BMC Genetics

BioMed Central

Research article

Open Access

Molecular characterization of the *singed wings* locus of *Drosophila melanogaster*

Yuri B Schwartz[1,2,3], Tatiana Boykova[1,2], Elena S Belyaeva[1], Michael Ashburner[2] and Igor F Zhimulev*[1]

Address: [1]Institute of Cytology and Genetics, Russian Academy of Sciences, Novosibirsk, 630090, Russia, [2]Department of Genetics, University of Cambridge, Cambridge, CB2 3EH, UK and [3]Department of Zoology, University of Geneva, Geneva, 1205, Switzerland

Email: Yuri B Schwartz - yuri.schwartz@zoo.unige.ch; Tatiana Boykova - boykova@bionet.nsc.ru; Elena S Belyaeva - belyaeva@bionet.nsc.ru; Michael Ashburner - ma11@gen.cam.ac.uk; Igor F Zhimulev* - zhimulev@bionet.nsc.ru

* Corresponding author

Published: 09 June 2004

BMC Genetics 2004, **5**:15 doi:10.1186/1471-2156-5-15

Received: 03 March 2004
Accepted: 09 June 2004

This article is available from: http://www.biomedcentral.com/1471-2156/5/15

© 2004 Schwartz et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Hormones frequently guide animal development *via* the induction of cascades of gene activities, whose products further amplify an initial hormonal stimulus. In *Drosophila* the transformation of the larva into the pupa and the subsequent metamorphosis to the adult stage is triggered by changes in the titer of the steroid hormone 20-hydroxyecdysone. *singed wings (swi)* is the only gene known in *Drosophila melanogaster* for which mutations specifically interrupt the transmission of the regulatory signal from early to late ecdysone inducible genes.

Results: We have characterized *singed wings* locus, showing it to correspond to *EG:171E4.2* (*CG3095*). *swi* encodes a predicted 68.5-kDa protein that contains N-terminal histidine-rich and threonine-rich domains, a cysteine-rich C-terminal region and two leucine-rich repeats. The SWI protein has a close homolog in *D. melanogaster*, defining a new family of *SWI*-like proteins, and is conserved in *D. pseudoobscura*. A lethal mutation, swi^{476}, shows a severe disruption of the ecdysone pathway and is a C>Y substitution in one of the two conserved CysXCys motifs that are common to SWI and the *Drosophila Toll-4* protein.

Conclusions: It is not entirely clear from the present molecular analysis how the SWI protein may function in the ecdysone induced cascade. Currently all predictions agree in that SWI is very unlikely to be a nuclear protein. Thus it probably exercises its control of "late" ecdysone genes indirectly. Apparently the genetic regulation of ecdysone signaling is much more complex then was previously anticipated.

Background

Hormones often exert their effects on animal development *via* the induction of cascades of gene activities, whose products further amplify an initial hormonal stimulus. In *Drosophila* the transformation of the larva into the pupa and the subsequent metamorphosis to the adult stage is triggered by changes in the titer of the steroid hormone 20-hydroxyecdysone (referred to below as ecdysone). Ecdysteroids act through a regulatory cascade, first discovered by observations of changes in the morphology of the giant polytene chromosomes of larval salivary glands [1]. Transcribed genes in polytene chromosomes are often represented by puffs, providing an opportunity

to visualize the process of hormonally triggered sequential gene activation.

The sharp increase in ecdysone titer in the late third instar larva leads to the rapid induction of a few "early" puffs. These regress after reaching their maximum size, and after a lag period of several hours, many "late" puffs become active. The drop in ecdysone concentration following puparium formation leads to the regression of these "late" puffs and the induction of novel genes characteristic of the "mid-prepupal" period; subsequently the prepupal ecdysone pulse re-induces many of the early and late puffs seen active earlier in late larval development [1,2].

Ashburner and colleagues [1] proposed that ecdysone bound by a receptor protein directly induces the "early" puffs. The protein products of early puffs then both activate "late" puffs and repress the transcription of those genes that form the "early" puffs. The model also suggested that ecdysone directly represses the activation of the "late" puffs, preventing their premature induction by early puff products.

This model predicts that it should be possible to specifically disturb the sequence of puffing events by mutating the genes that are active as early puffs. Indeed, several mutations blocking the initiation of ecdysone-dependent puffing have been isolated. The affected genes include components of the functional ecdysone receptor, a heterodimer of the *EcR* and *usp* proteins [3,4], as well as transcription factors encoded by the *Broad-Complex* (*BR-C*), which forms an early puff [5]. In addition, mutations in the β*FTZ-F1* orphan receptor (which forms the 75D puff in mid-prepupae) impair the specific induction of early genes after the prepupal ecdysone pulse [6].

Surprisingly, only one mutation has been isolated that specifically disturbs the puffing sequence downstream of "early" puff induction in third instar larvae. This is the *t467* allele of the *swi* (*singed wings*) complementation group. The *swi*-complementation group is localized in region 2B6 – 2B8 and includes two hypomorphic mutations and four recessive semi-lethal or lethal mutations. Survivors have faded (singed) wings and swollen abdomens. swi^{t467} is fully lethal, most swi^{t467} homozygotes die as pharate adults [7]. Third instar larvae homozygous for swi^{t467} have normal sized salivary glands and polytene chromosomes. The induction and repression of "early" as well as "mid-prepupal" puffs in the polytene chromosomes of mutant flies remains unaffected by the mutation, but most of "late" puffs are either greatly reduced in size or completely absent [8].

Experiments using *in vitro* cultured larval salivary glands of swi^{t467}/swi^{t467} larvae showed that the premature removal of ecdysone from cultured glands partially rescues late puffing activity. These findings suggest that the product of the *swi* locus may control "late" puff induction and, in addition, that there are distinct pathways leading to early puff repression and late puff activation.

Interestingly, about 2.5% of swi^{t467} homozygous larvae pupariate only in the anterior part of the body. This phenotype is reminiscent of one described by Rayle [9] for the flies bearing the *halfway* (*hfw*) mutation. Since this mutation was mapped to the same cytological interval as *swi*, it is very likely to have been allelic to *swi* (see Methods).

The peculiarity of the *swi* mutant phenotype made it of interest to characterize this gene at the molecular level. This we have done, showing it to correspond *EG:171E4.2* (*CG3095*), encoding an uncharacterized protein that is similar to mammalian decorins and the *Drosophila Toll-4* protein; this gene also has at least one close homolog in *D. melanogaster* and a homolog in *D. pseudoobscura*.

Results and discussion

Genetic complexity of the 2B6/B7-8 region

Our earlier data placed the *swi* locus in region 2B6/B7-8 of the *X* chromosome, between the proximal break points of *Df(1)S39* and $Dp(1;Y)y^2 67g24.2Y$ [7]. This region corresponds to 1421800–1466900 bp of the Release 3.1 *X* chromosome sequence [10-12]. This interval is overlapped by the EDGP cosmids 73D1 and 9D2 (Fig. 1). Saturation mutagenesis of this region revealed only two loci [7]*deep orange (dor)* (see [13]) and *singed wings (swi)*. Subsequently, Makunin and co-authors [14] identified two further genes (*b6* and *a6*) (Fig. 1) and the computational analysis of the genomic sequence identified a further six [12] (Fig. 1). cDNA sequences are available for all of the genes predicted by sequence analysis, with the exception of *CG14797* [15].

swi *corresponds to EG:171E4.2*

In order to narrow the list of possible candidate loci for *swi* a genomic library made from the DNA of swi^+ flies in the CosPeR cosmid vector (see Methods) was screened using the DNA of the clone λ1487 as a probe. A 30 Kb clone (5581) was recovered and mapped to the genome (Fig. 1). Two lines independently transformed by this cosmid (5581.118 and 5581.4), each with a single independent autosomal insertion, were obtained and tested for complementation of swi^{t476}. Neither of them was able to complement swi^{t476}, implying that *a6*, *b6*, *CG14798* and *EG:9D2.4* are not *swi*.

To choose between remaining four loci we made a further set of transposons (Fig 1). *SWI1* contained the very 3'-end of *dor*, a complete *EG:171E4.4* gene and the 5'-half of *EG:171E4.2*. *SWI3* overlapped the 3'-end of *dor*, and the

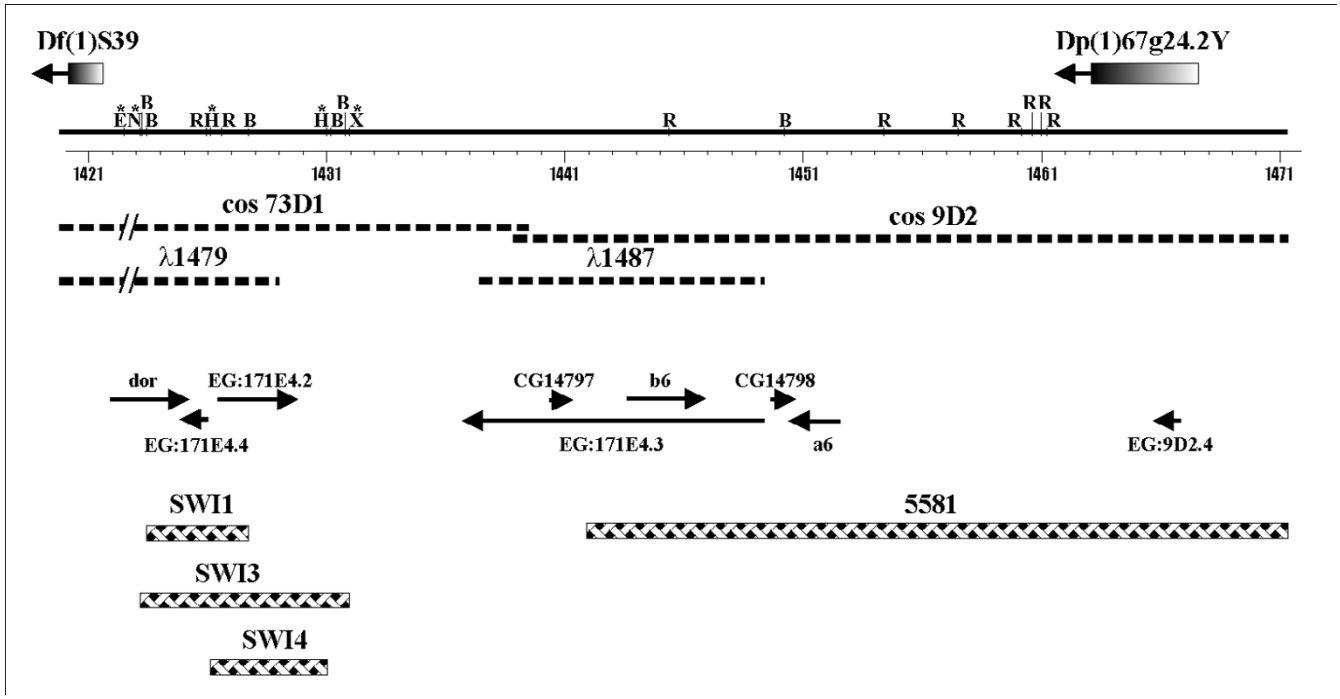

Figure 1
Organization of 2B6/B7-8 region The region is situated between the break points of *Df(1)S39* and the duplication *Dp(1;Y)y*^{2}*67g24.2Y* (shown by arrows). The shaded rectangles indicate the precision of rearrangement mapping. The restriction map of the region linked to the genomic sequence coordinates of FlyBase (in kb) is shown below. E-*Eag*I, N-*Nhe*I, B-*Bam*HI, X-*Xba*I, R-*Eco*R, H-*Hpa*I. In the case of those sites marked with an asterisk only ones relevant to this study are shown. The positions of genomic clones (indicated by dashed lines), genetic loci (shown with arrows) and pieces of genomic DNA used for transformation (cross-hatched lines) are presented under the restriction map. The genes *a6* and *b6* correspond to *CG3771* and *CG3100* respectively.

Table 1: Complementation test of SWI3 lines.

Transformed line		The number of flies in phenotypic classes of progeny from: $y^{1}swi^{t467}w^{pn}$/*FM6B*, $y^{+}w^{+}B$ **x** y^{1}, $Df(1)w^{67c23}$/Y; *SWI3*/+				
Name	Insertion chromosome	females		males		Total
		FM6, $y^{+}w^{+}B$	$y^{1}swi^{t467}w^{pn}$	*FM6*, $y^{+}w^{+}B$	$y^{1}swi^{t467}w^{pn}$	
swi 3.3.3	A[2]	234(46)[3]	209(41)	0	69(13)	512
swi 3.4.5	A	56(31)	102(56)	0	23(13)	181
swi 3.5.5	A	94(45)	83(40)	0	32(15)	209
swi 3.6.1	X	112(56)	87(44)	0	0	199
swi 3.7.2	A	83(45)	73(39)	0	30(16)	186
swi 3.8.16	A	95(41)	89(39)	0	46(20)	230
swi 3.9.5	A	195(43)	177(40)	0	75(17)	447
y^{1}, $Df(1)w^{67c23}$	(-)	44(52)	40(48)	0	0	84

[1] In the cross with autosomal insertions the eyes are darker then in w^{pn} flies because of the presence of an extra copy of w^{mW+} from a transposon. [2] A-autosomal insertion; X-insertion into the *X* chromosome; (-) – no insertion. [3] The numbers in parentheses indicates the percentage from the total of flies scored.

Table 2: Complementation test of SWI4 lines.

Transformed line		The number of flies in phenotypic classes of progeny from: $y^1 swi^{t467} w^{bn}$/*FM6B*, $y^+ w^+ B$ **x** y^1, $Df(1)w^{67c23}$/*Y*; *SWI4*/+				
Name	Insertion chromosome	females		males		Total
		FM6B, $y^+ w^+ B$	$y^1 swi^{t467} w^{bn}$	*FM6B*, $y^+ w^+ B$	$y^1 swi^{t467} w^{bn}$	
swi4-6.2	A	48(44)	39(35)	0	23(21)	110
swi4-11.1	A	54(32)	58(35)	0	56(33)	168
swi4-14.3	A	80(46)	68(39)	0	25(15)	173
swi4-15.2	A	59(42)	62(44)	0	19(14)	140
swi4-16.1	A	40(32)	60(48)	0	24(19)	124
swi4-16.3	A	70(40)	75(43)	0	30(17)	175
swi4-20.2	A	50(32)	72(46)	0	33(21)	155
swi4-21.1	A	59(46)	52(41)	0	16(13)	127
swi4-23.4	A	47(32)	50(34)	0	50(34)	147
swi4-24.3	A	49(35)	64(45)	0	28(20)	141
swi4-26.2	A	56(37)	68(45)	0	27(18)	151
swi4-27.3	A	32(34)	48(51)	0	15(16)	95
swi4-32.2	A	45(33)	68(50)	0	24(18)	137
swi4-39.2	A	41(41)	52(53)	0	6 (6)	99
swi4-39.3	A	58(42)	64(46)	0	17(12)	139
swi4-40.3	A	31(42)	38(51)	0	5 (7)	74
y^1, $Df(1)w^{67c23}$	(-)	68(48)	74(52)	0	0	142

full sequences of *EG:171E4.4* and *EG:171E4.2* while *SWI4* included only *EG:171E4.2*. Five independent transformed lines with insertions of *SWI1*, seven lines containing insertions of *SWI3* and sixteen lines with insertions of *SWI4* were established. All insertions of *SWI1* were autosomal; none complemented the swi^{t476} mutation. On the contrary, all six autosomal insertions of *SWI3* (Table 1) and all 16 insertions of *SWI4* completely rescued this mutation (Table 2). These results unambiguously indicate that the *swi* gene is identical to *EG:171E4.2*.

To confirm this conclusion we had isolated the DNA from homozygous $y^1 swi^{t476} w^{bn}/y^1 swi^{t47} w^{bn}$ third instar larvae and sequenced the region spanning from bases 1425864 to 1431743 (Release 3 coordinates). There are nineteen differences between this sequence (deposited in EMBL-Bank as AJ626646) and the reference sequence (AE003421). Only one of these differences affects the predicted protein product of *EG:171E4.2*, a G to A transition at base 1428762, that would change Cys379 to Tyr379.

The transformation rescue of the mutant *swi* allele by a construct that includes only *EG:171E4.2*, and the fact that the mutant allele carries a very non-conservative amino acid substitution prove that *swi* corresponds to this predicted gene.

Analysis of swi

The 4.9 kb *Hpa*I-*Hpa*I fragment from *SWI4* (Fig. 1) is capable of fully rescuing the swi^{t476} phenotype; this fragment, therefore, must include not only the coding sequence of *swi*, but also all of its necessary regulatory sequences. The *swi* (*EG:171E4.2*) mRNA as judged from the longest found cDNA sequence (RE03173) spans the region from position 1426365 to 1429883. We suspect that the true length of the transcript is slightly greater, since neither a consensus *Drosophila* Initiator (INR) element nor any known *Drosophila* promoter element could be found in the immediate vicinity of RE03173 5'-end. At the same time there is a perfect TATA box consensus (TATAAA), at position 1426228 (seen with PromH, [16]) and CA and T nucleotides at positions 30 and 36 downstream of this, characteristic of INR elements [17] (Additional file 1). The latter suggests that *swi* may have a strong TATA promoter, initiating 107 bp upstream of RE03173 and that is consistent with the fact that only a 150 bp upstream region is required for *swi* transgene function.

We have performed BLAST search of the *Drosophila pseudoobscura* genomic sequence database with the sequence of the 4.9 kb *Hpa*I-*Hpa*I fragment. This identified a region (coordinates 5225–8898) of contig 2140 that aligns with the *D. melanogaster* sequence from coordinate 1426109 to

1429698 and a short region of contig 2139 complementary sequence that aligns with the very 3' end of *D. melanogaster Hpa*I-*Hpa*I fragment (Additional file 1). The alignment of *D. melanogaster* sequence with that of contig 2140 corresponds to almost the entire *swi* locus with only last 185-bp of the 3'-UTR missing. The degree of conservation between these sequences is not uniform: there is high nucleotide conservation (78%) in the central coding region of *swi* (coordinates 1428116 – 1429529) (Additional file 1A), but less (44%) elsewhere (Additional file 1). In addition, there are short highly conserved regions within the first intron (coordinates 1427144 – 1427515) (Additional file 1). This region includes a number of short, but highly conserved, motifs, often AT rich and resembling the consensus homeobox binding consensus sequence (NNATTA) [18]. These motifs may well have a regulatory function.

Properties of the SWI protein

Translation of the *swi* open reading frame (ORF) predicts a 611 amino acid protein with threonine and histidine rich N-terminal domains, two leucine rich repeats and a cysteine rich C-terminus (Fig. 2A). The predicted amino acid sequence of the *Dpse*\SWI protein has 71% identity and 73% similarity to that of its *D. melanogaster* homolog.

The predicted SWI protein was used to search public protein databases for known or predicted proteins of similar sequence. The sequences of the protein of the predicted *CG14485* (TrEMBL:Q8SXT3) gene (e = 8e-24) and of the predicted *Anopheles gambiae* protein (agCP6178; protein_id:EAA11923.1) (e = 1e-14) showed significant similarity to that of SWI. It is also important to note that this proteins appeared to be more similar to each other than to SWI itself suggesting Q8SXT3 to be designated as SWI2, the second member of *Drosophila* SWI family of proteins, and EAA11923.1 as its *Anopheles gambiae* simalog *Agam*\SWI2.

The second large group of sequences showing much lower, but evident, similarity to SWI (e ~1e-07) consists of virtually all known homologs of decorin and biglycan proteins. These proteins are characteristic members of the Class I small leucine-reach proteoglycan protein family. These proteins possess a protein core substituted with a single glycosaminoglycan chain near their N-terminus [19]. The most salient feature of decorins and biglycans is the presence of 10 leucine-reach repeats flanked by cysteine-rich regions [20]. Despite the relatively small size of the protein core (~36 kDa), these proteins possess several distinct protein binding activities. Originally, they were shown to bind collagen fibers, playing a role in their assembly. More recently decorin and biglycan were also discovered to interact with the Transforming Growth Factor-β and Epidermal Growth Factor (EGF) receptors, participating in cellular proliferation control [19].

The third group of sequences similar to SWI contains two members of Toll receptor superfamily: Toll related protein *Ae*\Toll1B from *Aedes aegypti* (P_{rm} = 7e-04) and Toll-4 (P_{rm} = 9e-04) from *Drosophila melanogaster*. Both of these were discovered in course of the systematic characterization of Toll-like proteins [21,22]. Although some members of Toll protein family are transmembrane receptors, with a well established role in innate host defense, no evidence is available for the function of either Ae\Toll1B or Toll-4 [21].

We have made all possible pairwise combinations and multiple alignments of these proteins. As expected, the best multiple alignment is between the SWI, *Dpse*\SWI, SWI2 and *Agam*\ SWI2 protein sequences. (Fig. 2B). Examination of this alignment shows that the C-terminal 211 amino acids of SWI are conserved between four proteins (20% identity; 52% similarity) with much higher homology between SWI and *Dpse*\SWI (90% identity). This alignment also clearly shows that *Agam*\SWI2 is more similar to SWI2 than to SWI.

The region of significant similarity between SWI and the decorin/biglycan family proteins is narrower, it is limited to 90 amino acids between positions 396–485 of SWI. In decorins this region is N-terminal (residues 54–141 of murine decorin). This region contains the very end of N-terminal cysteine-rich flanking region and the first two leucine-rich repeats [20]. The region of homology between the proteins does not correspond to the leucine-rich repeats of SWI. The two stretches of amino acids especially conserved between SWI and decorins are also well conserved between SWI, SWI2, *Agam*\SWI2 and *Dpse*\SWI (Fig. 2B). Interestingly enough, such conserved regions have no documented role in known decorin functions, implying that decorins might have some yet undiscovered properties that these proteins have in common with SWI.

The region of homology between the SWI and *Ae*\Toll1B or Toll-4 proteins is situated between positions 249–580 of the SWI amino acid sequence. This corresponds to positions 307–626 of Toll-4 and includes one of its N-terminal cysteine-rich flanking region and two leucine-rich repeats; this region of similarity does not overlap the Toll/IL-1R (TIR) domain of Toll4. This cytoplasmic homology domain is a key feature common of all Toll-like proteins and is indispensable for the activation of the antifungal response in flies by Toll and 18-wheeler (Toll-2) [23]. Comparison of SWI and *Ae*\Toll1B or Toll-4 sequences revealed two distinct conserved CysXCys motifs (Fig. 2B). Both of them are also conserved in SWI2, *Agam*\SWI2 and *Dpse*\SWI. The swi^{t476} mutation affects one of the cysteines

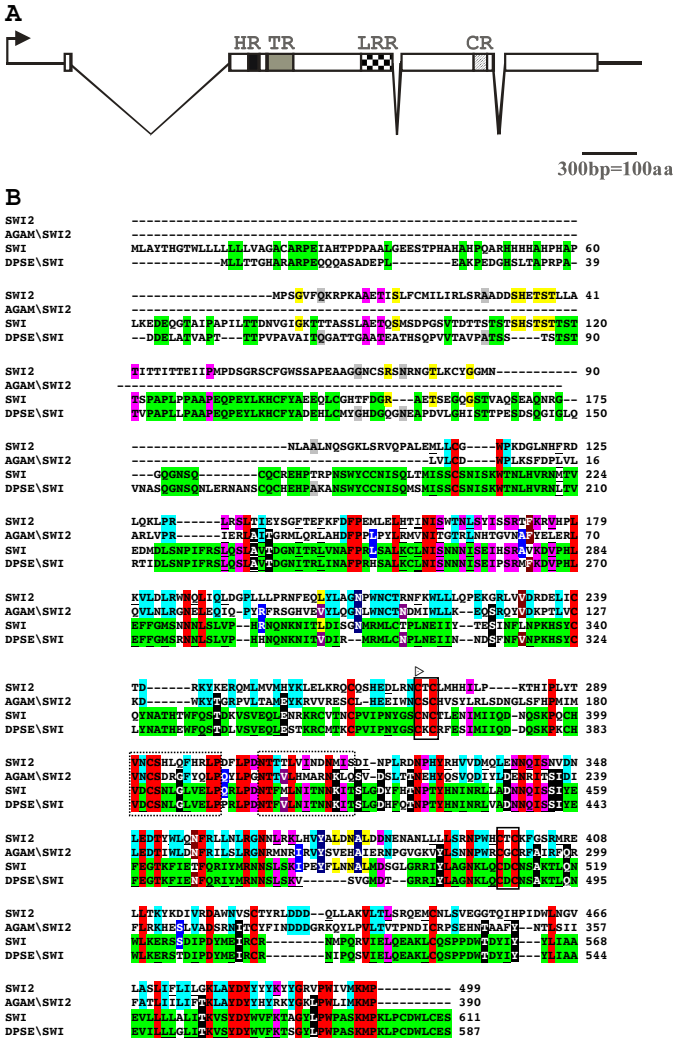

Figure 2

The products of *swi* gene A. The structure of the *swi* gene; exons are shown as boxes and introns as angled lines. Regions of exons encoding various domains are shaded: the histidine-rich domain (HR, black box), the threonine-rich domain (TR, grey box), the leucine-rich repeats (LRR, chess-desk box) and the cysteine-rich region (CR, shaded box). **B.** The alignment of the SWI amino acid sequence with amino acid sequences of its closest simalogs. For the better representation the following color code is applied. Residues in common between all proteins are marked in red, the ones common only between SWI, SWI2 and *Dpse*\SWI in pink and the ones identical only among the *Dpse* \SWI, SWI, and *Agam*\SWI2 are marked in black. Brown marks the amino acids common only between *Dpse* \SWI, *Agam* \SWI2 and SWI2 proteins, dark blue – the ones common only between SWI, *Agam* \SWI2 and SWI2. The pairs of amino acids identical between SWI and *Dpse* \SWI, SWI2 and *Agam* \SWI2, *Agam* \SWI2 and *Dpse* \SWI, *Dpse* \SWI and SWI2, SWI and *Agam* \SWI2 are marked in green, turquoise, violet, grey and blue, respectively. The amino acids similar between all four proteins are underlined. The CysXCys motives are boxed. Note the conserved Cys-379 (marked by a triangle) that is substituted in *swi*^{t476} mutant flies. The regions of the proteins with the highest similarity to decorins are marked by dotted boxes.

in the first CysXCys motif. These CysXCys motifs may be part of a yet uncharacterized functional protein domain (or two distinct but similar protein domains) indispensable for SWI activity and, perhaps, also important for Toll-4 function.

We have used a number of prediction programs in an attempt to learn more about the possible cellular location of the *swi* protein. von Heijne's algorithm [24], as implemented by PSORTII, indicates that the protein may have a cleavable signal sequence (residues 1–19) (a prediction confirmed by the TargetP Server v1.01 [25]. This is consistent with predictions from the Pastuer Institute's implementation of Claros and von Heijne's TopPred algorithm, which indicate a predominantly extracellular protein anchored by cytoplasmic N- and C-termini.

As an alternative approach we have searched the database of *Drosophila* protein interactions recently built on the basis of genome-wide yeast two-hybrid screening [26]. SWI was found to have only one interacting partner, the product of *CG9025* gene. This interaction seems to be very specific and has a very high confidence score (0.9085 with 1.0 being the maximum). *CG9025* is an uncharacterized protein similar to mammalian sex-determination protein encoded by *Fem1b* gene. *CG9025* in turn has eight high confidence interactors from which all with known cellular localization are cytoplasmic. It is interesting to note that one of *CG9025* interacting proteins, the product of *TBPH (CG10327)* has a putative RNA binding domain.

Conclusions

The present molecular analysis, unfortunately, throw remarkably little light on how the SWI protein may function in the ecdysone induced cascade. By far all predictions agree in that SWI is very unlikely to be a nuclear protein. Thus it probably exercises its control of "late" puffing indirectly. Clearly the genetic regulation of ecdysone signaling appears to be much more complex then was previously anticipated.

Methods

DNA clones and fly stocks

The phage genomic clones overlapping the 2B6/B7-8 region have been described by [10]. The coordinates of the FlyBase *Drosophila* genome annotation (Release 3.1) are used in this paper. The cosmid 73D1 was provided by F. Galibert. The genomic library made from DNA of isogenic y^1; $cn^1 bw^1 sp^1$ flies in the CosPeR cosmid vector, the generous gift of J. Tamkun. CosPeR contains *P*-element termini and the *white* gene as a marker, which allows the direct use of a clone for *Drosophila* transformation. The stock $y^1 swi^{t476} w^{pn}$/*FM6B*, carrying the lethal allele of *singed wings*, has been described earlier [7]. Lindsley and Zimm [27] designated an EMS-induced mutation swi^{t467} as hfw^7. However, complementation tests of the hfw^1 allele described by Rayle [9] with any of the *swi* alleles was never done, since the hfw^1 stock had been lost by the time any *swi* mutant allele was obtained. We use the swi^{t467} allele from the Novosibirsk stock collection. The $y^1 swi^{t476} w^{pn}$ mutant chromosome was also rebalanced over *FM7*, $y^{93j} sc^8 w$ *oc pt B*, *P*{w^{+mC}, *act::GFP=pActGFP*}, a "green" balancer (the generous gift of J.-M. Reichhart) to make the $y^1 swi^{t476} w^{pn}$/*FM7*, $y^{93j} sc^8 w$ *oc ptg B*, *P*{w^{+mC}, *act::GFP=pAct-GFP*} stock. This strain allows the easy selection of $y^1 swi^{t476} w^{pn}/y^1 swi^{t476} w^{pn}$ larvae as non-GFP animals.

DNA manipulations and sequence analysis

All general molecular methods were after Sambrook and co-authors [28]. Plasmid and cosmid DNA or PCR products were directly sequenced using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction kit (Perkin Elmer). To amplify fragments of genomic DNA isolated from $y^1 swi^{t476} w^{pn}/y^1 swi^{t476} w^{pn}$ flies the following pairs of primers were used:

Swi-m1 5'-CCGTGCGGCAGTATGACG-3', Swi-m2 5'-TCTGCAGGCTCGTTCGGT-3';

Swi-m3 5'-CAGCGAGTGCTCTGCTGC-3', Swi-m4 5'-GGCTTCGGCATGGAGTGG-3';

Swi-m5 5'-CCTCGTATCCTCACACGC-3', Swi-m6 5'-CTTCCTTCAGCGGCGCAT-3';

Swi-m7 5'-CAGGGCACATGGCTCCTG-3', Swi-m8 5'-GGCGTTCACCAGTCGAGT-3';

Swi-m9 5'-TGCCAGTGCAGGGAGCAC-3', Swi-m10 5'-TGGGGTAGCTCCTCCAGG-3';

Swi-m11 5'-CGAACTACGGCAGCTGCA-3', Swi-m12 5'-AGGGCAGATAGCCAGCCG-3';

Mut1 5'-GAACAGCCCGAGTATCTG-3', Mut2 5'-GTCG-TAGGACACCTTGGT-3';

Mut3 5'-TACCTGGCAGGCAACAAG-3', Mut4 5'-CAACAGTGAGCCGTATCC-3';

Mut5 5'-TCGAGCGCGCACCTAAGT-3', Mut6 5'-ATCCA-GAGGCCAGGACCT-3'.

Sequence analysis

The raw sequence data were analyzed using BLAST [29] and ClustalW [30] computer programs. Screening of *Drosophila pseudoobscura* genome sequence for homology to the *swi* sequence was done with BLAST at http://www.hgsc.bcm.tmc.edu/projects/drosophila/. To look for promoter elements, genomic DNA sequence correspond-

ing to the region of the putative transcription start, *i.e.* the 5' end of the BDGP cDNA clone RE03173 (accession number AY070931) was visually inspected for sequences corresponding to the TATA, INR and DPE consensus sequences, as described by Kutach and Kadonaga [17]. In addition, the entire sequence of *swi* locus was analyzed with PromH program [16]. Unless specially stated the default parameters of all software programs were employed. To browse *D. melanogaster* genome annotations the Apollo (v1.3.4) software was used [31]. The parameters of protein primary structure were calculated using the ProtParam tool at [32] and a protein motif search was performed with InterProScan software [33]. To predict protein localization the following programs were used. PSORTII at [34], the TargetP Server v1.01 [25] at [35] and the Pastuer Institute's implementation of Claros and von Heijne's TopPred algorithm [36]. The *Drosophila* protein yeast two-hybrid interaction network [26] was searched at [37].

Transposon construction

To make the SWI1 construct the 4.4 kb *Bam*HI-*Bam*HI fragment from the λ1479 clone (Fig. 1) was inserted into the CaSpeR3 vector [38] digested with *Bam*HI. The SWI3 transposon was constructed as follows: The 9.4 kb *Xba*I-*Eag*I fragment from the cosmid 73D1 clone was ligated into pBluescript KS (Stratagene), digested with *Xba*I and *Eag*I, to give 73D1-XE-pBl-KS. The 73D1-XE-pBl-KS construct was further digested with *Xba*I and *Nhe*I and the resulting 8.5 kb *Xba*I-*Nhe*I fragment was directly inserted into the CaSpeR2 vector (Pirrotta, 1988), digested with *Xba*I. The SWI4 transposon was constructed in similar way: The 73D1-XE-pBl-KS construct was cut with *Hpa*I and the resulting 4.9 kb fragment was inserted into the CaSpeR4 vector [38], digested with the same restriction enzyme. All constructs were injected into the host strain γ^1, $Df(1)w^{67c23}$ as described by Spradling [39].

Complementation analysis

To test complementation of the swi^{t476} mutation by various transformed fragments of genomic DNA $\gamma^1 swi^{t476} w^{pn}$/*FM6B* females were crossed to transformed γ^1, $Df(1)w^{67c23}/Y$ males carrying an autosomal insertion of the genomic fragment under study. The rescue of the *swi* lethal phenotype was judged by the appearance of *yellow* males in the progeny.

Authors' contributions

YBS performed molecular and genetic experiments, made bioinformatics studies, wrote the paper. TB made molecular and genetic experiments. ESB participated in the design of the study, made control cytogenetic experiments and participated in manuscript editing. MA edited the paper and provided guidance and funding for this project. IFZ conceived and supervised the project. All authors read and approved the final manuscript.

Additional material

Additional File 1

The alignment between the DNA sequence of the* D. melanogaster *swi gene and the sequence of* D. pseudoobscura *contigs 2140 and 2139. *For Drosophila melanogaster DNA sequence the coordinates of the FlyBase genome annotation (Release 3.1) were used. Identical nucleotides are marked with asterisks. The putative TATA box and characteristic CA and T nucleotides at the 5' region of* swi *are shown in red font. The 5' and 3' UTRs are marked in green, the coding part of the* swi *DNA sequence is marked in red. Note that the region of the first intron contains the number of small well conserved motifs. Those containing the NNATTA sequence, characteristic of homeobox protein binding sites, are marked in blue the others are marked in black. The very 3' part of the locus could be aligned with the complementary sequence of the 2139 contig due to the two stretches of extremely conserved DNA.*

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2156-5-15-S1.pdf]

Acknowledgements

Work in Novosibirsk was supported by a grant from Molecular and Cellular Biology Programme of the Russian Academy of Sciences. Work in Cambridge was supported by The Royal Society and a MRC Programme Grant to M. Ashburner, D. Gubb and S. Russell. We are grateful to John Roote, David Gubb, and Steven Russell for the fruitful discussions and advice.

References

1. Ashburner M, Chihara C, Meltzer P, Richards G: **Temporal control of puffing activity in polytene chromosomes.** *Cold Spring Harbor Symp Quant Biol* 1974, **38:**655-662.
2. Thummel CS: **Flies on steroids – *Drosophila* metamorphosis and the mechanisms of steroid hormone action.** *Trends Genet* 1996, **12:**306-310.
3. Bender M, Imam FB, Talbot WS, Ganetzky B, Hogness DS: **Drosophila Ecdysone Receptor mutations reveal functional differences among receptor isoforms.** *Cell* 1997, **91:**777-788.
4. Hall BL, Thummel CS: **The RXR homolog Ultraspiracle is an essential component of *Drosophila* ecdysone receptor.** *Development* 1998, **125:**4709-4717.
5. Belyaeva ES, Vlassova IE, Biyasheva ZM, Kakpakov VT, Richards G, Zhimulev IF: **Cytogenetic analysis of the 2B3-4 – 2B11 region of the X chromosome of *Drosophila melanogaster*. II. Changes in 20-OH ecdysone puffing caused by genetic defects of puff 2B5.** *Chromosoma* 1981, **84:**207-219.
6. Broadus J, McCabe JR, Endrizzi B, Thummel CS, Woodard CT: **The Drosophila βFTZ-F1 orphan nuclear receptor provides competence for stage-specific responses to the steroid hormone ecdysone.** *Mol Cell* 1999, **3:**143-149.
7. Belyaeva ES, Aizenzon MG, Semeshin VF, Kiss II, Koczka K, Baritcheva EM, Gorelova TD, Zhimulev IF: **Cytogenetic analysis of the 2B3-4 – 2B11 region of the X chromosome of *Drosophila melanogaster*. I. Cytology of the region and mutant complementation groups.** *Chromosoma* 1980, **81:**281-306.
8. Belyaeva ES, Zhimulev IF: **Cytogenetic analysis of the 2B3-4 – 2B11 region of the X chromosome of *Drosophila melanogaster*. IV. Mutation at the *swi* (singed wings) Locus Interfering with the Late 20-OH Ecdysone Puff System.** *Chromosoma* 1982, **86:**251-263.
9. Rayle RE: **A new mutant in Drosophila melanogaster causing an ecdysone-curable interruption of the prepupal molt.** *Genetics* 1967, **56:**583. (Abstract)

10. Protopopov MO, Belyaeva ES, Tretyakova IV, Zhimulev IF: **Molecular map of the 2B region of *D. melanogaster* X-chromosome.** *Dros Inf Serv* 1991, **70:**182-184.
11. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, Brandon RC, Rogers YH, Blazej RG, Champe M, Pfeiffer BD, Wan KH, Doyle C, Baxter EG, Helt G, Nelson CR, Gabor GL, Abril JF, Agbayani A, An HJ, Andrews-Pfannkoch C, Baldwin D, Ballew RM, Basu A, Baxendale J, Bayraktaroglu L, Beasley EM, Beeson KY, Benos PV, Berman BP, Bhandari D, Bolshakov S, Borkova D, Botchan MR, Bouck J, Brokstein P, Brottier P, Burtis KC, Busam DA, Butler H, Cadieu E, Center A, Chandra I, Cherry JM, Cawley S, Dahlke C, Davenport LB, Davies P, de Pablos B, Delcher A, Deng Z, Mays AD, Dew I, Dietz SM, Dodson K, Doup LE, Downes M, Dugan-Rocha S, Dunkov BC, Dunn P, Durbin KJ, Evangelista CC, Ferraz C, Ferriera S, Fleischmann W, Fosler C, Gabrielian AE, Garg NS, Gelbart WM, Glasser K, Glodek A, Gong F, Gorrell JH, Gu Z, Guan P, Harris M, Harris NL, Harvey D, Heiman TJ, Hernandez JR, Houck J, Hostin D, Houston KA, Howland TJ, Wei MH, Ibegwam C, Jalali M, Kalush F, Karpen GH, Ke Z, Kennison JA, Ketchum KA, Kimmel BE, Kodira CD, Kraft C, Kravitz S, Kulp D, Lai Z, Lasko P, Lei Y, Levitsky AA, Li J, Li Z, Liang Y, Lin X, Liu X, Mattei B, McIntosh TC, McLeod MP, McPherson D, Merkulov G, Milshina NV, Mobarry C, Morris J, Moshrefi A, Mount SM, Moy M, Murphy B, Murphy L, Muzny DM, Nelson DL, Nelson DR, Nelson KA, Nixon K, Nusskern DR, Pacleb JM, Palazzolo M, Pittman GS, Pan S, Pollard J, Puri V, Reese MG, Reinert K, Remington K, Saunders RD, Scheeler F, Shen H, Shue BC, Siden-Kiamos I, Simpson M, Skupski MP, Smith T, Spier E, Spradling AC, Stapleton M, Strong R, Sun E, Svirskas R, Tector C, Turner R, Venter E, Wang AH, Wang X, Wang ZY, Wassarman DA, Weinstock GM, Weissenbach J, Williams SM, Woodage T, Worley KC, Wu D, Yang S, Yao QA, Ye J, Yeh RF, Zaveri JS, Zhan M, Zhang G, Zhao Q, Zheng L, Zheng XH, Zhong FN, Zhong W, Zhou X, Zhu S, Zhu X, Smith HO, Gibbs RA, Myers EW, Rubin GM, Venter JC: **The genome sequence of *Drosophila melanogaster*.** *Science* 2000, **287:**2185-2195.
12. The FlyBase Consortium: **The FlyBase Database of the *Drosophila* Genome Projects and community literature.** *Nucl Acids Res* 2003, **31:**172-175.
13. Shestopal SA, Makunin IV, Belyaeva ES, Ashburner M, Zhimulev IF: **Molecular characterization of the *deep orange (dor)* gene of *Drosophila melanogaster*.** *Mol Gen Genet* 1997, **253:**642-648.
14. Makunin IV, Shestopal SA, Belyaeva ES, Ashburner M, Zhimulev IF: **Molecular and genetic map comparison of the 2B6-2B7-8 region of the *Drosophila melanogaster* X chromosome.** *Genetika (Moscow)* 1999, **35:**1071-1077.
15. Rubin GM, Hong L, Brokstein P, Evans-Holm M, Frise E, Stapleton M, Harvey DA: **A *Drosophila* complementary DNA resource.** *Science* 2002, **287:**2222-2224.
16. Solovyev V, Shahmuradov I: **PromH: promoters identification using orthologous genomic sequences.** *Nucl Acids Res* 2003, **31:**3540-3545.
17. Kutach AK, Kadonaga JT: **The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters.** *Mol Cell Biol* 2000, **20:**4754-4764.
18. Biggin MD, McGinnis W: **Regulation of segmentation and segmental identity by *Drosophila* homeoproteins: the role of DNA binding in functional activity and specificity.** *Development* 1997, **124:**4425-4433.
19. Iozzo RV: **The biology of small leucine-rich proteoglycans. Functional network of interactive proteins.** *J Biol Chem* 1999, **274:**18843-18846.
20. Scholzen T, Solursh M, Suzuki S, Reiter R, Morgan JL, Buchberg AM, Siracusa LD, Iozzo RV: **The murine Decorin. Complete cDNA cloning, genomic organization, chromosomal assignment, and expression during organogenesis and tissue differentiation.** *J Biol Chem* 1994, **269:**28270-28281.
21. Tauzig S, Jouanguy E, Hoffmann JA, Imler J-L: **Toll-related receptors and the control of antimicrobial peptide expression in *Drosophila*.** *Proc Natl Acad Sci USA* 2000, **97:**10520-10525.
22. Luna C, Hoa NT, Zhang J, Kanzok SM, Brown SE, Imler J-L, Knuds DL, Zheng L: **Characterization of three Toll-like genes from mosquito *Aedes aegypti*.** *Insect Mol Biol* 2003, **12:**67-74.
23. Imler J-L, Hoffmann JA: **Toll receptors in innate immunity.** *Trends Cell Biol* 2001, **11:**304-311.
24. von Heijne G: **A new method for predicting signal sequence cleavage sites.** *Nucleic Acids Res* 1986, **14:**4683-4690.
25. Emanuelsson O, Nielsen H, Brunak S, von Heijne G: **Predicting subcellular localization of proteins based on their N-terminal amino acid sequence.** *J Mol Biol* 2000, **300:**1005-1016.
26. Giot L, Bader JS, Brouwer C, Chaudhuri A, Kuang B, Li Y, Hao YL, Ooi CE, Godwin B, Vitols E, Vijayadamodar G, Pochart P, Machineni H, Welsh M, Kong Y, Zerhusen B, Malcolm R, Varrone Z, Collis A, Minto M, Burgess S, McDaniel L, Stimpson E, Spriggs F, Williams J, Neurath K, Ioime N, Agee M, Voss E, Furtak K, Renzulli R, Aanensen N, Carrolla S, Bickelhaupt E, Lazovatsky Y, DaSilva A, Zhong J, Stanyon CA, Finley RL Jr, White KP, Braverman M, Jarvie T, Gold S, Leach M, Knight J, Shimkets RA, McKenna MP, Chant J, Rothberg JM: **A protein interaction map of Drosophila melanogaster.** *Science* 2003, **302:**1727-1736.
27. Lindsley D, Zimm G: **The genome of *Drosophila melanogaster*.** San Diego: Academic Press; 1992.
28. Sambrook J, Frisch EF, Maniatis T: **Molecular cloning: a laboratory manual.** 2nd edition. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 1989.
29. 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.** *Nucl Acids Res* 1997, **25:**3384-3402.
30. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressivemultiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22:**4673-4680.
31. Lewis SE, Searle SMJ, Harris N, Gibson M, Iyer V, Richter J, Wiel C, Bayraktaroglu L, Birney E, Crosby MA, Kaminker JS, Matthews B, Prochnik SE, Smith CD, Tupy JL, Rubin GM, Misra S, Mungall CJ, Clamp ME: **Apollo: a sequence annotation editor.** *Genome Biology* 2002, **3:**research0082.
32. **ExPASy Proteomics Server** [http://www.expasy.org]
33. Zdobnov EM, Apweiler R: **InterProScan-an integration platform for the signature-recognition methods in InterPro.** *Bioinformatics* 2001, **17:**847-848.
34. **PSORT WWW Server** [http://psort.nibb.ac.jp]
35. **TargetP Server v1.01** [http://www.cbs.dtu.dk/services/TargetP/]
36. **TopPred** [http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html]
37. **Drosophila interaction database** [http://portal.curagen.com/cgi-bin/interaction/flyHome.pl]
38. Pirrotta V: **Vectors for P-mediated transformation in *Drosophila*.** In *Vectors: A survey of molecular cloning vectors and their uses* Edited by: Rodriguez RL, Denhardt DT. Boston: Butterworth; 1988:437-456.
39. Spradling AC: **P element-mediated transformation.** In *Drosophila: A practical approach* Edited by: Roberts DB. Oxford: IRL Press Ltd; 1986:175-198.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

BioMedcentral