The moron comes of age

Nichole Cumby,1 Alan R. Davidson1,2 and Karen L. Maxwell1,*
1Department of Molecular Genetics; University of Toronto; Toronto, ON Canada; 2Department of Biochemistry; University of Toronto; Toronto, ON Canada

Prophage-encoded genes can provide a variety of benefits for their bacterial hosts. These beneficial genes are often contained within “moron” elements. Morons, thus termed as the insertion of the DNA encoding them adds “more on” the genome in which they are found, are independent transcriptional units disseminated among phage genomes through horizontal gene transfer. Morons have been identified in the majority of phage genomes and they have been found to play diverse roles in bacterial physiology. At present, we are only beginning to ascribe functions to the many proteins encoded within these ubiquitous genetic elements. Recently, we discovered that the first described moron-encoded protein, gp15 of phage HK97, is expressed from the HK97 prophage and functions as a superinfection exclusion protein, protecting its host from genome injection by other phages. This work and the growing body of data pertaining to other morons challenges the traditional view of phages as purely parasitic in nature.

This year, the Human Microbiome Project released a comprehensive description of the microbial communities present within humans.1,2 The increasing recognition that diverse interactions with these microbial communities plays a critical role in health and disease has led to the conclusion that an organism’s genes cannot be studied in isolation when evolution occurs in the context of complex communities. Like humans, bacteria possess their own unique microbiomes, of which bacteriophages (phages) are a major component. In fact, much of the diversity observed in closely related bacterial strains is a result of the incorporation of diverse prophages into the core bacterial genome.3 Also like the human microbiome, in which the human body is estimated to contain ten times more microbial cells than human cells, phages on earth are estimated to outnumber bacterial cells by an order of magnitude.4

While it is clear that predation by phages is a dominant factor in regulating bacterial population levels in diverse environments,5 it has also become apparent that phage-encoded genes play an important role in enhancing bacterial survival.6–8 This seeming paradox is explicable by the fact that once a phage inhabits a bacterial community its survival and propagation is dependent on the success of its host population. In fact, while some phages propagate exclusively by a lytic cycle, the majority of phages live a dual lifestyle; growing lytically some of the time, but also incorporating their genomes into their bacterial hosts and existing passively as prophages. A prophage is, in effect, a symbiont; thus, it is evolutionarily adaptive for it to express genes that enhance survival of the host. Given that the vast majority of bacterial genomes contain numerous prophages, it could be argued that most phage genomes spend the bulk of their existences as prophages. As such, the production of phage particles could be viewed as an escape mechanism from certain adverse conditions, analogous to spore production by fungal and bacterial strains. Prophage induction can be triggered by a variety of host cell stressors such as DNA damage,9 changes in pH and temperature,10 heavy metal exposure,10 oxidative stress,11 exposure to bacteriocins,12 and other environmental factors.
and quorum sensing.\textsuperscript{13} Since survival of the prophage depends on the fitness of the bacterial host, the most successful phages will not be those that integrate silently within the genome, but those that actively protect and enhance the fitness of their host by expressing useful genes. The simple addition of a prophage to a bacterial genome thus often has significant phenotypic consequences for the host.\textsuperscript{7,8}

For a prophage to maintain itself within the bacterial genome, many of its genes, such as those involved in particle formation and cell lysis, must be repressed. However, amidst this sea of repression, all prophage genomes contain small islands of expressed genes, the importance of which are becoming increasingly apparent. Many of these expressed genes lie within genetic elements that have been collectively termed “morons,” which were first described by Juhala et al. in a paper comparing the genomes of \textit{E. coli} phages HK97 and HK022.\textsuperscript{14} Within these and other related phage genomes, gene order is highly conserved such that one can identify gene function by virtue of gene position, even in the absence of sequence identity. This allowed the identification of inserted genetic elements in HK97 and HK022 that were unique to these phages. Since their presence resulted in more DNA than other related phage genomes, gene order is highly conserved such that one can identify gene function by virtue of gene position, even in the absence of sequence identity. This allowed the identification of inserted genetic elements in HK97 and HK022 that were unique to these phages.

The first use of the term moron related to the genetic element containing gene $I_5$ of phage HK97, which lies in the tail operon between the genes encoding the tail tube assembly chaperone and the tape measure protein. Recently, we discovered that the product of this moron element, gp15, mediates a useful function for the host cell when expressed from a prophage. By inserting into the inner membrane this small protein prevents superinfection by other phages, including HK97 and HK75, without blocking their bacterial cell surface receptors.\textsuperscript{15} Thus, gp15 appears to prevent phage infection without affecting host cell functions mediated by the outer membrane transporters used by these phages. This exclusion is unrelated to traditional repressor-mediated immunity mechanisms since the phages it blocks belong to different immunity groups.\textsuperscript{15} Interestingly, other prophages express unrelated proteins that exploit similar entry preventing strategies for preventing phage infection. For example, the Cor proteins of \textit{E. coli} phages HK022, N15 and $\phi80$ protect bacterial lysogens against other phages that require the FhuA outer membrane protein for entry by inactivating this receptor.\textsuperscript{16} This activity reduces ferrichrome uptake by cells expressing Cor,\textsuperscript{10} but it plays an additional positive role in providing protection against Colicin M, a protein toxin that is imported through the FhuA receptor.\textsuperscript{17} The \textit{Salmonella} phage P22 membrane-bound SieA protein blocks genome entry by phages L, MG178 and MG40 through an uncharacterized mechanism.\textsuperscript{18} We observed that gp15 inhibition is independent of periplasmic and inner membrane proteins required for HK97 entry, since other phages requiring the same factors are not inhibited by gp15. However, it is dependent on features within the phage; swapping the tail regions of HK97 and HK022 is sufficient to alleviate the inhibition phenotype.\textsuperscript{19} Thus we postulate that gp15 inserts into the inner membrane of \textit{E. coli} where it interacts directly or indirectly with the incoming phage tail to prevent it from forming a channel for the DNA through the membrane. To date, the precise mechanisms of action of gp15 are unknown.

Besides blocking phage infection and toxin import, some moron-encoded proteins provide other advantages to their bacterial hosts. For example, the prophage-expressed $\lambda$ Lom protein improves the ability of \textit{E. coli} strains to interact with and invade human epithelial cells,\textsuperscript{19} and the \textit{Salmonella} phage SopE$\Phi$ protein, SopE, secreted by the bacterial lysogen,\textsuperscript{20} leads to ruffling of the primate kidney cell membrane\textsuperscript{21} and improved invasion of human epithelial cells by \textit{Salmonella}.\textsuperscript{22} Several moron-encoded proteins have been discovered that increase the ability of bacteria to survive the host immune response and proliferate in vivo. For example, the Gifsy-2 \textit{Salmonella} prophage encodes a superoxide dismutase that protects the bacterium against free radicals generated by the mammalian immune system.\textsuperscript{23} The phage $\lambda$ Bor protein has similarly been demonstrated to increase the resistance of lysogenized \textit{E. coli} to serum killing\textsuperscript{25} and the \textit{S. aureus} phage $\phi15$ sak gene encodes staphylokinase,\textsuperscript{26} which disrupts the fibrin network that forms to keep infections localized, and degrades human IgG using a serine protease created by the activation of human plasminogen into plasmin at the bacterial cell surface.\textsuperscript{27} \textit{S. pyogenes} phage-encoded mitogenic factors $mf2$, $mf3$ and $mf4$ act as superantigens to trigger overproliferation of defective T-cells that are unable to bind a specific epitope of the antigen, and thus are ineffective in clearing the bacterial infection.\textsuperscript{28} Some morons encode toxin proteins that directly increase the virulence of the host. For example, shiga, diphtheria and cholera toxins which greatly enhance virulence of \textit{E. coli},\textsuperscript{29} \textit{C. diphtheria}\textsuperscript{30} and \textit{V. cholerae}\textsuperscript{31} infections, respectively, are all encoded by morons.\textsuperscript{6}

Other morons have been shown to help bacteria evade the immune response by modifying bacterial cell antigenicity. For example, Oac from \textit{S. flexneri} phage Sf6 acetylates the O-antigen of the bacterial lipopolysaccharide,\textsuperscript{32} thereby altering the bacterial cell surface and impairing the ability of the immune system to recognize it.\textsuperscript{33} This same alteration is also believed to protect against superinfecting phages that recognize the O-antigen.\textsuperscript{34} \textit{P. aeruginosa} phage D3 is also a serotype-converting prophage; it alters the glycosidic linkages and acetylates the O-antigen.\textsuperscript{35} These changes were demonstrated to be the result of a moron element encoding three genes: an $\alpha$-polymerase inhibitor ($iap$), an $O$-acetylase ($oac$) and a $\beta$-polymerase ($wzy$).\textsuperscript{36} Although the two Oac proteins encoded by phages Sf6 and D3 are only 22% identical, their appearance in two quite different species of bacteria illustrates that while moron functions are diverse, the mechanisms by which they function may be conserved among phages.

In addition to conservation among phages, many of the morons discussed here have bacterial homologs found outside of prophage-encoded regions. When PSI-BLAST searches were performed...
with HK79 gp15, all homologs identified were found in Enterobacteriaceae, and the majority were not phage or prophage-encoded. The bacterial sequences belong to the YebO family, a group of proteins of unknown function.\textsuperscript{13} While the protein sequence of gp15 bears a strong resemblance to these proteins in its N-terminal region, it is highly diverged at the C-terminus, and we found that overexpression or deletion of the \textit{E. coli} YebO protein had no effect on HK79 phage propagation. Thus, we postulate that \textit{yebO} was inserted into the phage genome by horizontal gene transfer, and diverged over time until it acquired the new function of providing phage resistance. Similar to gp15, \textit{oac} of \textit{Pseudomonas} phage D3 has a number of bacterial homologs, including \textit{Pseudomonas} WbpC proteins, with which it shares 32\% sequence identity.\textsuperscript{37} The phage D3 Oac protein plays the same role as that from \textit{S. flexneri} phage Sf6, but the D3 protein is approximately twice as large (687 vs. 333 residues). Analysis of the closest homologs of each of these phage proteins reveals that they are bacterial in origin. Coupled with the large difference in the size of the proteins, this suggests that these proteins were independently co-opted by phages to protect the bacterial lysogens from superinfection.

Recently, a number of effectors of the type III secretion system were discovered to be encoded in phage moron elements. For example, the \textit{S. enterica} Gifsy-1 phage encodes GogB,\textsuperscript{38} which is secreted by the type III secretion system into mammalian cells where it reduces the host inflammatory response that is designed to help clear the \textit{S. enterica} infection, thereby helping the bacteria establish a chronic infection.\textsuperscript{39} A second example is provided by the NleA-like type III effector protein of the shiga-toxin encoding phage BP-4795 which increases virulence in mouse models.\textsuperscript{40,41} Since the effector molecules are only useful when they can be secreted and phages do not encode the secretion apparatus, the genes encoding these effector proteins are likely bacterial in origin.

As the number of bacterial and phage genome sequences increases, the number and diversity of described moron elements also multiplies. In a recent study of \textit{ Burkholderia pseudomallei} phages, 17 unique moron-encoded protein sequences were identified, with some phages encoding as many as 12 different morons.\textsuperscript{42} These moron elements are predicted to encode toxin-antitoxin pairs, superinfection exclusion proteins, restriction endonucleases, and virulence factors. In addition to morons being newly identified through bacterial and phage genome sequencing efforts, there are likely previously unrecognized moron elements in sequenced and characterized phage genomes. A study examining the expression of unannotated small open reading frames (< 50 amino acids) in \textit{E. coli} identified and characterized a number of previously unrecognized proteins.\textsuperscript{43,44} Like gp15, many of these proteins were found to localize to the bacterial cell envelope, and a number of these were found to play a role in modifying the bacterial cell in response to stress conditions such as changes in osmolarity and pH.\textsuperscript{44} It is possible that a number of these small proteins may also be located within prophage genomes, but were previously not annotated due to their small size.

To conclude, it is now clear that morons are “smarter” than they originally appeared to be. First recognized merely as add-ons to phage genomes, we now realize that these frequently occurring genetic elements improve the fitness of their bacterial hosts in a wide variety of ways.

References

1. Human Microbiome Project Consortium. A framework for human microbiome research. Nature 2012; 486:215-21; PMID:22696910; http://dx.doi.org/10.1038/nature11209.
2. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature 2012; 486:207-14; PMID:22696909; http://dx.doi.org/10.1038/nature11234.
3. Ahmed SA, Awoiska J, Baldwin C, Bishop-Lilly KA, Biswas B, Broomall S, et al.; Threat Characterization Consortium. Genomic Comparison of Escherichia coli O104:H4 Isolates from 2009 and 2011 Reveals Plasmid, and Prophage Heterogeneity, Including Shiga Toxin Encoding Phage strx2. PLoS One 2012; 7:e48228; PMID:23133618; http://dx.doi.org/10.1371/journal.pone.0048228.
4. Bergh O, Børsheim KY, Bratbak G, Heldal M. High abundance of viruses found in aquatic environments. Nature 2012; 486:207-14; PMID:22699610; http://dx.doi.org/10.1038/njac.2012.77.
5. Borek E, Ryan A. Lysogenic induction. Prog Mol Biol Transl Sci 2014; 136:249-300; PMID:2467058; http://dx.doi.org/10.1016/6007-6603(08)60105-1.
6. Choi J, Kotay SM, Goel R. Various physico-chemical stress factors cause phage induction in \textit{Nitrosospira} multiformis 25196—an ammonia oxidizing bacteria. Water Res 2010; 44:4550-8; PMID:20630557; http://dx.doi.org/10.1016/j.watres.2010.04.040.
7. Suttle CA. Viruses in the sea. Nature 2005; 437:356-400; PMID:15353570; http://dx.doi.org/10.1038/nature04160.
8. Casjens S. Prophages and bacterial genomics: what have we learned so far? Mol Microbiol 2003; 49:277-300; PMID:12886937; http://dx.doi.org/10.1046/j.1365-2958.2003.03580.x.
9. Borek E, Ryan A. Lysozyme induction. Prog Nucleic Acid Res Mol Biol 1973; 13:249-300; PMID:4267058; http://dx.doi.org/10.1016/6007-6603(08)60105-1.
10. Choi J, Kotay SM, Goel R. Various physico-chemical stress factors cause phage induction in \textit{Nitrosospira} multiformis 25196—an ammonia oxidizing bacteria. Water Res 2010; 44:4550-8; PMID:20630557; http://dx.doi.org/10.1016/j.watres.2010.04.040.
11. Banks DJ, Lei B, Musser JM. Phage induction and expression of prophage-encoded virulence factors in group \textit{A Streptococcus} serotype M3 strain MGA3355. Infect Immun 2003; 71:7079-86; PMID:14638798; http://dx.doi.org/10.1128/IAI.71.12.7079-7086.2003.
12. Maderá C, García P, Rodríguez A, Martínez KL. The bacteriophage HK97 gp15 moron element encodes a novel superinfection exclusion protein. J Virology 2012; 194:5012-9; PMID:22797755; http://dx.doi.org/10.1128/JVI.00843-12.
13. Uc-Mass A, Loeza EJ, de la Garza M, Guarnéros G, Hernández-Sánchez J, Kamayama L. An orthologue of the cor gene is involved in the exclusion of temperate lambda phages. Evidence that Cor inactivates FhuA receptor functions. Virology 2004; 329:425-33; PMID:15188820; http://dx.doi.org/10.1016/j.virol.2004.09.005.
14. Harkness RE, Oischlager T. The biology of colicin M. FEMS Microbiol Rev 1991; 8:27-41; PMID:1931137; http://dx.doi.org/10.1128/IAI.71.12.7079-7086.2003.
15. Gisalt P, Charnay C, Cornillet C, Bouchier T, Barrios C, et al. The bacterial type III secretion system was discovered to be encoded in phage moron elements. For example, the \textit{S. enterica} Gifsy-1 phage encodes GogB, which is secreted by the type III secretion system into mammalian cells where it reduces the host inflammatory response that is designed to help clear the \textit{S. enterica} infection, thereby helping the bacteria establish a chronic infection. A second example is provided by the NleA-like type III effector protein of the shiga-toxin encoding phage BP-4795 which increases virulence in mouse models. Since the effector molecules are only useful when they can be secreted and phages do not encode the secretion apparatus, the genes encoding these effector proteins are likely bacterial in origin.

As the number of bacterial and phage genome sequences increases, the number and diversity of described moron elements also multiplies. In a recent study of \textit{ Burkholderia pseudomallei} phages, 17 unique moron-encoded protein sequences were identified, with some phages encoding as many as 12 different morons. These moron elements are predicted to encode toxin-antitoxin pairs, superinfection exclusion proteins, restriction endonucleases, and virulence factors. In addition to morons being newly identified through bacterial and phage genome sequencing efforts, there are likely previously unrecognized moron elements in sequenced and characterized phage genomes. A study examining the expression of unannotated small open reading frames (< 50 amino acids) in \textit{E. coli} identified and characterized a number of previously unrecognized proteins. Like gp15, many of these proteins were found to localize to the bacterial cell envelope, and a number of these were found to play a role in modifying the bacterial cell in response to stress conditions such as changes in osmolarity and pH. It is possible that a number of these small proteins may also be located within prophage genomes, but were previously not annotated due to their small size.

To conclude, it is now clear that morons are “smarter” than they originally appeared to be. First recognized merely as add-ons to phage genomes, we now realize that these frequently occurring genetic elements improve the fitness of their bacterial hosts in a wide variety of ways.
20. Mirol S, Rabes W, Rohde M, Stender S, Tschäpe H, Rissmann H, et al. Isolation of a temperate bacteriophage encoding the type III effector protein Sopf: from an epidemic Salmonella typhimurium strain. Proc Natl Acad Sci U S A 1999; 96:8484-50; PMID:10449782; http://dx.doi.org/10.1073/pnas.96.17.9845.

21. Hard WD, Chen LM, Schuebel KE, Bustelo XR, Galán JES. S. typhimurium encodes an activator of Rho GTPases that induces membrane ruffling and nuclear responses in host cells. Cell 1998; 93:815-26; PMID:9630225; http://dx.doi.org/10.1016/S0092-8674(00)81442-7.

22. Wood MW, Rosqvist R, Mullan PB, Edwards MH, Galovy EE. Sopf, a secreted protein of Salmonella dublin, is translocated into the target eukaryotic cell via a sip-dependent mechanism and promotes bacterial entry. Mol Microbiol 1996; 22:327-38; PMID:8930917; http://dx.doi.org/10.1111/j.1365-2958.1996.00116.x.

23. Figueroa-Bossi N, Bossi L. Inducible prophages contribute to Salmonella virulence in mice. Mol Microbiol 1999; 33:167-76; PMID:10411733; http://dx.doi.org/10.1046/j.1365-2958.1999.01461.x.

24. Barondess JJ, Beckwith J. bor gene of phage lambda, 21. Hardt WD, Chen LM, Schuebel KE, Bustelo XR, Galán JES. S. typhimurium encodes an activator of Rho GTPases that induces membrane ruffling and nuclear responses in host cells. Cell 1998; 93:815-26; PMID:9630225; http://dx.doi.org/10.1073/pnas.96.17.9845.

25. Winkler KC, de Waart J, Grootsen C. Lysogenic conversion of staphylokinase. Microbes Infect 2005; 7:476-83; PMID:15993432; http://dx.doi.org/10.1128/JB.00873-09.

26. Coleman DC, Sullivan DJ, Russell RJ, Arbuthnott JP, Carey BF, Pomeroy HM. Staphylococcus aureus encoding a lipopolysaccharide O-antigen acetylase maps adjacent to the integrase-encoding gene on the genome of Shigella flexneri bacteriophage Sf6. Gene 1991; 107:43-52; PMID:1720755; http://dx.doi.org/10.1126/science.1.365-2958.1996.00116.x.

27. Rooijakkers SH, van Wamel WJ, Ruyken M, van Kessel KP, van Strijp JA. Anti-opsonic properties of their monosaccharide constituents. Biochem J 1966; 98:903-8; PMID:5913535.

28. Ronning CM, Lousada L, Brinkac L, Inman J, Ulrich RL, Schell M, et al. Genetic and phenotypic diversity in Burkholderia: contributions by prophage and phage-like elements. BMC Microbiol 2010; 10:202; PMID:20667135; http://dx.doi.org/10.1186/1471-2180-10-202.

29. O’Brien AD, Newland JW, Miller SF, Holmes RK, Smith HW, Formal SB. Shiga-like toxin-encoding phages from Escherichia coli strains that cause hemorrhagic colitis or infantile diarrhea. Science 1984; 226:694-6; PMID:6387911; http://dx.doi.org/10.1126/science.6387911.

30. Freeman VJ. Studies on the virulence of bacteriophage-infected strains of Corynebacterium diphtheriae. J Bacteriol 1951; 61:675-88; PMID:14850426.

31. Waldor MK, Mekalanos JJ. Lysogenic conversion by a filamentous phage encoding cholera toxin. Science 1996; 272:1910-14; PMID:8658163; http://dx.doi.org/10.1126/science.272.5270.1910.

32. Clark CA, Belrame J, Manning PA. The oac gene encoding a lipopolysaccharide O-antigen acetylase contributes to virulence in Burkholderia: contributions by prophage and phage-like elements. BMC Microbiol 2010; 10:202; PMID:20667135; http://dx.doi.org/10.1186/1471-2180-10-202.

33. Simmons DA. The immunochrometry of Shigella flexneri lipopolysaccharides. A quantitative analysis of their monosaccharide constituents. Biochem J 1966; 98:903-8; PMID:5913535.

34. Lindberg AA, Wollin R, Gemski P, Wohlhieter JA. Three-component-mediated serotype conversion in Pseudomonas aeruginosa by bacteriophage D3. J Bacteriol 1983; 155:203-12; PMID:690794.

35. Kuzio J, Kropinski AM. O-antigen conversion in Pseudomonas aeruginosa PAO1 by bacteriophage D3. J Bacteriol 1983; 155:203-12; PMID:690794.

36. Newton GJ, Daniels C, Burrows LL, Kropinski AM, Clarke AJ, Lam JS. Three-component-mediated serotype conversion in Pseudomonas aeruginosa by bacteriophage D3. J Bacteriol 1983; 155:203-12; PMID:690794.

37. Burrows LL, Chatter DA, Lam JS. Molecular characterisation of the Pseudomonas aeruginosa serotype O5 (PA01) B-band lipopolysaccharide gene cluster. Mol Microbiol 1996; 22:481-95; PMID:8939432; http://dx.doi.org/10.1126/science.1.365-2958.1996.00116.x.

38. Coombes BK, Wickham ME, Brown NF, Lemire S, Bossi L, Hsiao WW, et al. Genetic and molecular analysis of GogB, a phage-encoded type III translocated virulence factor of enterohaemorrhagic Escherichia coli: O57:H7. Mol Microbiol 2004; 51:1233-49; PMID:14982621; http://dx.doi.org/10.1109/1471-2180-10-202.

39. Hemm MR, Paul BJ, Miranda-Rios J, Zhang A, Solanazad N, Storz G. Small stress response proteins in Escherichia coli: proteins missed by classical proteomic studies. J Bacteriol 2010; 192:46-58; PMID:19734316; http://dx.doi.org/10.1128/JB.00873-09.

40. Hobbs EC, Astarita JL, Storz G. Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in Escherichia coli: analysis of a bar-coded mutant collection. J Bacteriol 2010; 192:59-67; PMID:19734312; http://dx.doi.org/10.1128/JB.00873-09.

41. Ronning CM, Lousada L, Brinkac L, Inman J, Ulrich RL, Schell M, et al. Genetic and phenotypic diversity in Burkholderia: contributions by prophage and phage-like elements. BMC Microbiol 2010; 10:202; PMID:20667135; http://dx.doi.org/10.1186/1471-2180-10-202.

42. Hobbs EC, Astarita JL, Storz G. Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in Escherichia coli: analysis of a bar-coded mutant collection. J Bacteriol 2010; 192:59-67; PMID:19734312; http://dx.doi.org/10.1128/JB.00873-09.

43. Ronning CM, Lousada L, Brinkac L, Inman J, Ulrich RL, Schell M, et al. Genetic and phenotypic diversity in Burkholderia: contributions by prophage and phage-like elements. BMC Microbiol 2010; 10:202; PMID:20667135; http://dx.doi.org/10.1186/1471-2180-10-202.

44. Hobbs EC, Astarita JL, Storz G. Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in Escherichia coli: analysis of a bar-coded mutant collection. J Bacteriol 2010; 192:59-67; PMID:19734312; http://dx.doi.org/10.1128/JB.00873-09.