Molecular Plant Pathology

MOLECULAR PLANT PATHOLOGY (2016) 17(5), 651–653 DOI: 10.1111/mpp.12380

Opinion Piece

The influence of funding sources on the scientific method

LAURA H. OKAGAKI AND RALPH A. DEAN*

Center for Integrated Fungal Research, Department of Plant Pathology, North Carolina State University, Raleigh, 27606, NC, USA

INTRODUCTION

Funding for scientific research in the USA has become harder to obtain, especially for discovery-based science, requiring the analysis of large datasets to find patterns and correlations. It is through such analysis that new hypotheses are formulated and dogma refuted, particularly as it pertains to science in the '-omics' era. The battle for funding has led to a change in the types of experiment that are being proposed by scientists. Grant panels are most commonly interested in proposals that are focused on solving problems, such as novel targets to prevent pathogen infection, or are guaranteed to be successful in the short term. Although it is understandable that grant panels would want to ensure that the money spent on projects will yield immediate results, targeting of virulence factors and drug targets has led to a narrowed view of pathogens and pathogenesis.

How does funding influence scientific ideas and hypotheses? It is common to design studies to examine genes already associated with virulence. Previous studies are the backbone of grant writing and help to solidify the logic of scientific studies. However, focusing on previously studied virulence traits and genes narrows our view of the biological processes that occur between host and pathogen. Experiments to find novel genes associated with virulence are often discouraged in grant writing, and such 'fishing expeditions' typically account for a very small proportion of proposed experiments. With the advent of novel algorithms for large-scale *in silico* analysis of genes and genomes, it is possible to derive a great deal of information from RNAseq, genome sequencing and proteomics. Such large-scale analyses go beyond 'fishing expeditions' and have shed light on biological functions that may or may not be involved in pathogenesis. Recent developments in bioinformatics have allowed scientists to test whether or not hypotheses derived from traditional experiments can be extrapolated to the whole genome. In addition, '-omics' studies can provide data that can be mined by an entire field for many years. Thus, large-scale bioinformatics and '-omics' projects should be viewed as a 'good investment' during grant funding.

* *Correspondence*: Email: radean2@ncsu.edu

TWO-SPEED GENOME EVOLUTION

Hypothesis-driven research has been key to grant writing and successful funding acquisition. However, a lack of large-scale analysis has skewed the research to a point at which it is influencing global hypotheses. For example, the two-speed genome evolution hypothesis in fungal plant pathogens, in which virulence genes are considered to be undergoing faster evolution than other genes, has grown from a hypothesis of how one small set of genes is evolving, to a global hypothesis, and almost dogma in the field.

In fungal plant pathogens, effector proteins have been implicated in early host–pathogen interactions that result in the evasion of the host immune response. Effector proteins are secreted proteins less than 250 amino acids in length and are typically expressed during host infection. Often, effectors are recognized by the host plant's innate immune system to activate host defences (reviewed in Jones and Dangl, 2006). Well-studied examples of effector proteins include AVR-Pita and AVR-Pitz-t in *Magnaporthe orzyae*, which are expressed during the early invasion of host cells (reviewed in Liu *et al.*, 2013; Petre and Kamoun, 2014; Stergiopoulos and de Wit, 2009; Yasuda *et al.*, 2008). *Magnaporthe oryzae* is a hemibiotrophic pathogen, requiring live plant tissue during the early stages of infection. Thus, *M. oryzae* has developed mechanisms to inhibit host cell death caused by activation of the immune system. For example, AVR-Piz-t targets the host ubiquitination system by suppressing the E3 ligase APIP6, which inhibits the ability of the host to induce an immune response and subsequent host cell death (Li *et al.*, 2009; Park *et al.*, 2012). However, the host plant is not left completely defenceless. The ever-adapting immune system will eventually recognize the fungal effector proteins, resulting in pathogen recognition and eventual resistance (Jones and Dangl, 2006). The pathogen must adapt to the host's new defences by altering the effector proteins that are recognized by the host. The idea of an 'arms race' between the host and pathogen is seen in many species of microbial pathogens and their hosts. In the malaria parasite *Plasmodium falciparum*, the outer membrane protein, MSP, undergoes rapid changes in order to evade recognition by antibodies, macrophages and neutrophils (Pearce *et al.*, 2004).

The 'Zig-Zag model' of pathogen and host evolution encompasses the idea that, as the host evolves to recognize the

pathogen, the pathogen must, in turn, evolve to evade the host immune response (Jones and Dangl, 2006). The selection pressure placed on the pathogen by the host immune response results in some genes undergoing more rapid evolution than others. The idea that virulence genes are undergoing more rapid evolution than other genes has become popular in phytopathogenic fungal and oomycete research, and has developed into the concept of a two-speed genome.

Hypothesized mechanisms for the two-speed genome include the proximity to repetitive elements, such as retrotransposons and DNA transposons, that are commonly found littering fungal genomes (Dhillon *et al.*, 2014; Rehmeyer *et al.*, 2006; Santana *et al.*, 2014; Starnes *et al.*, 2012; Wöstemeyer and Kreibich, 2002). Evolution of the sexual cycle in other fungi, including *Cryptococcus* species, has been attributed to high concentrations of repetitive sequence and inverted repeats found at the sex locus (Idnurm *et al.*, 2005). With *M. oryzae*, hundreds of copies of the retrotransposon, Maggy, can be found throughout the genome (Starnes *et al.*, 2012). In addition, Maggy has been shown to induce repeat induced point mutations (RIP), in which G–C pairings are mutated to A–T pairings, resulting in A–T-enriched regions (Ikeda *et al.*, 2002). Thus, RIP may contribute to increased mutation rate and evolutionary speed in genes that are nearby (Ikeda *et al.*, 2002).

HOW FUNDING INFLUENCES THE TWO-SPEED GENOME EVOLUTION HYPOTHESIS

There are critical shortcomings to the two-speed genome evolution hypothesis that could have been addressed with genome-wide '-omics' analyses as algorithms were developed, when coupled with large-scale screens for novel virulence factors. First, the determination of two speeds of evolution requires the comparison of two groups: virulence genes and genes not associated with virulence. How do we determine what is and what is not a virulence gene? In most fungal genomes, annotated genes are often homologues for conserved genes found in model organisms, such as *Saccharomyces cerevisiae*, or other fungal pathogens. Studies that focus on known homologues typically leave out unannotated genes or genes that are unique to a specific species, which are numerous in fungal genomes. Molecular techniques can be used to identify novel and unique virulence genes through mutant screens, but such large-scale screening is often limited by technical problems and cost. Most are far from exhaustive (such as Cai *et al.*, 2013; Huser *et al.*, 2009; Yu *et al.*, 2015) with a few exceptions (Betts *et al.*, 2007; Chen *et al.*, 2011; Jeon *et al.*, 2007). Many mutant screens are funded by alternative sources, such as start-up funds or small internal grants, rather than through external public funding agencies, because hypothesis-driven experiments are funded before 'fishing expeditions'. These small grants are often inadequate to cover the personnel and materials costs for large screens.

Second, there is a grey area in defining whether a gene function is or is not involved in virulence. An inability to determine the contribution of a gene to pathogenesis eliminates a large portion of unstudied genes from the two-speed genome evolution hypothesis. For example, what attribution should be given to genes that are needed for proper nutrient acquisition and metabolism during infection? Are these grey area genes also undergoing selection pressure by the host environment? Are they considered as virulence genes? When examining non-synonymous/synonymous mutation rates (dN/dS) with algorithms, such as phylogenetic analysis by maximum likelihood (PAML) (Yang, 2007), researchers can identify rates of evolution and identify genes that may be under diversifying or purifying selection. However, there is no clear delineation between genes undergoing diversifying selection and those that are not; rather, there is a wide range, or gradient, of dN/dS rates. Pairing large-scale bioinformatics work with traditional bench work could result in better annotation and functional analysis of the unstudied genes.

New computational methods for the examination of evolutionary rates as well as the identification of repetitive elements in fungal genomes can be used to examine the two-speed genome hypothesis from the top down. Whole-genome analysis and comparisons between closely related species can identify genes that are under diversifying selection and their proximity to repetitive elements. Using bioinformatics methods, researchers can remove the bias seen in traditional and targeted studies. Such global analyses should reveal whether or not there are certain classes of proteins (such as small secreted proteins) that are undergoing diversifying selection, and whether, indeed, they are associated with repetitive elements, across a variety of fungal species. Although bioinformatics will not definitively identify genes that are involved in virulence, computational approaches can be used to form the basis upon which to build studies at the benchtop. Scientists must be aware of the weaknesses in both bioinformatics and in bench work when developing hypotheses; however, both methods can be used to strengthen the science when used together.

TWO-PRONGED APPROACH TO RESEARCH

To prevent bias when developing hypotheses, a two-pronged approach to research needs to be implemented in laboratories that focus solely on bench work. Blending bioinformatics with bench research can be achieved by multiple methods. First, staffing bioinformaticists and computation biologists in traditionally bench work-based laboratories and departments is mutually beneficial. A full-time bioinformaticist can target a project in a top-down manner, allowing laboratories to see overall trends in genome, transcriptome and proteome, whereas the bench scientist can focus on the application of key finds from large datasets

and the biological relevance of trends found in *in silico* analyses. In addition, bioinformaticists working closely with laboratories focused on bench work could benefit from understanding the experimental techniques and limitations in order to develop new algorithms to predict and analyse biological datasets.

Second, biologists should receive additional training in basic bioinformatics techniques. Not only would this benefit the biologists, in that they would be able to run analyses themselves, but they would also have a better understanding of how algorithms analyse large datasets. Thus, biologists trained in bioinformatics could make informed decisions when choosing analyses to use for their work.

CONCLUSIONS

Funding of science requires projects to be relevant or important to society. For researchers who study fungal pathogens, simple examination of interesting biology is no longer sufficient to ensure funding. Instead, researchers rely on projects examining pathogenesis that can result in the treatment of fungal infections. Thus, targeted gene studies that are guaranteed to yield results are favoured over large-scale studies that may find function for both genes involved in virulence and genes not involved in virulence. The development of bioinformatics and *in silico* methods, coupled with functional studies for large-scale genome analysis, now allow for a melding of top-down and traditional targeted approaches to pathogenesis research. Integration of bioinformatics with bench work can be mutually beneficial and should be encouraged in grant writing and funding. Together, wet and dry laboratory work can come together to provide the greatest benefit to the development of scientific hypotheses. A balanced approach can prevent hypotheses that focus on single gene groups, such as the two-speed genome hypothesis, from becoming dogma, when large-scale studies may suggest that the hypothesis cannot be applied globally.

ACKNOWLEDGEMENTS

LHO was supported by a Tri-Institutional Molecular Mycology and Pathogenesis training grant (NIH 5T32-AI052080-09). The College of Agriculture and Life Sciences at North Carolina State University, Raleigh, NC, USA provided support to RAD. The authors have no conflicts of interest to declare.

REFERENCES

Betts, M.F., Tucker, S.L., Galadima, N., Meng, Y., Patel, G., Li, L., Donofrio, N., Floyd, A., Nolin, S., Brown, D., Mandel, M.A., Mitchell, T.K., Xu, J.R., Dean, R.A., Farman, M.L. and Orbach, M.J. (2007) Development of a high throughput transformation system for insertional mutagenesis in *Magnaporthe oryzae*. *Fungal Genet. Biol.* **44**, 1035–1049.

Cai, Z., Li, G., Lin, C., Shi, T., Zhai, L., Chen, Y. and Huang, G. (2013) Identifying pathogenicity genes in the rubber tree anthracnose fungus *Colletotrichum gloeosporioides* through random insertional mutagenesis. *Microbiol. Res.* **168**, 340–350.

Chen, X.-L., Yang, J. and Peng, Y.-L. (2011) Large-scale insertional mutagenesis in *Magnaporthe oryzae* by *Agrobacterium tumefaciens*-mediated transformation. *Methods Mol. Biol.* **722**, 213–224.

Dhillon, B., Gill, N., Hamelin, R.C. and Goodwin, S.B. (2014) The landscape of transposable elements in the finished genome of the fungal wheat pathogen *Mycosphaerella graminicola*. *BMC Genomics*, **15**, 1132.

Huser, A., Takahara, H., Schmalenbach, W. and O'Connell, R. (2009) Discovery of pathogenicity genes in the crucifer anthracnose fungus *Colletotrichum higginsianum*, using random insertional mutagenesis. *Mol. Plant–Microbe Interact.* **22**, 143–156.

Idnurm, A., Bahn, Y.-S., Nielsen, K., Lin, X., Fraser, J.A. and Heitman, J. (2005) Deciphering the model pathogenic fungus *Cryptococcus neoformans*. *Nat. Rev. Microbiol.* **3**, 753–764.

Ikeda, K.I., Nakayashiki, H., Kataoka, T., Tamba, H., Hashimoto, Y., Tosa, Y. and Mayama, S. (2002) Repeat-induced point mutation (RIP) in *Magnaporthe grisea*: implications for its sexual cycle in the natural field context. *Mol. Microbiol.* **45**, 1355–1364.

Jeon, J., Park, S.-Y., Chi, M.-H., Choi, J., Park, J., Rho, H.S., Kim, S., Goh, J., Yoo, S., Choi, J., Park, J.Y., Yi, M., Yang, S., Kwon, M.J., Han, S.S., Kim, B.R., Khang, C.H., Park, B., Lim, S.E., Jung, K., Kong, S., Karunakaran, M., Oh, H.S., Kim, H., Kim, S., Park, J., Kang, S., Choi, W.B., Kang, S. and Lee, Y.H. (2007) Genome-wide functional analysis of pathogenicity genes in the rice blast fungus. *Nat. Genet.* **39**, 561–565.

Jones, J.D.G. and Dangl, J.L. (2006) The plant immune system. *Nature*, **444**, 323–329.

Li, W., Wang, B., Wu, J., Lu, G., Hu, Y., Zhang, X., Zhang, Z., Zhao, Q., Feng, Q., Zhang, H., Wang, Z., Wang, G., Han, B., Wang, Z. and Zhou, B. (2009) The *Magnaporthe oryzae* avirulence gene AvrPiz-t encodes a predicted secreted protein that triggers the immunity in rice mediated by the blast resistance gene Piz-t. *Mol. Plant–Microbe Interact.* **22**, 411–420.

Liu, W., Liu, J., Ning, Y., Ding, B., Wang, X., Wang, Z. and Wang, G.-L. (2013) Recent progress in understanding PAMP- and effector-triggered immunity against the rice blast fungus *Magnaporthe oryzae*. *Mol. Plant*, **6**, 605–620.

Park, C.-H., Chen, S., Shirsekar, G., Zhou, B., Khang, C.H., Songkumarn, P., Afzal, A.J., Ning, Y., Wang, R., Bellizzi, M., Valent, B. and Wang, G.-L. (2012) The *Magnaporthe oryzae* effector AvrPiz-t targets the RING E3 ubiquitin ligase APIP6 to suppress pathogen-associated molecular pattern-triggered immunity in rice on JSTOR. *Plant Cell*, **24**, 4748–4762.

Pearce, J.A., Triglia, T., Hodder, A.N., Jackson, D.C., Cowman, A.F. and Anders, R.F. (2004) *Plasmodium falciparum* merozoite surface protein 6 is a dimorphic antigen. *Infect. Immun.* **72**, 2321–2328.

Petre, B. and Kamoun, S. (2014) How do filamentous pathogens deliver effector proteins into plant cells? *PLoS Biol.* **12**, e1001801.

Rehmeyer, C., Li, W., Kusaba, M., Kim, Y.-S., Brown, D., Staben, C., Dean, R. and Farman, M. (2006) Organization of chromosome ends in the rice blast fungus, *Magnaporthe oryzae*. *Nucleic Acids Res.* **34**, 4685–4701.

Santana, M.F., Silva, J.C., Mizubuti, E.S., Araújo, E.F., Condon, B.J., Turgeon, B.G. and Queiroz, M. V (2014) Characterization and potential evolutionary impact of transposable elements in the genome of *Cochliobolus heterostrophus*. *BMC Genomics*, **15**, 536.

Starnes, J.H., Thornbury, D.W., Novikova, O.S., Rehmeyer, C.J. and Farman, M.L. (2012) Telomere-targeted retrotransposons in the rice blast fungus *Magnaporthe oryzae*: agents of telomere instability. *Genetics*, **191**, 389–406.

Stergiopoulos, I. and de Wit, P.J.G.M. (2009) Fungal effector proteins. *Annu. Rev. Phytopathol.* **47**, 233–263.

Wöstemeyer, J. and Kreibich, A. (2002) Repetitive DNA elements in fungi (Mycota): impact on genomic architecture and evolution. *Curr. Genet.* **41**, 189–198.

Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591.

Yasuda, N., Noguchi, M.T. and Fujita, Y. (2008) Avirulence genes. *Plant Dis.* **92**, 1144–1149.

Yu, M., Yu, J., Hu, J., Huang, L., Wang, Y., Yin, X., Nie, Y., Meng, X., Wang, W. and Liu, Y. (2015) Identification of pathogenicity-related genes in the rice pathogen Ustilaginoidea virens through random insertional mutagenesis. *Fungal Genet. Biol.* **76**, 10–19.