BMC Evolutionary Biology

Research article

Evolutionary relationships of Aurora kinases: Implications for model organism studies and the development of anti-cancer drugs

James R Brown*[1], Kristin K Koretke[1], Marian L Birkeland[1], Philippe Sanseau[1] and Denis R Patrick[2]

Address: [1]Bioinformatics Division, Genetics Research, GlaxoSmithKline, 1250 South Collegeville Road, UP1345, P.O. Box 5089, Collegeville Pennsylvania 19426-0989, USA and [2]Dept. of Molecular Oncology, Microbial, Musculoskeletal and Proliferative Disease Center for Excellence in Drug Discovery, GlaxoSmithKline, 1250 South Collegeville Road, UP1345, P.O. Box 5089, Collegeville Pennsylvania 19426-0989, USA

Email: James R Brown* - James.R.Brown@gsk.com; Kristin K Koretke - Kristin.K.Koretke@gsk.com; Marian L Birkeland - Marian.L.Birkeland@gsk.com; Philippe Sanseau - Philippe.X.Sanseau@gsk.com; Denis R Patrick - Denis.R.Patrick@gsk.com

* Corresponding author

Published: 12 October 2004

BMC Evolutionary Biology 2004, **4**:39 doi:10.1186/1471-2148-4-39

Received: 12 May 2004
Accepted: 12 October 2004

This article is available from: http://www.biomedcentral.com/1471-2148/4/39

Abstract

Background: As key regulators of mitotic chromosome segregation, the Aurora family of serine/threonine kinases play an important role in cell division. Abnormalities in Aurora kinases have been strongly linked with cancer, which has lead to the recent development of new classes of anti-cancer drugs that specifically target the ATP-binding domain of these kinases. From an evolutionary perspective, the species distribution of the Aurora kinase family is complex. Mammals uniquely have three Aurora kinases, Aurora-A, Aurora-B, and Aurora-C, while for other metazoans, including the frog, fruitfly and nematode, only Aurora-A and Aurora-B kinases are known. The fungi have a single Aurora-like homolog. Based on the tacit assumption of orthology to human counterparts, model organism studies have been central to the functional characterization of Aurora kinases. However, the ortholog and paralog relationships of these kinases across various species have not been rigorously examined. Here, we present comprehensive evolutionary analyses of the Aurora kinase family.

Results: Phylogenetic trees suggest that all three vertebrate Auroras evolved from a single urochordate ancestor. Specifically, Aurora-A is an orthologous lineage in cold-blooded vertebrates and mammals, while structurally similar Aurora-B and Aurora-C evolved more recently in mammals from a duplication of an ancestral Aurora-B/C gene found in cold-blooded vertebrates. All so-called Aurora-A and Aurora-B kinases of non-chordates are ancestral to the clade of chordate Auroras and, therefore, are not strictly orthologous to vertebrate counterparts. Comparisons of human Aurora-B and Aurora-C sequences to the resolved 3D structure of human Aurora-A lends further support to the evolutionary scenario that vertebrate Aurora-B and Aurora-C are closely related paralogs. Of the 26 residues lining the ATP-binding active site, only three were variant and all were specific to Aurora-A.

Conclusions: In this study, we found that invertebrate Aurora-A and Aurora-B kinases are highly divergent protein families from their chordate counterparts. Furthermore, while the Aurora-A family is ubiquitous among all vertebrates, the Aurora-B and Aurora-C families in humans arose from a gene duplication event in mammals. These findings show the importance of understanding evolutionary relationships in the interpretation and transference of knowledge from studies of model organism systems to human cellular biology. In addition, given the important role of Aurora kinases in cancer, evolutionary analysis and comparisons of ATP-binding domains suggest a rationale for designing dual action anti-tumor drugs that inhibit both Aurora-B and Aurora-C kinases.

Background

The Auroras are a conserved family of serine/threonine kinases which have essential functions in cell division [1,2]. In mitosis, Aurora kinases are required for chromosome segregation, condensation and orientation in the metaphase plate, spindle assembly, and the completion of cytokinesis.

Model organism studies have played a pivotal role in functional characterization of Aurora kinases. Aurora kinases were first identified as mutant alleles in *Drosophila melanogaster* (fruitfly) that caused defective spindle-pole formation [3]. Subsequently, *Drosophila* was found to have a second Aurora homolog [4], and the nematode, *Caenorhabditis elegans*, similarly has two Aurora-like genes [5,6]. The fungi, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, have a single Aurora, known as increase-in-ploidy 1 (Ipl1) [7] and Aurora-related kinase 1 (Ark1) [8], respectively. Among cold-blooded vertebrates, Aurora kinases have been most widely studied in the frog, *Xenopus laevis*, which has two kinases; Aurora-A and Aurora-B [9,10]. More recently discovered is a third Aurora kinase called Aurora-C in rodents and humans [11].

The Aurora kinases are mitotic kinases that generally associate with chromosomes, often in complexes with other proteins, and interact with cytoskeletal components in cell division. The three mammalian Aurora kinases appear at specific locations during mitosis. Aurora-A, the "polar kinase", primarily associates with the separating centrosomes while Aurora-B, the "equatorial kinase", is a chromosomal passenger protein [1]. The least studied Aurora kinase, Aurora-C, appears to be localized to the centrosome from anaphase to telophase and is highly expressed in the testis [11,12].

Recent studies indicate that all three Aurora kinases have strong associations with cancer. Aurora-A has been mapped to a region in the human chromosome (20q13.2-13.3) that is amplified in cancer cell lines and primary tumors [13,14]. Transfected mouse cell lines with Aurora-A have been shown to cause tumors when injected into nude mice [14,15] and a polymorphic variant (amino acid substitution Phe31Ile) has been associated with human colon tumors [16]. Expression levels of Aurora-B [17] and Aurora-C [12] were elevated in several cancer cell lines relative to normal fibroblasts. Aurora-C is located on chromosome 19q13.2 to 13.4, a region associated with loss of heterozygosity in ovarian cancer [18] and pancreatic carcinomas [19]. Thus, the inhibition of one or more Aurora kinases might be a novel chemotherapeutic strategy against cancer [20]. Recently, several reports by research groups in pharmaceutical and biotechnology companies describe small molecules that target the ATP-binding domain of Aurora kinases, and have effects in human tumor cell lines [21-23].

Despite the importance of model organisms in understanding Aurora kinase function, the evolutionary relationships among these variants are unclear. Two previous phylogenetic analyses of Aurora kinases were incomplete because the contemporary complement of Aurora kinases was unavailable [4] or certain family members, namely the Aurora-C kinases, were excluded [1]. Here, we present an evolutionary analysis of all known Aurora kinases. We show that vertebrate Aurora kinases evolved through a series of gene duplication events from a chordate ancestor, and that they are highly distinct from invertebrate homologs. Moreover, the recent divergence, thus high level of sequence similarity, of human Aurora-B and Aurora-C suggests a novel anti-cancer strategy which might simultaneously target the ATP-binding domains of this kinase pair with dual action inhibitors.

Results and discussion

Aurora Evolution in Chordates

In order to construct a comprehensive phylogenetic tree, GenBank was searched for all possible Aurora kinases. In addition to previously published Aurora kinase sequences, further chordate and urochordate Aurora homologs were found by using mammalian Aurora-A, Aurora-B and Aurora-C protein sequences as queries in BLASTP or TBLASTN [24] searches of the genomes of the pufferfish, *Takifugu rubripes* [25], the zebrafish, *Danio rerio* [26], and the ascidian, *Ciona intestinalis* [27].

Multiple sequence alignments show that the Aurora kinase family is highly conserved among species (Fig. 1). Pairwise sequence comparisons estimate that the mean proportion of similar amino acids (based on the Blosum62 matrix) is much higher among all the different families of Aurora-A, Aurora-B and Aurora-C of vertebrates (0.84 ± 0.5) than within the same family (Aurora-A or Aurora-B) between vertebrates and invertebrates species (0.69 ± 0.3 for both families). This would suggest a recent evolutionary radiation of Aurora families within vertebrates.

Phylogenetic trees constructed using four methodologies, all rooted using polo-like kinases type 4 (PLK4), show that all vertebrate Auroras form a clade distinct from those of invertebrates (Fig. 2). The phylogenetic tree constructed by the neighbour-joining distance method shows moderate boot-strap support (67%) for the evolution of all vertebrate Auroras from a urochordate ancestor, represented by the ascidian, *C. intestinalis*. The use of alternative kinase families, other than PLK4, to root the tree did not alter the internal topology of the Aurora clade. Although its genome sequence is incomplete, *C. intestinalis* likely has

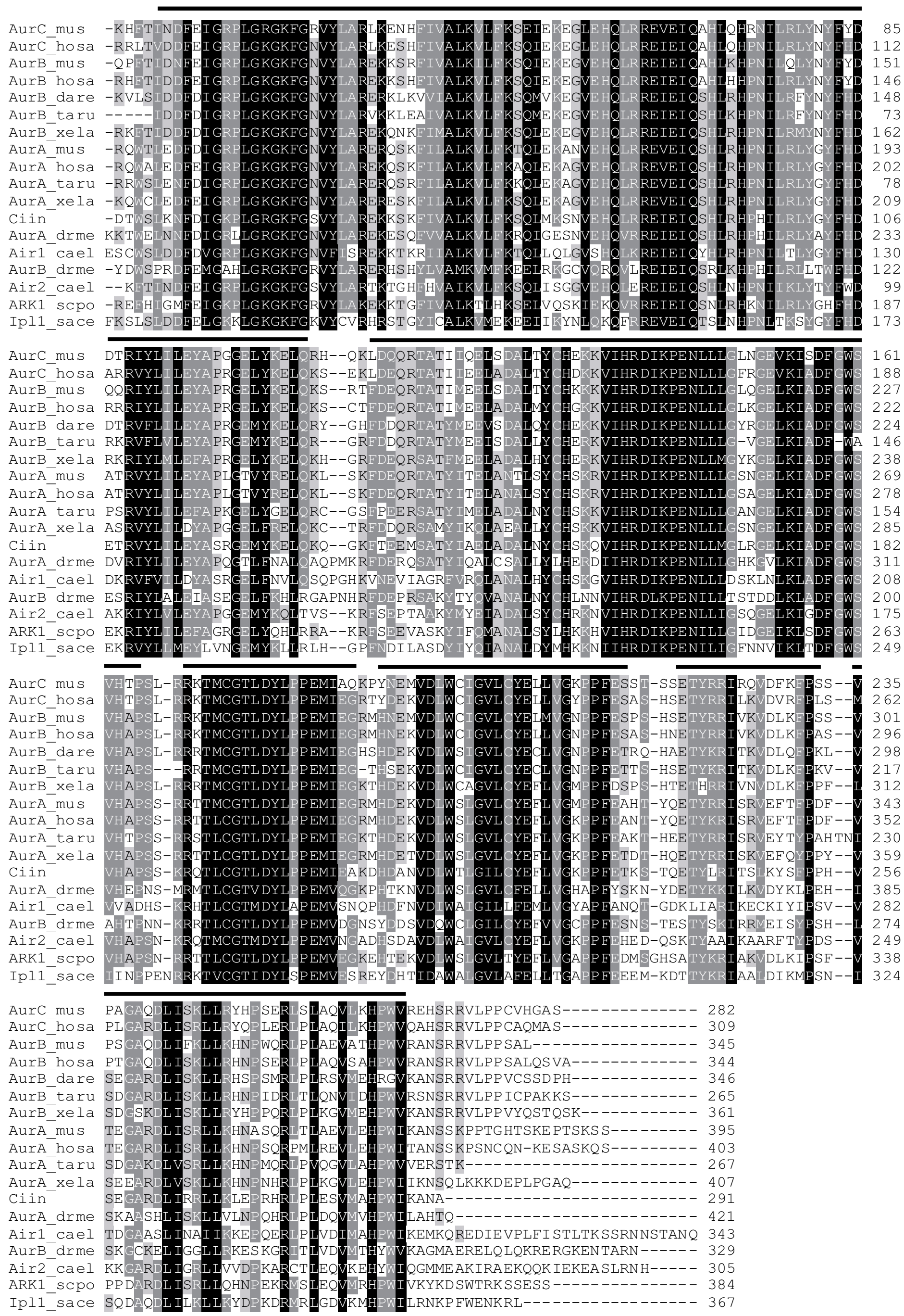


Figure 1
Multiple sequence alignment of representative Aurora-A (AurA), Aurora-B (AurB), and Aurora-C (AurC) kinases, and their homologs (Air1, Air2, ARK1 and Ipl1). N-terminal regions which are species-specific and could not be accurately aligned are excluded, although the numbering of residues begins at the starting amino acid for that particular peptide. Progressive darker shading indicates conservation of amino acid residues in 60%, 80% and 100% of the sequences, respectively. Dark line at the top of the sequence blocks indicates those regions used in the phylogenetic analyses (Also see additional file 1 and 2). Species include *Homo sapiens* (hosa), *Mus musculus* (mus), *Danio rerio* (dare), *Takifugu rubripes* (taru), *Xenopus laevis* (xela), *Ciona intestinalis* (ciin), *Drosophila melanogaster* (drme), *Caenorhabditis elegans* (cael), *Saccharomyces cerevisiae* (sace) and *Schizosaccharomyces pombe* (scpo). The program CLUSTALW [41] was used to constructed the initial alignment which was subsequently refined manually.

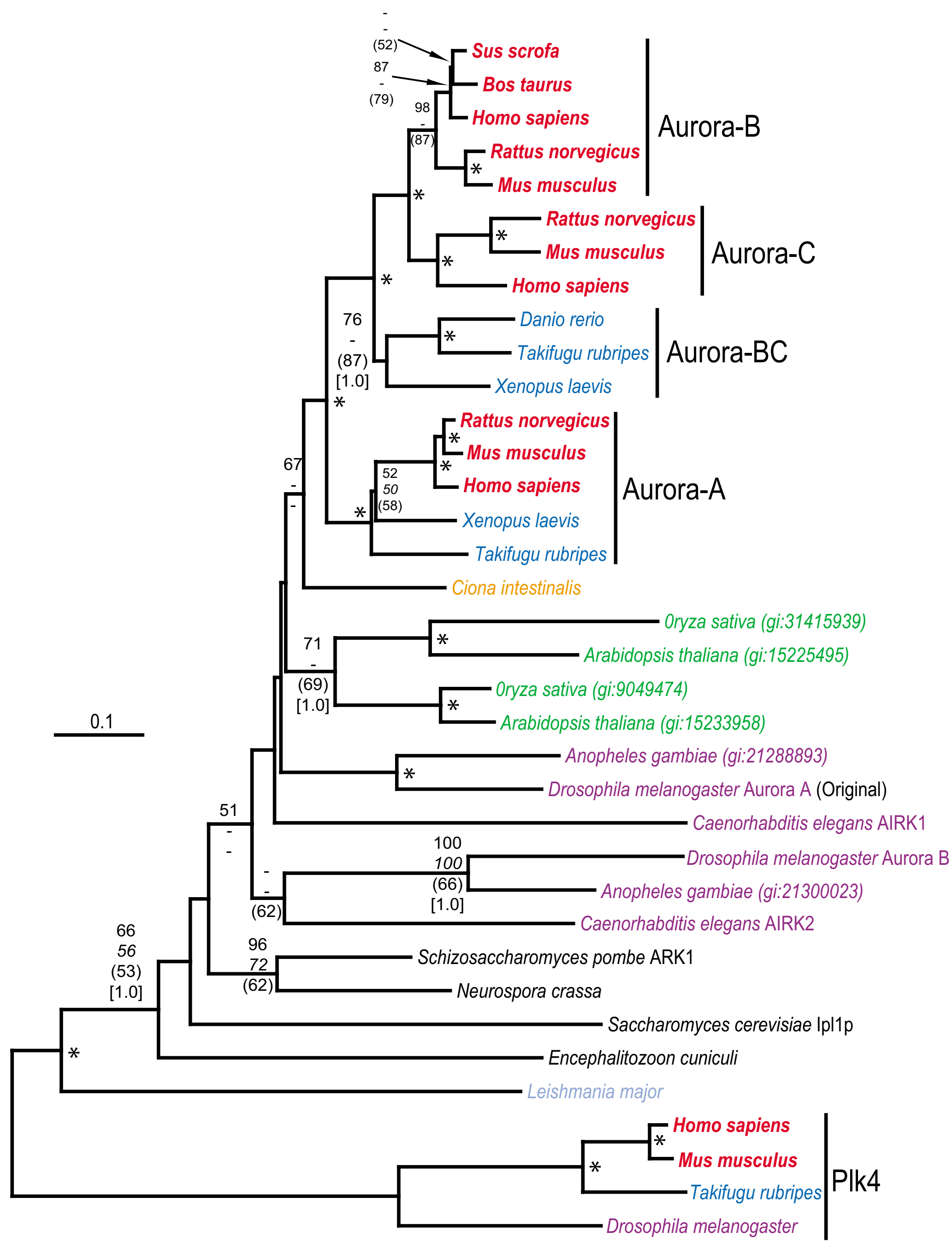


Figure 2

Phylogenetic tree of Aurora-A, Aurora-B, and Aurora-C kinases rooted by PLK4 kinases. Major organism groups (with colours, fonts) are mammals (red, bold italic), cold-blooded vertebrates (deep blue, italic), urochordates (orange, italic), invertebrates, (purple, italic), plants (green, italic), fungi (black, italic) and protists (light blue, italic). "Original" indicates the first Aurora identified from *Drosophila melanogaster* [3]. Plant sequences are identified by their Genbank accession number. Stacks of numbers show, in descending order, the percent occurrence of nodes in greater than 50% of 1000 bootstrap replicates of neighbor joining (plain text) and maximum parsimony (italicized text) analyses or greater than 50% of 10000 quartet puzzling steps of maximum likelihood analysis (in curved parentheses) or Bayesian posterior probability (only 0.90 or greater, in square parentheses). Asterisks ("*") indicate those nodes supported 70% or greater by the first three tree-building methods and 0.90 Bayesian posterior probability. Nodes with one or two values less than 50% have dashes ("-") while values less than 50% are unmarked. Scale bar represents 0.1 expected amino acid residue substitutions per site.

only a single Aurora homolog since other probable kinase open reading frames associated with the next top five BLASTP [24] hits did not cluster with Auroras from other species in phylogenetic trees.

Among true vertebrates, our phylogenetic tree shows that the Aurora kinases underwent two major gene duplication events. The first split in cold-blooded vertebrates lead to the formation of two Aurora subfamilies. One branch encompasses all known vertebrate Aurora-A sequences in a single orthologous lineage that includes fishes, amphibians and mammals. This family includes previously identified Aurora-A kinases in *Xenopus laevis*, rodents and humans as well as a new putative ortholog in *T. rubripes*.

The second family, previously known as Aurora-B [2,28] consists of cold-blooded vertebrate and mammalian Aurora-B as well as mammalian Aurora-C. Mammalian Aurora-B and Aurora-C are similarly related to the cold-blooded vertebrate Aurora presently known as "Aurora-B" in amphibians (*X. laevis*) and fish (*D. rerio* and *T. rubripes*). Searches of *T. rubripes* and *D. rerio* protein and DNA sequence databases detected several other putative serine/threonine kinase homologs but none were Auroras according to phylogenetic analyses. Thus, cold-blooded vertebrates appear to have only a single Aurora-A ortholog, a single Aurora-B-like homolog, and lack an Aurora-C ortholog. Conversely, Aurora-A, Aurora-B and Aurora-C appear to be ubiquitous to mammals (at least placentals) where they are encoded by separate chromosomal loci. It would appear that mammalian Aurora-B and Aurora-C evolved from a duplication event involving the ancestral Aurora-B found in cold-blooded vertebrates. This depiction of the evolutionary relationships of vertebrate Auroras was consistently determined by four different phylogenetic methods with high bootstrap or Bayesian posterior probability values (Fig. 2).

Comparisons of human Aurora-B and Aurora-C sequences to the resolved 3D structure of human Aurora-A [29] lends further support to the evolutionary scenario that vertebrate Aurora-B and Aurora-C are closely related paralogs (Fig. 3a). Of the 26 residues lining the ATP-binding active site, only three vary among the different human Aurora kinases; Leu215, Thr217 and R220 (numbering and residue identity based on Aurora-A), and all of these variants were specific to Aurora-A (Fig. 3b). Aurora-B and Aurora-C did not vary in their active site residues. Furthermore, all three Auroras have a carboxy-terminal destruction box (D-box) but only Aurora-A has the necessary amino-terminal A-box (also known as the D-box activating-domain) for its functional activation [30,31]. Collectively, these comparisons of structure and motifs support the phylogeny depicting an early divergence of Aurora-A from an Aurora-B / Aurora-C clade.

Non-chordate Evolution

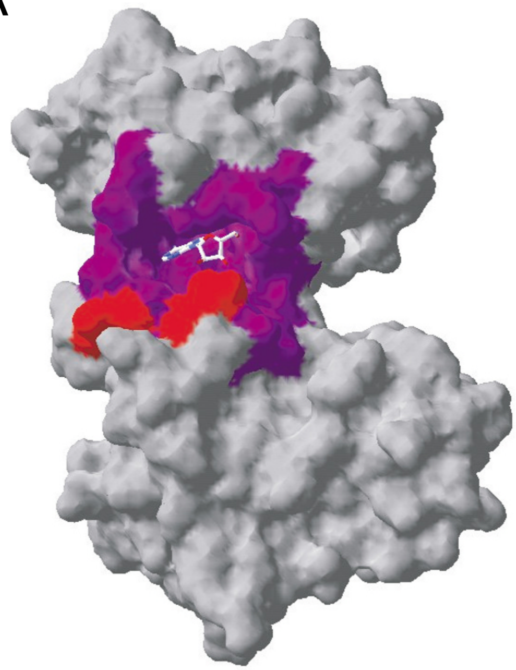
The Aurora kinases of plants and invertebrates are all outgroup lineages to chordates / urochordates (Fig. 2). Although all phylogenetic methods strongly support the monophyly of chordate Aurora kinases, the exact ordering among nodes leading to the various plant and invertebrate clusters were not resolved with similarly high bootstrap or probability values. Placement of plant Aurora kinases between chordates and invertebrates might be an artifact of tree construction methods. (Plant, protist, fungal and invertebrate lineages were all highly diverged from vertebrate Aurora kinases as witnessed by their longer branch lengths.) The earliest lineages of the Aurora tree are those fungal model organisms with a single Aurora-like homolog *S. cerevisiae* (Ipl1) and *S. pombe* (Ark1). Other basal branches are the amitochondrial fungi, *Encephalitozoon cuniculi*, and the kinetoplast protist, *Leishmania major* [32].

Invertebrate Aurora kinases, including those of the model organisms *C. elegans* and *D. melanogaster*, occupy separate early branches and are not, as their current names suggest, orthologs to either Aurora-A or Aurora-B of vertebrates. An unrooted phylogenetic tree with only model organism species shows the same topology of vertebrate Auroras as the more species-rich tree rooted by PLK4 kinases (Fig. 4). However, similar kinases from *C. elegans* and *D. melanogaster* now cluster together. The unrooted tree suggests that the invertebrate Aurora-B kinase family evolved prior to the invertebrate Aurora-A kinase family although further examples from other species are desirable to confirm this hypothesis. The consensus scenario in both rooted and unrooted trees is that vertebrate Aurora kinases are paralogous, rather than orthologous, to their invertebrate counterparts.

Model Organisms in Context

Aurora-B and Aurora-C, as specific innovations in mammals, might have distinct protein-binding partners and cellular functions from those of Aurora-B kinases in amphibians. The perturbation of Aurora-B function in different systems suggests variable kinetochore-microtubule interactions [33]. Transfection of normal rat kidney cells with a kinase-inactive, dominant negative form of Aurora-B caused multiple defects in mitosis [34] while an Aurora-B kinase inactivating antibody seemed to have milder effects in *Xenopus* tissue culture cells [35]. *Xenopus* Aurora-A functions in the extrusion of the first polar body [36] while in *C. elegans* Aurora B plays a similar role [5]. Also, *C. elegans* Aurora-B binds to a protein CSC-1 which has no homolog in other studied systems [37]. While these studies used different experimental methods, the lack of direct orthology among vertebrate and invertebrate Aurora-A and Aurora-B might also account for functional differences in these systems.

A



B

```
AuroraA   FEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGV 174
AuroraB   FEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGV 118
AuroraC   FEIGRPLGKGKFGNVYLARLKESHFIVALKVLFKSQIEKEGL  84
              * ****    *            * *

AuroraA   EHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLG 216
AuroraB   EHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRG 160
AuroraC   EHQLRREIEIQAHLQHPNILRLYNYFHDARRVYLILEYAPRG 126
                             *               ***** *

AuroraA   TVYRELQKLSKFDEQRTANLYNRIANALSYCHSKRVIHRDIK 258
AuroraB   ELYKELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIK 202
AuroraC   ELYKELQKSEKLDEQRTATIIEEVADALTYCHDKKVIHRDIK 168
            *

AuroraA   PENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
AuroraB   PENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEM 244
AuroraC   PENLLLGFRGEVKIADFGWSVHTPLPERKTMCGTLDYLPPEM 210
           ** **        * * *

AuroraA   IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRIS 342
AuroraB   IEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIV 286
AuroraC   IEGRTYDEKVDLWCIGVLCYELLVGYPPFESASHSETYRRIL 252

AuroraA   RVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPW   382
AuroraB   KVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPW   326
AuroraC   KVDVRFPLSMPLGARDLISRLLRYQPLERLPLAQILKHPW   292
```

Figure 3

Comparisons of the catalytic domains of human Aurora-A, Aurora-B and Aurora-C kinases. A. Crystal structure of the catalytic domain of Human Aurora kinase with an adenosine molecule shown in the binding pocket (PDB ID 1muoA) [29]. Residues lining the active site are colored purple when invariant and red when variant. **B**, Multiple sequence alignment of Auroras. Using the same color scheme as the structure in panel A, residues identified to be lining the active site are identified with invariant residues among all three Auroras marked with an asterisk. Of the 26 residues lining the active site, only three vary among the different human Aurora kinases; Leu215, Thr217 and R220 (numbering and residue identity based on Aurora-A), and all of this variation was found in Aurora-A.

The evolutionary analysis presented here also suggests revisiting the present Aurora nomenclature. Adams *et al.* [28] proposed a naming scheme where, irrespective of species, the original Aurora is known as Aurora-A (also called AIRK1, Aurora, Aurora-2, AIK, BTAK, human STK15, mouse STK6 and others), followed by Aurora-B (also known as AIRK-2, IAL, Aurora-1, AIK2, STK12 and others) and Aurora-C (or STK13). However, the proposed nomenclature fails to reflect evolutionary, and possibly functional, relationships among the Auroras. We suggest that Aurora-A be retained as the name for all orthologs in mammals and cold-blooded vertebrates. While Aurora-B and Aurora-C seem appropriate for mammalian versions, the ancestral cold-blooded vertebrate "Aurora-B" might be renamed "Aurora-BC". As for invertebrates, the so-called Aurora-A or Aurora-B genes are clearly not orthologs to their respective vertebrate counterparts. However, introducing a new nomenclature here might simply add further confusion to the field.

Evolution of an Anti-Cancer Target

There have been several recent reports of Aurora kinase inhibitors that are under development by pharmaceutical or biotechnology companies for cancer treatment. The compounds Hesperadin (Boehringer Ingelheim [21]) and ZM447439 (AstraZeneca [22]) are suggested to be targeted to Aurora-B. While both studies show lesser levels of compound inhibition of Aurora-A as well as several other kinases, neither report included Aurora-C in their kinase profile. Selective inactivation of multiple kinases is not an undesirable pharmaceutical profile for a small molecule inhibitor and, in fact, could be the best strategy to achieve maximal clinical efficacy of an anti-cancer agent [38]. Indeed, an intense area of anti-cancer research is the development of small molecular ATP analogues that generally target the kinase domain of protein kinases [39]. For example, Gleevec (also known as imatinib and made by Novartis) for chronic myelogenous leukemia, is a small-molecule inhibitor that targets BCR-ABL, c-Kit and

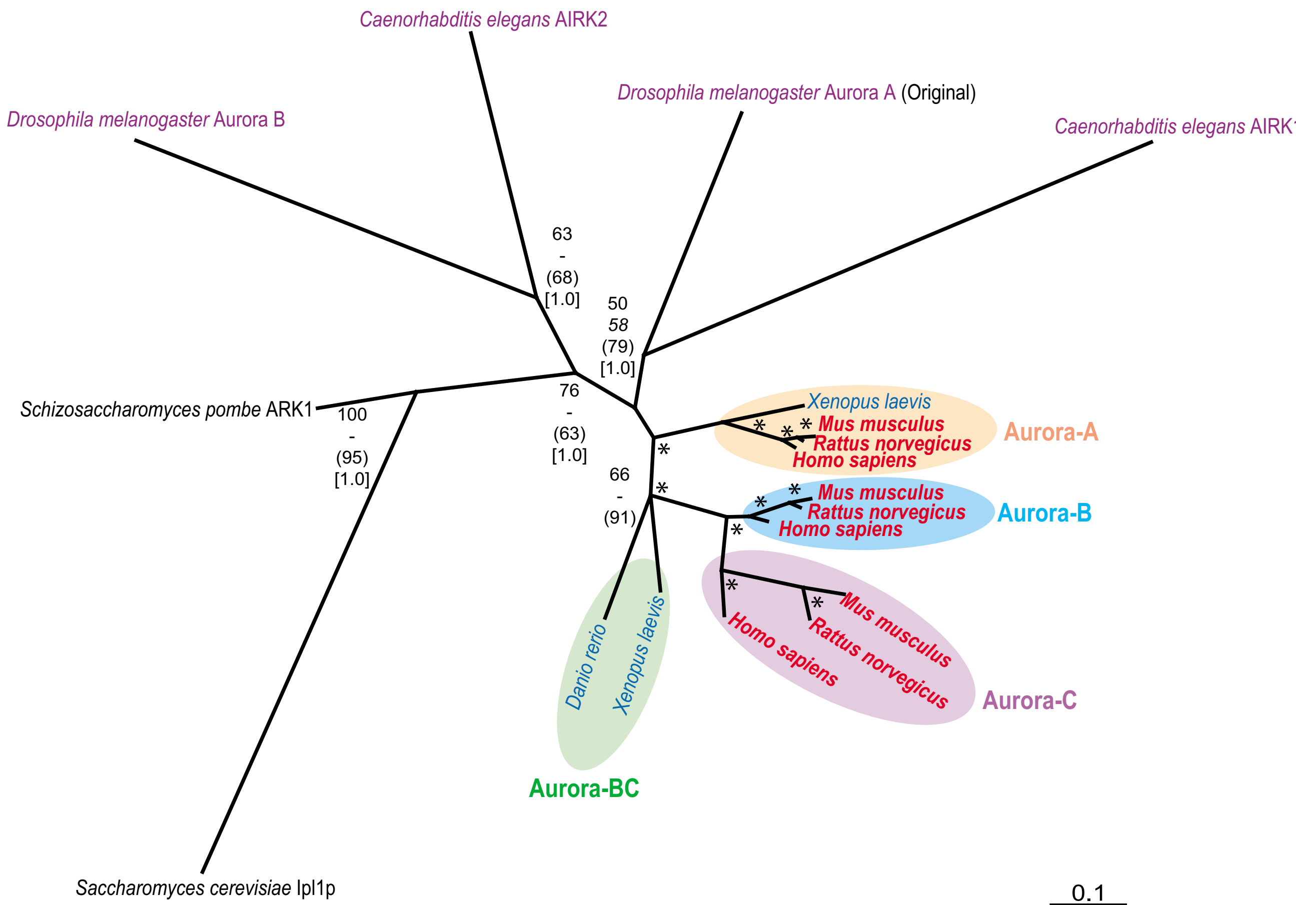


Figure 4
Unrooted phylogenetic tree of Aurora kinases from human and model organisms. Tree was constructed using the maximum likelihood quartet puzzling method [43]. Scale bar represents 0.1 expected amino acid residue substitutions per site. Confidence estimates of nodes, fonts, and colours of species names correspond to Fig. 2.

platelet-derived growth factor receptor kinases [40]. Recently, a selective inhibitor of all three Aurora kinases, VX-680 (made by Vertex Pharmaceuticals), was reported to inhibit cell-cycle progression and induce apoptosis in various human tumor cell types and in vivo xenograft models [23]. Interestingly, although VX-680 is a potent inhibitor of all three Aurora kinases, its apparent inhibition constant is much lower for Aurora-A (0.6 nM) than for either Aurora-B (18 nM) or Aurora-C (4.6 nM). Again, the compound's greater affinity for Aurora-A, relative to Aurora-B and Aurora-C is compatible with the proposed evolutionary scenario of mammalian Auroras.

Conclusions

Evolutionary analysis shows that cell division Aurora kinases, while consistent in theme throughout eukaryotes, have undergone lineage-specific expansions and specialization in metazoans. Aurora-C is the least known of the Aurora kinases. Yet as an evolutionary innovation in mammals, further studies are very much warranted from the perspectives of better understanding its potential roles in both cell replication and tumor progression. A better functional understanding of Aurora-C would help clarify the evolutionary relationships of Aurora-B and Aurora-C in mammals relative to the ancestral Aurora-BC in cold-blooded vertebrates. Additionally, the close evolutionary and structural relationships between mammalian Aurora-B and Aurora-C offers the tantalizing opportunity to design dual kinase inhibitors that might circumvent potential tumor cell resistance to mono-target chemotherapeutics.

Methods

Database searches

All Aurora kinase orthologs and paralogs were initially collected from GenBank nonredundant protein database by performing separate searches using BLASTP [24] with human Aurora-A, Aurora-B and Aurora-C proteins as query sequences and a cut-off E-value of 1.0e-10. Since this dataset included additional kinases to the Auroras, preliminary multiple sequence alignments and phylogenetic analysis using program CLUSTALW v1.7 [41] served to identify the clade of all known Aurora kinases. *Takifugu rubripes*, *Danio rerios* and *Ciona intestinalis* homologs were obtained by BLASTP and TBLASTN [24] of species-specific protein and DNA sequence databases, respectively. The top five homologs from each species retrieved from separate searches with human Aurora-A, Aurora-B and Aurora-C were entered into a preliminary phylogenetic analysis using all retrieved Aurora kinases from Genbank. These analyses revealed that *T. rubripes* and *D. rerios* had orthologs to *X. laevis* Aurora-A and Aurora-B but not mammalian Aurora-C. *C. intestinalis* had a single Aurora-like kinase.

Phylogenetic and structure analysis

PLK4 kinases were selected as the outgroup for phylogenetic analyses because they were the most similar non-Aurora kinases to either human Aurora-A, Aurora-B or Aurora-C in multiple BLASTP [24] searches of the non-redundant protein database of GenBank. Using alternative kinases as outgroups made no difference to the topology of the Aurora clade. Initial multiple sequence alignments were performed using the program CLUSTALW v1.7 [41] with default settings and subsequently, refined manually using the program SEQLAB of the GCG Wisconsin Package v11.0 software package (Accelrys, San Diego, CA, USA). We removed regions with residues that could not be unambiguously aligned or that contained insertions or deletions. The final multiple sequence alignment was 240 amino acids in length. Pairwise comparisons for the proportion of similar residues were estimated from the length of the shortest sequence without gaps and the Blosum62 weighting matrix as implemented in the program OLDDISTANCES in GCG.

We constructed phylogenetic trees using distance neighbor-joining (NJ), maximum parsimony (MP), maximum likelihood quartet puzzling (QP), and Bayesian posterior probabilities (BP). NJ trees were based on pair wise distances between amino acid sequences using the programs NEIGHBOR and PROTDIST (Dayhoff option) of the PHYLIP 3.6 package [42]. The programs SEQBOOT and CONSENSE were used to estimate the confidence limits of branching points from 1000 bootstrap replications. ML tree topologies were constructed using the software PUZZLE 4.0 [43], employing 1000 puzzling steps, the JTT substitution matrix, estimation of rate heterogeneity using the gamma distribution model with eight rate categories, and the gamma-parameter estimation from the dataset. MP analysis was performed using PAUP4.0b5 software [44] where the number and lengths of minimal trees were estimated from 100 random sequence additions, while confidence limits of branch points were estimated by 1000 bootstrap replications. BP trees were constructed using the software MrBayes v3.0B4 [45,46]. Bayesian analysis used the mixed model of sequence evolution with random starting trees. Markov chains were run for 10^6 generations, burn-in values were set for 10^4 generations, and trees sampled every 100 generations. All trees were visualized using the program TREEVIEW v1.6.6 [47].

For the Aurora kinase phylogeny rooted with PLK4 kinases shown in Fig. 2, the log likelihood of the final ML tree was -8059.78. Four minimal length MP trees were recovered, 1522 steps in length with a consistency index (CI) of 0.5802 and a retention index (RI) of 0.6016. The variable branch arrangements were terminal nodes (human, pig and cow Aurora-B) which did not affect the central findings.

For the unrooted phylogeny of Aurora kinases of model organisms shown in Fig. 3, the log likelihood of the final ML tree was -4469.20. A single minimal length MP trees were recovered, 779 steps in length with a consistency index (CI) of 0.7214 and a retention index (RI) of 0.2786.

The SwissPDBviewer program [48] was used to obtain the surface representation of human Aurora-A kinase (PDB ID 1muoA). The active site residues, defined as being within 5A of the ADP cofactor, were identified using the program CAST [49]. The multiple sequence alignment for the three human Aurora kinase proteins was obtained using CLUSTALW [41]. Multiple sequence alignment and sequence GenBank accession numbers are available as Supplementary Information [see additional file 1 and 2].

Authors' contributions

JRB conceived the study, performed the phylogenetic analysis and drafted the manuscript. KKK performed the 3D structural analyses and contributed to the draft manuscript. MLB participated in the preliminary phylogenetic analysis and contributed to the draft manuscript. PS and DRP both contributed to the draft of the manuscript and added key references.

Additional material

Additional File 1
Multiple sequence alignment of edited Aurora and Plk4 kinases used to produce the phylogeny shown in Fig. 2.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-4-39-S1.doc]

Additional File 2

Accession numbers (gi), Locus link Ids (LID) or genome predicted protein Ids for Aurora and PLK kinases as labelled in the Supplementary Data multiple sequence alignment file S1.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-4-39-S2.xls]

Acknowledgements

This work was supported by Bioinformatics, Genetics Research, and the Microbial, Musculoskeletal and Proliferative Disease Center for Excellence in Drug Discovery, GlaxoSmithKline.

References

1. Carmena M, Earnshaw WC: **The cellular geography of aurora kinases.** *Nat Rev Mol Cell Biol* 2003, **4**:842-854.
2. Nigg EA: **Mitotic kinases as regulators of cell division and its checkpoints.** *Nat Rev Mol Cell Biol* 2001, **2**:21-32.
3. Glover DM, Leibowitz MH, McLean DA, Parry H: **Mutations in aurora prevent centrosome separation leading to the formation of monopolar spindles.** *Cell* 1995, **81**:95-105.
4. Reich A, Yanai A, Mesilaty-Gross S, Chen-Moses A, Wides R, Motro B: **Cloning, mapping, and expression of ial, a novel Drosophila member of the Ipl1/aurora mitotic control kinase family.** *DNA Cell Biol* 1999, **18**:593-603.
5. Schumacher JM, Golden A, Donovan PJ: **AIR-2: An Aurora/Ipl1-related protein kinase associated with chromosomes and midbody microtubules is required for polar body extrusion and cytokinesis in Caenorhabditis elegans embryos.** *J Cell Biol* 1998, **143**:1635-1646.
6. Schumacher JM, Ashcroft N, Donovan PJ, Golden A: **A highly conserved centrosomal kinase, AIR-1, is required for accurate cell cycle progression and segregation of developmental factors in Caenorhabditis elegans embryos.** *Development* 1998, **125**:4391-4402.
7. Francisco L, Wang W, Chan CS: **Type 1 protein phosphatase acts in opposition to IpL1 protein kinase in regulating yeast chromosome segregation.** *Mol Cell Biol* 1994, **14**:4731-4740.
8. Petersen J, Paris J, Willer M, Philippe M, Hagan IM: **The S. pombe aurora-related kinase Ark1 associates with mitotic structures in a stage dependent manner and is required for chromosome segregation.** *J Cell Sci* 2001, **114**:4371-4384.
9. Roghi C, Giet R, Uzbekov R, Morin N, Chartrain I, Le Guellec R, Couturier A, Doree M, Philippe M, Prigent C: **The Xenopus protein kinase pEg2 associates with the centrosome in a cell cycle-dependent manner, binds to the spindle microtubules and is involved in bipolar mitotic spindle assembly.** *J Cell Sci* 1998, **111 (Pt 5)**:557-572.
10. Adams RR, Wheatley SP, Gouldsworthy AM, Kandels-Lewis SE, Carmena M, Smythe C, Gerloff DL, Earnshaw WC: **INCENP binds the Aurora-related kinase AIRK2 and is required to target it to chromosomes, the central spindle and cleavage furrow.** *Curr Biol* 2000, **10**:1075-1078.
11. Bernard M, Sanseau P, Henry C, Couturier A, Prigent C: **Cloning of STK13, a third human protein kinase related to Drosophila aurora and budding yeast Ipl1 that maps on chromosome 19q13.3-ter.** *Genomics* 1998, **53**:406-409.
12. Kimura M, Matsuda Y, Yoshioka T, Okano Y: **Cell cycle-dependent expression and centrosome localization of a third human aurora/Ipl1-related protein kinase, AIK3.** *J Biol Chem* 1999, **274**:7334-7340.
13. Bischoff JR, Anderson L, Zhu Y, Mossie K, Ng L, Souza B, Schryver B, Flanagan P, Clairvoyant F, Ginther C, Chan CS, Novotny M, Slamon DJ, Plowman GD: **A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers.** *EMBO J* 1998, **17**:3052-3065.
14. Zhou H, Kuang J, Zhong L, Kuo WL, Gray JW, Sahin A, Brinkley BR, Sen S: **Tumour amplified kinase STK15/BTAK induces centrosome amplification, aneuploidy and transformation.** *Nat Genet* 1998, **20**:189-193.
15. Littlepage LE, Wu H, Andresson T, Deanehan JK, Amundadottir LT, Ruderman JV: **Identification of phosphorylated residues that affect the activity of the mitotic kinase Aurora-A.** *Proc Natl Acad Sci U S A* 2002, **99**:15440-15445.
16. Ewart-Toland A, Briassouli P, de Koning JP, Mao JH, Yuan J, Chan F, MacCarthy-Morrogh L, Ponder BA, Nagase H, Burn J, Ball S, Almeida M, Linardopoulos S, Balmain A: **Identification of Stk6/STK15 as a candidate low-penetrance tumor-susceptibility gene in mouse and human.** *Nat Genet* 2003, **34**:403-412.
17. Ota T, Suto S, Katayama H, Han ZB, Suzuki F, Maeda M, Tanino M, Terada Y, Tatsuka M: **Increased mitotic phosphorylation of histone H3 attributable to AIM-1/Aurora-B overexpression contributes to chromosome number instability.** *Cancer Res* 2002, **62**:5168-5177.
18. Bicher A, Ault K, Kimmelman A, Gershenson D, Reed E, Liang B: **Loss of heterozygosity in human ovarian cancer on chromosome 19q.** *Gynecol Oncol* 1997, **66**:36-40.
19. Hoglund M, Gorunova L, Andren-Sandberg A, Dawiskiba S, Mitelman F, Johansson B: **Cytogenetic and fluorescence in situ hybridization analyses of chromosome 19 aberrations in pancreatic carcinomas: frequent loss of 19p13.3 and gain of 19q13.1-13.2.** *Genes Chromosomes Cancer* 1998, **21**:8-16.
20. Mahadevan D, Bearss DJ, Vankayalapati H: **Structure-based design of novel anti-cancer agents targeting aurora kinases.** *Curr Med Chem Anti -Canc Agents* 2003, **3**:25-34.
21. Hauf S, Cole RW, LaTerra S, Zimmer C, Schnapp G, Walter R, Heckel A, van Meel J, Rieder CL, Peters JM: **The small molecule Hesperadin reveals a role for Aurora B in correcting kinetochore-microtubule attachment and in maintaining the spindle assembly checkpoint.** *J Cell Biol* 2003, **161**:281-294.
22. Ditchfield C, Johnson VL, Tighe A, Ellston R, Haworth C, Johnson T, Mortlock A, Keen N, Taylor SS: **Aurora B couples chromosome alignment with anaphase by targeting BubR1, Mad2, and Cenp-E to kinetochores.** *J Cell Biol* 2003, **161**:267-280.
23. Harrington EA, Bebbington D, Moore J, Rasmussen RK, Ajose-Adeogun AO, Nakayama T, Graham JA, Demur C, Hercend T, Diu-Hercend A, Su M, Golec JM, Miller KM: **VX-680, a potent and selective small-molecule inhibitor of the Aurora kinases, suppresses tumor growth in vivo.** *Nat Med* 2004, **10**:262-267.
24. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**:3389-3402.
25. Aparicio S, Chapman J, Stupka E, Putnam N, Chia JM, Dehal P, Christoffels A, Rash S, Hoon S, Smit A, Gelpke MD, Roach J, Oh T, Ho IY, Wong M, Detter C, Verhoef F, Predki P, Tay A, Lucas S, Richardson P, Smith SF, Clark MS, Edwards YJ, Doggett N, Zharkikh A, Tavtigian SV, Pruss D, Barnstead M, Evans C, Baden H, Powell J, Glusman G, Rowen L, Hood L, Tan YH, Elgar G, Hawkins T, Venkatesh B, Rokhsar D, Brenner S: **Whole-genome shotgun assembly and analysis of the genome of Fugu rubripes.** *Science* 2002, **297**:1301-1310.
26. Centre The Sanger Sequencing: **The Danio rerio Sequencing Project.** 2004.
27. Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, De Tomaso A, Davidson B, Di Gregorio A, Gelpke M, Goodstein DM, Harafuji N, Hastings KE, Ho I, Hotta K, Huang W, Kawashima T, Lemaire P, Martinez D, Meinertzhagen IA, Necula S, Nonaka M, Putnam N, Rash S, Saiga H, Satake M, Terry A, Yamada L, Wang HG, Awazu S, Azumi K, Boore J, Branno M, Chin-Bow S, DeSantis R, Doyle S, Francino P, Keys DN, Haga S, Hayashi H, Hino K, Imai KS, Inaba K, Kano S, Kobayashi K, Kobayashi M, Lee BI, Makabe KW, Manohar C, Matassi G, Medina M, Mochizuki Y, Mount S, Morishita T, Miura S, Nakayama A, Nishizaka S, Nomoto H, Ohta F, Oishi K, Rigoutsos I, Sano M, Sasaki A, Sasakura Y, Shoguchi E, T Shin-i, Spagnuolo A, Stainier D, Suzuki MM, Tassy O, Takatori N, Tokuoka M, Yagi K, Yoshizaki F, Wada S, Zhang C, Hyatt PD, Larimer F, Detter C, Doggett N, Glavina T, Hawkins T, Richardson P, Lucas S, Kohara Y, Levine M, Satoh N, Rokhsar DS: **The draft genome of Ciona intestinalis: insights into chordate and vertebrate origins.** *Science* 2002, **298**:2157-2167.
28. Adams Richard R., Carmena Mar, Earnshaw William C.: **Chromosomal passengers and the (aurora) ABCs of mitosis.** *Trends in Cell Biology* 2001, **11**:49-54.

29. Cheetham GM, Knegtel RM, Coll JT, Renwick SB, Swenson L, Weber P, Lippke JA, Austen DA: **Crystal structure of aurora-2, an oncogenic serine/threonine kinase.** *J Biol Chem* 2002, **277:**42419-42422.
30. Castro A, Vigneron S, Bernis C, Labbe JC, Prigent C, Lorca T: **The D-Box-activating domain (DAD) is a new proteolysis signal that stimulates the silent D-Box sequence of Aurora-A.** *EMBO Rep* 2002, **3:**1209-1214.
31. Arlot-Bonnemains Y, Klotzbucher A, Giet R, Uzbekov R, Bihan R, Prigent C: **Identification of a functional destruction box in the Xenopus laevis aurora-A kinase pEg2.** *FEBS Lett* 2001, **508:**149-152.
32. Siman-Tov MM, Ivens AC, Jaffe CL: **Identification and cloning of Lmairk, a member of the Aurora/Ipl1p protein kinase family, from the human protozoan parasite Leishmania.** *Biochim Biophys Acta* 2001, **1519:**241-245.
33. Shannon KB, Salmon ED: **Chromosome dynamics: new light on Aurora B kinase function.** *Curr Biol* 2002, **12:**R458-R460.
34. Murata-Hori M, Wang YL: **The kinase activity of aurora B is required for kinetochore-microtubule interactions during mitosis.** *Curr Biol* 2002, **12:**894-899.
35. Kallio MJ, McCleland ML, Stukenberg PT, Gorbsky GJ: **Inhibition of aurora B kinase blocks chromosome segregation, overrides the spindle checkpoint, and perturbs microtubule dynamics in mitosis.** *Curr Biol* 2002, **12:**900-905.
36. Frank-Vaillant M, Haccard O, Thibier C, Ozon R, Arlot-Bonnemains Y, Prigent C, Jessus C: **Progesterone regulates the accumulation and the activation of Eg2 kinase in Xenopus oocytes.** *J Cell Sci* 2000, **113 (Pt 7):**1127-1138.
37. Romano A, Guse A, Krascenicova I, Schnabel H, Schnabel R, Glotzer M: **CSC-1: a subunit of the Aurora B kinase complex that binds to the survivin-like protein BIR-1 and the incenp-like protein ICP-1.** *J Cell Biol* 2003, **161:**229-236.
38. Adams Jerry, Huang Pearl, Patrick Denis: **A strategy for the design of multiplex inhibitors for kinase-mediated signalling in angiogenesis.** *Current Opinion in Chemical Biology* 2002, **6:**486-492.
39. Fischer OM, Streit S, Hart S, Ullrich A: **Beyond Herceptin and Gleevec.** *Curr Opin Chem Biol* 2003, **7:**490-495.
40. Capdeville R, Buchdunger E, Zimmermann J, Matter A: **Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug.** *Nat Rev Drug Discov* 2002, **1:**493-502.
41. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22:**4673-4680.
42. Felsentein J: **PHYLIP (Phylogenetic Inference Package).** Seattle, Department of Genetics, University of Washington; 2000.
43. Strimmer K, von Haesler A: **Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies.** *Mol Biol Evol* 1996, **13:**964-969.
44. Swofford DL: **PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods).** Sunderland, Massachusetts, Sinauer Associates; 1999.
45. Huelsenbeck JP, Ronquist F: **MRBAYES: Bayesian inference of phylogenetic trees.** *Bioinformatics* 2001, **17:**754-755.
46. Ronquist F, Huelsenbeck JP: **MrBayes 3: Bayesian phylogenetic inference under mixed models.** *Bioinformatics* 2003, **19:**1572-1574.
47. Page RD: **TreeView: an application to display phylogenetic trees on personal computers.** *Comput Appl Biosci* 1996, **12:**357-358.
48. Guex N, Peitsch MC: **SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling.** *Electrophoresis* 1997, **18:**2714-2723.
49. Liang J, Edelsbrunner H, Woodward C: **Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design.** *Protein Sci* 1998, **7:**1884-1897.