

SRK. Since SLG resembles the extracellular domain of SRK, and since the efficiency of SI increases as the homology between the two increases, it is reasonable to suppose that, while both molecules must interact with SCR emitted by the pollen [perhaps binding it as SLG and SLR1 bind PCPs (Ref. 7)], some feature of this interaction must regulate the kinase activity of SRK.

A role for SLG in the activation of SRK?

One possibility is that SLG acts as a shuttle vector, binding SCR at the outer face of the papillar wall, and transporting it to the extracellular domain of SRK below. Differences in affinity for SCR between SLG and SRK could then encourage SCR to abandon SLG for the receptor kinase. The results of Takasaki *et al.*¹² might therefore be explained in terms of differing affinities between SLG and SCR. Thus, where homology between SLG and SRK is very high, such as occurs between alleles within the same class (either class I or II), and particularly between the SRK and SLG of individual alleles, the affinity between SCR and SLG would be strong (reflecting the presumed high affinity of SCR for the extracellular domain of the SRK of that particular allele). As a consequence, large numbers of SCR molecules would

bind to SLG and would subsequently be presented to the SRK. Where SCR/SLG homology was lower, for example between class I and class II alleles, this affinity would be lessened, resulting in the shuttling of fewer SCR molecules to the SRK at the stigmatic plasma membrane, and the consequent activation of a smaller proportion of the receptor kinases (Fig. 2). Alternatively, the SCR/SLG complex might itself associate with the extracellular domain of SRK (itself also binding one or more SRC molecules) in a 'pseudo-dimerization' event resulting in the activation of the kinase. SRK can certainly dimerize¹³ and extra 'copies' of the extracellular domain in the form of SLG could markedly increase the sensitivity of the system to the presence of the ligand (Fig. 3), and thereby increase the efficiency of the SI response.

Concluding remarks

The key to understanding SLG function clearly lies in the nature of its interaction with SCR, SRK and perhaps even the PCPs (Ref. 6). Hopefully, imaginative approaches to the study of these interactions *in vitro* (as in Ref. 13) will begin to throw light on the role of this suddenly redundant stigmatic factor.

References

- Schopfer, C.R. *et al.* (1999) The male determinant of self-incompatibility in *Brassica*. *Science* 286, 1697–1700
- Stone, S.L. *et al.* (1999) A breakdown of self-incompatibility in *ARC1* antisense plants. *Science* 286, 1729–1731
- Nasrallah, J.B. *et al.* (1985) A cDNA clone encoding an S-locus specific glycoprotein from *Brassica oleracea*. *Nature* 318, 263–267
- Stein, J.C. *et al.* (1991) Molecular cloning of a putative receptor protein kinase encoded the self-incompatibility locus of *Brassica oleracea*. *Proc. Natl. Acad. Sci. U. S. A.* 88, 8816–8820
- Kusaba, M. *et al.* (1997) Striking sequence similarity in inter- and intra-specific comparisons of class I *SLG* alleles from *Brassica oleracea* and *Brassica campestris*: Implications for the evolution and recognition mechanism. *Proc. Natl. Acad. Sci. U. S. A.* 94, 7673–7678
- Doughty, J. *et al.* (1998) *PCPA1*, a defensin-like *Brassica* pollen coat protein that binds the S-locus glycoprotein, is the product of gametophytic gene expression. *Plant Cell* 10, 1333–1347
- Hiscock, S. *et al.* (1995) A 7-kDa pollen coating-borne peptide from *Brassica napus* interacts with S-locus glycoprotein and S-locus-related glycoprotein. *Planta* 196, 367–374
- Stephenson, A.G. *et al.* (1997) The male determinant of self-incompatibility in *Brassica oleracea* is located in the pollen coating. *Plant J.* 12, 1351–1359
- Takayama, S. *et al.* (2000) The pollen determinant of self-incompatibility in *Brassica campestris*. *Proc. Natl. Acad. Sci. U. S. A.* 97, 1920–1925
- Dickinson, H.G. *et al.* (2000) Pollen coatings – chimaeric genetics and new functions. *Sexual Plant Reprod.* 12, 302–310
- Dickinson, H.G. and Lewis, D. (1973) The formation of the tryphine coating the pollen grains of *Raphanus* and its properties relating to the self-incompatibility system. *Proc. R. Soc. London Ser. B* 184, 149–165
- Takasaki, T. *et al.* (2000) The S receptor kinase determines self-incompatibility in *Brassica stigma*. *Nature* 403, 913–916
- Cock, J.M. *et al.* (2000) Investigating the molecular mechanism of the self incompatibility response in *Brassica*. *Ann. Botany* 85 (Suppl. A), 147–153

Graphical genome comparison rearrangements and replication origin of *Helicobacter pylori*

The comparative genomics of bacterial strains can reveal interesting facts about their genome organization, evolution and pathogenicity. By comparing the sequences of *Helicobacter pylori* strains 26695 (Ref. 1) and J99, Alm *et al.*² have inferred a set of rearrangements that can account for the divergent orders of open reading frames (ORFs) in these strains. These authors mentioned that one

of the similarities of the two genomes was the absence of an identifiable replication origin (*oriC*), because no regions of sequence organization typical for bacterial *oriC* sites³ could be found in either strain.

Here I present a graphical analysis of the genome rearrangements, which provides clues to the location of the putative *oriC* and can serve as a useful tool for

examining genome reorganization and evolutionary history.

This approach is based on the observations that bacterial chromosomes can be divided into two halves of different base composition⁴⁻⁸. In one DNA strand, an excess of guanines vs cytosines ($G > C$) is detected in one half, and $G < C$ in the other. Cumulative skew diagrams⁷, or plots of numerically integrated GC skew $[(G - C)/(G + C)]$ in small adjacent windows over the chromosome length, are a convenient method for identifying the points where base composition changes. Singular global minima on these diagrams coincide with the location of *oriC* in all genomes with known origins; thus the leading strand in replication is G-richer.

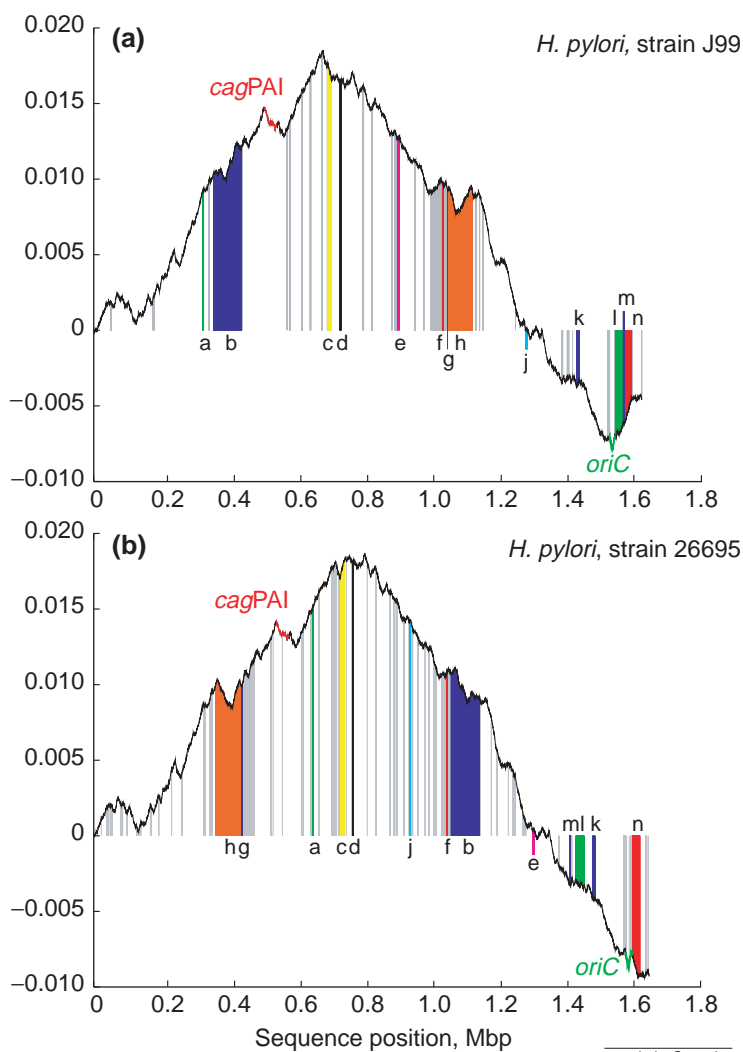
This has suggested the possibility of predicting the location of *oriC* for the less-studied genomes. Some of the predictions have already been proved experimentally, for example for the bacteria *Borrelia burgdorferi*⁹ and *Mycobacterium tuberculosis*¹⁰. This approach has also shed light on archaeal replication. Cumulative diagrams clearly indicated a singular (as in bacteria) *oriC* in *Methanobacterium thermoautotrophicum*⁷. That finding, further supported by other *in silico* analyses^{11,12}, was somewhat unexpected, given that archaea and eukaryotes share many more similarities in their replication machinery than either do with bacteria¹³. Additional evidence for the single-origin replication model has been recently reported for another archaeon, *Pyrococcus abyssi*, whose *oriC* position suggested by cumulative diagrams is in agreement with experimental data¹⁴.

In *H. pylori*, such prediction may seem problematic since the GC skew diagrams for two strains are not identical. The global minimum is clearly seen in strain J99 at position 1.557 Mbp (Fig. 1a), but in strain 26695 there are several adjacent minima (Fig. 1b: to the right of 1.63 Mbp). For a circular chromosome, the base pair positions of the right- and left-hand ends of the plot are the same, and the minimum in strain 26695 has been assigned to 1 bp^{5,7}. Another prediction, using genome-wide skew of the oligomer pattern RRTAGGGG, placed *oriC* in strain 26695 between positions 1.566 Mbp and 1.64 Mbp, with the probability $P = 4.9 \times 10^{-29}$ of the observed skew occurring by chance¹¹. With the 18 most skewed oligomers (a handful of which had the pattern RRTAGGGG), the centre of the predicted region was at 1.59 Mbp, but the result was not as significant¹¹.

What is the cause of these differences in the skew plots? The 'jagged' diagram of *H. pylori*, in contrast to many other genomes, has been hypothesized⁷ to result from frequent rearrangements and horizontally transferred DNA. For example, an inversion would swap the corresponding fragments of the leading and lagging strand and create a local minimum and maximum at the inversion boundaries. Using that fact and highlighting rearrangements on the plot (Fig. 1), one can transform cumulative diagrams into a convenient tool for comparative genomics.

This display illustrates interstrain differences in a novel way (like a genome-wide dot-plot combined with a composition analysis–rearrangement predictor). Inversions are clearly seen to result in mirror symmetry transformations of the respective diagram fragments (e.g. b, h, c and n). Several local extrema are in the 'plasticity zone'², containing the majority of strain-specific genes (areas shaded gray near b, f, g and h). The *cag* pathogenicity island¹⁵, possibly acquired in a single ancestral insertion event¹⁶, marks a

FIGURE 1. Graphical genome comparison



Cumulative GC skew diagrams of (a) strain J99 and (b) strain 26695 of *Helicobacter pylori* (black curve, green tip around the global minimum corresponds to the putative *oriC*-containing fragment, red curve fragment shows the *cag* pathogenicity island¹⁵, *cagPAI*). The y axis gives the cumulative GC skew⁷, multiplied by window size (120 bp) and divided by genome length. These are integral plots, where a minimum corresponds to the point where mean GC skew changes from negative to positive (and vice versa for a maximum). Genome rearrangements (designated with letters a–h, j–n) are shown as coloured areas under the curve. All fragments represent inversions (and, in most cases, translocations), except for the rearrangements designated 'a' (only translocation), 'j' and 'e' (reciprocal exchange). Areas shaded in light grey show strain-specific genes, retaining short ORFs annotated in strain 26695 and not in strain J99. A small number of strain-specific genes are not shown; these reside inside larger rearrangements. Note the mirror symmetry of the curve fragments corresponding to inversions designated by the same letters in the two strains.

prominent local maximum. Further studies of this organism are likely to reveal rearrangements corresponding to other local peaks of the diagrams.

What can be inferred from this graphical genome comparison? A notable feature of the plots is that three sequence fragments (l, m and n) form a contiguous block in strain J99. All these fragments are inverted and two of them are shifted to different positions to the left in strain 26695. This suggests that the inversions and translocations took place in the strain 26695 sequence, disrupting the original arrangement of these fragments. As a result,

Andrei Grigoriev
andrei.grigoriev@
gpc-biotech.com

GPC Biotech,
Fraunhoferstrasse 20,
Martinsried 82152,
Germany

the total skew (the right-hand end of the plot) in that strain is further away from zero than in strain J99.

Moreover, GC skew across the contiguous block l-m-n in strain J99 is practically uniformly positive. Assuming that was the original order of these fragments in the common ancestor, the cumulative GC skew diagram of *H. pylori* appears to be very similar to diagrams of other bacteria. Despite the above rearrangements near the putative origin at the global minimum, the immediate surroundings of *oriC* (green tip, Fig. 1) are conserved between the two strains. The minimum of the green tip resides over the putative ORF JHP1416 in strain J99 (HP1527 in strain 26695), next to the *dnaA* gene, which encodes the chromosome replication initiation protein. This is an additional indication that the minimum is close to *oriC* of *H. pylori*, because *dnaA* is found near the global minimum of cumulative GC skew (and the assigned origin) in many other bacterial genomes, including *B. burgdorferi*, *Bacillus subtilis*, *Mycoplasma genitalium*, *Mycoplasma pneumoniae*, *M. tuberculosis* and *Treponema pallidum*. Taken together, these facts point to a putative *oriC* location, which can be tested experimentally. The presence of *dnaA* in this region has also been noted by Salzberg *et al.*¹¹, although the midpoint of their predicted *oriC*-containing interval (with 18 oligomers) is some 20 kbp to the left of this gene.

Interestingly, the *cag* island and the plasticity zone are practically devoid of the ubiquitous and most skewed

octamer RRTAGGGG (not shown). This implies that the plasticity zone is not a result of differential retention of ancestral DNA in these strains but is a product of horizontal transfer, in agreement with Alm *et al.*², who suggested this region might represent pathogenicity islands.

Restoring the order l-m-n increases the number of genes transcribed in the direction of replication (given the *oriC* location next to *dnaA*). Such 'co-directional' organization is seen in many bacterial genomes and has been interpreted in terms of minimizing 'head-on' collisions between DNA and RNA polymerases¹⁷. However, despite these differences in gene orientation, 26695 is a viable strain and it remains to be seen if the rearrangements so close to *oriC* have decreased its fitness.

In conclusion, the described method of compact representation of large-scale genome differences can be applied for the graphical comparison of other genomes and strains to help in linking their organizational features (such as *oriC*), rearrangements and base composition. Perl scripts for calculating cumulative diagrams are available on request from the author.

Acknowledgements

I thank the anonymous referees, and also H. Loferer and I. Ivanov for their comments on earlier versions of the manuscript.

References

- 1 Tomb, J.-F. *et al.* (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388, 1020–1022
- 2 Alm, R.A. *et al.* (1999) Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* 397, 176–180
- 3 Marczyński, G.T. and Shapiro, L. (1993) Bacterial chromosome origins of replication. *Curr. Opin. Genet. Dev.* 3, 775–782
- 4 Lobry, J.R. (1996) Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol. Biol. Evol.* 13, 660–665
- 5 Freeman, J.M. *et al.* (1998) Patterns of genome organization in bacteria. *Science* 279, 1827a
- 6 Mrazek, J. and Karlin, S. (1998) Strand compositional asymmetry in bacterial and large viral genomes. *Proc. Natl. Acad. Sci. U. S. A.* 95, 3720–3725
- 7 Grigoriev, A. (1998) Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res.* 26, 2286–2290
- 8 McLean, M.J. *et al.* (1998) Base composition skews, replication orientation, and gene orientation in 12 prokaryote genomes. *J. Mol. Evol.* 47, 691–696
- 9 Picardeau, M. *et al.* (1999) Physical mapping of an origin of bidirectional replication at the centre of the *Borrelia burgdorferi* linear chromosome. *Mol. Microbiol.* 32, 437–445
- 10 Qin, M.H. *et al.* (1999) Characterization of the functional replication origin of *Mycobacterium tuberculosis*. *Gene* 233, 121–130
- 11 Salzberg, S.L. *et al.* (1998) Skewed oligomers and origins of replication. *Gene* 217, 57–67
- 12 Lopez, P. *et al.* (1999) Identification of putative chromosomal origins of replication in Archaea. *Mol. Microbiol.* 32, 883–886
- 13 Edgell, D.R. and Doolittle, W.F. (1997) Archaea and the origin(s) of DNA replication proteins. *Cell* 89, 995–998
- 14 Myllykallio, H. *et al.* (2000) Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science* 288, 2212–2215
- 15 Censini, S. *et al.* (1996) *cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. *Proc. Natl. Acad. Sci. U. S. A.* 93, 14648–14653
- 16 Akopyants, N.S. *et al.* (1998) Analyses of the *cag* pathogenicity island of *Helicobacter pylori*. *Mol. Microbiol.* 28, 37–53
- 17 Brewer, B.J. (1990) Replication and the transcriptional organization of the *Escherichia coli* chromosome. In *The Bacterial Chromosome* (Drlica, K. and Riley, M., eds), pp. 61–83, ASM



Influential passengers come of age

1st International *Wolbachia* Conference, Orthodox Academy, Kolymbari, Crete, Greece, 7–12 June 2000

To a good approximation, in 1990 nobody worked on *Wolbachia*. By contrast, a mere ten years later, 110 biologists from 21 countries assembled for the first meeting dedicated to these master manipulators of arthropod reproduction^{1,2}. *Wolbachia*, which infect a quarter of all insects^{1,2}, are not-so-distant relatives of

Rickettsia, and belong to the bacterial division from which mitochondria are descended. Perhaps the ability to 'cross-talk' with mitochondrial pathways underlies the success of these obligate intracellular parasites in altering host biology to enhance their own transmission^{1,2} (Fig. 1).