# **BMC Genomics**



Research article Open Access

# Identification of *Bacillus anthracis* specific chromosomal sequences by suppressive subtractive hybridization

Kathleen G Dwyer<sup>1</sup>, Janine M Lamonica<sup>1</sup>, Jennifer A Schumacher<sup>1</sup>, Leanne E Williams<sup>2</sup>, Joanne Bishara<sup>1</sup>, Anna Lewandowski<sup>1</sup>, Rajendra Redkar<sup>2,3</sup>, Guy Patra<sup>2,4</sup> and Vito G DelVecchio\*<sup>2</sup>

Address: <sup>1</sup>Loyola Hall of Science, University of Scranton, Scranton, Pennsylvania 18510, USA, <sup>2</sup>Institute of Molecular Biology and Medicine, University of Scranton, Scranton, Pennsylvania 18510, USA, <sup>3</sup>Present address: Schott Glass Technologies Inc. 400 York Avenue, Duryea, PA 18642, USA and <sup>4</sup>Present address: Vital Probes, Inc., 1300 Old Plank Road, Mayfield, PA 18433, USA

Email: Kathleen G Dwyer - kgd301@scranton.edu; Janine M Lamonica - lamonicaj2@scranton.edu; Jennifer A Schumacher - jschumac@mail.med.upenn.edu; Leanne E Williams - williamsl3@scranton.edu; Joanne Bishara - jab419@endeavor.med.nyu.edu; Anna Lewandowski - annalew22@hotmail.com; Rajendra Redkar - rajendra.redkar@us.schott.com; Guy Patra - gpatra@vitalprobes.com; Vito G DelVecchio\* - vdelvecchio@vitalprobes.com

\* Corresponding author

Published: 12 February 2004 BMC Genomics 2004, **5**:15

Received: 16 October 2003 Accepted: 12 February 2004

This article is available from: http://www.biomedcentral.com/1471-2164/5/15

© 2004 Dwyer 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:** Bacillus anthracis, Bacillus thuringiensis and Bacillus cereus are closely related members of the B. cereus-group of bacilli. Suppressive subtractive hybridization (SSH) was used to identify specific chromosomal sequences unique to B. anthracis.

**Results:** Two SSH libraries were generated. Genomic DNA from plasmid-cured *B. anthracis* was used as the tester DNA in both libraries, while genomic DNA from either *B. cereus* or *B. thuringiensis* served as the driver DNA. Progressive screening of the libraries by colony filter and Southern blot analyses identified 29 different clones that were specific for the *B. anthracis* chromosome relative not only to the respective driver DNAs, but also to seven other different strains of *B. cereus* and *B. thuringiensis* included in the process. The nucleotide sequences of the clones were compared with those found in genomic databases, revealing that over half of the clones were located into 2 regions on the *B. anthracis* chromosome.

**Conclusions:** Genes encoding potential cell wall synthesis proteins dominated one region, while bacteriophage-related sequences dominated the other region. The latter supports the hypothesis that acquisition of these bacteriophage sequences occurred during or after speciation of *B. anthracis* relative to *B. cereus* and *B. thuringiensis*. This study provides insight into the chromosomal differences between *B. anthracis* and its closest phylogenetic relatives.

#### **Background**

Bacillus anthracis, the etiological agent of anthrax, is responsible for a serious and often fatal disease of mammalian livestock and humans [8,22]. Its spore-forming capability and highly pathogenic nature have made it one

of the most effective bioterrorism agents. Pathogenic strains of *B. anthracis* contain two plasmids that confer virulence; the toxin-encoding 181.6-kb pXO1 [21,27,29] and the capsule-encoding 94.8-kb pXO2 [12,27,42]. Loss of either plasmid attenuates virulence. Environmental

isolates of virulent *B. anthracis* have been found to contain some pXO2<sup>-</sup> cells indicating that spontaneous loss of this plasmid may occur outside the selective pressures of the host [11,41].

Bacillus cereus is commonly isolated from soil and has been implicated in food-related diseases [9]. Bacillus thuringiensis contains plasmids that encode an array of gene products including the insecticidal crystal proteins [39]. B. anthracis, B. cereus and B. thuringiensis are closely-related members of what has been referred to as the Bacillus cereus-group. Comparison of their 16S rRNA sequences place them within a range expected for members of the same species [3]. In addition, DNA-DNA hybridization [19,37] and pulsed field gel electrophoresis [14] studies have revealed great homology between their chromosomal DNAs. Recent data obtained from multilocus enzyme electrophoresis (MEE) and by multiple-locus sequence typing suggested that B. anthracis, B. cereus, and B. thuringiensis belong to one and the same species, and that salient distinguishing of each member is due to plasmid-borne genes [15].

Suppressive subtractive hybridization (SSH) is a PCRbased DNA subtraction method that enables the identification of genomic sequence differences between various strains or species of bacteria. Tester DNA is defined as that in which the differences are being sought relative to a driver DNA. For example, SSH has been successfully used to identify unique genomic sequences for specific strains of Helicobacter pylori [2], uropathogenic E. coli [18,43], and colonizing Campylobacter jejuni [1]. This technique has also been used to detect unique genomic sequences found in E. coli but not in Salmonella typhimurium [4], in pathogenic Burkholderia pseudomallei but not in nonpathogenic B. thailandensis [36], in pathogenic Burkholderia mallei but not in B. thailandensis [7] and in B. anthracis but not in B. cereus 14579T, B. thuringiensis Al Hakam, or B. cereus 3A [32]. This methodology has not only identified genetic differences among different strains or species of bacteria but also defined genes that contribute to the virulence of an organism. These virulence-related genes are often localized on the chromosome. Such clusters have been referred to as pathogenicity islands [13].

Herein we report the application of SSH to elucidate genomic differences of *B. anthracis* in comparison to *B. cereus* and *B. thuringiensis*, thereby identifying DNA sequences that are unique to the *B. anthracis* chromosome. In contrast with *B. anthracis* described as one of the most monomorphic species within the *B. cereus* group, *B. cereus* and *B. thuringiensis* present a vast range of genetic diversity. For this study we used the genomic DNA of *B. cereus* 14579T or *B. thuringiensis* 10792T as driver DNA because these are representive of the respective species. In

the subsequent screening of the *B. anthracis*-specific chromosomal sequences, we used several *Bacillus* sp. strains more closely-related to *B. anthracis* [15,31,33]. The *B. anthracis*-specific chromosomal sequences found elucidate genomic differences of these three closely-related bacteria. These sequences can be exploited for the development of more stable and robust chromosome-specific markers important for accurate identification of *B. anthracis* and unique from those previously reported [30-32].

#### Results

# Screening of B. anthracis/B. cereus and B. anthracis/B. thuringiensis SSH libraries

SSH libraries were generated using RsaI-digested genomic DNA from B. anthracis strain 7700 (pX01-, pX02-) as the tester and either RsaI-digested genomic DNA from B. cereus 14579T or B. thuringiensis 10792T as the drivers. All clones derived from the B. anthracis/B. cereus (Ba/Bc) library were given a numerical designation, whereas those of the B. anthracis/B. thuringiensis (Ba/Bt) library commenced with an alphabetical designation. The subsequent screening process used for selection of clones containing B. anthracis-specific chromosomal DNA is outlined in Table 1.

Forty four hundred Ba/Bc and Ba/Bt clones were prescreened for B. anthracis-specific sequences by comparing duplicate sets of colony filters. One set was probed with 32 P-labeled tester sequences, while the other set was probed with the corresponding 32 P-labeled driver sequences. Two hundred and twenty-four Ba/Bc and 189 Ba/Bt clones displayed more intense hybridization when probed with B. anthracis genomic DNA than with that of B. cereus or B. thuringiensis. Recombinant plasmid DNA of the 413 SSH clones selected was digested with RsaI, electrophoresed on agarose gels, Southern blotted, and then probed with the corresponding 32 P-labeled driver genomic DNA. Insert DNA sizes ranged between 100-2000 bp, as predicted from analysis of the frequency with which RsaI digests each of these Bacillus genomic DNAs. Ninety-six Ba/Bc and 59 Ba/Bt clones had inserts that were undetected by the driver DNA probe.

The insert DNA from each of these 155 candidate tester-specific clones was isolated, <sup>32</sup> P-labeled and used to probe DNA blots containing the *Hin*dIII-digested genomic DNAs of (1) *B. anthracis* strain 7700, (2) *B. thuringiensis* strain 97-27, (3) *B. cereus* strain 14579T and (4) *B. thuringiensis* strain 10792T. *B. thuringiensis* strain 97-27 genomic DNA was used because it is extremely similar to that of *B. anthracis* [15,16,28,32,33]. Thirty-two of *the Ba/Bc* clones and 20 *Ba/Bt* clones detected sequences of *B. anthracis* 7700 DNA only (data not shown).

Table I: Progressive screening of SSH libraries.

SSH library	B. anthracis 7700 vs. B. cereus 14579T	B. anthracis 7700 vs. B. thuringiensis 10792T
Number SSH clones tested	2400	2000
<b>Comparative colony filters selection</b> (detection by tester probe > detection by driver probe)	224	189
Clone DNA blot analysis (selection of clones not detected by driver probe)	96	59
Initial genomic blot selection (blot contained <i>B. anthracis</i> 7700, <i>B. cereus</i> 14579T, <i>B. thuringiensis</i> 10792T and <i>B. thuringiensis</i> 97-27 genomic DNA probed with SSH clone insert DNA)	32	20
Survey genomic DNA blot selection of B. anthracis- specific clones (blot contained 4 strains of B. anthracis and 8 other strains of B. cereus and B. thuringiensis – see Figure 1 and 1)	27	12

Survey Southern blots containing samples of the *Hin*dIII-digested genomic DNAs of *B. anthracis* 7700, *B. anthracis* AC1, *B. anthracis* RA3, *B. anthracis* 7611, *B. thuringiensis* 97-27, *Bacillus* Ba813+ S8553/2, *Bacillus* Ba813+ III-BL, *B. cereus* F17289, *B. cereus* 14579T, *B. thuringiensis* CT43, *B. thuringiensis* huazhongensis, *and B. thuringiensis* 10792T were probed with the <sup>32</sup> P-labeled insert DNA of each of the 52 clones found to be *B. anthracis*-specific after initial genomic DNA blot screening. An example of this more extensive analysis is provided in Figure 1, demonstrating detection of sequence homology in the four *B. anthracis* genomic DNA samples only when probed with the insert of *Ba/Bt* clone E4. Clone E4 did not detect any homology in any of the *B. cereus*, *B. thuringiensis*, or other *Bacillus* species included on this blot.

A summary of the 52 clones tested in this manner is presented in Table 2 (see Additional file 1). Thirty-nine of the 52 clones were observed to be completely specific for *B*. anthracis (highlighted in bold type). All probes but one consistently hybridized and detected a similar pattern of restriction fragments found in all four B. anthracis genomic DNA samples which included pathogenic B. anthracis strains RA3 and 7611. Clone 140 was the exception in that it detected some polymorphism in the restriction fragment pattern displayed in the B. anthracis genomic DNA samples (data not shown). Thirteen SSH clones did detect nucleotide sequences in strains other than B. anthracis. Clones 14, 34, 140, O4, Z6, L7, and O7 hybridized with one other; clones 151, 227, G and R3 hybridized with two other; and clones A4 and K4 hybridized with three other B. cereus and / or B. thuringiensis genomic DNAs on the survey blot, exclusive of the driver strains.

# Sequence analysis of B. anthracis-specific clones

Sequence determination and chromosomal location of B. anthracis-specific clones

The nucleotide sequence was determined for each insert of the 39 clones that were found to be *B. anthracis*-specific. Seventeen of the 39 clones demonstrated double, triple and as much as quadruple redundancy, finally sorting into 7 different RsaI-fragments (Table 2 and 3, see Additional file 1 and 2). Therefore, these 39 clones delineate 29 different RsaI fragments of the B. anthracis 7700 chromosome. BLAST analysis [24] sorted these 29 RsaI fragments into 6 different regions on the A2012 strain [34] chromosome (Figure 2). Areas on the chromosome delineated either by 13 or more clones, 3 clones, or one clone were designated as regions, clusters, or orphans, respectively. Regions I and II encompassed eight and thirteen different B. anthracis-specific RsaI fragments that were identified by thirteen and sixteen clones, respectively. Three other locations were found to encompass two or three different B. anthracis-specific RsaI fragments. Clones 81/231 and F7 define Cluster I; clones 230, 84 and 135 define Cluster II; and clones 9/49 and 221 define Cluster III. Single B. anthracis-specific orphan clone M7 identified the remaining chromosomal area.

Annotation of Region I and II delineated by the B. anthracis-specific clones

Open reading frames (ORFs) overlapping and/or bordering *B. anthracis*-specific sequences were identified [26] and the amino acid sequences derived from these ORFs were compared to other protein sequences [25]. The results of these analyses are summarized in Table 3 (see Additional file 2).

Region I, delineated by 8 *B. anthracis*-specific *Rsa*I-fragments and consisting of an area of 14.5 kb, contains 10 potential ORFs. Nine of these ORFs demonstrated

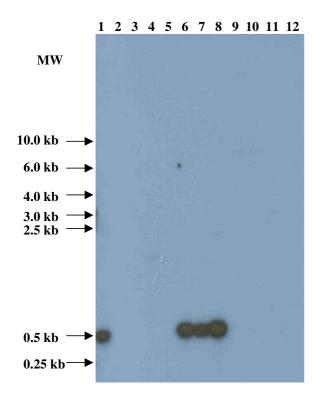


Figure I
Survey genomic DNA blot probed with clone E4 insert
DNA. A total of I μg of genomic DNA digested with 70 U of
Hindll1 to completion was used per lane. Lanes I, Bacillus
anthracis 7700; 2, Bacillus thuringiensis (97-27); 3, Bacillus cereus
F17289; 4, Bacillus thuringiensis huazhongensis; 5, Bacillus thuringiensis CT43; 6, Bacillus anthracis AC1; 7, Bacillus anthracis
RA3; 8, Bacillus anthracis 7611; 9, Bacillus spp. Ba813 (S8553/
2); 10, Bacillus spp. Ba813 (III-BL); 11, Bacillus cereus 14579;
12, Bacillus thuringiensis 10792. The positions of molecular
weight markers are indicated on the left side.

significant sequence similarity with genes encoding enzymes and proteins involved in cell wall biosynthesis or export of polysaccharides in other bacterial species. The G + C content of the DNA encompassing the SSH-selected sequences in this region was 32.1% which is not significantly different from that of the entire *B. anthracis* chromosome, reported to be 35.4% [35].

Region II, which encompassed an area of 24.5 kb flanked by clones H4 and 150b, contains 40 potential ORFs. Nineteen of these ORFs had sequence homology with a variety of bacteriophage-related gene products. Ten of the 13 B. anthracis-specific RsaI-fragments overlapped with these bacteriophage-related sequences. No functional match could be assigned to 21 of the potential ORFs interspersed with the 19 bacteriophage-related ORFs. However, detection of 13 of these potential ORFs by overlapping SSH clones suggests they contain B. anthracisspecific DNA. Analysis of the flanking chromosomal sequences reveals other bacteriophage-related ORFs 7.5 kb upstream and 5 kb downstream of this region, including a third ORF homologous to phage integrase gene sequences. Outside of these limits, potential ORFs had homology with housekeeping genes found to be conserved in other Bacillus species. For example, ORF BA\_1004, homologous with the YfjO gene (RNA methyltransferase) of B. subtilis, was found closest upstream, while ORF BA\_1064 homologous to the RocR gene (transcriptional activator of the arginine utilization operon) of B. subtilis was found downstream. The G + C content of the chromosomal region defined by these B. anthracis-specific clones was 35.3%. Again this is not significantly different from the average G + C content of the B. anthracis chromosomal DNA. In the recently published genome sequence of B. anthracis Ames strain Region II is contained within the area designated " B. anthracis phage lambda Ba01" [35].

# Annotation of clustered and isolated genes delineated by the B. anthracis-specific clones

Cluster III, encompassing an area of 30 kb, was defined by 2 B. anthracis-specific RsaI-fragments. This region contains 43 potential ORFs. Nineteen of the ORFs had sequence homology with bacteriophage-related gene products. Bacillus anthracis-specific RsaI-fragments derived from SSH clones 9/49 and 221, overlapped 2 of these bacteriophage-related sequences. No sequence similarity could be assigned to 17 of the 43 potential ORFs located in this region, three of which had sequences that overlapped with SSH clones. Also found in this cluster were 3 ORFs (BA\_4580/ YndN gene B. subtilis; BA\_4583/ dUTPase; BA\_4585/ RspF gene B. subtilis) encoding B. anthracis homologues for gene functions probably conserved in most species of Bacillus. No SSH-selected clones mapped to these genes. Analysis of the chromosomal sequences bordering Cluster III revealed other bacteriophage-related ORFs up to 2.6 kb upstream and 7 kb downstream. Two ORFs that were homologous to a phage integrase gene sequences were found, one in each flank. Open reading frames outside of these limits demonstrated homology with housekeeping genes found conserved in other Bacillus species. For instance, ORF BA 4538, the FtsK SpoIIIE gene, was found closest upstream; while ORF BA\_4601, homologous to the YlbM gene of B. subtilis was situated closest downstream. The G + C content of the DNA encompassing the SSH clone sequences mapped to this

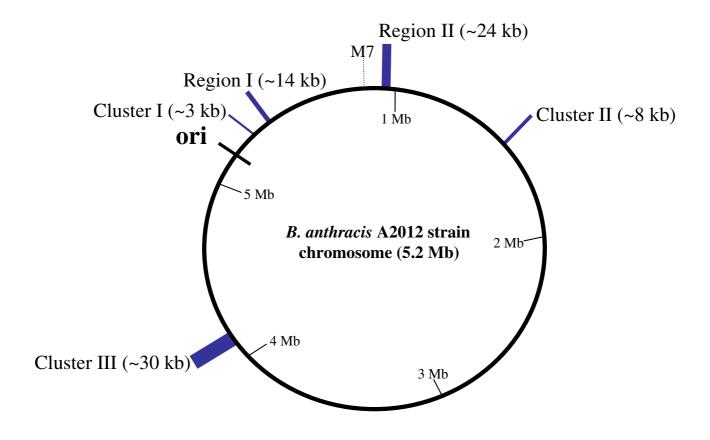


Figure 2
Distribution of the SSH-selected region, cluster and orphan areas on the *B. anthracis* A2012 strain chromosome. Megabase (Mb) distances are estimated on the inside of the circular chromosome.

third cluster was determined to be 35.0%. Again, this is not significantly different from the average G + C content of the remainder of the *B. anthracis* chromosome. Cluster III is contained within the area designated " *B. anthracis* phage lambda Ba03" of the *B. anthracis* Ames strain genome [35].

Cluster I contains a bacteriophage-related sequence that was detected by *B. anthracis*-specific clones 81/231, while clone F7 sequences overlapped with 3 potential ORFs encoding hypothetical proteins found to have no homology with any known proteins. These clones are contained within the area designated " *B. anthracis* phage lambda Ba02" of the *B. anthracis* Ames strain genome [35].

In Cluster II clone 230 was derived from an ORF displaying full-length homology with the wapA gene of B. subtilus

encoding a cell wall-associated protein. Clones 84 and 135, also located in Cluster II, overlapped potential ORFs encoding hypothetical proteins with no known function.

Finally, orphan clone M7 overlapped with an ORF demonstrating homology to a gene encoding a Nitrogen regulation protein in *B. subtilis*.

Comparison with published B. anthracis-specific SSH-selected clones A comparable SSH analysis was recently reported using the genomic DNAs of *B. anthracis* as the tester and of either *B. cereus* 14579T, *B. thuringiensis* Al Hakam or *B. cereus* 3A as the driver [32]. Of the 28 *B. anthracis*-specific clones reported only 3 share nucleotide sequences with 3 of the 29 *B. anthracis*-specific *Rsa*I fragments presented in this study (clones M. Ctg032, dhp61.183, and dhp64.188 with clones 55/67/212/235, 81/231 and 9/49,

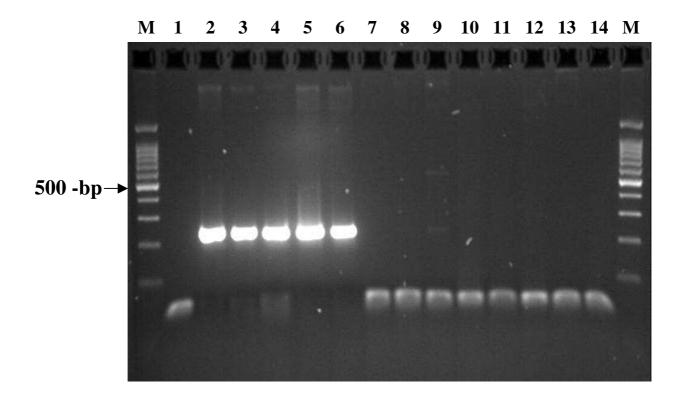


Figure 3
Gel electrophoresis analysis of PCR products generated after reaction of various Bacillus genomic DNAs with clone E4 derived primers. Lanes, M: Molecular weight markers,100-bp DNA ladder; I, Negative Control (no DNA); The samples are: 2, Bacillus anthracis A5; 3, Bacillus anthracis Sterne; 4, Bacillus anthracis A58; 5, Bacillus anthracis A16; 6, Bacillus anthracis 0074; 7, Bacillus cereus 14579; 8, Bacillus cereus 23261; 9, Bacillus thuringiensis 10792; 10, Bacillus thuringiensis B8; 11, Bacillus spp. Ba813 (9594/3); 12, Bacillus spp. Ba813 (S8553/2); 13, Bacillus mycoides 21929; 14, Bacillus thuringiensis 35646.

respectively). Therefore, 26 of the SSH clones reported here offer newly defined B. anthracis-specific chromosomal sequences. Five clones found in the other study (dhp61.178, dhp77.002, M. Ctg037, R. Ctg177, and dhp73.022) were proximal to (within the same ORF) 5 of the B. anthracis-specific RsaI fragments listed in this paper (clones 03, 5,198,150b, and 9/49, respectively). This comparison is summarized in the last column of Table 3 (see Additional file 2), as well as the letter designation of the corresponding genomic island. Thus, the genomic islands designated as A, B, C, and E [32] correlate with the locations of Cluster I, Region I, Region II and Cluster III, respectively. In addition, clone dhp77.047 is located in Region II upstream of clone H4, but within the confines of the bacteriophage-related sequences (phage lambda Ba01). Finally, clones dhp61.181 and R. Ctg152 are found within the area of Cluster III with phage lambda Ba03 sequences [35].

# Diagnostic screening of Bacillus strains by PCR amplification using clone E4-specific primers

B. anthracis-specific clone E4 sequences were used to design oligonucleotide primers for PCR screening of a larger sampling of Bacillus strains. The expected size of the anticipated PCR product was 235 bp. PCR amplifications were performed on (i) 14 B. anthracis strains isolated from different geographical locations and harboring both, one, or neither of the pX01 and pX02 plasmids, (ii) 34 closely-related B. cereus and B. thuringiensis strains, including 14 Bacillus sp. Ba813+ transition strains, and (iii) a B. mycoides strain (see Table 2). The same set of strains was used to evaluate a B. anthracis chromosome-specific PCR assay that exploited polymorphisms found in the rpoB

Table 4: Bacillus species used for SSH, DNA blot and PCR screening, respectively.

Bacillus Species	Strain ID	p <b>X</b> 01	pX02
Bacillus anthracis	7700	-	-
Bacillus cereus	14579T*	N/A	N/A
Bacillus thuringiensis	10792T*	N/A	N/A
Bacillus anthracis	ACI	+	- -
Bacillus anthracis	RA3	+	+
Bacillus anthracis	7611	+	+
Bacillus cereus	F17289*	_	_
Bacillus thuringiensis	huazhongensis*	N/A	N/A
Bacillus thuringiensis	CT43*	N/A	N/A
Bacillus thuringiensis	97-27*	-	-
Bacillus sp. Ba813+	S8553/2*	_	_
Bacillus sp. Ba813+	III-BL	_	_
Bacillus anthracis	Sterne	+	<u> </u>
Bacillus anthracis	A5	+	+
Bacillus anthracis	AI2	+	+
Bacillus anthracis	AI6	+	+
Bacillus anthracis	AI8	· -	+
Bacillus anthracis	A18 A19	-	+
Bacillus anthracis	A19 A30	- +	+
Bacillus anthracis	A32	+	+
Bacillus anthracis	A33	+	+
Bacillus anthracis	A39	+	+
Bacillus anthracis	A43	+	+
Bacillus anthracis	A58	-	-
Bacillus anthracis	9066	+	+
Bacillus anthracis	0074	+	<del>-</del>
Bacillus cereus	13472	N/A	N/A
Bacillus cereus	19637	N/A	N/A
Bacillus cereus	23261	N/A	N/A
Bacillus cereus	F17202	-	-
Bacillus cereus	F16959	-	-
Bacillus cereus	UW-85	N/A	N/A
Bacillus thuringiensis	35646	N/A	N/A
Bacillus thuringiensis	B8	N/A	N/A
Bacillus thuringiensis	HI	N/A	N/A
Bacillus thuringiensis	H3	N/A	N/A
Bacillus thuringiensis	H4	N/A	N/A
Bacillus thuringiensis	HI3	N/A	N/A
Bacillus thuringiensis	H29	N/A	N/A
Bacillus thuringiensis	H39	N/A	N/A
Bacillus thuringiensis	H43	N/A	N/A
Bacillus sp. Ba813+	#4	-	-
Bacillus sp. Ba813+	#13	-	-
Bacillus sp. Ba813+	#17	-	-
Bacillus sp. Ba813+	#18	-	-
Bacillus sp. Ba813+	#2 l	-	-
Bacillus sp. Ba813+	#29	-	-
Bacillus sp. Ba813+	#3403	-	-
Bacillus sp. Ba813+	#6681	-	-
Bacillus sp. Ba813+	9594/3	-	<del>-</del>
Bacillus sp. Ba813+	11614/2	-	<u>-</u>
Bacillus sp. Ba813	IB/A	_	-
Bacillus sp. Ba813-	Zimbabwe	_	_
Bacillus mycoides	21929	_	<u>-</u>
Ducinus Hiycoldes	21/2/	-	-

<sup>\*</sup>used for both DNA blot and PCR assay; N/A – not applicable

gene [31]. Without exception, all 14 *B. anthracis* strains analyzed resulted in the amplification of the expected size PCR product, while all 35 closely-related *B. cereus, B. thuringiensis* and *B. mycoides* strains analyzed did not yield any significant amplification product. An electrophoretic representation of these results in presented in Figure 3.

#### **Discussion**

Two SSH libraries were generated that identified chromosomal DNA sequences found in *B. anthracis* and not in *B.* cereus or B. thuringiensis. B. anthracis 7700 DNA was used as the tester DNA, and total genomic DNA of either B. cereus 14579T or B. thuringiensis 10792T served as the drivers. Progressive screening by colony filter and Southern blot analyses identified 39 clones that were specific for the *B. anthracis* chromosome relative to the driver strains of *B.* cereus and B. thuringiensis and seven other more closelyrelated strains of B. cereus and B. thuringiensis. Since B. anthracis 7700 is cured of all plasmids, only chromosomal sequences were identified. The B. anthracis-specific clones' insert DNAs were sequenced, mapped to the B. anthracis chromosome, and annotated. Seventeen clones shared insert sequence, resulting in a final count of 29 B. anthracis-specific RsaI-fragment inserts. Interestingly, four of these seven redundant inserts were derived from both driver libraries. The 29 insert sequences are distributed over 6 different areas on the B. anthracis chromosome, with 8 and 13 of them being found in Regions I and II, respectively. Seven insert sequences were distributed into 3 clusters. The remaining clone was located singly on the chromosome.

Region I contains ORFs that had significant similarity with genes encoding enzymes and proteins involved in cell wall polysaccharide biosynthesis. Typically in this process, phosphorylated sugar molecules are first activated by conversion into nucleotide sugar derivatives. Enzymes, such as epimerases and dehydrogenases, may then modify these nucleotide sugar precursors into the distinctive sugar subunit(s) characteristic of the cell wall of a particular bacterial species. Transferases may then catalyze the sequential transfers of the sugar residues from their nucleotide carriers to a membrane-bound acceptor, usually a lipid carrier, forming the characteristic carbohydrate repeating unit. This first stage of cell wall polysaccharide biosynthesis usually occurs in the cytoplasm at the cytoplasmic side of the plasma membrane. At this point, an ABC transporter cassette may permit the lipidlinked carbohydrate-repeating unit to pass through the membrane. This two-component system consists of a translocation permease protein with membrane-spanning domains and an associated ATP-binding protein that energizes the permease.

In Region I potential ORFs BA-0356/7/8 and BA\_0361/2, demonstrating sequence homology to genes encoding an epimerase and a dehydrogenase, respectively, were identified by SSH clones 190/N, Y4 and 142/I6. Potential ORFs BA\_0365/6 and BA\_0371, detected by SSH clones 55/67/ 212/235, 66 and P3, demonstrated sequence similarity to genes encoding different types of transferases. Clone 190/ N and Y4 inserts also overlapped with potential ORFs BA 0359 and BA 0360, which displayed significant sequence similarity to other known ABC transporter cassette proteins. The best match was with B. subtilis genes tagG and tagH, involved in the translocation of cell wall component teichoic acid. However, the corresponding *B*. anthracis potential ORFs cannot encode equivalent functional homologs. B. subtilis cell wall-associated teichoic (TA) and teichuronic acid (TUA) are absent in *B. anthracis* [23]. The remaining potential ORFs (BA\_0363/4 and BA\_0368/9 and BA\_0370 detected by clones 201, 55/67/ 212/235, O3, G and P3) in this region were found to share homology with proteins involved in cell wall synthesis whose specific functions are currently unknown. These same ORFs were also found to be specific for B. anthracis A2012 by whole genome comparison with B. cereus 14579T [17].

The Gal-NAG polysaccharide, derived from sugar components D-galactose and N-acetyl-D-glucosamine, has been demonstrated by monoclonal antibodies to be uniquely associated with the cell wall of B. anthracis and a few B. cereus strains [10]. This result was also obtained for B. anthracis pX01- and pX02-strains, indicating that the enzymes and proteins necessary for synthesis and export of this B. anthracis-specific polysaccharide are located on the chromosome. It is possible that the Gal-NAG polysaccharide, or other components associated uniquely with the cell wall of B. anthracis, are synthesized by enzymes and exported by proteins encoded by genes specific to B. anthracis. Such specific genes might be detected by SSH as the genes associated with Region I have been. Substantiation of this hypothesis will require functional analysis of the Region I genes. It is interesting to note though that genes encoding cell wall biosynthesis proteins have been selected in many other subtractive hybridization experiments [1,4,5,7,18,36].

Bacteriophage-related sequences dominated *B. anthracis*-specific Regions II and Cluster III delineated by 15 SSH-selected *RsaI* fragments. Potential ORFs within and flanking Regions II and Cluster III included some identified as bacteriophage integrases/recombinases. Such enzymes could have facilitated the integration of these sequences into the chromosome of *B. anthracis*. The putative ORFs of Regions II and Cluster III encoding proteins of unknown function may be bacteriophage sequences that have degenerated beyond recognition by database comparison.

PCR analysis indicated that the clone E4 sequence of Region II is conserved in and is specific for *B. anthracis*. Also, phage lambda Ba01 and Ba03 sequences, which include Regions II and Cluster III, respectively, were found by comparative genome hybridization between the B. anthracis Ames strain and 19 other members of the B. cereus group to be absent in all but B. anthracis [35]. These results suggest that acquisition of bacteriophage-related Region II and Cluster III occurred during or after speciation of B. anthracis relative to B. cereus and B. thuringiensis, i.e. during or after the acquisition of plasmids pX01 and pX02, the distinguishing attributes of B. anthracis. The G+ C content of Region II (35.3 %) and Cluster III (35.0%) do not differ significantly from the genome-wide average of 35.4%. This indicates that these regions may have been (i) acquired from bacteriophage with similar G + C content or (ii) acquired so long ago that they amalgamated with the B. anthracis genome, attaining an indistinguishable G + C content by amelioration [20].

For the most part the remaining *B. anthracis*-specific clones correlated again with ORFs encoding (*i*) cell surface molecules (detected by clone 230), (*ii*) other bacteriophage-related sequences (detected by clone 81/231), (*iii*) conserved hypothetical proteins (detected by clones 135 and M7) or (iv) hypothetical proteins of unknown function (detected by clones 81/231, F7, 84,135, and M7).

In conclusion, the present study revealed differences between the chromosome of *B. anthracis* and those of eight strains of *B. cereus* and *B. thuringiensis*. The general differences were seen in genes that may encode proteins related to cell wall synthesis and bacteriophages as well as ORFs encoding proteins of unknown function.

#### **Methods**

### Bacterial strains and genomic DNA preparation

The *B. anthracis, B. cereus* and *B. thuringiensis* strains used in this study are listed in Table 2. The bacterial genomic DNA was extracted by a method described elsewhere [6].

### Suppressive subtractive hybridization

SSH was performed using the PCR-Select Bacterial Genome Subtraction Kit as per manufacturer's instructions (Clontech, Palo Alto, CA). Generation of SSH libraries also required the use of the TA Cloning Kit (Invitrogen, Carlsbad, CA) with vector pCR2.1 as per manufacturers instructions.

#### Screening of suppressive subtractive libraries

Transformed *E. coli* TOP10F' cells were plated out on selective LB (Luria-Bertani) medium containing X-Gal, IPTG and either ampicillin (100  $\mu$ g/ml) or kanamycin (50  $\mu$ g/ml) [38]. White colonies were streaked onto master

plates and two sets of nitrocellulose filters (Millipore, Billerica, MA) on LB/kanamycin plates. Filters were treated with 10% (w/v) SDS, denature (0.4 N NaOH, 0.6 M NaCl), neutralize (0.5 M tris-HCl, pH 7.5, 1.5 M NaCl) and 2X SSPE (20 mM EDTA, 0.2 M sodium phosphate, pH 7.4, 3.6 M NaCl) solutions for 5 min. each, air dried and baked at 80°C under vacuum for 2 hrs. Filters were prehybridized in 5 × Denhardt solution at 65°C for at least four hours before hybridization in 5 × SET (0.75 M NaCl, 0.1 M Tris-HCl, pH 7.8, 5 mM EDTA), 1 × Denhardt, 0.5 % (w/v) SDS, 12.5% dextran sulfate, and 25 µg/ml denatured salmon sperm DNA at 65°C overnight. RsaIdigested B. anthracis, B. cereus, and B. thuringiensis genomic DNAs were labeled with 32 P (Random-Primer Labeling System, Roche Molecular Biochemicals, Indianapolis, IN) and used to probe the two sets of colony filters for each SSH library.

# Plasmid DNA preparation and insert fragment isolation

Plasmid DNA was isolated from 1–5 ml LB or TB/kanamycin cultures by either the boiling miniprep method [38] or the Plasmid Mini Kit (Qiagen, Valencia, CA). Insert fragments to be used as probes for genomic DNA blots were separated from vector sequences on 0.7% agarose gels and extracted by electroelution [38].

### **DNA** blot analysis

One microgram of DNA was digested with the appropriate restriction enzyme (Life Technologies, Rockville, Md.), subjected to electrophoresis in 0.7% (w/v) agarose gels and transferred to Genescreen Plus nylon membranes (New England Nuclear Life Science Products, Inc., Boston, MA) as per manufacturer's instructions. DNA blots were prehybridized and hybridized at 65°C in 330 mM sodium phosphate, pH 7, 10 mM EDTA, 5% (w/v) SDS, 10% (w/v) dextran sulfate and salmon sperm DNA at 25  $\mu$ g/ml. Filters were washed twice for 30 min. intervals at 65°C in 2 × SET, 0.5% SDS and autoradiographed.

#### Sequencing and bioinformatics

Each clone was sequenced in duplicate with the M13 forward and reverse primers using an ABI377 automated sequencer (Applied Biosystems, Boston, MA) and the BigDye terminator ready reaction kit (Perkin-Elmer Applied Biosystems, Foster City, CA). The nucleotide sequences were edited and assembled with the Sequencing Analysis 3.0 and AutoAssembler 3.1.2 programs. The SSH DNA sequences were subsequently localized on the A2012 strain chromosome [24] and annotated [25,26].

# PCR amplification of the sequences identified by SSH clone E4

The sequences of primer E4F (5'-AGC AAA AGC TAA CTG CTC GGG-3') and E4R (5'-TCT ACC CCT TCA GGA TTA GCG-3') were derived from SSH clone E4 sequences

found 235 bp apart. The PCR reaction of 25 µl consisted of 10 mM Tris-HCl (pH 9.2), 75 mM KCl, 1.5 mM MgCl<sub>2</sub> , 0.2 mM of dNTPs (Boehringer Mannheim Corp., Indianapolis, IN), 1  $\mu$ M of each of primers E4F and E4R, 0.05 U/ μl of AmpliTaq® DNA polymerase (Perkin Elmer Corp., Foster City, CA), and 100 ng of genomic DNA template. Amplification was performed in a GeneAmp PCR System 2400 (Perkin Elmer) and the cycling conditions were: initial denaturation at 94°C for 5 min, followed by 35 cycles of 94°C for 30 sec., 64°C for 30 sec., and 72°C for 30 sec., and finishing with a final extension at 72°C for 7 min. The PCR assays were performed on 14 B. anthracis strains isolated from different geographical locations and harboring any combination of the pXO1, pXO2 plasmids, as well as 35 closely-related Bacillus strains, including 14 Bacillus spp. Ba813+ strains (Table 2). The amplicons were electrophoresed in 2% (w/v) SeaKem GTG agarose (FMC Bioproducts, Rockland, ME) with 40 mM Tris-acetate-1 mM EDTA (pH 8.3) running buffer and visualized by ethidium bromide staining.

#### **Authors' contributions**

KGD solely made both SSH libraries, screened the first 800 clones, did all <sup>32</sup> P hybridizations for colony filters and Southern blots, performed all of the annotation analysis, and served as primary author on the text of this paper. JAS was the primary co-researcher for screening of the second 800 SSH clones batch. AL acted as primary co-researcher for screening of the third 800 SSH clones batch. JML acted as primary co-researcher for screening of the last 2000 SSH clones. JB performed SSH clone E4 PCR detection analysis of 49 *Bacillus* strains. GP provided the *Bacillus* DNA with the help of LW. RR served as consultant early in the project. GP and VGD supervised the sequencing and acted as consultants.

#### **Additional** material

### **Additional File 1**

Table 2. Survey genomic DNA blot results. Indication of detection (+) or no detection (-) by each <sup>32</sup> P-labeled insert of the 52 SSH-selected clones used to probe survey Southern blots containing samples of HindIII-digested genomic DNAs of B. anthracis 7700, B. anthracis AC1, B. anthracis RA3, B. anthracis 7611, B. thuringiensis 97-27, Bacillus Ba813+ S8553/2, Bacillus Ba813+ III-BL, B. cereus F17289, B. cereus 14579T, B. thuringiensis CT43, B. thuringiensis huazhongensis, and B. thuringiensis 10792T.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-5-15-S1.doc]

# Additional File 2

Table 3. Annotation of the 29 B. anthracis-specific SSH clones. For each of the 29 B. anthracis-specific SSH clones reported in this paper this table provides their B. anthracis A2012 chromosome coordinates, overlapping ORF designations, closest homolog genes with putative function, e value, accession number and comparative results with other recently published B. anthracis SSH study [32].

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-5-15-S2.doc]

### **Acknowledgements**

This project was supported by research grants no. DABT63-99-I-002 from the Defense Advanced Research Projects Agency and DE-FGO2-00ER62773 from the Department of Energy. The authors wish to thank Dr. Tabbi L. Miller for critically reading and editing the manuscript, and Dr. Mildred A. Donlon for her assistance and encouragement. The authors also thank G. Bolus, M. L. Ferguson and T. Horn for their valuable technical assistance in sequencing of the clones. Supporting sequence data was obtained from The Institute for Genomic Research <a href="http://www.tigr.org">http://www.tigr.org</a>. The sequencing of the B. anthracis genome was accomplished with support from the Office of Naval Research (ONR), the Department of Energy (DOE) and the National Institute of Allergy and Infectious Diseases (NIAID). We also thank: W. Beyer, R. Böhm, T. N. Brahmbhatt, J. Burans, A. Cataldi, J. Ezzel, Z. Liu, M. Mock, C. L. Turnbough, J. Vaissaire and R. J. Zabransky for kindly providing Bacillus strains.

#### References

- Ahmed IH, Manning G, Wassenaar TM, Cawthraw S, Newell DG: Identification of genetic differences between two Campylobacter jejuni strains with different colonization potentials. Microbiology 2002, 148:1203-12.
- Akopyants NS, Fradkov A, Diatchenko L et al.: PCR-based subtractive hybridization and differences in gene content among strains of Helicobacter pylori. Proc Natl Acad Sci USA 1998, 95-13108-13
- 3. Ash C, Farrow JA, Dorsch M, Stackebrandt E, Collins MD: Comparative analysis of *Bacillus anthracis*, *Bacillus cereus* and related species on the basis of reverse transcriptase sequencing of 16S rRNA. Int J Syst Bacteriol 1991, 41:343-6.
- Bogush ML, Velikodvorskaya TV, Lebedev YB et al.: Identification and localization of differences between Escherichia coli and Salmonella typhimurium genomes by suppressive subtractive hybridization. Mol Gen Genet 1999, 262:721-9.
- Bonacorsi SP, Clermont O, Tinsley C et al.: Identification of regions of the Escherichia coli chromosome specific for neonatal meningitis-associated strains. Infect Immun 2000, 68:2096-101.
- Brumlik MJ, Szymajda U, Zakowska D et al.: Use of long-range repetitive element polymorphism-PCR to differentiate Bacillus anthracis strains. Appl Environ Microbiol 2001, 67:3021-8.
- DeShazer D, Waag DM, Fritz DL, Woods DE: Identification of a Burkholderia mallei polysaccharide gene cluster by subtractive hybridization and demonstration that the encoded capsule is an essential virulence determinant. Microb Pathog 2001, 30:253-69.
- Dixon TC, Meselson M, Guillemin J, Hanna PC: Anthrax. N Engl J Med 1999, 341:815-26.
- Drobniewski FA: Bacillus cereus and related species. Clin Microbiol Rev 1993, 6:324-38.
- Ezzell JW Jr, Abshire TG, Little SF, Lidgerding BC, Brown C: Identification of Bacillus anthracis by using monoclonal antibody to cell wall galactose-N-acetylglucosamine polysaccharide. J Clin Microbiol 1990, 28:223-31.
- Fellows P: A survey of worldwide strains of Bacillus anthracis. Proceeding of the International Workshop on Anthrax. Salisbury Med Bull 1996, 87:31-33. Ref Type: Conference Proceeding

- Green BD, Battisti L, Koehler TM, Thorne CB, Ivins BE: Demonstration of a capsule plasmid in Bacillus anthracis. Infect Immun 1985. 49:291-7.
- Hacker J, Blum-Oehler G, Muhldorfer I, Tschape H: Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. Mol Microbiol 1997, 23:1089-97.
- Harrell LJ, Andersen GL, Wilson KH: Genetic variability of Bacillus anthracis and related species. J Clin Microbiol 1995, 33:1847-50.
- Helgason E, Okstad OA, Caugant DA et al.: Bacillus anthracis, Bacillus cereus, and Bacillus thuringiensis-one species on the basis of genetic evidence. Appl Environ Microbiol 2000, 66:2627-30.
- Hill KK, Ticknor LO, Richardson A, Pardington PE, Jackson PJ: Fluorescent amplified fragment length polymorphism (AFLP) analysis of Bacillus anthracis, Bacillus cereus, and Bacillus thuringiensis isolates. In 4th International Conference on Anthrax.; Annapolis, Maryland 2001.
- Ivanova N, Sorokin A, Anderson I et al.: Genome sequence of Bacillus cereus and comparative analysis with Bacillus anthracis. Nature 2003, 423:87-91.
- Janke B, Dobrindt U, Hacker J, Blum-Oehler G: A subtractive hybridization analysis of genomic differences between the uropathogenic E. coli strain 536 and the E. coli K-12 strain MG1655. FEMS Microbiol Lett 2001, 199:61-6.
- Kaneko T, Nozaki R, Aizawa K: Deoxyribonucleic acid relatedness between Bacillus anthracis, Bacillus cereus and Bacillus thuringiensis. Microbiol Immunol 1978, 22:639-41.
- Lawrence JG, Ochman H: Amelioration of bacterial genomes: rates of change and exchange. J Mol Evol 1997, 44:383-97.
- Mikesell P, Ivins BE, Ristroph JD, Dreier TM: Evidence for plasmidmediated toxin production in Bacillus anthracis. Infect Immun 1983, 39:371-6.
- 22. Mock M, Fouet A: Anthrax. Annu Rev Microbiol 2001, 55:647-71.
- Molnar J, Pragai B: Attempts to detect the presence of teichoic acid in Bacillus anthracis. Acta Microbiol Acad Sci Hung 1971, 18:105-8.
- 24. National Center for Biotechnology Information Blastn Database [http://www.ncbi.nlm.nih.gov/BLAST/]
- 25. National Center for Biotechnology Information Blastp Database [http://www.ncbi.nlm.nih.gov/blast/Blast.cgi]
- National Center for Biotechnology Information Open Reading Frame Finder Analysis Tool [http://www.ncbi.nlm.nih.gov/gorf/gorf.html]
- Okinaka R, Cloud K, Hampton O et al.: Sequence, assembly and analysis of pX01 and pX02. J Appl Microbiol 1999, 87:261-2.
- Okinaka R, Longmire J, Svensson R, Doggett N, Bruce D, Mundt M, Jackson P: Whole-genome sequencing Bacillus thuringiensis Strain 97-27, a close relative of Bacillus anthracis. In 4th International Conference on Anthrax: Annapolis, Maryland 2001.
- Okinaka RT, Cloud K, Hampton O et al.: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harboring the anthrax toxin genes. J Bacteriol 1999, 181:6509-15.
- 30. Patra G, Vaissaire J, Weber-Levy M, Le Doujet C, Mock M: Molecular characterization of Bacillus strains involved in outbreaks of anthrax in France in 1997. J Clin Microbiol 1998, 36:3412-4.
- 31. Qi Y, Patra G, Liang X et al.: Utilization of the rpoB gene as a specific chromosomal marker for real-time PCR detection of Bacillus anthracis. Appl Environ Microbiol 2001, 67:3720-7.
- Radnedge L, Agron PG, Hill KK et al.: Genome differences that distinguish Bacillus anthracis from Bacillus cereus and Bacillus thuringiensis. Appl Environ Microbiol 2003, 69:2755-64.
- Ramisse V, Patra G, Vaissaire J, Mock M: The Ba813 chromosomal DNA sequence effectively traces the whole Bacillus anthracis community. J Appl Microbiol 1999, 87:224-8.
- Read TD, Salzberg SL, Pop M et al.: Comparative genome sequencing for discovery of novel polymorphisms in Bacillus anthracis. Science 2002, 296:2028-33.
- Read TD, Peterson SN, Tourasse N et al.: The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria. Nature 2003, 423:81-6.
- Reckseidler SL, DeShazer D, Sokol PA, Woods DE: Detection of bacterial virulence genes by subtractive hybridization: identification of capsular polysaccharide of Burkholderia pseudomallei as a major virulence determinant. Infect Immun 2001, 69:34-44

- Roloff H, Glockner P, Mistele K, Bohm R: The taxonomic relationship between B. anthracis and the B. cereus-group, investigated by DNA-DNA hybridization and DNA amplification fingerprinting (DAF). 2001. Ref Type: Conference Proceeding
- Sambrook J, Fritch EF, Maniatis T: Molecular Cloning: a Laboratory Manual. Cold Spring Harbor, NY: Laboratory Press; 1989.
- Schnepf E, Crickmore N, Van Rie, J et al.: Bacillus thuringiensis and its pesticidal crystal proteins. Microbiol Mol Biol Rev 1998, 62:775-806
- Schraft H, Griffiths MW: Specific oligonucleotide primers for detection of lecithinase-positive Bacillus spp. by PCR. Appl Environ Microbiol 1995, 61:98-102.
- 41. Turnbull PC, Hutson RA, Ward MJ et al.: Bacillus anthracis but not always anthrax. J Appl Bacteriol 1992, 72:21-8.
- Uchida I, Sekizaki T, Hashimoto K, Terakado N: Association of the encapsulation of Bacillus anthracis with a 60 megadalton plasmid. J Gen Microbiol 1985, 131(Pt 2):363-7.
- Zhang L, Foxman B, Manning SD, Tallman P, Marrs CF: Molecular epidemiologic approaches to urinary tract infection gene discovery in uropathogenic Escherichia coli. Infect Immun 2000, 68:2009-15.

Publish with **Bio Med 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
- $\bullet$  yours you keep the copyright

Submit your manuscript here: http://www.biomedcentral.com/info/publishing\_adv.asp

