AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

COMMENTARY

How Bioinformatic Tools Guide Experiments To Resolve the Chaos of Apparently Unlimited Metabolic Variation

Jeffrey A. Cole[a]

[a]School of Biosciences and Institute of Microbiology and Infection, University of Birmingham, Birmingham, United Kingdom

ABSTRACT Hexuronic acids, oxidation products of common sugars, are widespread in eukaryotic cells. Galacturonic acid is the main carbohydrate component of pectin found in plant cell walls and glucuronic acid is a component of proteoglycans in animals. However, despite their importance as carbohydrate substrates, metabolism of hexuronic acids has long remained a poorly studied corner of the bacterial metabolic map. In the current issue of *Journal of Bacteriology*, Bouvier and coworkers present a detailed analysis of genes involved in hexuronate utilization in various proteobacteria and report the verification of their bioinformatics predictions by carefully designed experiments (J. T. Bouvier et al., J Bacteriol 201:e00431-18, 2019, https://doi.org/10.1128/JB.00431-18). This study provides a solid basis for understanding hexuronate metabolism and its regulation in other bacterial phyla.

KEYWORDS bioinformatics, hexuronates, proteobacteria

Between 2000 and 2011, a series of comparative genomics papers were published from Michael Gelfand's group with Dmitry Rodionov as first author. Some of these papers were the starting point for almost 100 comparative genomics papers that included supporting experimental verification in collaboration with other groups (reviewed in reference 1). Topics covered included carbohydrate utilization pathways and their regulation in various microorganisms (for some examples, see references 2 to 7). One of the first papers in the series that focused on hexuronate metabolism was followed 11 years later by a more detailed account from the Ozoline group of the transcription factors that control these pathways (8, 9). Given the similarity of that topic to the current paper by Bouvier et al. (10), what makes the current paper worthy of special comment?

The aim of many research projects is to identify the physiological roles of genes of unknown function. Fragments of information are often available, but there are critical gaps in our knowledge often because key genes appear to be missing or have been incorrectly annotated. Various bioinformatics tools provide reliable predictions to guide experimental verification. The current paper is a tour de force showing how intelligent use of bioinformatics predictions can—and should—be substantiated by detailed experimental analysis. However, it also illustrates our time-honored experience that every research project reveals further unanswered questions.

Most microbiology and biochemistry university courses include basic accounts of glycolysis, the citric acid cycle, and how bacteria selectively metabolize glucose in preference to lactose. Roles of LacI and Crp are often mentioned but are soon forgotten by students once they have passed their examination. Few courses include analysis of ribose, rhamnose, or galactose metabolism, let alone glucuronic or galacturonic acids, despite their importance in cell structure and function. In the metabolomics era, metabolism is out of vogue. Galacturonic acid is the main carbohydrate component of pectin found in plant cell walls, and glucuronic acid is a component of proteoglycans in animals; both compounds are components of hemicellulose in plant cell walls.

Citation Cole JA. 2019. How bioinformatic tools guide experiments to resolve the chaos of apparently unlimited metabolic variation. J Bacteriol 201:e00628-18. https://doi.org/10.1128/JB.00628-18.

Editor Michael Y. Galperin, NCBI, NLM, National Institutes of Health

Address correspondence to j.a.cole@bham.ac.uk.

For the article discussed, see https://doi.org/10.1128/JB.00431-18.

Accepted manuscript posted online 29 October 2018
Published 20 December 2018

Nevertheless, until recently little was known about how they are metabolized in bacteria other than *Escherichia coli* and *Bacillus subtilis*. At the start of this project, hexuronic acids were known to be metabolized essentially by two different pathways. In *E. coli* and *B. subtilis*, they are imported by the major facilitator superfamily protein ExuT and converted by the isomerase pathway via 2-keto-3-deoxy-D-gluconate to glyceraldehyde-3-phosphate. In contrast, in various genera of proteobacteria, such as *Agrobacterium*, *Bradyrhizobium*, *Burkholderia*, *Polaromonas*, and *Pseudomonas*, hexuronates are imported by proteins of the tripartite ATP-independent periplasmic (TRAP) transporter family, UxuPQM (11, 12). In *Agrobacterium tumefaciens*, hexuronates are then metabolized by an oxidative pathway in which the key intermediate is 5-keto-4-deoxy-D-galactarate (KDG) and the final product is 2-ketoglutarate (13, 14). The second step in the production of KDG requires one of two ring-opening lactone isomerases, Gli or Gci. These two enzymes were identified and characterized experimentally by Andberg and colleagues in Finland (14) and the Almo and Gerlt groups in the United States (15).

The starting point for the current project was the realization that although most of the genes, including *uxuPQM*, required for the hexuronate oxidative pathway are present in annotated genomes of *Pseudomonaceae*, *Burkholderiaceae*, and *Comomonadaceae*, the expected D-lactaro-1,5-lactone ring-opening enzyme (lactonase) gene appeared to be missing. Bioinformatics analysis of genes colocalized with *uxuPQM* revealed genes for two candidate lactonases, UxuL and UxuF, which are unrelated to Gli and Gci. A screen of other bacteria revealed that they are consistently colocalized with genes for hexuronic acid metabolism. Genes for four putative UxuF orthologues and two UxuL orthologues were cloned and overexpressed. Kinetic analysis of the purified proteins confirmed their lactone isomerase functions. The catalytic efficiency of the UxuL lactonase was higher with glucarolactone than with galactarolactone; the converse was true for UxuF. Both of the newly identified lactonases are members of protein families that require a divalent metal ion, usually Ca^{2+}, for activity. The lactonase activity of UxuL from *Ralstonia pickettii* was shown to be inhibited by EDTA but was far more active with Zn^{2+} than with Ca^{2+} as the divalent metal ion.

The genome context analysis also revealed the presence of a gene for a novel transcription factor, GguR, in four families investigated. GguR is only distantly related to the *E. coli* glucuronate and fructuronate regulators, ExuR and UxuR, and its gene is almost invariably located with genes for aldarate catabolism rather than with genes for hexuronic acid metabolism. Two copies of *gguR* were found in several genomes from the *Pseudomonas*, *Burkholderia*, and *Halomonas* lineages. Phylogenetic trees of these proteins revealed two divergent branches, which suggested their early emergence via lineage-specific gene duplication. A highly conserved palindromic sequence was confirmed experimentally to be the binding site for GguR from *Burkholderiaceae*, *Comamonadaceae*, and *Pseudomonadaceae*, but a different binding sequence was found in the *Halomonadaceae* family. These consensus sequences were used to construct the GguR regulons; this confirmed that the dominant role for GguR is to control aldarate metabolism. The dissociation constant for GguR in the regulatory regions of some of its operons was slightly high, 50 to 100 μM. However, binding was disrupted by 150 μM KDG but not by D-glucuronate, D-galacturonate, D-glucarate, or *meso*-galactarate. Genes predicted to be regulated by GguR were shown by quantitative reverse transcription-PCR (RT-PCR) to be induced 57- to 6,000-fold during growth in the presence of D-glucarate compared with growth with glucose as carbon source. The conclusion drawn is that GguR is a global regulator of aldarate metabolism.

Some of the operons predicted to be required for hexuronate or aldarate metabolism lacked a GguR binding site but included genes for other novel transcription factors. They included two proteins, UdhR and GudR, that are related to LacI, and the LysR family regulator, GulR. Palindromic DNA-binding sites for UdhR and GudR were deduced.

Orthologues of the two known hexuronate transporters ExuT and UxuPQM were identified in 35 other genomes studied. In most cases they were components of the

GguR or GulR regulons. The GguR regulons in *Halomonas* and two *Pseudomonas* species include genes for a novel tripartite tricarboxylate (TTT) family transporter, TctABC. As these genes often cluster with aldarate dehydrogenase genes, TctABC was proposed to be required for aldarate uptake. One of these proteins was screened by differential scanning fluorometry against a 189-compound library that included all variants of 3- to 6-carbon aldonic, aldaric, and uronic acids. Only D-glucarate bound significantly.

A cynic might dismiss the most recent paper from the Rodionov group as yet another story about "same pathway, different organism" and question the value of the bioinformatics guesswork on which it is based. The point of this commentary is to highlight that such superficial analysis fails to equip the reader with tools that might solve their own problems of why attempts to exploit bacteria for particular applications are unsuccessful. The genomics era has long since invalidated the quip that "what is true of *E. coli* is also true of an elephant." Indeed, what is true of one strain of *E. coli* is often not even true for another strain. This paper demonstrates how bioinformatics can be used to design laboratory experiments targeted to bring order out of the chaos of apparently unlimited metabolic variation and its associated regulatory networks.

REFERENCES

1. Rodionov DA, Dubchak IL, Arkin AP, Alm EJ, Gelfand MS. 2005. Dissimilatory metabolism of nitrogen oxides in bacteria: comparative reconstruction of transcriptional networks. PLoS Comput Biol 1:e55. https://doi.org/10.1371/journal.pcbi.0010055.eor.
2. Rodionov DA, Novichkov PS, Stavrovskaya ED, Rodionova IA, Li X, Kazanov MD, Ravcheev DA, Gerasimova AV, Kazakov AE, Kovaleva GY, Permina EA, Laikova ON, Overbeek R, Romine MF, Fredrickson JK, Arkin AP, Dubchak I, Osterman AL, Gelfand MS. 2011. Comparative genomic reconstruction of transcriptional networks controlling central metabolism in the *Shewanella* genus. BMC Genomics 12(Suppl 1):S3. https://doi.org/10.1186/1471-2164-12-S1-S3.
3. Rodionova IA, Scott DA, Grishin NV, Osterman AL, Rodionov DA. 2012. Tagaturonate-fructuronate epimerase UxaE, a novel enzyme in the hexuronate catabolic network in *Thermotoga maritima*. Environ Microbiol 14:2920–2934. https://doi.org/10.1111/j.1462-2920.2012.02856.x.
4. Leyn SA, Gao F, Yang C, Rodionov DA. 2012. N-acetylgalactosamine utilization pathway and regulon in proteobacteria: genomic reconstruction and experimental characterization in *Shewanella*. J Biol Chem 287:28047–28056. https://doi.org/10.1074/jbc.M112.382333.
5. Ravcheev DA, Godzik A, Osterman AL, Rodionov DA. 2013. Polysaccharides utilization in human gut bacterium *Bacteroides thetaiotaomicron*: comparative genomics reconstruction of metabolic and regulatory networks. BMC Genomics 14:873. https://doi.org/10.1186/1471-2164-14-873.
6. Khoroshkin MS, Leyn SA, Van Sinderen D, Rodionov DA. 2016. Transcriptional regulation of carbohydrate utilization pathways in the Bifidobacterium genus. Front Microbiol 7:120.
7. Leyn SA, Suvorova IA, Kazakov AE, Ravcheev DA, Stepanova VV, Novichkov PS, Rodionov DA. 2016. Comparative genomics and evolution of transcriptional regulons in *Proteobacteria*. Microb Genom 2:e000061.
8. Rodionov DA, Mironov AA, Rakhmaninova AB, Gelfand MS. 2000. Transcriptional regulation of transport and utilization systems for hexuronides, hexuronates and hexonates in gamma purple bacteria. Mol Microbiol 38:673–683. https://doi.org/10.1046/j.1365-2958.2000.02115.x.
9. Suvorova IA, Tutukina MN, Ravcheev DA, Rodionov DA, Ozoline ON, Gelfand MS. 2011. Comparative genomic analysis of the hexuronate metabolism genes and their regulation in gammaproteobacteria. J Bacteriol 193:3956–3963. https://doi.org/10.1128/JB.00277-11.
10. Bouvier JT, Sernova NV, Ghasempur S, Rodionova IA, Vetting MW, Al-Obaidi NF, Almo SC, Gerlt JA, Rodionov DA. 2019. Novel metabolic pathways and regulons for hexuronate utilization in *Proteobacteria*. J Bacteriol 201:e00431-18. https://doi.org/10.1128/JB.00431-18.
11. Vetting MW, Al-Obaidi N, Zhao S, San Francisco B, Kim J, Wichelecki DJ, Bouvier JT, Solbiati JO, Vu H, Zhang X, Rodionov DA, Love JD, Hillerich BS, Seidel RD, Quinn RJ, Osterman AL, Cronan JE, Jacobson MP, Gerlt JA, Almo SC. 2015. Experimental strategies for functional annotation and metabolism discovery: targeted screening of solute binding proteins and unbiased panning of metabolomes. Biochemistry 54:909–931. https://doi.org/10.1021/bi501388y.
12. Zhao J, Binns AN. 2016. Involvement of *Agrobacterium tumefaciens* galacturonate tripartite ATP independent periplasmic (TRAP) transporter GaaPQM in virulence gene expression. Appl Environ Microbiol 82:1136–1146. https://doi.org/10.1128/AEM.02891-15.
13. Pick A, Schmid J, Sieber V. 2015. Characterization of uronate dehydrogenase catalyzing the initial step in an oxidative pathway. Microb Biotechnol 8:633–643. https://doi.org/10.1111/1751-7915.12265.
14. Andberg M, Maaheimo H, Boer H, Penttilä M, Koivula A, Richard P. 2012. Characterization of a novel *Agrobacterium tumefaciens* galactarolactone cycloisomerase enzyme for direct conversion of D-galactarolactone to 3-deoxy-2-keto-L-threo-hexarate. J Biol Chem 287:17662–17671. https://doi.org/10.1074/jbc.M111.335240.
15. Bouvier JT, Groninger-Poe FP, Vetting M, Almo SC, Gerlt JA. 2014. Galactaro δ-lactone isomerase: lactone isomerization by a member of the amidohydrolase superfamily. Biochemistry 53:614–616. https://doi.org/10.1021/bi5000492.