginine (5 %), glutamine (4.8 %), and serine (4.8 %).

The G>A and C>T substitutions form the largest classes of changes, ~24 % (Table 5.4). The types of base changes were investigated more closely. The changes from amino to keto base and vice versa are much more frequent than substitutions within these groups. There is clearly a higher frequency of transitions (purine to purine and pyrimidine to pyrimidine, 66 %) than transversions (34 %). The higher rate of transitions agrees with the higher rate (~70 %) of transitions found to be typical for human genes [25]. The strong to weak base substitutions are by far the biggest category, containing 60 % of the variations. This was also found in the VariSNP variant datasets [21].

Table 5.4 Nucleotide substitutions in unique families (%)

All					
→	a	c	g	t	Total
a	0	3.5	8.3	2.8	14.6
c	5	0	2.1	23.5	30.7
g	23.7	5.5	0	7.8	36.9
t	3.6	10	4.2	0	17.8
Total	32.3	19	14.6	34.1	100

Summarized into amino and keto categories

→	Amino	Keto	Total
Amino	8.5	36.7	45.2
Keto	42.8	12	54.8
Total	51.3	48.7	100

Summarized into weak and strong categories

→	Weak	Strong	Total
Weak	6.4	26	32.4
Strong	60	7.6	67.6
Total	66.4	33.6	100

Summarized into purine and pyrimidine categories

→	Purines	Pyrimidines	Total
Purines	31.9	19.6	51.5
Pyrimidines	14.9	33.6	48.5
Total	46.9	53.1	100

5.2.2 Structural Consequences

BTK consists of five domains which, except for the SH3 domain, contain amino acid substitutions (Fig. 5.2). The effects and consequences of the variations vary widely. A recent study revealed that about two thirds of all kinase domain variations originating from a single nucleotide change likely lead to XLA [27]. This is not to say that two thirds of all possible amino acid changes were harmful since the majority of them do not originate from single base changes (because of the organization of the genetic code). Numerous variants affect functional sites, such as ligand- and substrate-binding regions at the domains. Stability-affecting changes are common. There are putative explanations available for the consequences of all the 1495 substitutions studied. These results are well in line with previous studies and predictions of BTK variants [5, 7–9, 11, 12, 14, 16, 20, 22, 24, 28, 32, 35–41].

Minor changes can be accommodated without major structural alterations. As has been seen in especially the PH domain, changes to electrostatics are common [16]. When the charge is reversed, added, or removed, the properties of the site are modified. If this happens on the protein surface of the binding site, then the interactions with partners are impaired or weakened.

Structural variations appear frequently in secondary structural elements. Although there are some variations at loops connecting these elements, the α - and

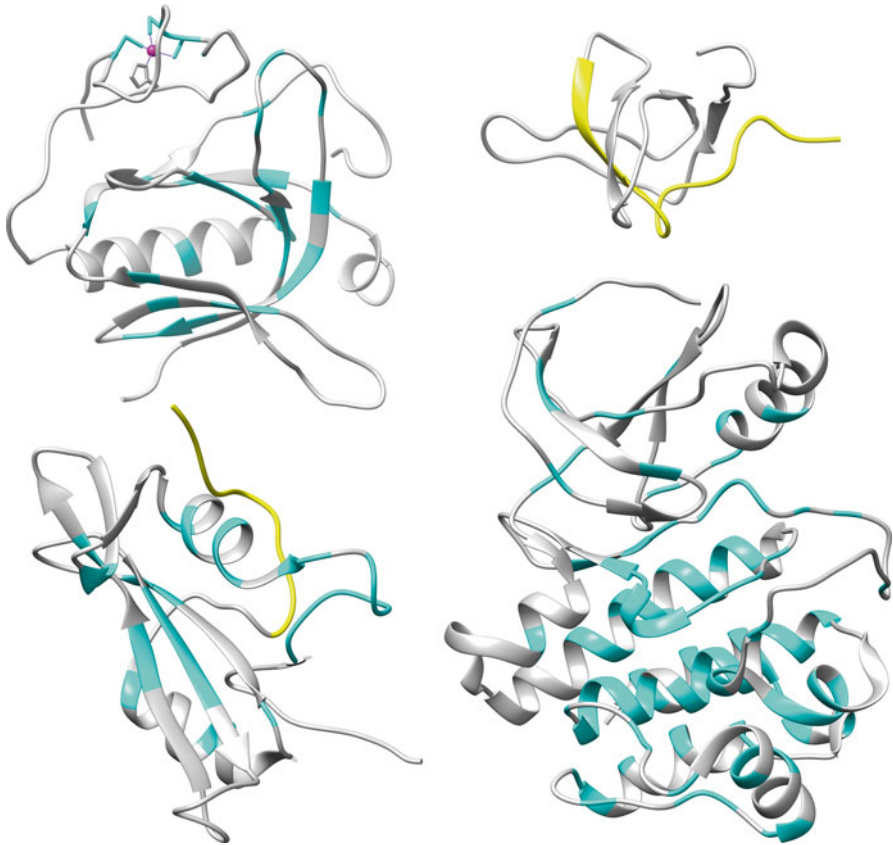


Fig. 5.2 Distribution of amino acid substitutions to BTK domains. Affected amino acids are shown in yellow. PH domain is on top left (PDB code 1BTK [7]). The first part of the TH domain including the BTK motif binding Zn^{2+} (magenta) is on the top of the domain. In the SH3 (1AWX [4]), top right, and SH2 domain (2GE9 [6]), bottom left, an in-frame deletion of 21 residues is indicated. The kinase domain (1K2P [12]) is at *bottom to the right*. Amino acid substitutions appear throughout all the domains except the SH3 domain where there are none

β -structures are more sensitive for substitutions. Structural variants are frequent on the protein core where there is no space for larger side chains due to tight packing. Further, introduction of charged or polar residues to the protein core, even if sterically possible, is usually harmful. Much more variation is allowed on the protein surface in areas not involved in intra- or intermolecular interactions. Some of these interactions are known; however, we do not even know the three-dimensional organization of the entire BTK. The domains are independently folding and connected by loops, which can be quite long. It is likely that the domain interactions are different in different folds of the entire protein. There is structural information for the entire BTK in elongated conformation [13]; however, this conformation is not likely the only one.

BTK variation information has been collected already for two decades into BTKbase, which has been a central resource for research and diagnosis. The database is constantly growing; however, the recent explosion in sequencing activities has not contributed much to the increased numbers in the database. That is presumably because many cases remain in laboratories and are never published or submitted to a database. It is in the interest of the entire community to share information about variations, especially in rare diseases such as XLA.

Acknowledgments Financial support from Lund University, Medical Faculty, and Swedish Research Council Vetenskapsrådet is gratefully acknowledged.

References

1. Berglöf A, Turunen JJ, Gissberg O, Bestas B, Blomberg KE, Smith CI (2013) Agammaglobulinemia: causative mutations and their implications for novel therapies. *Expert Rev Clin Immunol* 9:1205–1221
2. Celli J, Dalgleish R, Vihinen M, Taschner PE, den Dunnen JT (2012) Curating gene variant databases (LSDBs): toward a universal standard. *Hum Mutat* 33:291–297
3. den Dunnen JT, Antonarakis SE (2000) Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum Mutat* 15:7–12
4. Hansson H, Mattsson PT, Allard P, Haapaniemi P, Vihinen M, Smith CIE, Härd T (1998) Solution structure of the SH3 domain from Bruton tyrosine kinase. *Biochemistry* 37:2912–2924
5. Holinski-Feder E, Weiss M, Brandau O, Jedele KB, Nore B, Bäckesjö CM, Vihinen M, Hubbard SR, Belohradsky BH, Smith CIE, Meindl A (1998) Mutation screening of the BTK gene in 56 families with X-linked agammaglobulinemia (XLA): 47 unique mutations without correlation to clinical course. *Pediatrics* 101:276–284
6. Huang KC, Cheng HT, Pai MT, Tzeng SR, Cheng JW (2006) Solution structure and phosphopeptide binding of the SH2 domain from the human Bruton tyrosine kinase. *J Biomol NMR* 36:73–78
7. Hyvönen M, Saraste M (1997) Structure of the PH domain and Btk motif from Bruton tyrosine kinase: molecular explanations for X-linked agammaglobulinemia. *Embo J* 16:3396–3404
8. Jin H, Webster ADB, Vihinen M, Sideras P, Vořechovský I, Hammarström L, Bernatowska-Matuszkiewicz E, Smith CIE, Bobrow M, Vetrie D (1995) Identification of Btk mutations in 20 unrelated patients with X-linked agammaglobulinemia (XLA). *Hum Mol Genet* 4:693–700
9. Korpi M, Väliäho J, Vihinen M (2000) Structure-function effects in primary immunodeficiencies. *Scand J Immunol* 52:226–232
10. Lindvall JM, Blomberg KE, Väliäho J, Vargas L, Heinonen JE, Berglöf A, Mohamed AJ, Nore BF, Vihinen M, Smith CIE (2005) Bruton tyrosine kinase: cell biology, sequence conservation, mutation spectrum, siRNA modifications, and expression profiling. *Immunol Rev* 203:200–215
11. Maniar HS, Vihinen M, Webster ADB, Nilsson L, Smith CIE (1995) Structural basis for X-linked agammaglobulinemia (XLA): mutations at interacting Btk residues R562, W563, and A582. *Clin Immunol Immunopathol* 76:S198–S202
12. Mao C, Zhou M, Uckun FM (2001) Crystal structure of Bruton tyrosine kinase domain suggests a novel pathway for activation and provides insights into the molecular basis of X-linked agammaglobulinemia. *J Biol Chem* 276:41435–41443
13. Márquez JA, Smith CIE, Petoukhov MV, Lo Surdo P, Mattsson PT, Knekt M, Westlund A, Scheffzek K, Saraste M, Svergun DI (2003) Conformation of full-length Bruton tyrosine kinase (Btk) from synchrotron X-ray solution scattering. *Embo J* 22:4616–4624

14. Mattsson PT, Lappalainen I, Bäckesjö CM, Brockmann E, Lauren S, Vihinen M, Smith CIE (2000) Six X-linked agammaglobulinemia-causing missense mutations in the Src homology 2 domain of Bruton tyrosine kinase: phosphotyrosine-binding and circular dichroism analysis. *J Immunol* 164:4170–4177
15. Mohamed AJ, Yu L, Bäckesjö CM, Vargas L, Faryal R, Aints A, Christensson B, Berglöf A, Vihinen M, Nore BF and others (2009) Bruton tyrosine kinase (Btk): function, regulation, and transformation with special emphasis on the PH domain. *Immunol Rev* 228:58–73
16. Okoh MP, Vihinen M (1999) Pleckstrin homology domains of tec family protein kinases. *Biochem Biophys Res Commun* 265:151–157
17. Okoh MP, Vihinen M (2002) Interaction between Btk TH and SH3 domain. *Biopolymers* 63:325–334
18. Ollila J, Lappalainen I, Vihinen M (1996) Sequence specificity in CpG mutation hotspots. *FEBS Lett* 396:119–122
19. Piirilä H, Väliäho J, Vihinen M (2006) Immunodeficiency mutation databases (IDbases). *Hum Mutat* 27:1200–1208
20. Saha BK, Curtis SK, Vogler LB, Vihinen M (1997) Molecular and structural characterization of five novel mutations in the Bruton tyrosine kinase gene from patients with X-linked agammaglobulinemia. *Mol Med* 3:477–485
21. Schaafsma GCP, Vihinen M (2015) VarisNP, a benchmark database for variations from dbSNP. *Hum Mutat* 36:161–166
22. Shen B, Vihinen M (2004) Conservation and covariance in PH domain sequences: physicochemical profile and information theoretical analysis of XLA-causing mutations in the Btk PH domain. *Protein Eng Des Sel* 17:267–276
23. Smith CIE, Baskin B, Humire-Greff P, Zhou JN, Olsson PG, Maniar HS, Kjellén P, Lambris JD, Christensson B, Hammarström L, Bentley D, Vetrie D et al (1994) Expression of Bruton agammaglobulinemia tyrosine kinase gene, BTK, is selectively down-regulated in T lymphocytes and plasma cells. *J Immunol* 152:557–565
24. Speletas M, Kanariou M, Kanakoudi-Tsakalidou F, Papadopoulou-Alataki E, Arvanitidis K, Pardali E, Constantopoulos A, Kartalis G, Vihinen M, Sideras P, Ritis K (2001) Analysis of Btk mutations in patients with X-linked agammaglobulinaemia (XLA) and determination of carrier status in normal female relatives: a nationwide study of Btk deficiency in Greece. *Scand J Immunol* 54:321–327
25. Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, Stanley SE, Jiang R, Messer CJ, Chew A, Han JH and others (2001) Haplotype variation and linkage disequilibrium in 313 human genes. *Science* 293:489–493
26. Tsukada S, Saffran DC, Rawlings DJ, Parolini O, Allen RC, Klisak I, Sparkes RS, Kubagawa H, Mohandas T, Quan S, Belmont JW, Cooper MD et al (1993) Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia. *Cell* 72:279–290
27. Väliäho J, Faisal I, Ortutay C, Smith CIE, Vihinen M (2015) Characterization of all possible single nucleotide change –caused amino acid substitutions in the kinase domain of Bruton tyrosine kinase. *Hum Mutat* 36:638–647
28. Väliäho J, Smith CIE, Vihinen M (2006) BTKbase: the mutation database for X-linked agammaglobulinemia. *Hum Mutat* 27:1209–1217
29. Vetrie D, Vořechovský I, Sideras P, Holland J, Davies A, Flinter F, Hammarström L, Kinnon C, Levinsky R, Bobrow M, Smith CIE, Bentley DR (1993) The gene involved in X-linked agammaglobulinaemia is a member of the src family of protein-tyrosine kinases. *Nature* 361:226–233
30. Vihinen M (2014) Variation Ontology for annotation of variation effects and mechanisms. *Genome Res* 24:356–364
31. Vihinen M, Belohradsky BH, Haire RN, Holinski-Feder E, Kwan SP, Lappalainen I, Lehväsäläho H, Lester T, Meindl A, Ochs HD, Ollila J, Vořechovský I et al (1997) BTKbase, mutation database for X-linked agammaglobulinemia (XLA). *Nucleic Acids Res* 25:166–171
32. Vihinen M, Cooper MD, de Saint Basile G, Fischer A, Good RA, Hendriks RW, Kinnon C, Kwan SP, Litman GW, Notarangelo LD, Ochs HD, Rosen FS et al (1995) BTKbase: a database of XLA-causing mutations. *Immunol Today* 16:460–465

33. Vihinen M, den Dunnen JT, Dalglish R, Cotton RG (2012) Guidelines for establishing locus specific databases. *Hum Mutat* 33:298–305
34. Vihinen M, Lehtälä H, Cotton RDH (1999) Immunodeficiency mutation databases. In: Ochs HD, Smith CIE, Puck JM (eds) *Primary immunodeficiency diseases. A molecular and genetic approach*. Oxford University Press, New York/Oxford, pp 443–447
35. Vihinen M, Nilsson L, Smith CIE (1994) Structural basis of SH2 domain mutations in X-linked agammaglobulinemia. *Biochem Biophys Res Commun* 205:1270–1277
36. Vihinen M, Nilsson L, Smith CI (1994) Tec homology (TH) adjacent to the PH domain. *FEBS Lett* 350:263–265
37. Vihinen M, Nore BF, Mattsson PT, Backesjö CM, Nars M, Koutaniemi S, Watanabe C, Lester T, Jones A, Ochs HD, Smith CI (1997) Missense mutations affecting a conserved cysteine pair in the TH domain of Btk. *FEBS Lett* 413:205–210
38. Vihinen M, Vetrie D, Maniar HS, Ochs HD, Zhu Q, Vořechovský I, Webster ADB, Notarangelo LD, Nilsson L, Sowadski JM, Smith CIE (1994) Structural basis for chromosome X-linked agammaglobulinemia: a tyrosine kinase disease. *Proc Natl Acad Sci U S A* 91:12803–12807
39. Vihinen M, Zvelebil MJ, Zhu Q, Brooimans RA, Ochs HD, Zegers BJ, Nilsson L, Waterfield MD, Smith CIE (1995) Structural basis for pleckstrin homology domain mutations in X-linked agammaglobulinemia. *Biochemistry* 34:1475–1481
40. Vořechovský I, Vihinen M, de Saint Basile G, Honsova S, Hammarström L, Müller S, Nilsson L, Fischer A, Smith CIE (1995) DNA-based mutation analysis of Bruton tyrosine kinase gene in patients with X-linked agammaglobulinaemia. *Hum Mol Genet* 4:51–58
41. Zhu Q, Zhang M, Rawlings DJ, Vihinen M, Hagemann T, Saffran DC, Kwan SP, Nilsson L, Smith CIE, Witte ON, Chen S-H, Ochs HD (1994) Deletion within the Src homology domain 3 of Bruton tyrosine kinase resulting in X-linked agammaglobulinemia (XLA). *J Exp Med* 180:461–470