Kinase mutations in human disease: interpreting genotype–phenotype relationships

Piya Lahiry, Ali Torkamani‡, Nicholas J. Schork‡ and Robert A. Hegele**

Abstract | Protein kinases are one of the largest families of evolutionarily related proteins and comprise one of the most abundant gene families in humans. Here we survey kinase gene mutations from the perspective of human disease phenotypes and further analyse the structural features of mutant kinases, including mutational hotspots. Our evaluation of the genotype–phenotype relationship across 915 human kinase mutations — that underlie 67 single-gene diseases, mainly inherited developmental and metabolic disorders and also certain cancers — enhances our understanding of the role of kinases in development, kinase dysfunction in pathogenesis and kinases as potential targets for therapy.

Amino acids
Amino acids contain a basic amino (NH_2) group, an acidic carboxyl (COOH) group and a side chain attached to an alpha carbon atom. The 20 amino acids can be classified based on the charge of their side chain, which can be neutral non-polar, neutral polar, acidic or basic.

Apoptosis
The process of programmed cell death that does not involve the release of harmful substances into the surrounding area. It has crucial function in division and differentiation by eliminating cells that are unnecessary for appropriate embryonic development.

**Robarts Research Institute and Departments of Medicine and Biochemistry, Schulich School of Medicine and Dentistry, University of Western Ontario, London, Ontario N6A 5C1, Canada.*
‡Scripps Genomic Medicine, Scripps Research Institute, La Jolla, California 92037, USA.
Correspondence to R.A.H.
e-mail: hegele@robarts.ca
doi:10.1038/nrg2707

Protein kinases comprise one of the largest families of evolutionarily related proteins and >500 distinct kinases are encoded by ~2% of all human genes[1]. Kinases transiently phosphorylate specific amino acids on ~30% of all human proteins, including molecules that govern complex cellular processes, such as growth, differentiation, proliferation and apoptosis[2]. Given the importance of these cellular activities, the catalytic activity of kinases involved in these pathways is stringently regulated[2]. Over the past 20 years, mutations in kinase genes have been found to underlie many human diseases, particularly developmental and metabolic disorders, as well as certain cancers. The isolated nature and large volume of the individual gene-mapping studies has tended to impede efforts to produce a comprehensive survey of kinase mutations. However, an overview of kinase mutations from the perspective of human disease phenotypes could reveal patterns of structure–function relationships that govern development and pathogenesis and that could, in turn, complement experiments from *in vitro* and *in vivo* model systems.

For this Review, we have curated inherited germline kinase gene mutations by phenotype according to organ system involvement. We also briefly review the relationship between kinase gene mutations in somatic cells and cancer phenotypes. We next consider the consequences of gain- and loss-of-function mutations in kinase genes, genetic pleiotropy and locus heterogeneity of certain disease phenotypes. The analysis of the distribution of mutations has helped to define crucial functional domains of kinases, mutational hotspots across the kinase gene family and the relevance of evolutionarily mediated lineage-specific variations (BOX 1). An integrated understanding of kinase structure and function will improve the diagnosis and prognosis of diseases that are related to kinase gene variation. Because kinases comprise ~20% of all putative drug targets[3], we also discuss how defining phenotype–genotype relationships will help the development of future therapies.

Basics of kinase structure and function

Kinase structure and classification. Most kinases, excluding the atypical kinases, contain a conserved catalytic domain, which transfers a phosphate from ATP to a target protein[4]. This phosphorylation can modulate enzyme activity or affect the interaction of the kinase with its protein targets[4], which in turn can specify the downstream response. Regulatory domains are present either within or outside the kinase catalytic domain and help both to localize the kinase and to modulate its activity in response to various stimuli (FIG. 1). The two main classes of kinase are tyrosine kinases (TKs) and serine–threonine kinases (STKs), which phosphorylatetyrosine and serine or threonine residues on a substrate, respectively[5]. Both TKs and STKs can be membrane-bound and nuclear; in addition, TKs can be transmembrane receptors whereas STKs can also be cytoplasmic. BOX 2 shows the diverse functions of archetypal STKs and TKs.

The catalytic domain of protein kinases consists of 250–300 amino acids and contains 12 conserved subdomains that fold into a common catalytic core structure[5]. The amino-terminal (N)-lobe of the catalytic

Genetic pleiotropy
The effect of a single gene on multiple phenotypic traits. The underlying mechanism is related to the effects of the gene product on various targets.

Locus heterogeneity
This occurs when a phenotype is caused by mutations at more than one gene locus, which suggests that the products of the genes belong to the same metabolic pathway.

Mutational hotspots
A region in which the frequency of mutation is greater than expected, owing to specific structural and/or functional features of the protein or gene.

Kinome
The set of protein kinases in the genome of an organism.

domain contains a glycine-rich stretch of residues (GxGxxG) that is crucial for ATP binding and phosphoryl transfer. By contrast, the C-terminus of the catalytic domain is involved in substrate binding, and the N-terminus of this domain contains a conserved aspartic acid that is important for the catalytic activity of the enzyme[5].

Phylogenetic relationships. The human kinome can be subdivided into seven major groups according to the sequence and structure of the catalytic domain[1], including TK, TKL (tyrosine kinase-like), STE (homologues of the yeast sterile 7, sterile 11 and sterile 20 kinases), CK1 (casein kinase 1), AGC (family of protein kinases A, G and C), CAMKs (calcium/calmodulin-dependent protein kinases) and the highly conserved CMGC subgroup, which contains GSK3, CLK, cyclin-dependent kinases (CDKs) and mitogen-activated protein kinases (MAPKs). The CMGC subgroup contains both MAPKs, which control cellular processes across all eukaryotic phyla, and also less-understood kinases, such as intestinal cell kinase (ICK).

As major kinase groups and most kinase family members have a conserved structure and function throughout the evolution of metazoans, which include all multicellular eukaryotes, it is not surprising that these proteins are under high selective pressure to resist the accumulation of variation, such that any variation that leads to the abnormal activation or suppression of kinase activity will result in severe phenotypic consequences. However, there is no obvious relationship between the kinase subgroup and patterns of disease susceptibility[6,7].

Box 1 | Kinase genomic mutation databases and resources

Several genomic databases and web resources were accessed during the preparation of this Review, some of which are described here. The Kinase Sequence Database is a collection of protein kinase sequences grouped into families by the homology of their catalytic domains. Kinase.com explores the function, evolution and diversity of the protein kinases that comprise the kinome. The Protein Kinase Resource provides a compendium of information on the protein kinases, including tools for structural and computational analyses as well as links to related databases. Kinweb provides a comprehensive analysis of the functional domains of each kinase gene product and a collection of conserved sequence elements that have been identified by the comparative analysis of human kinase genes and their murine counterparts. The Kinase Pathway Database classifies protein kinases and their functions. The COSMIC database contains data on somatic mutations in human cancer, and combines curation of the scientific literature with tumour resequencing data from the Cancer Genome Project at the Sanger Institute, UK.

The characterization of mutations and naturally occurring genetic mutations that affect kinase protein structure, function and, ultimately, clinical phenotypic endpoints will be greatly facilitated through several recent genomics initiatives. Large-scale sequencing initiatives, such as The Cancer Genome Atlas and the 1000 Genomes Project will identify hundreds of thousands of coding variations, many of which will be rare and many of which will be present in kinases. Although Online Mendelian Inheritance in Man is not a specific curated resource for kinase gene mutations, it contains extensive clinical descriptions and mutation lists for many of the disorders discussed in this Review. There are also disease- or gene-specific kinase databases, such as the multiple endocrine neoplasia type 2 (MEN2) *RET* database. Once mutations have been identified, their functional effects need to be assessed using model organisms, such as transgenic mice, and *in vitro* functional studies. The ability to overexpress and inactivate (knockout) genes can be valuable to understand the complex changes in phenotype due to genomic alterations. The technology of knockout mice has been such an invaluable tool to study human development that the National Institutes of Health and the International Knockout Mouse Consortium have created a public resource of mouse embryonic stem cells containing null or conditional mutations in every gene in the mouse genome[85]. Finally, the recent restructuring of the Protein Structure Initiative to emphasize biomedical applications will enhance the characterization of normal and mutant kinase structures, leading to greater insight into the molecular effects of genetic variations.

Human diseases due to germline mutations

Careful phenotypic characterization — which we have called 'phenomic analysis' (REFS 8,9) — of patients with rare monogenic diseases can indicate patterns of organ system involvement that suggest the presence of bottlenecks and redundancies in tissue-specific kinase function and expression. These *in vivo* patterns may in turn suggest novel biological associations[10]. As a first step towards a curated database of kinasopathies, we have collected phenotype and genotype data from 67 kinase-related germline disorders in humans. The following criteria were used to construct this list: STKs and TKs were included; the corresponding kinasopathies were derived from the Online Mendelian Inheritance in Man (OMIM) and UniProt databases; the mutations underlying each kinasopathy listed had to have been validated functionally *in vitro* using biochemical assays or *in vivo* using induced mutant animal models. An abbreviated version of this list is shown in TABLE 1 and Supplementary information S1 (table) shows more complete details.

Our survey found 50 kinases that underlie 67 distinct single-gene clinical entities. Approximately half of these disease-associated kinases were TKs. Of the 915 curated disease-associated mutations, 77% were missense mutations, 19% were nonsense mutations and 4% were associated with altered splicing. More than 80% of mutations directly affected or encompassed the catalytic domain of the respective kinase gene. Virtually all organ systems were affected by kinase mutations, and although many disorders involved more than one tissue, most kinasopathies could be classified according to the predominant organ involvement. Some illustrative kinasopathies, which are grouped by organ system involvement, are discussed below.

Kinase signalling pathways can also be indirectly activated by mutations in negative regulators or downstream signalling components, such as mutations in the protein tyrosine phosphatase in Noonan syndrome[11]; however, such mutational mechanisms that do not directly involve kinase genes will not be further discussed here.

Neurological disorders. Autosomal recessive kinase mutations are the predominant cause of disparate neurological diseases, which range from degenerative and encephalopathic disorders to epilepsies, myasthenia and ataxia. For example, lethal congenital contracture syndrome type 2 (LCCS2) is an autosomal recessive neurodegenerative disorder characterized by the degeneration of anterior horn neurons and joint contractures[12]. LCCS2 results from a loss-of-function splicing mutation in *ERBB3*, which encodes a member of the epidermal growth factor receptor (EGFR) family of receptor TKs.

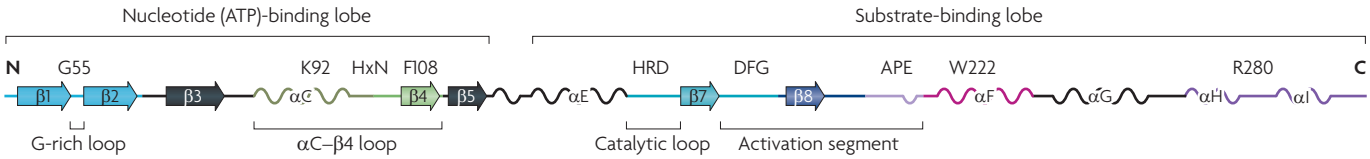


Figure 1 | **Generic catalytic domain of protein kinases depicting subdomain structure and function.** The subdomains and functional elements of a generic protein kinase catalytic domain are displayed as a linear ribbon diagram. Block arrows correspond to β-sheets (β1–5, β7 and β8), waves correspond to α-helices (αC and αE–I) and intervening lines correspond to unstructured loops (the elements β6 and αD are very short and not depicted in the diagram). The G-rich loop is one of the most flexible elements of the protein kinase catalytic core and plays a key part in the phosphoryl-transfer reaction. The HxN is a conserved motif in the amino (N)-terminal αC–β4 loop, which serves as a hinge point for αC-helix movement. The position of the αC-helix varies in response to the activation state of the protein kinase, such that it swings outward when the kinase is inactive and inward when the kinase is active. The HRD is a conserved motif that includes a conserved aspartate involved directly in catalysis and a regulatory arginine that coordinates with the phosphorylated residue in the activation loop after autophosphorylation. Autophosphorylation governs the activation state of protein kinases. The DFG is a conserved motif that chelates the magnesium ion involved in catalysis. Movements of this loop are required for adoption of the active conformation. Finally, APE is a conserved motif in the C-terminal substrate recognition pocket of kinases. Conserved functional residues (G55, K92, F108, W222 and R280) or motifs are displayed above the ribbon, the numbering corresponds to residues of protein kinase A. Motifs, named according to the amino acids involved, are depicted above the ribbon and functional regions are annotated below the ribbon.

In a second example, 18 different missense mutations in *PRKCG*, which encodes the STK protein kinase Cγ, cause the autosomal dominant neurodegenerative disorder spinocerebellar ataxia type 14 (REF. 13). PRKCG requires activation by calcium and diacylglycerol before phosphorylating several signalling protein targets[14], and *Prkcg* deficiency leads to learning deficits in rodents[15,16]. However, spinocerebellar ataxia type 14 is an exceptional neurodegenerative kinasopathy that arises from a dominant activating mutation. The fact that most other neurological kinasopathies result from homozygous loss-of-function mutations suggests that these diseases arise from the failure to adapt to an impaired — rather than an augmented — signalling network.

Skeletal and craniosynostosis disorders. Kinasopathies that affect the skeleton are usually caused by autosomal dominant gain-of-function mutations in TKs. For example, gain-of-function mutations that affect the fibroblast growth factor receptor (FGFR) family underlie several dysplasias that are characterized by the premature fusing of skull sutures in infancy[17]. Binding of FGFs to their receptors normally activates the TK domain of the FGFR; this is followed by phosphorylation of downstream signalling, including the activation of the RAS (a small GTPase) and MAPK cascade, which ultimately induces mitogenesis and differentiation[18] (BOX 2).

In contrast to TK mutations, STK mutations seem to affect bone development more generally. For example, missense mutations in *ACVR1* (activin A receptor type 1) have been associated with fibrodysplasia ossificans progressiva, which is characterized by extensive ossification[19] and has a homologous phenotype in a mouse model[19,20]. Activins are growth and differentiation factors in the transforming growth factor (TGF)-β superfamily[21].

The range of genes and preponderance of germline activating kinase mutations that result in skeletal phenotypes reinforces the prominent role of kinases in bone development and differentiation.

Haematological and vascular disorders. Although there are a few germline mutations that cause haematological disease (TABLE 1 and Supplementary information S1 (table)), the discovery of somatic mutations — including large chromosomal rearrangements and gene fusion events — in some leukaemias and myeloproliferative disorders provided early evidence of the importance of kinase gene mutations in human disease. The archetypal example of such rearrangements is the B cell receptor (*BCR*)–*ABL* fusion gene that underlies chronic myeloid leukaemia (CML) and acute lymphocytic leukaemia[22]. The demonstration that other fusion genes involving kinases underlie leukaemia suggested a more general association, and has led to the development of TK inhibition using agents such as imatinib[23] as a therapeutic approach for CML. Interestingly, the response to imatinib is modulated by the BCR–ABL genotype of the patient[23]. Another example of a kinase fusion gene is a somatic *BCR–FGFR1* fusion gene that is seen occasionally in patients with CML[24].

Somatic fusion events involving the Janus kinase gene *JAK2*, *ETV6* (REFS 25,26) and pericentriolar material 1 (*PCM1*)[27] are also associated with leukaemia, and a somatic missense mutation (V617F) in *JAK2* is seen in a large proportion of patients with polycythemia vera[28]. Interestingly, a *JAK2* germline mutation has recently been reported to affect disease susceptibility in primary myelofibrosis regardless of the V617F mutational status[29].

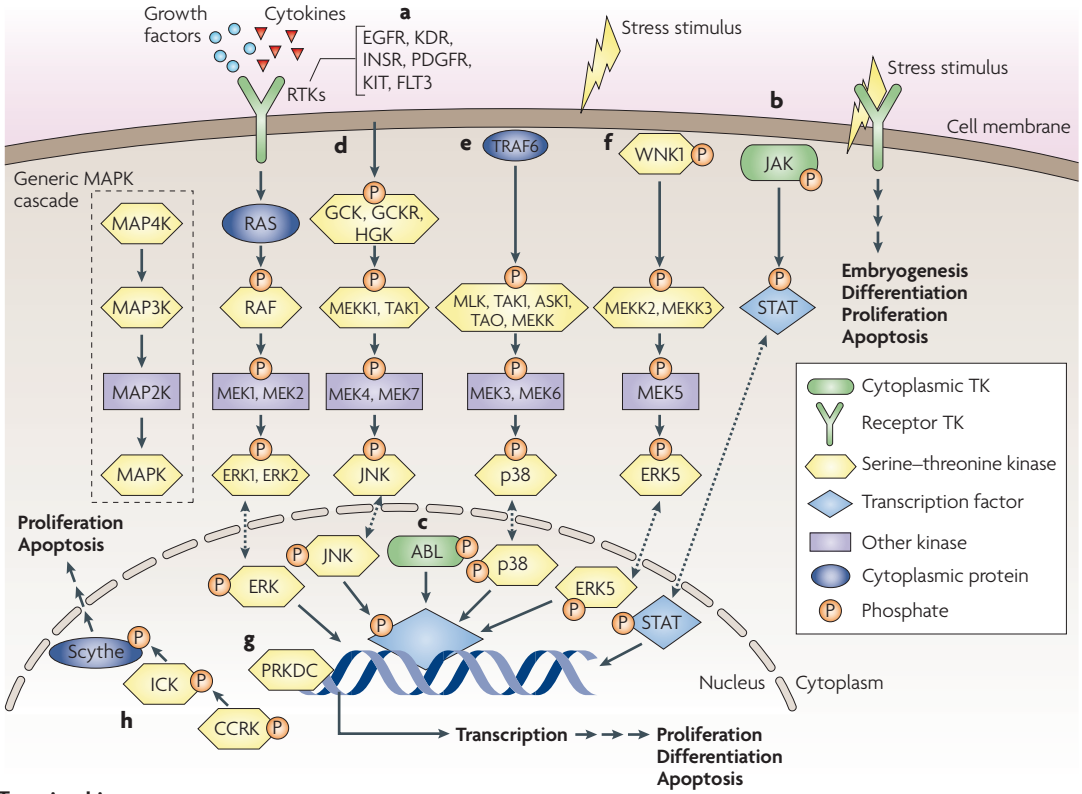
Germline kinase mutations can also affect the heart and blood vessels. For example, almost 30 mutations in the gene encoding the STK bone morphogenetic protein receptor type 2 (*BMPR2*) are associated with autosomal dominant primary pulmonary hypertension[30], implicating BMPR2 in the development of the pulmonary vasculature due to its interaction with c-SRC[31]. Knockout of *Bmpr2* is embryonic lethal in mice due to aberrant mesoderm development[32]. In addition, mutant BMPR2 has an inhibitory effect on breast cancer cells, suggesting it may be used as a potential anti-cancer therapeutic agent[33].

Kinasopathy
A clinical phenotype that is caused by germline mutations in the kinase domain of functional proteins that lead to a loss-of-function or gain-of-function of the protein.

Myasthenia
A general term for an inherited neuromuscular disorder characterized by fluctuating muscle weakness and fatiguability that is often caused by one of several types of functional molecular defects at the neuromuscular junction.

Polycythemia vera
A blood disorder in which the bone marrow overproduces red blood cells (and sometimes other blood components). The resulting increase in blood viscosity can lead to health problems, especially enhanced blood clotting.

Box 2 | **Classification and role diversity of tyrosine and serine–threonine kinases**



Tyrosine kinases

Tyrosine kinases (TKs) can be subclassified as receptor and non-receptor proteins[1]: receptor TKs (RTKs) transduce extracellular signals to the cytoplasm[108] whereas non-receptor TKs are intracellular proteins that relay intracellular signals[108].

RTKs (**a** in the figure above) contain a ligand-binding extracellular domain, an intracellular catalytic domain and a transmembrane domain, which may contain a disulphide bond that connects the extracellular and intracellular regions of the receptor[108]. The highly conserved catalytic domain is responsible for TK activity and several regulatory functions, such as acting as a binding site for proteins containing SRC homology 2 (SH2) domains[108]. RTKs can phosphorylate RAS, which initiates the RAF–MEK–ERK phosphorylation cascade, and leads to a direct effect on gene expression through phosphorylated ERK; this function is important for cell division, migration and survival[109,110]. RTKs can use several other pathways to regulate events such as embryogenesis and overall cellular homeostasis.

Non-receptor TKs, which can be membrane-bound or nuclear-specific, have diverse roles in cell signalling (**b** in the figure). For example, after Janus kinase (JAK) — a membrane-bound TK — is phosphorylated by extracellular stimuli, such as interleukin 6, it can activate signal transducer and activator of transcription (STAT), which is involved in regulation of gene expression[111]. Once nuclear TKs such as ABL are stimulated they can activate the transcription factor RB, which can lead to cellular growth inhibition (**c** in the figure)[112].

Serine–threonine kinases

Serine–threonine kinases (STKs) phosphorylate diverse target substrates, such as transcription factors, cell cycle regulators and an array of cytoplasmic and nuclear effector molecules[113]. Moreover, STK activity is regulated by specific triggers, including growth factors, cytokines and physical or chemical stressors[114]. For example, extracellular stimuli can activate cytoplasmic STKs, such as the JNK pathway of the mitogen-activated protein kinase (MAPK) signalling cascade (**d** in the figure), and trigger the translocation of phosphorylated JNK into the nucleus where, through the transcription factor JUN, it can stimulate cellular apoptosis[110,115]. Another manner by which cytoplasmic MAPK can be activated is through other cytoplasmic proteins. For example, the stress-activated MAPK, p38, is phosphorylated through the cascade that begins with the inner membrane-bound TNF-receptor-associated factor 6 (TRAF6) and ends with phosphorylated p38 localizing in the nucleus to induce effects such as cell motility, inflammation, osmoregulation, chromatin remodelling and apoptosis (**e** in the figure)[109,110,115]. An example of an inner membrane-bound STK activating a MAPK signalling cascade is WNK1. WNK1 has been seen to be involved in the MEKK3–MEK5–ERK5 cascade that causes the phosphorylation of ERK5, which ultimately leads to changes in gene expression (**f** in the figure)[116]. Nuclear STKs such as PRKDC rely on DNA binding to trigger their kinase activity and ability to phosphorylate several transcription factors (**g** in the figure)[117]. Intestinal cell kinase (ICK), a MAP-like kinase, is an example that requires nuclear localization for the initiation of kinase activity (by cell cycle-related kinase (CCRK)) and has been reported to exert its effects on Scythe, an anti-apoptotic protein that resides in the cytoplasm and nucleus (**h** in the figure)[118–120]. ASK1, apoptosis signal-regulating kinase 1; EGFR, epidermal growth factor receptor; GCK, germinal centre kinase; HGK, haematopoietic progenitor kinase/germinal centre kinase-like kinase; INSR, insulin receptor; KDR, kinase insert domain receptor; MLK, mixed lineage kinase; PDGFR, platelet-derived growth factor receptor.

Immunological disorders. Both autosomal dominant and recessive mutations implicate certain kinases in the normal function of the immune system[34]. For example, Bruton X-linked agammaglobulinaemia is a severe immunodeficiency resulting from a failure to produce mature B lymphocytes. Almost 150 (mostly missense) mutations have been reported in Bruton tyrosine kinase (*BTK*), which encodes a key regulator of B cell development[35]. BTK interacts with TKs involved in B lymphocyte signalling pathways[36], and mice with a *Btk* mutation located outside the kinase domain failed to produce B cells[37]. Furthermore, missense mutations in *IKBKG* underlie two X-linked diseases, incontinentia pigmenti and hypohidrotic ectodermal dysplasia with immune deficiency, which indicate the importance of this STK in both T and B cell function[38]. *Ikbkg*-deficient mouse embryos fail to survive due to severe liver damage resulting from excessive apoptosis[39].

Endocrine and metabolic disorders. A preponderance of autosomal dominant mutations underlies endocrine and metabolic kinasopathies, and highlights the existence of genetic pleiotropy. For example, FGFR1 is important in bone development, and also in gonadogenesis and reproduction (REF. 40) (FIG. 2a). Several loss-of-function mutations in *FGFR1* are associated with hypogonadotrophic hypogonadism[40]. Also, a mutation in *AKT2*, which encodes a ubiquitous STK that has a key downstream function in activating the insulin receptor, causes autosomal dominant insulin resistance with type 2 diabetes[41], metabolic dyslipidemia, lipodystrophy and hepatic steatosis[42]. These findings, together with corroborating evidence in *Akt2*-knockout mice[43], indicate the central importance of AKT signalling to insulin sensitivity. As highlighted by the example of FGFR1, a kinase can affect different organs based on the mutation

Table 1 | **Inherited kinasopathies**

Gene symbol (OMIM number)	Locus	Disease name (OMIM number)	MOI
Tyrosine kinases			
ALK (105590)	2p23	Neuroblastoma (316014)	AD
BTK (300300)	Xq21.3–q22	Agammaglobulinemia (300300)	XL
EFNB1 (300035)	Xq12	Craniofrontonasal syndrome (304110)	XL
EIF2AK3 (604032)	2p12	Wolcott–Rallison syndrome (226980)	AR
ERBB3 (190151)	12q13	Lethal congenital contractural syndrome type 2 (607598)	AR
FGFR1 (136350)	8p11.2–p11.1	Osteoglophonic dysplasia (166250), Pfeiffer syndrome (101600) and hypogonadotropic hypogonadism (146110)	AD
FGFR2 (176943)	10q26	Pfeiffer syndrome (101600), Apert syndrome (101200), Crouzon syndrome (123500) and lacrimoauriculodentodigital syndrome (149730)	AD
FGFR3 (134934)	4p16.3	Achondrodysplasia (100800), thanatophoric dysplasia types 1 and 2 (187600 and 187601) and Muenke syndrome (602849)	AD
FLT4 (136352)	5q35.3	Hereditary lymphedema type 1A (153100)	AD
GRK1 (180381)	13q34	Oguchi disease 1 (258100)	AR
INSR (147670)	19p13.2	Insulin-resistant diabetes with acanthosis (610549) and Donahue syndrome (246200)	AD, AR
JAK3 (600173)	19p13.1	Severe combined immunodeficiency (600802)	AR
KIT (164920)	4q12	Piebaldism (172800)	AD
LTK (151520)	15q15.1–q21.1	Systemic lupus erythematosus (152700)	AR
MERTK (604705)	2q14.1	Retinitis pigmentosa 38 (268000)	AR
MUSK (601296)	9q31.3–q32	Myasthenic syndrome (608931)	AR
NTRK1 (191315)	1q23–q24	Congenital insensitivity to pain with anhidrosis (256800)	AR
NTRK2 (600456)	9q22.1	Early obesity, hyperphagia and developmental delay (600456)	AD
PANK2 (606157)	20p13–p12.3	Pantothenate kinase-type neurodegeneration (234200)	AR
PDGFRA (173490)	4q12	Gastrointestinal stromal tumour (606764)	AD
RET (164761)	10q11.2	Multiple endocrine neoplasia type 2B (162300), familial medullary thyroid carcinoma (155240), familial pheochromocytoma (171300) and Hirschsprung disease (142623)	AD
RET (164761)	10q11.2	Congenital failure of autonomic control (209880) and renal dysplasia (191830)	ND
ROR2 (602337)	9q22	Robinow syndrome (268310)	AR
TEK (600221)	9p21	Cutaneous and mucosal venous malformations (600195)	AD
ZAP70 (176947)	2q12	Severe-combined immunodeficiency (T cell-negative) (176947)	AR

Table 1 (cont.) | **Inherited kinasopathies**

Gene symbol (OMIM number)	Locus	Disease name (OMIM number)	MOI
Serine–threonine kinases			
ACVR1 (102576)	2q23–q24	Fibrodysplasia ossificans progressiva (135100)	AD
ACVRL1 (601284)	12q13	Hereditary haemorrhagic telangiectasia type 2 (600376)	AD
AKT2 (164731)	19q13.1–q13.2	Atypical lipodystrophy (125853)	AD
BMPR1B (603248)	4q23–q24	Brachydactyly type 2A (112600)	AD
BMPR2 (600799)	2q33	Primary pulmonary hypertension 1 (178600)	AD
BRAF (164757)	7q34	Cardio–facio–cutaneous syndrome (115150)	AD
CDKL5 (300203)	Xp22	Early infantile epileptic encephalopathy type 2 (300672)	XL
CHEK2 (604373)	22q12.1	Li–Fraumeni syndrome 2 (609265)	AD
ICK (612325)	6p12.3	Endocrine–cerebro–osteodysplasia (612651)	AR
IKBKG (300248)	Xq28	Hypohidrotic ectodermal dysplasia (300291) and incontinentia pigmenti type 2 (308300)	XL
IRAK4 (606883)	12q12	Invasive pneumococcal disease (610799) and pyogenic bacterial infections (607676)	AR
MAPK10 (602897)	4q21.3	Epileptic encephalopathy Lennox–Gaustaut type (606369)	AR
MEK1 (176872)	15q21	Cardio–facio–cutaneous syndrome (115150)	AD
MEK2 (601263)	7q32	Cardio–facio–cutaneous syndrome (115150)	AD
PHKA2 (306000)	Xp22.2–p22.1	Glycogen storage disease type 9A (types 1 and 2) (306000)	XL
PHKG2 (172471)	16p12.1–p11.2	Glycogen storage disease type 9C (172471)	AR
PINK1 (608309)	1p36	Early-onset Parkinson disease 6 (605909)	AR
PRKAR1A (188830)	17q23–q24	Primary pigmented nodular adrenocortical disease (610489) and Carney complex (160980)	AD
PRKCG (176980)	19q13.4	Spinocerebellar ataxia type 14 (605361)	AD
RAF1 (164760)	3p25	Noonan syndrome type 5 (611553) and LEOPARD syndrome type 2 (611554)	AD
RPS6KA3 (300075)	Xp22.2–p22.1	Coffin–Lowry syndrome (303600)	XL
STK11 (602216)	19p13.3	Peutz–Jeghers syndrome (175200)	AD
TGFBR1 (190181)	9q22	Loeys–Dietz syndrome (types 1A and 2A) (609192 and 608967)	AD
TGFBR2 (190182)	3p23	Loeys–Dietz syndrome (types 1B and 2B) (610168 and 610380)	AD
TRPM7 (605692)	15q21	Amyotrophic lateral sclerosis–Parkinsonism (105500)	ND
WNK4 (601844)	17q21–q22	Pseudohypo-aldosteronism type 2 (145260)	AD

ACVR, activin A receptor; AD, autosomal dominant; *ALK*, anaplastic lymphoma kinase; AR, autosomal recessive; BMPR, bone morphogenetic protein receptor; *BTK*, Bruton tyrosine kinase; *CDKL5*, cyclin-dependent kinase-like 5; *EFNB1*, ephrin B1; FGFR, fibroblast growth factor receptor; *GRK1*, G protein-dependent receptor kinase 1; *ICK*, intestinal cell kinase; *INSR*, insulin receptor; *JAK3*, Janus kinase 3; *LTK*, leukocyte tyrosine kinase; *MAPK10*, mitogen-activated protein kinase 10; MOI, mode of inheritance; *MUSK*, muscle, skeletal, receptor tyrosine kinase; ND, not determined; NTRK, neurotrophic tyrosine kinase receptor; *PANK2*, pantothenate kinase 2; *PDGFRA*, platelet-derived growth factor receptor α; PHK, phosphorylase kinase; *ROR2*, receptor tyrosine kinase-like orphan receptor 2; TGFBR, transforming growth factor-β receptor; XL, X-linked.

type. However, mutations in certain kinases may affect multiple organs simultaneously, as seen in the cases in the following section.

Multi-organ disorders. Diseases involving multiple organs can implicate a role for a single kinase in development across a range of tissues. Mutations in *PRKAR1A* in Carney syndrome show the importance of this gene in the heart, endocrine, cutaneous and neuronal tissues (REF. 44) (FIG. 2b). Also, the involvement of *RPS6KA3* in Coffin–Lowry syndrome indicates its crucial role in skeletal, growth and cognitive development[45]. Among other STKs, ICK was implicated in endocrine–cerebro–osteodysplasia, a multi-organ neonatal lethal condition (REF. 46) (FIG. 2b). Although ICK has a developmental role across a range of tissues, perhaps early in development before extensive differentiation of embryonic germ layers, its signalling pathway and direct interactions with other proteins have yet to be fully characterized.

Germline and somatic mutations in cancer

Protein kinases are the most frequently mutated family of genes that contribute to neoplastic malignancies, with an approximately fourfold overrepresentation compared with a random selection of the same number

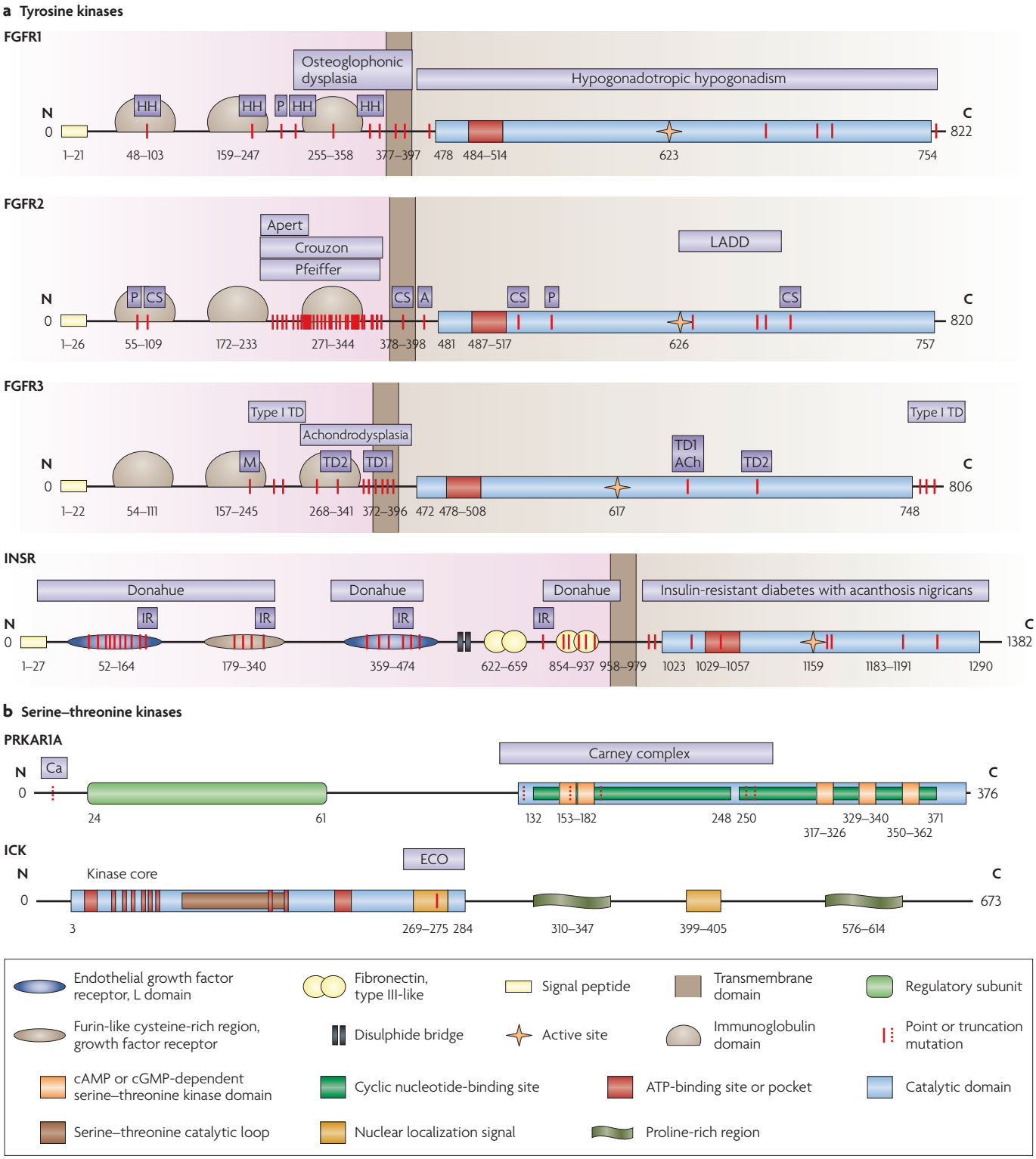


Figure 2 | **Genetic pleiotropy.** Germline point mutations in selected kinases illustrate genetic pleiotropy. Each disease name is displayed on the schematic diagram of its corresponding kinase on which point mutations are depicted with red lines. **a** | The four examples of tyrosine kinases (TKs) are the three proteins from the fibroblast growth factor receptor (FGFR) family (FGFR1, FGFR2 and FGFR3) and the insulin receptor (INSR). Point mutations in the FGFR family are associated with craniosynotosis syndromes, as well as osteo- and endocrine dysplasias. INSR mutations are associated with endocrine disorders. **b** | cAMP-dependent protein kinase type I-α regulatory subunit (PRKAR1A) and intestinal cell kinase (ICK) are shown as examples of serine–threonine kinases (STKs) that are associated with syndromes according to the location of point mutations in conserved regions. A, Apert syndrome; Ach, achondrodysplasia; C, carboxyl-terminus; Ca, Carney complex; CS, Crouzon syndrome; ECO, endocrine–cerebro–osteodysplasia; HH, hypgonadotropic hypogonadism; LADD, lacrimoauriculodentodigital syndrome; M, Muenke syndrome; N, amino-terminus; P, Pfeiffer syndrome; TD1, thanatophoric dysplasia type 1; TD2, TD type 2.

of genes[47]. Some overrepresentation of kinases might be attributable to ascertainment bias, given the large size of this protein family. However, mutations in most protein kinases were identified through positional cloning approaches, so ascertainment bias is unlikely to completely account for their overrepresentation among cancer genes[47].

Kinases as tumour suppressors or proto-oncogenes. Protein kinases may act as tumour suppressors or proto-oncogenes in normal, healthy cells. Therefore, mutations in protein kinases may lead to tumorigenesis through numerous mechanisms, including the activation of proliferative pathways, genomic instability, reduction of the DNA damage response, deactivation of apoptotic pathways and/or the promotion of angiogenesis and cell motility. As mentioned above, several somatic chromosomal translocations may also lead to constitutive kinase activation, but these will not be further considered here.

TABLE 2 provides examples of kinases that cause cancers due to somatic mutations, mainly from the Catalogue of Somatic Mutations in Cancer (COSMIC) database (REF. 48) (BOX 1). The kinases in this table have mutations in ≥1% of samples analysed and have been functionally evaluated in murine models (a more detailed list of kinases is provided in Supplementary information S2 (table)). TABLE 2 and Supplementary information S2 (table) indicate that many cancers are due to acquired somatic mutations and that most kinases are proto-oncogenes that develop these cancer-causing somatic mutations. However, some cancer syndromes are caused by inherited germline mutations in both proto-oncogenes and tumour suppressor genes (TABLE 1). Examples of cancers caused by inherited mutations in kinase proto-oncogenes include multiple endocrine neoplasia type 2, thyroid carcinoma, pheochromocytoma due to *RET* (the 'rearranged during transfection' gene) mutations, hereditary gastrointestinal stromal tumours due to platelet-derived growth factor receptor α (*PDGFRA*)[21] mutations and Li–Fraumeni syndrome due to *CHEK2* (REF. 49) mutations. Examples of tumour suppressor genes that lead to cancer due to both somatic and germline mutations are *MAPK10* in brain tumours and *STK11* in Peutz–Jeghers syndrome.

Structure–function relationships. The role of specific kinase structural mutations in both inherited cancers and those acquired through somatic mutations is seen in neuroblastoma[50]. Neuroblastoma is a devastating childhood cancer that has a low survival rate. Both germline and somatic mutations in anaplastic lymphoma receptor tyrosine kinase (*ALK*) have been identified in neuroblastoma patients[50–53], highlighting the role of both germline and somatic mutations of proto-oncogenes in tumorigenesis. Mosse *et al.* examined the distribution of the various mutations that have been identified over the ALK protein structure and inferred that most of the mutations were likely to be activating[50]. *In vitro* cell growth assays verified the growth advantage of the activating mutations such that in cell lines that contained either ALK amplification or ALK mutation, growth reduction occurred when levels of ALK were reduced by siRNA[50].

Specific malignancies mediated by protein kinases have been studied extensively to identify the functional roles and tissue specificities of the responsible kinase. For example, the primary FLT3-based malignancy is ALL[54]. This correlates with the observation that FLT3 is mainly found in normal haematopoietic cells[55]. By contrast, receptor tyrosine kinase-like orphan receptor 2 (ROR2)-based cancers, such as renal cell and colorectal-specific carcinoma[21,56], involve tissues in which ROR2 was not originally identified — that is, in the chondrocyte lineage, telencephelon, heart and dermis[57]. Therefore, a specific cancer phenotype provides insight into the biological functions regulated by the responsible mutant or perturbed protein kinase. Additionally, studying protein kinases involved in particular forms of cancer can shed light on the underlying pathological mechanisms. For example, TEK, which is a TK localized to endothelia, has been observed to cause breast, ovarian and renal cancers. This insight into the kinase localization pattern leads to better diagnosis and treatment as the malignant trigger is recognized to lie in the endothelial cell lineage rather than another tissue in the affected organ.

Genotype–phenotype relationships

Mode of inheritance in kinasopathies. TABLE 1 shows a preponderance of autosomal dominant kinase-related disorders compared with autosomal recessive or X chromosome-linked diseases. Interestingly, most of the disorders that involve the nervous and immune systems are autosomal recessive, whereas most of the disorders that affect the skeletal, haematological, vascular and endocrine and metabolic systems are autosomal dominant. Although the explanation for the division of mutation types across organ systems is not apparent, it may be relevant that the distribution of inheritance according to organ system involvement mirrors that seen with mutations in lamin A/C (*LMNA*), which encodes a structural non-kinase nuclear protein. Neurological laminopathies can show autosomal recessive inheritance, but myopathies, cardiomyopathies and endocrinopathies tend to show autosomal dominant inheritance[58]. Perhaps the correlation between the distribution of phenotypes and the mode of inheritance in kinasopathies — as well as in other disorders, such as laminopathies — reflects the robustness of the signalling networks in the course of normal development of particular tissues or organ systems, like the nervous system.

Mutations resulting in gain- or loss-of-function. A gain-of-function mutation increases constitutive kinase activity, sometimes leading to unrestrained cellular signalling and may trigger oncogenesis or cause rare inherited dominant phenotypes. Conversely, loss-of-function mutations can lead to a loss in cell signalling, which can affect cell growth and tissue or organ development. This is illustrated by mutations in the proto-oncogene *RET*, which encodes a receptor TK that belongs to the cadherin superfamily[59]. As RET is

Tumour suppressor
A molecule that inhibits uncontrolled cell growth such that its loss- or reduction-of-function mutation favours the formation of tumours.

Proto-oncogene
A gene that promotes the specialization and division of cells; however, when it is mutated or expressed at high levels, it causes abnormal cellular growth.

Neuroblastoma
A childhood cancer derived from immature neurons of the sympathetic nervous system.

Table 2 | **Selected kinases associated with cancer and related findings in induced mutant animal models**

Gene	Examples of somatic mutations associated with tumorigenesis*‡	Examples of effects of gene target deletion in transgenic animal models§
ACVR1	Melanoma	Embryonic KO is lethal at E9.5; essential in normal mesoderm formation and development after gastrulation; conditional KO in the ectoderm causes a reduction in lens size due to decreased lens cell proliferation
ALK	Ovary, breast and lung cancers, neuroblastomas and glioblastomas	Behavioural and neurochemical alterations, KO viable without any gross alterations
BMPR1B	Gastric adenocarcinoma and melanoma	Forms intact cartilaginous elements
BMPR2	Inhibitory effect on breast cancer cells by dominant-negative protein	Embryonic KO is lethal, fails to form organized structures and lacks mesoderm
BRAF	Widespread with greater incidence in ovary, skin, colon and thyroid cancers and glioblastomas	Midgestation lethality with vascular defects due to endothelial apoptosis
BTK	Lung carcinoma	Missense mutation leads to failure of mature B lymphocyte production and of immunoglobulin heavy chain rearrangement
CHEK2	Glioblastomas	Embryonic stem cells fails to maintain γ-irradiation-induced arrest in G2 phase
ERBB3	Prostate, bladder and breast cancers and glioblastomas	Prenatal lethality due to lack of Schwann cells and precursors
FGFR1	Stem cell leukemia lymphoma (FGFR1–ZNF198 chimerism), pancreatic adenocarcinomas, glioblastoma, breast carcinomas and lung cancers	Embryonic KO is embryonic lethal due to lack of embryonic growth and mesodermal patterning; in conditional KO mouse limb development is affected
FGFR2	Glioblastoma, breast, gastric, lung, ovarian, cervical and endometrial cancer	Full or partial gene KO is embryonic lethal as post-implantation development is disrupted
FGFR3	Prostrate and cervical cancer, lung, bladder, upper digestive tract and intestinal carcinoma and plasma cell myeloma	Embryonic KO has defects in long bones, vertebrae growth and inner-ear development; transgenic mice with a missense mutation have retarded endochondral bone growth
FLT4	Increase in metastasis in adenocarcinoma and lymph node cancer, glioblastoma, kidney and ovary carcinoma and melanoma	Missense mutation in the catalytic domain leads to chylous ascites accumulation and limb swelling
INSR	Stomach and skin cancers and glioblastomas and colorectal cancer	Normal birth but postnatal fatal diabetic ketoacidosis
IRAK4	Prostate cancer	Severely impaired interleukin 1 and Toll-like receptor signalling
JAK3	Acute megakaryoblastic leukemia and gastric adenocarcinoma	Knockout of the catalytic domain leads to reduced number of thymocytes and severe B cell and T cell lymphopenia
KIT	Testicular and ovarian tumours	Two mutations cause protein deficiency, which leads to white coat colour, sterility and anaemia due to migration and/or proliferation failure of stem cell populations
MAPK10	Loss of expression in some brain tumours	Reduced stress-induced JNK activity; protection from brain injury after cerebral ischaemia or hypoxia
MERTK	Renal, head and neck carcinoma	C-terminal truncation of the protein leads to macrophages that are deficient in the clearance of apoptotic thymocytes
MUSK	Lung cancer	Failure to induce neuromuscular synapse formation
NTRK1	Thyroid carcinoma	Sensory and sympathetic neuropathies
NTRK2	Skin and lung cancer	Conditional KO in the postnatal forebrain has reduced hippocampal-mediated learning and overall synaptic strengthening
PDGFRA	Neuroblastoma, stomach, soft tissue, small intestine, lung and gastrointestinal cancer, glioblastoma, melanoma and haematopoietic and lymphoid myeloma	Posterolateral diaphragmatic defects
PINK1	Glioblastoma and ovary carcinoma	Impaired mitochondrial function
PRKAR1A	Soft tissue myxoma, adrenocortical tumours and thyroid carcinomas	Embryonic KO is lethal due to failure of mesodermal structure development; heterozygous knockout has osteoblast neoplasia, and Schwann cell and thyroid tumours
RAF1	Ovarian and lung adenocarcinoma	Cardiac muscle-specific conditional deletion leads to left ventricular dysfunction and heart dilation
RET	Lung, ovarian, bladder, large intestinal carcinomas, pheochromocytoma, thyroid tumours and glioblastomas	Severe defects in kidney and enteric nervous system development
ROR2	Renal cell and colorectal carcinoma	Perinatal lethal with cardiac septal defects and skeletal defects

Table 2 (cont.) | **Selected kinases associated with cancer and related findings in induced mutant animal models**

Gene	Examples of somatic mutations associated with tumorigenesis*‡	Examples of effects of gene target deletion in transgenic animal models§§
RPS6KA3	Prostate cancer, breast cancer, gastric adenocarcinoma and glioblastoma	Progressive osteopenia due to impaired osteoblast and osteoclast function
STK11	Widespread with greater incidence in lung, cervical and pancreatic cancers and melanoma	Midgestation lethality due to neural tube defects, mesenchymal cell death and vascular abnormalities; heterozygous mice have multiple gastric adenomatous polyps
TEK	Breast, ovarian and renal cancer	Embryonic lethal due to cardiac dysfunction and vascular haemorrhaging
TGFBR1	TGFBR1*6A/9A polymorphism is associated with increased risk of breast and ovarian cancer.	Midgestation lethality due to abnormal angiogenesis yet intact haematopoietic potential in the yolk sac
TGFBR2	Hereditary non-polypopsis colon, oesophageal, stomach and lung cancers and glioblastomas	Embryonic lethality due to defects in yolk sac haematopoiesis and vasculogenesis
TRPM7	Inhibitory on head and neck carcinoma cells when expression suppressed; breast, ovary and stomach cancers	Embryonic knockout is lethal; conditional T cell-specific deletion shows disrupted thymopoiesis
WNK4	Lung and ovarian carcinoma and melanoma	Transgenic mouse has hypertension, hyperkalemia, hypercalciuria and marked hyperplasia of the distal convoluted tubule
ZAP70	Lung and head and neck carcinoma	Missense mutation in the SH2 domain causes autosomal chronic arthritis

*Cancer type listed for all somatic mutations recorded in the COSMIC database for each kinase. ‡Increased kinase expression or function, unless otherwise stated. §Germline gene knockout, unless otherwise stated. *ACVR1*, activin A receptor type 1; *ALK*, anaplastic lymphoma kinase; BMPR, bone morphogenetic protein receptor; *BTK*, Bruton tyrosine kinase; FGFR, fibroblast growth factor receptor; *INSR*, insulin receptor; *IRAK4*, interleukin 1 receptor-associated kinase 4; *JAK3*, Janus kinase 3; KO, knockout; *MAPK10*, mitogen-activated protein kinase 10; *MUSK*, muscle, skeletal, receptor tyrosine kinase; NTRK, neurotrophic tyrosine kinase receptor; *PDGFRA*, platelet-derived growth factor receptor α; *ROR2*, receptor tyrosine kinase-like orphan receptor 2; TGFBR, transforming growth factor-β receptor.

widely expressed and has a crucial role in neural crest lineages and regulates cell proliferation, migration, differentiation and survival during embryogenesis[59]. *RET* mutations cause markedly diverse phenotypes depending on their nature. Constitutive RET activity, due to a gain-of-function mutation, leads to various types of human cancer, including multiple endocrine neoplasia type 2B[60] and its individual components, such as familial medullary thyroid cancer[61] and familial pheochromocytoma[62]. However, human *RET* loss-of-function mutations and *Ret* disruption in mice[63] have been associated with Hirschsprung disease, renal agenesis[64] and central hypoventilation syndrome[65]. Similarly *BRAF*, which encodes a STK that regulates the ERK signalling pathway[66], has diverse physiological consequences based on the mutation type. Somatic gain-of-function *BRAF* mutations are associated with non-Hodgkin lymphoma, colorectal cancer, malignant melanoma, thyroid carcinoma and lung carcinoma[67]. By contrast, germline loss-of-function mutations in *BRAF* result in cardio–facio–cutaneous syndrome (CFC syndrome) syndrome in humans[68], and *Braf*-null mice die mid-gestation, indicating a key role for BRAF in the regulation of programmed cell death[69].

Genetic pleiotropy. Several kinase genes show pleiotropy, in which several phenotypes result from mutations in the same gene (FIG. 2). For example, mutations in the insulin receptor (*INSR*) cause either insulin-resistant diabetes mellitus with acanthosis nigricans or Donohue syndrome (also called leprechaunism) (FIG. 2a). These phenotypes indicate the importance of INSR in growth and metabolism through its ability to activate the p21–Ras pathway, which in turn controls protein synthesis, glycogenesis, lipogenesis and apoptosis[70].

Similarly, the FGFR genes show pleiotropy (FIG. 2). Gain-of-function mutations in the region of the immunoglobulin domain of *FGFR1* result in Pfeiffer syndrome and osteoglophonic dysplasia[71], and loss-of-function mutations throughout the protein cause hypogonadotrophic hypogonadism[40]. Mutations in the region of the immunoglobulin domain of *FGFR2* cause Crouzon[72], Pfeiffer[73] or Apert syndromes[74]. Mutations throughout *FGFR3* (REF. 18) cause syndromes that range from achondroplasia[75] and hypochondroplasia with acanthosis nigricans[76] to thanatophoric dysplasias[77], and Meunke syndrome (REF. 78) (FIG. 2a), which is characterized by skull and skeletal dysplasia, is found with *FGFR3* mutations in the region of the immunoglobulin domain. Although the complete explanation for such pleiotropy among FGFR mutations remains elusive, it is likely that in each case a balance of factors is involved, such as the position of the mutation, its effect on catalytic or non-catalytic function and the absolute change in kinase activity that is imparted by the mutation, all in the context of the genetic background of the individual, the stage of cellular development and the anatomical site of the affected tissue(s)[79].

The location of a mutation in a kinase gene may also determine the severity of the syndrome[80] (see the 'Insights from structural studies' section below). For example, mutations in *PRKAR1A* underlie a range of diseases depending on the location of the mutation (FIG. 2b). PRKAR1A is a widely expressed STK that binds to cAMP and regulates catalytic function in heart, endocrine tissue, skin and neurons. Intronic mutations in *PRKAR1A* that cause aberrant splicing are associated with a less severe phenotype, primary pigmented nodular adrenocortical disease[81]. Conversely, mutations in conserved domains of PRKAR1A, such as the nucleotide-binding

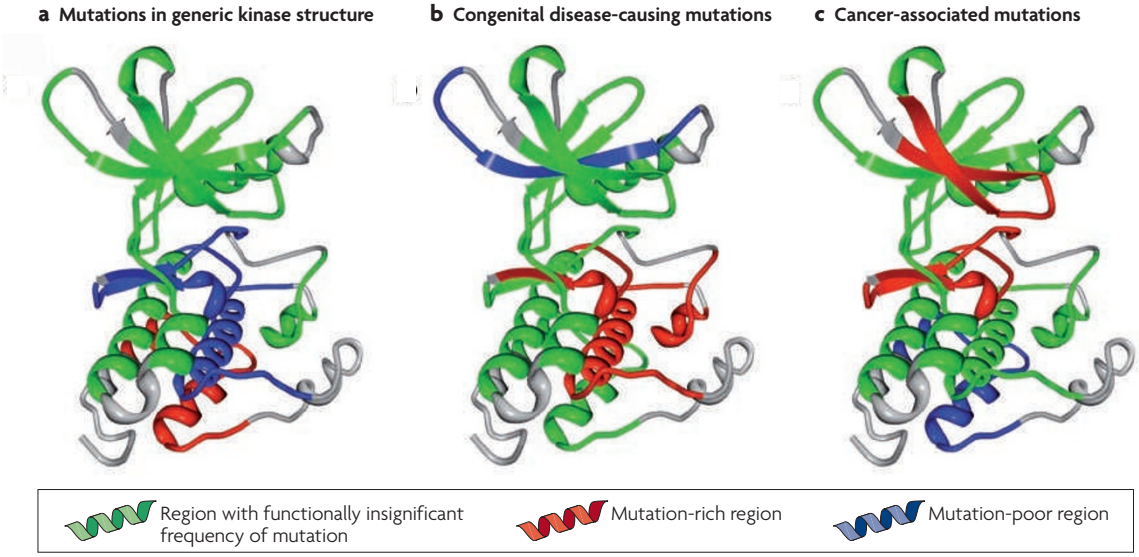


Figure 3 | **Structural distribution of kinase mutations. a** | The structural distribution of common mutations depicted in a generic kinase structure. **b** | The structural distribution of congenital disease-causing mutations. **c** | The structural distribution of cancer-associated mutations. Green represents regions that have a mutation frequency equivalent to what would be expected at random, blue represents regions statistically devoid of mutations and red depicts regions statistically enriched in mutations. The 5% significance level was determined by the general binomial distribution based on data obtained from general mutation and genetic variation databases such as dbSNP, OMIM, HGVBase and COSMIC[47,48,50,88,89,93]. Note that regions that are either enriched or devoid of mutations differ across the three mutation types and show minimal overlap.

site or the dimerization and phosphorylation regions, lead to Carney complex, a severe multi-organ tumour disorder[44]. Also, as noted above, pleiotropy has been observed for *IKBKG* mutations, and X-linked hypohidrotic ectodermal dysplasia with immune deficiency and incontinentia pigmenti[38] result from mutations in the *IKBKG* coding sequence. However, X-linked hypohidrotic ectodermal dysplasia with immune deficiency is milder and results from mutations in the coiled-coil domains, and incontinentia pigmenti is more severe and results from mutations that affect protein–protein interactions.

Locus heterogeneity. The same disease phenotype can sometimes result from mutations in different kinase genes, a phenomenon known as locus heterogeneity. This is observed for CFC syndrome, caused mainly by gain-of-function mutations in one of three STK genes — *BRAF*, *MEK1* and *MEK2* — which all belong to the RAS–MAPK pathway (REFS 68,82,83) (BOX 2). The mutations in these genes show that activation of the RAS–MAPK pathway, regardless of the specific mutation, results in the same developmental phenotype, and that mutation in any one of these three genes cannot be rescued by the preserved function of the remaining two STK genes[84].

Mutations in related proteins can also result in similar phenotypes, as illustrated by the FGFR family. *FGFR2* and *FGFR3* mutations can each underlie Crouzon syndrome[72,85]. In general, mutations in *FGFR1* and *FGFR2* cause most syndromes involving craniosynostosis, and the dwarfing syndromes are mainly associated with *FGFR3* mutations. Osteoglophonic dysplasia is an 'overlap' disorder with skeletal features that are seen with *FGFR1*, *FGFR2* and *FGFR3* mutations. It results from missense mutations in *FGFR1* that alter highly conserved residues in the ligand-binding and transmembrane domains, indicating that this receptor is a negative regulator of long-bone growth[86].

Insights from structural analysis. Given that kinases are partly characterized by a unique array of protein structural elements, it is possible to equate some kinase functions with particular structural features (FIG. 3). This is clear from an assessment of the kinase protein structural features that are perturbed in different disease states: common neutral mutations (FIG. 3a), inherited (that is, congenital) germline disease-causing mutations (FIG. 3b) and acquired somatic cell mutations that ultimately contribute to tumorigenesis (FIG. 3c) occupy or affect different structural elements across the protein kinase catalytic core.

In general, common neutral mutations tend to occupy the C-terminal regions of the catalytic core and substrate-binding or catalytic residues are avoided. The C-terminal region is thought to have a basic structural role; therefore, the amino acids in this region may not be as important for function as in other regions of the catalytic core. By contrast, inherited germline disease-causing mutations, most of which result in loss-of-function developmental and/or metabolic disorders, tend to cluster in regions of the catalytic core involved in regulation and substrate binding, especially residues that participate in protein–protein and allosteric interactions. These germline disease mutations only rarely occupy

Common neutral mutation
A non-synonymous SNP present in at least 1% of the human population that is either overtly neutral or not known to influence disease in appreciable ways.

Allosteric interaction
In an enzyme with at least two binding sites (an active site and another binding site that binds an allosteric effector), the binding of an allosteric effector alters the structure of the enzyme and increases or decreases catalytic activity.

regions involved in ATP binding or catalysis. A possible explanation for this observation is simple organism viability: perturbations in ATP binding or catalysis may act as complete loss-of-function mutations and cause embryonic lethality rather than a disease phenotype[87]. From TABLE 2, knockouts of many protein kinases result in embryonic or perinatal lethality. Therefore, it seems that at least some residual function of the protein kinase must be maintained to allow viability, although severe biological deficits may result. This requirement for residual function may also explain the disease pleiotropy observed for many protein kinases (at least in the cases in which pleiotropy is not explained by kinase activation versus loss-of-function mutations): the amount of residual protein function may explain the severity of the disease phenotype.

By contrast, acquired somatic mutations that cause or contribute to cancer tend to populate ATP binding and catalytic residues. Cancer-causing somatic mutations can either activate oncogenes or deactivate tumour suppressors, as discussed briefly above. Activating mutations increase the catalytic activity of the protein kinase, and therefore tend to alter residues involved in the regulation of catalysis. Conversely, deactivating mutations — at least in the context of cancer — may not require residual catalytic function, as is the case for inherited germline disease mutations. Therefore, the direct and complete inactivation of catalytic functions may be one hallmark feature of cancer-causing acquired somatic mutations.

Inherited mutational hotspots. In addition to a more general differential pattern in the kinase protein structural features associated with inherited and acquired somatic mutations, hotspots in germline disease mutations and somatic cancer mutations have been observed in protein kinases[87,88]. By assessing the frequency at which disease-associated mutations are found in different kinase structural features, relationships between specific kinase perturbations and diseases can be obtained. We and others have characterized specific positions in kinases that seem to harbour mutations across several diseases and cancers[88,89]; many of these mutations presumably lead to activation of the kinase (FIG. 3), and as such may be used as targets for future pharmaceutical intervention[90]. Many of these disease-associated kinase mutational hotspots have also been studied in statistical models of oncogenesis[91].

Inherited disease mutation hotspots tend to occupy residues of the catalytic core that are specific to eukaryotic protein kinases (EPKs) and are not observed in eukaryotic-like kinases (ELKs), which are prokaryotic small-molecule kinases that fold into the same general structure as EPKs. For example, the third glycine of the G-loop (G55), the histidine of the HxN motif (H100), and the putative regulatory molecule docking sites K92 and F108 (the residue numbers correspond to protein kinase A (PKA) residues), which cap the αC–β4 region, have been shown (G55, K92 and H100) or are likely (F108) to be key players in movements of the C-helix from the inactive to active conformation in EPKs (REF. 87) (FIG. 1). By contrast, the C-helix is held in a constitutively active conformation in ELKs[92]. Most disease hotspot residues are involved in the side-chain network formed by the APE motif, W222 and R280, which is a unique feature of EPKs[93]. Distantly related ELKs in prokaryotes that phosphorylate small metabolites lack these residues[94], suggesting a role for the EPK-specific network in substrate-binding function and allosteric regulation. Consistent with this notion, mutation of the APE glutamate to lysine in integrin-linked kinase may reduce substrate affinity or, alternatively, may reduce affinity for the associated kinase that is responsible for substrate phosphorylation[95]. Likewise, mutation of the arginine of subdomain XII in yeast PKA was shown to affect the binding and release of protein substrates[96].

Furthermore, inherited disease hotspots tend not to involve ultra-conserved residues in both EPKs and ELKs. These residues are likely to be key to kinase catalytic activities, and validate the previous observation that catalytic activity is somewhat preserved in disease states. Rather, the kinase structural hotspots associated with inherited, congenital diseases suggest a role for the lineage-specific variations that underlie certain biological functions in disease. This lineage-specific regulation is a secondary level of regulatory complexity layered on top of the more ancient catalytic machinery.

Acquired mutational hotspots. Acquired somatic cancer-causing kinase mutations also cluster into structural hotspots. These hotspots include the 'gatekeeper' residue, residues that are C-terminal to the DFG motif and residues adjacent to the nucleotide-binding pocket. These residues are key players in ATP binding, catalytic regulation and, in some cases, binding of ATP-mimetic drugs. These hotspots tend to be kinase-activating mutations as a much wider range of mutations could knockout kinase activity by destabilizing the structure of the protein, whereas increasing catalytic activity requires more precise fine tuning of the kinase regulatory machinery.

In addition to patterns in the specific kinase residues in the catalytic core that are perturbed in disease-causing states, there are mutational trends involving accessory domains and specific types of amino acid substitutions that probably reflect more subtle molecular physiological effects. For example, disease-causing mutations in kinase accessory domains that include regulatory domains, such as nuclear localization signals and SH2-binding domains, are due to point mutations in the amino acid sequence or to splicing aberrations in the kinase-encoding transcript. Overall, our observations suggest that transitions from arginine, cysteine, glycine, proline and serine were most commonly seen to underlie disease. Furthermore, transitions to arginine, proline, cysteine, histidine and serine were most commonly seen to underlie disease, which suggests a further level of complexity by which kinase mutations elicit pathogenic effects[4,87]. After grouping amino acids according to the polarity and charge of their side chain, the most frequent amino acid transitions occur in the non-polar group, followed by transitions in the polar group in which basic-polar amino acids are frequently substituted to neutral-polar amino acids. Other prevalent intergroup transitions include basic-polar

Ultra-conserved residue
An amino acid in a protein that has virtually 100% sequence identity across many species spanning hundreds of millions of years of evolution, suggesting that it has some essential role(s) in ontogeny and development.

'Gatekeeper' residue
A residue in the ATP-binding site of a protein kinase that controls the access of ATP or ATP-mimetic inhibitors to the binding pocket.

to neutral-non-polar and neutral-polar to non-polar amino acids. A previously generated prediction model for disease-causing amino acid transitions[4] had similar results to our pattern of kinasopathy-based amino acid transitions.

Overall, our observations as well as those of others[4,97] show that mutations involving amino acids with specific structural functions in proteins — such as salt bridges for cysteine, turns for proline or structural flexibility for glycine — are more likely to cause disease.

Kinase genes and genome-wide association studies. An exciting new phase for kinase gene involvement in human diseases has emerged from recent genome-wide association (GWA) studies, which have shown association signals of SNPs at loci that contain genes encoding kinases and related proteins with numerous complex and common disease phenotypes. These associations include MAP2K5 and restless legs syndrome[98], glucokinase regulatory protein (GCKR) and both plasma triglycerides and type 2 diabetes[99], PXK-domain-containing serine–threonine kinase and systemic lupus erythematosis[100], both FGFR2 and MAP3K1 and breast cancer[101], FGFR2 and schizophrenia[102], choline kinase-β (CHKB) and narcolepsy[103], STK38 with hypertension[104], TYK2 and multiple sclerosis[105], and CDK inhibitor 2A (CDKN2A) and CDKN2B with coronary artery disease[106] and type 2 diabetes[107]. Additional associations will doubtless be detected as GWA studies expand to evaluate more phenotypes and larger, more diverse samples. Although it is too soon to predict the clinical relevance of such findings, it seems likely that some of the associations will help identify new mechanisms underlying the disease traits, which in turn will help to inform biological understanding and the development of new interventions.

Conclusions and future directions

Human mutations in members of the kinase gene family underlie a broad range of disease phenotypes. The tissue and organ specificity seen with specific kinase mutations can provide clues into organogenesis or system development in early embryogenesis and can suggest situations in which there is insufficient redundancy in expression or function among related proteins to rescue a detrimental phenotype. The identification of key mutational hotspots in amino acid and protein domains as well as the range of distinct phenotypes in a single kinase gene brings clinical insight into various germline disorders and cancers. Human kinase mutations and associated phenotypes can complement experiments from other model systems and can provide support for the physiological role of particular kinases. Additional mapping of disease genes and the characterization of the amino acids affected are required to fully understand kinase structure, function and their role in disease. GWA studies herald a new wave of discoveries by expanding the involvement of the germline genetic variation that affects kinases involved in diseases ranging from rare syndromes to cancer and other complex diseases that affect adults.

1. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912–1934 (2002). **This landmark article established the size, range and implications of the protein kinase family in the human genome.**
2. Hubbard, M. J. & Cohen, P. On target with a new mechanism for the regulation of protein phosphorylation. *Trends Biochem. Sci.* **18**, 172–177 (1993).
3. Hopkins, A. L. & Groom, C. R. The druggable genome. *Nature Rev. Drug Discov.* **1**, 727–730 (2002).
4. Torkamani, A. & Schork, N. J. Distribution analysis of nonsynonymous polymorphisms within the human kinase gene family. *Genomics* **90**, 49–58 (2007).
5. Hanks, S. K. & Hunter, T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576–596 (1995).
6. Manning, G., Plowman, G. D., Hunter, T. & Sudarsanam, S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem. Sci.* **27**, 514–520 (2002).
7. Lopez-Bigas, N. & Ouzounis, C. A. Genome-wide identification of genes likely to be involved in human genetic disease. *Nucleic Acids Res.* **32**, 3108–3114 (2004).
8. Hegele, R. A. Phenomics, lipodystrophy, and the metabolic syndrome. *Trends Cardiovasc. Med.* **14**, 133–137 (2004).
9. Schork, N. J. Genetics of complex disease: approaches, problems, and solutions. *Am. J. Respir. Crit. Care Med.* **156**, S103–109 (1997).
10. Hegele, R. A. Phenomics, lamin A/C, and metabolic disease. *J. Clin. Endocrinol. Metab.* **92**, 4566–4568 (2007).
11. Tartaglia, M. *et al.* Mutations in *PTPN11*, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nature Genet.* **29**, 465–468 (2001).
12. Narkis, G. *et al.* Lethal congenital contractural syndrome type 2 (LCCS2) is caused by a mutation in ERBB3 (Her3), a modulator of the phosphatidylinositol-3-kinase/Akt pathway. *Am. J. Hum. Genet.* **81**, 589–595 (2007).
13. Chen, D. H. *et al.* Missense mutations in the regulatory domain of PKCγ: a new mechanism for dominant nonepisodic cerebellar ataxia. *Am. J. Hum. Genet.* **72**, 839–849 (2003).
14. Coussens, L. *et al.* Multiple, distinct forms of bovine and human protein kinase C suggest diversity in cellular signaling pathways. *Science* **233**, 859–866 (1986).
15. Aronowski, J. & Grotta, J. C. Ca^{2+}/calmodulin-dependent protein kinase II in postsynaptic densities after reversible cerebral ischemia in rats. *Brain Res.* **709**, 103–110 (1996).
16. Aronowski, J., Grotta, J. C., Strong, R. & Waxham, M. N. Interplay between the gamma isoform of PKC and calcineurin in regulation of vulnerability to focal cerebral ischemia. *J. Cereb. Blood Flow Metab.* **20**, 343–349 (2000).
17. Stenirri, S. *et al.* Integrated strategy for fast and automated molecular characterization of genes involved in craniosynostosis. *Clin. Chem.* **53**, 1767–1774 (2007).
18. Chen, L. & Deng, C. X. Roles of FGF signaling in skeletal development and human genetic diseases. *Front. Biosci.* **10**, 1961–1976 (2005).
19. Shore, E. M. *et al.* A recurrent mutation in the BMP type I receptor ACVR1 causes inherited and sporadic fibrodysplasia ossificans progressiva. *Nature Genet.* **38**, 525–527 (2006).
20. Yu, P. B. *et al.* BMP type I receptor inhibition reduces heterotopic ossification. *Nature Med.* **14**, 1363–1369 (2008).
21. Greenman, C. *et al.* Patterns of somatic mutation in human cancer genomes. *Nature* **446**, 153–158 (2007). **This report provided the first in-depth exploration of the frequency of somatic mutation in protein kinases in humans.**
22. Maru, Y. & Witte, O. N. The *BCR* gene encodes a novel serine/threonine kinase activity within a single exon. *Cell* **67**, 459–468 (1991).
23. Gorre, M. E. *et al.* Clinical resistance to STI-571 cancer therapy caused by BCR–ABL gene mutation or amplification. *Science* **293**, 876–880 (2001).
24. Demiroglu, A. *et al.* The t(8;22) in chronic myeloid leukemia fuses BCR to FGFR1: transforming activity and specific inhibition of FGFR1 fusion proteins. *Blood* **98**, 3778–3783 (2001).
25. Peeters, P. *et al.* Fusion of *TEL*, the ETS-variant gene 6 (*ETV6*), to the receptor-associated kinase *JAK2* as a result of t(9;12) in a lymphoid and t(9;15;12) in a myeloid leukemia. *Blood* **90**, 2535–2540 (1997).
26. Lacronique, V. *et al.* A TEL–JAK2 fusion protein with constitutive kinase activity in human leukemia. *Science* **278**, 1309–1312 (1997).
27. Reiter, A. *et al.* The t(8;9)(p22;p24) is a recurrent abnormality in chronic and acute leukemia that fuses *PCM1* to *JAK2*. *Cancer Res.* **65**, 2662–2667 (2005).
28. Baxter, E. J. *et al.* Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet* **365**, 1054–1061 (2005).
29. Tefferi, A. *et al. JAK2* germline genetic variation affects disease susceptibility in primary myelofibrosis regardless of V617F mutational status: nullizygosity for the *JAK2* 46/1 haplotype is associated with inferior survival. *Leukemia* 22 Oct 2009 (doi: 10.1038/leu.2009.225).
30. Sankelo, M. *et al.* BMPR2 mutations have short lifetime expectancy in primary pulmonary hypertension. *Hum. Mutat.* **26**, 119–124 (2005).
31. Wong, W. K., Knowles, J. A. & Morse, J. H. Bone morphogenetic protein receptor type II C-terminus interacts with c-Src: implication for a role in pulmonary arterial hypertension. *Am. J. Respir. Cell. Mol. Biol.* **33**, 438–446 (2005).
32. Beppu, H. *et al.* BMP type II receptor is required for gastrulation and early development of mouse embryos. *Dev. Biol.* **221**, 249–258 (2000).
33. Pouliot, F., Blais, A. & Labrie, C. Overexpression of a dominant negative type II bone morphogenetic protein receptor inhibits the growth of human breast cancer cells. *Cancer Res.* **63**, 277–281 (2003).
34. Elder, M. in *Primary Immunodeficiency Diseases: A Molecular and Genetic Approach* (eds Ochs, H. D., Edvard Smith, C. I. & Puck, J.) 203–211 (Oxford University Press, New York, 2006).

35. Vorechovsky, I. *et al.* Molecular diagnosis of X-linked agammaglobulinaemia. *Lancet* **341**, 1153 (1993).
36. Cheng, G., Ye, Z. S. & Baltimore, D. Binding of Bruton's tyrosine kinase to Fyn, Lyn, or Hck through a Src homology 3 domain-mediated interaction. *Proc. Natl Acad. Sci. USA* **91**, 8152–8155 (1994).
37. Thomas, J. D. *et al.* Colocalization of X-linked agammaglobulinemia and X-linked immunodeficiency genes. *Science* **261**, 355–358 (1993).
38. Zonana, J. *et al.* A novel X-linked disorder of immune deficiency and hypohidrotic ectodermal dysplasia is allelic to incontinentia pigmenti and due to mutations in *IKK-gamma* (*NEMO*). *Am. J. Hum. Genet.* **67**, 1555–1562 (2000).
39. Rudolph, D. *et al.* Severe liver degeneration and lack of *NF-κB* activation in *NEMO/IKKγ*-deficient mice. *Genes Dev.* **14**, 854–862 (2000).
40. Pitteloud, N. *et al.* Mutations in fibroblast growth factor receptor 1 cause both Kallmann syndrome and normosmic idiopathic hypogonadotropic hypogonadism. *Proc. Natl Acad. Sci. USA* **103**, 6281–6286 (2006).
41. George, S. *et al.* A family with severe insulin resistance and diabetes due to a mutation in *AKT2*. *Science* **304**, 1325–1328 (2004).
42. Semple, R. K. *et al.* Postreceptor insulin resistance contributes to human dyslipidemia and hepatic steatosis. *J. Clin. Invest.* **119**, 315–322 (2009).
43. Cho, H. *et al.* Insulin resistance and a diabetes mellitus-like syndrome in mice lacking the protein kinase Akt2 (PKBβ). *Science* **292**, 1728–1731 (2001).
44. Kirschner, L. S. *et al.* Mutations of the gene encoding the protein kinase A type I-α regulatory subunit in patients with the Carney complex. *Nature Genet.* **26**, 89–92 (2000).
45. Kesler, S. R. *et al.* Altered neurodevelopment associated with mutations of *RSK2*: a morphometric MRI study of Coffin–Lowry syndrome. *Neurogenetics* **8**, 143–147 (2007).
46. Lahiry, P. *et al.* A multiplex human syndrome implicates a key role for intestinal cell kinase in development of central nervous, skeletal, and endocrine systems. *Am. J. Hum. Genet.* **84**, 134–147 (2009).
47. Futreal, P. A. *et al.* A census of human cancer genes. *Nature Rev. Cancer* **4**, 177–183 (2004).
48. Forbes, S. A. *et al.* The Catalogue of Somatic Mutations in Cancer (COSMIC). *Curr. Protoc. Hum. Genet.* Chapter 10, Unit 10.11 (2008).
49. Ingvarsson, S. *et al.* Mutation analysis of the *CHK2* gene in breast carcinoma and other cancers. *Breast Cancer Res.* **4**, R4 (2002).
50. Mosse, Y. P. *et al.* Identification of *ALK* as a major familial neuroblastoma predisposition gene. *Nature* **455**, 930–935 (2008).
51. Chen, Y. *et al.* Oncogenic mutations of ALK kinase in neuroblastoma. *Nature* **455**, 971–974 (2008).
52. George, R. E. *et al.* Activating mutations in *ALK* provide a therapeutic target in neuroblastoma. *Nature* **455**, 975–978 (2008).
53. Janoueix-Lerosey, I. *et al.* Somatic and germline activating mutations of the ALK kinase receptor in neuroblastoma. *Nature* **455**, 967–970 (2008).
54. Reilly, J. T. FLT3 and its role in the pathogenesis of acute myeloid leukaemia. *Leuk. Lymphoma* **44**, 1–7 (2003).
55. Petzer, A. L., Hogge, D. E., Landsdorp, P. M., Reid, D. S. & Eaves, C. J. Self-renewal of primitive human hematopoietic cells (long-term-culture-initiating cells) *in vitro* and their expansion in defined medium. *Proc. Natl Acad. Sci. USA* **93**, 1470–1474 (1996).
56. Wright, T. M. *et al.* Ror2, a developmentally regulated kinase, promotes tumor growth potential in renal cell carcinoma. *Oncogene* **28**, 2513–2523 (2009).
57. Takeuchi, S. *et al.* Mouse Ror2 receptor tyrosine kinase is required for the heart development and limb formation. *Genes Cells* **5**, 71–78 (2000).
58. Hegele, R. *LMNA* mutation position predicts organ system involvement in laminopathies. *Clin. Genet.* **68**, 31–34 (2005).
59. Plaza-Menacho, I., Burzynski, G. M., de Groot, J. W., Eggen, B. J. & Hofstra, R. M. Current concepts in RET-related genetics, signaling and therapeutics. *Trends Genet.* **22**, 627–636 (2006).
60. Mulligan, L. M. *et al.* Germ-line mutations of the *RET* proto-oncogene in multiple endocrine neoplasia type 2A. *Nature* **363**, 458–460 (1993).
61. Xue, F. *et al.* Germline RET mutations in MEN 2A and FMTC and their detection by simple DNA diagnostic tests. *Hum. Mol. Genet.* **3**, 635–638 (1994).
62. Eng., C. *et al.* Mutations in the *RET* proto-oncogene and the von Hippel–Lindau disease tumour suppressor gene in sporadic and syndromic phaeochromocytomas. *J. Med. Genet.* **32**, 934–937 (1995).
63. Arighi, E., Borrello, M. G. & Sariola, H. RET tyrosine kinase signaling in development and cancer. *Cytokine Growth Factor Rev.* **16**, 441–467 (2005).
64. Skinner, M. A., Safford, S. D., Reeves, J. G., Jackson, M. E. & Freemerman, A. J. Renal aplasia in humans is associated with *RET* mutations. *Am. J. Hum. Genet.* **82**, 344–351 (2008).
65. Amiel, J. *et al.* Mutations of the RET–GDNF signaling pathway in Ondine's curse. *Am. J. Hum. Genet.* **62**, 715–717 (1998).
66. Solit, D. B. *et al. BRAF* mutation predicts sensitivity to MEK inhibition. *Nature* **439**, 358–362 (2006).
67. Wan, P. T. *et al.* Mechanism of activation of the RAF–ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* **116**, 855–867 (2004).
An example of the subtle way in which somatic mutations in cancer alter the catalytic activity of a protein kinase.
68. Rodriguez-Viciana, P. *et al.* Germline mutations in genes within the MAPK pathway cause cardio–facio–cutaneous syndrome. *Science* **311**, 1287–1290 (2006).
69. Wojnowski, L. *et al.* Endothelial apoptosis in Braf-deficient mice. *Nature Genet.* **16**, 293–297 (1997).
70. Dudek, H. *et al.* Regulation of neuronal survival by the serine–threonine protein kinase Akt. *Science* **275**, 661–665 (1997).
71. Muenke, M. *et al.* A common mutation in the fibroblast growth factor receptor 1 gene in Pfeiffer syndrome. *Nature Genet.* **8**, 269–274 (1994).
72. Reardon, W. *et al.* Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. *Nature Genet.* **8**, 98–103 (1994).
73. Rutland, P. *et al.* Identical mutations in the *FGFR2* gene cause both Pfeiffer and Crouzon syndrome phenotypes. *Nature Genet.* **9**, 173–176 (1995).
74. Wilkie, A. O. *et al.* Apert syndrome results from localized mutations of *FGFR2* and is allelic with Crouzon syndrome. *Nature Genet.* **9**, 165–172 (1995).
75. Shiang, R. *et al.* Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. *Cell* **78**, 335–342 (1994).
76. Alatzoglou, K. S., Hindmarsh, P. C., Brain, C., Torpiano, J. & Dattani, M. T. Acanthosis nigricans and insulin sensitivity in patients with achondroplasia and hypochodroplasia due to *FGFR3* mutations. *J. Clin. Endocrinol. Metab.* **94**, 3959–3963 (2009).
77. Tavormina, P. L. *et al.* Thanatophoric dysplasia (types I and II) caused by distinct mutations in fibroblast growth factor receptor 3. *Nature Genet.* **9**, 321–328 (1995).
78. Bellus, G. A. *et al.* Identical mutations in three different fibroblast growth factor receptor genes in autosomal dominant craniosynostosis syndromes. *Nature Genet.* **14**, 174–176 (1996).
79. Wilkie, A. O. Craniosynostosis: genes and mechanisms. *Hum. Mol. Genet.* **6**, 1647–1656 (1997).
80. Ortutay, C., Valiaho, J., Stenberg, K. & Vihinen, M. KinMutBase: a registry of disease-causing mutations in protein kinase domains. *Hum. Mutat.* **25**, 435–442 (2005).
This paper describes KinMutBase, a comprehensive database for human disease-related mutations in protein kinase domains.
81. Groussin, L. *et al.* Mutations of the *PRKAR1A* gene in Cushing's syndrome due to sporadic primary pigmented nodular adrenocortical disease. *J. Clin. Endocrinol. Metab.* **87**, 4324–4329 (2002).
82. Denayer, E. & Legius, E. What's new in the neuro–cardio–facial–cutaneous syndromes? *Eur. J. Pediatr.* **166**, 1091–1098 (2007).
83. Niihori, T. *et al.* Germline *KRAS* and *BRAF* mutations in cardio–facio–cutaneous syndrome. *Nature Genet.* **38**, 294–296 (2006).
84. Seger, R. & Krebs, E. G.The MAPK signaling cascade. *FASEB J.* **9**, 726–735 (1995).
85. Meyers, G. A., Orlow, S. J., Munro, I. R., Przylepa, K. A. & Jabs, E. W. Fibroblast growth factor receptor 3 (*FGFR3*) transmembrane mutation in Crouzon syndrome with acanthosis nigricans. *Nature Genet.* **11**, 462–464 (1995).
86. White, K. E. *et al.* Mutations that cause osteoglophonic dysplasia define novel roles for FGFR1 in bone elongation. *Am. J. Hum. Genet.* **76**, 361–367 (2005).
87. Torkamani, A., Kannan, N., Taylor, S. S. & Schork, N. J. Congenital disease SNPs target lineage specific structural elements in protein kinases. *Proc. Natl Acad. Sci. USA* **105**, 9011–9016 (2008).
This paper provides the basis for an analysis of the structural distribution of disease-causing and common genomic variants in protein kinases.
88. Torkamani, A., Verkhivker, G. & Schork, N. J. Cancer driver mutations in protein kinase genes. *Cancer Lett.* **281**, 117–127 (2009).
89. Dixit, A., Torkamani, A., Schork, N. J. & Verkhivker, G. Computational modeling of structurally conserved cancer mutations in the RET and MET kinases: the impact on protein structure, dynamics, and stability. *Biophys. J.* **96**, 858–874 (2009).
90. Garber, K. The second wave in kinase cancer drugs. *Nature Biotechnol.* **24**, 127–130 (2006).
91. Torkamani, A. & Schork, N. J. Prediction of cancer driver mutations in protein kinases. *Cancer Res.* **68**, 1675–1682 (2008).
This paper provides the basis for an analysis of the structural distribution of somatic mutations in protein kinases.
92. Kannan, N., Haste, N., Taylor, S. S. & Neuwald, A. F. The hallmark of AGC kinase functional divergence is its C-terminal tail, a *cis*-acting regulatory module. *Proc. Natl Acad. Sci. USA* **104**, 1272–1277 (2007).
93. Kannan, N., Taylor, S. S., Zhai, Y., Venter, J. C. & Manning, G. Structural and functional diversity of the microbial kinome. *PLoS Biol.* **5**, e17 (2007).
This paper is a survey of the microbial kinome and compares the structural elements of eukaryotic-like kinases with eukaryotic protein kinases
94. Kannan, N. & Neuwald, A. F. Did protein kinase regulatory mechanisms evolve through elaboration of a simple structural component? *J. Mol. Biol.* **351**, 956–972 (2005).
This paper describes some of the mechanistic details of protein kinase activity and their evolution from more primitive eukaryotic-like kinases.
95. Lynch, D. K., Ellis, C. A., Edwards, P. A. & Hiles, I. D. Integrin-linked kinase regulates phosphorylation of serine 473 of protein kinase B by an indirect mechanism. *Oncogene* **18**, 8024–8032 (1999).
96. Deminoff, S. J., Howard, S. C., Hester, A., Warner, S. & Herman, P. K. Using substrate-binding variants of the cAMP-dependent protein kinase to identify novel targets and a kinase domain important for substrate interactions in *Saccharomyces cerevisiae*. *Genetics* **173**, 1909–1917 (2006).
97. Vitkup, D., Sander, C. & Church, G. M. The amino-acid mutational spectrum of human genetic disease. *Genome Biol.* **4**, R72 (2003).
98. Winkelmann, J. *et al.* Genome-wide association study of restless legs syndrome identifies common variants in three genomic regions. *Nature Genet.* **39**, 1000–1006 (2007).
99. Sparso, T. *et al.* The *GCKR* rs780094 polymorphism is associated with elevated fasting serum triacylglycerol, reduced fasting and OGTT-related insulinaemia, and reduced risk of type 2 diabetes. *Diabetologia* **51**, 70–75 (2008).
100. Harley, J. B. *et al.* Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in *ITGAM*, *PXK*, *KIAA1542* and other loci. *Nature Genet.* **40**, 204–210 (2008).
101. Easton, D. F. *et al.* Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* **447**, 1087–1093 (2007).
This paper reports one of the first three-stage genome-wide association studies in breast cancer that identified six highly significantly associated SNPs, with the most notable gene being *FGFR2*.
102. O'Donovan, M. C. *et al.* Analysis of 10 independent samples provides evidence for association between schizophrenia and a SNP flanking fibroblast growth factor receptor 2. *Mol. Psychiatry* **14**, 30–36 (2009).
103. Miyagawa, T. *et al.* Variant between *CPT1B* and *CHKB* associated with susceptibility to narcolepsy. *Nature Genet.* **40**, 1324–1328 (2008).
104. Wang, Y. *et al.* Whole-genome association study identifies *STK39* as a hypertension susceptibility gene. *Proc. Natl Acad. Sci. USA* **106**, 226–231 (2009).
105. The Australia and New Zealand Multiple Sclerosis Genetics Consortium (ANZgene). Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nature Genet.* **41**, 824–828 (2009).
106. McPherson, R. *et al.* A common allele on chromosome 9 associated with coronary heart disease. *Science* **316**, 1488–1491 (2007).

107. Zeggini, E. *et al.* Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* **316**, 1336–1341 (2007).
108. Pawson, T. Regulation and targets of receptor tyrosine kinases. *Eur. J. Cancer* **38,** (Suppl. 5), 3–10 (2002). **A key overview of receptor tyrosine kinase biology and regulation.**
109. Gerthoffer, W. T. & Singer, C. A. MAPK regulation of gene expression in airway smooth muscle. *Respir. Physiol. Neurobiol.* **137**, 237–250 (2003).
110. Johnson, G. L. & Lapadat, R. Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases. *Science* **298**, 1911–1912 (2002).
111. Heinrich, P. C. *et al.* Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *Biochem. J.* **374**, 1–20 (2003).
112. Welch, P. J. & Wang, J. Y. A C-terminal protein-binding domain in the retinoblastoma protein regulates nuclear c-Abl tyrosine kinase in the cell cycle. *Cell* **75**, 779–790 (1993).
113. Edelman, A. M., Blumenthal, D. K. & Krebs, E. G. Protein serine/threonine kinases. *Annu. Rev. Biochem.* **56**, 567–613 (1987).
114. Capra, M. *et al.* Frequent alterations in the expression of serine/threonine kinases in human cancers. *Cancer Res.* **66**, 8147–8154 (2006).
115. Guicciardi, M. E. & Gores, G. J. AIP1: a new player in TNF signaling. *J. Clin. Invest.* **111**, 1813–1815 (2003).
116. Xu, B. E. *et al.* WNK1 activates ERK5 by an MEKK2/3-dependent mechanism. *J. Biol. Chem.* **279**, 7826–7831 (2004).
117. Anderson, C. W. & Lees-Miller, S. P. The nuclear serine/threonine protein kinase DNA-PK. *Crit. Rev. Eukaryot Gene Expr.* **2**, 283–314 (1992).
118. Cheung, E. C. *et al.* Dissociating the dual roles of apoptosis-inducing factor in maintaining mitochondrial structure and apoptosis. *EMBO J.* **25**, 4061–4073 (2006).
119. Desmots, F., Russell, H. R., Michel, D. & McKinnon, P. J. Scythe regulates apoptosis-inducing factor stability during endoplasmic reticulum stress-induced apoptosis. *J. Biol. Chem.* **283**, 3264–3271 (2008).
120. Fu, Z. *et al.* Identification of yin-yang regulators and a phosphorylation consensus for male germ cell-associated kinase (MAK)-related kinase. *Mol. Cell Biol.* **26**, 8639–8654 (2006).

Acknowledgements
P.L. is supported by the Canadian Institutes of Health Research (CIHR) Scriver Family M.D./Ph.D. studentship award and the Heart and Stroke Foundation of Ontario Vascular Training Program. R.A.H. holds the Edith Schulich Vinet Canada Research Chair (Tier I) in Human Genetics and the Jacob J. Wolfe Distinguished Medical Research Chair. This work was supported by operating grants from the Heart and Stroke Foundation of Ontario (NA 6,018), the Canadian Institutes for Health Research (MOP 13,430 and 79,533), the Jean Davignon Distinguished Cardiovascular-Metabolic Research Award (Pfizer, Canada) and Genome Canada through the Ontario Genomics Institute. N.J.S. and A.T. are supported in part by the Scripps Translational Science Institute Clinical Translational Science Award (NIH U54RR02504-01). A.T. is also supported in part by a Scripps Dickinson Fellowship. We thank Natarajan Kannan for providing an earlier version of Figure 1.

Competing interests statement
The authors declare no competing financial interests.

DATABASES
Entrez Gene: http://www.ncbi.nlm.nih.gov/gene
ALK | *BCR* | *BMPR2* | *BRAF* | *BTK* | *FGFR1* | *FGFR2* | *FGFR3* | *IKBKG* | *JAK2* | *PRKAR1A* | *PRKCG* | *RET*
OMIM: http://www.ncbi.nlm.nih.gov/omim
Carney syndrome | CFC syndrome | CML | Pfeiffer syndrome
UniProtKB: http://www.uniprot.org
ICK

FURTHER INFORMATION
Robert Hegele's homepage: http://www.robarts.ca/hegele
1000 Genomes: http://www.1000genomes.org
Catalogue of Somatic Mutations in Cancer (COSMIC): http://www.sanger.ac.uk/genetics/CGP/cosmic
Kinase.com: http://kinase.com
Kinase Pathway Database: http://kinasedb.ontology.ims.u-tokyo.ac.jp:8081
Kinase Sequence Database: http://sequoia.ucsf.edu/ksd/
KinMutBase: http://bioinf.uta.fi/KinMutBase
Kinweb Kinase Database: http://www.itb.cnr.it/kinweb **or** http://kinweb.ceinge.unina.it/
Multiple endocrine neoplasia type 2 (MEN2) *RET* database: http://www.arup.utah.edu/database/MEN2/MEN2_welcome.php
Protein Kinase Resource: http://www.nih.go.jp/mirror/Kinases
Protein Structure Initiative: http://kb.psi-structuralgenomics.org
The Cancer Genome Atlas: http://cancergenome.nih.gov

SUPPLEMENTARY INFORMATION
See online article: S1 (table) | S2 (table)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF