AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology

Genome Analysis of the Fruiting Body-Forming Myxobacterium *Chondromyces crocatus* Reveals High Potential for Natural Product Biosynthesis

Nestor Zaburannyi,[a,d] Boyke Bunk,[b,d] Josef Maier,[c,e] Jörg Overmann,[b,d] Rolf Müller[a,d]

Department of Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Pharmaceutical Biotechnology, Saarland University, Saarbrücken, Germany[a]; Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany[b]; ATG:biosynthetics GmbH, Merzhausen, Germany[c]; German Center for Infection Research, Hannover-Braunschweig Site, Braunschweig, Germany[d]; IStLS, Information Services to Life Science, Oberndorf am Neckar, Germany[e]

Here, we report the complete genome sequence of the type strain of the myxobacterial genus *Chondromyces*, *Chondromyces crocatus* Cm c5. It presents one of the largest prokaryotic genomes featuring a single circular chromosome and no plasmids. Analysis revealed an enlarged set of tRNA genes, along with reduced pressure on preferred codon usage compared to that of other bacterial genomes. The large coding capacity and the plethora of encoded secondary metabolite biosynthetic gene clusters are in line with the capability of Cm c5 to produce an arsenal of antibacterial, antifungal, and cytotoxic compounds. Known pathways of the ajudazol, chondramide, chondrochloren, crocacin, crocapeptin, and thuggacin compound families are complemented by many more natural compound biosynthetic gene clusters in the chromosome. Whole-genome comparison of the fruiting-body-forming type strain (Cm c5, DSM 14714) to an accustomed laboratory strain which has lost this ability (nonfruiting phenotype, Cm c5 fr−) revealed genetic changes in three loci. In addition to the low synteny found with the closest sequenced representative of the same family, *Sorangium cellulosum*, extensive genetic information duplication and broad application of eukaryotic-type signal transduction systems are hallmarks of this 11.3-Mbp prokaryotic genome.

Myxobacteria represent a branch of *Deltaproteobacteria* with exceptionally large and GC-rich genomes. Usually isolated from soil, these bacteria are best known for a complex development cycle and as prolific producers of secondary metabolites exhibiting a broad variety of biological activities (1). The chemical diversity of the produced compounds has motivated scientists to study more myxobacterial strains, with an emphasis on phylogenetic variety (2, 3, 4). Several genomes of secondary metabolite-producing myxobacteria have been completely sequenced (5, 6). They exhibit genome sizes ranging from 9.2 to 14.7 Mbp, including the currently largest known bacterial genome of *Sorangium cellulosum* So0157-2 (7). The genomic diversity of myxobacteria is vast, and even strains of the same species may undergo genome reduction or expansion at an order of millions of base pairs (6, 7). Such diversity coincides with deviations in the genomic content of secondary metabolite biosynthetic gene clusters of known genomes, ranging from a prediction of 19 such clusters in *Sandaracinus amylolyticus* DSM 53668 to an estimated 38 clusters in *Sorangium cellulosum* So0157-2. Therefore, to better understand the biosynthetic capabilities of the myxobacteria and the relative frequency of secondary metabolites encoded in their genomes, more representatives are required to be sequenced. We here report our analysis of the genome sequence of *Chondromyces crocatus* Cm c5, the first representative of the genus *Chondromyces* belonging to the family of the *Sorangiineae*.

MATERIALS AND METHODS

Genome sequencing. The genome sequence of *Chondromyces crocatus* Cm c5 was determined through a combination of next- and third-generation technologies available at the Helmholtz Centre for Infection Research (Braunschweig, Germany), LGC Genomics GmbH (Berlin, Germany), and the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). Illumina and Roche 454 platforms were used to generate several small-insert (300- to 400-bp) and long-insert (6,000-bp) genomic libraries, respectively. Next-generation technologies yielded 3.3 Gb of data, which represent ≈300-fold genome coverage. These data were assembled with the help of Abyss-pe, version v3.1.4, and Newbler, version 2.6, software. Additionally, a cosmid library of 1,152 molecules was end sequenced, which represents 4× mean genome coverage, at LGC Genomics GmbH (Berlin, Germany). Primer walking, subcloning of DNA, physical mapping of selected cosmid clones, and Sanger sequencing at GATC Biotech AG (Constance, Germany) were used to fill the remaining gaps. Finally, the sequence of the chromosome was verified by an assembly of data derived from three single-molecule, real-time (SMRT) cells in SMRT Portal, version 2.2.0, software (Pacific Biosciences, Menlo Park, CA, USA) after sequencing at the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). An SMRTbell template library was prepared according to the instructions from Pacific Biosciences (Menlo Park, CA, USA) by following the procedures and checklist for preparation and sequencing of greater than 10 kb of template. Briefly, for preparation of 10-kb libraries, ~10 μg of genomic DNA was end repaired and ligated overnight to hairpin adapters by ap-

Received 14 September 2015 **Accepted** 10 January 2016

Accepted manuscript posted online 15 January 2016

Citation Zaburannyi N, Bunk B, Maier J, Overmann J, Müller R. 2016. Genome analysis of the fruiting body-forming myxobacterium *Chondromyces crocatus* reveals high potential for natural product biosynthesis. Appl Environ Microbiol 82:1945–1957. doi:10.1128/AEM.03011-15.

Editor: C. Vieille, Michigan State University

Address correspondence to Rolf Müller, rolf.mueller@helmholtz-hzi.de.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.03011-15.

plying components from DNA/Polymerase Binding Kit P4 from Pacific Biosciences (Menlo Park, CA, USA). Reactions were carried out according to the manufacturer's instructions. SMRTbell template was exonuclease treated for removal of incompletely formed reaction products. Conditions for annealing of sequencing primers and binding of polymerase to the purified SMRTbell template were assessed with the calculator in RS Remote (Pacific Biosciences, Menlo Park, CA, USA). SMRT sequencing was carried out on a PacBio RS II system (Pacific Biosciences, Menlo Park, CA, USA) taking one 180-min movie for each SMRT cell. In total three SMRT cells were run. The absence of additional replicating elements was confirmed by total DNA extraction in 1% agarose blocks (8) and subsequent pulsed-field gel electrophoresis in a Rotaphor device (Biometra GmbH, Goettingen, Germany).

Genome annotation and characterization. Coding sequences were predicted using Prodigal, version 2.6.2, software (9). In addition to the prediction by Prodigal, to improve the detection of short coding sequences and genes with atypical nucleotide compositions, sequences were compared to Swiss-Prot database release 2015_02 with blastx, version 2.2.26+ (10) (E value of $<1e-5$). Annotation of tRNA genes was based on tRNAscan-SE, version 1.3.1 (11). Functional assignment of the putative coding sequence (CDS) functions was based on kClust, version 1.0, protein clustering (12) with 1,139 RefSeq prokaryotic genomes. Conserved domains of the protein sequences were identified with the Conserved Domain Database (CDD) (13). For a genome comparison, the annotated version of *Sorangium cellulosum* So ce56 was downloaded from the GenBank database, and both genomes were checked for bidirectional blastp hits with an E value of $<1e-10$. A whole-genome synteny plot was generated by Circoletto, version 01.11.13 (14); homologous hits with an E value of $<1e-10$ were considered significant. A cumulative GC skew plot was generated using Gnumeric, version 1.12.17, by dividing the genome sequence into 1,000 equally spaced windows and applying for each window the following formula: $\text{skew}_n = \text{skew}_{n-1} + (G - C)/(G + C)$, where n is the window sequential number. The phylogenetic tree for 13 myxobacterial strains was constructed using Geneious Tree Builder, version 8.1.14 (15), the Tamura-Nei evolution model TN93 (16), and the *Bdellovibrio bacteriovorus* 16S rRNA gene (BD_RS04005) sequence as an outgroup. Bootstrap support values are shown beside interior nodes (except for nodes fully supported by 100 out of 100 runs). The scale bar represents the phylogenetic distance in nucleotide substitutions per site. General sequence statistics were calculated using Geneious, version 8.1.14 (15). To calculate the tRNA/rRNA gene ratio, the GenBank FTP resource containing complete bacterial genomes (ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/) was replicated on 11 December 2014 and processed to obtain values from 2,787 genomes. Mean intergenic distances were calculated as the sum of distances between genes divided by the number of gene features. Orthologs of protein sequences for *C. crocatus* Cm c5, *Myxococcus xanthus* DK 1622, and *S. cellulosum* So ce56 were estimated in Proteinortho, version 5.11 (17), with synteny comparison turned on and an algebraic connectivity threshold of 0.1.

Comparison of Cm c5 fr− and Cm c5 fr+ genomes. Sequencing of the Cm c5 fruiting body-forming strain (Cm c5 fr+) was performed using Illumina paired-end technology at Helmholtz Centre for Infection Research (Braunschweig, Germany). Obtained paired sequencing reads with the length of 301 bp were assembled to the Cm c5 fr+ genome reference in Geneious software, version 8.0.4. The result of the assembly was exported and processed in the GATK toolkit, version 3.3-0 (18) (Unified Genotyper variant caller, ploidy of 1) and SVDetect tool, version 0.7 (19) (window size, 1,000; step length, 1,000; read length, 301; mates orientation, forward-reverse [FR]).

Estimation of biosynthetic potential. The genome sequence of Cm c5 was screened for gene clusters involved in natural product biosynthesis with the antiSMASH, version 2.0.2, tool (20). The fraction of genome that is devoted to secondary metabolism was calculated as the sum of lengths of all detected biosynthetic gene clusters divided by genome length. Loci without known correlation to any particular natural product were submitted for homology searches using tblastx, version 2.2.26+ (10), to a subset of the NCBI nucleotide database. The subset database was created by issuing the query "gene cluster [TITL] 5000:500000 [SLEN]" to the NCBI nucleotide database on 26 November 2014, downloading the 3,117 resulting records, and formatting them into the local nucleotide blast database by the "makeblastdb" command from the BLAST package. The potential of nonribosomal peptide synthetase (NRPS)/type 1 polyketide synthase [PKS(I)] systems encoded in the genome was estimated by comparing the protein sequences of MbtH from *Mycobacterium tuberculosis* H37Rv (GenBank accession number NP_216893) and the phosphopantetheinyl transferase from *Sorangium cellulosum* So ce56 (GenBank accession number CAN95050.1 to all protein sequences of Cm c5 by using blastp, version 2.2.26+ (10).

Analysis of codon adaptation. (i) Source of coding sequences. For comparative analysis CDS sequences of four genomes were used: *Escherichia coli* K-12 MG1655 (NCBI Genome RefSeq NC_000913.3 of 15 May 2014; 4,140 CDSs), *Myxococcus xanthus* DK 1622 (GenBank accession number CP000113.1; 23 June 2014; 7,331 CDSs), *Sorangium cellulosum* So ce56 (GenBank accession number AM746676.1; 23 June 2014; 9,375 CDSs) and *Chondromyces crocatus* Cm c5 (this work; 8,339 CDSs).

(ii) Extraction of CDS subsets. *(a) Hxp2 subsets.* A set of 258 highly translated CDSs covering important cellular functions were selected from *E. coli* K-12 MG1655 based on the ribosome profiling data from Li et al. (21). Homologous sequences (encoding putative orthologs and paralogs) were subsequently selected from three myxobacterial genomes (353 from *M. xanthus* DK 1622 [see File S2, *M. xanthus* DK 1622 Hxp2 tab, in the supplemental material], 388 from *S. cellulosum* So ce56 [see File S2, *S. cellulosum* So ce56 Hxp2 tab], and 352 from *C. crocatus* Cm c5 [see File S2, *C. crocatus* Cm c5 Hxp2 tab]). For *E. coli* several well-translated CDSs with low codon adaptation (e.g., *cspA*) and those with no homologous CDSs in myxobacterial genomes (e.g., the *E. coli*-specific murein lipoprotein CDS *lpp*) were omitted. Paralogous CDSs and functionally related proteins of several basic functional categories (see File S2, CDS categories tab, in the supplemental material) were added to the list. The selected CDSs from *E. coli* comprise 70% of all translation events during logarithmic growth (21). CDSs conserved in all three myxobacterial species and showing a relatively high codon adaptation according to Hxp2 codon usage tables were also added to the myxobacterial Hxp2 sets. Also, 23 CDSs coding for ribosomal proteins, PheS, and InfA (initiation factor A) with extremely low codon adaptation (relative to the codon usage tables calculated from either all CDSs or Hxp2 CDSs) were deleted from the respective lists (see File S2, *C. crocatus* Cm c5 Hxp2 tab, in the supplemental material). The genes for these omitted CDSs cluster within two chromosomal regions, in which all the occurring CDSs show very atypical low codon adaptation.

(b) Msy subsets. For megasynthase (Msy) subsets all CDSs equal to or longer than 2,000 bp were selected if they contained motifs typical for PKSs or NRPSs detected by antiSMASH, version 2.0.2 (20).

(iii) Calculation of codon usage adaptation. Codon usage tables for all CDSs and the Hxp2 subset were calculated with the EMBOSS program cusp (22) after FASTA format headers and line breaks within sequences from the FASTA-formatted set of CDS sequences were removed. The EMBOSS program cai was used for calculation of codon adaptation indices (CAI) (23) for each CDS for both codon usage tables (all CDSs and Hxp2 CDSs). The box plot function of R (R Development Core Team, 2011) was used for display of statistic parameters of CAI distributions, where the boxes around the median line extend to the first and third quartiles of the data, and the whiskers extend to the last value within a range covering 1.5 times the distance from the median to the quartile edges. Outliers from these ranges are indicated by additional circles. Significant differences in the location parameters of the CAI distributions were evaluated with a Mann-Whitney test (function wilcox.test for two samples in R).

(iv) Quantification of codon bias. For analysis of codon bias within genomes and CDS sets (all, Hxp2, and Msy) two different mea-

sures of codon bias were applied: *(a)* measuring $\Delta mCAI_{Hxp2-All}$ and *(b)* measuring ΔH.

(a) Measuring $\Delta mCAI_{Hxp2-All}$. The $\Delta mCAI_{Hxp2-All}$ values represent the difference between the median CAI (mCAI) values of the Hxp2 CDS set and those of all CDSs in the genome, where the CAI values for both of the CDS sets were calculated either with the codon usage table derived from the Hxp2 CDS set or with the codon usage table derived from all sequences. The significance for the distributions of CAI values having different median values was evaluated with a Mann-Whitney test (function wilcox.test for two samples in R). This measure describes the overall differences in codon usage but does not quantify selectively enriched codons in an optimized set of CDSs. $\Delta mCAI_{Msy-All}$ represents the difference between the median CAI values of the Msy CDSs set and those of all CDSs in the genome, respectively.

(b) Measuring ΔH. The codon bias measure ΔH (24) is the improvement in the log likelihood per codon of an evolutionarily optimized set of CDSs when the synonymous codon usage fractions of the optimized set (observation) instead of the fractions of a typical control set (expectation) are used. The measure is identical with the *G* test statistic (likelihood ratio of goodness of fit), normalized by the number of observations (codons) and divided by 2.

The *G* test statistic provides *P* values for significant differences between the codon count distributions of a CDS set and the control set by comparison with the χ^2 distribution with 41 degrees of freedom (the sum of the number of synonymous codons for each amino acid minus 1). The respective *G* statistic for a *P* value of 0.001 is 74.7; for a *P* value of 0.01 it is 65.0, and for a *P* value of 0.05 it is 56.9. ΔH quantifies the enrichment with preferred codons, whether they are most used or not, and represents the strength of translational selection on synonymous codon usage acting on a set of CDSs. For calculation of ΔH we used the Hxp2 CDS set or the Msy CDS set as an optimized CDS set and a control set, which comprises 80% of all CDSs, reduced by the removal of that 20% of CDSs which have the most extreme distance to the typical codon usage. A reference set of CDSs resembling 80% of all CDSs with the most typical codon usage was originally suggested by A. C. Retchless and J. G. Lawrence for their codon bias measure, adaptive codon enrichment (ACE) (25). For ranking the CDSs on typical codon usage, the measure of the distance to native codon usage, D_{ncu} (26), was used, which is the weighted geometrical mean of the Euclidian distances in synonymous codon usage of all amino acids, except for single-codon amino acids Trp and Met, in a CDS.

(v) Calculation of intragenic leucine codon localization. Intragenic distribution of leucine codons was calculated by dividing coding sequences into 100 windows and calculating the leucine codon usage frequencies for each window. Afterwards, this codon usage table was loaded and imported into Gnumeric, version 1.12.17, and subjected to correlation analysis (by selecting statistics → descriptive statistics → correlation).

HGT analysis. To detect putative horizontally transferred genes, the implicit phylogenetic method of protein sequence comparison by the blastp, version 2.2.26+ (10), program was used. Protein sequence queries from the *C. crocatus* Cm c5 genome were used in homology searches against the nonredundant (NR) protein database downloaded from NCBI on 26 November 2014. High-scoring pairs were filtered by E value (<1e−30) and query-to-hit and hit-to-query sequence overlap (>75%). Only the first occurrence of each genus was used to evaluate candidate genes for horizontal gene transfer (HGT). The top 10 genera were tested for the presence of sequences belonging to the *Myxococcales* order, with the exception of sequences belonging to the *Chondromyces* genus. Proteins were assigned as candidates for being horizontally transferred based on the presence of members of *Myxococcales* in the top 20 unique genera of the BLAST hits. The number of 20 top hits was chosen to be comparable with the number of complete and draft myxobacterial genomes available in the downloaded NR database, excluding the genus on which ancestry assessment is being performed (*Chondromyces*): five *Myxococcus* (GenBank accession numbers NC_008095, NC_015711, NC_020126, NZ_AKYI00000000, and NZ_AOBT00000000), four *Anaeromyxobacter* (accession numbers NC_007760, NC_009675, NC_011145, and NC_011891), and two *Sorangium* (accession numbers NC_010162 and NC_021658) genomes and one *Stigmatella* (NC_014623), one *Corallococcus* (NC_017030), one *Cystobacter* (NZ_ANAH00000000), one *Plesiocystis* (NZ_ABCS00000000), and one *Haliangium* (NC_013440) genome.

Nucleotide sequence accession number. The nucleotide sequence of the *C. crocatus* Cm c5 fr+ chromosome was submitted to the GenBank database under accession number CP012159.

RESULTS

General characteristics of *C. crocatus* Cm c5 genome. *Chondromyces crocatus* Cm c5 features a single, circular chromosome, containing 11.3 Mbp and more than 8,000 predicted coding sequences (CDSs), making it one of the largest sequenced prokaryotic genomes to date. No additional extrachromosomal replicating elements are present, a trait which is consistent among myxobacteria. The placement of the genus *Chondromyces* within the high-GC branch of the *Deltaproteobacteria* is in line with the genomic GC content (68.7 mol% G+C) and 16S rRNA gene sequence phylogeny (Fig. 1). The closest sequenced relative is the cellulose-degrading myxobacterium *Sorangium cellulosum* So ce56 (6). As customary, the start of the chromosome and locus numbering was assigned to the first base of *dnaA* coding sequence (position 1), whereas the origin of replication is located 725 kb downstream. The spatial separation between the *dnaA* gene and *oriC* is clearly reflected by the cumulative GC skew pattern (Fig. 2; see also Fig. S1 in the supplemental material for noncumulative GC skew). Interestingly, other highly conserved genes (e.g., *dnaN* and *gyrB*) typically found near the *oriC* site also occur in the vicinity of *oriC* in strain Cm c5. The effect of similar *dnaA* gene translocations on replication is not known (27). It most likely occurred in the ancestor species of *Sorangium* and *Chondromyces* before the two genera diverged and has subsequently been maintained in these particularly slow-growing myxobacteria.

For genome comparison within the order of *Myxococcales*, we chose the closest sequenced myxobacterial relatives, the soil-dwelling, cellulose-degrading *S. cellulosum* So ce56 and the predatory model organism *Myxococcus xanthus* DK 1622. Both reference organisms are motile and capable of forming fruiting bodies. In spite of the rather close phylogenetic relatedness of the strains Cm c5 and So ce56, the global synteny of their genome sequences is low (Fig. 3). This confirms a considerable degree of genome divergence that has been reported within myxobacterial suborders (6) but complicates the elucidation of genome evolution in myxobacteria. Clustering of protein sequences from *C. crocatus* Cm c5, *S. cellulosum* So ce56, and *M. xanthus* DK 1622 into orthologous groups through the Proteinortho software resulted in just 50 to 55% overlap between the proteomes of the first two of these strains (Fig. 4). This indicates that substantial parts of genomes are either not of common ancestral origin or have diverged rapidly. As much as 30% of predicted CDSs from Cm c5, the smaller of the two genomes, have no homologs in So ce56. Therefore, genome reduction is not the single major driving force of genome evolution in these bacteria. A dot plot analysis identified numerous deletions and the presence of extensive nonhomologous loci but also revealed some remaining synteny of the entire genomes (Fig. 3). Overall, our findings on the limited number of myxobacterial genome sequences available thus far reflect the fact that currently

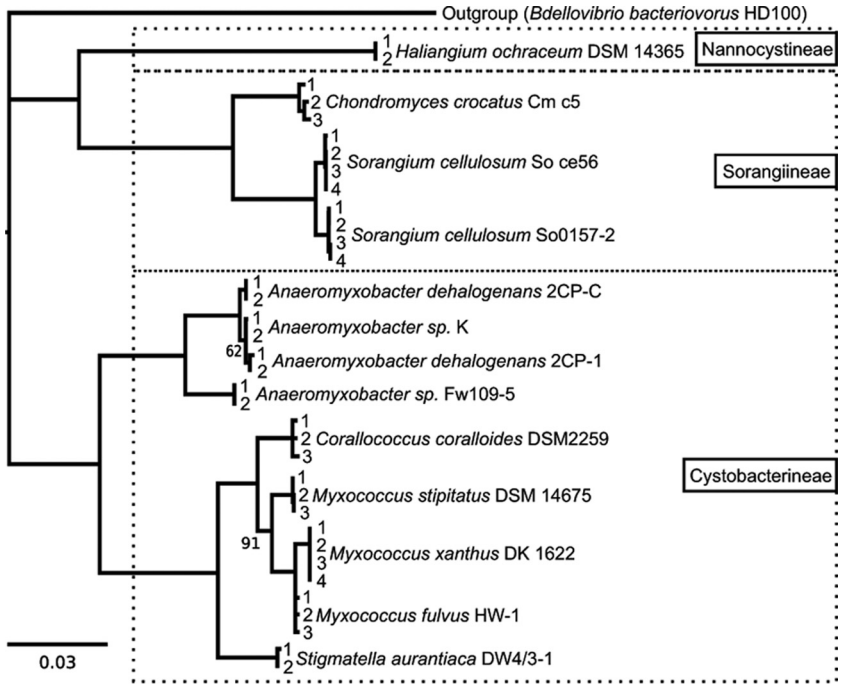


FIG 1 Phylogenetic tree of 16S rRNA gene sequences from 13 myxobacterial genomes of the suborders of *Nannocystineae*, *Sorangiineae*, and *Cystobacterineae*. Between two and four rRNA gene clusters are present in each of the analyzed genomes (indicated by the numbers beside the terminal nodes).

available isolates in pure cultures represent deep-branching lineages with considerable diversification of their genome sequences.

Putative functions were annotated for 42% of the predicted proteins; the remaining protein sequences had no significant similarity to any functionally assigned record in the RefSeq database. The chromosome of Cm c5 contains 93 predicted tRNA genes, 10 of which are presumably pseudogenes, and three rRNA gene clusters. This high tRNA gene/rRNA gene ratio is unusual among prokaryotes (Fig. 5). Out of the 93 tRNA genes, 30 (5 of which are presumably pseudogenes) are clustered in a ≈10-kb region. In the myxobacterial genomes described so far, tRNA genes are rarely aggregated into clusters, with the largest known to contain eight genes. In contrast, numerous long stretches of tRNA genes are characteristic of the smaller, streamlined genomes of pathogenic bacteria, such as *Staphylococcus aureus* (28). Some of the possible implications of the enlarged tRNA gene pool in Cm c5 are discussed further below. The features of the chromosomes of the three myxobacterial strains are summarized in Table 1.

Specific features of codon usage in Cm c5. The observed increased number of tRNA genes in the Cm c5 genome might reflect evolutionary processes that operate on the codon level. We therefore investigated the codon composition of the Cm c5 genome using several approaches. First, the codon adaptation indices (CAI) for the complete CDS set (all), the presumably highly expressed CDS subset (Hxp2), and the megasynthase CDS subset (Msy) of the three myxobacterial genomes were compared to those of *Escherichia coli* K-12 MG1655 (Fig. 6; see also Table S1 in

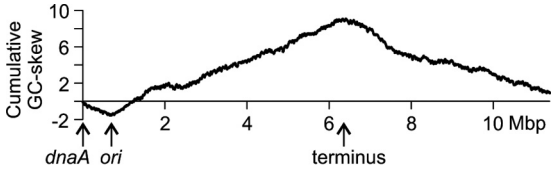


FIG 2 Cumulative GC skew of the Cm c5 chromosome. Each dot represents a window length of 1/1,000 of the chromosome. Arrows indicate the locations of the *dnaA* gene and the origin and terminus of replication.

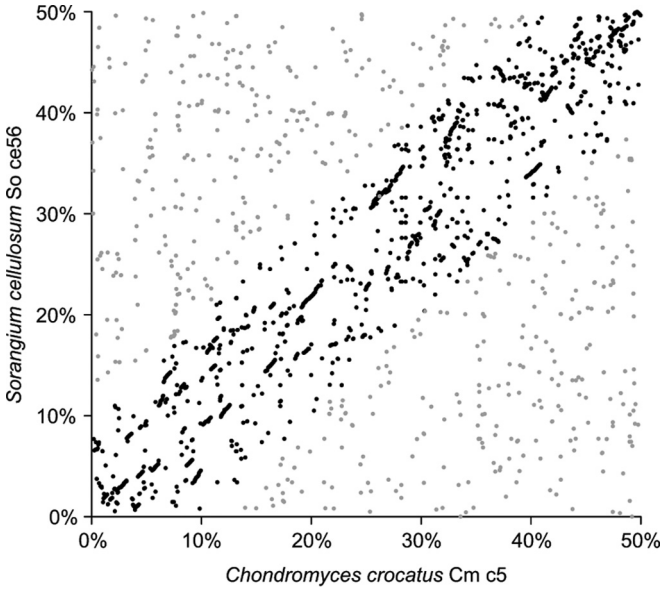


FIG 3 Dot plot comparison of homologous regions between chromosomes of *Chondromyces crocatus* Cm c5 and *Sorangium cellulosum* So ce56. Each dot represents a nucleotide blastn homology hit with an E value of <1e−10. Scales display the homology region's physical distance to *ori* as a percentage of the chromosome (maximum, 50%). Homologies with translocations of no more than 10% are shown in black.

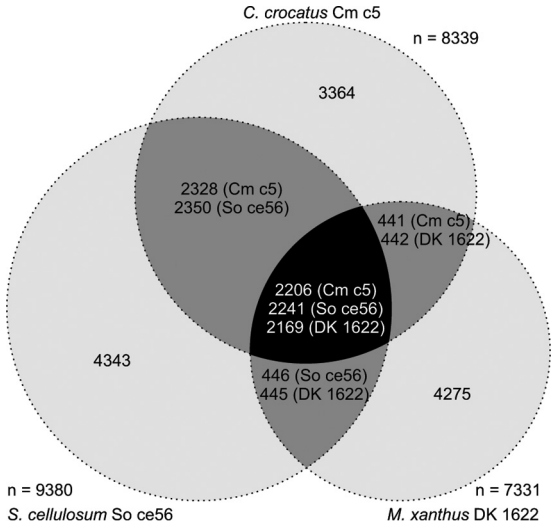


FIG 4 Summary of clustering of protein sequences derived from three genomes of Cm c5, So ce56, and DK 1622 by the Proteinortho program. Overlapping areas of the Venn diagram display common proteins assigned into orthologous clusters in two (dark gray area) or all three (black area) organisms.

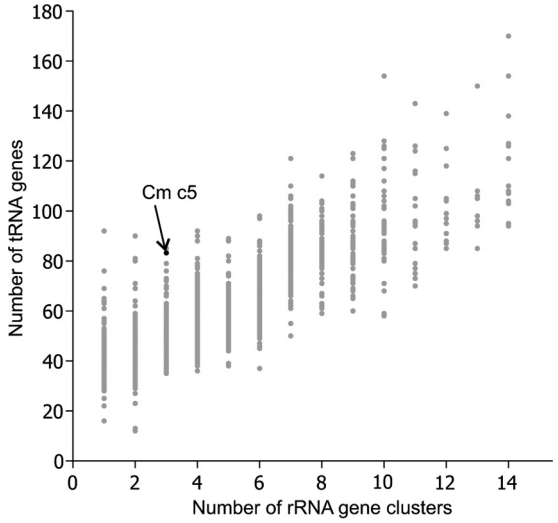


FIG 5 Correspondence between the number of tRNA genes and rRNA gene clusters in bacterial genomes. Data were obtained from sequences of 2,787 complete prokaryotic genomes, including *C. crocatus* Cm c5.

the supplemental material). Next, we used the bias measure ΔH (24) that quantifies the evolutionary selection for codons of a CDS subset (Hxp2 or Msy) compared with that of a control set. The control set consisted of the 80% of all CDSs with the most typical codon usage (26). The CDSs of the three myxobacterial genomes were found to exhibit a wider distribution of CAI values than *E. coli*. This is also true for the respective Hxp2 CDS sets. The distance of the median CAI values, calculated with the codon usage table of all CDSs (Fig. 6, white boxes) between both CDS sets ($\Delta mCAI_{Hxp2-All}$) (see Table S1 in the supplemental material) are similar for the four genomes; however, the distribution of the Hxp2 set of *E. coli* is more focused. The CAI distributions of the Hxp2 sets are significantly different from those of all CDSs for all investigated species. The codon usage bias, expressed as the difference in the median CAI values ($\Delta mCAI_{Hxp2-All}$), when calculated with the codon usage table of the Hxp2 CDSs (Fig. 6, gray boxes) is larger for *E. coli* (0.200) and smaller for the myxobacterial species, decreasing from *M. xanthus* DK 1622 (0.131), to *S. cellulosum* So ce56 (0.100), and *C. crocatus* Cm c5 (0.080). The respective values are all significantly different from those of the control set compared by *G* test statistics. The CAI distributions of the megasynthase (Msy) CDS sets are significantly different from those of all CDSs only for *M. xanthus* DK 1622 (median difference, −0.040), whereas for *S. cellulosum* So ce56 (0.012) and *C. crocatus* Cm c5 (−0.015) the distributions are found to be similar. When calculated with the Hxp2-specific codon usage frequencies, differences of the median CAI values are significant for all three species, with the largest value for *M. xanthus* DK 1622 (−0.185) and this value about halved for both of the other myxobacterial species. The ΔH measure for selective codon enrichment of the Msy CDS set over the 80% control set is lower than that of the respective ΔH values of the Hxp2 CDS set. The *G* test statistics show high significance in selective codon enrichment for all species; they decrease from *M. xanthus* DK 1622 to *C. crocatus* Cm c5.

Despite a growth rate comparable to the rates of *S. cellulosum* So ce56 and *M. xanthus* DK 1622, *C. crocatus* Cm c5 has more tRNA genes predicted by tRNAscan-SE (93; 10 of them are pseudogenes) than the two other myxobacterial species (*S. cellulosum* So ce56, 60 [1 pseudogene]; *M. xanthus* DK 1622, 65 [1 pseudogene]). Many of these additional nonorthologous tRNA genes (18 of 28) are located in one cluster of 30 (pseudogenes) putative tRNA genes (10-kbp genomic region on the minus strand between genes CMC5_054990 and CMC5_055280). These additional tRNA genes either have anticodons for preferred codons or represent the most efficient tRNAs for the respective codon set with the greatest wobble capacity. The additional tRNAs may therefore have been acquired and retained not only to optimize the usage of preferred codons but also to allow better initial expression of foreign genetic material.

For instance, four synonymous codons exist for alanine (GCA, GCC, GCG, and GCU), which are decoded by tRNAs having three different anticodons (GGC, CGC, and UGC). *S. cellulosum* So ce56 and *C. crocatus* Cm c5 share four orthologous tRNA genes, one with anticodon GGC, two with CGC, and one with UGC. The UGC anticodon is generally modified by U_{34} methyl transferases, which enables pairing of its modified form (5-hydroxyuridine derivative, or xo^5U) with GCA, GCG, and GCU codons and, albeit weakly, to the GCC codon. It was shown that tRNA genes harboring anticodons with U at the first position of the anticodon (U_{34}) are generally preferred in the bacterial kingdom, presumably because their wobbling capacity makes decoding much more efficient and faster (29). *C. crocatus* Cm c5 has five additional alanine-specific tRNA genes compared to the genome of *S. cellulosum* So ce56, and of these three are found in the large tRNA gene cluster mentioned above. Four of the new tRNA genes harbor the advantageous anticodon UGC (CMC5_055140, CMC5_055130, CMC5_010950, and CMC5_054770) with enhanced wobbling, whereas one harbors the anticodon GGC (CMC5_055120), which decodes the codon GCC, whose wobble base is not so well decoded by tRNAs with the modified UGC anticodon. Moreover, the four new tRNAs with the UGC anticodon most likely arose by gene duplication of one of the existing conserved tRNA genes, as indicated by their high similar-

TABLE 1 General features of the genomes of *C. crocatus* Cm c5, *S. cellulosum* So ce56, and *M. xanthus* DK 1622

Strain	Chromosome length (bp)	GC content (%)	No. of CDSs	rRNA operons (no.)	No. of tRNA genes (pseudogenes)[a]	Aminoacyl-tRNA synthetases (no.)	Largest CDS (bp)	Avg CDS length (bp)	Avg intergenic distance (bp)	Coding potential (%)	Biosynthetic gene clusters (no.)	Biosynthetic potential (%)
DK 1622	9,139,763	68.9	7,331	4	65 (1)	27	42,825	1,141	106	90.5	24	10.9
Cm c5 fr+	11,388,132	68.7	8,339	3	93 (10)	30	24,477	1,213	150	88.6	35	11
Cm c5 fr−	11,387,322	68.7	8,337	3	93 (10)	30	24,477	1,213	150	88.6	35	11
So ce56	13,033,779	71.4	9,380	4	60 (1)	26	25,254	1,206	176	86.6	29	6.5

[a] The number of pseudogenes included in the total number of tRNA genes is shown.

ity (CMC5_010950 and CMC5_054770 are nearly identical to the conserved UGC anticodon tRNA gene CMC5_060730). One additional tRNA with a similar potential for more efficient decoding and reduction of codon bias was also found for two further synonymous codon sets, the proline-specific UGG anticodon (CMC5_055080), which is also present in the large tRNA gene cluster, and the threonine-specific UGU anticodon (CMC5_056460). Moreover, *C. crocatus* Cm c5 has four additional nonorthologous aminoacyl-tRNA synthetase genes compared with the genome of *S. cellulosum* So ce56 (CMC5_058260 prolyl-tRNA synthetase, CMC5_049140 arginyl-tRNA synthetase, CMC5_038260 threonyl-tRNA synthetase, and CMC5_018350 lysine-tRNA ligase). An enlarged set of tRNAs could empower *C. crocatus* Cm c5 to express a wider set of originally foreign biosynthetic gene clusters.

The importance of specific tRNA genes for expression of secondary metabolite clusters is also demonstrated by the *bldA* gene of streptomycetes coding for the tRNA of the extremely avoided leucine UUA codon (30). *bldA* is specifically expressed under constrained growth conditions and has evolved to control streptomycete secondary metabolite gene clusters, which possess TTA codons in the CDSs of their regulatory transcription factors and in some CDSs of the gene clusters. Interestingly, the TTA codon is not present in any of the 351 CDSs of the presumably highly expressed Hxp2 CDSs of *C. crocatus* Cm c5 and occurs only 330 times in all other CDSs. As in distantly related streptomycetes, TTA codon localization within coding sequences significantly differs from that of five other, synonymous leucine codons. The latter are positioned evenly and have very high correlation coefficients (R of >0.999) to theoretical codon distribution (see Fig. S2 in the supplemental material), while the codon TTA is more often used in the beginning of coding sequences (R of 0.915). It remains to be investigated if the gene coding for a rare leucine tRNA (CMC5_026400) has a specific function in Cm c5 or if lack of the TTA codon in highly expressed sets of genes and uneven intragenic distribution in other coding sequences merely reflect a general rare codon usage tendency of this organism.

In contrast to the enlarged set of tRNA genes, *C. crocatus* Cm c5 has one less ribosomal rRNA gene cluster than *S. cellulosum* So ce56. Half of the ribosomal protein genes, mainly located at two clusters in the genome (CMC5_007430 to CMC5_007650 and CMC5_075030 to CMC5_075050), have unusually low CAI values compared to the respective orthologous genes in *S. cellulosum* So ce56 and *M. xanthus* DK 1622, suggesting also substantially fewer translation events for these genes. The average CAI of these genes was so low that we omitted them tentatively from the Hxp2 list of highly expressed genes. The average CAI, calculated with the codon usage of the Hxp2 CDS set, for these 23 genes from both clusters is 0.430, whereas that for the homologous set of genes in *S.*

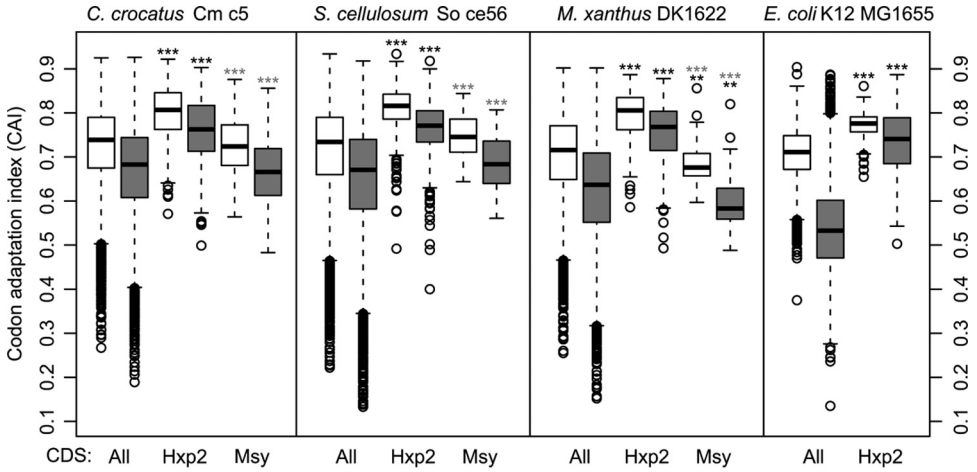


FIG 6 Distribution of CAI of classes of coding sequences (CDSs) based on synonymous codon usage tables constructed either from all CDSs (in white) or from a subset encoding the presumably highly expressed proteins (in gray). The shown CDS classes are as follows: All, all CDSs; Hxp2, presumably highly expressed; Msy, megasynthases. Significant differences between CAI distributions of the classes and the respective distribution for all CDSs (black asterisks) and that of the Msy CDS set and the Hxp2 CDS set (gray asterisks) were calculated using a Mann-Whitney test (***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$). The boxes around the median line extend to the first and third quartiles of the data, and the whiskers extend to the last value within a range covering 1.5 times the distance from the median to the quartile edges. Outliers from these ranges are indicated by additional circles.

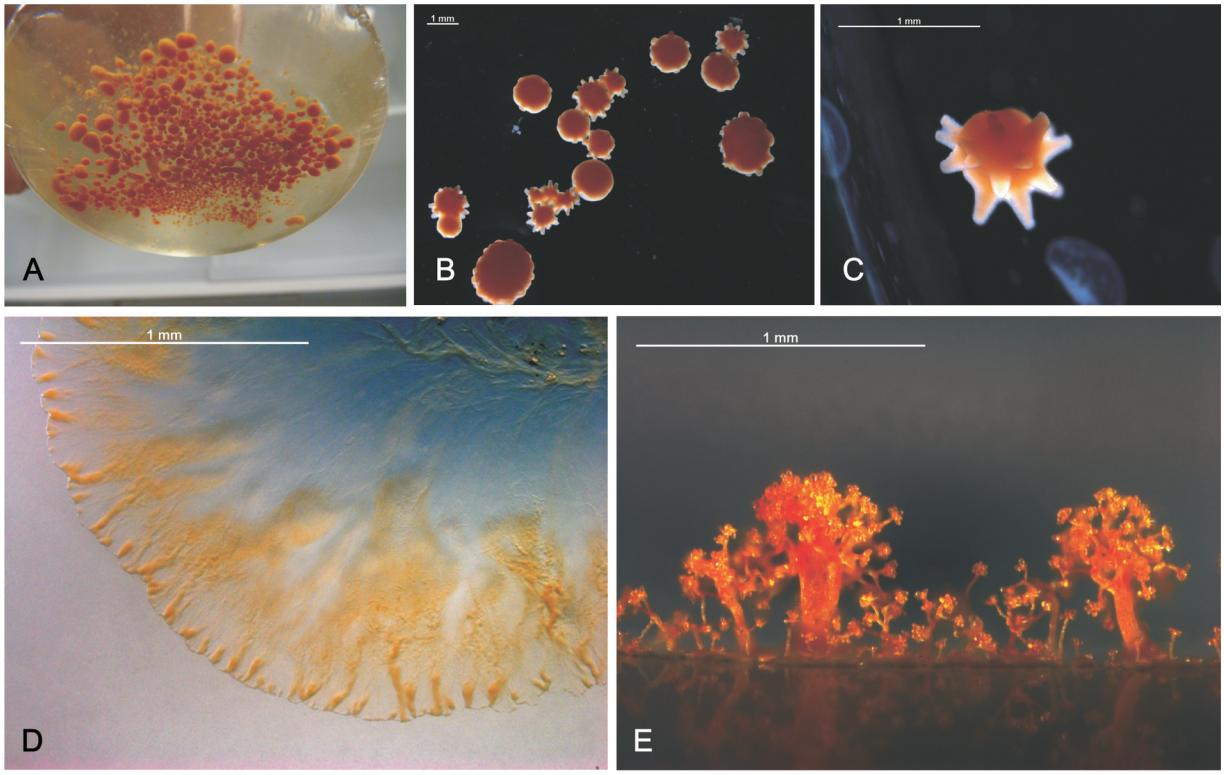


FIG 7 Growth of *C. crocatus* Cm c5 in liquid Pol03 medium and on solid VY/2 medium. (A) Shake flask culture of Cm c5 fr− at 14 days after inoculation. (B) Clumps of Cm c5 fr− of various sizes. (C) A clump of Cm c5 fr−. (D) Swarming of Cm c5 fr− on agar plate. (E) Side view of fruiting bodies of Cm c5 fr+. Scale bar, 1 mm.

cellulosum So ce56 is 0.654, and that for *M. xanthus* DK 1622 is 0.676.

Sensory network and impaired development cycle. One of the major differences of myxobacterial regulatory systems from those of other bacteria is the abundance of eukaryotic-like serine/threonine protein kinases (ELKs), which after this finding were reconsidered as also relevant for prokaryotes (31). Myxobacterial ELKs often comprise a unique and complex domain organization, leading to the presumption that they may be involved in cell signaling pathways. We estimate the number of ELKs in Cm c5 to be 250 (see Table S2 in the supplemental material), which is half the amount of all protein kinases encoded by the human genome (32). Encoded protein kinases, apart from ELK conserved catalytic domains, contain ATPase and ligand-binding and multiple types of repeat domains (TPR, WD40, and Ank), which are thought to be important in both protein-protein and protein-ligand interactions (33). The presence of CheY-binding, cyclase, and transmembrane domains is also consistent with the explanation that the primary function of ELK in myxobacteria is in response to nutrient and developmental signals (34). Apart from ELK-like protein kinases, the predicted proteome contains 97 histidine kinase ATP-binding domains and 88 histidine kinase phosphate acceptor domains. Other elements of transcriptional regulation are abundantly encoded, too, including an estimated 3 repressors of the MarR family, 28 LysR-family and 2 GntR-family regulators, and 25 genes of the TetR class of transcriptional regulators.

Extensive sensory and regulatory networks are a prerequisite to myxobacterial lifestyle, which includes adventurous (A) and social (S) motilities. For the latter system to function, dozens of genes are required, most of which could be tracked in the Cm c5 genome. Three copies of the pilin gene, for example, are found in Cm c5 (CMC5_052100 to CMC5_052120); other genes responsible for motor function of the pili were also found, as well as those responsible for regulation of the extension/retraction of the pilus assembly. Unlike the genes of other myxobacteria (35), these genes are not strictly clustered in Cm c5 (see Table S3 in the supplemental material). A-motility, on the other hand, relies on a large set of myxobacterium-specific genes and extensive cell-to-cell signaling. As in *S. cellulosum* (6), previous knowledge on chemosensory and developmental processes in the model strain *M. xanthus* cannot be readily applied to Cm c5 and its genome since key genes required for A-motility could not be identified. The presence of normal swarm expansion and aggregation during growth on solid and in liquid media, respectively (Fig. 7; see also Fig. S3 in the supplemental material), suggests, however, that both S- and A-motility systems of Cm c5 are functional even if genes responsible for A-motility diverge beyond recognition from those of the model species *M. xanthus*.

Myxobacteria and the genus *Chondromyces*, in particular, are remarkable for their ability to form multicellular structures under stress conditions, the so-called fruiting bodies; components of both motility systems are required in this process. However, during long-term cultivation, our laboratory culture of Cm c5 fr− apparently lost this ability and instead evolved to become a more stable producer of many secondary metabolites. After repeated cultivation trials on Pol03 (Probion Me 069 [Hoechst AG], 0.3%;

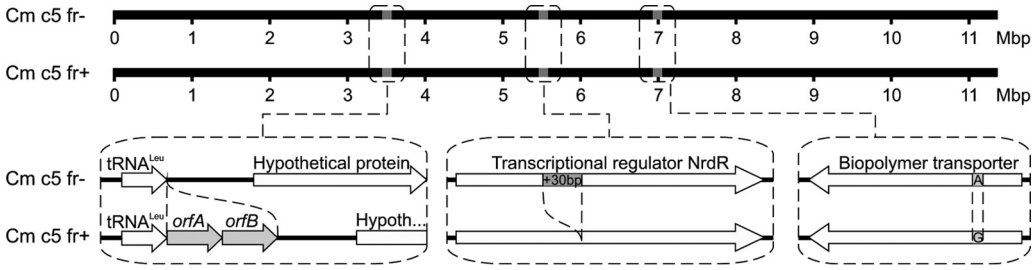


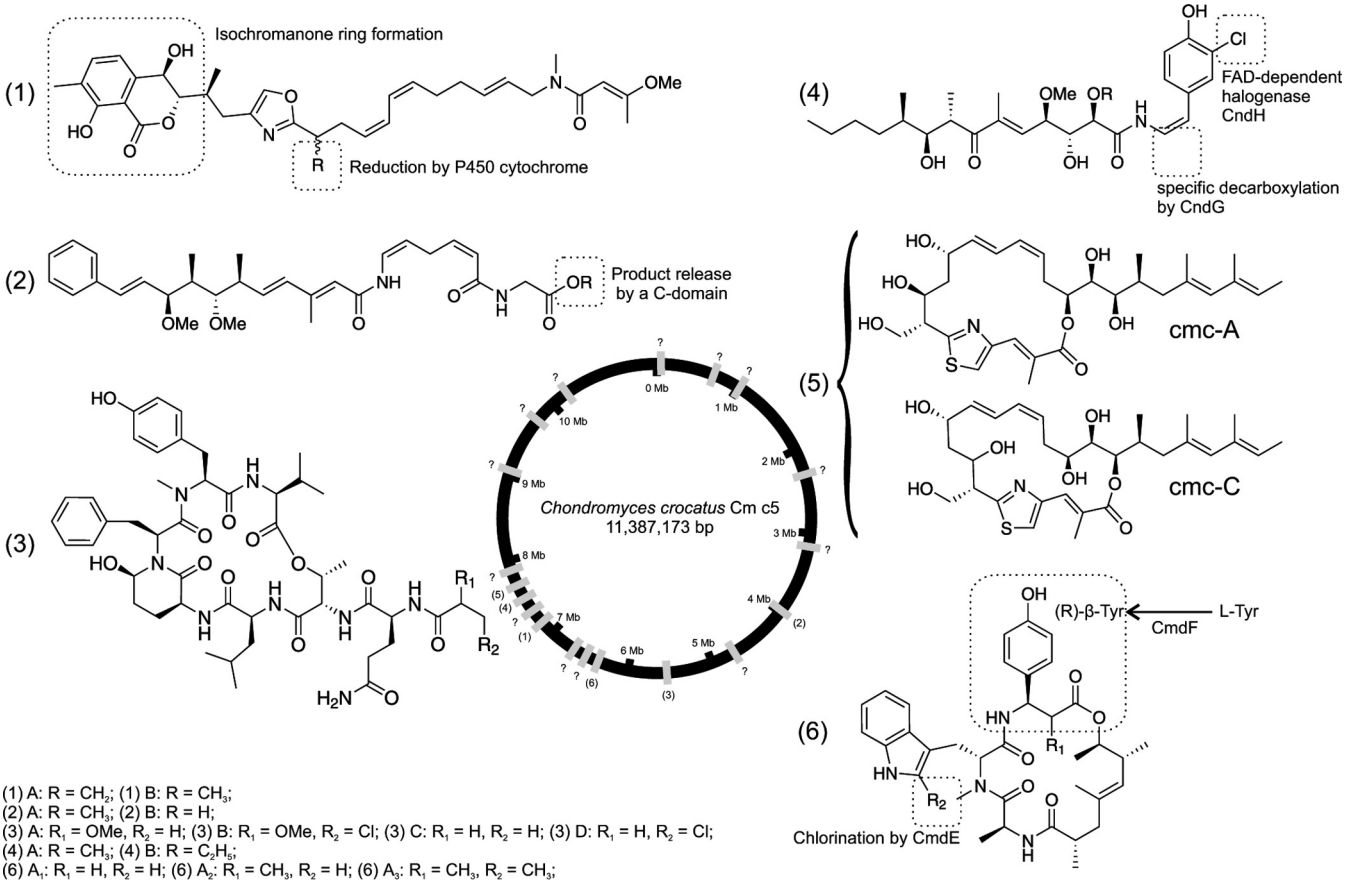
FIG 8 Schematic representation of the three genomic features distinguishing *C. crocatus* Cm c5 fr− from *C. crocatus* Cm c5 fr+. Chromosomes are shown as linear for simplicity. Magnified areas around 3.5 Mbp, 5.5 Mbp, and 7 Mbp represent two genomic rearrangements and one SNP.

soluble starch, 0.3%; $CaCl_2 \cdot 2H_2O$, 0.05%; $MgSO_4 \cdot 7H_2O$, 0.2%; HEPES, 1.19%; pH 7.2 [by addition of KOH]) and VY/2 (73) solid media in order to trigger aggregation and differentiation, we did not observe any signs of Cm c5 fr− forming fruiting bodies.

Comparative genome analysis of several myxobacterial species have recently suggested a shortage of development-specific genes as a possible reason for the nonfruiting phenotype (36). To determine possible reasons for the failure of Cm c5 fr− to produce fruiting bodies, we decided to sequence and analyze another Cm c5 strain (DSM 14714) obtained directly from the DSMZ-German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany), here referred to as Cm c5 fr+. This strain was first confirmed to exhibit a development cycle that includes aggregation into fruiting bodies on VY/2 solid medium (Fig. 7E). Three genetic differences between the strains were detected (Fig. 8): a mobile element transposition in the vicinity of CMC5_026400, a 30-bp insertion into CMC5_041190, and, finally, a single nucleotide polymorphism (SNP) in CMC5_051880. The first genetic difference is a missing/additional copy of the sequence, which resembles bacterial mobile elements and contains two putative coding sequences *orfA* and *orfB*. In the genome of Cm c5 fr−, this mobile element was found in three copies, while in Cm c5 fr+, an additional fourth copy is integrated between genes CMC5_026400 and CMC5_026410. Integration occurred several nucleotides downstream of the CMC5_026400 gene, which functions as $tRNA^{Leu}$ (UAA). Thus, integration of the mobile element is likely to have an effect on tRNA processing around the 3′ CCA motif, which in turn might have significant influence on the global gene expression profile. The second change is an insertion of 30 bp in the CMC5_041190 gene sequence coding for an NrdR family transcriptional regulator in Cm c5 fr−. This insertion does not shift the reading frame of *nrdR*, and, as such, the gene product in Cm c5 fr− is elongated by 10 amino acids in comparison to the sequence in Cm c5 fr+. Repressors of the NrdR family regulate the expression of essential ribonucleotide reductase genes and are thus present in most bacterial species. Highly conserved *nrdR* genes share a structure of one zinc finger domain plus one nucleotide-binding domain. The insertion adds 10 amino acids to the nucleotide-binding domain, which is required for allosteric control (see Fig. S4 in the supplemental material). If a deregulation of deoxynucleoside triphosphate (dNTP) synthesis indeed occurs in the nonfruiting Cm c5 fr− strain, then it may have a significant effect on cell metabolism. Mutation in *nrdR* might thus explain the cause of the impaired growth cycle of the Cm c5 fr− strain. A third genetic variation represents an SNP in the coding sequence of the biopolymer transporter ExbD (CMC5_051880). A homolog of this membrane-bound protein is required for gliding motility in *M. xanthus* (37). The SNP differentiates the respective amino acid residue 43 either into leucine in Cm c5 fr− or proline in Cm c5 fr+. An alignment of known ExbD homologs from myxobacteria shows the proline residue in that position to be inconsistent and found only in Cm c5 fr+, while the leucine residue in the nonfruiting strain seems to be the common variant (see Fig. S5 in the supplemental material). Therefore, the third mutation is thought to have occurred in Cm c5 fr+ and has likely no implications for the lack of fruiting body formation in Cm c5 fr−.

HGT. Several outstanding features of Cm c5 in comparison to its closest sequenced relatives have prompted us to seek for evidence of a recent horizontal gene transfer (HGT) into the genome. Indeed, single genes and whole gene clusters were identified in Cm c5 that are likely to have been gained through HGT from distant bacterial taxonomic groups. This includes a complete aerobic benzoate catabolic pathway consisting of seven genes, probably recently acquired as a whole 9-kb stretch from xenobiotic-degrading *Betaproteobacteria* (38) (CMC5_025720 to CMC5_025780). Similarly, a malonate decarboxylase enzyme complex (CMC5_081220 to CMC5_081280) which catalyzes the conversion of malonate into acetate as well as its coenzyme A (CoA)-activated forms, may provide an alternative pathway for the synthesis of these building blocks for the secondary metabolism of Cm c5. Also, an operon of six genes encoding the monovalent ion antiporter MrpA-MrpG (CMC5_018120 to CMC5_018170) highly similar to antiporters of *Alphaproteobacteria*, along with a uracil-DNA glycosylase repair protein (CMC5_026110) likely from a similar bacterial source, was found. The proposed horizontal transfer events are further supported by the fact that these locus combinations are not present in any of the other myxobacterial genomes sequenced to date and, where applicable, by gene synteny.

While repeated HGTs between bacteria coexisting in soil were to be expected, the Cm c5 genome appears to contain few genes presumably derived from nonbacterial sources. In addition to several bacterial chitinases and chitinase-like proteins (CMC5_022560, CMC5_022660, CMC5_025630, CMC5_025640, and CMC5_067020), the product of another Cm c5 gene (CMC5_079670) has full-length identity of >50% to glycoside hydrolase 19 (GH19) family class 1 and 2 chitinases of higher plants. Another, even more intriguing, example of foreign gene capture may be CMC5_013480, which codes for a bifunctional 6-phosphofructo-2-kinase/fructose-2,6-bisphosphate-2-phosphatase. Known as one of the key regulators of primary carbohydrate metabolism (39), this enzyme is currently believed to be found predominantly in eukaryotic organisms but also in terminal clades of *Deltaproteobacteria*, *Desulfovibrio* (40), and, as shown now, in Cm c5, implying recent horizontal gene transfer event(s).



(1) A: R = CH_2; (1) B: R = CH_3;
(2) A: R = CH_3; (2) B: R = H;
(3) A: R_1 = OMe, R_2 = H; (3) B: R_1 = OMe, R_2 = Cl; (3) C: R_1 = H, R_2 = H; (3) D: R_1 = H, R_2 = Cl;
(4) A: R = CH_3; (4) B: R = C_2H_5;
(6) A_1: R_1 = H, R_2 = H; (6) A_2: R_1 = CH_3, R_2 = H; (6) A_3: R_1 = CH_3, R_2 = CH_3;

FIG 9 Natural compound families isolated from cultures of *C. crocatus* Cm c5: ajudazols (1), crocacins (2), crocapeptins (3), chondrochlorens (4), thuggacins (5), and chondramides (6). Unusual features for bacterial secondary metabolism and newly described reactions or enzymes are highlighted here and referred to in the text. Biosynthetic gene clusters are shown on the schematic representation of the chromosome as gray rectangles. The 13 uncorrelated gene clusters of PKS/NRPS type are shown with question marks. FAD, flavin adenine dinucleotide.

The possible role of this strictly regulatory eukaryotic enzyme in these prokaryotes remains obscure. The HGT therefore undoubtedly contributes to the genetic makeup of *C. crocatus* Cm c5, possibly including some of its secondary metabolite biosynthetic gene clusters. A list of HGT candidate genes, the majority of which link Cm c5 to phylogenetically distant prokaryotes, is shown in Table S4 in the supplemental material.

Biosynthesis of natural products. Natural compounds derived from myxobacteria exhibit unusual and diverse activities (41). The number and activity spectra of small molecules produced by *C. crocatus* Cm c5 underline the valuable biosynthetic capacity of these bacterial genomes once again. To estimate the biosynthetic potential, the genome sequence of *C. crocatus* Cm c5 was analyzed by antiSMASH (20), which identified 35 putative biosynthetic gene clusters (see Table S5 in the supplemental material). Of these 35 clusters, several are known to yield natural products (Fig. 9): ajudazols A and B (42) and crocacins A and B (43), both antifungal inhibitors of the mitochondrial respiratory chain (44, 45); chondramides A to D (46), which block eukaryotic cell division by interacting with the actin system and are active against fungi (47), algae (48), and several tumor cell lines (49); chondrochlorens A and B (50), with weak antibacterial activity (51); thuggacins cmc-A and cmc-C (52), late-step respiratory chain inhibitors of *Mycobacterium* spp. and related Gram-positive bacteria (53); and the potent protease inhibitors of the crocapeptin family (54). The biosynthesis of these and many other natural products is directed by gene clusters encoding large, multienzyme complexes of two principal classes, polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs). Furthermore, owing to their similar and compatible biosynthetic logic (55), PKS and NRPS systems can interchange substrates in both directions (NRPS ↔ PKS), yielding hybrid assembly lines and therefore even more natural product variability. Of the known compound families synthesized by Cm c5, most are in fact assembled on hybrid systems, except for crocapeptins, which originate from a pure NRPS enzyme complex.

The genome of Cm c5, in addition to six known biosynthetic pathways, encodes pathways for at least a dozen other compounds. Knowledge of the biosynthetic gene cluster sequences of the complete genome allows for more accurate characterization of the secondary metabolite profile of an organism by genome mining (56). An uncommon biosynthetic profile of a genome increases chances to encounter compounds that have not bee previously described and that therefore may represent novel activities (57). This is also the case with strain Cm c5, which is phylogenetically highly divergent from known myxobacteria, including the closest relative *S. cellulosum* So ce56. None of the identified biosynthetic clusters from Cm c5 seem to correspond to those known

from So ce56. However, the reconstruction of genomic loci involved in secondary metabolism is inherently difficult because of their repetitive structures (58). This fact, along with the exceptional genome size, also adds to difficulties of complete genome reconstruction of natural product-rich myxobacteria. The availability of high-quality genomic sequences is therefore still one of the barriers for the widespread application of genome mining to metabolite-producing strains of myxobacteria.

Apart from the set of known natural compounds and their underlying biosynthetic genes, 29 other gene loci identified are putatively involved in natural product biosynthesis, including two NRPSs, five type I PKSs, and six combined PKS(I)/NRPS systems. The remaining 17 non-PKS(I)/NRPS clusters comprise five terpene, two bacteriocin, one lantipeptide, two lantipeptide-bacteriocin, one type III PKS, one siderophore, one phenazine, and four clusters of other types. However, great care must be taken before relying on these purely bioinformatic predictions. The 29 unassigned clusters (see Table S5 in the supplemental material) displayed no significant homology to a subset of the NCBI nucleotide database containing gene cluster sequences, emphasizing the likelihood for these gene clusters to encode the formation of novel secondary metabolites and the need to continue research on these novel biosynthetic pathways. Noteworthy, the only siderophore cluster in Cm c5 containing four genes, CMC5_023400 to CMC5_023430, does not correspond to the usual and highly conserved myxochelins (59) among myxobacteria, and thus iron uptake is likely mediated via an alternate chelator.

The biosynthetic potential of NRPS/PKS systems in bacterial genomes can also be estimated by the number and diversity of several small marker proteins, such as MbtH homologs (60), or the number of 4′-phosphopantetheinyl transferases (61). For a total of 18 biosynthetic gene clusters embodying NRPS and type I PKS domains, only one MbtH homolog could be identified (CMC5_078140), and two genes encoding 4′-phosphopantetheinyl transferases (CMC5_018000 and CMC5_023730) are found in the 11.4-Mb chromosome of *C. crocatus*. These facts alone would have placed the strain in the "not-so-gifted" category (60) of natural producer strains, which hardly seems to be the case. Obviously, otherwise successful approaches cannot be readily applied for the analysis of myxobacterial genomes.

DISCUSSION

Many genomes, even those of uncultured bacteria have become accessible in recent years. One group largely missing in the current inventory are the myxobacteria, ubiquitous but slowly growing and motile microorganisms that often act collectively and are valuable producers of bioactive natural compounds with diverse bioactivities. The relatively slow pace of studying their genomes can be attributed to the complex, highly repetitive genome structures, as well as to isolation and cultivation challenges. This complete genome sequence of the type strain *C. crocatus* Cm c5 allows better insight into the genome structure, large coding capacity, and plentiful secondary metabolism of the genus *Chondromyces*. A large portion of the genome consisting of hundreds of encoded proteins is devoted to regulatory functions, possibly even approaching practical limits (62).

The specific bacterial lifestyle has direct implications on codon usage. Since faster growth enhances the selection for more rapidly translated codons, they are often more in use. It is therefore reasonable to suggest that the different degrees of codon bias, which we have observed here, reflects the different life styles of the respective organisms: rapid growth of *E. coli* with strongly biased codon usage versus slower growth in the predatory myxobacterium *M. xanthus* DK 1622 and a further decrease of codon bias in the even more slowly growing *S. cellulosum* So ce56 and *C. crocatus* Cm c5, which are more specialized to utilize complex carbohydrates. To compensate for unbalanced expression of highly translated CDSs, which would be caused by more tRNA genes, one ribosomal rRNA gene cluster may have been lost, and the expression of many ribosomal protein genes was adjusted by reduction of the adaptation of their codon usage. The combined effects have also reduced the general codon bias of the whole genome and allowed *C. crocatus* Cm c5 to harbor and express more gene clusters of presumed foreign origin than its competitors.

Cell motility is another important part of the myxobacterial lifestyle and is required to form myxobacterial fruiting bodies in *M. xanthus*. It was because of these complex, sporangium-like structures that the genus *Chondromyces* was originally considered a fungus until it was recognized as a bacterium (63). The genome comparison of the nonfruiting to the fruiting strain revealed mutations in three loci, including the vicinity of the $tRNA^{Leu}$ (UAA), in an NrdR regulatory protein and in an ExbD membrane protein. It is well known that tRNA genes are among the favored loci for mobile elements, as well as phage and plasmid integration in bacteria (64). A similar scenario, where a transposon had inserted near a tRNA gene, was observed in the myxobacterium *Stigmatella aurantiaca* DW4/3-1 and resulted in a mutant strain unable to form fruiting bodies (65). However, in the case reported here, the strain with the affected $tRNA^{Leu}$ (UAA) gene is capable of fruiting (Cm c5 fr+), not vice versa. The second mutation discovered in the allosteric part of the NrdR regulator could also have a significant effect on DNA synthesis. The third genetic modification affects one amino acid residue in the ExbD membrane protein but is not thought to cause the deficient phenotype of the Cm c5 fr− strain. Obviously, further work is required to establish the reason for the nonfruiting phenotype of Cm c5 fr−.

Small natural bioactive molecules derived from bacteria constitute one of the pillars of drug discovery in pharmaceutical sciences, providing remarkable chemical diversity and application range (58). Over the years, bacteria of the genus *Chondromyces* have proven to be valuable resources for biosynthetic diversity of natural products and the concomitant discovery of unique enzyme functions. Several uncommon features of natural product biosynthesis and biosynthetic enzymes were identified (66, 67, 68, 69), and two previously unknown functions were attributed to the multimodular megasynthases of Cm c5 (43, 70). Intriguing biosynthetic capabilities of Cm c5 are further supported by the structural similarity of chlorinated terrestrial chondramides derived from *C. crocatus* Cm c5 and brominated marine jasplakinolides (see Fig. S6 in the supplemental material) isolated from sponge *Jaspis* sp. (71). This similarity suggests a shared underlying structure of biosynthetic proteins and/or gene clusters. One possible explanation for this finding might be the presence of highly similar biosynthetic pathways in yet uncharacterized microorganisms inhabiting sponges. Structural similarities between crocapeptins isolated from Cm c5 and cyanopeptolins isolated from cyanobacteria (72) also suggest horizontal transfer of ancestral biosynthetic gene clusters.

The secondary metabolite profile of Cm c5 is likely to be extended by further natural products and possibly more connec-

tions to the marine environment. Therefore, the genome sequence of this already valuable natural producer strain will allow a better understanding of its biosynthetic potential. Indeed, the majority of the unassigned biosynthetic pathways are encoded in gene clusters without any significant sequence similarity to other bacterial genomes, and these are therefore the most promising candidates for further studies of the strain's rich secondary metabolism. To achieve the goal of unraveling the structures of these currently unknown metabolites, studies into the physiology and regulation of Cm c5 will have to be connected with natural product research in the future.

ACKNOWLEDGMENTS

We thank Silke Wenzel, Cathrin Spröer, and the GMAK group at the Helmholtz Centre for Infection Research for support during genome sequencing and valuable discussions. We thank Ronald Garcia for providing microscopic images. We thank Gregor Zipf and Hubert Bernauer from ATG:biosynthetics GmbH for contributions to the analysis. We thank Sara Andes, Simone Severitt, and Nicole Mrotzek for technical assistance.

We declare that we have no conflicts of interest.

REFERENCES

1. **Reichenbach H, Höfle G.** 1993. Biologically active secondary metabolites from myxobacteria. Biotechnol Adv **11:**219–277. http://dx.doi.org/10.1016/0734-9750(93)90042-L.
2. **Plaza A, Garcia R, Bifulco G, Martinez JP, Hüttel S, Sasse F, Meyerhans A, Stadler M, Müller R.** 2012. Aetheramides A and B, potent HIV-inhibitory depsipeptides from a myxobacterium of the new genus "*Aetherobacter*." Org Lett **14:**2854–2857. http://dx.doi.org/10.1021/ol3011002.
3. **Wenzel SC, Müller R.** 2009. The impact of genomics on the exploitation of the myxobacterial secondary metabolome. Nat Prod Rep **26:**1385–1407. http://dx.doi.org/10.1039/b817073h.
4. **Garcia RO, Krug D, Müller R.** 2009. Chapter 3. Discovering natural products from myxobacteria with emphasis on rare producer strains in combination with improved analytical methods. Methods Enzymol **458:**59–91. http://dx.doi.org/10.1016/S0076-6879(09)04803-4.
5. **Goldman BS, Nierman WC, Kaiser D, Slater SC, Durkin AS, Eisen JA, Ronning CM, Barbazuk WB, Blanchard M, Field C, Halling C, Hinkle G, Iartchuk O, Kim HS, Mackenzie C, Madupu R, Miller N, Shvartsbeyn A, Sullivan SA, Vaudin M, Wiegand R, Kaplan HB.** 2006. Evolution of sensory complexity recorded in a myxobacterial genome. Proc Natl Acad Sci U S A **103:**15200–15205. http://dx.doi.org/10.1073/pnas.0607335103.
6. **Schneiker S, Perlova O, Kaiser O, Gerth K, Alici A, Altmeyer MO, Bartels D, Bekel T, Beyer S, Bode E, Bode HB, Bolten CJ, Choudhuri JV, Doss S, Elnakady YA, Frank B, Gaigalat L, Goesmann A, Groeger C, Gross F, Jelsbak L, Jelsbak L, Kalinowski J, Kegler C, Knauber T, Konietzny S, Kopp M, Krause L, Krug D, Linke B, Mahmud T, Martinez-Arias R, McHardy AC, Merai M, Meyer F, Mormann S, Muñoz-Dorado J, Perez J, Pradella S, Rachid S, Raddatz G, Rosenau F, Rückert C, Sasse F, Scharfe M, Schuster SC, Suen G, Treuner-Lange A, Velicer GJ, Vorhölter F-J, et al.** 2007. Complete genome sequence of the myxobacterium *Sorangium cellulosum*. Nat Biotechnol **25:**1281–1289. http://dx.doi.org/10.1038/nbt1354.
7. **Han K, Li Z, Peng R, Zhu L, Zhou T, Wang L, Li S, Zhang X, Hu W, Wu Z, Qin N, Li Y.** 2013. Extraordinary expansion of a *Sorangium cellulosum* genome from an alkaline milieu. Sci Rep **3:**2101. http://dx.doi.org/10.1038/srep02101.
8. **Kieser T, Bibb MJ, Buttner KF, Chater KF, Hopwood DA.** 2000. Practical *Streptomyces* genetics. The John Innes Foundation, Norwich, England.
9. **Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ.** 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics **11:**119. http://dx.doi.org/10.1186/1471-2105-11-119.
10. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J Mol Biol **215:**403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.
11. **Lowe TM, Eddy SR.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res **25:**955–964. http://dx.doi.org/10.1093/nar/25.5.0955.
12. **Hauser M, Mayer CE, Söding J.** 2013. kClust: fast and sensitive clustering of large protein sequence databases. BMC Bioinformatics **14:**248. http://dx.doi.org/10.1186/1471-2105-14-248.
13. **Marchler-Bauer A, Zheng C, Chitsaz F, Derbyshire MK, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, Zhang D, Bryant SH.** 2013. CDD: conserved domains and protein three-dimensional structure. Nucleic Acids Res **41:**D348–D352. http://dx.doi.org/10.1093/nar/gks1243.
14. **Darzentas N.** 2010. Circoletto: visualizing sequence similarity with Circos. Bioinformatics **26:**2620–2621. http://dx.doi.org/10.1093/bioinformatics/btq484.
15. **Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A.** 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics **28:**1647–1649. http://dx.doi.org/10.1093/bioinformatics/bts199.
16. **Tamura K, Nei M.** 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol **10:**512–526.
17. **Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ.** 2011. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics **12:**124. http://dx.doi.org/10.1186/1471-2105-12-124.
18. **McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA.** 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res **20:**1297–1303. http://dx.doi.org/10.1101/gr.107524.110.
19. **Zeitouni B, Boeva V, Janoueix-Lerosey I, Loeillet S, Legoix-né P, Nicolas A, Delattre O, Barillot E.** 2010. SVDetect: a tool to identify genomic structural variations from paired-end and mate-pair sequencing data. Bioinformatics **26:**1895–1896. http://dx.doi.org/10.1093/bioinformatics/btq293.
20. **Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T.** 2013. antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res **41:**W204–W212. http://dx.doi.org/10.1093/nar/gkt449.
21. **Li G-W, Burkhardt D, Gross C, Weissman JS.** 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell **157:**624–635. http://dx.doi.org/10.1016/j.cell.2014.02.033.
22. **Rice P, Longden I, Bleasby A.** 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet **16:**276–277. http://dx.doi.org/10.1016/S0168-9525(00)02024-2.
23. **Sharp PM, Li WH.** 1987. The codon Adaptation Index—a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res **15:**1281–1295. http://dx.doi.org/10.1093/nar/15.3.1281.
24. **Ran W, Higgs PG.** 2012. Contributions of speed and accuracy to translational selection in bacteria. PLoS One **7:**e51652. http://dx.doi.org/10.1371/journal.pone.0051652.
25. **Retchless AC, Lawrence JG.** 2012. Ecological adaptation in bacteria: speciation driven by codon selection. Mol Biol Evol **29:**3669–3683. http://dx.doi.org/10.1093/molbev/mss171.
26. **Qian W, Yang J-R, Pearson NM, Maclean C, Zhang J.** 2012. Balanced codon usage optimizes eukaryotic translational efficiency. PLoS Genet **8:**e1002603. http://dx.doi.org/10.1371/journal.pgen.1002603.
27. **Mackiewicz P, Zakrzewska-Czerwińska J, Zawilak A, Dudek MR, Cebrat S.** 2004. Where does bacterial replication start? Rules for predicting the *oriC* region. Nucleic Acids Res **32:**3781–3791.
28. **Green CJ, Vold BS.** 1993. *Staphylococcus aureus* has clustered tRNA genes. J Bacteriol **175:**5091–5096.
29. **Novoa EM, Pavon-Eternod M, Pan T, Ribas de Pouplana L.** 2012. A role for tRNA modifications in genome structure and codon usage. Cell **149:**202–213. http://dx.doi.org/10.1016/j.cell.2012.01.050.
30. **Chater KF, Chandra G.** 2008. The use of the rare UUA codon to define "expression space" for genes involved in secondary metabolism, develop-

ment and environmental adaptation in streptomyces. J Microbiol **46:**1–11. http://dx.doi.org/10.1007/s12275-007-0233-1.
31. **Pérez J, Castañeda-García A, Jenke-Kodama H, Müller R, Muñoz-Dorado J.** 2008. Eukaryotic-like protein kinases in the prokaryotes and the myxobacterial kinome. Proc Natl Acad Sci U S A **105:**15950–15955. http://dx.doi.org/10.1073/pnas.0806851105.
32. **Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S.** 2002. The protein kinase complement of the human genome. Science **298:**1912–1934. http://dx.doi.org/10.1126/science.1075762.
33. **Grove TZ, Cortajarena AL, Regan L.** 2008. Ligand binding by repeat proteins: natural and designed. Curr Opin Struct Biol **18:**507–515. http://dx.doi.org/10.1016/j.sbi.2008.05.008.
34. **Pereira SFF, Goss L, Dworkin J.** 2011. Eukaryote-like serine/threonine kinases and phosphatases in bacteria. Microbiol Mol Biol Rev **75:**192–212. http://dx.doi.org/10.1128/MMBR.00042-10.
35. **Kaiser D, Robinson M, Kroos L.** 2010. Myxobacteria, polarity, and multicellular morphogenesis. Cold Spring Harb Perspect Biol **2:**a000380. http://dx.doi.org/10.1101/cshperspect.a000380.
36. **Huntley S, Hamann N, Wegener-Feldbrügge S, Treuner-Lange A, Kube M, Reinhardt R, Klages S, Müller R, Ronning CM, Nierman WC, Søgaard-Andersen L.** 2011. Comparative genomic analysis of fruiting body formation in *Myxococcales*. Mol Biol Evol **28:**1083–1097. http://dx.doi.org/10.1093/molbev/msq292.
37. **Youderian P, Burke N, White DJ, Hartzell PL.** 2003. Identification of genes required for adventurous gliding motility in *Myxococcus xanthus* with the transposable element *mariner*. Mol Microbiol **49:**555–570. http://dx.doi.org/10.1046/j.1365-2958.2003.03582.x.
38. **Gescher J, Ismail W, Olgeschlager E, Eisenreich W, Worth J, Fuchs G.** 2006. Aerobic benzoyl-coenzyme A (CoA) catabolic pathway in *Azoarcus evansii*: conversion of ring cleavage product by 3,4-dehydroadipyl-CoA semialdehyde dehydrogenase. J Bacteriol **188:**2919–2927. http://dx.doi.org/10.1128/JB.188.8.2919-2927.2006.
39. **Wu C, Khan SA, Peng L-J, Lange AJ.** 2006. Roles for fructose-2,6-bisphosphate in the control of fuel metabolism: beyond its allosteric effects on glycolytic and gluconeogenic enzymes. Adv Enzyme Regul **46:**72–88. http://dx.doi.org/10.1016/j.advenzreg.2006.01.010.
40. **Michels PA, Rigden DJ.** 2006. Evolutionary analysis of fructose 2,6-bisphosphate metabolism. IUBMB Life **58:**133–141. http://dx.doi.org/10.1080/15216540600688280.
41. **Weissman KJ, Müller R.** 2010. Myxobacterial secondary metabolites: bioactivities and modes-of-action. Nat Prod Rep **27:**1276–1295. http://dx.doi.org/10.1039/c001260m.
42. **Buntin K, Rachid S, Scharfe M, Blöcker H, Weissman KJ, Müller R.** 2008. Production of the antifungal isochromanone ajudazols A and B in *Chondromyces crocatus* Cm c5: biosynthetic machinery and cytochrome P450 modifications. Angew Chem Int Ed Engl **47:**4595–4599. http://dx.doi.org/10.1002/anie.200705569.
43. **Müller S, Rachid S, Hoffmann T, Surup F, Volz C, Zaburannyi N, Müller R.** 2014. Biosynthesis of crocacin involves an unusual hydrolytic release domain showing similarity to condensation domains. Chem Biol **21:**855–865. http://dx.doi.org/10.1016/j.chembiol.2014.05.012.
44. **Kunze B, Jansen R, Höfle G, Reichenbach H.** 1994. Crocacin, a new electron transport inhibitor from *Chondromyces crocatus* (myxobacteria). Production, isolation, physico-chemical and biological properties. J Antibiot **47:**881–886.
45. **Kunze B, Jansen R, Höfle G, Reichenbach H.** 2004. Ajudazols, new inhibitors of the mitochondrial electron transport from *Chondromyces crocatus*. Production, antimicrobial activity and mechanism of action. J Antibiot **57:**151–155.
46. **Rachid S, Krug D, Kunze B, Kochems I, Scharfe M, Zabriskie TM, Blöcker H, Müller R.** 2006. Molecular and biochemical studies of chondramide formation—highly cytotoxic natural products from *Chondromyces crocatus* Cm c5. Chem Biol **13:**667–681. http://dx.doi.org/10.1016/j.chembiol.2006.06.002.
47. **Kunze B, Jansen R, Sasse F, Höfle G, Reichenbach H.** 1995. Chondramides A ~ D, new antifungal and cytostatic depsipeptides from *Chondromyces crocatus* (myxobacteria). Production, physico-chemical and biological properties. J Antibiot **48:**1262–1266.
48. **Holzinger A, Lütz-Meindl U.** 2001. Chondramides, novel cyclodepsipeptides from myxobacteria, influence cell development and induce actin filament polymerization in the green alga *Micrasterias*. Cell Motil Cytoskeleton **48:**87–95. http://dx.doi.org/10.1002/1097-0169(200102)48:2<87::AID-CM1000>3.0.CO;2-C.
49. **Sasse F, Kunze B, Gronewold TM, Reichenbach H.** 1998. The chondramides: cytostatic agents from myxobacteria acting on the actin cytoskeleton. J Natl Cancer Inst **90:**1559–1563. http://dx.doi.org/10.1093/jnci/90.20.1559.
50. **Rachid S, Scharfe M, Blöcker H, Weissman KJ, Müller R.** 2009. Unusual chemistry in the biosynthesis of the antibiotic chondrochlorens. Chem Biol **16:**70–81. http://dx.doi.org/10.1016/j.chembiol.2008.11.005.
51. **Jansen R, Kunze B, Reichenbach H, Höfle G.** 2003. Chondrochloren A and B, new β-amino styrenes from *Chondromyces crocatus* (myxobacteria). Eur J Org Chem **2003:**2684–2689. http://dx.doi.org/10.1002/ejoc.200200699.
52. **Buntin K, Irschik H, Weissman KJ, Luxenburger E, Blöcker H, Müller R.** 2010. Biosynthesis of thuggacins in myxobacteria: comparative cluster analysis reveals basis for natural product structural diversity. Chem Biol **17:**342–356. http://dx.doi.org/10.1016/j.chembiol.2010.02.013.
53. **Irschik H, Reichenbach H, Höfle G, Jansen R.** 2007. The thuggacins, novel antibacterial macrolides from *Sorangium cellulosum* acting against selected Gram-positive bacteria. J Antibiot **60:**733–738. http://dx.doi.org/10.1038/ja.2007.95.
54. **Viehrig K, Surup F, Harmrolfs K, Jansen R, Kunze B, Müller R.** 2013. Concerted action of P450 plus helper protein to form the amino-hydroxy-piperidone moiety of the potent protease inhibitor crocapeptin. J Am Chem Soc **135:**16885–16894. http://dx.doi.org/10.1021/ja4047153.
55. **Cane DE, Walsh CT.** 1999. The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. Chem Biol **6:**R319–R325. http://dx.doi.org/10.1016/S1074-5521(00)80001-0.
56. **Nett M.** 2014. Genome mining: concept and strategies for natural product discovery. Prog Chem Org Nat Prod **99:**199–245. http://dx.doi.org/10.1007/978-3-319-04900-7_4.
57. **Bachmann BO, Lanen SGV, Baltz RH.** 2014. Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? J Ind Microbiol Biotechnol **41:**175–184. http://dx.doi.org/10.1007/s10295-013-1389-9.
58. **Koehn FE, Carter GT.** 2005. The evolving role of natural products in drug discovery. Nat Rev Drug Discov **4:**206–220. http://dx.doi.org/10.1038/nrd1657.
59. **Silakowski B, Kunze B, Nordsiek G, Blöcker H, Höfle G, Müller R.** 2000. The myxochelin iron transport regulon of the myxobacterium *Stigmatella aurantiaca* Sg a15. Eur J Biochem **267:**6476–6485. http://dx.doi.org/10.1046/j.1432-1327.2000.01740.x.
60. **Baltz RH.** 2014. MbtH homology codes to identify gifted microbes for genome mining. J Ind Microbiol Biotechnol **41:**357–369. http://dx.doi.org/10.1007/s10295-013-1360-9.
61. **Sunbul M, Zhang K, Yin J.** 2009. Using phosphopantetheinyl transferases for enzyme posttranslational activation, site specific protein labeling and identification of natural product biosynthetic gene clusters from bacterial genomes. Methods Enzymol **458:**255–275. http://dx.doi.org/10.1016/S0076-6879(09)04810-1.
62. **Croft LJ, Lercher MJ, Gagen MJ, Mattick JS.** 2003. Is prokaryotic complexity limited by accelerated growth in regulatory overhead? Genome Biol **5:**P2. http://dx.doi.org/10.1186/gb-2003-5-1-p2.
63. **Thaxter R.** 1892. On the *Myxobacteriaceae*, a new order of Schizomycetes. Bot Gaz **17:**389–406.
64. **Reiter W-D, Palm P, Yeats S.** 1989. Transfer RNA genes frequently serve as integration sites for prokaryotic genetic elements. Nucleic Acids Res **17:**1907–1914. http://dx.doi.org/10.1093/nar/17.5.1907.
65. **Müller S, Shen H, Hofmann D, Schairer HU, Kirby JR.** 2006. Integration into the phage attachment site, *attB*, impairs multicellular differentiation in *Stigmatella aurantiaca*. J Bacteriol **188:**1701–1709. http://dx.doi.org/10.1128/JB.188.5.1701-1709.2006.
66. **Krug D, Müller R.** 2009. Discovery of additional members of the tyrosine aminomutase enzyme family and the mutational analysis of CmdF. Chembiochem **10:**741–750. http://dx.doi.org/10.1002/cbic.200800748.
67. **Rachid S, Krug D, Weissman KJ, Müller R.** 2007. Biosynthesis of (*R*)-beta-tyrosine and its incorporation into the highly cytotoxic chondramides produced by *Chondromyces crocatus*. J Biol Chem **282:**21810–21817. http://dx.doi.org/10.1074/jbc.M703439200.
68. **Buedenbender S, Rachid S, Müller R, Schulz GE.** 2009. Structure and action of the myxobacterial chondrochloren halogenase CndH: a new

variant of FAD-dependent halogenases. J Mol Biol **385:**520–530. http://dx.doi.org/10.1016/j.jmb.2008.10.057.
69. **Rachid S, Revermann O, Dauth C, Kazmaier U, Müller R.** 2010. Characterization of a novel type of oxidative decarboxylase involved in the biosynthesis of the styryl moiety of chondrochloren from an acylated tyrosine. J Biol Chem **285:**12482–12489. http://dx.doi.org/10.1074/jbc.M109.079707.
70. **Buntin K, Weissman KJ, Müller R.** 2010. An unusual thioesterase promotes isochromanone ring formation in ajudazol biosynthesis. Chembiochem **11:**1137–1146. http://dx.doi.org/10.1002/cbic.200900712.
71. **Scott VR, Boehme R, Matthews TR.** 1988. New class of antifungal agents: jasplakinolide, a cyclodepsipeptide from the marine sponge, *Jaspis* species. Antimicrob Agents Chemother **32:**1154–1157. http://dx.doi.org/10.1128/AAC.32.8.1154.
72. **Martin C, Oberer L, Ino T, König WA, Busch M, Weckesser J.** 1993. Cyanopeptolins, new depsipeptides from the cyanobacterium *Microcystis* sp. PCC 7806. J Antibiot **46:**1550–1556.
73. **Shimkets LJ, Dworkin M, Reichenbach H.** 2006. The myxobacteria, p 31–115. *In* Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E (ed), The prokaryotes. Springer, New York, NY.