Deletion of the *moe*A gene in *Flavobacterium* IR1 drives structural color shift from green to blue and alters polysaccharide metabolism

Álvaro Escobar Doncel[1,2], Constantinos Patinios[3], Alexandre Campos[4], Maria Beatriz Walter Costa[2], Maria V. Turkina[5], Maria Murace[6], Raymond H.J. Staals[3], Silvia Vignolini[6,7], Bas E. Dutilh[2,8], Colin J. Ingham[7]*

Affiliations

1. Hoekmine BV, Verenigingstraat 36, 3515 GJ, Utrecht, the Netherlands
2. Institute of Biodiversity, Faculty of Biological Sciences, Cluster of Excellence Balance of the Microverse, Friedrich Schiller University Jena, Rosalind Franklin Straße 1, D-07745 Jena, Germany
3. Laboratory of Microbiology, Wageningen University and Research, Wageningen, The Netherlands
4. CIIMAR, Interdisciplinary Centre of Marine and Environmental Research, Terminal de Cruzeiros do Porto de Leixões, Av. General Norton de Matos s/n, 4450-208 Matosinhos, Portugal
5. Department of Biomedical and Clinical Sciences, Faculty of Medicine and Health Sciences, Linköping University, SE-581 85 Linköping, Sweden
6. Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom
7. Max Planck Institute of Colloids and Interfaces, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany
8. Theoretical Biology and Bioinformatics, Science4Life, Utrecht University, Padualaan 8, 3584 CH Utrecht, the Netherlands

* Corresponding author

ABSTRACT

Structural colors (SC), generated by light interacting with nanostructured materials, is responsible for the brightest and most vivid coloration in nature. Despite being widespread within the tree of life, there is little knowledge of the genes involved. Partial exceptions are some *Flavobacteriia* in which genes involved in a number of pathways, including gliding motility and polysaccharide metabolism, have been linked to SC. A previous genomic analysis of SC and non-SC bacteria suggested that the pterin pathway is involved in the organization of bacteria to form SC. Thus here, we focus on the *moe*A molybdopterin molybdenum transferase. When this gene was deleted from *Flavobacterium* IR1, the knock-out mutant showed a strong blue shift in SC of the colony, different from the green SC of the wild-type. The *moe*A mutant showed a particularly strong blue shift when grown on kappa-carrageenan and was upregulated for starch degradation. To further analyze the molecular changes, proteomic analysis was performed, showing the upregulation of various polysaccharide utilization loci, which supported the link between *moe*A and polysaccharide metabolism in SC. Overall, we demonstrated that single-gene mutations could change the optical properties of bacterial SC, which is unprecedented when compared to multicellular organisms where structural color is the result of several genes and can not yet be addressed genetically.

MAIN TEXT

INTRODUCTION

Structural color (SC) is the result of the interaction of light with nanoscale structures, causing selective, angle-dependent light reflectance, an optical mechanism distinct from pigmentation which is a property of differential light reflection in molecules. This phenomenon can have a bright, metallic and iridescent appearance, where the color seen is often highly dependent on viewing and illumination angles. SC has been reported in many eukaryotes, including vertebrates, invertebrates, plants, and *Myxomycota*, as well as in bacteria, but not in *Eumycota* or *Archaea* (Brodie et al., 2021). Among bacteria, SC from colonies of the phylum Bacteroidetes is the best characterized (Kientz et al., 2012b; Kientz et al.,, 2016, Johansen et al., 2018). SC in bacteria results from the periodic organization of the rod-shaped cells packed in a regular hexagonal lattice, forming a two-dimensional photonic crystal that reflects light (Scherte et al., 2020). Interestingly, the ecological role of bacterial SC is yet to be determined. Hypotheses point at predation (Hamidjaja et al., 2019) and polysaccharide metabolism optimization (van der Kerkhof et al., 2022), but further research is needed to elucidate its biological significance.

Information on genes and pathways involved in bacterial SC is limited but growing. Transposon mutagenesis suggests the involvement of cellular functions including the

stringent response, plant metabolite modification, carbohydrate metabolism, and Bacteroidetes-specific gliding motility (Johansen et al., 2018). A recent bioinformatic study has shown the possible link of some metabolic pathways to bacterial SC (Zomer et al., 2024). In *Flavobacterium* iridescence species 1 (IR1), SC has been linked to interactions with microalgae, particularly through the metabolism of algal polysaccharides such as kappa-carrageenan and fucoidan (Johansen et al., 2018, van de Kerkhof et al., 2022). IR1's colony organization, which underlies SC, may play a role in interbacterial competition, such as predation, but this has no obvious link to the photonic properties of the bacteria (Hamidjaja et al., 2020).

A bioinformatic analysis of 117 bacterial genomes (87 with SC and 30 without) identified genes potentially involved in SC by comparing gene presence/absence, providing a SC-score. By this method, pterin pathway genes were strongly predicted to be involved in SC (Zomer et al., 2024). Pterins mainly work as enzyme cofactors in various functions, such as aerobic/anaerobic metabolism and detoxification. In eukaryotes, pterins contribute to pigmentary colors, such as in the scale structures of pierid butterfly wings (Wijnen et al., 2007), and appear in insects, fish, amphibians, and reptiles (Daubner et al., 2018). While pigment coloration is different from SC, structurally organized pterins can function as refractive index dopants (Wilts et al., 2016; Sai et al., 2023) and function in UV protection, phototaxis, and intracellular signaling (Feirer et al., 2017).

We focused on one specific pterin, the molybdenum cofactor (MoCo), due to its predicted involvement in bacterial SC (Zomer et al., 2024). MoCo is a cofactor in a group of enzymes known as molybdoenzymes which are key enzymes in nitrogen, purine, and sulfur metabolism. These enzymes in bacteria fall into three families: xanthine oxidases, dimethyl sulfoxide reductases, and sulphite oxidases (Wootton et al., 1991; Zhang and Gladyshev 2008). To study the link between MoCo and SC, we use IR1 as a model organism for bacterial SC due to the availability of genome engineering tools and its intense coloration (Johansen et al., 2018; Patinios et al., 2021). Using the SIBR-Cas (Self-splicing Intron-Based Riboswitch-Cas) genome engineering tool (Patinios et al., 2021), we deleted the molybdopterin molybdenum transferase *moe*A gene, one of the most important genes for predicting bacterial SC (Zomer et al., 2024), as its protein is crucial in the final MoCo pathway reaction.

MATERIALS AND METHODS

Bioinformatics analysis of the molybdopterin pathway operon in *Flavobacterium* IR1. Synteny and homology of the proteins related to SC were visualized with gggenomes 1.0.0 (Hackl et al., 2024) in RStudio 1.1.456. First, sequences of genomes and SC proteins were obtained from a previous work (Zomer et al., 2024). Proteins were predicted in the genomes with Prodigal 2.6.3 (Hyatt et al., 2010). Proteins of interest were matched with BLAST

2.14.0+ blastp (Altschul et al., 1990) against Prodigal's predicted proteins to find genomic coordinates. Operon start coordinate matches the start of the first gene of the putative operon and operon end coordinate matches the end of the last gene of the putative operon. Python 3.12.4 and jupyter notebook 7.2.1 were used to adapt file formats and create objects compatible to gggenomes. The corresponding phylogenetic tree was made from aligned 16S rRNA genes using Barrnap 0.9 (Seemann 2024), BEDtools 2.31.0 getfasta (Quinlan et al 2010), MAFFT 7.505 (Katoh et al., 2013), iqtree v1.6.2 (Minh et al., 2020) and iToL online v6 (Letunic and Bork, 2024). The final figure containing synteny, homology, and the tree was done in Inkscape 1.3.2. The tutorial and scripts for reproducing the figure were stored in a GitHub repository: https://github.com/MGXlab/genes_synteny. Tools were used with their default parameters and exact commands can be found in the GitHub repository.

Bacterial strains and growth conditions. Bacteria strains used in this study are described in Table S1. *Flavobacterium* iridescence species 1 (IR1) was the target strain used in this project. IR1 was grown in Artificial Sea Water (ASW) medium composed of 5g·L^{-1} peptone (Sigma-Aldrich), 1g·L^{-1} yeast extract (Sigma-Aldrich), and 10g·L^{-1} sea salt (Lima), at 25°C and grown in an orbital incubator at 200rpm (Johansen et al., 2018). *Escherichia coli* DH5α (New England Biolab, NEB) was used for general plasmid propagation and standard molecular techniques. *E. coli* was grown in Luria-Bertani (LB) medium composed of 10g·L^{-1} tryptone (Sigma-Aldrich), 5g·L^{-1} yeast extract, and 10g·L^{-1} NaCl (Sigma-Aldrich), at 37°C shaken at 200rpm. IR1 was plated on ASW with 1% agar (Invitrogen) with or without 0.25g·L^{-1} nigrosine (Sigma-Aldrich) (Johansen et al., 2018). *E. coli* was plated on LB medium containing 1.5% agar (Invitrogen). Media were supplemented with 50µg·mL^{-1} spectinomycin (Sigma-Aldrich), 100µg·mL^{-1} ampicillin or 200µg·mL^{-1} erythromycin (Sigma-Aldrich) when necessary. All the strains were stored in 25% glycerol solution at -80°C.

Plasmid construction. All the plasmids used for SIBR-based gene knockout (KO) were constructed from pSIBR048 (Table S2) following the previously described protocol by Patinios and coworkers (Patinios et al., 2021). In brief, to introduce the *moe*A homologous arms (HA) and mediate the deletion of *moe*A, pSIBR048 was linearized using MluI (NEB) and the phosphorylated ends were removed using Shrimp Alkaline Phosphatase (NEB). 1500 bp HA corresponding upstream and downstream of *moe*A were amplified from the IR1 genome by PCR with Dream Taq DNA Polymerase (Thermo Fisher). The amplicons were resolved on 1% agarose (Eurogentec) electrophoresis gel and purified using GenElute PCR Clean-Up Kit (Sigma-Aldrich). The PCR products were introduced to the linearized pSIBR048 using NEBuilder HiFi DNA Assembly Master Mix (NEB), resulting in the pMoeA_NT. Following, the *moe*A targeting spacer was introduced in the pMoeA_S1 plasmid as previously described (Patinios et al., 2023). The DNA sequence of each newly created plasmid was verified by Sanger sequencing. Oligonucleotides used in this study are listed in Table S3.

***E.* coli DH5α competent cell preparation and transformation.** Competent cells of *E. coli* DH5α, for chemical transformation, were prepared following the $CaCl_2$ method described by Sambrook (Sambrook et al., 1989). The cells were aliquoted ready to be used or stored at -80ºC. The transformation of the competent DH5α cells was done by heat-shock following the High Efficiency Transformation Protocol of NEB. For this protocol, LB medium was used instead of SOC medium. The cells were plated on LB 1.5% agar supplemented with 50µg·mL^{-1} spectinomycin and incubated at 37ºC for 1 day.

IR1 competent cell preparation, transformation, and SIBR-Cas genetic engineering assay. The methods used for the preparation of the electro-competent cells of IR1, transformation with plasmids and SIBR-Cas genetic engineering were as previously described (Patinios et al., 2021). Mutant colonies were identified through colony PCR using primers cFwd *moe*A and cRev *moe*A, and Sanger sequencing (Eurofins).

Effects of nutrient composition on SC in IR1 WT and *Δmoe*A. The visual phenotype of the mutant in comparison to the WT was first checked on agar plates under different nutrient conditions. ASWB agar contains ASW medium with 1% agar and 0.25g·L^{-1} nigrosine (Johansen et al., 2018). ASWB low nutrient medium (ASWBLow) contains the same nutrients as ASWB but without peptone (Johansen et al., 2018). Minimal medium (MM) contains 0.5% sea salt, 0.1% $MgSO_4$, 0.25% kappa-carrageenan (Special Ingredients) and 1% agar. ASWB kappa-carrageenan (ASWBKC) contains the same nutrients as ASWB, but with kappa-carrageenan instead of agar (ASWBC modified from Johansen et al., 2018). ASWB fucoidan (ASWBF) contains the same nutrients as ASWB plus 1% fucoidan (Absonutrix) (Johansen et al., 2018). ASWB starch (ASWBS) contains the same nutrients as ASWB plus 1% starch (Sigma-Aldrich) (Johansen et al., 2018). Before studying the effects of the nutrient composition, both strains were cultivated overnight at 25ºC on an ASWB plate from which some bacterial biomass was collected, resuspended in 1% sea salt and 10µL of the bacteria suspension was spotted on the plates. The mutant was observed after 2 days by eye to check the display of SC.

Imaging. Photographs of colonies were taken with a Canon digital camera equipped with a RF 100 mm macro lens or using a KEYENCE VHX-7000 Digital Microscope using defined angles of illumination and data capture (Figure 2B).

Determining colony spread. Colonies of IR1 WT and Δ*moeA* were grown as a spot on ASWB, ASWBKC, ASWBF, ASWBS, ASWBLow, and MM for 6 days at 25ºC. The diameter of the colonies was measured at two time points, just after the spot was inoculated and after 6 days. These data were measured in triplicates for each condition and strain.

Angle-resolved spectroscopy (goniometry). The optical properties of the bacteria colonies were studied following the method previously described (Johansen et al., 2018).

Angle-dependent reflectance spectra were measured using a custom-built goniometer setup (Vignolini et al., 2013) both in scattering and specular configuration. The samples were illuminated from a fixed direction by a Xenon lamp (Ocean Optics HPX-2000), and the reflected light was collected at different detection angles (resolution 1°) using a rotating arm connected to a spectrometer (Avantes HS2048) via an optical fiber. Data presented in this work were normalized against a white diffuser (Labsphere SRS-99-010).

Analysis of the optical response. Angle-resolved reflectance spectra show peculiar features caused by the two-dimensional structural organization. In scattering configuration, diffraction spots are visible that can be correlated to the diffraction grating formed by the bacteria on the surface (Schertel 2020, Johansen 2018). More specifically, the angles of constructive interference from a diffraction grating can be expressed by the grating equation,

$$\theta_m = arcsin\left(\frac{m\lambda}{d} - sin\theta_i\right) \qquad [1]$$

where m ∈ [0, − 1, + 1, − 2, + 2, …] is the diffraction order, λ is the wavelength of light, d is the period of the structure, θ_i is the angle of incidence and θ_m is the reflection angle for a given order. This equation can be used to determine the period (d) of the bacteria organization, and deviation from the predicted diffraction spots can quantitatively inform about the degree of disorder compared to an ideal periodic structure. Information on the effective refractive index can be obtained from goniometry data acquired in specular configuration. In this case, reflectance peaks arise from the constructive interference of light with the multilayer structure and depend on various parameters. Considering both Bragg's law and Snell's law, the peak reflection wavelength λ_B and corresponding incident angles θ_{in} at which constructive interference occur are linked via the following equation:

$$\lambda_B = 2n_{avg}d\ cos(arcsin\left(\frac{sin\theta_{in}}{n_{avg}}\right)) \qquad [2]$$

where θ_{in} is the illumination angle and n_{avg} is the volume average effective refractive index of the total material composite in the photonic crystal. For construction, the angle of observation θ_{out} equals θ_{in}.

Intracellular and extracellular proteome sample preparation. WT IR1 and the *moe*A mutant were selected for intracellular and extracellular proteomics analysis. Cells were grown for 2 days at 25°C completely covering ASWBKC plates. To prepare the whole cell fractions, cultures were harvested and centrifuged at 12,000 rpm for 15 min at 4°C in 2mL tubes. Cells were washed with 1% KCl solution, centrifuged at 12,000rpm for 15min at 4°C and cell pellets were stored at -80°C. For preparation of extracellular protein fractions, supernatants were collected after the first cell centrifugation, the supernatants were

transferred into new 2 mL tubes, and centrifuged at 12,000rpm for 25 min at 4°C. To ensure reproducibility, both preparations were performed in biological triplicates.

Peptides originating from IR1 intracellular and extracellular proteins were extracted according to the protocol described by Campos and coworkers (Campos et al., 2015, 2016). The resulting dried peptides were resuspended in 0.1% formic acid in deionized water followed by bath-sonication for 5 min and 5 min centrifugation at 12,000rpm at 25°C. Peptide concentration was assessed at A280 using ND-1000 Nanodrop spectrophotometer (Thermo Scientific) peptide concentrations were adjusted to 0.1mg/ml to normalize samples prior to LC-MS/MS analyses.

Proteome sample analysis. For the LC-MS/MS analyses, peptides were separated by EASY-nLC II system (Thermo Scientific) at flow rate of 300nl/min on a precolumn (Acclaim PepMap 100, 75µm × 2cm, Thermo Scientific) followed byEASY-Spray C18 reversed-phase nano LC column (PepMap RSLC C18, 2µm, 100A 75µm × 25cm, Thermo Scientific) thermostated at 55°C. A 90 min gradient of 0.1% formic acid in water (A) and 0.1% formic acid in 80% acetonitrile (B) was distributed as follows: from 6% B to 30% B in 65min; from 30% B to 100% B in 20min and hold at 100% B for 5min. Automated online analyses were performed in positive ionization mode by a Q Exactive HF mass spectrometer (Thermo Scientific) equipped with a nano-electrospray. Full scans were performed at resolution 120,000 in a range of 380–1,400 m/z and the top 15 most intense multiple charged ions were isolated (1.2m/z isolation window) and fragmented at a resolution of 30,000 with a dynamic exclusion of 30s. The generated raw files were analyzed using Sequest HT in Proteome Discoverer software (Thermo Fisher Scientific, San Jose, CA, USA, CS version 2.5.0.400). *Flavobacterium* (NCBI Taxonomy ID 2026304) protein sequence database used for protein identification was acquired from NCBI (https://www.ncbi.nlm.nih.gov/; downloaded on 10th of February 2023; 5468 entries. The following search parameters were used: trypsin as a digestion enzyme; maximum number of missed cleavages 2; fragment ion mass tolerance 0.08Da; parent ion mass tolerance 10ppm; carbamidomethylation of cysteine as fixed modification and methionine oxidation as variable modifications.

Proteome bioinformatics. Scaffold (version Scaffold_5.3.0, Proteome Software Inc., Portland, OR) was used to validate protein identifications and for relative quantification of proteins. Peptide identifications were accepted if they could be established at greater than 90% probability by the Scaffold Local FDR algorithm. Protein identifications were considered correct if they could be established at a greater than 95% probability and contained at least 1 unambiguously identified peptide. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al., 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into

clusters. These clusters were associated to a specific protein of IR1 within the GenBank database giving the following default identity name: PAM9XXXX.

Proteome data analysis. The quantitative protein abundance levels were analyzed in the proteins that had a difference between the sample groups when applying the Student's t-test, using the multiple test correction Benjamin-Hochberg, and a cut-off p-value lower than 0.01 was chosen for statistically significant quantitative difference in relative proteins amount between *moe*A and WT sample groups. A protein was considered downregulated when the log2 of the fold change (Δ*moe*A/WT) was lower than -1, and upregulated when was higher than 1.

The identified differentially expressed proteins were bioinformatically analyzed using the KEGG tool BlastKOALA for functional characterization and the InterProScan software (Kanehisa et al., 2016; Jones et al., 2014). The proteins identified in extracellular fractions were also analyzed using SecretomeP (identifies signal independent secreted proteins) and SignalP (predict signal peptides) software to confirm that they were predicted to be potentially secreted and to exclude possible contamination by intracellular proteins (Bendtsen et al., 2005; Teufel et al., 2023).

Monitoring starch degradation by iodine staining assay. Colonies of IR1 WT and Δ*moeA* were grown on plates with ASWS (ASWBS without nigrosin) for 2 days at 25ºC. Iodine crystals were deposited on the lid of the plate and incubated upside down overnight to expose the agar to the iodine vapor. The plates were checked for starch degradation which corresponds to the zones of clearing , with dark, stained areas indicating presence of undegraded starch (Kasana and Salwan, 2008). These were measured in triplicates for each condition and strain.

RESULTS

Bioinformatic analysis of the molybdopterin operon

We reanalyzed a recent bioinformatic analysis on SC to specifically investigate genes involved in molybdopterin cofactor (MoCo) synthesis (Zomer et al., 2024). MoCo genes are typically clustered in SC bacteria and are consecutively encoded on the IR1 genome (Figure 1A), probably forming an operon comprising molybdopterin molybdenum transferase (*moe*A), molybdenum cofactor guanylyl transferase (*mob*A), uroporphyrinogen-III C-methyltransferase (*sum*T), molybdopterin synthase sulfur carrier unit (*moa*D), adenylyl transferase/sulfur transferase (*moe*Z), molybdopterin synthetase catalytic unit (*moa*E), cyclic pyranopterin monophosphate synthase 2 (*moa*C2), and GTP 3',8-cyclase (*moa*A). In 117 bacterial genomes (87 SC and 30 non-SC) analyzed (Table 1), most bacteria showing SC contained all these genes, except *mob*A and *moa*D. Meanwhile, in non-SC bacteria,

these genes appeared less frequently. Overall, 61 of 87 SC genomes had a complete MoCo pathway, 10 lacked one gene, and 16 lacked two. Conversely, only 6 of 30 non-SC genomes had a full pathway, while others showed partial gene loss, with 6 missing the entire pathway.

The genetic structure of this putative operon for molybdopterin synthesis was compared across 8 strains with variable SC (Figure 1BC). Using the SC classifier, these strains were scored for SC based on the presence/absence of specific genes in their genomes (Zomer et al., 2024), revealing that the predictions were consistent with our experimental results SC (Figure 1D).

Synteny analysis of selected genomes revealed the organization of MoCo synthesis genes. IR1 contains a putative MoCo synthesis operon consisting of *moe*A, *mob*A, *sum*T, *moa*D, *moe*Z, *moa*E, *moa*C2, and *moa*A, and an additional *sum*T homolog. UW101 has a similar operon without the *sumT* duplication (Figure 1C). Other strains show different gene orders, loci, or variations like missing or duplicated genes. Notably, *Flavobacteriaceae* strains DSM15718 and DD5b, which only contain *sumT*, and HM20, lacking *sumT* but retaining most MoCo genes, do not exhibit SC. Thus, while the MoCo synthesis pathway is crucial for SC, its structure and organization vary among SC strains and are not the sole determinants of SC.

Phenotyping the Δ*moe*A mutant

To study the role of *moe*A in SC, we generated a clean knock-out (KO) of *moe*A in IR1 using the SIBR-Cas tool (Patinios et al., 2021). After successfully deleting *moe*A, we compared the colors of the Δ*moe*A colonies with those of the wild-type (WT) strain under three nutrient conditions: (1) ASWB, a standard peptone/yeast extract medium; (2) ASWBLow, a low-nutrient medium with yeast extract as the sole nutrient; and (3) minimal medium (MM), with the minimum nutrients required for IR1 growth.

On ASWB agar plates, the WT strain colony showed a vivid brilliant green SC with a red ring, while the Δ*moe*A colony displayed a dull green-blue SC with a blue ring. On ASWBLow, the WT's SC shifted to a shiny green-yellow-orange color, whereas Δ*moe*A displayed a dull green center with an intense green ring. On MM, both the WT and Δ*moe*A had weaker SC than when grown on higher nutrient media, showed dispersed clusters of cells, and maintained their green and blue hues, respectively. Additionally, Δ*moe*A colonies spread more slowly than the WT under all conditions evaluated. In summary, deleting *moe*A produced a general SC shift from green to blue, and a reduction of colony spreading (Figure 2A).

The optical properties of the Δ*moe*A colony were checked by growing as a spot on ASWB plates and observing its color from different angles to capture the full optical response of its

photonic structure. When photographed from directly above the light source at position (X,Y,Z coordinates 0,-1,1.1 respectively), Δ*moe*A displayed a primarily green SC (Figure 3A), albeit duller than when photographed from positions (1,-1,1) (Figure 3B), and (1,-1,0.36) (Figure 3C). From the positions in Figure 3BC, two distinct colored rings were visible: an inner blue ring and an outer green-yellow ring. Although, when they were photographed from positions (1,0,0.36) (Figure 3D), and (1,0,0.18) (Figure 3E), the SC shifted to predominantly blue, with a highly reflective blue ring and a green ring. SC was lost when photographed from position (0.58,1,1), displaying a gray-brown color (Figure 3F). Overall, we confirmed the angle-dependency of the SC in the Δ*moeA*, showing variations in color and intensity with changes in viewing angle.

When IR1 was grown with the polysaccharides fucoidan (from brown algae), or kappa-carrageenan (from red algae), its SC shifted to dark purple and shinier green, respectively (van de Kerkhof et al., 2022). To investigate how nutrient supply affects SC in IR1 WT and Δ*moe*A strains, they were grown as spot on ASW medium gelled with kappa-carrageenan instead of agar (ASWBKC), fucoidan and agar (ASWBF), or starch and agar (ASWBS). On ASWBKC plates, both strains exhibited more intense SC than on ASWB, with Δ*moe*A displaying a brilliant, blue shifted color compared to the WT's structural green. The WT strain also displayed a dark green ring, and a thin red outer ring as observed in ASWB (Johansen et al., 2018; Hamidjaja et al., 2020). On ASWBF plates, the WT displayed a dull blue-purple SC, while Δ*moe*A showed a dull green SC with a dull green-yellow ring and a red thin outer ring. On ASWBS, the colonies displayed a mix of colors rather than the mostly monochromatic patterns seen on agar (Figure 2), kappa-carrageenan or fucoidan (Figure 4A). The WT showed a dull green center, a green-yellow ring, and a shiny red outer ring. In contrast, Δ*moe*A displayed a dull blue center with a shiny blue ring, and a shiny green outer ring. Overall, polysaccharides significantly influenced SC, with both strains showing the most intense colors on kappa-carrageenan.

Deletion of the *moe*A gene reduces colony expansion

During the analysis of the colors displayed by IR1 WT and Δ*moe*A, differences in the colony spreading were observed indicating variations in gliding motility. To quantify this, both strains were grown for an extended period, and colony expansion was measured (Figure 4B). The Δ*moe*A showed slower colony expansion, reaching about half the size of the WT in most conditions, except on ASWBF, where colony expansion was similar to the WT. Interestingly, Δ*moe*A colony expansion was faster on ASWBLow, and especially on MM, compared to other conditions, which also happened for the WT. Thus, the lack of nutrients is an enhancer of colony expansion.

The organization and motility of groups of cells at the colony edges were visualized using a digital stereo microscope with full coaxial light. Both strains were grown as a spot on ASWB,

and the colony edges were visualized for 1 hour (Figure 6). IR1 WT showed high motility of the bacterial layers at the edge of the colony, with dispersed cell layers forming 'vortex' patterns (Figure 5, yellow arrows). In contrast, Δ*moe*A exhibited limited motility, with a more tightly packed cell organization and a fine, slow-moving layer at the edge (Figure 5, blue arrows). This suggests that *moe*A deletion significantly impairs cell motility and colony expansion.

Quantification of the optical responses of IR1 WT and Δ*moe*A colonies

IR1 WT and Δ*moe*A grown on ASWBKC plates were studied using an optical goniometer to understand the optical characteristics of their displayed colors. We selected this media due to the uniform, vibrant blue coloration of the Δ*moe*A colony.

The complex optical response of both IR1 strains observed in the heatmaps in Figure 6 can be attributed to a polycrystalline two-dimensional structure with hexagonal packing, as previously described (Schertel et al., 2020). In particular, the specular reflection data (Figure 6AB) allowed us to extrapolate an effective refractive index of 1.38 for both strains, consistent with earlier studies (Schertel et al., 2020). In a diffraction configuration, intense diffraction peaks are observed in the visible range around a detection angle of -30° for wavelengths of 550 nm (green) for IR1 WT colonies (Figure 6C) and 480 nm (blue) for Δ*moe*A (Figure 6D), coherent with the primary colors observed qualitatively in Figure 2A.

In addition, other two bright diffraction spots are present in both cases outside of the visible range. For IR1 WT, such spots are present around 550 nm, 400 nm, and 350 nm; in Δ*moe*A, these diffraction spots shift to a lower wavelength around 480 nm, 350 nm, and 300 nm. By matching the diffraction grating equation with the observed spots (white dashed lines in Figure 6), the inter-bacterial distance can be obtained (Schertel et al., 2020). The periodicity was therefore estimated to be 410 nm for IR1 WT, and 365 nm for Δ*moe*A. This optical analysis aligns with visual observations, confirming the blue shift in Δ*moe*A, and suggested that this change in SC is caused by cells which are still highly ordered but narrower.

Changes in the proteome due to the deletion of *moe*A

To further investigate the effects of *moe*A deletion, we performed a characterization and quantitative comparison of cellular (Figure 7A) and extracellular (Figure 7B) proteomes of IR1 WT and Δ*moe*A strain using a mass spectrometry-based proteomic approach. We identified 203 intracellular proteins that significantly changed their abundance upon deletion of *moe*A (Table S4 and S5), and 268 differentially abundant extracellular proteins (Table S6 and S7). The following pathway analysis provided insight into how these proteins might be related to SC.

Peptides derived from molybdopterin molybdenum transferase, encoded by *moe*A, were only detected in the WT strain, confirming a successful knock out in Δ*moe*A. The intra- and extracellular proteome analysis showed some differentially expressed proteins involved in the MoCo pathway or containing molybdopterin-binding motif. The deletion of *moe*A produced different regulatory effects on the peptides encoded from the genes within its putative operon. The proteins encoded from *moa*A, *moa*C2, and *mob*A were upregulated in the mutant, while from *moa*E and *moe*Z were unaffected, and from *sum*T and *moa*D were undetected in both strains (Figure 8). Proteins with a molybdopterin-binding motif were differentially expressed. The downregulated proteins included xanthine dehydrogenase *yag*S and *yag*R (involved in purine catabolism), an alanine dehydrogenase involved (amino acid biosynthesis), and a nitrite reductase (nitrogen assimilation) (Table S4). An upregulated protein was NAD(P)H-nitrite reductase, also involved in nitrogen assimilation (Table S5).

Of the 5,471 known proteins in IR1, 58.1% (3,181 proteins) intracellular proteins were identified, 10.2% (324) showed significant differences ($p<0.01$), with 34.3% (111) considered downregulated, and 27.2% (88) upregulated in the Δ*moe*A (Figure 7A). The downregulated subset included 29 hypothetical proteins, while the upregulated subset had 31.

Downregulated intracellular proteins were involved in amino acid metabolism (10), RNA processing (10), transport (9), DNA transcription (8), translation (6), fatty acid metabolism (4), antimicrobial resistance (4), nucleotide metabolism (4), cofactor biosynthesis (4), proteolysis (4), biofilm formation (3), homeostasis (3), carbohydrate metabolism (3), lipid metabolism (3), phospholipid transformation (2), stress response (2), transport (2), and various metabolic processes (Table S4). Additionally, 28 proteins with unknown functions were identified (Table S4). Some downregulated proteins, such as an ABC transporter ATP-binding protein and a membrane assembly protein (involved in phospholipid transformation), as well as an alanine dehydrogenase (amino acid metabolism), and some hypothetical proteins with unknown role, were completely repressed (found only in the WT).

Upregulated intracellular proteins were involved in transport (12), non-ribosomal peptide synthesis (11), stress response (6), carbohydrate metabolism (5), proteolysis (4), signaling (4), electron transport (3), glycosylation (3), DNA repair (3), amino acid metabolism (2), antibiotic resistance (2), cofactor biosynthesis (2), metabolism (2), RNA processing (2), and various metabolic processes (Table S5). Additionally, 20 proteins with unknown roles were identified (Table S5). Notably, among the most upregulated proteins, we observed a 23S rRNA (adenine(1618)-N(6))-methyltransferase (involved in RNA processing), a hypothetical protein (unknown role), a chalcone isomerase (stress response), an aminopeptidase (proteolysis), and a transcriptional regulator (regulation of DNA transcription).

Of the total known proteins in IR1, 27.5% (1,504 proteins) extracellular proteins were identified, 60.4% (909) were statistically significant ($p<0.01$), with 20.5% (186) considered downregulated, and 20% (182) upregulated in Δ*moeA* (Figure 7B). The downregulated subset included 44 hypothetical proteins, while the upregulated subset had 70. Although fewer proteins were identified in the extracellular space compared to the intracellular space, a higher proportion were statistically significant and differentially regulated.

Analysis of downregulated proteins using SecretomeP showed that 29.6% (55) were likely secreted through a non-classical way, lacking typical secretion sequence motifs in their N-terminus. Additionally, SignalP analysis revealed that 31.7% (59) had a putative signal peptide, suggesting they are Sec (general secretory pathway) substrates and likely to be secreted. The downregulated proteins likely to be secreted (69) included those involved in carbohydrate metabolism (7), transport (7), stress response (4), antibiotic resistance (3), lipopolysaccharide assembly (3), protein modification (3), electron transport (2), lipid metabolism (2), purine metabolism (2), motility (2), and several other functions (Table S6). Additionally, 27 proteins with unknown roles were identified (Table S6). Notably, among the most highly downregulation proteins included a flagellin biosynthesis protein (involved in motility), probably misannotated as the pathways for flagella synthesis are absent in *Flavobacterium* IR1, a murein hydrolase activator (cell division), a hypothetical protein (lipopolysaccharide assembly), a peptidylprolyl isomerase (protein modification), and an azurin (electron transport).

Analysis of upregulated proteins using SecretomeP revealed that 47.3% (86) potentially follow a non-classical secretion pathway. SignalP analysis indicated that 54,4% (99) of the upregulated proteins possessed a signal peptide. The upregulated proteins likely to be secreted (109) included those involved in transport (19), carbohydrate metabolism (18), proteolysis (12), stress response (5), fatty acid metabolism (4), cell division (1), electron transport (1), iron acquisition (1), motility (1), protein modification (1), and signaling (1) (Table S7). Additionally, 45 proteins with unknown roles were identified (Table S7). Notably, the most upregulated proteins included two hypothetical proteins (involved in unknown roles), a hypothetical protein (cell division), an acetyl-CoA carboxylase biotin carboxyl carrier protein (fatty acid biosynthesis), a glycoside hydrolase (carbohydrate metabolism), and a hypothetical protein (transport).

The combination of protein analysis and genomic data from the IR1 genome provided insights into the putative operons or gene clusters affected by the deletion of *moe*A (Figure 8). Intracellular proteomic analysis suggested the downregulation of putative operons associated with antimicrobial drug resistance, fatty acid biosynthesis, purine catabolism, and phospholipid transformation. Conversely, putative operons involved in respiratory electron transport, carbohydrate metabolism, non-ribosomal peptide synthesis, antioxidant stress, and cell wall synthesis were upregulated. In the extracellular proteomic analysis, a putative

operon with an unknown function was downregulated, while putative operons involved in fatty acid biosynthesis, carbohydrate metabolism, and unknown functions were upregulated. Notably, the deletion of *moe*A created a cascade of regulation effects that affected pathways not previously linked to molybdopterin synthesis.

Previous studies, alongside the results of this investigation, have shown the importance of complex polysaccharides degradation in the development of SC (Johansen et al., 2018; van de Kerkhof et al., 2022). In the Bacteroidetes phylum, polysaccharides utilization loci (PUL) operons facilitate the uptake and processing of these polysaccharides. Typically, PUL operons consist of a tandem pair of genes resembling *sus*CD, which encode a transport and substrate-binding complex, and various carbohydrate active enzymes (CAZymes), such as glycosyl hydrolases and pectate lyases (Terrapon et al., 2015).

Our intracellular and extracellular protein analysis revealed the upregulation of three putative PUL operons with similar organization (Figure 8): (1) PAM95095-90, which includes a glycoamilase, a glycosyl hydrolase family 3 (GH3) involved in cellulose degradation, a glycerophosphoryl diester phosphodiesterase, and a GH43 that degrades hemicellulose and pectin polymers (Ara et at., 2020; Mewis et al., 2016); (2) PAM95448-51, which includes an unidentified GH, and a GH35 enzyme that hydrolyzes terminal non-reducing β-D-galactose residues (Tanthanuch et al., 2008); (3) PAM95391-88, which includes a hypothetical protein, and two GH16, one of which was not detected, involved in the degradation of various polysaccharides such as agar and kappa-carrageenan (Viborg et al., 2019). Additionally, other carbohydrate metabolism-related proteins were upregulated in the Δ*moe*A, including a GH18 enzyme involved in chitin degradation (Chen et al., 2020), and a pectate lyase involved in starch degradation (Table S8) (Aspeborg et al., 2012).

***moe*A deletion affects metabolism of complex carbohydrates**

As previously described, the IR1 WT and Δ*moe*A strain were grown on various complex polysaccharides, showing different color phenotypes. The Δ*moe*A colony displayed a strong blue SC phenotype on ASWBKC, a dull green on ASWBF, and a dull blue center with a blue internal ring and green external ring on ASWBS (Figure 4). These results suggest a connection between SC, moeA, and polysaccharide metabolism. Proteins linked to carbohydrate metabolism were also highly regulated, reinforcing this link (Tables S6 and S8). Both strains were grown on ASWS, and starch degradation was visualized using iodine vapor (Kasana and Salwan, 2008). The colonies were photographed from the front and the back (Figure 9). The WT strain showed a duller and smaller starch degradation zone (0.58±0.12 cm) compared to Δ*moe*A (1.17±0.17 cm). In contrast to other media where Δ*moe*A colony expansion was less than WT, the Δ*moe*A showed stronger starch degradation, supporting a role of *moe*A in complex polysaccharides metabolism.

DISCUSSION

SC in biological systems is well-studied optically, but less well understood genetically. This study aimed to expand and deepen the knowledge of genes involved in bacterial SC, focusing on the predicted SC-related gene, *moe*A (Zomer et al., 2024). By deleting *moe*A from the IR1 genome, a model for bacterial SC, we conducted microbiological, optical, proteomic, and comparative genomic analyses of the mutant. The results demonstrated the possibility of engineering SC by targeting specific pathways.

The *moe*A gene is part of the molybdenum cofactor (MoCo) synthesis pathway, which is not exclusive to bacteria, but also found in archaea, animals, and plants, tracing back to the last universal common ancestor (Allen et al., 1994; Weiss et al., 2016). MoCo is essential for molybdoenzymes that catalyze oxo-transfer and hydroxylation reactions such as nitrate reductase, xanthine dehydrogenase, and aldehyde oxidase (Wootton et al., 1991; Zhang and Gladyshev, 2008). In the Δ*moe*A proteome, some molybdoenzymes like xanthine hydrogenases, aldehyde oxidase, and nitrite reductase were downregulated, suggesting their synthesis depends on MoCo availability. However, one nitrite reductase protein (PAM94801) was upregulated, potentially independent of MoCo. Additionally, proteins from *moaA*, *moaC2*, and *mob*A genes which are present in the same operon as *moe*A, were upregulated, possibly to boost molybdopterin availability for MoCo synthesis.

The presence of *moe*A in the genome or the putative operon structure for MoCo pathway alone does not determine a bacterial strain's ability to form SC colonies. For example, the genomes of the Bacteroidetes strains *Flavobacterium* IR1, *F. johnsoniae* UW101, and *C. lytica* HI1, which contain all the genes for the synthesis of MoCo, display SC. Meanwhile, *M. algicola* HM30 and *Z. galactinovorans* DSM1208, lacking *moa*D and *moa*C, respectively, also show SC. Interestingly, the corresponding proteins of *moa*D and *sum*T were not detected in the proteomic analysis of IR1 WT and Δ*moe*A. Additionally, *moa*D and *mob*A were not present in all SC strains. Thus, we concluded that the presence of *moa*C, *moa*D, *mob*A and *sum*T are not essential for SC formation (Zomer et al., 2024).

The predominantly green SC of IR1 WT has been studied using transposon mutagenesis, cultivation, and optical characterization, revealing additional structural colors like yellow, orange, red, blue, and purple (Johansen et al., 2018; van de Kerkhof et al., 2022). Here, *moe*A was deleted from the IR1 genome using SIBR-Cas (Patinios et al., 2021), resulting in a strong blue shift in the colony color, confirmed and quantified by goniometry. The WT and Δ*moe*A colonies show variations in color, color pattern and intensity depending on three conditions: 1) observation angle, displaying green, yellow and blue hues with different intensities; 2) the presence of peptone and yeast extract, affecting color and motility; and 3) the type of polysaccharides present in the media, which significantly altered color and

motility. These findings showed that SC color hue, pattern, and intensity can be modified by genetic engineering, observation angle, and nutrient changes.

Previously, mutations in *trm*D tRNA methyltransferase, and a *clb*B triosephosphate isomerase were described (Johansen et al., 2018). A transposon insertion in *trm*D led to the loss of SC, while preserving growth and motility, and a *clb*B disruption produced a dull green/blue SC. Here, TrmD was downregulated, and ClbB upregulated in the Δ*moe*A proteomics analysis. Deleting *moe*A also caused downregulation of GldL, a protein essential for gliding motility and secretion in *F. johnsoniae* (Shrivastava et al., 2013). The reduced motility in the Δ*moe*A mutant may have resulted from the combined downregulation of *trmD*, GldL, ribosomal proteins, and other uncharacterized proteins. Additionally, the upregulation of ClbB and other regulated proteins may contribute to the SC shift from green to blue.

Polysaccharide metabolism in IR1 has been linked to changes in colony color and motility (van de Kerkhof et al., 2022). Although *moe*A has not been previously linked to polysaccharide degradation (Hasona et al., 1998; Tao et al., 2005; Leimkühler, 2017), its deletion led to the upregulation of proteins from three PUL operons and others involved in polysaccharide metabolism, likely causing the color shift from green (WT) to blue (Δ*moe*A). The identified proteins were involved in degrading cellulose, hemicellulose, pectin, galactose polymers, agar, kappa-carrageenan, mannanose, and starch. The polysaccharide degradation versatility was supported by checking the starch degradation in both strains, with the Δ*moe*A capable of degrading starch faster and more efficiently than WT, producing larger and clearer halos with iodine staining.

On different polysaccharide media, the Δ*moe*A strain showed varied SC and colony expansion patterns: green/blue SC and low colony expansion on agar, intense blue SC and low colony expansion on kappa-carrageenan, dull green SC and low colony expansion on fucoidan, and blue/green SC with higher colony expansion on starch. While reduced motility has been associated with dull or absent SC, and reduced polysaccharide metabolism (Kientz et al., 2012a; Johansen et al., 2018), Δ*moe*A showed reduced motility, but an intense blue SC, and high polysaccharide metabolism. Based on these results, we established a link among polysaccharide metabolism, MoCo biosynthesis, and SC, showing that intense SC is not strictly dependent on motility.

Ecologically, we hypothesize that dense, highly structured bacterial colonies, such as necessary for the SC phenotype, could limit the loss, by diffusion, of the metabolic degradation products of complex polysaccharides. These large macromolecules are often partially hydrolyzed extracellularly because they are too large to pass through bacterial cell membranes. Bacteria secrete enzymes into the surrounding environment to break these polysaccharides down into more easily absorbable monosaccharides or oligosaccharides. The colony structure could create a physical barrier that keeps these products concentrated

and near the cells, allowing the colony to efficiently access and utilize these products, and preventing them from leaking into the surrounding environment. While SC may also yield other ecological benefits associated with growth in biofilms, the highly structured colonies that characterize SC may be more resistant against invasion by competitor species scavenging for degradation products, than an unstructured biofilm. This model is consistent with the observation that SC is associated with polysaccharide metabolism genes, and with the recent observation that SC is mainly localized on surface and interface environments such as air-water interfaces, tidal flats, and marine particles (Zomer et al., 2024).

SC bacteria like *Cellulophaga lytica* (Sullivan et al., 2023) and *Flavobacterium* IR1 (Groutars and Risseeuw, 2022) have been recently studied to be used as colorful biomaterials, making genetic engineering to modify SC a potential next step for developing new colorants. Similar to IR1, *C. lytica* belongs to the *Flavobacteriaceae* family, exhibiting gliding motility, similar SC, and has diverse polysaccharide metabolism genes, though it lacks genetic engineering tools (Kientz et al., 2012a; Kientz et al., 2016; Lysov et al., 2022). Genetic engineering SC in IR1 opens the way to synthetic biology of SC and its application in biomaterials, offering a sustainable alternative to traditional pigments.

CONCLUSIONS

Our results demonstrate the involvement of bioinformatically predicted genes in bacterial SC and suggested that such genes could be targeted to modify the optical characteristics of SC colonies. The simple deletion of one gene, *moe*A, shifted the SC of IR1 colony from green to blue, while nutrient and polysaccharide availability emerged as key factors affecting SC color and motility. Proteomics analysis revealed polysaccharide metabolism as a driver of SC hue changes, hinting at its ecological significance. Additionally, several uncharacterized proteins were differentially expressed, providing exciting new leads for further exploration of bacterial SC. This study marks a step forward in the synthetic biology of SC, with promising applications in biomaterials.

ACKNOWLEDGEMENTS

This project is supported by the European Union’s Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant agreement No 860125 (S.V., C.J.I., A.E.D., and M.M.), the European Research Council (ERC) Consolidator grant 865694: DiversiPHI, the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC 2051 – Project-ID 390713860, the Alexander von Humboldt Foundation in the context of an Alexander von Humboldt-Professorship founded by the German Federal Ministry of Education and Research, VIDI grant (VI.Vidi.203.074) from The Netherlands Organization for Scientific Research (NWO) (R.H.J.S), and B-INK

Proof of Concept grant from the ERC 101188114 (S.V.). Open access funding provided by Max Planck Society.

REFERENCES

Allen, R.M., Chatterjee, R., Madden, M.S., Ludden, P.W. & Shah, V.K. (1994). Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Critical Reviews in Biotechnology*, *14*(3), 225–249. doi:10.3109/07388554409079834

Altschul S.F., Gish, W., Miller W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. Journal of molecular biology 215(3), 403-410.

Andrade, P. & Carneiro, M. (2021). Pterin-based pigmentation in animals. *Biology Letters*, *17*(8), 20. doi:10.1098/rsbl.2021.0221

Ara, K.Z.G., Månberger, A., Gabriško, M., Linares-Pastén, J.A., Jasilionis, A., Friðjónsson, Ó.H. & Nordberg Karlsson, E. (2020). Characterization and diversity of the complete set of GH family 3 enzymes from Rhodothermus marinus DSM 4253. *Scientific Reports*, *10*(1), 1329. doi:10.1038/s41598-020-58015-5

Aspeborg, H., Coutinho, P.M., Wang, Y., Brumer, H. & Henrissat, B. (2012). Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC Evolutionary Biology*, *12*(1). doi:10.1186/1471-2148-12-186

Batianis, C., Kozaeva, E., Damalas, S.G., Martín-Pascual, M., Volke, D.C., Nikel, P.I. & Martins Dos Santos, V.A.P. (2020). An expanded CRISPRi toolbox for tunable control of gene expression in Pseudomonas putida. *Microbial Biotechnology*, *13*(2), 368–385. doi:10.1111/1751-7915.13533

Bendtsen, J. D., Kiemer, L., Fausbøll, A., & Brunak, S. (2005). Non-classical protein secretion in bacteria. *BMC Microbiology*, *5*(1). doi:10.1186/1471-2180-5-58

Brodie, J., Ingham, C. J. & Vignolini, S. (2021). Does structural color exist in true fungi? *Journal of Fungi*, *7*(2), 141. doi:10.3390/jof7020141

Campos, A., Apraiz, I., da Fonseca, R. R. & Cristobal, S. (2015). Shotgun analysis of the marine mussel *Mytilus edulis* hemolymph proteome and mapping the innate immunity elements. *Proteomics*, *15*(23–24), 4021–4029. doi:10.1002/pmic.201500118

Campos, A., Danielsson, G., Farinha, A. P., Kuruvilla, J., Warholm, P. & Cristobal, S. (2016). Shotgun proteomics to unravel marine mussel (Mytilus edulis) response to long-term exposure to low salinity and propranolol in a Baltic Sea microcosm. *Journal of Proteomics*, *137*, 97–106. doi:10.1016/j.jprot.2016.01.010

Chen, W., Jiang, X., & Yang, Q. (2020). Glycoside hydrolase family 18 chitinases: The known and the unknown. Biotechnology Advances, 43(107553), 107553. doi:10.1016/j.biotechadv.2020.107553

Colette Daubner, S. & Lanzas, R. O. (2018). Pteridines. *Reference Module in Biomedical Sciences*.

Doucet, S. M. & Meadows, M. G. (2009). Iridescence: a functional perspective. *Journal of the Royal Society, Interface*, *6 Suppl 2*(suppl_2), S115-32. doi:10.1098/rsif.2008.0395.focus

Dushkina, N. & Lakhtakia, A. (2013). Structural Colors. *Engineered Biomimicry*, 267–303. Elsevier.

Feirer, N. & Fuqua, C. (2017). Pterin function in bacteria. *Pteridines*, *28*(1), 23–36. doi:10.1515/pterid-2016-0012

Groutars, E. G., Risseeuw, C. C., Ingham, C., Hamidjaja, R., Elkhuizen, W. S., Pont, S. C., & Karana, E. (2022). Flavorium: An exploration of Flavobacteria's living aesthetics for living color interfaces. CHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM.

Hackl, T., Ankenbrand, M. & van Adrichem, B. (2024). gggenomes: A Grammar of Graphics for Comparative Genomics. R package version 1.0.0. Available from: https://github.com/thackl/gggenomes

Hamidjaja, R., Capoulade, J., Catón, L. & Ingham, C.J. (2020). The cell organization underlying structural colour is involved in *Flavobacterium* IR1 predation. *The ISME Journal*, *14*(11), 2890–2900. doi:10.1038/s41396-020-00760-6

Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. *Journal of Molecular Biology*, *166*(4), 557–580. doi:10.1016/s0022-2836(83)80284-8

Hasona, A., Ray, R.M. & Shanmugam, K.T. (1998). Physiological and genetic analyses leading to identification of a biochemical role for the moeA (molybdate metabolism) gene product in Escherichia coli. *Journal of Bacteriology*, *180*(6), 1466–1472. doi:10.1128/JB.180.6.1466-1472.1998

Hyatt D., Chen, G., Locascio, P.F., Land, M.L., Larimer, F.W. & Hauser L.J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC bioinformatics 11, 1-11.

Ingham, C.J., Sprenkels, A., Bomer, J., Molenaar, D., van den Berg, A., van Hylckama Vlieg, J.E.T. & de Vos, W.M. (2007). The micro-Petri dish, a million-well growth chip for the culture and high-throughput screening of microorganisms. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(46), 18217–18222. doi:10.1073/pnas.0701693104

Johansen, V.E., Catón, L., Hamidjaja, R., Oosterink, E., Wilts, B.D., Rasmussen, T.S. & Vignolini, S. (2018). Genetic manipulation of structural color in bacterial colonies. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(11), 2652–2657. doi:10.1073/pnas.1716214115

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C. & Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics (Oxford, England)*, *30*(9), 1236–1240. doi:10.1093/bioinformatics/btu031

Kanehisa, M., Sato, Y. & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology*, *428*(4), 726–731. doi:10.1016/j.jmb.2015.11.006

Kasana, R.C., Salwan, R., Dhar, H., Dutt, S. & Gulati, A. (2008). A rapid and easy method for the detection of microbial cellulases on agar plates using gram's iodine. *Current Microbiology*, *57*(5), 503–507. doi:10.1007/s00284-008-9276-8

Katoh, K. & Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution, 30(4):772-80.

Kientz, B., Agogué, H., Lavergne, C., Marié, P. & Rosenfeld, E. (2013). Isolation and distribution of iridescent Cellulophaga and other iridescent marine bacteria from the Charente-Maritime coast, French Atlantic. *Systematic and Applied Microbiology*, *36*(4), 244–251. doi:10.1016/j.syapm.2013.02.004

Kientz, B., Ducret, A., Luke, S., Vukusic, P., Mignot, T. & Rosenfeld, E. (2012). Glitter-like iridescence within the bacteroidetes especially Cellulophaga spp.: optical properties and correlation with gliding motility. *PloS One*, *7*(12). doi:10.1371/journal.pone.0052900

Kientz, B., Luke, S., Vukusic, P., Péteri, R., Beaudry, C., Renault, T. & Rosenfeld, E. (2016). A unique self-organization of bacterial sub-communities creates iridescence in Cellulophaga lytica colony biofilms. *Scientific Reports*, *6*(1). doi:10.1038/srep19906

Kientz, B., Marié, P. & Rosenfeld, E. (2012). Effect of abiotic factors on the unique glitter-like iridescence of Cellulophaga lytica. *FEMS Microbiology Letters*, *333*(2), 101–108. doi:10.1111/j.1574-6968.2012.02614.x

Kientz, B., Vukusic, P., Luke, S. & Rosenfeld, E. (2012). Iridescence of a marine bacterium and classification of prokaryotic structural colors. *Applied and Environmental Microbiology*, *78*(7), 2092–2099. doi:10.1128/AEM.07339-11

Kinoshita, S., Yoshioka, S. & Miyazaki, J. (2008). Physics of structural colors. *Reports on Progress in Physics. Physical Society (Great Britain)*, *71*(7). doi:10.1088/0034-4885/71/7/076401

Letunic, I. & Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Research.

Leimkühler, S. (2017). Shared function and moonlighting proteins in molybdenum cofactor biosynthesis. Biological Chemistry, 398(9), 1009–1026. doi:10.1515/hsz-2017-0110

Lisov, A.V., Kiselev, S.S., Trubitsina, L.I., Belova, O.V., Andreeva-Kovalevskaya, Z.I., Trubitsin, I.V. & Leontievsky, A.A. (2022). Multifunctional enzyme with endoglucanase and alginase/glucuronan lyase activities from bacterium Cellulophaga lytica. *Biochemistry. Biokhimiia*, *87*(7), 617–627. doi:10.1134/s0006297922070045

Mamiatis, T., Fritsch, E.F., Sambrook, J. & Engel, J. (1985). Molecular cloning–A laboratory manual. New York: Cold Spring Harbor Laboratory. 1982, 545 S., 42 $. *Acta Biotechnologica*, *5*(1), 104–104. doi:10.1002/abio.370050118

Mewis, K., Lenfant, N., Lombard, V. & Henrissat, B. (2016). Dividing the large glycoside hydrolase family 43 into subfamilies: A motivation for detailed enzyme characterization. *Applied and Environmental Microbiology*, *82*(6), 1686–1692. doi:10.1128/aem.03453-15

Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. *Analytical Chemistry*, *75*(17), 4646–4658. doi:10.1021/ac0341261

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A. & Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Molecular biology and evolution, 37(5), 1530-1534.

Parker, A.R. (2000). 515 million years of structural colour. *Journal of Optics A Pure and Applied Optics*, *2*(6), R15–R28. doi:10.1088/1464-4258/2/6/201

Patinios, C., Creutzburg, S.C.A., Arifah, A.Q., Adiego-Pérez, B., Gyimah, E.A., Ingham, C.J. & Staals, R.H.J. (2021). Streamlined CRISPR genome engineering in wild-type bacteria using SIBR-Cas. *Nucleic Acids Research*, *49*(19), 11392–11404. doi:10.1093/nar/gkab893

Quinlan, A.R. & Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26.6, 841-842.

Sai, T., Froufe-Pérez, L.S., Scheffold, F., Wilts, B.D. & Dufresne, E.R. (2023). Structural color from pigment-loaded nanostructures. *Soft Matter*, *19*(40), 7717–7723. doi:10.1039/d3sm00961k

Schertel, L., van de Kerkhof, G.T., Jacucci, G., Catón, L., Ogawa, Y., Wilts, B.D. & Johansen, V.E. (2020). Complex photonic response reveals three-dimensional self-organization of structural coloured bacterial colonies. *Journal of the Royal Society, Interface*, *17*(166). doi:10.1098/rsif.2020.0196

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675. doi:10.1038/nmeth.2089

Seemann, T. (2024). Barrnap: rapid ribosomal RNA prediction. Available from: https://github.com/tseemann/barrnap

Shrivastava, A., Johnston, J.J., van Baaren, J.M. & McBride, M.J. (2013). *Flavobacterium* johnsoniae GldK, GldL, GldM, and SprA are required for secretion of the cell surface gliding motility adhesins SprB and RemA. *Journal of Bacteriology*, *195*(14), 3201–3212. doi:10.1128/jb.00333-13

Sullivan, C.J., Brown, K., Hung, C.-S., Tang, J.K.-H., DeSimone, M., Chen, V., Kelley-Loughnane, N. (2023). Iridescent biofilms of Cellulophaga lytica are tunable platforms for scalable, ordered materials. Scientific Reports, 13(1). doi:10.1038/s41598-023-38797-0

Tanthanuch, W., Chantarangsee, M., Maneesan, J. & Ketudat-Cairns, J. (2008). Genomic and expression analysis of glycosyl hydrolase family 35 genes from rice (Oryza sativa L.). *BMC Plant Biology*, *8*(1). doi:10.1186/1471-2229-8-84

Tao, H., Hasona, A., Do, P. M., Ingram, L. O., & Shanmugam, K. T. (2005). Global gene expression analysis revealed an unsuspected deo operon under the control of molybdate sensor, ModE protein, in Escherichia coli. Archives of Microbiology, 184(4), 225–233. doi:10.1007/s00203-005-0039-7

Terrapon, N., Lombard, V., Gilbert, H.J. & Henrissat, B. (2015). Automatic prediction of polysaccharide utilization loci in Bacteroidetes species. *Bioinformatics (Oxford, England)*, *31*(5), 647–655. doi:10.1093/bioinformatics/btu716

Teufel, F., Almagro Armenteros, J.J., Johansen, A.R., Gíslason, M.H., Pihl, S.I., Tsirigos, K.D. & Nielsen, H. (2022). SignalP 6.0 predicts all five types of signal peptides using protein

language models. *Nature Biotechnology*, *40*(7), 1023–1025. doi:10.1038/s41587-021-01156-3

van de Kerkhof, G.T., Schertel, L., Catòn, L., Parton, T.G., Müller, K.H., Greer, H.F. & Vignolini, S. (2022). Polysaccharide metabolism regulates structural colour in bacterial colonies. *Journal of the Royal Society, Interface*, *19*(190). doi:10.1098/rsif.2022.0181

Viborg, A.H., Terrapon, N., Lombard, V., Michel, G., Czjzek, M., Henrissat, B. & Brumer, H. (2019). A subfamily roadmap of the evolutionarily diverse glycoside hydrolase family 16 (GH16). *The Journal of Biological Chemistry*, *294*(44), 15973–15986. doi:10.1074/jbc.ra119.010619

Vignolini, S., Moyroud, E., Glover, B.J. & Steiner, U. (2013). Analysing photonic structures in plants. *Journal of the Royal Society, Interface*, *10*(87). doi:10.1098/rsif.2013.0394

Vogler-Neuling, V.V., Saba, M., Gunkel, I., Zoppe, J.O., Steiner, U., Wilts, B.D. & Dodero, A. (2023). Biopolymer photonics: From nature to nanotechnology. *Advanced Functional Materials*. doi:10.1002/adfm.202306528

Weiss, M.C., Sousa, F.L., Mrnjavac, N., Neukirchen, S., Roettger, M., Nelson-Sathi, S. & Martin, W.F. (2016). The physiology and habitat of the last universal common ancestor. *Nature Microbiology*, *1*(9). doi:10.1038/nmicrobiol.2016.116

Wijnen, B., Leertouwer, H.L. & Stavenga, D.G. (2007). Colors and pterin pigmentation of pierid butterfly wings. *Journal of Insect Physiology*, *53*(12), 1206–1217. doi:10.1016/j.jinsphys.2007.06.016

Wilts, B.D., Wijnen, B., Leertouwer, H.L., Steiner, U. & Stavenga, D.G. (2017). Extreme refractive index wing scale beads containing dense pterin pigments cause the bright colors of pierid butterflies. *Advanced Optical Materials*, *5*(3). doi:10.1002/adom.201600879

Wootton, J.C., Nicolson, R.E., Mark Cock, J., Walters, D.E., Burke, J.F., Doyle, W.A. & Bray, R.C. (1991). Enzymes depending on the pterin molybdenum cofactor: sequence families, spectroscopic properties of molybdenum and possible cofactor-binding domains. *Biochimica et Biophysica Acta. Bioenergetics*, *1057*(2), 157–185. doi:10.1016/s0005-2728(05)80100-8

Zhang, Y. & Gladyshev, V.N. (2008). Molybdoproteomes and evolution of molybdenum utilization. *Journal of Molecular Biology*, *379*(4), 881–899. doi:10.1016/j.jmb.2008.03.051

Zomer, A., Ingham, C.J., von Meijenfeldt, F.A.B., Escobar Doncel, Á., van de Kerkhof, G.T., Hamidjaja, R. & Dutilh, B.E. (2024). Structural color in the bacterial domain: The ecogenomics of a 2-dimensional optical phenotype. *Proceedings of the National Academy of Sciences of the United States of America*, *121*(29). doi:10.1073/pnas.2309757121

FIGURES

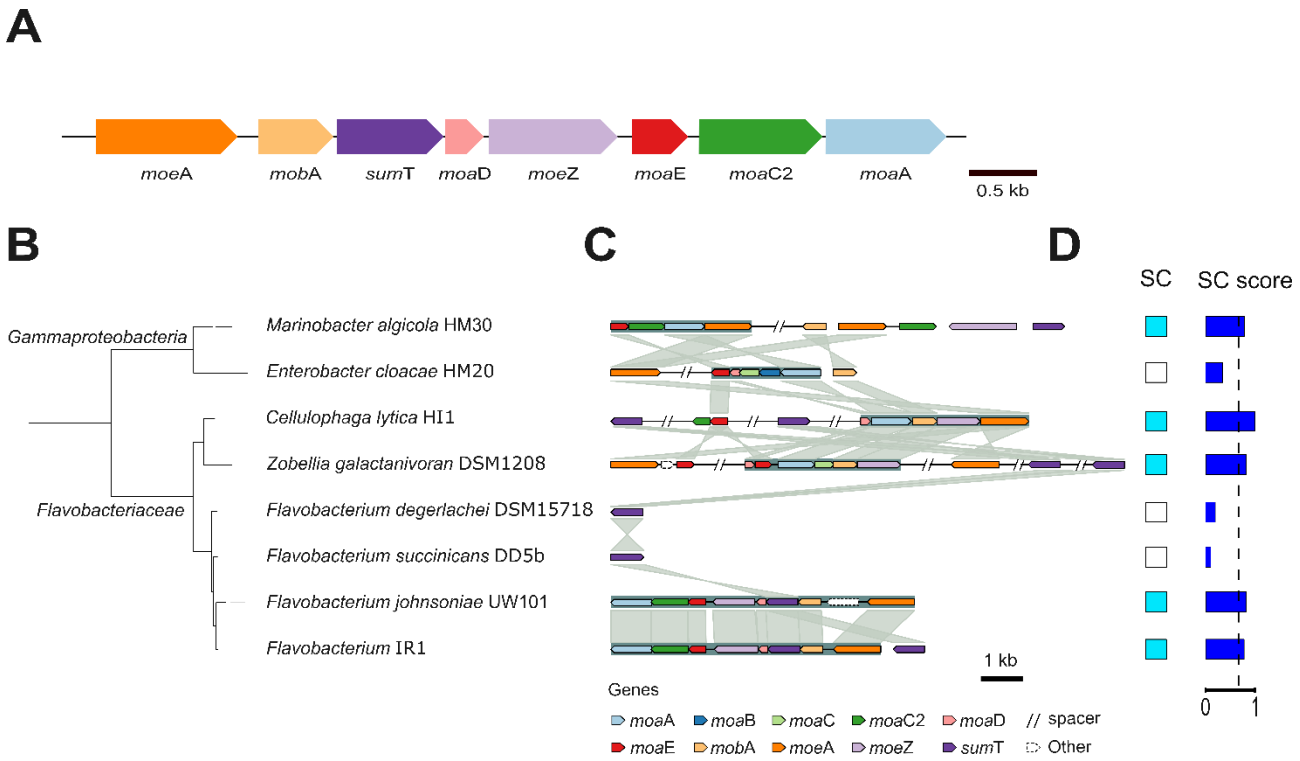


Figure 1. A) Schematics of the putative molybdopterin synthesis operon in the IR1 genome. In blue, the target gene: *moe*A. B) Phylogenetic tree of the 16S ribosomal RNA gene, showing IR1 and other 7 selected strains. C) Synteny and homology visualization of genes that are putatively involved in molybdopterin synthesis. Spacers indicated with // represent stretches longer than 5kb on the same contig, which may encode unshown genes. Whitespaces separate different contigs. D) Presence of SC in the selected strains and its SC score based on the SC classifier software (Zomer et al., 2024). The suggested cut-off value (0.68) for presence of SC is shown as dashed vertical line.

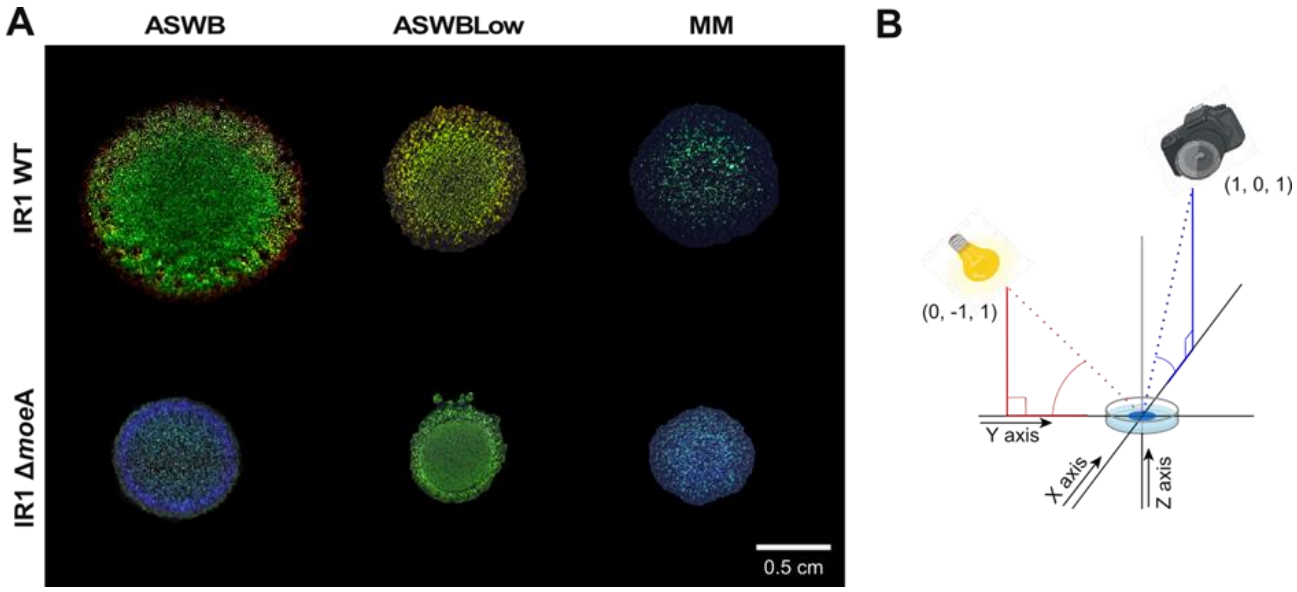


Figure 2. A) Colonies of IR1 WT and IR1 Δ*moe*A grown on agar plates with three different nutrient conditions: ASWB, ASWBLow and MM. B) Schematic of how the colony image was taken. It shows the position of the incident light and the camera as X, Y and Z coordinates (X,Y,Z). The colony is positioned at position (0,0,0), the light source at (0,-1,1), and the camera at (1,0,1). The red dotted line represents the light direction, the blue dotted line represents the camera direction, and the red and blue lines the position of the light and the camera.

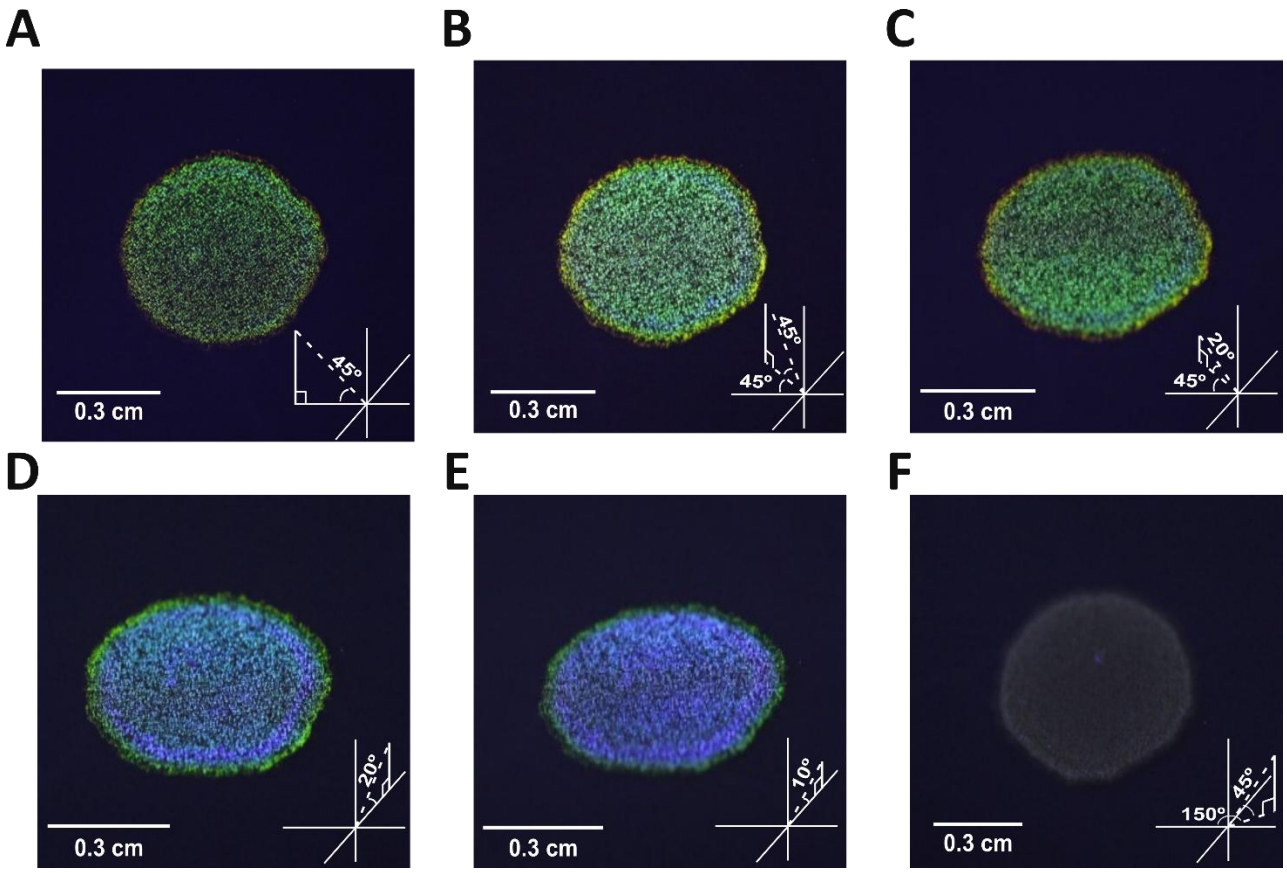


Figure 3. Δ*moe*A colonies grown on ASWB and photographed from different angles. The location of the camera is shown in the bottom right of each panel, following the scheme on Figure 2B. The camera coordinates are A) (0,-1,1.1), B) (1,-1,1), C) (1,-1,0.36), D) (1,0,0.36), E) (1,0,0.18), and F) (0.58,1,1). The light was always positioned at (0,-1,1).

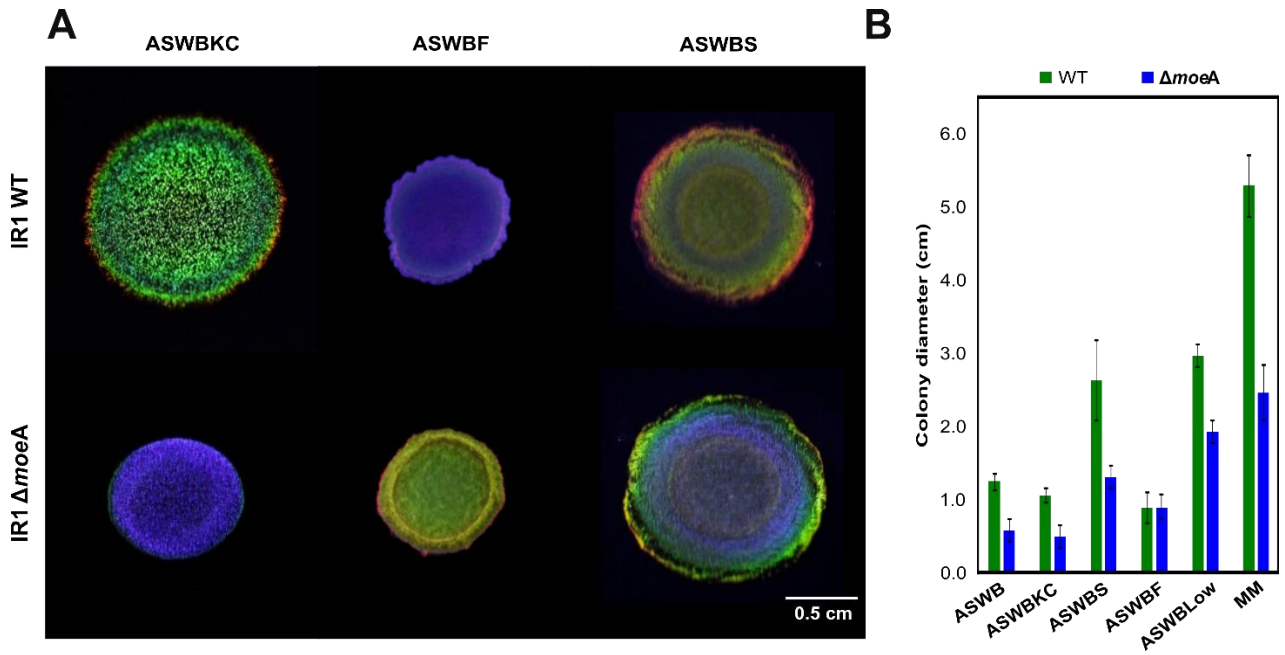


Figure 4. A) Colonies of IR1 WT and Δ*moe*A are grown for 2 days with 1% of 3 different polysaccharides: Artificial Sea Water Black with Kappa-Carrageenan instead of agar (ASWBKC), ASWB with agar and Fucoidan (ASWBF), and ASWB with agar Starch (ASWBS). All the photos were taken from position (1,0,1), following the scheme on Figure 2B. B) Colony diameter in centimeters of IR1 WT and Δ*moe*A grown on different media after 6 days, as mean ± standard deviation of three biological replicates.

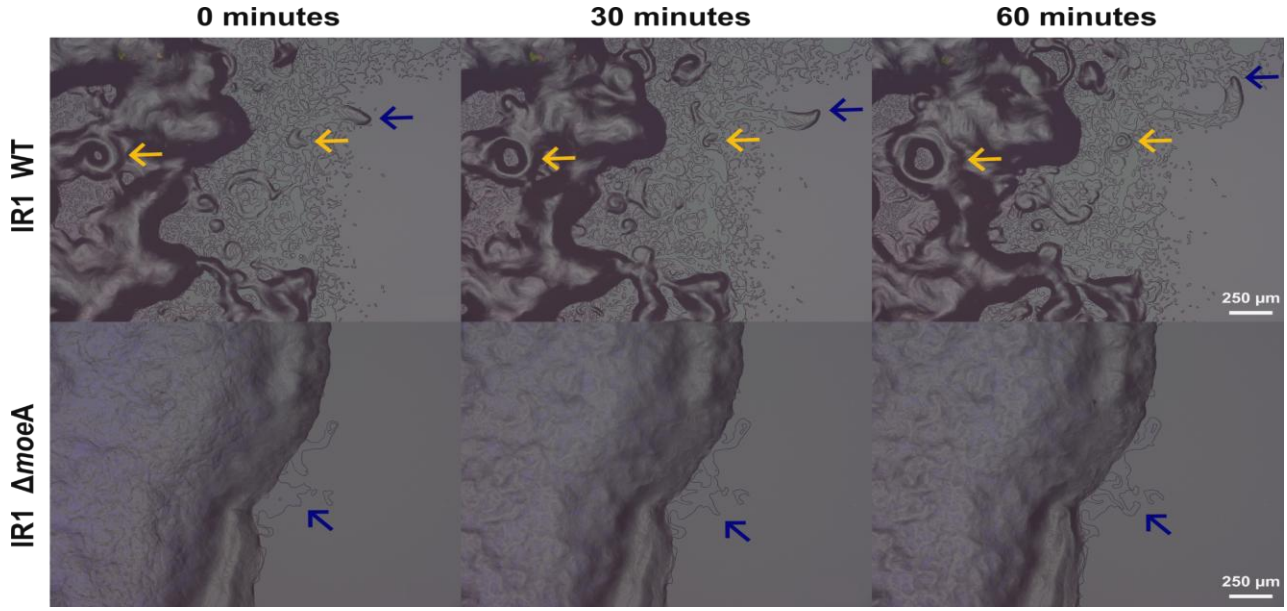


Figure 5. Images taken with a KEYENCE microscope using full coaxial light at the edge of the colony of IR1 WT and Δ*moe*A growing on ASWB. These are frames at 0 minutes, 30 minutes and 60 minutes from the respective 1-hour time-lapse videos. The blue arrows indicated the motility of a group of cells, and the yellow arrows indicated the forming of circular 'vortex' patterning and movement.

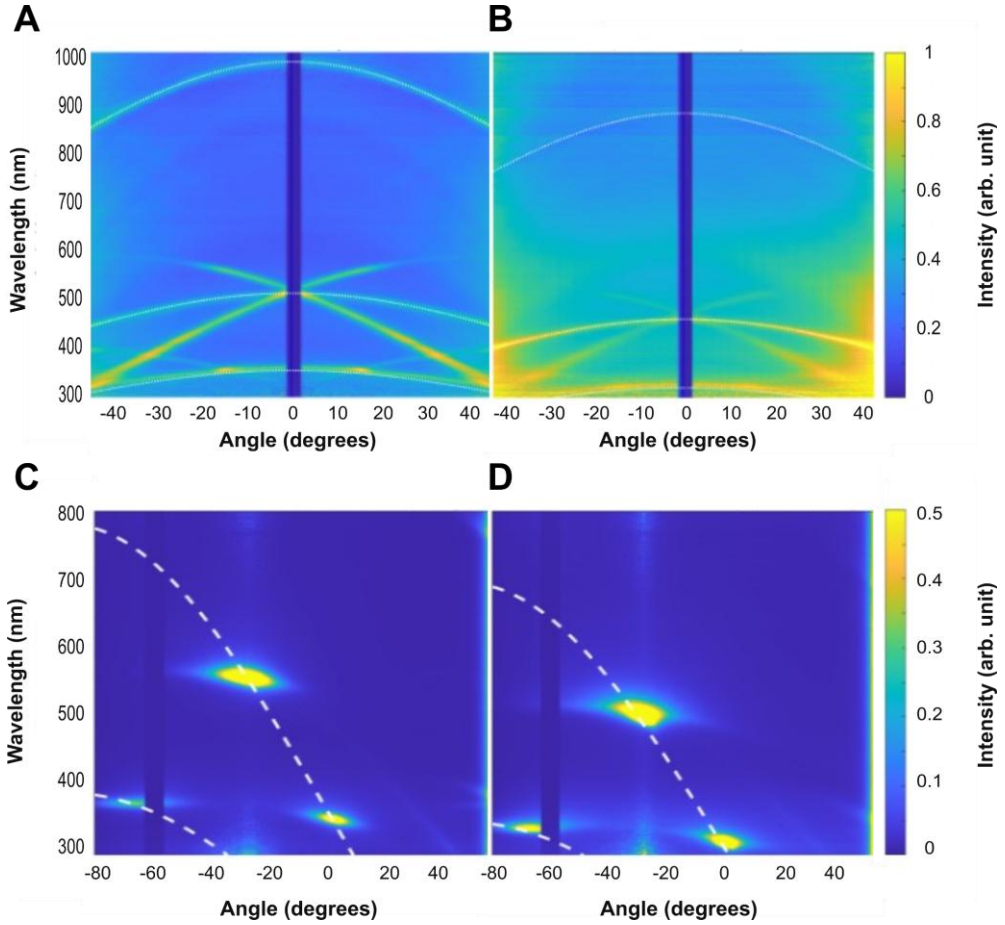


Figure 6. Goniometry analysis of IR1 WT and Δ*moe*A strains grown as a film layer on ASWBKC medium. Specular reflection analysis of A) WT, and B) Δ*moe*A, and scattering (light illumination with an angle of 60º) of C) WT, and D) Δ*moe*A. The dotted lines represent the values of the grating equation.

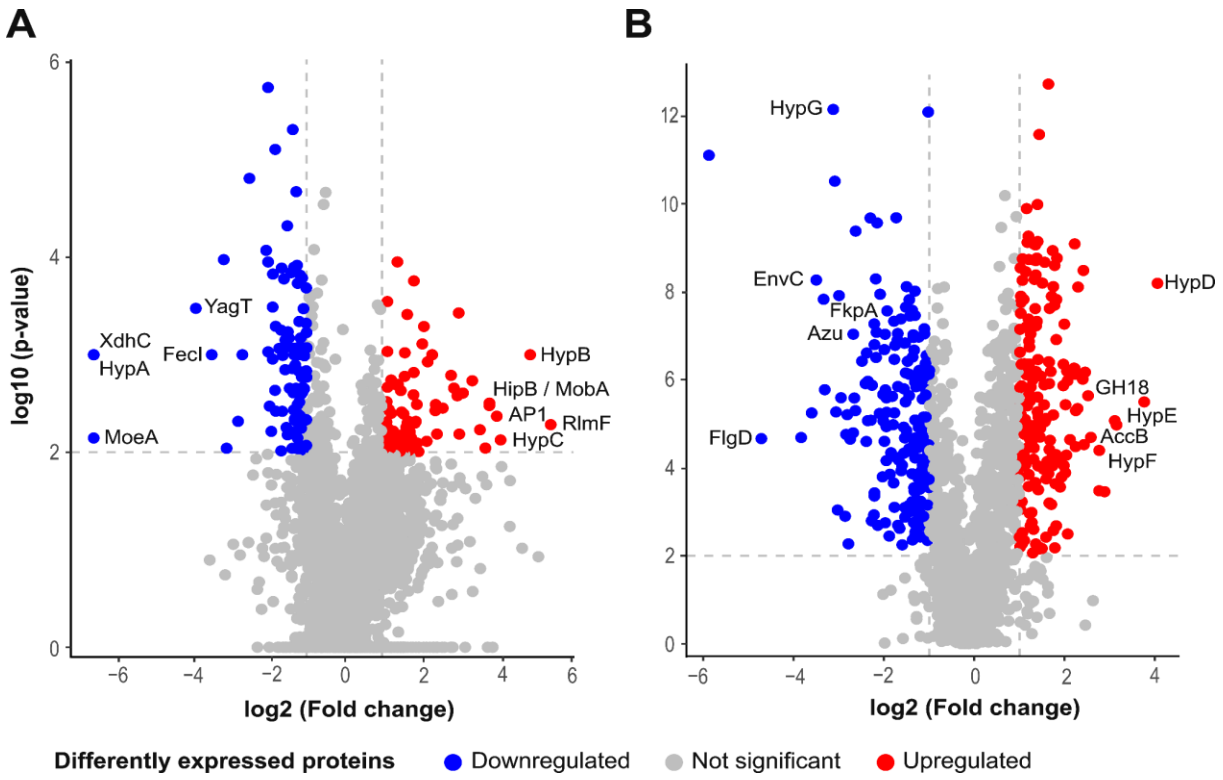


Figure 7. Volcano plots of the peptides identified in A) the intracellular protein analysis, and in B) the extracellular protein analysis. Some of the most regulated proteins are shown in the plots. The horizontal dashed lines represent the cut-off value for the p-value (2), and the vertical dashed lines represent the cut-off value for the fold change (-1 and 1).

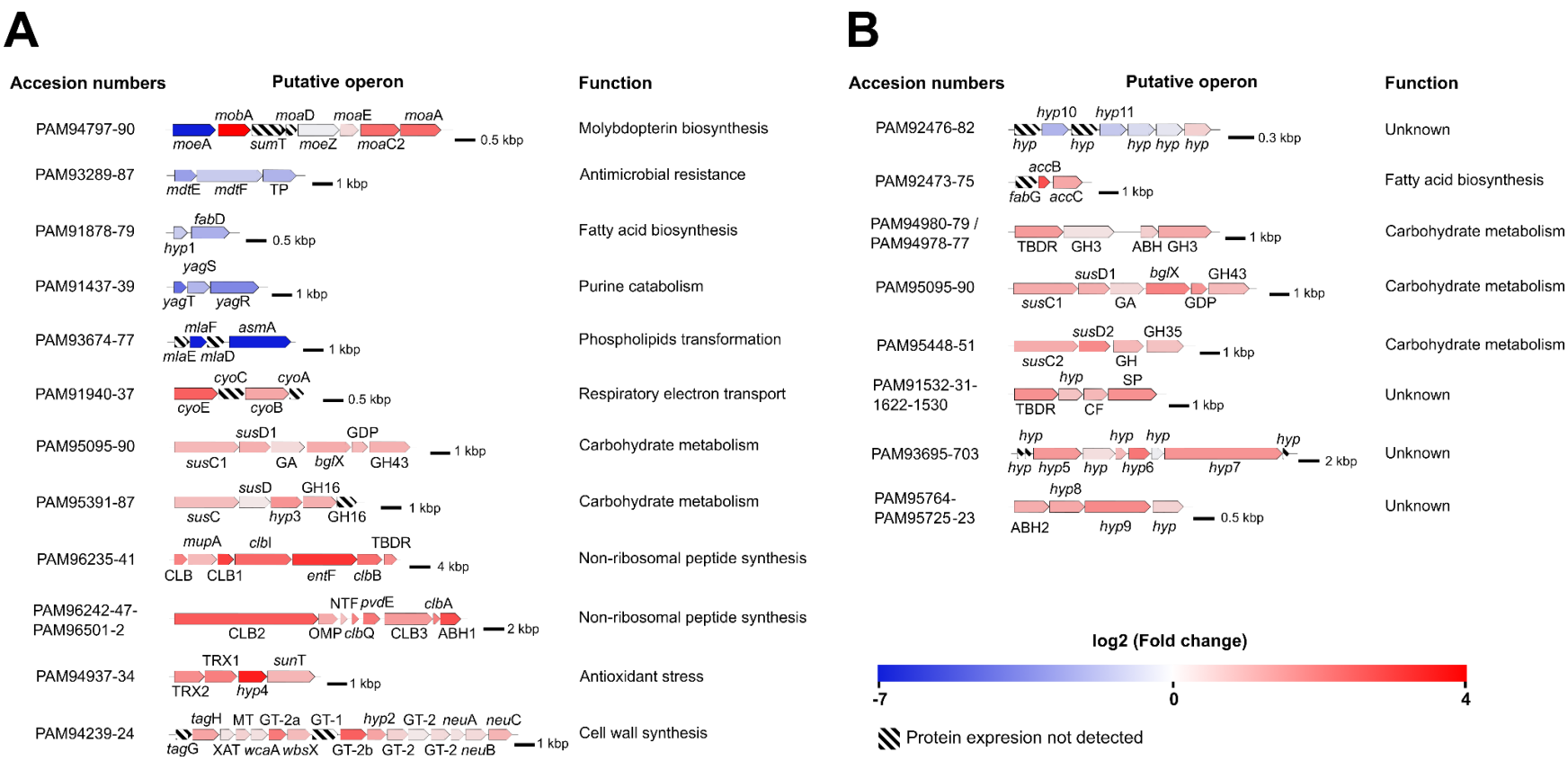


Figure 8. Putative operons or gene clusters with differentially expressed proteins identified in A) intracellular and B) extracellular proteomic analyses, based on function and proximity. To the left of each operon are the accession numbers of the translated proteins, to the right is the predicted function. Gene or protein names are indicated. Genes are colored based on the fold change of the encoded proteins. The black bars show the scale in kilobase pairs (kbp). TP: transporter, *hyp*: hypothetical protein, GA: glycoamylase, GHX: glycosyl hydrolase family X, GDP: glycerophosphoryl diester phosphodiesterase, CLB: colibactin biosynthesis, TBDR: TonB-dependent receptor, OMP: outer membrane protein, NTF: nuclear transport factor 2, ABH: alpha/beta hydrolase, TRX: thioredoxin domain-containing protein, XAT: xenobiotic acyltransferase, MT: SAM-dependent methyltransferase, GT-X: glycosyltransferase family X, CF: cell surface protein, SP: secretion protein.

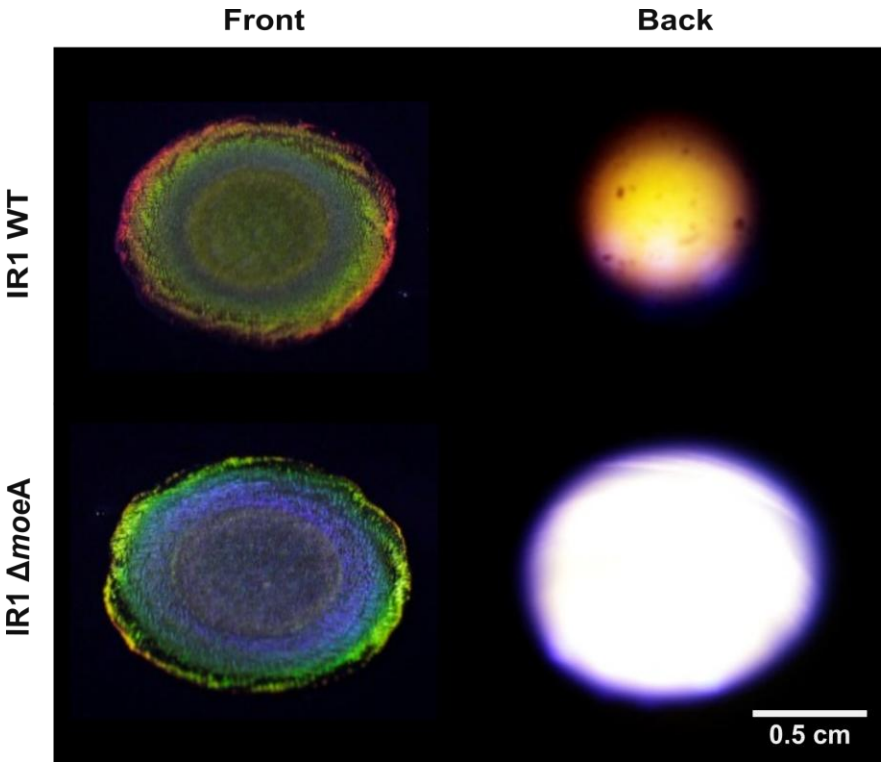


Figure 9. Colonies of IR1 WT (top row) and Δ*moe*A (bottom) grown on ASWS. Iodine vapor was used to dye the starch remaining in the media. The zones of starch degradation are seen as the lighter areas under the colonies. The images were taken at the same 90º angle from the front (left column) and back (right) of the plate.

TABLES

Table 1. Analysis of 117 bacterial genomes (87 SC and 30 non-SC) for the presence of the genes involved in molybdopterin cofactor synthesis.

	***moe*A**	***mob*A**	***sum*T**	***moa*D**	***moe*Z**	***moa*E**	***moa*C2**	***moa*A**
SC bacteria	100%	87%	100%	70%	100%	100%	100%	100%
Non-SC bacteria	40%	37%	63%	20%	70%	40%	40%	50%

SUPPLEMENTARY INFORMATION

Table S1. Bacterial strains used in this study.

Bacterial strains	**Relevant characteristics**	**Source**
Escherichia coli DH5α	Strain used for general cloning	NEB

Flavobacterium IR1	Wild-type (WT)	Johansen et al., 2018
Flavobacterium IR1 Δ*moe*A	IR1 *moe*A knock-out (KO) via SIBR	This study

Table S2. Plasmids used in this study.

Plasmid name	**Description**	**Source**
pSIBR048	*omp*Ap-Int3 FnCas12a-mapt, Hup-NT spacer-*omp*At, Spec^R (*E. coli*), Erm^R (IR1)	(Patinios et al., 2021)
pMoeA_NT	pSIBR048 containing homologous arms of *moe*A	This study
pMoeA_S1	pSIBR048 containing homologous arms of *moe*A and targeting spacer 1 for *moe*A	This study

Table S3. Oligonucleotides used in this study.

Oligo ID	**Sequence**	**Description**
cHA fwd	TCCTCCTTAGCTCAGTTGGTTAG	To check the introduction of the homologous arms into the plasmid, forward
cHA rev	CAGGAAACAGCTATGACCATG	To check the introduction of the homologous arms into the plasmid, reverse
cSp fwd	CAGGAAACAGCTATGACCATG	To check the introduction of the spacer into the plasmid, forward
cSp rev	CCAACACTTGCAAGGAACGG	To check the introduction of the spacer into the plasmid, reverse
*moe*A US fwd	GGCCTCGAGATCTCCATGGATATTT CCCAAGATGAATTTG	Upstream homologous arm for *moe*A, forward
*moe*A US rev	GCTATTTTATAAGGTAAGCA	Upstream homologous arm for *moe*A, reverse
*moe*A DS fwd	TGCTTACCTTATAAAATAGCAGCAGT GTGTAAATTTAAAC	Downstream homologous arm for *moe*A, forward
*moe*A DS rev	CCTGCAATAAATCCTGCAGT	Downstream homologous arm for *moe*A, reverse
cHA inner *moe*A	TAAAGATGCAGGCGTTTACG	To check the insertion of the homologous arms of *moe*A
*moe*A S1 fwd	TGGTCTCTTAGACATTATTGCGCAA AATAGTACATCTATGAGACCT	*moe*A spacer insertion 1, forward
*moe*A S1 rev	AGGTCTCATAGATGTACTATTTTGC GCAATAATGTCTAAGAGACCA	*moe*A spacer insertion 1, reverse
cFwd *moe*A	GCTGTATAGGATGTAAAGCC	To check the deletion of the *moe*A gene in the genome, forward

cRev *moe*A	TAAAGATGCAGGCGTTTACG	To check the deletion of the *moe*A gene in the genome, reverse

Table S4. The most downregulated intracellular proteins in the Δ*moe*A mutant, and proteins mentioned in the main text and in Figure 8.

Role	**ID protein (GenBank)**	**Protein name**	**Gene name***	**Fold change**
Molybdenum cofactor (MoCo) synthesis	PAM94797	Molybdopterin molybdenum transferase	*moe*A	-6.64
Purine catabolism	PAM91437	2Fe-2S ferredoxin	*yag*T	-3.94
	PAM91438	FAD-binding molybdopterin dehydrogenase	*yag*S	-1.82
	PAM91439	Aldehyde oxygenase	*yag*R	-3.20
Fatty acid biosynthesis	PAM91878	DUF983 domain-containing protein	*hyp*1	-1.18
	PAM91879	[acyl-carrier-protein] S-malonyltransferase	*fab*D	-2.02
Phospholipid transformation	PAM93675	ABC transporter ATP-binding protein	*mla*F	-6.64
	PAM93677	Membrane assembly protein	*asm*A	-6.64
Nitrogen assimilation	PAM94801	Nitrite reductase	-	-1.15
Translation	PAM92040	Ribosomal protein L27	*rpm*A	-1.57
	PAM92136	Ribosomal protein S15	*rps*O	-1.54
	PAM94293	Ribosomal protein L16	*rpl*P	-1.35
RNA processing	PAM93099	tRNA(guanosine(37)-N1)-methyltransferase	*trm*D	-1.38
	PAM93366	Proline—tRNA ligase	*pro*S	-1.25
	PAM96417	tRNA(adenosine(37)-N6)-threonylcarbamoyltransferase complex ATPase subunit type 1	*tsa*E	-2.70
DNA transcription	PAM92509	RNA polymerase sigma-54 factor	*rpo*N	-1.23
	PAM93777	RNA polymerase sigma-19 factor	*fec*I	-3.52
Amino acid biosynthesis	PAM91434	Alanine dehydrogenase	*xdh*C	-6.64
	PAM94076	Type II 3-dehydroquinate dehydratase	*aro*Q	-1.41
	PAM95008	Acetylornithine deacetylase	*arg*E	-1.33
	PAM95600	Glutamine synthetase	*gln*A	-1.89
Antimicrobial resistance	PAM93287	Transporter	TP	-1.90
	PAM93288	Hydrophobe/amphiphite efflux-1 family RND transporter	*mdt*F	-1.17
	PAM93289	Efflux transporter periplasmic adaptor unit	*mdt*E	-2.52
Unknown	PAM92103	Hypothetical protein	*hyp*A	-6.64
	PAM93709	Hypothetical protein	-	-6.64

* Gene name or protein with the designation that gives the most information. "-" means no gene name; *hyp*: hypothetical; TP: transporters.

Table S5. The 5 most upregulated intracellular proteins in the Δ*moe*A mutant, and proteins mentioned in the main text and in Figure 8.

Role	**ID protein (GenBank)**	**Protein name**	**Gene name***	**Fold change**
Molybdenum cofactor (MoCo) synthesis	PAM94790	GTP 3',8-cyclase	*moa*A	2.18
	PAM94791	Cyclic pyranopterin monophosphate synthase 2	*moa*C2	2.15
	PAM94796	Molybdenum cofactor guanylyl transferase	*mob*A	3.84
Nitrogen assimilation	PAM94787	NAD(P)H-nitrite reductase	*nir*B	1.05
Cell wall synthesis	PAM94230	Hypothetical protein	*hyp*2	1.23
	PAM94231	Hypothetical protein	GT-2a	2.32
	PAM94234	Hypothetical protein	GT-2b	1.92
	PAM94238	ABC transporter ATP-binding protein	*tag*H	1.24
Respiratory electron transport	PAM91938	Cytochrome oxidase subunit III	*cyo*B	1.00
	PAM91940	Protoheme IX farnesyltransferase	*cyo*E	2.32
Carbohydrate metabolism	PAM95090	Xylosidase	GH43	1.03
	PAM95092	Beta-glucosidase	*bgl*X	1.11
	PAM95094	Nutrient uptake outer membrane protein	*sus*D1	1.06
	PAM95388	Laminarase	GH16	1.07
	PAM95389	Hypothetical protein	*hyp*3	1.52
Stress response	PAM94360	Chalcone isomerase	-	4.14
	PAM94935	DUF6734 family protein	TRX1	3.39
	PAM94936	Hypothetical protein	TRX2	1.84
	PAM94937	Hypothetical protein	*hyp*4	1.62
Signal transduction	PAM95501	Response regulator	-	3.59
RNA processing	PAM93105	23S rRNA (adenine(1618)-N(6))-methyltransferase	*rlm*F	5.47
Regulation of DNA transcription	PAM92290	TetR family transcriptional regulator	*acr*R	1.38
	PAM93723	Transcriptional regulator	*omp*R	2.41
	PAM94944	Transcriptional regulator	*hip*B	3.84
Proteolysis	PAM96640	Aminopeptidase	AP1	4.04
Non-ribosomal peptide synthesis	PAM96235	Hypothetical protein	CLB	1.78
	PAM96237	Hypothetical protein	CLB1	2.83
	PAM96238	Hypothetical protein	*clb*I	2.21
	PAM96239	Hypothetical protein	*ent*F	3.03
	PAM96240	Hypothetical protein	*clb*B	2.11
	PAM96241	Hypothetical protein	TBDR	1.60
	PAM96242	Hypothetical protein	CLB2	2.42
	PAM96243	Hypothetical protein	OMP	1.14

	PAM96245	Thioesterase	*clb*Q	1.77
	PAM96247	Non-ribosomal peptide synthetase	CLB3	1.40
	PAM96501	4-phosphopantetheinyl transferase	*clb*A	1.83
	PAM96502	Alpha/beta hydrolase	ABH1	2.61
Unknown	PAM96476	Hypothetical protein	*hyp*B	4.92

* Gene name or protein with the designation that gives the most information. "-" means no gene name; *hyp*: hypothetical; GT-2:glycosyltransferase family 2; GHX: glycosyl hydrolase family X; TRX: thioredoxin domain-containing protein; TBDR: TonB-dependent receptor; OMP: outer membrane protein; CLB: colibactin biosynthesis ABH: alpha/beta hydrolase.

Table S6. The most 5 downregulated extracellular proteins in the Δ*moe*A mutant, and proteins mentioned in the main text and in Figure 8.

Role	**ID protein (GenBank)**	**Protein name**	**Gene name***	**Secretion pathway****	**Fold change**
Protein modification	PAM93429	Peptidylprolyl isomerase	*fkp*A	SP	-3.00
	PAM93873	Glutamine cyclotransferase	-	SP	-2.18
Carbohydrate metabolism	PAM91916	Galactose oxidase	-	SP	-1.48
	PAM91646	Nutrient uptake outer membrane protein	*sus*C	SP	-1.03
Stress response	PAM94872	Superoxide dismutase	*sod*C	SP	-1.51
Cell division	PAM92714	Murein hydrolase activator	*env*C	SP	-3.51
Antibiotic resistance	PAM96491	Serine hydrolase	-	SP	-2.40
	PAM91787	Serine hydrolase	-	SP	-1.71
Lipopolysaccharide assembly	PAM94009	Hypothetical protein	*hyp*G	NC	-3.13
	PAM95883	Hypothetical protein	-	SP	-2.19
Electron transport	PAM91962	Cytochrome C	*cta*D	SP	-1.13
	PAM95417	Azurin	*azu*	SP	-2.82
Motility	PAM91688	Flagellin biosynthesis protein	*flg*D	SP	-4.72
	PAM91755	Gliding motility protein	*gld*L	NC	-2.22
Transport	PAM92093	TonB-dependent receptor	*fep*A	SP	-1.98
	PAM93287	RND transporter	*tol*C	SP	-1.95
Unknown	PAM92477	Hypothetical protein	*hyp*10	SP	-1.93
	PAM92479	Hypothetical protein	*hyp*11	SP	-1.32
	PAM96055	Hypothetical protein	-	SP	-2.63

* Gene name or protein with the designation that gives the most information. "-" means no gene name; *hyp*: hypothetical.

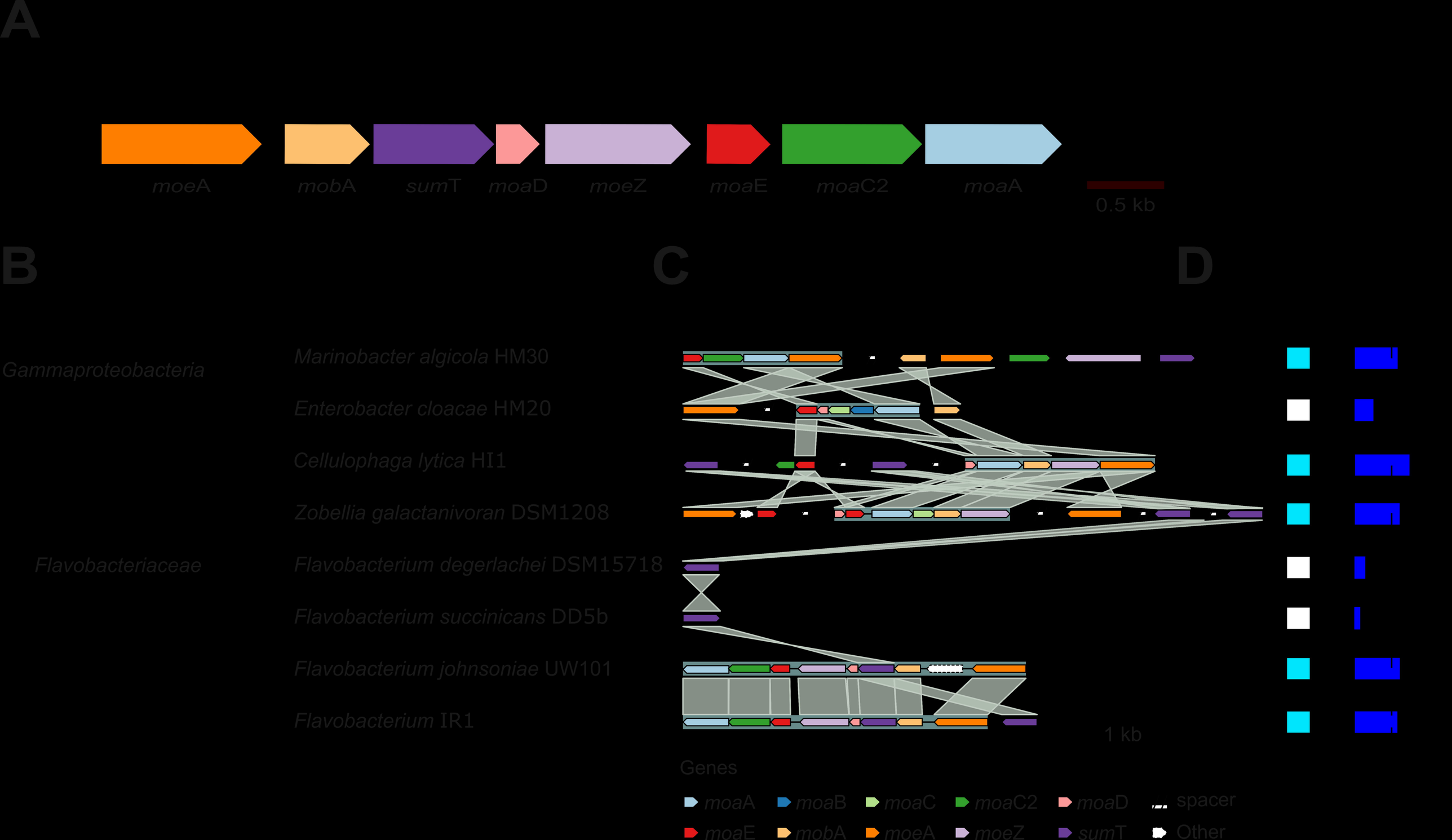
** SP: signal peptide; NC: non-classical; -: not secreted.

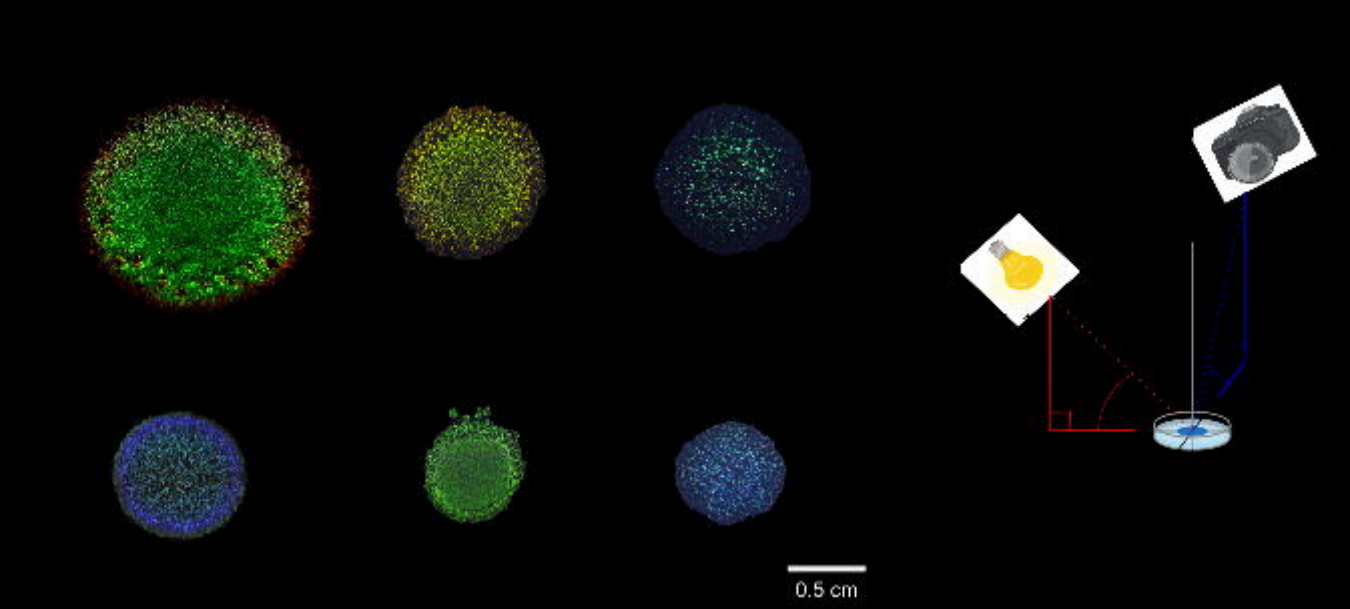
Table S7. The 5 most upregulated extracellular proteins in the Δ*moeA* mutant, and proteins mentioned in the main text and in Figure 8.

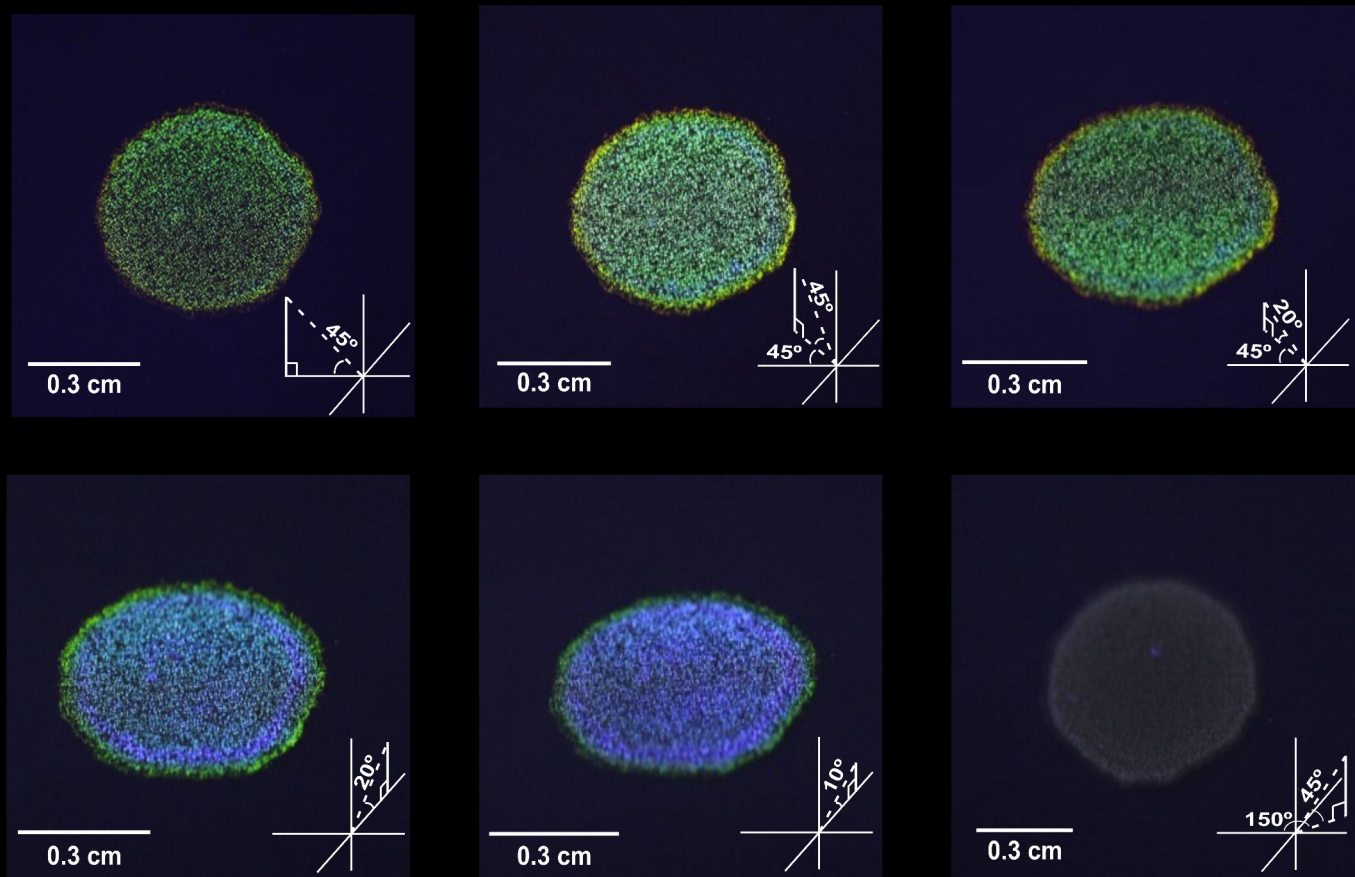
Role	**ID protein (GenBank)**	**Protein name**	**Gene name***	**Secretion pathway****	**Fold change**
Transport	PAM94786	Hypothetical protein	-	SP	2.28
	PAM96649	Hypothetical protein	-	SP	2.41
Proteolysis	PAM93900	Zinc metalloprotease	-	SP	2.29
	PAM95531	Hypothetical protein	-	SP	2.22
Carbohydrate metabolism	PAM93863	Glycoside hydrolase family 18	GH18	SP	2.42
	PAM94980	TonB-dependent receptor	TBDR	SP	1.39
	PAM95091	Glycerophosphodiester phosphodiesterase	GDP	SP	1.49
	PAM95092	Beta-glucosidase	*bgl*X	SP	1.78
	PAM95094	Nutrient uptake outer membrane protein	*sus*D1	SP	1.10
	PAM95095	TonB-linked outer membrane protein	*sus*C1	SP	1.13
	PAM95273	Pectate lyase	*pel*B	SP	1.41
	PAM95448	TonB-linked outer membrane protein	*sus*C2	SP	1.12
	PAM95449	Nutrient uptake outer membrane protein	*sus*D2	SP	1.66
	PAM95764	Alpha/beta hydrolase	ABH2	-	1.10
Fatty acid biosynthesis	PAM92474	Acetyl-CoA carboxylase, biotin carboxyl carrier protein	*acc*B	NC	2.56
	PAM92475	Acetyl-CoA carboxylase biotin carboxylase subunit	*acc*C	-	1.20
Cell division	PAM92348	Hypothetical protein	*hyp*F	SP	2.76
Unknown	PAM91530	Secretion protein	SP	SP	1.55
	PAM91622	Cell surface protein	CF	SP	1.09
	PAM92266	Hypothetical protein	*hyp*E	NC	3.11
	PAM92573	Hypothetical protein	*hyp*D	SP	4.05
	PAM93697	Hypothetical protein	*hyp5*	SP	1.43
	PAM93700	Hypothetical protein	*hyp*6	NC	1.97
	PAM93702	Hypothetical protein	*hyp*7	SP	1.50
	PAM95724	Hypothetical protein	*hyp*8	SP	1.65
	PAM95725	Hypothetical protein	*hyp*9	SP	1.18

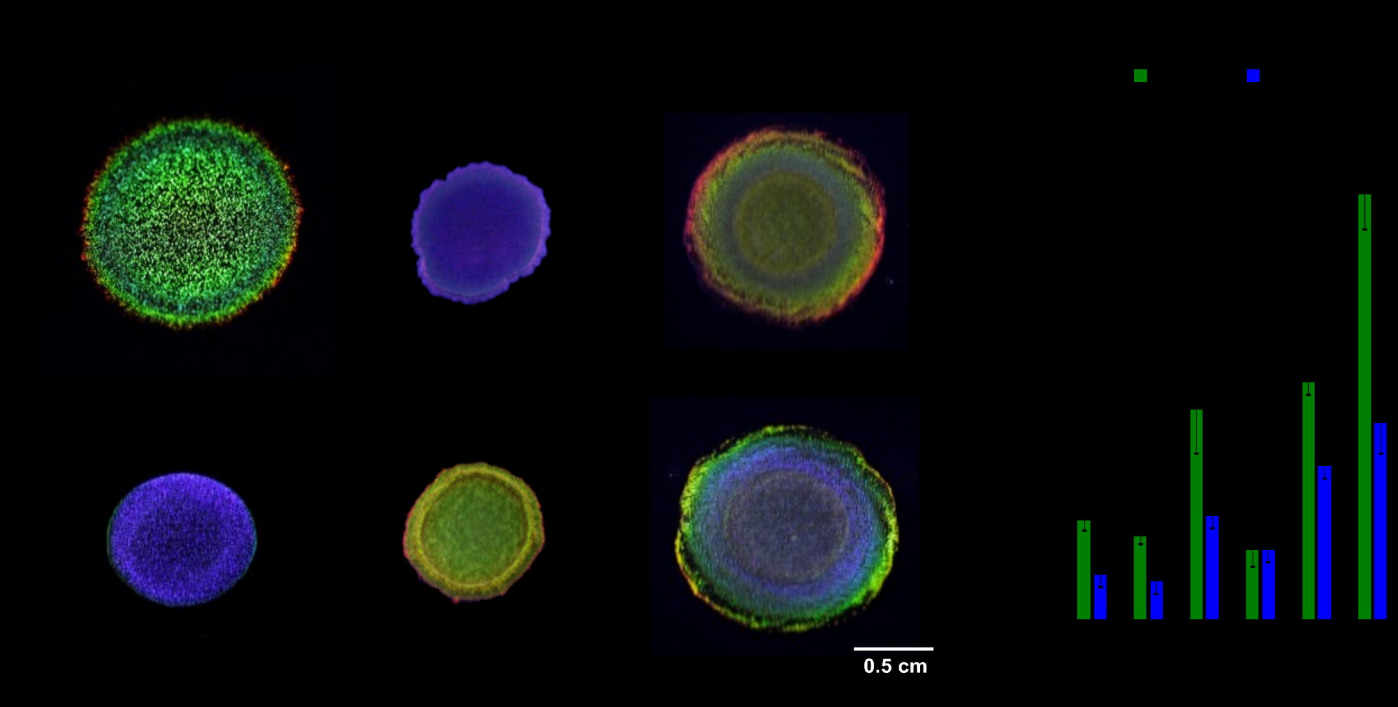
* Gene name or protein with the designation that gives the most information. ABH: alpha/beta hydrolase; SP: secreted protein; CF: cell surface.

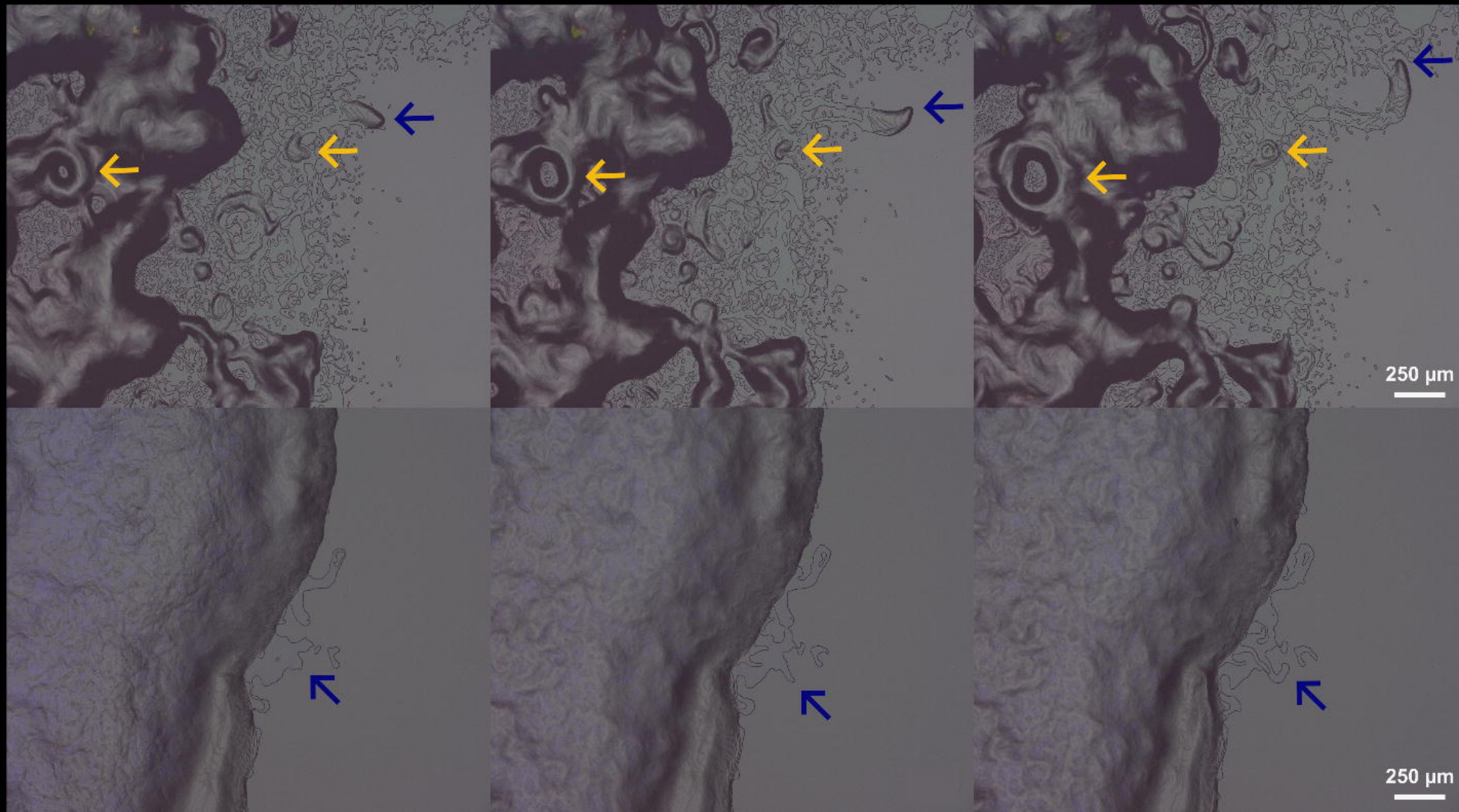
** SP: signal peptide; NC: non-classical; -: not secreted.

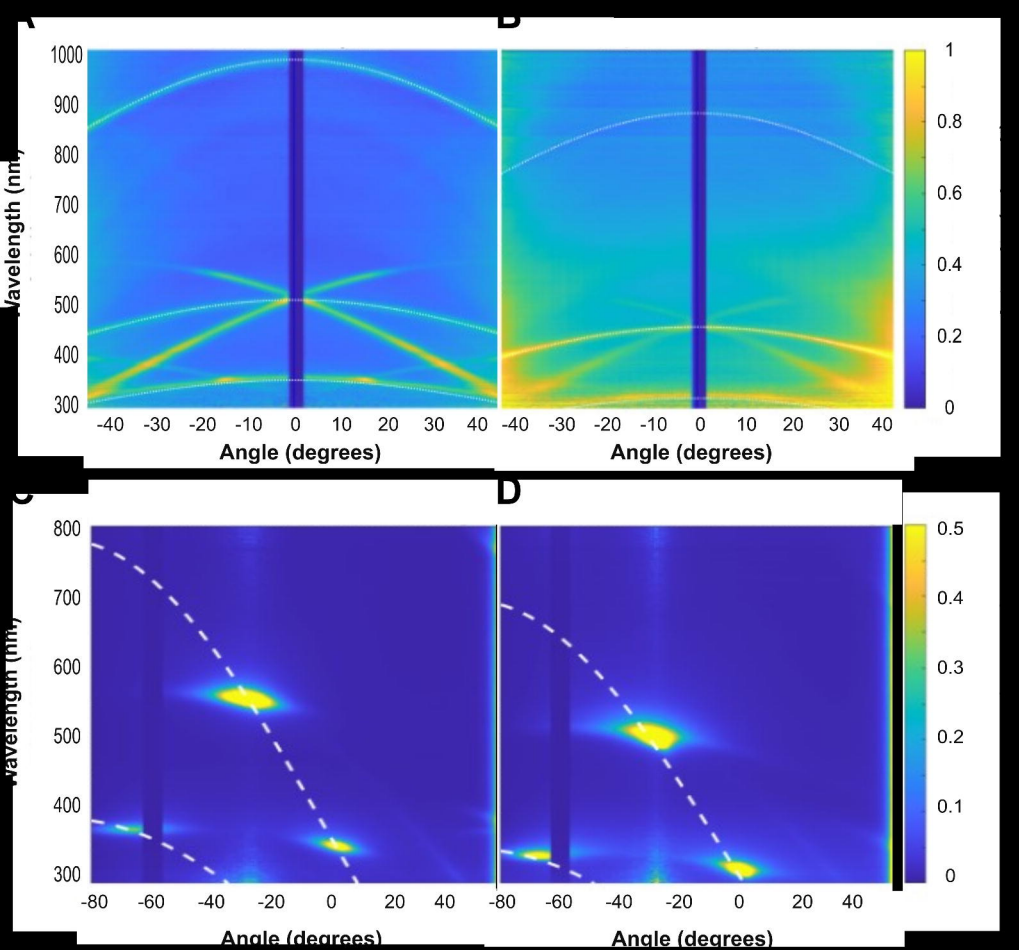

A
moeA
mobA
sumT
moaD
moeZ
moaE
moaC2
moaA
0.5 kb
B
C
D
Gammaproteobacteria
Flavobacteriaceae
Marinobacter algicola HM30
Enterobacter cloacae HM20
Cellulophaga lytica HI1
Zobellia galactanivoran DSM1208
Flavobacterium degerlachei DSM15718
Flavobacterium succinicans DD5b
Flavobacterium johnsoniae UW101
Flavobacterium IR1
1 kb
Genes
moaA
moaB
moaC
moaC2
moaD
spacer
moaE
mobA
moeA
moeZ
sumT
Other

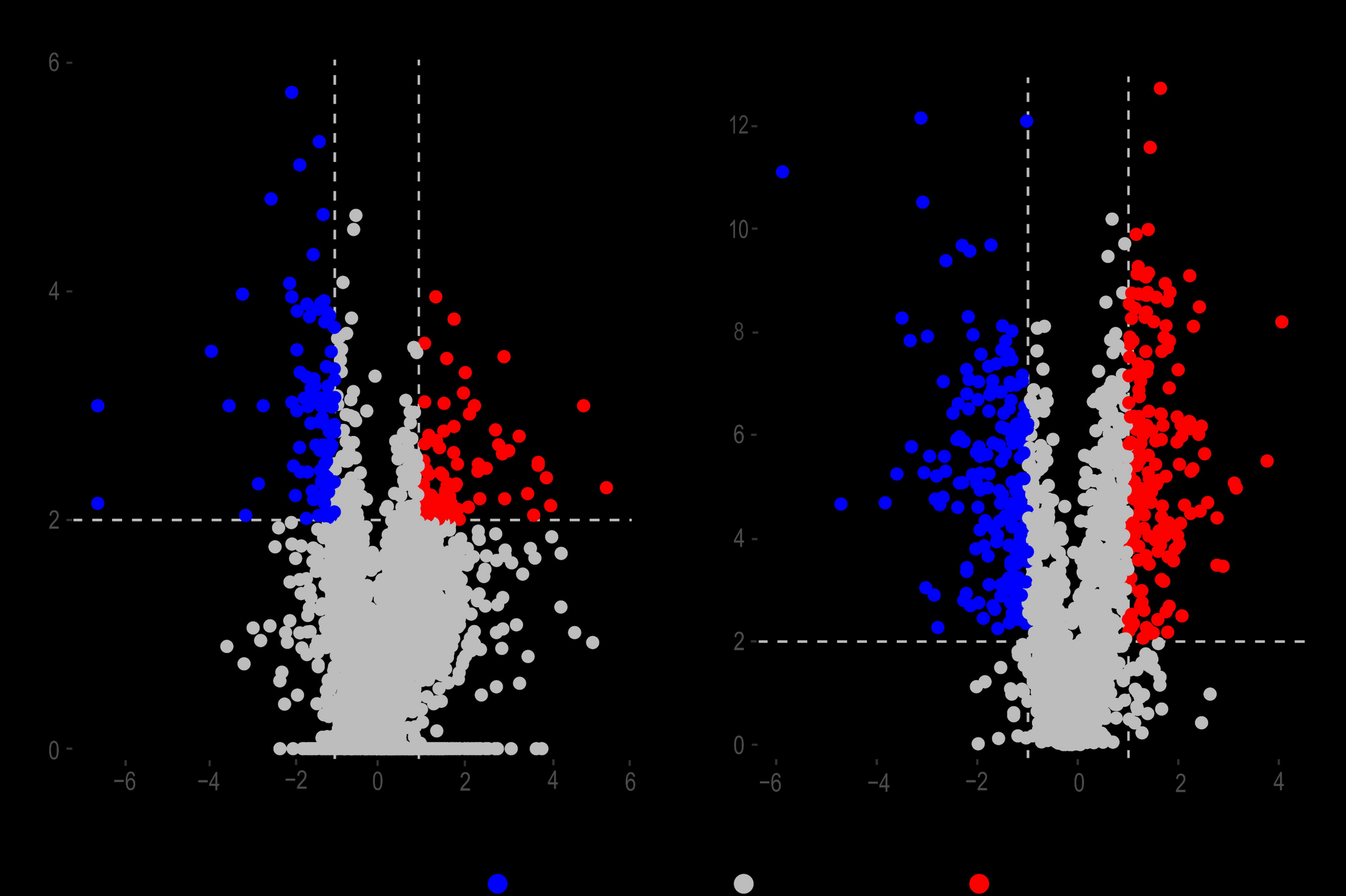

0.5 cm

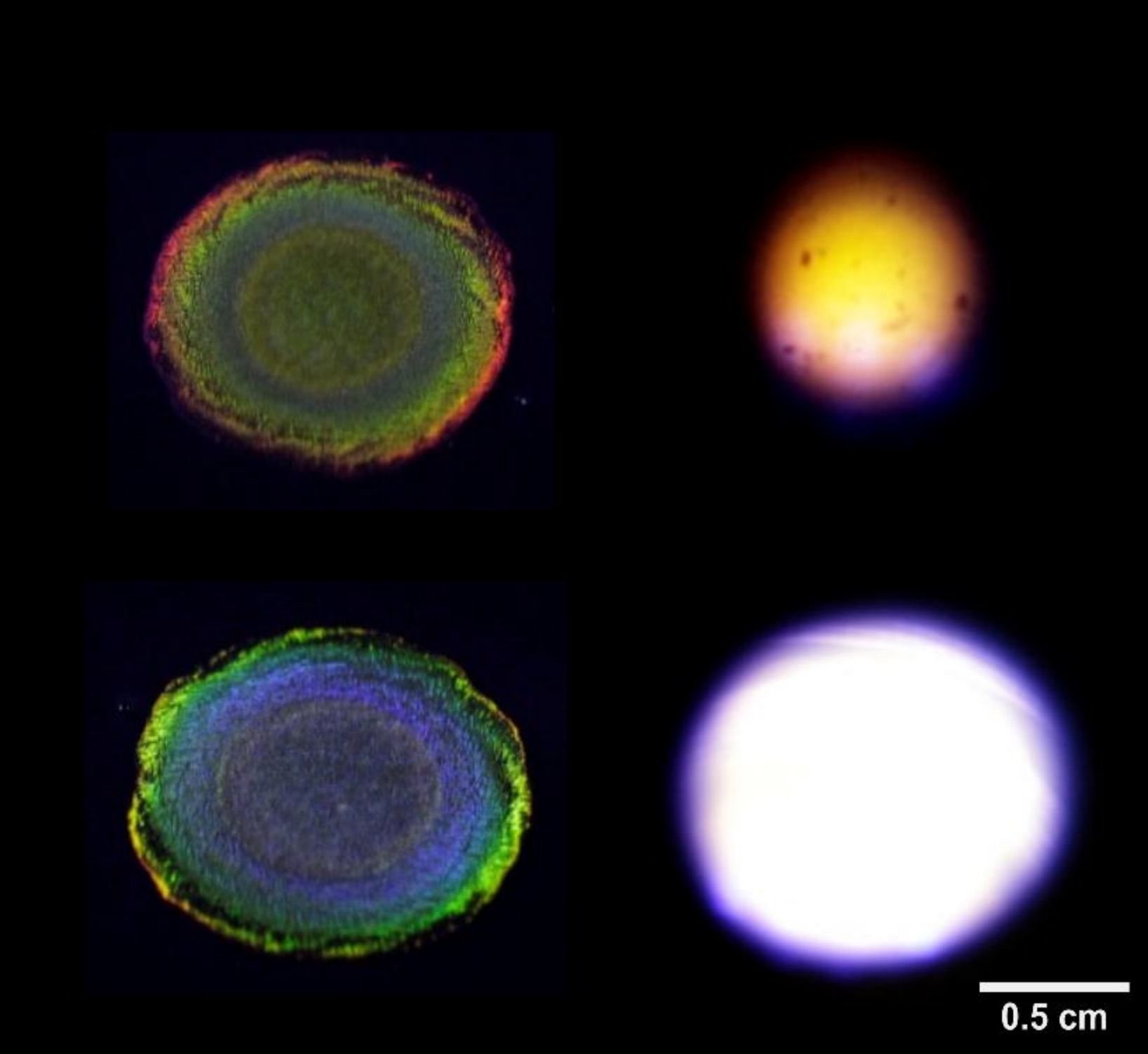

0.3 cm
45°
0.3 cm
45°
45°
0.3 cm
20°
45°
0.3 cm
20°
0.3 cm
10°
0.3 cm
45°
150°


0.5 cm


250 µm
250 µm


A
B
C
D
Wavelength (nm)
Angle (degrees)
1000
900
800
700
600
500
400
300
-40 -30 -20 -10 0 10 20 30 40
1
0.8
0.6
0.4
0.2
0
-80 -60 -40 -20 0 20 40
0.5
0.4
0.3
0.2
0.1
0


6
4
2
0
−6
−4
−2
0
2
4
6
12
10
8
6
4
2
0
−6
−4
−2
0
2
4


0.5 cm