Comparative Genomic Analysis of Fruiting Body Formation in Myxococcales

Stuart Huntley,1 Nils Hamann,1 Sigrun Wegener-Feldbrügge,1 Anke Treuner-Lange,1 Michael Kube,2 Richard Reinhardt,2 Sven Klages,2 Rolf Müller,3 Catherine M. Ronning,4 William C. Nierman,4 and Lotte Søgaard-Andersen*.1

1Department of Ecophysiology, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany
2Max Planck Institute for Molecular Genetics, Berlin, Germany
3Department Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland, Saarland University, Saarbrücken, Germany
4Infectious Diseases Program, J. Craig Venter Institute, Rockville, MD

*Corresponding author: E-mail: sogaard@mpi-marburg.mpg.de.
Associate Editor: James McInerney

Abstract

Genetic programs underlying multicellular morphogenesis and cellular differentiation are most often associated with eukaryotic organisms, but examples also exist in bacteria such as the formation of multicellular, spore-filled fruiting bodies in the order Myxococcales. Most members of the Myxococcales undergo a multicellular developmental program culminating in the formation of spore-filled fruiting bodies in response to starvation. To gain insight into the evolutionary history of fruiting body formation in Myxococcales, we performed a comparative analysis of the genomes and transcriptomes of five Myxococcales species, four of these undergo fruiting body formation (Myxococcus xanthus, Stigmatella aurantiaca, Sorangium cellulosum, and Haliangium ochraceum) and one does not (Anaeromyxobacter dehalogenans). Our analyses show that a set of 95 known M. xanthus development-specific genes—although suffering from a sampling bias—-are overrepresented and occur more frequently than an average M. xanthus gene in S. aurantiaca, whereas they occur at the same frequency as an average M. xanthus gene in S. cellulosum and in H. ochraceum and are underrepresented in A. dehalogenans. Moreover, genes for entire signal transduction pathways important for fruiting body formation in M. xanthus are conserved in S. aurantiaca, whereas only a minority of these genes are conserved in A. dehalogenans, S. cellulosum, and H. ochraceum. Likewise, global gene expression profiling of developmentally regulated genes showed that genes that upregulated during development in M. xanthus are overrepresented in S. aurantiaca and slightly underrepresented in A. dehalogenans, S. cellulosum, and H. ochraceum. These comparative analyses strongly indicate that the genetic programs for fruiting body formation in M. xanthus and S. aurantiaca are highly similar and significantly different from the genetic program directing fruiting body formation in S. cellulosum and H. ochraceum. Thus, our analyses reveal an unexpected level of plasticity in the genetic programs for fruiting body formation in the Myxococcales and strongly suggest that the genetic program underlying fruiting body formation in different Myxococcales is not conserved. The evolutionary implications of this finding are discussed.

Key words: Sporulation, fruiting body formation, bacterial development, bacterial differentiation, Myxococcus xanthus, comparative genomics.

Introduction

To optimize their chances of survival in an ever-changing environment, bacteria have several strategies to sense and respond to these changes (Perez and Groisman 2009). One strategy involves adaptive changes in gene expression without evident changes in cell morphology. In an alternative strategy, changes in gene expression result in cellular differentiation with the formation of cells with altered morphologies and properties. One of the best-studied examples of cell differentiation in bacteria in response to changes in the environment is the formation of a heat and desiccation-resistant cell morphology, referred to as spores and cysts (from here on referred to collectively as spores). Spore formation is generally induced by nutrient depletion and spores are generally less metabolically active than vegetative cells, are more resistant to various types of chemical and physical stresses than vegetative cells, and can survive extended periods of starvation (Shimkets and Brun 2000).

Bacterial species from widely separated taxonomic clades have evolved the ability to form spores in response to starvation (Shimkets and Brun 2000), suggesting that the ability to form spores has emerged independently several times in bacteria. Spore-forming bacteria include members of the low G + C Gram-positive bacteria (e.g., Bacillus subtilis and Clostridium tetani), high G + C Gram-positive bacteria (e.g., Streptomyces coelicolor), filamentous cyanobacteria (e.g., Nostoc punctiforme) as well as members of the Gram-negative Alphaproteobacteria (e.g., Methylosinus trichosporium and Rhodospirillum capsulatus), Gammaproteobacteria (e.g., Azotobacter vinelandii) and Deltaproteobacteria (e.g,
Myxococcus xanthus) (Shimkets and Brun 2000). Spore formation has been most thoroughly studied in B. subtilis, S. coelicolor, and M. xanthus. Supporting the idea that sporulation has evolved several times independently in bacteria, the mechanisms underlying spore formation in these three phylogenetically widely separated species are very different (Shimkets and Brun 2000). In B. subtilis (Kroos 2007), spore formation is initiated by an asymmetric cell division in a rod-shaped cell resulting in the formation of a large mother cell and a small prespore. The prespore is subsequently engulfed by the mother cell before the mother cell lyses resulting in the release of the spherical mature spore. In S. coelicolor (Flärhd and Buttner 2009), chains of spherical spores are formed by synchronous cell divisions in aerial hyphae. In M. xanthus, spore formation is coupled to the formation of multicellular fruiting bodies, and individual rod-shaped cells that have accumulated inside a fruiting body directly differentiate into spherical spores without a prior cell division (Kroos 2007). The availability of complete genome sequences of related bacteria provide the opportunity to explore what extent the genetic programs governing spore formation in a taxonomically well-defined group of bacteria are conserved. Here, we have focused on an evolutionary analysis of the genetic program underlying spore formation in the Myxococcales.

Members of the Deltaproteobacteria order Myxococcales are typically found in topsoil where they form cooperatively feeding colonies and grow as saprophytes by decomposing degradable polymers or as predators by preying on other microorganisms (Shimkets et al. 2006). They share three traits that distinguish them from other Deltaproteobacteria. First, most members are strictly aerobic, whereas most Deltaproteobacteria are anaerobic (Shimkets et al. 2006). Second, in response to starvation, most Myxococcales initiate a developmental program that culminates in the formation of multicellular fruiting bodies inside which the rod-shaped cells differentiate into spores (Shimkets et al. 2006). Third, most members of the Myxococcales have large genome sizes around 10 Mb (Goldman et al. 2006; Schneiker et al. 2007; Thomas et al. 2008; Wu et al. 2009; Natalia et al. 2010), whereas other Deltaproteobacteria have smaller genome sizes ranging from 2.9 Mb to 6.5 Mb (supplementary table S1, Supplementary Material online).

Myxococcus xanthus has emerged as the model organism to understand the molecular mechanisms underlying fruiting body formation (Kroos 2007). Fruiting body formation is initiated by the RelA-dependent stringent response and proceeds in distinct morphological stages. The first signs of fruiting body formation are evident 4–6 h after starvation as cells aggregate to form small aggregation centers. As they accumulate more cells, the centers increase in size and eventually become mound shaped. By 24 h, the aggregation process is complete and each nascent fruiting body contains approximately $10^5$ cells. Inside the nascent fruiting bodies, the rod-shaped cells differentiate into spherical spores, resulting in mature fruiting bodies. Of the total population present at the onset of starvation, only 10% undergo sporulation and normally only those cells that have accumulated inside the fruiting bodies. Up to 30% of the cells remain outside the fruiting bodies. These cells remain rod shaped and differentiate to a cell type termed peripheral rods (O’Connor and Zusman 1991a, 1991b). Finally, the remaining cells undergo lysis (Wireman and Dworkin 1977; Rosenbluh et al. 1989). Recently, this cell lysis was suggested to reflect programmed cell death (Nariya and Inouye 2008). Fruiting body formation depends on intercellular signaling (Kroos 2007) and highly coordinated changes in gene expression (Kroos 2007), as well as on the coordinated motility behavior of cells (Leonardy et al. 2008). Thus, fruiting body formation includes many of the characteristics associated with developmental programs in metazoa such as cell–cell signaling, cells adopting different cell fates, programmed cell death, morphogenetic cell movements, and coupling of cell differentiation to multicellular morphogenesis. Because spore formation occurs inside the fruiting bodies, these structures are essentially spor-filled assemblies. In the presence of nutrients, spores germinate to give rise to rod-shaped, motile, metabolically active vegetative cells. With each fruiting body consisting of approximately $10^6$ spores, germination immediately gives rise to a spreading, cooperatively feeding colony. Thus, it has been argued that fruiting bodies are optimally designed to ensure that a new vegetative cycle is initiated by a community of cells rather than by a single cell (Dworkin 1972).

All Myxococcales tested—with the exception of one (see below)—initiate fruiting body formation in response to starvation, suggesting that the last common ancestor of the Myxococcales harbored a genetic program for fruiting body formation and that fruiting Myxococcales would share in common a genetic program underlying fruiting body formation. Here, we examined this hypothesis using comparative and functional genomics on four complete genomes of Myxococcales and one genome of the only known nonfruiting Myxococcales species. Our analyses suggest that the genetic program underlying fruiting body formation in different Myxococcales is not conserved, that is, different species of Myxococcales harbor different programs for fruiting body formation. The evolutionary implications of this finding are discussed.

Materials and Methods

Phylogenetic Analyses

Full-length 16S rRNA sequences of type strains representing diverse Myxococcales genera, as well as sequences from several non-Myxococcales Proteobacteria were collected from the Ribosomal Database Project (http://rdp.cme.msu.edu/). The sequences were aligned using ClustalW1.83 (Thompson et al. 1994), ends of the resulting alignment were trimmed so that all aligned sequences extended fully to both ends of the alignment. Gaps were removed and the sequences were realigned using ClustalW1.83 to ensure that the modifications from the first alignment did not alter the relationships between the remaining sequences. The
resulting alignment was used to generate a bootstrapped (1,000 iterations) neighbor joining tree file from which the image shown in Figure 1 was generated using the Treedyn software (Chevenet et al. 2006). The above analyses were repeated using the 16S rRNA sequences from all Deltaproteobacteria with a completed genome sequence to produce the tree figure shown in supplementary table S1 (Supplementary Material online).

Sequencing and Annotation of Stigmatella aurantiaca DW4/3-1

Stigmatella aurantiaca DW4/3-1 DNA was isolated with the Genomic DNA kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. The genomic sequences were determined by whole-genome shotgun sequencing using Sanger-based sequencing technology and pyrosequencing. The genome was covered by short-insert shotgun libraries and a fosmid library with 37-kb inserts (CopyControl Fosmid Library Production Kit; Epicentre, Madison). End sequencing of clone-derived inserts was performed using BigDye 3.1 chemistry and 3730XL capillary sequencers (ABI, Darmstadt, Germany) resulting in 90,783 reads (plasmid-derived reads 61,770 and 27,683 fosmid-derived reads) reaching a 7-fold sequencing coverage. In addition, pyrosequencing was performed using the GS20 sequencer (454 life science/Roche) resulting in 292,874 reads (241 bp average read length), which provided an additional 7-fold sequencing coverage. The assembly of Sanger and 454 derived reads was performed within two steps. GS20 data were initially assembled by Newbler (454 life science/Roche) and the resulting contigs were fragmented using PERL scripts resulting in overlapping sequence fragments, which were assigned as forward and reverse reads in fasta format and the corresponding fasta quality files. These “faked reads” were assembled together with the Sanger-derived processed reads using PhredPhrap (http://www.phrap.org). GS FLX data and Sanger-derived reads were assembled using the Celera assembler v4.3 (Myers et al. 2000). Assembled data were imported into Consed (Gordon 2003), edited, and verified. Finishing experiments included primer walking on bridging fosmid clones and polymerase chain reaction products to improve sequence quality and gap closure to a single closed circular sequence. Stigmatella aurantiaca protein-coding sequences were predicted by Glimmer3 (Delcher et al. 1999), manually curated in Artemis (Rutherford et al. 2000), and annotated with HTGA (Rabus et al. 2002). Structural rRNAs and tRNAs were determined using RNAmer (Lagesen et al. 2007) and tRNAscan-SE (Lowe and Eddy 1997). An unfinished version of the S. aurantiaca DW4/3-1 genome was previously reported (NZ_AAMD00000000). The genome sequence reported here represents the complete genome sequence of the DW4/3-1 strain. The genome sequence is deposited in GenBank with the accession number CP002271.

Functional Annotation and Ortholog Identification

Sequences and annotation for all publicly available bacterial species for which complete genome data available were obtained from the NCBI site ftp://ftp.ncbi.nih.gov/genomes/Bacteria/, files “All.faa.tar.gz and All.ptt.tar.gz (6 May 2009). Based on our information concerning known developmentally important genes, we included two M. xanthus loci (MXAN_3217, actD; and MXAN_5430, desS) to those included in the publicly available annotation. Functional annotations used in this study were generated by adaptation of available cluster of orthologous group (COG) functional assignments. First, locus descriptions from the four annotations were combined and sorted based on product description. All loci bearing the same description were assigned a single functional assignment code, chosen based on the code most prevalent among those provided in the COG assignments. Loci for which no COG assignment were available, as well as those loci for which the ambiguous COG categories R (general prediction only) and S (conserved) were originally assigned, were combined into a new category “X,” to indicate that no clear function has been identified for these proteins. Following the reciprocal Blast analyses (see below), final adjustments to the functional assignments were made manually such that orthologs shared the same functional category assignment.

To identify orthologs and homologs for each Myxococcales protein, complete, independent protein sequence databases were generated for each of the Myxococcales genome, as well as each of 973 complete, non-Myxococcales bacterial genomes available at the time of the analyses (from NCBI data files listed above). Each Myxococcales protein was used as a query sequence in a BlastP analysis against each database to identify homologs of the protein, using an expect score cutoff of $1 \times 10^{-5}$. The BlastP-identified matches were filtered, saving only those
matches that spanned at least one-third of the query sequence length. The expect scores used and the alignment length filter size were empirically determined to produce BlastP results with a maximum number of queries producing significant hits while limiting random noise in the results. For the purposes of this study, we considered all BlastP hits that passed these requirements to be homologs of the initial query protein. The protein corresponding to the highest scoring (bit score) match was then used as a query in a second BlastP analysis, this time against a complete protein database for the organism from which the original query protein was found. If the highest scoring match in this second BlastP analysis was to the original query protein, the two proteins were considered orthologs for the purposes of this study. To estimate the false-negative discovery rate of orthologs, we determined the frequency with which the original query protein was among the five best hits but not the best scoring hit in the second (reciprocal) BlastP analysis.

### Results

#### Synteny Plots

To visualize the conservation of genomic context of orthologs between species, a Perl script was written, which reads the genomic coordinates of each pair of orthologs from their respective genomes. Subsequently, the coordinates of all pairs of orthologs were plotted in a diagram having the coordinates of one species on the X axis and the coordinates of the other species on the Y axis giving rise to a dot plot-style image.

#### Comparisons of Five Completely Sequenced Myxococcales Genomes

At the time of this study, complete genome sequence information was available for *M. xanthus* strain DK1622 (Goldman et al. 2006), *Sorangium cellulosum* strain Soce56 (Schneiker et al. 2007), *Haliangium ochraceum* strain SMP-2 (Natalia et al. 2010), *Anaeromyxobacter dehalogenans* strains 2CP-C (Thomas et al. 2008) and 2CP-1, and *Anaeromyxobacter* species 109-5 and K. To decide which *Anaerobacter* genome to use in our analyses, we compared the four *Anaeromyxobacter* genomes. The genomes of *Anaeromyxobacter* 2CP-C, 2CP-1, Fw109-5, and K contain 4,346, 4,473, 4,466, and 4,457 protein-coding sequences, respectively (supplementary fig. S1, Supplementary Material online). Possible orthologs for each *Anaeromyxobacter* gene were identified using a reciprocal best BlastP hit method (see below). About 3,041 genes are conserved in all four genomes and represent 70%, 68%, 68%, and 68% of the genes in 2CP-C, 2CP-1, Fw109-5, and K, respectively (supplementary fig. S1, Supplementary Material online). About 7%, 6%, 26%, and 5% of all genes in the genomes of 2CP-C, 2CP-1, Fw109-5, and K, respectively, have no orthologs in any of the other *Anaeromyxobacter* genomes. Because the 2CP-C genome is the best-characterized of these four genomes (Thomas et al. 2008) and appear to represent well all four genomes, we used the *A. dehalogenans* 2CP-C genome for our further analyses. To increase the resolving power of comparative genome analyses, we determined the complete nucleotide sequence of the *S. aurantiaca* DW4-3-1 genome (table 1).

These five species cover all three Myxococcales suborders: *M. xanthus*, *S. aurantiaca*, and *A. dehalogenans* belong to the Cystobacterineae, *S. cellulosum* belong to the Sorangineae, and *H. ochraceum* belong to the Nanocystineae (Fig. 1 and table 1) (Shimkets et al. 2006). In all three suborders, species have been isolated, which initiate a developmental program resulting in fruiting body formation (Shimkets et al. 2006).

*Anaeromyxobacter dehalogenans* is significantly different from the four other species in several ways (table 1). Although the others are strict aerobes, *A. dehalogenans* is

---

**Table 1. Characteristics of the Five Myxococcales Species Analyzed in This Study.**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th><em>Myxococcus xanthus</em></th>
<th><em>Stigmatella aurantiaca</em></th>
<th><em>Anaeromyxobacter dehalogenans</em></th>
<th><em>Sorangium cellulosum</em></th>
<th><em>Haliangium ochraceum</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Suborder</td>
<td>Cystobacterineae</td>
<td>Cystobacterineae</td>
<td>Cystobacterineae</td>
<td>Sorangineae</td>
<td>Nannocystineae</td>
</tr>
<tr>
<td>Genome size (bp)*a</td>
<td>9,139,732</td>
<td>10,260,755</td>
<td>5,013,478</td>
<td>13,033,778</td>
<td>9,446,314</td>
</tr>
<tr>
<td>Protein-coding</td>
<td>7,333</td>
<td>8,352</td>
<td>4,346</td>
<td>9,384</td>
<td>6,719</td>
</tr>
<tr>
<td>sequences*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genome GC content*</td>
<td>68.9</td>
<td>67.5</td>
<td>74.9</td>
<td>71.4</td>
<td>69.0</td>
</tr>
<tr>
<td>Metabolism*</td>
<td>Amino acids; respiration</td>
<td>Amino acids; respiration</td>
<td>Short fatty acids; respiration</td>
<td>Cellulose; respiration</td>
<td>DNA, amino acids, starch; respiration</td>
</tr>
<tr>
<td>Oxygen tolerance/requirement*</td>
<td>Aerobic</td>
<td>Aerobic</td>
<td>Microaerophilic</td>
<td>Aerobic</td>
<td>Aerobic</td>
</tr>
<tr>
<td>Fruiting body</td>
<td>Hay-stack shaped</td>
<td>Sporangioles on stalk</td>
<td>Not fruiting</td>
<td>Assembly of sporangioles, not on stalk</td>
<td>Assembly of sporangioles, not on stalk</td>
</tr>
<tr>
<td>Spore morphology*</td>
<td>Spherical, optically refractile</td>
<td>Short bent rods, optically refractile</td>
<td>NA</td>
<td>Short rods, optically refractile</td>
<td>Spherical, optically refractile</td>
</tr>
</tbody>
</table>

**Notes:**

* References for genome information: *M. xanthus* (Goldman et al. 2007), *S. aurantiaca* (this study), *A. dehalogenans* (Thomas et al. 2008), *S. cellulosum* (Schneiker et al. 2007), and *H. ochraceum* (Natalia et al. 2010).

* References for physiology, fruiting body, and spore morphology: *M. xanthus*, *S. aurantiaca*, and *S. cellulosum* (Shimkets et al. 2006), *A. dehalogenans* (Sanford et al. 2002), and *H. ochraceum* (Fudou et al. 2002).
microaerophilic and its genome size is roughly half that of the other species’. The most significant difference, for the purposes of this study, is that A. dehalogenans has not been demonstrated to initiate fruiting body formation and sporulation in response to starvation (Sanford et al. 2002; Thomas et al. 2008). The morphology of the fruiting bodies formed by the four fruiters varies significantly (table 1): Myxococcus xanthus fruiting bodies are haystack shaped, S. aurantiaca produces elaborate fruiting bodies consisting of a stalk carrying several spore-filled sporangioles, and S. cellulosum as well as H. ochraceum generate assemblies of sporangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative cells in all four species are long and rod shaped, spororangioles, which are not carried by a stalk. Although the vegetative cells in all four species are long and rod shaped, spor morphology varies from spherical in vegetative ce

### Gene Conservation Within the Myxococcales

To be able to analyze the significance of the conservation of genes directly involved in fruiting body and sporulation in M. xanthus in the four other species or in other bacteria, we first established a baseline for these comparisons by systematically identifying possible orthologs for each Myxococcales gene (from here on predicted orthologs are referred to as orthologs). To this end, we used a reciprocal best BlastP hit method (Overbeek et al. 1999; Fang et al. 2010) in which each protein from each species was tested against the protein repertoires of each of the four other species as well as against the protein repertoires of 973 bacterial species for which the complete, annotated genome was available at the time of this study (Materials and Methods). From these analyses, Myxococcales proteins were sorted into two categories: Myxococcales specific and non-Myxococcales specific, based on whether or not an ortholog is found in a non-Myxococcales genome. The M. xanthus, S. aurantiaca, A. dehalogenans, S. cellulosum, and H. ochraceum genomes contain 1,240, 1,501, 707, 1,731, and 1,165 pairs of paralogous genes, respectively. Paralogs represent a potential problem in the identification of orthologs using a reciprocal best BlastP hit method and may result in an underestimation of orthologs (Overbeek et al. 1999; Fang et al. 2010). Using a less-stringent definition of orthologs (Materials and Methods), we found that the false-negative discovery rate for all pairwise genome comparisons varied between 2% and 4% (data not shown). We conclude that our method results in a valid estimation of orthologs.

In agreement with the phylogenetic analysis, the protein repertoires of M. xanthus and S. aurantiaca appear more

<table>
<thead>
<tr>
<th>Functional category</th>
<th>Myxococcus xanthus</th>
<th>Stigmatella aurantiaca</th>
<th>Anaeromyxobacter dehalogenans</th>
<th>Sorangium cellulosum</th>
<th>Haliangium ochraceum</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5(0.1)</td>
<td>4(0.1)</td>
<td>2(0.1)</td>
<td>3(0.1)</td>
<td>2(0.1)</td>
</tr>
<tr>
<td>B</td>
<td>5(0.1)</td>
<td>3(0.1)</td>
<td>2(0.1)</td>
<td>2(0.1)</td>
<td>3(0.1)</td>
</tr>
<tr>
<td>C</td>
<td>262(3.6)</td>
<td>257(3.1)</td>
<td>296(6.8)</td>
<td>364(3.9)</td>
<td>238(3.5)</td>
</tr>
<tr>
<td>D</td>
<td>41(0.6)</td>
<td>49(0.6)</td>
<td>38(0.9)</td>
<td>62(0.7)</td>
<td>45(0.7)</td>
</tr>
<tr>
<td>E</td>
<td>315(4.3)</td>
<td>324(3.9)</td>
<td>231(5.3)</td>
<td>365(3.9)</td>
<td>294(4.4)</td>
</tr>
<tr>
<td>F</td>
<td>94(1.3)</td>
<td>102(1.2)</td>
<td>70(1.6)</td>
<td>97(1.0)</td>
<td>72(1.1)</td>
</tr>
<tr>
<td>G</td>
<td>192(2.6)</td>
<td>306(3.7)</td>
<td>143(3.3)</td>
<td>343(3.7)</td>
<td>181(2.7)</td>
</tr>
<tr>
<td>H</td>
<td>156(2.1)</td>
<td>168(2.0)</td>
<td>120(2.8)</td>
<td>209(2.2)</td>
<td>139(2.1)</td>
</tr>
<tr>
<td>I</td>
<td>204(2.8)</td>
<td>188(2.3)</td>
<td>127(2.9)</td>
<td>198(2.1)</td>
<td>173(2.6)</td>
</tr>
<tr>
<td>J</td>
<td>225(3.1)</td>
<td>228(2.7)</td>
<td>198(4.6)</td>
<td>231(2.5)</td>
<td>186(2.8)</td>
</tr>
<tr>
<td>K</td>
<td>340(4.6)</td>
<td>352(4.2)</td>
<td>214(4.9)</td>
<td>475(5.1)</td>
<td>249(3.7)</td>
</tr>
<tr>
<td>L</td>
<td>290(4.0)</td>
<td>326(3.9)</td>
<td>169(3.9)</td>
<td>323(3.4)</td>
<td>237(3.5)</td>
</tr>
<tr>
<td>M</td>
<td>340(4.6)</td>
<td>329(3.9)</td>
<td>257(5.9)</td>
<td>335(3.6)</td>
<td>235(3.5)</td>
</tr>
<tr>
<td>N</td>
<td>570(7.8)</td>
<td>58(0.7)</td>
<td>59(1.4)</td>
<td>39(0.4)</td>
<td>26(0.4)</td>
</tr>
<tr>
<td>O</td>
<td>273(3.7)</td>
<td>272(3.3)</td>
<td>187(4.3)</td>
<td>313(3.3)</td>
<td>205(3.1)</td>
</tr>
<tr>
<td>P</td>
<td>268(3.7)</td>
<td>254(3.0)</td>
<td>178(4.1)</td>
<td>298(3.2)</td>
<td>180(2.6)</td>
</tr>
<tr>
<td>Q</td>
<td>241(3.3)</td>
<td>261(3.1)</td>
<td>93(2.1)</td>
<td>288(3.1)</td>
<td>192(2.8)</td>
</tr>
<tr>
<td>R</td>
<td>566(7.7)</td>
<td>674(8.1)</td>
<td>339(7.8)</td>
<td>831(8.9)</td>
<td>556(8.3)</td>
</tr>
<tr>
<td>S</td>
<td>94(1.3)</td>
<td>93(1.1)</td>
<td>63(1.4)</td>
<td>69(0.7)</td>
<td>59(0.8)</td>
</tr>
<tr>
<td>T</td>
<td>96(1.3)</td>
<td>89(1.1)</td>
<td>70(1.6)</td>
<td>98(1.0)</td>
<td>78(1.2)</td>
</tr>
<tr>
<td>Z</td>
<td>1(0.1)</td>
<td>1(0.1)</td>
<td>0(0)</td>
<td>0(0)</td>
<td>1(0.1)</td>
</tr>
<tr>
<td>X</td>
<td>3268(44.5)</td>
<td>4014(48.1)</td>
<td>1490(34.3)</td>
<td>4441(47.3)</td>
<td>3370(50.2)</td>
</tr>
</tbody>
</table>

**Table 2.** COG Categorization of Genes of the Five Myxococcales Species Analyzed in This Study.

*Numbers indicate total number of genes in each COG in that species, with numbers in parentheses indicating percent of all proteins in that species.

*The category hypothetical combines COG categories “general prediction” and “conserved hypothetical” and proteins for which no functional assignment was made.
similar to each other than any other species pairing, sharing well over half of their protein repertoires (Table 3). In contrast, less than half of either *M. xanthus* or *S. aurantiaca* proteins are conserved in *A. dehalogenans*. Given the small size of the *A. dehalogenans* genome, this was not surprising. In agreement with earlier comparisons of *M. xanthus* and *S. cellulosum* (Schneiker et al. 2007), only about 30% of *M. xanthus* and *S. aurantiaca* proteins are conserved in *S. cellulosum* and in *H. ochraceum* and vice versa. Finally, only approximately 30% of the protein repertoires of *S. cellulosum* and *H. ochraceum* are shared.

Pairwise comparisons of the position of orthologous genes on their respective genomes revealed a high degree of conserved synteny between *M. xanthus* and *S. aurantiaca* genomes and a lesser degree of synteny between these two genomes and that of *A. dehalogenans* (Fig. 2). The evident X-shaped pattern in the *M. xanthus* versus *S. aurantiaca* plot—which is also evident at a lower resolution in the plots comparing *M. xanthus* or *S. aurantiaca* with *A. dehalogenans*—likely reflects large chromosomal inversions that reverse gene order symmetrically around the origin of replication (Eisen et al. 2000). In all other pairwise

![Fig. 2. Synteny plot for orthologous genes in pairwise combinations of genomes of Myxococcales species. Each dot indicates the position of an orthologous gene pair on the indicated chromosomes. Chromosome orientations and sizes in megabases (Mb) are indicated on the appropriate axis for each species. The dnaA gene is defined as gene 1 in all five species.](https://academic.oup.com/mbe/article-abstract/28/2/1083/1220803)
comparisons, no discernable synteny was observed (Fig. 2). Thus, in agreement with the phylogenetic relatedness of the five Myxococcales, gene organization is significantly more conserved between M. xanthus, S. aurantiaca, and A. dehalogenans compared with S. cellulosum and H. ochraceum. Moreover, gene organization in S. cellulosum and H. ochraceum does not show a high degree of conservation.

To analyze the divergence in the gene repertoires of the five Myxococcales since their last common ancestor, we identified genes conserved in all five species and genes having lineage specificity within the order. About 1,056 genes are conserved between all five Myxococcales genomes, and most of these genes (1,046; 99%) have orthologs outside the Myxococcales (Segment IV in table 4 and supplementary table S2, Supplementary Material online). The 1,056 genes correspond to the Myxococcales core genome. The core genome was found to represent from a high of 24% of all genes in a species (A. dehalogenans) to a low of 11% (S. cellulosum). In agreement with the idea that the core genome is expected to represent functions required for these bacteria to thrive despite differences in their respective ecological niches, 85% of the core genes have inferred functions with the largest functional categories including housekeeping proteins involved in translation and ribosome structure (J); amino acid transport and metabolism (E); energy production and conversion (C); replication, recombination, and DNA repair (L); and cell wall, membrane, and envelope biogenesis (M). Interestingly, the largest COG category in the core genome is that of proteins of unknown function (X) (15%). About 53 genes (5%) for signal transduction (T) are also part of the core genome. The 10 genes function (X) (15%). About 53 genes (5%) for signal transduction (T) are also part of the core genome. The 10 genes

### Table 4. Gene Conservation in Five Myxococcales Genomes.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>1506</td>
<td>2290</td>
<td>254</td>
<td>133</td>
<td>221</td>
</tr>
<tr>
<td>III</td>
<td>159</td>
<td>39</td>
<td>30</td>
<td>14</td>
<td>27</td>
</tr>
<tr>
<td>II</td>
<td>159</td>
<td>39</td>
<td>30</td>
<td>14</td>
<td>27</td>
</tr>
<tr>
<td>I</td>
<td>159</td>
<td>39</td>
<td>30</td>
<td>14</td>
<td>27</td>
</tr>
</tbody>
</table>

**Table 4. Gene Conservation in Five Myxococcales Genomes.**

<table>
<thead>
<tr>
<th>Gene repertoire</th>
<th>Five Myxococcales Genomes</th>
<th>Development</th>
<th>Downregulated during M.x. Development</th>
<th>Upregulated during M.x. Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>M.x.</td>
<td>1506</td>
<td>2290</td>
<td>254</td>
<td>133</td>
</tr>
<tr>
<td>S.a.</td>
<td>159</td>
<td>39</td>
<td>30</td>
<td>14</td>
</tr>
<tr>
<td>A.d.</td>
<td>159</td>
<td>39</td>
<td>30</td>
<td>14</td>
</tr>
<tr>
<td>S.c.</td>
<td>159</td>
<td>39</td>
<td>30</td>
<td>14</td>
</tr>
<tr>
<td>H.o.</td>
<td>159</td>
<td>39</td>
<td>30</td>
<td>14</td>
</tr>
</tbody>
</table>

*Note.—M.x., Myxococcus xanthus; S.a., Stigmatella aurantiaca; A.d., Anaeromyxobacter dehalogenans; S.c., Sorangium cellulosum; and H.o., Halothiobacillus oshimae. The four segments labeled I, II, III, and IV highlight M. xanthus–specific conservation, M. xanthus/S. aurantiaca–specific conservation, fruiter-specific conservation, and conservation in all five Myxococcales species, respectively. Gray shading indicates the presence of a gene in the genome of that species. The left number indicates the number of genes shared between the species marked with gray shading for four different gene repertoires. The right number indicates the subset of those proteins for which an ortholog was found in at least one non-Myxococcales species.*
the genus *Prochlorococcus*, the core genome includes approximately 1,250 genes corresponding to 40–67% of the genes of the 12 fully sequenced *Prochlorococcus* genomes, and 81% of the genes have an inferred function (Kettler et al. 2007). Finally, in the order Lactobacillales, the core genome is made up of 567 genes corresponding to 20–33% of the genes in fully sequenced Lactobacillales genomes, and 84% of these genes have an inferred function (Makarova et al. 2006).

About 425 genes are conserved in *M. xanthus*, *S. aurantiaca*, *S. cellulosum*, and *H. ochraceum* (Segment III in table 4 and supplementary table S2, Supplementary Material online). These genes could, in principle, represent the signature genes encoding the functions needed to be a member of the fruiting body-forming Myxococcales. Among the 425 genes, those involved in ion transport and metabolism (P) (8%) as well as proteins of unknown function (X) (30%) represent the two largest groups. In addition, these putative fruiting body–specific genes are enriched for signal transduction (T), amino acid transport and metabolism (E), as well as energy production and conversion (C), functional categories. Interestingly, most of the 425 genes (403; 95%) have orthologs outside the Myxococcales. The 22 Myxococcales-specific genes are enriched for those involved in signal transduction (T) as well as proteins of unknown function (X).

Between the five Myxococcales species, 1,548 genes are shared specifically by *M. xanthus* and *S. aurantiaca* (Segment II in table 4 and supplementary table S2, Supplementary Material online), and 57% of this group (886) have orthologs outside the Myxococcales. Regardless of the presence or absence of non-Myxococcales orthologs, genes involved in signal transduction (T) as well as those of unknown function (X) are enriched in this group. Finally, 26%, 35%, 35%, 59%, and 34% of all genes from *M. xanthus*, *S. aurantiaca*, *A. dehalogenans*, *S. cellulosum*, and *H. ochraceum*, respectively, have no ortholog in any of the other Myxococcales. Approximately, 50% of these genes have an ortholog in a non-Myxococcales species. Like the genes shared between *M. xanthus* and *S. aurantiaca*, *M. xanthus*–specific genes with and without orthologs in non-Myxococcales species (Segment I in table 4 and supplementary table S2, Supplementary Material online) are enriched for those involved in signal transduction (T) as well as proteins of unknown function (X).

Several patterns emerge from these analyses. First, *M. xanthus* genes at the four different levels of conservation (Segments I–IV in table 4) are differentially enriched for genes specifying different functional categories. Those conserved in all five species or by the four fruiters are dominated by genes for general cellular processes, whereas those conserved in *M. xanthus* and *M. xanthus* plus *S. aurantiaca* are dominated by signal transduction proteins and proteins of unknown function. Second, there is a gradient of conservation with respect to orthologs outside the Myxococcales with most genes (99%) shared by all five species having orthologs in non-Myxococcales but only 43% of the genes specific to *M. xanthus* having orthologs in non-Myxococcales.

### Conservation of *M. xanthus* Development Genes

The previous analyses established a baseline from which to analyze the extent of conservation of *M. xanthus* genes directly involved in fruiting body and sporulation in the three other species. We identified 95 genes in *M. xanthus* that are directly involved in fruiting body formation and/or sporulation but do not appear to influence cells growing vegetatively (from here on this gene set is referred to as *M. xanthus* development genes) ( supplementary table S3, Supplementary Material online). It should be emphasized that 60 (64%) of these genes belong to the signal transduction (T) category ( supplementary table S3, Supplementary Material online), whereas only 7.7% of the total *M. xanthus* proteins belong to this COG. Thus, the *M. xanthus* development genes are subject to a sampling bias. Nevertheless, as a first approximation, we used this set of genes to analyze the possible conservation of the *M. xanthus* genetic program for fruiting body formation and sporulation in other Myxococcales.

About 78 (82%) of the 95 *M. xanthus* development genes are conserved in *S. aurantiaca*, 28 (29%) in *A. dehalogenans*, 34 (36%) in *S. cellulosum*, and 25 (26%) in *H. ochraceum* (table 4). The *M. xanthus* development genes shared with *S. aurantiaca* are overrepresented, compared with the overall frequency of *M. xanthus* genes having orthologs in *S. aurantiaca* (82% vs. 63%) ( tables 3 and 5 for summary). The frequency of *M. xanthus* development genes shared with *A. dehalogenans* (29% vs. 31%) or with *S. cellulosum* (36% vs. 36%) equals the frequency of all *M. xanthus* genes shared with either of these two organisms. *Myxococcus xanthus* development genes shared with *H. ochraceum* (31% vs. 26%) are slightly underrepresented. With the caveat that the *M. xanthus* development-specific genes are subject to a sampling bias, these comparisons suggest that *M. xanthus* development genes are overrepresented in *S. aurantiaca* and not more likely to be conserved than...
any other gene or slightly underrepresented in S. cellulosum and H. ochraceum, respectively.

Interestingly, only 22 of the 95 M. xanthus development-specific genes are specific to the Myxococcales with the remaining 77 having orthologs in non-Myxococcales species (table 4 and supplementary table S3, Supplementary Material online). About 15 development-specific genes are conserved in all five Myxococcales and they all have orthologs outside Myxococcales. Only three genes are shared by the four fruiters, two have orthologs outside the Myxococcales. Three (M. xanthus, S. aurantiaca, H. ochraceum), zero (M. xanthus, S. cellulosum, H. ochraceum), and nine genes (M. xanthus, S. aurantiaca, S. cellulosum) are specifically shared by the three different combinations of three fruiters. About 36 genes are conserved only in M. xanthus and S. aurantiaca, and of these, 24 are found outside the Myxococcales. Finally, 11 are specific to M. xanthus, and of these, seven are found in a non-Myxococcales.

By parsimony, the M. xanthus development-specific genes shared by all five Myxococcales likely existed in their common ancestor. These 15 genes are all conserved outside Myxococcales, suggesting that they could be part of an ancestral response to starvation rather than being specifically involved in fruiting body formation and/or sporulation. In agreement with this notion, the RelA protein is found in this group. Notably, RelA is also important for fruiting body formation in S. cellulosum (Knauber et al. 2008). RelA is required for induction of the stringent response in M. xanthus and in S. cellulosum and is a universally conserved protein in bacteria. The stringent response is a ubiquitiously conserved mechanism in bacteria that links nutrient availability to accumulation of the second messenger guanosine-5′-[(tri)di-3′-diphosphate (ppGpp) by means of the RelA and/or SpoT proteins (Potrykus and Cashel 2008). Similarly, genes shared by the four fruiters could be relics of an ancient mechanism for fruiting body formation and sporulation and which was lost in A. dehalogenans. In contrast, genes specific to M. xanthus and S. aurantiaca may be part of a development mechanism specific to those species’ lineage and might reflect an adaptation to their particular lifestyle or environment.

The above interpretation is based on numbers. We thought that a more biological approach to determine the relevance of the differential conservation of M. xanthus development-specific genes would be to determine to which extent genes for specific signal transduction pathways important for development in M. xanthus are conserved in the four other species. Therefore, we analyzed the conservation of the genes in two of the best-studied developmental pathways in M. xanthus, the MrpC/FruA pathway, and the C-signal transduction pathway (fig. 3). Briefly, in the MrpC/FruA pathway, expression of the mrpAB genes encoding a two-component system is induced by starvation (Sun and Shi 2001a, 2001b). Phosphorylated MrpC induces expression of mrpC that encodes a transcriptional regulator of the CRP family (Sun and Shi 2001a, 2001b). In vegetative cells, MrpC is phosphorylated by the Pkn8/Pkn14 cascade both of which are Ser/Thr protein kinases (Nariya and Inouye 2005). During starvation, phosphorylation of MrpC is thought to cease and instead MrpC is thought to be cleaved by the LonD protease to MrpC2 (Ueki and Inouye 2003). The EspA histidine protein kinase by an unknown mechanism inhibits accumulation of MrpC (Higgs et al. 2008). MrpC2 activates transcription of fruA (Ueki and Inouye 2003) encoding a DNA-binding response regulator (Ogawa et al. 1996; Ellehauge et al. 1998) and mazF encoding a toxin protein involved in programmed cell death during development (Nariya and Inouye 2008). Intriguingly, all the genes for the proteins in this pathway except Pkn8 are conserved in S. aurantiaca (fig. 3A), whereas these genes are largely absent in A. dehalogenans, S. cellulosum, and H. ochraceum.

The same picture emerges from an analysis of the conservation of proteins in the C-signal transduction pathway (fig. 3B): Transcription of the csgA gene is induced in response to starvation in a RelA-dependent manner (Crawford and Shimkets 2000). The full-length CsgA protein (p25) is exported and anchored in the outer membrane (Lobedanz and Søgaard-Andersen 2003). Here, p25 is cleaved by the PopC protease to generate p17, the actual C-signal (Kim and Kaiser 1990; Lobedanz and Søgaard-Andersen 2003; Rolbetzki et al. 2008). Transcription of csgA is also regulated by the four ActABCD proteins (Gronewold and Kaiser 2001). C-signal transmission likely induces phosphorylation of FruA (Ellehauge et al. 1998). Phosphorylated FruA inhibits the activity of the Frz chemosensory system (Søgaard-Andersen and Kaiser 1996). Subsequently, the activity of the Ras-like GTPase MglA decreases and cells stop reversing (Jelsbak and Søgaard-Andersen 2002) and eventually aggregate into nascent fruiting bodies. FruA in combination with LadA (Viswanathan et al. 2007b) stimulates transcription of the devTRS operon (Viswanathan et al. 2007b), which encodes three clustered, regularly interspaced short palindromic repeat proteins-associated proteins (Viswanathan et al. 2007a), and alone induces transcription of fdgA (Ueki and Inouye 2005). Several other pathways converge on the C-signal transduction pathway at the level of FruA phosphorylation. These pathways are defined by proteins of two-components systems including SdeK (Pollack and Singer 2001), TodK (Rasmussen and Søgaard-Andersen 2003), RodK (Rasmussen et al. 2005), and RokA (Wegener-Feldbrugge and Søgaard-Andersen 2009) as well as the transcription factor MXAN4899 (Jelsbak et al. 2005). Intriguingly, with the exception of the genes for DevTRS, all these genes are conserved in S. aurantiaca. On the other hand, only a few of the genes in the C-signal transduction pathway are present in A. dehalogenans, S. cellulosum, and H. ochraceum.

The conservation of essentially all the genes for the proteins in these two pathways in S. aurantiaca suggests that these pathways may also exist in S. aurantiaca and could be important for development in this organism. On the other hand, the almost complete absence of the genes for the proteins of these two pathways in S. cellulosum and H. ochraceum strongly suggests that these two pathways do not exist to regulate development in these two organisms.
Because several of the *M. xanthus* development-specific genes present only in *M. xanthus* and *S. aurantiaca* are also conserved in non-Myxococcales species (supplementary table S3, Supplementary Material online), we sought to determine whether these genes were lost from *A. dehalogenans*, *S. cellulosum*, and *H. ochraceum* or transferred by horizontal gene transfer (HGT) to a common ancestor of *M. xanthus* and *S. aurantiaca*. Therefore, we analyzed the phylogenetic distribution of orthologs and homologs of each *M. xanthus* development-specific gene. As a criterion for HGT, we used disagreements between the phylogenetic trees of a gene and that of the species (Beiko et al. 2005).

As shown in supplementary figure S2 (Supplementary Material online), most *M. xanthus* development-specific genes have orthologs or homologs in the four other Myxococcales as well as in non-Myxococcales making conclusions about HGT difficult. Only in the case of hthA (MXAN_6889), SdeK, TodK, and RodK/RokA pathways converge with the C-signal transduction pathway are unknown.

**Fig. 3.** Two signal transduction pathways important for development in *Myxococcus xanthus*. (A) The MrpC/FruA pathway. See the text for details. Genes and proteins are color coded according to conservation as shown in the box. Proteins not circled by a black line are also important for vegetative growth and are, therefore, not included in the list of *M. xanthus*-specific developmental genes. (B) The C-signal transduction pathway. See the text for details. The color code is as in (A). The stippled line indicates the cell envelopes of two neighboring cells. The gray box around FruA~P indicate that the mechanisms by which the MXAN_4899, SdeK, TodK, and RodK/RokA pathways converge with the C-signal transduction pathway are unknown.
Conservation of Developmentally Regulated Genes in M. xanthus

Because there is a sampling bias in our set of M. xanthus development-specific genes, we sought an alternative method to evaluate the extent of conservation of genes important for M. xanthus development in other species. Many genes with a function in M. xanthus development are transcriptionally regulated during fruiting body formation (e.g., see Fig. 3). Therefore, to obtain an unbiased set of genes potentially required for development in M. xanthus, we identified at a global scale genes that are transcriptionally regulated in response to starvation, using data from a previously published M. xanthus DNA microarray experiment (Shi et al. 2008). In this experiment, total RNA was isolated from M. xanthus cells after 0, 2, 4, 6, 9, 12, 15, 18, and 24 h of starvation on a solid surface. Using as a criterion for developmental regulation, a 1.5-fold up- or downregulation in at least one time point during development, we found that, over the course of development, 424 genes were downregulated and 410 upregulated (supplementary tables S5 and S6, Supplementary Material online) and that the COG categories represented by up- and downregulated genes are significantly different (supplementary table S7 and S8, Supplementary Material online).

In total, 338 (80%) of the 424 downregulated genes are conserved in S. aurantiaca, 225 (53%) in A. dehalogenans, 229 (54%) in S. cellulosum, and 220 (52%) H. ochraceum (table 4). Compared with the overall percentages of M. xanthus genes having orthologs in the three other Myxococcales (table 3; summary table 5), downregulated genes are overrepresented in all species, suggesting the existence of a common transcriptional response to starvation in these five species. Moreover, a total of 350 (83%) of all downregulated genes have orthologs outside the Myxococcales (table 4). Nearly half (160) of the downregulated genes are conserved in all five Myxococcales and also conserved outside the order. Three of the largest groups of downregulated genes shared by all five Myxococcales include the categories translation and ribosome structure (J), energy production and conversion (C), and posttranslational modification, protein turnover, chaperones (O), and, thus, likely represent the stringent response in M. xanthus (supplementary table S7, Supplementary Material online). Moreover, a significant fraction of these genes represent proteins of unknown function. There are relatively few (20) downregulated genes conserved in the four fruiting body forming bacteria and, of those, most are involved in posttranslational modification, protein turnover, chaperones (O). Conversely, of the M. xanthus-specific and M. xanthus and S. aurantiaca-specific genes downregulated during development, more than 50% represent proteins of unknown function (X).

In all, 280 (68%) of the 410 M. xanthus genes upregulated during development have orthologs outside the Myxococcales; 281 (69%) of the 410 upregulated genes are found in S. aurantiaca, 105 (26%) in A. dehalogenans, 125 (30%) in S. cellulosum, and 116 (28%) in H. ochraceum (table 4). Compared with the overall percentages of M. xanthus genes having orthologs in the four other Myxococcales (table 5 for summary), M. xanthus upregulated genes shared with S. aurantiaca are overrepresented (69% vs. 63%). As expected, M. xanthus upregulated genes shared with the non-developer A. dehalogenans are slightly underrepresented (26% vs. 31%). Strikingly, the percentage of M. xanthus upregulated genes shared with S. cellulosum (30% vs. 36%) or with H. ochraceum (28% vs. 31%) are also slightly underrepresented.

In agreement with these observations, the conservation of upregulated genes between the four Myxococcales is very different compared with that of downregulated genes: Half of the upregulated genes are found either only in M. xanthus (87; 21%) or are shared with S. aurantiaca (117; 29%) (table 4). Most of these genes are classified as signal transduction proteins or proteins of unknown function (supplementary table S8, Supplementary Material online). Only 26 (6%) of the upregulated genes are conserved in all four fruiters and 40 (10%) genes are conserved in all five species (table 4). Nearly all these more broadly conserved genes within the Myxococcales are also found outside the order (table 4), suggesting that these may be genes involved in starvation-induced stress response mechanisms common to many bacteria.

Discussion

The aim of the work presented here was to investigate the hypothesis that a common a genetic program underlies fruiting body formation in Myxococcales. We used three comparative genomics approaches to address this hypothesis. All approaches strongly indicate that the genetic programs for fruiting body formation in M. xanthus and S. aurantiaca may share many components and perhaps even entire signal transduction pathways in common. All approaches also suggest that these programs in S. cellulosum and H. ochraceum are significantly different from that in M. xanthus. Thus, the overall conclusion from these analyses is that fruiting Myxococcales do not share a common genetic program directing fruiting body formation. This conclusion is based on the following observations: First, a set of 95 M. xanthus development-specific genes—although suffering from a sampling bias—are overrepresented in S. aurantiaca, whereas they occur at the same frequency as an average M. xanthus gene in A. dehalogenans and in S. cellulosum and are underrepresented in H. ochraceum. Second, essentially all proteins in the MrpC/FruA and C-signal signal transduction pathways in M. xanthus are conserved in S. aurantiaca, whereas only a minority of the proteins in these pathways are present in S. cellulosum, A. dehalogenans, or in H. ochraceum. Third, genes upregulated during development in M. xanthus are overrepresented in S. aurantiaca and slightly underrepresented in S. cellulosum, A. dehalogenans, and in H. ochraceum. In contrast, genes downregulated during development in M. xanthus are overrepresented in S. aurantiaca as well as in S. cellulosum, A. dehalogenans, and H. ochraceum. These
differences among up and downregulated genes suggest that the five species studied could share in common a transcriptional response to starvation involving down-regulated genes. The COG classification of the corresponding genes and their high frequency of conservation outside the Myxococcales strongly suggest that this response represents the stringent response.

The overall conclusion that fruiting Myxococcales do not share a common genetic program directing fruiting body formation has two interpretations. One would be that the genetic programs governing fruiting body formation evolved more than once in the Myxococcales. In other words, the last common ancestor of the Myxococcales was not a fruiting bacterium, fruiting bodies are not homologous structures, and fruiting body formation in the Myxococcales is an example of convergent evolution. All five Myxococcales share *M. xanthus* development-specific proteins and upregulated genes in common. If fruiting body formation did in fact evolve more than once, these shared genes could be part of a general response to starvation. A likely precedent for multicellularity evolving more than once within an order are provided by the evolution of multicellularity in the green algae Volvocales, which form planar or spherical colonies. In this group of organisms, it has been suggested that multicellularity has likely evolved several times during evolution (Kirk 1999, 2005).

An alternative explanation for the overall conclusion that fruiting Myxococcales do not share a common genetic program directing fruiting body formation would be that the last common ancestor of the Myxococcales was a fruiting bacterium. Accordingly, this program would have been lost in *A. dehalogenans* probably in parallel with a reduction in genome size. A precedent for an apparently similar loss is provided by the loss of the sporulation program in the order Lactobacillales after the divergence from the last common ancestor of Bacilli (Makarova et al. 2006). In *M. xanthus* and *S. aurantiaca*, the extant version of this program would largely be shared because their separation into separate species occurred relatively late. In particular, the overrepresentation of *M. xanthus* development-specific proteins and genes upregulated during development genes in *S. aurantiaca* and the possible conservation of entire signal transduction pathways in *S. aurantiaca* support this evolutionary scheme. Moreover, the overrepresentation of *M. xanthus* development-specific genes and upregulated genes in *S. aurantiaca* suggest that selection acts to maintain these genes in *M. xanthus* and *S. aurantiaca*. In addition, *M. xanthus* may have evolved development-specific functions as suggested by the existence of development-specific proteins, which are not conserved in *S. aurantiaca*. We speculate that the same is true for *S. aurantiaca*. This idea is supported by the observation that fruiting body formation in *S. aurantiaca* depends on the intercellular signaling molecule stigmolone whereas this molecule has no effect on *M. xanthus* development (Plaga et al. 1998). On the other hand, the *S. cellulosum* and the *H. ochraceum* lineages separated much earlier from the *M. xanthus* and *S. aurantiaca* lineage and during the course of evolution the initially shared program may have diverged to such an extent that the extant overlap with the programs in *M. xanthus* and *S. aurantiaca* is minor. According to this scheme, fruiting bodies are homologous structures generated from nonhomologous proteins. On the basis of the existing data, we cannot distinguish between the two scenarios, that is, convergent evolution or divergent evolution. Likewise, our data do not allow conclusions about the extent to which the genetic programs for fruiting body formation in *S. cellulosum* and *H. ochraceum* may overlap. Regardless, our findings suggest considerable plasticity in and innovation of developmental pathways in the Myxococcales.

The last common ancestor of the Myxococcales is thought to have existed 2.0 Gyr ago (Hedges et al. 2006). The lack of conservation of a genetic program for fruiting body formation and sporulation in Myxococcales is in stark contrast to the conservation of the genetic programs for sporulation in bacilli and clostridia (both of which are low G + C Gram-positive bacteria) and cell cycle control in Alphaproteobacteria. Sporulation in bacilli, which are generally facultative anaerobes, and clostridia, which are obligate anaerobes, has been mostly studied in *B. subtilis*. Except for the components of the phosphorelay that regulates initiation of sporulation in *B. subtilis*, most other regulatory proteins in this program are conserved in sporulating Bacilli and Clostridia (Paredes et al. 2005). Clostridia are thought to have diverged from other low G + C Gram-positive bacteria 2.7 Gyr ago and bacilli 2.3 Gyr ago (Paredes et al. 2005), suggesting that the genetic program for sporulation in these bacteria has been largely conserved for 2.7 Gyr. Cell cycle control in Alphaproteobacteria has been mostly studied in *Caulobacter crescentus*. A recent study of 13 proteins involved in cell cycle control in *C. crescentus* showed that many of these are conserved in other Alphaproteobacteria (Brilli et al. 2010), which had their last common ancestor 2.0 Gyr ago (Hedges et al. 2006), suggesting that the genetic program for cell cycle control in the Alphaproteobacteria was largely in place 2.0 Gyr ago. Cell cycle control is an essential process, and therefore, the possibility of significant evolutionary changes in this program is minor. On the other hand, sporulation is not absolutely essential for survival, thus possibly allowing for a larger degree of change. However, this does not explain the differences in level of conservation of proteins required for sporulation in the Myxococcales and low G + C Gram-positive bacteria.

The evolution of morphological variations in metazoans has been suggested to have mainly involved cis-regulatory mutations affecting gene expression patterns rather than protein function (Stern and Orgogozo 2008; Tirosli et al. 2009) and gene duplication followed by neo- or subfunctionalization (Lynch and Conery 2000). Nevertheless, there are examples in which the formation of homologous structures are regulated by nonhomologous pathways (De Robertis 2008; Sommer 2009). Likewise, orphan genes have been suggested to have important functions in the generation of morphological diversity (Khalturin et al. 2009). Bacterial genomes are thought to evolve mainly...
by incorporation of novel genes by HGT, by intragenomic duplications and by gene loss (Ochman et al. 2000; Lerat et al. 2005; Dagan et al. 2008). It has previously been suggested that several genes important for development in M. xanthus were acquired by HGT and “fashioned into a unique developmental cycle and presumably played a different role in their previous host” (Goldman et al. 2007). Our analyses do not support HGT as a significant source for genes in developmental pathways in M. xanthus. These different conclusions are likely based on the much larger set of bacterial genomes available for our analyses.

In summary, our analyses reveal an unexpected level of plasticity in the genetic programs for fruiting body formation and sporulation in the Myxococcales. Future work will be directed at elucidating whether this plasticity is the result of divergent evolution of a homologous program or convergent evolution of similar structures.

**Supplementary Material**

Supplementary tables S1–S8 and Figures S1–S2 are available at Molecular Biology and Evolution online (http://mbe.oxfordjournals.org/).

**Acknowledgments**

We thank Penelope Higgs and Frank Müller for helpful discussions. US National Science Foundation grant number 0236595 (to W.C.N.) and the Max Planck Society supported this work.

**References**


