 Nucleic Acids Research, 2000, Vol. 28, No. 1 **369–371**

KinMutBase, a database of human disease-causing protein kinase mutations

Kaj A. E. Stenberg[1], Pentti T. Riikonen[2,3] and Mauno Vihinen[2,*]

[1]Department of Biosciences, Division of Biochemistry, PO Box 56, FIN-00014 University of Helsinki, Finland,
[2]Institute of Medical Technology, University of Tampere, PO Box 607, FIN-33101 Tampere, Finland and
[3]Department of Computer Science, University of Turku, Lemminkäisenkatu 14A, FIN-20520 Turku, Finland

Received October 4, 1999; Accepted October 6, 1999

ABSTRACT

KinMutBase (http://www.uta.fi/imt/bioinfo/KinMutBase/) is a registry of mutations in human protein kinases related to disorders. Kinases are essential cellular signaling molecules, in which mutations can lead to diseases, including immunodeficiencies, cancers and endocrine disorders. The first release of KinMutBase contained information for protein tyrosine kinases. The current release includes also serine/threonine protein kinases, as well as an update of the tyrosine kinases. There are 251 entries altogether, representing 337 families and 621 patients. Mutations appear both in conserved hallmark residues of the kinases as well as in non-homologous sites. The KinMutBase WWW pages provide plenty of information, namely mutation statistics and display, clickable sequences with mutations and changes to restriction enzyme patterns.

INTRODUCTION

The human genome has been estimated to encode for ~2000 protein kinases. These enzymes catalyze protein phosphorylation: the transfer of a phosphate group to serine, threonine, tyrosine, arginine or histidine residues of the target molecules. Protein phosphorylation constitutes one of the most frequent post-translational modifications of structural and regulatory proteins. Maybe about one third of the cellular proteins are phosphorylated at some stage. The activity of kinases is finely regulated by phosphorylation/dephosphorylation and interactions with other molecules. Abnormal phosphorylation is observed in numerous pathological situations including endocrine disorders, cancers, immunodeficiencies and cardiovascular diseases.

Defects in protein kinases have been implicated behind a number of diseases and knowledge of these disorders is of great clinical interest. Mutations can be either activating or inactivating in nature. Activating mutations turn the enzymes constitutively on, leading to constant activation of certain signaling cascades. These kind of oncogenic diseases are often observed in tumors. Inactivating mutations block signaling pathways and thereby prevent cells from responding to certain stimuli. In many cases the inactivation is not so severe, presumably due to redundant activities of related kinases. Certain kinase knock-out mice do not present any phenotype indicating redundancy in signaling pathways. In some instances, however, overlapping activities do not exist.

In order to facilitate the analysis of disease-causing protein kinase mutations, information has been extracted from the literature and compiled into a database, KinMutBase. Diseases present in the database include carcinomas, immunodeficiences and endocrinological disorders. Common to all these diseases is the association with mutations in the kinase domain of a protein kinase. A listing of the kinases so far implemented in KinMutBase can be found in Table 1.

The most studied of all the kinases harboring disease-causing mutations are protein tyrosine kinases (PTKs) Bruton's tyrosine kinase (Btk) and insulin receptor kinase (Irk). Mutations in Btk prevent B cell maturation and production of immunoglobulins in X-linked agammaglobulinemia (XLA) (1–3). KinMutBase is directly linked to BTKbase, the registry of XLA-causing mutations. Mutated insulin receptor kinase (IRK) is in charge of non-insulin dependent diabetes mellitus (NIDDM) and insulin resistance (4,5). Jak3 has in addition to a kinase domain (JH1) a kinase-like pseudokinase domain (JH2). Mutations in both the domains have been noted to cause severe combined immunodeficiency (SCID) (6). ZAP-70 kinase domain mutations also cause SCID (7).

Several mutated kinases are related to tumor development. Ret mutations can be either activating or inactivating and cause different disorders, namely Hirschprung's disease, multiple endocrine neoplasia type 2 (MEN2) a and b, and medullary thyroid carcinoma (FMTC) (8). MET mutations are connected to papillary renal carcinoma (HPRC) (9). The ser/thr-kinase (PSK) LKB1 has been linked to Peutz–Jeghers syndrome (10), and two other PSKs in the database, CDK4 and TGR2, are also involved in tumor formation (11,12).

Kkit is a mast/stem cell growth factor receptor kinase, which is responsible for mastocytosis, systematic mast cell disease and piebaldism when mutated (13,14). Another growth factor receptor, FGFR3, is connected to hypochondroplasia, thanatophoric dysplasia and skeletal dysplasia when mutated (15,16). The Coffin–Lowry syndrome (CLS), characterized by mental retardation and progressive skeletal deformations has been connected to the PSK RSK2 (17). The modified neurotrophic tyrosine kinase NTRK1 has been shown to be responsible for congenital insensitivity to pain with anhidrosis (CIPA) (18). Vascular dysplasia and errors in vascular morphogenesis have

*To whom correspondence should be addressed. Tel: +358 3 215 7735; Fax: +358 3 215 7710; Email: ltmavi@uta.fi

Table 1. Distribution of unique mutations/unrelated families/patients in kinase domains of protein kinases

Kinase	Missense	Nonsense	Insertion	Deletion	Total
Protein tyrosine kinases					
Btk	62/102/120	21/37/42	10/8/10	30/30/33	123/147/205
FGFR3	6/41/43	1/1/1	–	–	7/42/44
IRK	12/15/22	–	–	1/2/2	13/17/24
JAK3	1/1/1	1/1/1	–	1/1/1	3/3/3
KIT	8/10/13	–	1/2/2	2/2/4	11/14/19
MET	9/10/10	–	–	–	9/10/10
NTRK1	6/6/8	–	–	1/2/4	7/8/12
RET	21/25/150	1/1/1	–	–	22/26/151
ZAP	1/1/1	–	1/2/2	1/1/1	3/4/4
TIE2	1/2/65	–	–	–	1/2/65
Protein serine/threonine kinases					
PAK3	–	1/1/4	–	–	1/1/4
LKB1	9/9/9	6/6/6	6/8/8	2/2/2	23/25/25
KIR3	4/4/18	–	–	–	4/4/18
CDK4	3/13/16	–	–	–	3/13/16
TGR2	5/5/5	–	–	–	5/5/5
RSK2	7/7/7	6/6/6	–	3/3/3	16/16/16
Total	155/251/488	37/53/61	18/20/22	41/43/50	251/337/621

been connected to kinase domain mutations in the ser/thr-kinases KIR3 and TIE2 (19,20).

Two of these kinases, the PTK Btk and the PSK PAK3 have been shown to cause X-linked diseases, XLA and X-linked mental retardation, respectively (1–3,21). The X-linked nature of the diseases greatly simplifies the analysis of the genetics and structural bases for the diseases.

KinMutBase

KinMutBase contains information about disease-causing protein kinase domain mutations. The current version contains data about tyrosine and serine/threonine kinases. The data has been compiled from literature except for Btk and Jak3 mutations, which were extracted from BTKbase (22) and JAK3base (M.Vihinen, A.Villa, P.Mella, R.F.Schumacher, J.J.O'Shea, F.Candotti and L.Notarangelo, submitted), respectively.

The database contains the following information for each mutation. The identification and plain English description of the mutation is followed by reference and formal characterization of the mutation. The same mutation may have occurred several times. The frequency of mutations, as a number of affected patients and families, is included to facilitate statistical analyses. MEDLINE and OMIM (23) links are included when available, as well as links to BTKbase and JAK3base. Mutations are also linked to reference sequences at genomic, cDNA and amino acid level, depending on availability.

Mutation data can be submitted to KinMutBase by filling in a questionnaire on the web pages. The program writes the data into the database format. However, new entries are added only after curator inspection. KinMutBase includes several features for the analysis of the effects of mutations. The localization of the mutations can be displayed separately for each kinase or simultaneously for all of them. By clicking the figures the mutation(s) affecting certain positions are listed. This can be performed either at genomic, cDNA or amino acid level. Several tables provide information about the distribution of mutations. Several clickable pages are related to restriction patterns and modifications in them. These analyses include all the restriction enzymes in the latest version of REBASE (24).

CONCLUSIONS

There are currently 251 mutations in the database (Table 1) representing 621 patients in 337 unrelated families. About half of the mutations (123) appear in Btk, the most extensively studied kinase with disease causing alterations (2). A significant number of mutations have been found also in the Irk, Kit, Met and Ret kinases, as well as in the PSKs LKB1 and RSK2. Most of the kinases contain mutational hotspots, sites where mutations have been identified several times. In Btk and in many other diseases such sites are in CpG dinucleotides, which generally have pyrimidines 5′ and purines on their 3′ side (25). The total number of different mutations in Ret is much lower than, for example, in Btk, but a single site, M918, harbors by far the greatest number of mutations. The M918T mutation, which accounts for almost all MEN2B cases, has been shown to occur on the paternal chromosome (26). The exact number of unrelated Ret families is not possible to determine based on the published articles. The same also applies to some other cases. The numbers of families in the database include only those that can be verified from the literature.

The distribution of mutations in Btk is similar to many other diseases with respect to mutation types. In the other kinases, missense mutations dominate and nonsense mutations, for example, are clearly underrepresented. This can be due to several reasons, e.g., due to detection methods used. In Jak3, a number of mutations affect also the inactive pseudokinase domain. The majority of these are deletions, which affect the folding of the whole protein, but there are also some missense mutations.

The structural basis for the kinase related diseases have been studied in the case of Irk (27), Btk (16,28–30) and Jak3 (M.Vihinen, A.Villa, P.Mella, R.F.Schumacher, J.J.O'Shea, F.Candotti and L.Notarangelo, submitted) by using crystal structures and molecular modeling. Some of the mutations affect the conserved hallmark residues, but there are also numerous alterations in non-conserved residues. The structural consequences range from impaired folding to catalytic or substrate-binding mutations to alterations in sites forming contacts with other domains or ligands.

CITATION

Users of KinMutBase are asked to cite this article in their publications, including the URL http://www.uta.fi/imt/bioinfo/KinMutBase

DISTRIBUTION OF THE DATABASE

The database is available on the WWW. Researchers are asked to submit new mutations preferably by using the form in the KinMutBase WWW pages or contacting the database curators.

SUPPLEMENTARY MATERIAL

See supplementary material available at NAR Online.

ACKNOWLEDGEMENTS

Financial support from the Finnish Academy, The National Technology Agency of Finland, The Medical Research Fund of Tampere University Hospital and Sigrid Juselius foundation is gratefully acknowledged.

REFERENCES

1. Vihinen,M., Mattsson,P. and Smith,C.I.E. (1997) *Frontiers Biosci.*, **2**, d384–d399.
2. Mattsson,P., Vihinen,M. and Smith,C.I.E. (1996) *Bioessays*, **18**, 825–834.
3. Smith,C.I.E., Bäckesjö,C.-M., Berglöf,A, Brandén,L.J., Islam,T., Mattsson,P.T., Mohamed,A.J., Müller,S., Nore,B. and Vihinen,M. (1998) *Springer Semin. Immunopathol.*, **19**, 369–381.
4. Cocozza,S., Porcellini,A., Riccardi,G., Monticelli,A., Condorelli,G., Ferrara,A., Pianese,L., Miele,C., Capaldo,B., Beguinot,F. and Varrone,S. (1992) *Diabetes*, **41**, 521–526.
5. Moller,D.E., Yokota,A., White,M.F., Pazianos,A.G. and Flier,J.S. (1990) *J. Biol. Chem.*, **265**, 14979–14985.
6. Candotti,F., O'Shea,J.J. and Villa,A. (1988) *Springer Semin. Immunopathol.*, **19**, 401–415.
7. Chan,A.C., Kadlecek,T.A., Elder,M.E., Filipovich,A.H., Kuo,W.L., Iwashima,M., Parslow,T.G. and Weiss,A. (1994) *Science*, **264**, 1599–1601.
8. Eng,C. and Mulligan,L.M. (1997) *Human Mut.*, **9**, 97–109.
9. Schmidt,L., Duh,F.-M., Chen,F., Kishida,T., Glenn,G., Choyke,P., Scherer,S.W., Zhuang,Z., Lubensky,I., Dean,M. *et al.* (1997) *Nature Genet.*, **16**, 68–73.
10. Hemminki,A., Markie,D., Tomlinson,I., Avizienyte,E., Roth,S., Loukola,A., Bignell,G., Warren,W., Aminoff,A., Höglund,P. *et al.* (1998) *Nature*, **391**, 184–187.
11. Wölfel,T., Hauer,M., Schneider,J., Serrano,M., Wölfel,C., Klehman-Hieb,E., De Plaen,E., Hankeln,T., zum Büschenfelde,K.-H.M. and Beach,D. (1995) *Science*, **269**, 1281–1284.
12. Knaus,P.I., Lindemann,D., DeCoteau,J.F., Perlman,R., Yankelev,H., Hille,M., Kadin,M.E. and Lodish,H.F. (1996) *Mol. Cell. Biol.*, **16**, 3480–3489.
13. Furitsu,T., Tsujimura,T., Tono,T., Ikeda,H., Kitayama,H., Koshimizu,U., Sugahara,H., Butterfield,J.H., Ashman,L.K., Kanayama,Y. *et al.* (1993) *J. Clin. Invest.*, **92**, 1736–1744.
14. Giebel,L.B. and Spritz,R.A. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 8696–8699.
15. Bellus,G.A., Mcintosh,I., Smith,E.A., Aylsworth,A.S., Kaitila,I., Horton,W.A., Greenhaw,G.A., Hecht,J.T. and Francomano,C.A. (1995) *Nature Genet.*, **10**, 357–359.
16. Tawormina,P.L., Shiang,R., Thompson,L.M., Zhu,Y-Z., Wilkin,D.J., Lachman,R.S., Wilcox,W.R., Rimoin,D.L., Cohn,D.H. and Wasmuth,J.J. (1995) *Nature Genet.*, **9**, 321–328.
17. Jacquot,S., Merienne,K., De Cesare,D., Pannetier,S., Mandel,J.-L., Sassone-Corsi,P. and Hanauer,A. (1998) *Am. J. Hum. Genet.*, **63**, 1631–1640.
18. Indo,Y., Tsuruta,M., Hayashida,Y., Karim,M.A., Ohta,K., Kawano,T., Mitsubishi,H., Tonoki,H., Awaya,Y. and Matsuda,I. (1996) *Nature Genet.*, **13**, 485–488.
19. Johnson,D.W., Berg,J.N., Baldwin,M.A., Gallione,C.J., Marondel,I., Yoon,S.-J., Stenzel,T.T., Speer,M., Pericak-Vance,M.A., Diamond,A., Guttmacher,A.E. *et al.* (1996) *Nature Genet.*, **13**, 189–195.
20. Vikkula,M., Boon,L.M., Carraway,K.L.,III., Calvert,J.T., Diamonti,A.J., Goumnerov,B., Pasyk,K.A., Marchuk,D.A., Warman,M.L., Cantley,L.C. *et al.* (1996) *Cell*, **87**, 1181–1190.
21. Allen,K.M., Gleeson,J.G., Bagrodia,S., Partington,M.W., MacMillan,J.C., Cerione,R.A., Mulley,J.C. and Walsh,C.A. (1998) *Nature Genet.*, **20**, 25–30.
22. Vihinen,M., Brandau,O., Brandén,L.J., Kwan,S.-P., Lappalainen,I., Lester,T., Noordzij,J.G., Ochs,H.D., Ollila,O., Pienaar,S.M., Riikonen,P., Saha,B.K. and Smith,C.I.E. (1998) *Nucleic Acids Res.*, **26**, 242–247.
23. McKussick,V.A. (1994) *Mendelian Inheritance in Man. Catalogs of Human Genes and Genetic Disorders*. 11th edition, John Hopkins University Press, Baltimore, MD.
24. Roberts,R.J. and Macelis,D. (1998) *Nucleic Acids Res.*, **26**, 338–350. Updated article in this issue: *Nucleic Acids Res.* (2000), **28**, 306–307.
25. Ollila,J., Lappalainen,I. and Vihinen,M. (1996) *FEBS Lett.*, **396**, 119–122.
26. Carlson,K.M., Bracamontes,J., Jackson,C.E., Clark,R., Lacroix,A., Wells,S.A. and Goodfellow,P.J. (1994) *Am. J. Hum. Genet.*, **55**, 1076–1082.
27. Hubbard,S.R., Wei,L., Ellis,L. and Hendrickson,W.A. (1994) *Nature*, **372**, 746–754.
28. Vihinen,M., Vetrie,D., Maniar,H.S., Ochs,H.D., Zhu,Q., Vorechovsky,I., Webster,A.D.A., Notarangelo,L.D., Nilsson,L., Sowadski,J.M. and Smith,C.I.E. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 12803–12807.
29. Vihinen,M., Brooimans,R.A., Kwan,S.-P., Lehväslaiho,H., Litman,G.W., Resnick,I., Ochs,H.D., Schwaber,J.H., Vorechovsky,I. and Smith,C.I.E. (1996) *Immunol. Today* **17**, 502–506.
30. Vihinen,M., Kwan,S.P., Lester,T., Resnick,I., Väliaho,J., Conley,M.E. and Smith,C.I.E. (1999) *Hum. Mutat.* **13**, 280–285.