Gene Expression to Genetical Genomics

REVIEW

Is Glycogen Synthase Kinase-3β an Ultraconserved Kernel Enzyme?

M. Fenger[1], S.B. Haugaard[2], O. Andersen[3], S. Madsbad[4], T. Werge[5] and A. Linneberg[6]

[1]University of Copenhagen, Denmark and Department of Clinical Biochemistry, Capital Hospitals of Denmark at Hvidovre. [2]Department of Endocrinology and Clinical Research Centre, Capital Hospitals of Denmark at Hvidovre. [3]Clinical Research Centre, Capital Hospitals of Denmark at Hvidovre. [4]Department of Endocrinology Capital Hospitals of Denmark at Hvidovre. [5]Psychiatric Research Centre, Skt. Hans Hospital, Roskilde, Denmark. [6]Research Centre for Prevention and Health, Glostrup, Denmark. Email: mogens.fenger@hvh.regionh.dk

Abstract: Glycogen synthase kinase 3β (GSK3β) was discovered as a major factor in glucose homeostasis. GSK3 β is a pivotal regulator of glycogen synthesis by altering the activity of glycogen synthase and perturbation in this process have been linked to at least some sub-entities of insulin resistance and diabetes mellitus. However, GSK3β has a central role in many biochemical and physiological processes apparently not related to insulin resistance or diabetes. The functionality of GSK3 β is vast including more than 33 substrates and 72 protein-protein interactions, and is involved in such diverse diseases as cancer, schizophrenia, inflamation, cardiac diseases and many more. Considering this scale of the involvement of GSK3β, many related to developmental processes, suggests that either is GSK3β a high-level network-hub or is a peripheral kinase in several non-connected networks. Here, evidence is provided that GSK3 β is an ultraconserved protein which would suggest that the enzyme qualify to be a kernel protein in the sense that life as we know it depends on conserved molecular entities to provide the essential functionality to an organism. The enzyme has evolved under strong purifying selection implicating that GSK3β is a functional hub.

Keywords: glycogen synthase kinase 3β, insulin resistance, diabetes mellitus

Gene Expression to Genetical Genomics 2009:2 1–27

This article is available from http://www.la-press.com.

Introduction

The Cambrian explosion, which happened approximately 540 million years ago (Mya), refers to the vast expansion of living species. Although extensive diversions of species between and within the phyla since then are encountered, only minor changes have occurred in the grand body plans of even apparently remote species. It has been suggested that concomitant preservation of the basic body plan and the divergent evolution of phyla and species is reflected in the hierarchical structure of modular structures of multigenic regulatory networks.[1] On top of this hierarchy are the ultraconserved regulatory subcircuits which comprise the upstream blue print of the body plan and are as such considered the kernel of the organism. Beneath this level are several evolving regulatory subcircuits and modullary structures determining the development and diversion of intra-phyla species. In the bottom of the hierarchy are the effectors of biological processes i.e. proteins involved in metabolic processes, signal transduction, and as structural cellular components. The concept of kernel structures thus refers to a collection of genes encoding transcription factors dedicated to execute the blue print of the basic body plan. The individual transcription factors may participate in subsequent regulatory processes, but the essential point is that the kernels are functionally integrated structures of genes and any perturbation of a kernel is detrimental to the developmental purpose of the kernel resulting in the lack of development of a particular body part or even being lethal to the organism, and hence the kernel structure and participating genes are extremely conserved.

Genes evolves as a consequence of selection pressures, which is intimately (but not exclusively) related to the functionality of the genes products i.e. the proteins. Proteins are involved in all aspects of cellular processes: organisation of the genome, transcription, translation, organelle constructions, mechanic structures and cellular skeletons, transport of metabolites, as enzymes, and signal transduction and regulation. They are engaged in complex interactions in networks with other proteins and metabolites, and are central to all physiological processes and phenotypic expression of unicellular as well as multicellular organisms. Proteins may not be restricted to a single functionality, but may differ in their function in context-specific ways.

A prime example of this multifunctionality is glycogen synthase kinase 3 beta (GSK3β) which phosphorylates at least 33 substrates and interacts with at least 72 proteins participating in numerous metabolic processes, signal transductions, and cellular structures (collated in Human Protein Reference Database, http://www.hprd.org). GSK3β was baptised so because it was initially described as the key enzyme in glycogen metabolism—which it is indeed.[2] As such the functionality and regulation GSK3β is a major contributor to the homeostasis of glucose, but as indicated above the functionality of this kinase goes far beyond the regulation of glucose homeostasis. Considering the numerous and diverse processes that involves GSK3β two extreme scenarios can be imagined: 1) GSK3β is evolutionary diverse and rather non-conserved allowing for extensive intra-species variation, and is not a hub in any of the networks governing the physiological processes it is part of; or 2) GSK3β is a component (even as a hub) in networks of proteins and is pivotal to essential processes in diverse and phylogenetically distant species, and perturbations of GSK3β may have detrimental effects on cells and organisms, hence GSK3β is highly conserved through evolution. After a short review of regulatory aspects of GSK3β, we argue and provide evidence for the latter scenario.

Glycogen Synthase Kinase 3β

GSK3β is located on chromosome 3q13.3-q21 and is expressed in almost all tissues including muscle, brain, kidney, testes, thymus, prostate and ovary in humans.[3] GSK3β is usually constitutively activated by phosphorylation of Tyr216,[4–6] which may be the result of an autophosphorylation process.[7] The ubiquitous mechanism of inhibition of GSK3β is by phosphorylation of serine 9 (Ser9), but other mechanisms are operative. GSK3β is involved in a plethora of biochemical processes, which raises the question how specificity is obtained. This is obtained by several signalling pathways, compartmentalizing of GSK3β and its substrates, and by mechanisms not related to the phosphorylation status of GSK3β. The major pathways and mechansms are summarized below. Much direct evidence have been present, some conclusions are by inference, and most of all, we still have much to learn.

1. *Akt (protein kinase B, PKB) phosphorylation*, as in the insulin signalling cascade, includes the insulin receptor substrate 1 (IRS1), phosphatidylinositol-3-kinase(PI3K),and3-phosphoinositide-dependent kinase (PDK1), resulting in activation of Akt and subsequently phosphorylation of Ser9 in the cytoplasmatic pool of GSK3β. A characteristic of this pathway is that the substrates have to be primed by phophorylation by other kinases. In addition, insulin stimulates protein phosphatases which dephosphorylates and thereby activates substrates of GSK3β such as glycogen synthase and the eukaryotic initiation factor 2B, eIF2B[8] promoting glycogen and protein synthesis. Thus, GSK3β is an anti-anabolic kinase.

   Recently it has been proposed that spingosine-1-phosphate (S1P) induces a time-dependent activation of the PI3K/PDK1/Akt pathway stimulating a migratory response in vascular smooth muscle cells. The activation of the pathway induces the anticipated decrease in GSK3β.[9] This action of S1P links GSK3β to the intricate balance of the ceramide/S1P signalling system with the apoptotic effects of ceramide opposing the S1P stimulation of cell proliferation.[10] It is not known if there is a feed-back regulation of the ceramide/S1P pathway by GSK3β.
2. *The Wnt-pathway.* The lipoproteins Wnt ligands (Wingless in *Drosophila melanogaster*) suppress the activity of GSK3β, through the activation of the Frizzled receptors[8] and the co-receptors lipoprotein receptor-related proteins LRP5 and LRP6. More than 10 Wnts and receptors have been identified.[11] In the absence of Wnt-signalling GSK3β is complexed with axin, β-catenin and APC (adenomatous polyposis coli protein). In this complex GSK3β phosphorylate and stabilize axin and APC increasing the ability of GSK3β to phosphorylate β-catenin. Here, axin plays the critical role in insulating GSK3β from other GSK3β-related pathways: non-axin-associated GSK3β has almost no activity toward β-catenin,[12] and axin-associated GSK3β is not regulated by insulin. Phosphorylation of β-catenin in the complex promotes its ubiquitin-mediated degradation by the proteasome.[8] Upon binding of Wnts to their receptors Dvl (Dishevelled in *Drosophila*) FRAT1 is mobilized disrupting the GSK3β/axin/APC/β-catenin complex by recruiting axin to the Wnt receptors.[7,13] In this process non-phosphorylated β-catenin is released from the complex and translocates to the nucleus, where it binds to transcription factors of the TCF/LEF family of transcription factors.

   As in the insulin/Akt signalling pathway the reduced phosphorylation of axin by Wnt-stimulation may in part be mediated by action of protein phosphatase 2A (PP2A), which binds to axin. Also, PP2A may dephosphorylate APC in its cental region, decreasing the facilitation of β-catenin phosphorylation.[7] However, apart from the above described mechanism of regulation, the Wnt-pathway differs from the insulin pathway in that there is probably no prerequisite for priming phosphorylation of axin, β-catenin or APC by other kinases prior to phosphorylation by GSK3. Also, GSK3β is not phosphorylated in Ser9 by the Wnt-signalling. Rather, GSK3β is made inefficient in this pathway by simply disrupting the phosphorylation complex.
3. *Hedgehog pathway.* Hedgehog is a group of related, secreted proteins.[14] Unique to hedgehog is the covalent coupling of cholesterol to the C-terminal end of the mature protein and the palmitoylation in a N-terminal Cys by an acyl-transferase. The cholesterol moity seems not to influence signalling but is important for secretion, whereas palmitoylation potentiates the signal activity.[15]

   The hedgehog receptor (Patched) is a twelve-pass transmembrane receptor that acts as a key inhibitor of the constitutively active seven-pass transmembrane G-protein-coupled receptor Smoothened. Binding of hedgehog relieves the inhibition of Smoothened activating the translocation of the transcription activator Cubitus interruptus (Ci) in flies (Gli in the mammalians).[16] In the absence of signalling in this pathway Ci is in part truncated and functions as transcriptional repressor, and is in part targeted to degradation. The latter is accomplished by a sequential process in which GSK3β along with casein kinase 1 and protein kinase A phosphorylates the 155-residue in Ci and this phosphorylation is the signal to degradation,[7,13] very much similar to the Wnt-pathway targeting of β-catenin for degradation.

Thus, in this and in the Wnt-pathway GSK3β appears to repress transcriptional activity. It must be stressed that the hedgehog-patched signalling pathway is very complex and far from resolved involving many proteins, kinases and phosphatases. In addition this signalling system has many targets including Patched itself, Wnt-gene family, and members of the TGFβ superfamily and their second messengers.[14]

4. *Notch pathway*. Notch signalling participates in several cellular processes the most important being preserving stem cells, cell-fate decision processes and induction of terminal differentiation, in all cellular ontogenese.[17,18] Notch is a type 1 integral, single-pass transmembrane receptor, which are activated in *trans* by membrane bound ligands.[19] Initiating Notch signalling by receptor-ligand interaction between two cells is thought to be accomplished by dissociation of the extracellular part of the receptor followed by *trans*-endocytosis into the opposing cell. This is triggered by an extracellular metalloprotease TACE (TNFα-converting enzyme/ADAM 17) cleavage of the receptor,[20–22] although other enzymes may be active in this process.[23] Cleavage of the intra-membrane portion of Notch by the γ-secretase complex follows, thereby releasing the cytoplasmic part of Notch (Notch-IC).[17,19,24,25] This fragment enters directly into the nucleus and binds to the transcription factor Suppressor of Hairless.[26,27] The nuclear transcription activity may be stabilized by GSK3-phosphorylation of intracellular/nuclear Notch-IC,[17] although this is debated. Recent studies demonstrates the cross-talk between Notch-signalling and the PIK3-Akt pathway, where activation of the latter (by e.g. insulin) inhibits GSK3β.[28] Termination of Notch signalling is accomplished by phosphorylation of the Notch intra-cellular signal protein and subsequent degradation by the ubiqutination-proteasome pathway.[27]
5. *Steroids*. Steroids stimulates the transcription of the serum- and glucocorticoid-regulated protein kinase (SGK). This kinase is then activated by phosphoinositide-dependent protein kinases (PDK) which in turn activates Akt. *In vitro* both kinases inactivates GSK3β, i.e. SGK and Akt may substitute or complement each other.[29] Glucocorticoid-induced degradation (through Akt) of GSK3β prevents degradation of the cytoplasmatic pool of β-catenin, which is a component of microtubules and the adherence junction complex.[30] GSK3β itself is also recruited to the to the cellular adhesion complex[31] promoting on-location regulatory phosphorylation. Thus the dual function of β-catenin, as a transcription factor and as structural cellular component is regulated in completely different ways.

Several other GSK3β-regulating pathways acting through phophorylation of Ser9 has been identified. This includes kinases such mitogen activated protein kinases (MAPK) regulated by growth factors, amino acids through a pathway involving the rapamycin target mTOR, and by cAMP-dependent kinase (PKA).[5,32] The mTOR pathway regulates the ribosomal p70/p85-S6 kinase, which in turn inactivates GSK3β.[33] Activation of the Fas receptor mediates apoptosis, but the mechanism is distinct from classical Akt/GSK3β pathway as it is the Fas-regulator c-FLIP, inhibiting the Fas-mediated recruitment of caspase-8/10, that reduces phosphorylation of both Akt and GSK3β at least in cardiomyocytes.[34]

GSK3β activity are also modulated by dihydrofolate reductase (DHFR), which is the target for the anti-inflamatory drug methotrexat. The effect is probably not direct, but via changes in DNA-methylation altering the transcription levels.[35] Folate depletion has been shown to alter several components of the Wnt-pathway including APC and β-catenin[36] which would influence GSK3β activity. In addition, the expression of GSK3β itself may be influenced by altered methylation patterns in it s promoter, although this is speculative at the moment.

Recently, it has been discovered that calpain may truncate neuronal GSK3β, which increase the GSK3β activity. The exact significance of the cleavage is not known.[37]

Physiology of GSK3β

Glycogen is the major storage form of glucose and its cellular level are tightly regulated by glycogen synthase, GSK3β, and glycogen phosphorylase.[38] GSK3β is regulated by insulin in a complex process as mentioned above and that involves phosphatidylinositol-3-kinase (PI3-K), 3-phosphoinositide-dependent kinase, PDK1, and protein kinase B (PKB/Akt) mediated phosphorylation of Ser9 in GSK3β.[5,39,40] However,

GSK3β is involved in many pivotal processes, where regulation of the glucose metabolism is just one and maybe not the most significant.

Cell cycle and embryogenesis

GSK3β is involved in the meiotic process in the oocyte presumably by modifying spindle activity.[41,42] Further, GSK3β is also of paramount importance in the cell cycle and embryogenesis,[43] but the outcome of the GSK3β activity is complex and far from elucidated. For instance, at least one of the isoforms of GSK3 are necessary for axon elongation in a critical period comprising at least the first 24 hours of axon formation. Inhibition of GSK3 after this period does not compromise elongation, but exacerbates axon branching.[44] GSK3β probably exerts its effect by phosphorylation of the microtubule-associated proteins MAP1B and tau reducing their binding to the microtubules. This renders the microtubules unstable and more dynamically active in axon growth.[45]

Ontogenesis of mesenchymal cells

The canonical Wnt/β-catenin pathway controls the differentiation of osteoblasts and chondrocytes. β-catenin is required for osteoblast development from the mesenchymal progenitor cell, and suppression of β-catenin induces differentiations to chondrocytes instead.[46,47] In established chondrocytes the dysregulation of Wnt-pathway increasing β-catenin expression may be involved in development of osteoarthritis.[48] The increase of β-catenin may be accomplished by increased expression of Smurf2. This protein interacts with GSK3β inducing ubiquitination of GSK3β itself and subsequential proteasomal degradation. The final result is deprivation of GSK3β and hence reduced phosphorylation and degradation of β-catenin.[49] Thus, is seems that dysregulation of the signal pathway at this stage reverts the chondrocytes to an osteoblast-like behaviour with ectopic calcification of cartilage.

Muscles

GSK3β has been established as a kinase with strong anti-hypertrophic properties in cardiac muscle,[50] probably related to the apoptotic action of GSK3β. It has been suggested that (excessive) Wnt-signalling is causative in cardiac hypertrophy by re-activating a fetal gene program and by inhibiting GSK3β, but this remains to be firmly established. Wnt-independent mechanisms involving GSK3β (and β-catenin) has also been suggested[51] in this context. Activation of the Fas receptor, usually mediating apoptosis, has been associated with cardiac hypertrophy by inhibition of GSK3β.[52,53] The natural regulator of Fas, c-FLIP, seems to have a protective role in the heart preventing cardiac hypertrophy. Overexpression of c-FLIP reduces phosphorylation of both Akt and GSK3β, suggesting that the hypertrophy is avoided at least partially by the action of GSK3β.[34] Interestingly, in zebra fish (Danio rerio) GSK3α, but not GSK3β is necessary for the survival of the cardiomyocytes, while GSK3β exerts its important functions in modulating right-left symmetry of the heart.[54] The signalling is through the PKC/ERK/RSK pathway and not through the Akt pathway.[55] Although our knowledge still is rather rudimentary, the anti-hypertrophic effect of GSK3β seems well established.

Apoptosis

GSK3β is pro-apoptotic and inhibition of GSK3β may prevent apoptosis.[56,57] The mechanism is not clear, but may involve the nuclear transcription promoter NF-κB. The potential role of GSK3β may be tissue-specific.[58] NF-κB is retained as an inactive transcription factor in the cytoplasm by binding to the inhibitory protein IκB. When IκB becomes phosphorylated it is targeted for ubiquitin-dependent degradation, and NF-κB may enter the nucleus. Although GSK3β is thought to be instrumental in this process, the mechanism has not been unequivocally established.[58,59] Another suggested mechanism is the inhibitory phosphorylation of translation initiation factor eIF2B by GSK3β.[60] Further support for the involvement of GSK3β in apoptosis comes from the findings that growth factors phosphorylation of GSK3 inhibits apoptosis.[61,62] Withdrawal of growth factors induces apoptosis, which in addition can be prevented by lithium, a well-established inhibitor of GSK3β, although not entirely specific.

Nervous tissue

GSK3β facilitates staurosporine- and heat shock-induced neuronal apoptosis.[57,63] The neuronal pro-apoptotic effects of GSK3β, measured as caspase-3 activity, is inhibited by valproate and lamotrigen supporting neuronal survival. This effects mimics the insulin/growth factor stimulation of Akt-inactivation of GSK3β.[64,65] Heat shock can activate Akt and this can

be inhibited completely by LY294002 and wortmannin, i.e. the activation of Akt is PI3K-dependent. Heat shock and other stimuli like cell injury, oxidation or osmotic stress, activates the heat shock factor-1 (HSF-1) transcription factor,[66] which in turn increases the expression of heat shock proteins conveying protection against lethal conditions. GSK3β suppress the activity of HSF-1 by phosphorylation of Ser303 in HSF-1. Also, inactivation of GSK3β (but not GSK3α) in the Wnt-pathway protects the neurons from apoptosis in the cerebellum.[57]

p53-induced apoptosis after DNA-damage involves binding of p53 to and activation of nuclear GSK3β. In this context GSK3β promotes p53-mediated transcription by phosphorylation of serine 33 in p53.[67–69] GSK3β accumulates in the nucleus during senescence of human fibroblast. Here the kinase is complexed with p53, and this interaction may promote senescence.[70]

These are just a few examples of the more well-studied effects of GSK3 and particular the β-form, but the ubiquitous espression of GSK3β in almost all tissues suggest involvement of the kinase in most cellular processes in development. However, much has to be done to establish the role of GSK3β.

GSK3β and Disease

Diabetes

As mentioned above GSK3 earned its name by the discovery of its role in glucose homeostasis. Indeed, GSK3, and particularly the β-form, has been shown to be instrumental in this context and aberrant activity of the kinase has been linked to diabetes mellitus. GSK3β have two effects in the liver: it inhibits glycogen synthase and increases the expression of the gluconeogenetic genes glucose-6-phosphatase and phospho-enol-pyruvate carboxypeptidase increasing the glucose output.[71] GSK3β Ser/Thr-phosphorylation of insulin related substrate IRS-1, inhibiting tyrosine-phosphorylation of IRS-1, may also be a part of the glucose homeostasis. Thus, increased GSK3β activity moves the physiological state of a subject in the direction of a diabetic phenotype. Inhibition of GSK3β improves the diabetic state through increased activity of the Akt-pathway stimulated by insulin.[7]

Alzheimers disease

GSK3α and -β are highly abundant in the brain. The kinases phosphorylates various nuclear, cytosolic and extracellular proteins and the three major regulatory pathways of GSK3 (Akt, Wnt, and growth factor pathways) are in operation in brain.[72] Processes such as synaptic transmission, axonal transport and cell-cell interactions are regulated by GSK3. Also, GSK3 is involved in brain development and apoptosis.[73]

Increased levels of GSK3β has been linked to the development of Alzheimer's disease and other neurodegenerative disorders[73] The mechanism includes modulation of interactions of β-amyloid, tau, presenilin and other proteins, as well as apoptosis and modulation of cholinergic signal transmission.[74] In particular, hyper-phosphorylation of tau resulting in neurofibrilar tangles has been implicated in Alzheimer's disease. Presenilin (PS1 and PS2) may function as a scaffold in a complex of presenilin, tau, GSK3β, the amyloid precursor protein (APP), and catenins. The mechanisms are however complex and far from elucidated.[75,76] PS1 binds both tau and GSK3β facilitating the interaction of tau with GSK3β.[77] PS1 also forms complexes with GSK3β and β-catenin.[78] In addition, presenilin activates γ-secretase, which is involved in the cleavage of APP to amyloid-β peptides (Aβ, the hallmark of Alzheimers disease) and is instrumental in the Notch pathway. The latter probably implicates a cross-talk with the Wnt-pathway, implicating the latter in development of Alzheimer's disease.[74] Thus, GSK3 is involved en formation of the two major components in Alzheimers disease, the amyloid plaque and neurofibrilar tangles. Finally, diabetes type 2 and Alzheimers disease (and schizophrenia) share overlapping pathology in the insulin resistance and amyloidogenesis through the PI3K/Akt/GSK3β pathway in the brain. Insulin receptor mRNA and receptor protein are present in discrete areas in the brain, but differs in several aspects compared to the peripheral insulin receptor. The co-occurrence of insulin resistance and amyloidogenesis mutually accelerates both, but the exact mechanism is not clear.[79]

Various other neurodegenerate conditions have been associated with dysregulation of specific Wnt- and Notch-pathways (see)[80] for a presentation).

Mental disorders

Dysregulated serotonin (5-hydroxy tryptophan, 5-HT) is involved in many mental conditions as schizophrenia and autism, depression and bipolar disorders, and anxiety. Stimulation of the two serotonin receptors

5-HT_1 and 5-HT_2 respectively enhances or reduces Ser9 phosphorylation of GSK3β, and hence reduces or enhances the activity of GSK3β, respectively. Mutations (artificial) in the rate-limiting enzyme for 5-HT synthesis, tryptophan hydroxylase 2 (Tph2) reduces the 5-HT level in frontal cortex in mice by ~80% with a concomitant activation of GSK3β,[81] but the exact involvement of GSK3β in the mental conditions is unclear. Activation of dopamine-like receptors (D2-like) activates GSK3β, through dephosphorylation of Akt in a complex involving β-arrestin-2 and the protein phosphatase PP2A. Thus, Akt is inactivated thereby releasing the Ser9 phosphorylation of GSK3β.[82]

There has been reports of dysfunction of the Wnt-pathway in schizophrenia resulting in altered GSK3-activity and reduction in β- and γ-catenin in the frontal cortex.[72,83,84] However, some studies does not find any alterations in β-catenin, Dishevelled or GSK3β in the prefrontal cortex in schizophrenia, bipolar disorders or major depression.[85]

Clearly, GSK3β is involved in both neurodegenerative conditions and mental disorders, but many issues are unsolved. It is difficult to study brain function, but the association to common conditions as diabetes mellitus may open new avenues to entangle at least some of the processes in the brain, despite the shortcomings, limitations and pitfalls there obviously may be.

Cancer

GSK3β is overexpressed in several colon cancer cell lines, which does not seem to influence β-catenin accumulation in the nucleus. Inhibition of GSK3β may promote apoptosis in these cells,[86] contrary to the supposed apoptotic effects of GSK3β in normal cells. In fact, the impact of GSK3β depends on the tumour: GSK3β may be a tumour suppressor in skin and mammary cancers, but functions as a tumour promoter in colon and pancreatic cancers.[87] Thus, the diverse actions of GSK3β is not a feature of the kinase *per se*, but rather is consequence of aberrant activity in the interactions with other proteins and metabolics in the various networks GSK3β is a part of.

Several oncogenes are modified by GSK3β,[88] such as MUC1 and c-Myc.[89] MUC1 is overexpressed in a variety of tumours and may sequester β-catenin, thereby inhibiting formation of the E-cadherin/β-catenin complex contributing to reduced cell-cell and cell-matrix interactions promoting invasiveness of tumour cells. Phosphorylation of MUC1 by GSK3β is proposed to reduce the MUC1-β-catenin interaction.[7] DNA-dependent protein kinase (DNA-PK) modulates the c-Myc oncogene by phosphorylation of Akt. As usual this results in reduced GSK3β and in this context reduced phosphorylation of c-Myc. Deficiency of DNA-PK kinase decreases the phosphorylation of Akt leading to increased activity of GSK3β, which in turn phosphorylate and stabilize c-Myc.[89] Adding to the list of GSK3β in cancer, breast cancer often displays abnormalities in the PTEN/PI3K/Akt/β-catenin pathway.[90]

Only a few genic germ-line mutations have been reported with unknown clinical impact (see below). Most interestingly, mutations in GSK3β has been detected in sputum from patients with lung carcinomas.[91] The mutations or polymorphisms are all present in introns, suggesting that the dysregulation of the kinase may be the important factor, and not functional mutations in the gene *per se*. However, none of the mutations could be established as monogenic and are also detected in non-carcinomatous tissue, hence must be put in context of a larger genetic network of genes related to cancer development. Nevertheless, these observations suggests that even cancers, often supposed to be sporadic, may in fact be polygenic in nature and at least for some sub-populations of cancers they may be inherited.

In summary, GSK3β is involved in such a large amount of processes, that you may ask if genetic variations of this kinase *per se* is the cause of aberrant behaviours in many physiological functions and disease. *A priori* you may anticipate that any genetic variation, at least in the functional part of the gene i.e. the protein, would have such a diverse impact that it would be incompatible with life. Below we will provide evidence that GSK3β is in fact a highly conserved protein, with only few mutations allowed, which would classify GSK3β as an evolutionary and functional ultraconserved protein.

Methods

Evolution of GSK3β

The species included in the analysis covers the phylogenetic evolution back to Precambrian periods. The phylogenetic timeline were deducted using the TimeTree programme,[92] which uses databases including molecular data from hundreds of studies of the

molecular clock. The programme makes extensive use of the NCBI Taxonomy browser, which in August 2009 contained nucleotide sequences from more than 300,000 organisms.[93,94] A particular advantage of the programme is that time estimates of divergence of two species are traced back to the same node irrespective of which member of the lineages are queried. Both nucleotide (genomic and mitochondrial) and protein data are included in the calculations, which necessarily introduces variability of the estimates. The number of estimates and the number of genes included in the calculation of molecular distances varies considerably and there may be ambiguities about the node of divergence.

Molecular details from recent studies were obtained from the sources (e.g.)[92,95–99] and see the links in the TimeTree programme), but in some instances a species could not be placed accurately in the phylogenetic tree and the first occurrence of the nearest ancestor was used instead. Also there may be differences in estimates between fossil and molecular trees. These inaccuracies is however of lesser importance as the issue here is the position on the timeline, not the exact evolutionary distance between species.

DNA and protein sequences of human GSK3β were aligned using the Blast algorithms[100] to all species directly accessible from National Centre for Biotechnology Information, NCBI (http://www.ncbi.nlm.nih.gov). The alignment was done for the full coding sequences and translated proteins as well as the exons and the exon-encoded peptides separately. The latter could be ambiguous in the terminals of the peptide as many codons are assembled by two exons. In this instances the amino acids were included in both exons in the search process. The most recent build of the sequence databases were used.

Insulin resistance

Due to the essential role in glucose homeostasis, it has been hypothesized that variants in the GSK3β genes are risk factors that predispose to insulin resistance and type-2 diabetes. Besides frank diabetics several other conditions have a co-morbidity with insulin resistance including schizophrenia[101] and 40%–60% of HIV-patients in antiviral treatment.[102,103] Therefore, all the exons, including the splice site regions were sequenced to detect any mutations/polymorphisms related to insulin resistance. The patients were sampled from three different biobanks: 20 subjects with diabetes mellitus type 2 (Diabetes Biobank, Hvidovre Hospital, Denmark); 26 patients with schizophrenia (Danish Psychiatric Biobank, Sct. Hans Hospital, Denmark); and 26 HIV patients from (The HIV Biobank, Hvidovre Hospital, Denmark). All subjects suffered from insulin resistance and all patients represent the extreme phenotypic expression of their primary disease. DNA was extracted from whole blood using standard methods. Details of the primer design can be found in the Supportive material.

Population

Several polymorphisms have been detected in human GSK3β, but most are not validated thoroughly. Therefore, we genotyped a large randomly sampled population (Monica 10) to validate the polymorphisms and if present, their relation to clinical outcomes. The Monica 10 population[104,105] were genotyped for most of the exonic GSK3β single nucleotide polymorphisms (SNP) included in the dbSNP database maintained by NCBI. The Monica 10 study population has been described in detail elsewhere. In brief, Monica 10 consisted of 2,656 re-invited participants (1993 to 1994) who originated from the Danish part of the monitoring of trends and determinants in cardiovascular disease (MONICA) health survey,[106] which was performed from 1982 to 1984. The present genotyping in addition included 173 participants from the original sample not included in the Monica 10 sample. In total 2,829 subjects were genotyped for 6 of the 9 reported SNPs in human GSK3β. The genotyping of Monica 10 was performed by KBioscience, Manchester, UK.

The genotyping of the patients and Monica was approved by the ethics committee for Copenhagen County, Denmark. Written informed consent was obtained from all participants.

Results

Table 1 summarizes the alignment of human GSK3β protein and DNA sequences with several animal species covering evolutionary time back to Cambrian period 540 million years ago (Mya). Table 2 summarize data for species originating in the Pre-Cambrian periods. (GSK3α is mainly included to illustrate the extent of similarity to GSK3β). The timeline is constructed as the average of estimates for the molecular divergence

of the species from Homo sapiens without trying to evaluate the validity of single studies, although TimeTree offer an expert evaluation. Mostly this timeline is in accord with fossil time lines. Some ambiguity are, however, present particularly in ordering the subphyla of chordates.[107] Chordates, a monophyletic group sharing a specific body plan, consists of three subphyla: urochordates (here represented by *Ciona intestinalis*), cephalochordates (here represented by *Branchiostoma floridae*), and craniata or vertebrates. The relationship between these three subphyla is however murky. Using genomic data the cephalochordates may be declared the basal chordate subphylum predating urochordates and vertebrates, the latter two being sisters.[108] However, mitochondrial data would suggest that urochordates are the basal subphylum.[107] Recent studies using protein data in the molecular clock estimates suggest that the urochordates diverged from cephlohcordates in the Precambrian period about 900 Mya and from vertebrates about 800 Mya, that is in the Cryogenian period.[97,99] The hemichordates and echinoderms (worms, sea urchins) as well as the protostomians (the other major clade of bilateral animals), which includes arthropods and nematodes, also dates back deep in the Precambrian period.[98] Thus, the crown group of animals diverged long before the ecological Cambrian explosion. The cradle of animal life lies in the Cryogenian period (635–850 Mya) and the emergence of bilaterian taxa arose in the Ediacaran period (from 635 and up to the Cambrian explosion), and since then the higher-level taxa have developed in separate lines. It should be noted that the molecular estimates are about 300 Mya earlier than the fossil records would suggest.

The platypus and marsupials/eutherian are descendent of ancestors in distinct geographical areas, and it has been disputed if these monophylic divisions are members of the same clade according to the Thera hypothesis.[109] The platypus together with the echidna are the only extant egg-laying mammals (monotremes). They diverged from birds 325 Mya exactly as Homo sapiens did. The important issue here is however, not the exact time of divergence of the species (not least considering the uncertainty in the estimates of the divergence),[110] but rather the hierarchical location on the time line and position in the phylogenetic tree, as when the split occurred evolutionary events in the species proceeded independently of the common ancestor. The prominent members of the primates, gorilla and orangutan are not included here due to lack of or only sparse sequence data are available in the NCBI databases.

Several striking results emerge. First, the entire GSK3β protein is highly conserved among vertebrates. This conservation dates back at least to the divergence of amniota into mammals and birds about 325 Mya, and only deteriorated slightly in amphibians and fish, the divergence to Homo sapiens dating back 455 Mya. The high degree of homology is reflected by the almost complete homology between most of the exon-coded peptides of GSK3β in mammals and to somewhat lesser extent for some of the remaining vertebrates in the table.

Second, peptides encoded by exon 3, 6, 7 and 8 are completely conserved in all species and this is also true for exon 1, 2, 5 and 10 in mammals, except for the old world monkey Cynomolgus and platypus. The complete conservation of peptides encoded by exon 1, 2 and 5 extents to birds, while the frog and fish peptides are next to complete for these exons. Thus, the conserved regions harbours the serine/threonine kinase active domain located in exon 5 and the tyrosine kinase activity mapped to exon 3, although the protein sequence 56–353, encoded by exon 2 to 10, is involved in the kinase activities. Thus, the functionally most important domains of the enzymes seems to have been conserved almost to completion since the Cambrian explosion.

Third, homologies of exon 1, 11 and 12 is high in vertebrates except in platypus. In non-vertebrates (Table 2) the homologies in these exons deteriorate rapidly, suggesting species-specific functionalities (if any) of these exons. Some divergence is detected for exon 11 and 12 peptides in vertebrates, but still the similarities are very high. The homologies to the two *Macaca* are rather low. These peptides are annotated to be GSK3α peptides and has a high similarity to several of the other vertebrate GSK3α including man, chimpanzee, dog, cow, rat and mouse. The actual sequence of the exon 11-encoded peptide for the *Macaca*'s is unresolved.

Exon 9, coding a complete 13 amino acid stretch, is somewhat ambiguous, in the sense that it were not detected in its full length in many of the species, including mammals. However, at least in vertebrates this may be caused by technical problems in the

Table 1. Homology of homo sapiens GSK3β with vertebrates estimated to have diverge after the cambrian explosion.

Mya f	Class	Order	Species	Common name	Length	Coding region only
	Mammalia	Primates	Homo sapiens	Man (GSK3β)	Protein	433
					Exon	*1302*
	Mammalia	Primates	Homo sapiens	Man (GSK3α)		71.1
						46.2
5.4	Mammalia	Primates	Pan troglodytes	Chimpanzee		98.15
						99.2
30.2	Mammalia	Primates	Macaca mulatta	Rhesus monkey		92.2
						91.4
30.2	Mammalia	Primates	Macaca fascicularis	Cynomolgus monkey		72.1
						47.8
98.2	Mammalia	Artiodactyla	Bos taurus	Cow		99.0
						96.0
98.2	Mammalia	Artiodactyla	Ovis aries	Domestic sheep		94.0
						93.8
98.2	Mammalia	Artiodactyla	Sus scrofa	Wild boar or hog		95.6
						93.0
98.2	Mammalia	Carnivora	Canis lupus familiaris	Dog		99.77
						97.0
98.2	Mammalia	Perissodactyla	Equus callabus	Horse		99.5
						96.7
103.7	Mammalia	Rodentia	Rattus novegicus	Rat		99.05
						89.9
103.7	Mammalia	Rodentia	Mus musculus	Mouse		99.05
						90.3
103.7	Mammalia	Rodentia	Spermophilus citellus	Squirrel		96.3
						93.9
180	Mammalia	Didelphimorphia	Monodelphis domestica	Opossum		99.5
						92.5
220	Mammalia	Monotremata	Ornithorhynchus anatinus	Platypus		56.1
						48.5
Divergence of amniota						
325	Aves	Galliformes	Gallus gallus	Chicken		96.77
						88.7
325	Aves	Passeriformes	Taeniopygia guttata	Zebra finch		97.0
						88.6
361	Amphibia	Anura	Xenopus laevis	Frog		88.0
						77.0

Exon 1	Exon 2	Exon 3	Exon 4	Exon 5	Exon 6	Exon 7	Exon 8	Exon 9	Exon 10	Exon 11	Exon 12	3'UTR
30	64	28	38	43	35	34	31	13	63	33	21	
88	*192*	*86*	*111*	*131*	107	98	96	39	187	*99*	*68*	87
low **a**	78.1	**100**	73.7	90.9	**100**	**100**	**100**	NAe	71.0	48.5	NA	
inc **b**	*72.0*	*88.4*	*72.1*	*inc*	*80.4*	*76.5*	*inc*	NA	*40.1*	NA	NA	
100	**100**	**100**	86.5c	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	
99.8	*100*	*100*	*85.6*	*100*	*99.1*	*100*	*100*	*97.4*	*100*	*99.0*	*100*	*100*
100	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	45.5	**100**	
98.2	*97.9*	*98.8*	*100*	*96.9*	*100*	*96.9*	*99.0*	*100*	*98.3*	*inc*	*98.5*	*inc*
low	79.4	**100**	73.6	93.0	**100**	**100**	**100**	7/7d	71.4	45.5	*low*	
inc	*70.6*	*79.1*	*73.0*	*inc*	*79.4*	*inc*	*inc*	*inc*	*39.6*	*inc*	*inc*	*inc*
100	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	94.0	85.7	
97.8	*95.8*	*96.5*	*96.4*	*96.9*	*99.0*	*93.9*	*93.8*	*100*	*97.9*	*93.9*	*86.8*	*95.4*
100	**100**	**100**	**100**	**100**	**100**	**100**	**100**	*low*	**100**	94.0	85.7	
98.9	*96.4*	*96.5*	*97.3*	*96.9*	*99.1*	*92.9*	*95.8*	*inc*	*97.9*	*93.9*	*94.1*	*inc*
100	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	97.0	85.7	
97.8	*96.9*	*95.3*	*99.1*	*97.7*	*100*	*92.9*	*94.8*	*100*	*94.1*	*94.9*	*85.3*	*98.7*
100	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	95.2	
98.9	*97.4*	*95.3*	*100*	*96.2*	*100*	*92.7*	*95.8*	*100*	*97.9*	*97.0*	*92.6*	*97.7*
100	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	97.0	95.2	
98.9	*96.9*	*95.3*	*98.2*	*96.2*	*97.2*	*94.8*	*94.8*	*100*	*98.4*	*94.9*	*95.6*	*97.6*
100	**100**	**100**	**100**	**100**	**100**	**100**	**100**	7/7 **d**	**100**	94.0	90.4	
94.3	*90.6*	*95.3*	*93.7*	*92.4*	*93.5*	*91.8*	*90.6*	*43.6*	*93.6*	*88.9*	*91.2*	*88.5*
100	**100**	**100**	**100**	**100**	**100**	**100**	**100**	7/7 **d**	**100**	94.0	90.4	
94.3	*91.1*	*94.2*	*93.7*	*95.4*	*94.4*	*91.8*	*91.7*	*76.9*	*94.7*	*85.9*	*94.1*	*88.5*
100	**100**	**100**	94.7	**100**	**100**	**100**	**100**	NA	**100**	**100**	90.4	
97.7	*95.9*	*97.7*	*95.5*	*93.9*	*98.1*	*96.9*	*92.7*	*inc*	*95.7*	*93.9*	*95.6*	*94.4*
100	**100**	**100**	97.3	**100**	**100**	**100**	**100**	**100**	**100**	**100**	95.2	
97.7	*90.1*	*96.5*	*91.9*	*91.6*	*91.6*	*91.8*	*89.6*	*82.1*	*92.5*	*93.9*	*88.2*	*inc*
46.7	**100**	**100**	**100**	**100**	**100**	**100**	*low*	7/7 **d**	*low*	NA	47.6	
inc	*81.4*	*86.0*	*82.0*	*90.8*	*90.7*	*88.8*	*inc*	*inc*	*inc*	*inc*	*inc*	*inc*
100	**100**	**100**	97.3	**100**	**100**	**100**	**100**	**100**	96.8	87.9	66.7	
90.8	*87.0*	*93.0*	*90.1*	*89.3*	*89.7*	*89.8*	*90.6*	*53.8*	*85.0*	*86.9*	*80.9*	*86.9*
100	**100**	**100**	97.3	**100**	**100**	**100**	**100**	**100**	98.4	87.9	66.7	
90.0	*87.5*	*89.5*	*91.0*	*91.6*	*91.6*	*89.8*	*88.5*	*89.7*	*83.4*	*86.9*	*79.4*	*inc*
96.7	96.9	**100**	94.7	97.7	**100**	**100**	**100**	NA	90.5	50.0	80.9	
85.6	*79.4*	*87.2*	*80.2*	*84.0*	*85.0*	*45.9*	*79.2*	*inc*	*75.4*	*inc*	*inc*	*inc*

(*Continued*)

Table 1. (*Continued*)

Mya f	Class	Order	Species	Common name	Length	Coding region only
455	Actinopterygii	Tetraodontiformes	Tetraodon nigroviridis	Pufferfish		70.4
						inc
455	Actinopterygii	Cypriniformes	Danio rerio	Zebra fish		89.8
						77.3
				Average similarity	Protein	91.2
					Exon	*86.0*

sequencing procedure (shot-gun sequencing), and may not be real. In non-vertebrates the alignments are poor for this exon, which very well may be real, but considering the rather high level of homologies detected for most of the other exons (particular the flanking exons 8 and 10), it could possible be that exon 9 for some unknown reason may be difficult to sequence. This is also reflected in the inconclusive and incomplete sequences alignments detected, particular in non-vertebrates. The homology of the nucleotide sequences may often seem rather low, but in many instances shorter stretches with perfect match are detected while the remaining sequence is missing most probably for technical reasons. Analysis of *Strongylocentrotus purpuratus* and *Nematostella vectensis* revealed a perfect similarity of 8 consecutive amino acids in the exon 9-encoded 13 amino acid peptide. None of them were annotated as being GSK3β related and for *Nematostella vectensis* the protein is simply quoted as predictive to be a protein. However, considering the probably general problem in sequencing this exon it may be suggested that the exon or at least the peptide coded by the exon 9 is highly conserved deep into the Precambrian period. Thorough re-sequencing is needed to confirm this hypothesis.

A survey of selected non-animal species revealed high similarities in the exon 5, 6, 7, and 8-encoded peptides. The homology to the human peptide is absolute for the exon 7-encoded peptide for *Triticum* (wheat) and *Aegilops* (flowering plants), 97.1% for Zea mays (all species with a predicted molecular distance from Homo sapiens of 1,400 Mya), and 96.7% for *Oscarella lobularis* (marine sponge) with a molecular distance of 1,020 Mya. The homologies are somewhat less for the other three peptides ranging from 80% to 91%. Nevertheless, the homology is extremely high for these peptides covering 1/3 of the entire enzyme, and most importantly includes the pivotal enzymatic active domain of GSK3β.

A closer look at the vertebrate protein sequences (Supporting material T2) reveal that many of the dissimilarities to Homo sapiens are deeply rooted. Almost all differences are caused by single nucleotide mutations and many may in fact be ambiguous as more than one amino acid substitution are predicted. Thus, the homologies are most probably on the conservative side and may increase as and if more comprehensive sequencing is performed. The amino acid substitutions collated (Supporting material T2, 115 in all) are caused by 57 (54.3%) nucleotide changes in the first position in the codons, 23 (21.9%) in position 2, 19 (18.1%) changes in both position 1 and 2, and only 7 (5.7%) in position 3 (the wobble position). The number of unique substitutions are 83 requiring 43 (51.8%), 19 (22.9%), 4 (4.8%), and 17 (20.5%) changes in nucleotide 1, 2 or 3, and changes in both position 1 and 2, respectively.

Table 1. (*Continued*)

Exon 1	Exon 2	Exon 3	Exon 4	Exon 5	Exon 6	Exon 7	Exon 8	Exon 9	Exon 10	Exon 11	Exon 12	3'UTR
96.6	**100**	**100**	94.7	97.7	**100**	**100**	**100**	8/8	93.5	44.1	38.1	
inc	*inc*	*NA*	*inc*	*inc*	*inc*	*inc*	*inc*	*inc*	*NA*	*inc*	*inc*	*inc*
96.7	**100**	**100**	94.7	**100**	**100**	**100**	**100**	69.2	93.7	75.8	85.7	
88.6	*81.3*	*84.9*	*82.0*	*83.2*	*79.4*	*81.6*	*81.3*	*43.6*	*75.9*	*inc*	*33.8*	*inc*
96.3	98.6	**100.0**	91.4	99.4	**100**	**100**	**100.0**	96.9	96.7	80.1	82.3	
95.3	*90.4*	*93.0*	*91.2*	*93.3*	*93.4*	*89.5*	*91.9*	*70.5*	*88.8*	*92.5*	*86.9*	*94.2*

In the alignment of exon 1 and 12 DNA-sequences only the protein coding parts are used.
[a]Low homology, doubtfull or short sequences, fragmented, or not annotated as GSK3β protein.
[b]Inc, incomplete, fragmented or inconclusive nucleoteide seqeunces, or sequences covering less than 1/3 of the human sequence.
[c]Thirtytwo of the 38 amino acids in the exon were detected by assmbling several sequences. The nucleotide sequence detected spanned 96 nt of the 111 nt in the exon and 95 of these 96 were identical to the human exon.
[d]The (S)SGTGHFT is common in (2)3–9 position in the 13 amino acid sequence of eson 9. This sequence is partially identical to the tetracycline repressor protein (Transposon Tn10).
[e]NA, No sequences were detected.
[f]Mya (million of years ago) since the divergence from Homo sapiens (molcular clock estimates). The weigthed avarage of all studies as collated in TimeTree is shown.

The only non-synonymous substitution detected to be heterozygotic in human is residue 152 in exon 4. This residue is an alanine in *Xenopus*, but a Thr->Met substitution is suggested to be present in man (a C->T changes in position 2 of the codon). None of the other substitutions in the table (Supporting material T2) has been detected as polymorphic in man.

The substitution rates for the amino acids are extremely low. In the vast majority of exon-coded peptides no substitutions are detected at all. The lowest substitution rate (disregarding *Macaca fascicularis*) is seen in exon 2 for *Danio rerio* (5.1 E-11 substitutions per year in the peptide), which is far lower than expected for neutral substitutions (see)[111] for a discussion). More to the point is that most exon-coded peptides are invariant deep in the phylogenetic tree, and, particularly for exon 7, the invariance extents 1,400 Mya to wheat and plants. Again, this may be on the conservative site due to not entirely reliable sequence data for many species (including *Macaca* and platypus).

The protostome (the other major bilaterian) annelid *Pltynereis dumerlilii* did reveal some modest homology to GSK3β, but the protein is annotated as a Jun-kinase, and the stretches of homology did not extend beyond four amino acids and have many gaps. The overall homology is 31% and only 41% for the highly conserved exon 7-coded peptide. This is in sharp contrast to some plants like *Tricitum* and *Aegilops* (100% amino acid homology to the exon 7-encoded peptide) and several other plants and trees with homologies in range of 84 to 97% for this peptide (not shown). All of these latter species are estimated to diverge from human apporximately 1,400 Mya (not shown).

The DNA sequences are not as conserved as the protein sequences, which is to be expected as many mutations results in synonymous substitutions of amino acids and the selection pressure is on the functionality of the proteins and not the DNA sequences *per se*. For gorilla and orangutan (Pongo) only sparse protein and DNA homologies were obtained. It is unlikely that these species, which diverged less than 15 Mya from Homo sapiens, do not posses the genes for GSK3β. Rather, sequencing may be incomplete at the moment. In fact, many of the alignments may be curtailed by less than optimal and thorough sequencing, and hence the homologies may be underestimated. The 3'UTR region is highly conserved in the vertebrate species where solid sequence data is available (Table 1).

Genotyping in homo sapiens

In the patient samples all 12 exons were successfully amplified and sequenced in all patients with the exception of exon 1 which could not be determined in one subject with diabetes mellitus type 2. We did not find any new mutations in any of the examined exons,

Table 2. Homology of homo sapiens GSK3β protein with non-vertebrates estimated to have diverge before the cambrian explosion.

Mya f	Phylum	Class	Order	Family	Species	Common name
	Chordata	Mammalia	Primates	Hominidae	Homo sapiens	Man (GSK3β)
794	Chordata	Ascidiacea	Phlebobranchia	Cionidae	Ciona intestinalis	Vase tunicate
868	Chordata	Cephalochordata		Branchiostomidae	Branchiostoma floridae	Florida lancelet
653	Platyhelminthes	Turbellaria	Tricladida	Planariidae	Dugesia japonica	Flatworm
768	Echinodermata	Echinoidea	Temnopleuroida	Toxopneustidae	Lytechinus variegatus	Green sea urchin
768	Echinodermata	Echinoidea	Echinoida	Echinidae	Paracentrotus lividus	Sea urchin
768	Echinodermata	Echinoidea	Echinoida	Strongylocentrotidae	Strongylocentrotus purpuratus	Purple sea urchin
856	Hemichordata	Enteropneusta		Harrimaniidae	Saccoglossus kowalevskii	Worm
867	Nematoda	Secemetea	Spirurida	Filariidae	Brugia malayi	Roundworm (nematode)
867	Nematoda	Secemetea	Rhabditida	Rhabditidae	Caenorhabditis elegans	Worm
980	Arthropoda	Insecta	Diptera	Drosophilidae	Drosophila melanogaster	Fruit fly
980	Arthropoda	Insecta	Diptera	Culicidae	Anopheles gambiae	African maliara mosquito
980	Arthropoda	Insecta	Diptera	Culicidae	Culex quinquefasciatus	Mosquito
980	Arthropoda	Insecta	Diptera	Culicidae	Aedes aegypti	Yellow fever mosquito
980	Arthropoda	Insecta	Diptera	Muscidae	Glossina morsitans morsitans	Tsetse fly
980	Arthropoda	Insecta	Hymenoptera	Apidae	Apis mellifera	European honey bee
980	Arthropoda	Insecta	Anoplura	Pediculidae	Pediculus humanus corporis	Lice
980	Arthropoda	Insecta	Lepidoptera	Danaidae	Danaus plexippus	Butterfly
980	Arthropoda	Insecta	Hemiptera	Aphididae	Acyrthosiphon pisum	Endosymbionts of insects
980	Arthropoda	Insecta	Coleoptera	Tenebrionidae	Tribolium castaneum	Red flour beetle
980	Arthropoda	Arachnida	Parasitiformes	Ixodidae	Ixodes scapularis	Deer tick
980	Arthropoda	Arachnida	Parasitiformes	Ixodidae	Rhipicephalus microplus	Cattle tick
988	Cnidaria	Anthozoa	Actiniaria	Edwardsiidae	Nematostella vectensis	Starlet sea anemone

Length	Coding region only	Exon 1	Exon 2	Exon 3	Exon 4	Exon 5	Exon 6	Exon 7	Exon 8	Exon 9	Exon 10	Exon 11	Exon 12
Protein	433	30	64	28	38	43	35	34	31	13	63	33	21
	73.9	*low*	82.8	96.4	78.9	95.3	**100**	**100**	93.4	NA	79.4	NA	42.9
	79.4	70.0	93.8	92.8	84.2	93.0	97.1	93.9	90.6		79.4	42.4	61.9
	63.0	33.3	82.8	85.7	71.1	86.4	91.4	**100**	84.4	NA	61.9	NA	*low*
	71.6	56.7	76.6	89.3	84.2	93.0	91.4	**100**	84.4	*low*	73.0	39.4	*low*
	71.4	56.7	82.8	89.3	78.9	95.3	94.3	**100**	84.4	*low*	71.4	39.4	*low*
	71.1	53.3	82.8	85.7	81.6	93.2	91.4	**100**	84.4	8/8d	69.8	33.3	66.7
	78.5	63.3	93.8	96.4	86.8	93.0	97.1	**100**	87.5	*low*	77.8	33.3	low
	57.5	*low*	75.0	85.7	76.3	88.6	82.9	93.9	58.1	*low*	73.0	39.4	42.9
	77.74	*low*	65.6	85.7	76.3	88.4	88.6	93.9	90.3	*low*	54.0	NA	*low*
	75.7	36.7	88.9	89.3	63.2	95.3	85.7	97.1	93.5	61.5	73.0	NA	47.6
	83.62	*low*	95.2	85.7	65.8	93.0	85.7	**100**	93.5	53.8	71.4	NA	low
	67.4	*low*	92.2	85.7	63.2	95.3	82.9	**100**	90.6	53.8	71.4	27.3	42.9
	70.9	60.0	93.7	83.1	63.2	95.3	82.9	**100**	90.3	61.5	71.4	NA	*low*
	42.3	*low*	85.7	89.3	57.9	95.3	85.7	60.6	*low*	*low*	*low*	33.3	*low*
	72.1	54.8	89.1	82.1	71.1	93.0	88.6	**100**	93.5	*low*	68.3	NA	*low*
	70.7	36.7	93.7	82.1	65.8	95.3	85.7	**100**	93.4	*low*	68.3	NA	*low*
	72.1	46.7	92.1	85.7	68.4	93.0	88.6	**100**	96.9	*low*	73.0	low	*low*
	71.0	40.0	90.5	82.1	71.1	93.0	85.7	**100**	93.4	46.2	68.3	NA	*low*
	67.9	*low*	93.1	85.7	78.9	93.0	85.7	**100**	87.0	*low*	68.3	30.3	*low*
	63.4	*low*	88.9	82.1	73.6	93.0	71.4	**100**	90.6	*low*	71.4	NA	*low*
	71.4	53.3	89.1	85.7	76.3	93.0	88.6	**100**	87.5	*low*	66.7	NA	*low*
	73.2	60.0	85.9	92.8	76.3	95.3	85.7	**100**	87.5	8/8d	77.8	30.3	*low*

(*Continued*)

Table 2. (*Continued*)

Mya f	Phylum	Class	Order	Family	Species	Common name
988	Cnidaria	Hydrozoa	Hydroida	Campanulariidae	Clytia hemisphaerica	
988	Cnidaria	Hydrozoa	Hydroida	Hydridae	Hydra vulgaris	Freshwater hydroid
						Average similarity

nor in the intron/exon splice site regions in any of the subjects.

Nine exonic GSK3β SNPs are included in the dbSNP database (Table 3), and no other exonic SNPs have been found. Detailed data is missing for some of the SNPs and some are not validated. Three of the reported SNPs are non-synonymous and 6 are synonymous. The polymorphisms in exon 2, 4 and 10 (rs72548709) have been detected in one Asian subject, while rs17183904 in exon 10 and rs72548719 in exon 11 were only detected in African Americans. All three missense SNPs were genotyped in the Monica population and three of the synonymous were genotyped to cover the all exons in which SNPs have been detected. Only SNP rs72546694 (synonymous) in exon 5 and rs34009575 (missense) in exon 9 were non-monomorph with two genotypes: the so-called wild-type and a few heterozygotes (7 for rs72546694 and 6 for rs34009575). The 6 subjects heterozygotes for the missense mutation in exon 9 were tracked down in the Central Disease and Death registers in Denmark. One have died from a bronchial neoplasma, one suffered from prostate cancer, one has a mixture of rheumatological symptoms, one has osteroporoses, and the two remaining subjects are healthy. All subjects are now 57 years old or older. Thus, the polymorphism cannot clearly be related to any specific clinical entity although the power of this conclusion is rather low.

Discussion

GSK3β is highly conserved in mammals and vertebrates with homologies on the protein level and to some lesser extent on the nucleotide level. Although the similarities to humans are somewhat less extensive the homologies extent back to the Cambrian explosion and even before. Particularly, some of exons are highly conserved way back to exoskeletonian species. The conservation of GSK3β is significant compared to average homology between species of e.g. transcription factors.[112] The only non-synomous mutation detected in the large Monica cohort could not be associated with any single disease, and in fact for two of the subjects no disease has been registered so far (August 2009). Also, no mutations were detected in the diabetes mellitus type 2 populations and the two populations (schizophenics and HIV-treated patients). The subjects in these populations were selected as the most extreme in their primary diseases arguing that these samples should present themeself with mutations in GSK3β if the kinase should in fact have any patophysiological relevans. This assumption may flawed, but nevertheless the results do not support GSK3β as a major factor in these diseases or the co-morbidity of schizophrenia or drug-induced insulin resistance in HIV-treated patients.

An interesting question is if the conservation of a gene or rather its encoded protein extents to the

Table 2. (*Continued*)

Length	Coding region only	Exon 1	Exon 2	Exon 3	Exon 4	Exon 5	Exon 6	Exon 7	Exon 8	Exon 9	Exon 10	Exon 11	Exon 12
	67.2	36.7	71.9	89.3	71.1	93.0	88.6	97.0	90.6	*low*	68.3	9/33	*low*
	67.9	*low*	82.8	85.7	76.3	81.8	82.9	**100**	83.9	*low*	61.9	NA	*low*
	70.0	46.6	86.2	87.2	73.4	92.7	87.9	97.4	84.6	39.7	70.4	31.7	43.6

In the alignment of exon 1 and 12 DNA-sequences only the protein coding parts are used.
[a]Low homology, doubtfull or short sequences, fragmented, or not annotated as GSK3$ protein.
[b]Inc, incomplete, fragmented or inconclusive nucleoteide seqeunces, or sequences covering less than 1/3 of the human sequence.
[c]Thirtytwo of the 38 amino acids in the exon were detected by assmbling several sequences. The nucleotide sequence detected spanned 96 nt of the 111 nt in the exon and 95 of these 96 were identical to the human exon.
[d]The (S)SGTGHFT is common in (2)3-9 position in the 13 amino acid sequence of eson 9. This sequence is partially identical to the tetracycline repressor protein (Transposon Tn10).
[e]NA, No sequences were detected.
[f]Mya (million of years ago) since the divergence from Homo sapiens (molcular clock estimates). The weigthed avarage of all studies as collated in TimeTree is shown.

network it is part of, here GK3β in the numerous signal transduction networks it is a part of. The conservation of selected up-stream pathway-proteins and β-cartenin were therefor evaluated (Table 4). Sequence comparisons were only done for the entire protein transcripts as it is not the scoop of this presentation to evaluate every single protein on the exon level. The conservation is substantial in the Post-Cambrian species, but rapidly decreases in the Pre-Cambrian species. The notably exceptions are Akt, β-catenin, and cAMP-dependent protein kinase (PKA), but in particular the catalytic subunit of the protein phosphatase 2 A (PPP2CA in Table 4) is conserved deep down in the Pre-Cambrian periods and even to a greater extent than GSK3β. Although Akt, β-catenin and PKA is not conserved to the same degree as GSK3β, some of central players in the major signal transduction pathways are very conserved. This is not the case for the more peripheral elements in the signal transduction pathways, but it is clear that the nidus of the GSK3β-related pathways is rooted deep in the Pre-Cambrian.

Considering the plethora of processes GSKβ is participating in and the conservation between species, suggest that GSK3β functions as a hub

Table 3. Single nucleotide polymorphisms in Homo sapiens GSK3β.

Exon	SNP-ID	AA position	Codon position	Residues	Heterozygocity[a]			
					dbSNP		Monica	Number subjects
2	rs34002644	65	3	Gly/Gly	0.176	Asian	0	2688
4	rs72546695	152	2	Met/Thr	0.004	Asian	0	2719
	rs17183890	152	3	Thr/Thr	0.004	Asian	ND	
5	rs72546694	184	3	Pro/Pro	0.006	3/124 white	0.0026	2696
9	**rs34009575**	**310**	**3**	**Gln/His**	ND[b]	No details	**0.0022**	**2704**
10	rs72548709	332	2	Arg/His	0.002	Asian	0	2705
	rs17183904	351	3	Ser/Ser	ND	No details	ND	
	rs17183897	370	3	Pro/Pro	0.012	African Americans	ND	
11	rs72548719	410	3	Ala/Ala	0.011	African Americans	0	2655

[a]Hetrozygoty is for the entire study populations.
[b]ND, not determined.

Table 4. Conservation of selected genes related to GSK3β function.

						PKB/Akt pathway		
						PI3K		
Phylum	**Class**	**Order**	**Species**	**Common name**	**GSK3β**	**IRS1**	**p85**	**p110**
Chordata	Mammalia	Primates	Homo sapiens	Number of amino acids	433	1242	732	1068
				Exon polymorphisms[d]	9/3	32/18	16/8	29/13
				Non-exonic polymorphisms[e]	994	280	471	379
				Non-synonomous subs./amino acid %[f]	0.692841	1.449275	1.092896	1.217228
Chordata	Mammalia	Primates	Pan troglodytes	Chimpanzee	98.2	99	98	100
Chordata	Mammalia	Primates	Macaca mulatta	Rhesus monkey	92.2	96	97	100
Chordata	Mammalia	Primates	Macaca fascicularis	Cynomolgus Monkey	72.1		54	
Chordata	Mammalia	Artiodactyla	Bos taurus	Cow	99.0	90	95	99
Chordata	Mammalia	Artiodactyla	Ovis aries	Domestic sheep	94.0			
Chordata	Mammalia	Artiodactyla	Sus scrofa	Wild Boar or hog	95.6	92	11	34
Chordata	Mammalia	Carnivora	Canis lupus familiaris	Dog	99.8	86	94	99
Chordata	Mammalia	Perissodactyla	Equus callabus	Horse	99.5	91	96	99
Chordata	Mammalia	Rodentia	Rattus novegicus	Rat	99.1	88	93	98
Chordata	Mammalia	Rodentia	Mus musculus	Mouse	99.1	88	94	98
Chordata	Mammalia	Rodentia	Spermophilus citellus	Squirrel	96.3			
Chordata	Mammalia	Didelphimorphia	Monodelphis domestica	Opossum	99.5	76	89	97
Chordata	Mammalia	Monotremata	Ornithorhynchus anatinus	Platypus	56.1	86	60	97
Chordata	Aves	Galliformes	Gallus gallus	Chicken	96.7	90	87	96
Chordata	Aves	Passeriformes	Taeniopygia guttata	Zebra finch	97.0	30	87	34
Chordata	Amphibia	Anura	Xenopus laevis	Frog	88.0	42	79	71
Chordata	Actinopterygii	Tetraodontiformes	Tetraodon nigroviridis	Pufferfish	70.4	48	53	85
Chordata	Actinopterygii	Cypriniformes	Danio rerio	Zebra fish	89.8	55	77	92
Pre-Cambrian								
Chordata	Ascidiacea	Phlebobranchia	Ciona intestinalis	Vase tunicate	73.9	11	24	35
Chordata	Cephalochordata	Branchiostoma floridae	Florida lancelet		79.4	16	45	61
Platyhelminthes	Turbellaria	Tricladida	Dugesia japonica	Flatworm	63.0			
Echinodermata	Echinoidea	Temnopleuroida	Lytechinus variegatus	Green sea urchin	71.6			
Echinodermata	Echinoidea	Echinoida	Paracentrotus lividus	Sea urchin	71.4			
Echinodermata	Echinoidea	Echinoida	Strongylocentrotus purpuratus	Purple sea urchin	71.1		43	36
Hemichordata	Enteropneusta		Saccoglossus kowalevskii	Worm	78.5			
Nematoda	Secemetea	Spirurida	Brugia malayi	Roundworm (nematode)	57.5		18	32
Nematoda	Secemetea	Rhabditida	Caenorhabditis elegans	Worm	77.7		23	29
Arthropoda	Insecta	Diptera	Drosophila melanogaster	Fruit fly	75.7	8	19	35
Arthropoda	Insecta	Diptera	Anopheles gambiae	African maliara mosquito	83.6		20	35

		Wnt-pathway					Miscelaneous kinases, catalytic subunits				
PDK1	**PKB/Akt**	**Dishevelled**	**Axin**	**APC**	**β-catenin**	**FRAT1/GBP**	**PKA**	**SGK**	**PPP2CA**	**p 70 S6K**	**p 90 S6K**
556	480	695	862	2843	781	279	351	431	309	525	735
1/1	20/11	19/11	13/7	44/23	16/10	3/2	7/5	10/7	3/2	3/1	20/12
107	289	56	494	513	268	7	133	57	240	285	235
0.179856	2.291667	1.58273381	0.812065	0.809005	1.28041	0.71684588	1.424501	1.62413	0.647249	0.190476	1.632653
?	*81?*	96	99[a]	99	*100*	98	*98*	99	*99*	100	81
?	*99*	62	98[c]	30	*89*	97	*92*	99	*99*	94	72
98	*?*	94			*99*		*42*	99	*100*	100	76
96	*96*	93	85	94	*99*	81	*99*	96	*99*	99	99
?	*96*			30	*?*		*99*	48		29	?
94	*97*				*99*		*99*	72	*100*	62	?
94	*96*	95	86	92	*99*	61	*98*	97	*99*	99	98
?	*97*	81	87	95	*99*		*97*	97	*97*	99	96
95	*98*	94	83	89	*99*	70	*98*	97	*99*	99	97
94	*98*	94	85	90	*99*	75	*98*	96	*99*	99	98
?											
62	*95*	84	78	86	*99*		*94*	97	*100*	97	93
51	*78*	45	43	86	*99*		*93*	95	*92*	98	83
90	*96*	79	77[b]	89	*9*		*93*	95	*99*	97	89
90	*96*	75	75[b]	83	*99*		*93*	95	*99*	97	82
67	*93*	66	69	74	*97*	35	*93*	88	*98*	92	85
79	*88*	62	59	43	*92*		*94*	76	*99*	96	81
77	*86*	70	67[b]	66	*97*	26	*93*	86	*97*	84	81
41	*66*	37	22	18	*69*		*83*	83	*93*	62	70
48	*?*	25	44	15	*72*		*88*	59	*94*	58	69
		26									
		48	36		*69*						
55	*64*	27	36	32	*70*		*78*	66	*94*	69	44
	70	51	40		*75*		*43*	52	*41*	29	31
38	*60*	33			*39*		*41*	41	*88*	65	43
36	*59*	33			*28*		*85*	52	*89*	46	61
37	*62*	36		13	*67*		*82*	56	*93*	51	56
39	*39*	29			*69*		*72*	43	*93*	51	60

(*Continued*)

Table 4. (*Continued*)

						PKB/Akt pathway		
						PI3K		
Phylum	**Class**	**Order**	**Species**	**Common name**	**GSK3β**	**IRS1**	**p85**	**p110**
Arthropoda	Insecta	Diptera	Culex quinquefasciatus	Mosquito	67.4	5	23	34
Arthropoda	Insecta	Diptera	Aedes aegypti	Yellow fever mosquito	70.9		22	35
Arthropoda	Insecta	Diptera	Glossina morsitans morsitans	Tsetse fly	42.3			
Arthropoda	Insecta	Hymenoptera	Apis mellifera	European honey bee	72.1	9	36	38
Arthropoda	Insecta	Anoplura	Pediculus humanus corporis	Lice	70.7	9	37	37
Arthropoda	Insecta	Lepidoptera	Danaus plexippus	Butterfly	72.1			
Arthropoda	Insecta	Hemiptera	Acyrthosiphon pisum	Endosymbionts of insects	71.0		25	37
Arthropoda	Insecta	Coleoptera	Tribolium castaneum	Red flour beetle	67.9		38	23
Arthropoda	Arachnida	Parasitiformes	Ixodes scapularis	Deer tick	63.4		25	22
Arthropoda	Arachnida	Parasitiformes	Rhipicephalus microplus	Cattle tick	71.4			
Cnidaria	Anthozoa	Actiniaria	Nematostella vectensis	Starlet sea anemone	73.2	7	6	42
Cnidaria	Hydrozoa	Hydroida	Clytia hemisphaerica		67.2			
Cnidaria	Hydrozoa	Hydroida	Hydra vulgaris	Freshwater hydroid	67.9			

integrating several metabolic networks. Hubs in networks are extremely fragile and sensitive to perturbations,[113–115] so at least heuristically it can be understood that functional mutations in GSK3β are strongly selected against. Recently a −50T/C transition in the promoter has bee suggested to be related to the onset of bipolar disease, but not the disease itself,[116] i.e. the coding sequence are preserved while variations may exist in the regulatory sequences. This would support the theory of facilitated variation put forward by Gerhart and Kirschner,[117] which suggest that the major genetic evolution and differences between species is caused by differences in regulatory structures rather than in protein-coding sequences, at least for GSK3β. It should be noted that polymorphisms outside the coding regions of GSK3β including the introns runs in the thousands, supporting the Gerhart-Kirschner theory (although the species-specific variations has not been examined).

The results presented here shows that ultraconserved proteins is in fact present, but the extent and number of ultraconserved proteins is not known.

Nearly 500 absolute conserved elements defined as segments of more than 200 bp in length with 100% identity (and no insertions or deletions) have been detected in the human and rodents (rat and mouse) genomes. These ultraconserved elements have in addition been found to be extremely, although not completely, conserved in chicken and dog (95% and 99%, respectively).[118] The elements are distributed in all human chromosomes except chromosome 21 and Y. One hundred and eleven of the 481 detected ultraconserved elements are partly overlapping exonic protein-coding regions, 256 are

Table 4. (*Continued*)

		Wnt-pathway					Miscelaneous kinases, catalytic subunits				
PDK1	**PKB/Akt**	**Dishevelled**	**Axin**	**APC**	**β-catenin**	**FRAT1/GBP**	**PKA**	**SGK**	**PPP2CA**	**p 70 S6K**	**p 90 S6K**
40	*64*	38	17	8	*66*		*81*	56	*78*	50	58
41	*65*	41	16		*70*		*82*	47	*93*	53	59
	70		26				*82*				
51	*65*	46	27	10	*70*			47	*96*	52	64
47	*63*	47	26	10	*71*		*83*	59	*96*	57	69
							83				
46	*66*	43	13	9	*66*		*49*	56	*95*	51	67
45	*64*	52	30	9	*53*			52	*95*	49	61
43	*?*	42	23		*69*		*54*	51	*94*	29	29
36	*38*	50	24		*69*		*54*	66	*91*	57	71
	52	43			*58*		*39*	41		25	

Empty fields denotes less than 5% of similarity, small stretches 3 amino acids of less in length or no sequnece was available.
The Question marks denotes ambiguities in the interpretation of the detected sequences or unexpected low similarities.
[a]Exon 10 is not included but several clones include exon 10 with modest similarity to the human form.
[b]Exon 8 is not detected.
[c]Only exon 9.
[d]Polymorphisms in exons (all/non-synonymous amino acid substitutions).
[e]Polymorphisms in intronsd and 5'- and 3'-UTR areas as collated in dbSNP.
[f]Rate of non-synonymous substitutions pr amino acid in the protein.

non-exonic, while the classification of the remaining elements is inconclusive. The non-exonic elements are located in introns and integenic regions and tends to cluster close to transcription factors and early developmental genes. The exonic elements are highly enriched in RNA binding, splicing regulating, and DNA binding genes, while the non-exonic elements are particularly enriched in DNA binding genes including transcription factors and homebox genes. In contrast to the extreme conservation of these elements in vertebrates, only 5% of the ultra-conserved elements could partially be traced back to *Ciona intestinalis*, *Drosophila melanogaster*, or *Caenorhabditis elegans*.[118] It should be noted that the ultraconserved elements in[1] are ultraconserved nucleotide sequences including only parts of exonic elements and not full genes. The results presented here extends the concept of kernel genes[1] not only to be transcription factors but also to include essential and highly preserved signal transduction pathways and entire genes.

In conclusion, GSK3β is ultraconserved,[118] and the functional diverse and essential modular processes GSK3β is part of, would suggest that GSK3β (and PPP2CA) qualifies to be a member of the probably very exclusive club of functional kernel structures. The data argue that the ultraconserved GSK3β and particular some of the exons are under a strong purifying selection pressure, and are functionally highly constrained as also suggested for the ultraconserved DNA elements.[119]

Disclosures

The authors report no conflicts of interest.

References

1. Davidson EH, Erwin DH. Gene regulatory networks and the evolution of animal body plans. *Science*. 2006;311:796–800.
2. Lee J, Kim MS. The role of GSK3 in glucose homeostasis and the development of insulin resistance. *Diabetes Res Clin Pract*. 2007;77 Suppl 1: S49–57.
3. Lau KF, Miller CC, Anderton BH, Shaw PC. Expression analysis of glycogen synthase kinase-3 in human tissues. *J Pept Res*. 1999;54:85–91.
4. Hughes K, Nikolakaki E, Plyte SE, Totty NF, Woodgett JR. Modulation of the glycogen synthase kinase-3 family by tyrosine phosphorylation. *EMBO J*. 1993;12:803–8.
5. Cohen P, Frame S. The renaissance of GSK3. *Nat Rev Mol Cell Biol*. 2001;2: 769–76.
6. Hartigan JA, Johnson GV. Transient increases in intracellular calcium result in prolonged site-selective increases in Tau phosphorylation through a glycogen synthase kinase 3beta-dependent pathway. *J Biol Chem*. 1999;274: 21395–401.
7. Dominguez I, Green JB. Missing links in GSK3 regulation. *Dev Biol*. 2001; 235:303–13.
8. Frame S, Cohen P. GSK3 takes centre stage more than 20 years after its discovery. *Biochem J*. 2001;359:1–16.
9. Roztocil E, Nicholl SM, Davies MG. Mechanisms of sphingosine-1-phosphate-induced akt-dependent smooth muscle cell migration. *Surgery*. 2009;145:34–41.
10. Morales A, Lee H, Goni FM, Kolesnick R, Fernandez-Checa JC. Sphingolipids and cell death. *Apoptosis*. 2007;12:923–39.
11. Cadigan KM, Liu YI. Wnt signaling: complexity at the surface. *J Cell Sci*. 2006;119:395–402.
12. Woodgett JR. Judging a protein by more than its name: GSK-3. *Sci STKE*. 2001:RE12-GSK3.
13. Thomas GM, Frame S, Goedert M, Nathke I, Polakis P, Cohen P. A GSK3-binding peptide from FRAT1 selectively inhibits the GSK3-catalysed phosphorylation of axin and beta-catenin. *FEBS Lett*. 1999;458:247–51.
14. Cohen MM Jr. The hedgehog signaling network. *Am J Med Genet*. 2003; 123A:5–28.
15. Ingham PW. Hedgehog signaling: a tale of two lipids. *Science*. 2001;294: 1879–81.
16. Doble BW, Woodgett JR. GSK-3: tricks of the trade for a multi-tasking kinase. *J Cell Sci*. 2003;116:1175–86.
17. Schweisguth F. Notch signaling activity. *Curr Biol*. 2004;14:R129–38.
18. Fortini ME. Notch signaling: the core pathway and its posttranslational regulation. *Dev Cell*. 2009;16:633–47.
19. Fortini ME. Gamma-secretase-mediated proteolysis in cell-surface-receptor signalling. *Nat Rev Mol Cell Biol*. 2002;3:673–84.
20. Mumm JS, Schroeter EH, Saxena MT, et al. A ligand-induced extracellular cleavage regulates gamma-secretase-like proteolytic activation of Notch1. *Mol Cell*. 2000;5:197–206.
21. Brou C, Logeat F, Gupta N, et al. A novel proteolytic cleavage involved in Notch signaling: the role of the disintegrin-metalloprotease TACE. *Mol Cell*. 2000;5:207–16.
22. Moro MA, Hurtado O, Cardenas A, et al. Expression and function of tumour necrosis factor-alpha-converting enzyme in the central nervous system. *Neurosignals*. 2003;12:53–8.
23. Allinson TM, Parkin ET, Turner AJ, Hooper NM. ADAMs family members as amyloid precursor protein alpha-secretases. *J Neurosci Res*. 2003;74:342–52.
24. Nickoloff BJ, Osborne BA, Miele L. Notch signaling as a therapeutic target in cancer: a new approach to the development of cell fate modifying agents. *Oncogene*. 2003;22:6598–608.
25. Maillard I, Pear WS. Notch and cancer: best to avoid the ups and downs. *Cancer Cell*. 2003;3:203–5.
26. Lai EC. Keeping a good pathway down: transcriptional repression of Notch pathway target genes by CSL proteins. *EMBO Rep*. 2002;3:840–5.
27. Baron M. An overview of the Notch signalling pathway. *Semin Cell Dev Biol*. 2003;14:113–9.
28. McKenzie G, Ward G, Stallwood Y, et al. Cellular Notch responsiveness is defined by phosphoinositide 3-kinase-dependent signals. *BMC Cell Biol*. 2006; 7:10.
29. Kobayashi T, Cohen P. Activation of serum- and glucocorticoid-regulated protein kinase by agonists that activate phosphatidylinositide 3-kinase is mediated by 3-phosphoinositide-dependent protein kinase-1 (PDK1) and PDK2. *Biochem J*. 1999;339(Pt 2):319–28.
30. Failor KL, Desyatnikov Y, Finger LA, Firestone GL. Glucocorticoid-induced degradation of glycogen synthase kinase-3 protein is triggered by serum- and glucocorticoid-induced protein kinase and Akt signaling and controls beta-catenin dynamics and tight junction formation in mammary epithelial tumor cells. *Mol Endocrinol*. 2007;21:2403–15.
31. Cortes JC, Montalvo EA, Muniz J, et al. Dp71f modulates GSK3-beta recruitment to the beta1-integrin adhesion complex. *Neurochem Res*. 2009;34: 438–44.
32. Fang X, Yu SX, Lu Y, Bast RC Jr, Woodgett JR, Mills GB. Phosphorylation and inactivation of glycogen synthase kinase 3 by protein kinase A. *Proc Natl Acad Sci U S A*. 2000;97:11960–5.
33. Frodin M, Gammeltoft S. Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction. *Mol Cell Endocrinol*. 1999;151:65–77.
34. Giampietri C, Petrungaro S, Musumeci M, et al. c-Flip overexpression reduces cardiac hypertrophy in response to pressure overload. *J Hypertens*. 2008;26:1008–16.
35. Klinghoffer RA, Frazier J, Annis J, et al. A lentivirus-mediated genetic screen identifies dihydrofolate reductase (DHFR) as a modulator of beta-catenin/GSK3 signaling. *PLoS One*. 2009;4:e6892.
36. Liu Z, Choi SW, Crott JW, et al. Mild Depletion of Dietary Folate Combined with Other B Vitamins Alters Multiple Components of the Wnt Pathway in Mouse Colon. *J Nutr*. 2007;137:2701–8.
37. Goni-Oliver P, Lucas JJ, Avila J, Hernandez F. N-terminal cleavage of GSK-3 by calpain: a new form of GSK-3 regulation. *J Biol Chem*. 2007;282:22406–13.
38. Ferrer JC, Favre C, Gomis RR, et al. Control of glycogen deposition. *FEBS Letters*. 2003;546:127–32.
39. van Weeren PC, de Bruyn KM, Vries-Smits AM, van Lint J, Burgering BM. Essential role for protein kinase B (PKB) in insulin-induced glycogen synthase kinase 3 inactivation. Characterization of dominant-negative mutant of PKB. *J Biol Chem*. 1998;273:13150–6.
40. Eldar-Finkelman H. Glycogen synthase kinase 3: an emerging therapeutic target. *Trends Mol Med*. 2002;8:126–32.
41. Pelech S, Jelinkova L, Susor A, et al. Antibody microarray analyses of signal transduction protein expression and phosphorylation during porcine oocyte maturation. *J Proteome Res*. 2008;7:2860–71.
42. Baluch DP, Capco DG. GSK3 beta mediates acentromeric spindle stabilization by activated PKC zeta. *Dev Biol*. 2008;317:46–58.
43. Ferkey DM, Kimelman D. GSK-3: new thoughts on an old enzyme. *Dev Biol*. 2000;225:471–9.
44. Garrido JJ, Simon D, Varea O, Wandosell F. GSK3 alpha and GSK3 beta are necessary for axon formation. *FEBS Lett*. 2007;581:1579–86.
45. Trivedi N, Marsh P, Goold RG, Wood-Kaczmar A, Gordon-Weeks PR. Glycogen synthase kinase-3{beta} phosphorylation of MAP1B at Ser1260 and Thr1265 is spatially restricted to growing axons. *J Cell Sci*. 2005;118: 993–1005.
46. Day TF, Guo X, Garrett-Beal L, Yang Y. Wnt/beta-catenin signaling in mesenchymal progenitors controls osteoblast and chondrocyte differentiation during vertebrate skeletogenesis. *Dev Cell*. 2005;8:739–50.
47. Hill TP, Spater D, Taketo MM, Birchmeier W, Hartmann C. Canonical Wnt/beta-catenin signaling prevents osteoblasts from differentiating into chondrocytes. *Dev Cell*. 2005;8:727–38.
48. Zhu M, Tang D, Wu Q, et al. Activation of beta-catenin signaling in articular chondrocytes leads to osteoarthritis-like phenotype in adult beta-catenin conditional activation mice. *J Bone Miner Res*. 2009;24:12–21.
49. Wu Q, Huang JH, Sampson ER, et al. Smurf2 induces degradation of GSK-3 beta and upregulates beta-catenin in chondrocytes: a potential mechanism for Smurf2-induced degeneration of articular cartilage. *Exp Cell Res*. 2009; 315:2386–98.
50. Blankesteijn WM, van de Schans VAM, ter Horst P, Smits JFM. The Wnt/frizzled/GSK-3[beta] pathway: a novel therapeutic target for cardiac hypertrophy. *Trends in Pharmacological Sciences*. 2008;29: 175–80.

51. Haq S, Michael A, Andreucci M, et al. Stabilization of beta-catenin by a Wnt-independent mechanism regulates cardiomyocyte growth. *Proc Natl Acad Sci U S A*. 2003;100:4610–5.
52. Badorff C, Ruetten H, Mueller S, et al. Fas receptor signaling inhibits glycogen synthase kinase 3 beta and induces cardiac hypertrophy following pressure overload. *J Clin Invest*. 2002;109:373–81.
53. Hardt SE, Sadoshima J. Glycogen synthase kinase-3beta: a novel regulator of cardiac hypertrophy and development. *Circ Res*. 2002;90:1055–63.
54. Lee HC, Tsai JN, Liao PY, et al. Glycogen synthase kinase 3 alpha and 3 beta have distinct functions during cardiogenesis of zebrafish embryo. *BMC Dev Biol*. 2007;7:93.
55. Markou T, Cullingford TE, GiraldoA, et al. Glycogen synthase kinases 3 alpha and 3 beta in cardiac myocytes: regulation and consequences of their inhibition. *Cell Signal*. 2008;20:206–18.
56. Pap M, Cooper GM. Role of glycogen synthase kinase-3 in the phosphatidylinositol 3-Kinase/Akt cell survival pathway. *J Biol Chem*. 1998;273: 19929–32.
57. Hongisto V, Vainio JC, Thompson R, Courtney MJ, Coffey ET. The Wnt pool of glycogen synthase kinase 3beta is critical for trophic-deprivation-induced neuronal death. *Mol Cell Biol*. 2008;28:1515–27.
58. Jope RS, Bijur GN. Mood stabilizers, glycogen synthase kinase-3beta and cell survival. *Mol Psychiatry*. 2002;7 Suppl 1:S35–45.
59. Hoeflich KP, Luo J, Rubie EA, Tsao MS, Jin O, Woodgett JR. Requirement for glycogen synthase kinase-3beta in cell survival and NF-kappaB activation. *Nature*. 2000;406:86–90.
60. Pap M, Cooper GM. Role of translation initiation factor 2B in control of cell survival by the phosphatidylinositol 3-kinase/Akt/glycogen synthase kinase 3beta signaling pathway. *Mol Cell Biol*. 2002;22:578–86.
61. Somervaille TC, Linch DC, Khwaja A. Growth factor withdrawal from primary human erythroid progenitors induces apoptosis through a pathway involving glycogen synthase kinase-3 and Bax. *Blood*. 2001;98: 1374–81.
62. Scalia P, Heart E, Comai L, Vigneri R, Sung CK. Regulation of the Akt/Glycogen synthase kinase-3 axis by insulin-like growth factor-II via activation of the human insulin receptor isoform-A. *J Cell Biochem*. 2001;82: 610–8.
63. Bijur GN, De Sarno P, Jope RS. Glycogen synthase kinase-3beta facilitates staurosporine- and heat shock-induced apoptosis. Protection by lithium. *J Biol Chem*. 2000;275:7583–90.
64. Weihl CC, Ghadge GD, Kennedy SG, Hay N, Miller RJ, Roos RP. Mutant presenilin-1 induces apoptosis and downregulates Akt/PKB. *J Neurosci*. 1999;19:5360–9.
65. Li X, Bijur GN, Jope RS. Glycogen synthase kinase-3beta, mood stabilizers, and neuroprotection. *Bipolar Disord*. 2002;4:137–44.
66. Bijur GN, Jope RS. Proapoptotic stimuli induce nuclear accumulation of glycogen synthase kinase-3 beta. *J Biol Chem*. 2001;276:37436–42.
67. Turenne GA, Price BD. Glycogen synthase kinase3 beta phosphorylates serine 33 of p53 and activates p53‘s transcriptional activity. *BMC Cell Biol*. 2001;2:12.
68. Watcharasit P, Bijur GN, Zmijewski JW, et al. Direct, activating interaction between glycogen synthase kinase-3beta and p53 after DNA damage. *Proc Natl Acad Sci U S A*. 2002;99:7951–5.
69. Watcharasit P, Bijur GN, Song L, Zhu J, Chen X, Jope RS. Glycogen synthase kinase-3beta (GSK3beta) binds to and promotes the actions of p53. *J Biol Chem*. 2003;278:48872–9.
70. Zmijewski JW, Jope RS. Nuclear accumulation of glycogen synthase kinase-3 during replicative senescence of human fibroblasts. *Aging Cell*. 2004;3:309–17.
71. Lochhead PA, Coghlan M, Rice SQ, Sutherland C. Inhibition of GSK-3 selectively reduces glucose-6-phosphatase and phosphatase and phosphoenolypyruvate carboxykinase gene expression. *Diabetes*. 2001;50: 937–46.
72. Kozlovsky N, Belmaker RH, Agam G. GSK-3 and the neurodevelopmental hypothesis of schizophrenia. *Eur Neuropsychopharmacol*. 2002;12: 13–25.
73. Kaytor MD, Orr HT. The GSK3 beta signaling cascade and neurodegenerative disease. *Curr Opin Neurobiol*. 2002;12:275–8.
74. Bhat RV, Budd SL. GSK3beta signalling: casting a wide net in Alzheimer‘s disease. *Neurosignals*. 2002;11:251–61.
75. Agarwal-Mawal A, Qureshi HY, et al. 14-3-3 connects glycogen synthase kinase-3 beta to tau within a brain microtubule-associated tau phosphorylation complex. *J Biol Chem*. 2003;278:12722–8.
76. Adachi A, Kano F, Saido TC, Murata M. Visual screening and analysis for kinase-regulated membrane trafficking pathways that are involved in extensive beta-amyloid secretion. *Genes Cells*. 2009;14:355–69.
77. Takashima A, Murayama M, Murayama O, et al. Presenilin 1 associates with glycogen synthase kinase-3beta and its substrate tau. *Proc Natl Acad Sci U S A*. 1998;95:9637–41.
78. Tesco G, Tanzi RE. GSK3 beta forms a tetrameric complex with endogenous PS1-CTF/NTF and beta-catenin. Effects of the D257/D385A and FAD-linked mutations. *Ann N Y Acad Sci*. 2000;920:227–32.
79. Zhao WQ, Townsend M. Insulin resistance and amyloidogenesis as common molecular foundation for type 2 diabetes and Alzheimer‘s disease. *Biochim Biophys Acta*. 2009;1792:482–96.
80. Wei J, Hemmings GP. The NOTCH4 locus is associated with susceptibility to schizophrenia. *Nat Genet*. 2000;25:376–7.
81. Beaulieu JM, Zhang X, Rodriguiz RM, et al. Role of GSK3 beta in behavioral abnormalities induced by serotonin deficiency. *Proc Natl Acad Sci U S A*. 2008;105:1333–8.
82. Beaulieu JM, Gainetdinov RR, Caron MG. Akt/GSK3 signaling in the action of psychotropic drugs. *Annu Rev Pharmacol Toxicol*. 2009;49: 327–47.
83. Kozlovsky N, Belmaker RH, Agam G. Low GSK-3 activity in frontal cortex of schizophrenic patients. *Schizophr Res*. 2001;52:101–5.
84. Cotter D, Kerwin R, al Sarraji S, et al. Abnormalities of Wnt signalling in schizophrenia—evidence for neurodevelopmental abnormality. *Neuroreport*. 1998;9:1379–83.
85. Beasley C, Cotter D, Everall I. An investigation of the Wnt-signalling pathway in the prefrontal cortex in schizophrenia, bipolar disorder and major depressive disorder. *Schizophr Res*. 2002;58:63–7.
86. Shakoori A, Ougolkov A, Yu ZW, et al. Deregulated GSK3beta activity in colorectal cancer: its association with tumor cell survival and proliferation. *Biochem Biophys Res Commun*. 2005;334:1365–73.
87. Luo J. Glycogen synthase kinase 3beta (GSK3beta) in tumorigenesis and cancer chemotherapy. *Cancer Lett*. 2009;273:194–200.
88. Manoukian AS, Woodgett JR. Role of glycogen synthase kinase-3 in cancer: regulation by Wnts and other signaling pathways. *Adv Cancer Res*. 2002;84:203–29.
89. An J, Yang DY, Xu QZ, et al. DNA-dependent protein kinase catalytic subunit modulates the stability of c-Myc oncoprotein. *Mol Cancer*. 2008; 7:32.
90. Korkaya H, Paulson A, Charafe-Jauffret E, et al. Regulation of mammary stem/progenitor cells by PTEN/Akt/beta-catenin signaling. *PloS Biol*. 2009;7:e1000121.
91. Hosgood HD, Menashe I, Shen M, et al. Pathway-based evaluation of 380 candidate genes and lung cancer susceptibility suggests the importance of the cell cycle pathway. *Carcinogenesis*. 2008;29:1938–43.
92. Hedges SB, Dudley J, Kumar S. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics*. 2006;22:2971–2.
93. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res*. 2009;37:D26–31.
94. Sayers EW, Barrett T, Benson DA, et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res*. 2009;37: D5–15.
95. Kumar S, Filipski A, Swarna V, Walker A, Hedges SB. Placing confidence limits on the molecular age of the human-chimpanzee divergence. *Proc Natl Acad Sci U S A*. 2005;102:18842–7.
96. Woodburne MO, Rich TH, Springer MS. The evolution of tribospheny and the antiquity of mammalian clades. *Mol Phylogenet Evol*. 2003;28:360–85.
97. Blair JE, Hedges SB. Molecular phylogeny and divergence times of deuterostome animals. *Mol Biol Evol*. 2005;22:2275–84.
98. Blair JE, Shah P, Hedges SB. Evolutionary sequence analysis of complete eukaryote genomes. *BMC Bioinformatics*. 2005;6:53.

99. Peterson KJ, Cotton JA, Gehling JG, Pisani D. The Ediacaran emergence of bilaterians: congruence between the genetic and the geological fossil records. *Philos Trans R Soc Lond B Biol Sci*. 2008;363:1435–43.
100. Altschul SF, Madden TL, Schaffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25:3389–402.
101. Elias AN, Hofflich H. Abnormalities in glucose metabolism in patients with schizophrenia treated with atypical antipsychotic medications. *Am J Med*. 2008;121:98–104.
102. Haugaard SB, Andersen O, Storgaard H, et al. Insulin secretion in lipodystrophic HIV-infected patients is associated with high levels of nonglucose secretagogues and insulin resistance of beta-cells. *Am J Physiol Endocrinol Metab*. 2004;287:E677–85.
103. Haugaard SB, Andersen O, Dela F, et al. Defective glucose and lipid metabolism in human immunodeficiency virus-infected patients with lipodystrophy involve liver, muscle tissue and pancreatic beta-cells. *Eur J Endocrinol*. 2005;152:103–12.
104. Jorgensen T. Prevalence of gallstones in a Danish population. *Am J Epidemiol*. 1987;126:912–21.
105. Rasmussen SL, Torp-Pedersen C, Borch-Johnsen K, Ibsen H. Normal values for ambulatory blood pressure and differences between casual blood pressure and ambulatory blood pressure: results from a Danish population survey. *J Hypertens*. 1998;16:1415–24.
106. Tunstall-Pedoe H, Kuulasmaa K, Mahonen M, Tolonen H, Ruokokoski E, Amouyel P. Contribution of trends in survival and coronary-event rates to changes in coronary heart disease mortality: 10-year results from 37 WHO MONICA project populations. Monitoring trends and determinants in cardiovascular disease. *Lancet*. 1999;353:1547–57.
107. Swalla BJ, Smith AB. Deciphering deuterostome phylogeny: molecular, morphological and palaeontological perspectives. *Philos Trans R Soc Lond B Biol Sci*. 2008;363:1557–68.
108. Putnam NH, Butts T, Ferrier DEK, et al. The amphioxus genome and the evolution of the chordate karyotype. *Nature*. 2008;453:1064–71.
109. Killian JK, Buckley TR, Stewart N, Munday BL, Jirtle RL. Marsupials and Eutherians reunited: genetic evidence for the Theria hypothesis of mammalian evolution. *Mamm Genome*. 2001;12:513–7.
110. Roger AJ, Hug LA. The origin and diversification of eukaryotes: problems with molecular phylogenetics and molecular clock estimation. *Philos Trans R Soc Lond B Biol Sci*. 2006;361:1039–54.
111. Kumar S. Molecular clocks: four decades of evolution. *Nat Rev Genet*. 2005;6:654–62.
112. Odom DT, Dowell RD, Jacobsen ES, et al. Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat Genet*. 2007;39:730–2.
113. He X, Zhang J. Why do hubs tend to be essential in protein networks? *PloS Genet*. 2006;2:e88.
114. Lu X, Jain VV, Finn PW, Perkins DL. Hubs in biological interaction networks exhibit low changes in expression in experimental asthma. *Mol Syst Biol*. 2007;3:98–103.
115. Yu H, Kim PM, Sprecher E, Trifonov V, Gerstein M. The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Comput Biol*. 2007;3:713–20.
116. Benedetti F, Bernasconi A, Lorenzi C, et al. A single nucleotide polymorphism in glycogen synthase kinase 3-beta promoter gene influences onset of illness in patients affected by bipolar disorder. *Neurosci Lett*. 2004;355:37–40.
117. Gerhart J, Kirschner M. The theory of facilitated variation. *Proc Natl Acad Sci U S A*. 2007;104 Suppl 1:8582–9.
118. Bejerano G, Pheasant M, Makunin I, et al. Ultraconserved elements in the human genome. *Science*. 2004;304:1321–5.
119. Katzman S, Kern AD, Bejerano G, et al. Human genome ultraconserved elements are ultraselected. *Science*. 2007;317:915.

Supporting Material

Table S1. Primer sequences for PCR of GSK3β (HomoSapiens).

	Forward primer	Reverse primer
Exon 1	5′-TTTAGGATTTGTCCTCTC-3′	5′-AGATCCTGGTATTTCGAGGTT-3′
Exon 2	5′-ACCTTGTATATTTTGGCAGTT-3′	5′-AAAATGCGCACAATAGA-3′
Exon 3	5′-CATCTTGTTCAATATGATGA-3′	5′-AAACTGTCTATGAGCG-3′
Exon 4	5′-TTTAAAAATCCGAATAAG-3′	5′-AGTTTTTCTCCCTGAGTATTA-3′
Exon 5	5′-TCTGTTACTTTATTGACTGCT-3′	5′-AATTACTGTGTTTAGGCT-3′
Exon 6	5′-GATACTGTGAAAGGATAGCAGTC-3′	5′-AACATTATGGGATATATCCA-3′
Exon 7	5′-ATTGTATAATCTTTTCCTGTGAT-3′	5′-AAAAATTCCTTCCAGTTCAA-3′
Exon 8	5′-CCTGAAAACATGGAGACTAAAT-3′	5′-ATACAACTATCTTTACATGCAGA
Exon 9	5′-CCACACCATGGATTTCTTATAAC-3′	5′-TGAAACCTATGCCATACAGT-3′
Exon 10	5′-ACAGGGTATTTGCTTAGATTCTT-3′	5′-GTACTTAATATGTCCGTTTTTG-3′
Exon 11	5′-CCCGGCATAAACTGGTAGT-3′	5′-ACACCTGGCTAAGTTTTT-3′
Exon 12	5′-CTTGGAAGGCCATACAGT-3′	5′-CTTTTTAATATTCTTTCCAAACG-3

Table S2. Alignment of amino acid sequences of vertebrates for stretches with less than absolute similarity.

Exon 1	
Homo sapiens	MSGRPRTTSFAESCKPVQQPSAFGSMKVSR
Xenopus laevis	MSGRPRTTSFAESCKPVQQPSSFGSMKVSR
Ornithorhynchus	SERPRTPSFGES
Tetraodon	MSGRPRTTSFAESCKPVPQPSAFGSMKVSR
Danio rerio	MSGRPRTTSFAESCKPVPQPSAFGSMKVSR
Exon 2	
Homo sapiens	DKDGSKVTTVVATPGQGPDRPQEVSYTDTKVIGNGSFGVVYQAKLCDSGEL
Xenopus	DKDGSKVTTVVATPGQGPDRQQEVTYTDTKVIGNGSFGVVYQAKLCDSGEL
Macaca fascic	DRDSGKVTTVVATLGQGPERSQEVAYTD I KVIGNGSFGVVYQARLAETREL
Homo sapiens	VAIKKVLQDKRFK
Xenopus	VAIKKVPQDKRFK
Macaca fascic	VAIKKVLQDKRFK
Exon 3	
No dissimilarities	
Exon 4	
Homo sapiens	KDEVYLNLVLDYVPETVYRVARHYSRAKQTLPVIYVK
Pan troglodytes	KDELYLNLVLDYVPETVYRVARHYSRAKLT I PILYIK
Macac fascic	KDELYLNLVLEYVPETVYRVARHFTKAKLT I PILYIK
Spermophilus	KDVVYLNLVLDYVPETVYRVARHYSRAKQTLPVIYVK
Monodelphis	KDEVYLNLVLDYVPETVYRVARHYSRAKQTLPLIYVK
Taeniopygia	KDEVYLNLVLDYVPETVYRVARHYSRAKQTLPLIYVK

(*Continued*)

Table S2. (*Continued*)

Xenopus	KDEVYLNLVLDYVPETVYRVARHYSRAKQALP IIYVK
Tetraodon	
and Danio	KDEVYLNLVLDYVPETVYRVARHYSRAKQTLPMVYVK
Exon 5	
Homo sapiens	LYMYQLFRSLAYIHSFGICHRDIKPQNLLLDPDTAVLKLCDFGS
Macaca fasc	VYMYQLFRSLAYIHSQGVCHRDIKPQNLLVDPDTAVLKLCDFGS
Xenopus	LYMYQLFRSLAYIHSFGICHRDIKPQNLLLDPETAVLKLCDFGS
Tetraodon	
and Danio	LYMYQLFRSLAYIHSFGICHRDIKPQNLLLDPETAVLKLCDFGS
Exon 6–8	
No dissimilarities	
Exon 9	
See text	
Exon 10	
Homo sapiens	VFRPRTPPEAIALCSRLLEYTPTARLTPLEACAHSFFDELRDPNVKLPNGRDTP
Gallus	VFRPRTPPEAIALCSRLLEYTPTARLTPLEACAHSFFDELRDPNVKLPNGREKP
Taeniopygia	VFRPRTPPEAIALCSRLLEYTPTARLTPLEACAHSFFDELRDPNVKLPNGRDKP
Xenopus	VFRARTPPEAIALCSRLLEYTPTSRLTPLEACAHSFFDELRDPNLKLPNGREFP
Danio rerio	VFRPRTPPEAIALCSRLLEYTPTARLTPLEACAHSFFDELREPNVKLPNGREKP
Homo sapiens	ALFNFTTK
Gallus	ALFNFTTQ
Taeniopygia	ALFNFTTQ
Xenopus	ALFNFTTQ
Danio rerio	SLFNFTTQ
Exon 11	
Homo sapiens	ELSSNPPLATILIPPHARIQAAASTPTNATAASD
Macaca	ELS IQPSLNAILIPPHLRSPAGTTTLTQSSQA
Bos taurus	ELSSNPPLATILIPPHARIQAAASTPSNTTAASD
Ovis aries	ELSSNPPLATILIPPHARIQAAASTPSNTTAASD
Sus scrofa	ELSSNPPLATILIPPHARIQAAASTPSNATAASD
Equus	ELSSNPPLATILIPPHARIQATASTPTNATAASD
Rattus norv	ELSSNPPLATILIPPHARIQAAASPPANATAASD
Mus musc	ELSSNPPLATILIPPHARIQAAASPPANATAASD
Gallus	ELSSNPSLASILIPAHARNQAAASTPTNATAASD
Taeniopygia	ELSSNPSLASILIPAHARNQAAASTPTNATAASD
Xenopus	ELSSNPSLSSILIPAHARNQAAASTPTNATAASD
Tetraodon	ELSIQPQLNSILIPPHARITYTTASS
Danio rerio	ELSSNPTLASILIPAHARNQAGASTPTNPSATSD
Exon 12	
Homo sapiens	ANTGDRGQTNNAASASASNST
Bos taurus	ANAGDRGQTNNAASASASDS

(*Continued*)

Table S2. (*Continued*)

Ovis aries	ANAGDRGQTNNAASASASDS
Sus scrofa	ANAGDRGQTNNTASASASNST
Canis	ANAGDRGQTNNAASASASNST
Equus	ANAGDRGQTNNAASASASNST
Rattus norv	ANAGDRGQTNNAASASASNST
Mus musc	ANAGDRGQTNNAASASASNST
Spermophilu	ANAGDRGQTNNAAFASASNST
Monodelphis	ANAGDRGQTNNAASASASNST
Gallus	ANAGERVQTNSVATASASNST
Taeniopygia	ANAGERVQTNSVATASASNST
Xenopus	ANTGDRGQTNNAASASASNS
Tetraodon	Degenerate, uncertain coalescent of fragments
Danio rerio	ANSGDRSQTTTAASASASNTST (*N* in italics is an insertion)

Empty spaces relative to the human sequences in the ends of the peptides signifies lack of sequence information.