Population Genomic- and Phylogenomic-Enabled Advances to Increase Insight Into Pathogen Biology and Epidemiology

Remco Stam,^{1,†} Pierre Gladieux,² Boris A. Vinatzer,³ Erica M. Goss,^{4,5} Neha Potnis,⁶ Thierry Candresse,⁷ and Marin Talbot Brewer⁸

¹Phytopathology, Technical University Munich, 85354 Freising, Germany

²UMR BGPI, University of Montpellier, INRA, CIRAD, Montpellier SupAgro, 34398 Montpellier, France

³School of Plant and Environmental Sciences, Virginia Tech, Blacksburg, VA 24061, U.S.A.

⁴Department of Plant Pathology, University of Florida, Gainesville, FL 32611, U.S.A.

⁵Emerging Pathogens Institute, University of Florida, Gainesville, FL 32611, U.S.A.

⁶Department of Entomology and Plant Pathology, Rouse Life Science Building, Auburn University, Auburn, AL 36849, U.S.A.

⁷University of Bordeaux, INRAE, UMR BFP, Villenave d'Ornon, France

⁸Department of Plant Pathology, University of Georgia, Athens, GA 30602, U.S.A.

Accepted for publication 2 December 2020.

ABSTRACT

Population genetics has been a key discipline in phytopathology for many years. The recent rise in cost-effective, high-throughput DNA sequencing technologies, allows sequencing of dozens, if not hundreds of specimens, turning population genetics into population genomics and opening up new, exciting opportunities as described in this Focus Issue. Without the limitations of genetic markers and the availability of whole

or near whole-genome data, population genomics can give new insights into the biology, evolution and adaptation, and dissemination patterns of plant-associated microbes.

Keywords: fungal pathogens, genomics, microbe-genome sequencing, nematodes, oomycetes, population biology, virology

In this Focus Issue, we present work representing the many different facets of population genomics in the field of phytopathology. The research articles in the issue cover the major types of plant-pathogenic organisms, ranging from viruses and bacteria to fungi, oomycetes, and nematodes. We also present a series of reviews and perspectives on various topics. These range from articles that briefly describe the basic steps of the population genomics analysis pipeline for newcomers to the field to technical reviews on phylogeographic approaches. They also include topical reviews providing an overview of the current status of phytopathology in plant-pathogenic viruses, bacteria, and nematodes, and the impact of human activities on plant health. This article briefly summarizes all articles appearing in this Focus Issue.

Population genetics has been a key discipline in phytopathology for many years. With the advent of cost-effective, high-throughput DNA sequencing technologies, population genetics has morphed into population genomics and opened up new, exciting opportunities. While sequencing the genome of a single representative of a pathogen species was a daunting task 20 years ago, analyzing the genomes of dozens or even hundreds of individuals of a population is possible today. As a result, population genomics can give unprecedented insights into the biology, evolution and adaptation, and dissemination patterns of plant-associated microbes. Biological questions can now be addressed at much higher resolution than with population genetics because the number of markers is no longer limiting. As another major advantage, a population genomic approach allows for exploration of mechanisms behind population processes at the gene level (Everhart et al. 2021).

Population genomics also has direct, practical applications in preventing, or at least curtailing, pathogen dissemination by

supporting early interventions to limit economic damage. For example, identifying dissemination routes using genomic epidemiology can inform practices to limit pathogen spread, or identifying the core virulence genes of a species can help target plant breeding for disease resistance.

While population genomics is often limited to a single population or species, the inclusion of multiple genomes of multiple species using phylogenomics and comparative genomics can reveal genes that form the basis of phenotypic differences between species to gain new insights into fundamental questions of host range evolution, adaptation, and other phenotypes.

In this Focus Issue, we present state-of-the-art work representing the many different facets of population genomics in the field of phytopathology. The articles cover the major types of plant-pathogenic organisms, ranging from viruses and bacteria to fungi, oomycetes, and nematodes. Some of the articles are prime examples of recent advances in technology and all articles address exciting questions regarding basic pathogen biology, questions about migration and epidemiology, or both. We discuss all of them briefly below. The issue opens with a series of reviews and perspectives on how population genomic approaches have been used to track disease emergence and spread, to study the impact of humans on plant pathogen evolution and adaptation, and to understand better the evolution and emergence of plant-pathogenic fungi, oomycetes, viruses, bacteria, and nematodes.

REVIEWS HIGHLIGHTING THE MANY FACETS OF POPULATION GENOMICS

In the first article in this issue, Everhart et al. (2021) provide an overview of several main aspects in population genomics, using examples drawn from studies on filamentous plant pathogens. They discuss different types of population genomics-inspired research questions, discuss important considerations in the various steps in the population genomic workflow, and highlight some of the hurdles and technical challenges that need to be born in mind when conducting population genomic analyses. Rasmussen and

[†]Corresponding author: R. Stam; stam@wzw.tum.de

The author(s) declare no conflict of interest.

Grünwald (2021) discuss phylogeographic approaches useful in tracking plant disease emergence and pathogen spread. The abundance of genomic data now available for hundreds of plant pathogens has provided greater resolution for studying pathogen evolution. The authors introduce phylogeographic theory, discuss methodological considerations and assumptions, and present examples of studies where phylogeographic approaches were used to study plant pathogens. The impact of human activities on plant health are discussed by Hessenauer et al. (2021). The authors summarize studies illuminating how genome sequencing and population genomics provide a better understanding of plant pathogen emergence in agroecosystems. They provide solutions on how to minimize the impact of anthropogenic activities on plant diseases and reduce the occurrence of global epidemics. To address risks of virus emergence, McLeish et al. (2021) analyze the use of high-throughput technologies, of both traditional genetic methods and of new computational tools for big data, to monitor changes in plant virus populations from single cells to communities and ecosystems. They present an overview of studying virus populations across scales, including evolution within an individual host, transmission to new hosts, and factors outside the host affecting viral populations. Current genomic methods for addressing questions on bacterial populations are summarized by Straub et al. (2021). They focus on recent population genomic studies covering topics of critical importance to plant pathology, including tracking pathogen transmission pathways and origins during bacterial outbreaks, identifying the genetic basis of host specificity and virulence, and determining how bacterial populations are shaped by agricultural practices. And lastly, Montarry et al. (2021) provide a comprehensive overview of the advances made in population genomics of nematodes in the past 5 years. In their review, the authors provide information on current methods and resources of interest to nematologists who may be considering population genomic approaches or genome sequencing studies. Furthermore, they highlight recent papers where these approaches were used to identify the molecular basis of pathogenicity, understand genomic variation, and study the evolution of plant-parasitic nematodes.

TECHNICAL ADVANCES

This Focus Issue highlights the opportunities gained from recent technological advances and how they can be exploited to generate relevant data on plant pathogens. Weldon et al. (2021) developed transcriptome-derived amplicon sequencing (AmpSeq) markers to elucidate the population structure of the hop powdery mildew pathogen and track virulent populations that overcome the resistance conferred by the R6 gene deployed in widely planted hop cultivars. Their genotyping strategy has the advantage of being high-throughput, inexpensive, and widely applicable, also to obligate biotrophs. We are also delighted to see major advances in nematode population genomics. Until recently, as highlighted in the review by Montarry et al. (2021), nematode population genetics or genomics had to be performed on large number of individual worms at once, for the simple reason that a single nematode did not contain sufficient DNA or RNA to work with. Advances in single-cell sequencing now allow the sequencing of individual nematodes. Ste-Croix et al. (2021) use a nematode-adapted single-cell RNA-seq approach to identify genes in the soybean cyst nematode that are potentially involved in resistance breakdown in various soybean cultivars. An innovative combination of marker-based genotyping and whole-genome sequencing of select individuals was used by Pordel et al. (2021) to move beyond genotype classifications of the fungus *Pyricularia oryzae* to identify essential genes involved in adaptation processes. Finally, Vogel et al. (2021) make innovative use of genotyping by sequencing (GBS); by employing one single nucleotide polymorphism (SNP) every 600 bp and by genotyping 252 isolates, the authors were able to robustly quantify aneuploidy and polyploidy in populations of the oomycete *Phytophthora capsici*.

POPULATION DIVERSITY AND BIOLOGY

Several of the studies on filamentous pathogens provide answers to fundamental biological questions by using GBS. In order to understand the biology of two closely related maize pathogens whose distinctness was a matter of debate, Fumero et al. (2021) studied populations of *Fusarium subglutinans* and *F. temperatum* isolated in Argentina. They found that the two species have distinct biological properties, with more evidence for sexual recombination in *F. subglutinans*. The species split seems to predate the domestication of maize, but subsequent between-species gene flow has occurred in both directions, indicating that the exact species boundaries merit further study. In the paper by Vogel et al. (2021), we learn about an investigation of the genetic determinants of mating type and fungicide sensitivity in *Phytophthora capsici* using phenotypic assays and GBS. In addition to identifying several candidate genes for mefenoxam sensitivity and for mating type, their study also shows that limited gene flow occurs among *Phytophthora capsici* populations on different farms and that a substantial proportion of isolates are either aneuploids or genome-wide polyploids. Bautista Jalon et al. (2021) asked the question whether *Verticillium dahliae* from symptomatic and asymptomatic hosts differ in their genetic make-up. They used GBS with 26,934 SNPs across 126 isolates and found a highly clonal structure. Isolates from unrelated symptomatic and asymptomatic hosts often belonged to the same lineage. Thus, asymptomatic weeds and rotational hosts may be potential reservoirs for *V. dahliae*. This example shows that studies on genetic differences of co-occurring pathogen populations can shed light on their basic biology, as well as provide pointers to inform management practices.

High-throughput sequencing of viral pathogens also sheds light on their biology and yields insights in transmission pathways. The complete genome sequences of seven *Cytorhabdovirus* isolates from China are reported by Zhang et al. (2021b). These sequences allow to clarify their taxonomy and to ascribe them to a single species infecting a range of hosts in China and South America. Remarkably, the isolates from citrus have nonfunctional glycoprotein (G) and P4 genes as a consequence of isolate-specific mutations, indicative of the host-specific selection of such defective variants in this host species. Since the G protein is required for cell recognition and entry into the arthropod vectors, this suggests an original evolutionary scenario in which propagation of the citrus isolates is through vegetative propagation of the citrus hosts while isolates in other hosts likely retain their ability to be transmitted by vectors.

Population genomics has also accelerated studies investigating the evolutionary and epidemiological processes associated with disease emergence and spread by characterizing signature footprints in pathogen genomes. When pathogen populations are clonal or partially clonal, genomic characterization of few strains can provide insight at a population scale. For example, the late blight pathogen, *Phytophthora infestans*, is notorious for overcoming major gene resistance in potato. In China, “supervirulent” populations of *Phytophthora infestans* that can overcome resistance across the R1-R11 differential set have increased in frequency. Zhang et al. (2021a) used the genome and transcriptome of two field isolates representing these emerging populations to determine the status of *Avr* genes and their expression and the mechanisms by which these isolates evade recognition by host *R* proteins. They found multiple strategies of evasion, including gene deletion and nucleotide polymorphism, and transcriptional silencing and effector-encoding genes show distinct transcription between the isolates. Importantly, the authors showed that several *R* genes are still effective against these isolates. These might recognize particular RXLR effectors that have been maintained by the pathogen.

The power of comparative genomics was employed by Ares et al. (2021) to infer pathogen populations associated with disease outbreaks. The kiwifruit pathogen, *Pseudomonas syringae* pv. *actinidiae*, is subject to emergency measures in Europe. It

represents a classic example where introduction of a novel pathogen in an area is characterized by clonal expansion followed by diversification into genetically distinct lineages. The authors present draft genomes of *Pseudomonas syringae* pv. *actinidiae* strains representing distinct genetic lineages based on BOX profiles, with three strains simultaneously isolated from a single plant. Such genome resources provide opportunities to identify genomic signatures responsible for divergence and also allow us to pinpoint the risk factors associated with disease emergence.

In order to address fundamental questions about the gain of virulence of the soybean cyst nematode *Heterodera glycines* on various soybean cultivars, Ste-Croix et al. (2021) investigated gene expression changes of nematodes originating from different host plants to understand host adaptation. Using single-cell RNA-seq, the authors showed that also in nematodes several candidate effector genes may be associated with a gain of virulence. Interestingly, results further showed that the genetic diversity between virulent and avirulent strains depends on the genetic background of the host, with smaller variation on soybean line PI 88788 than on line Peking, suggesting that the genetic background of the host populations leads to distinct pathways of virulence evolution.

EPIDEMIOLOGY, MIGRATION, AND EVOLUTION

Several studies in this Focus Issue are centered around defining the origin or migration of specific plant pathogens. By studying pathogen population structures and placing them in an anthropological or agricultural context, this can also lead to valuable insights to improve pathogen management in agriculture or in natural reserves. For example, new tree diseases continue to emerge around the globe and threaten the health of forests and fruit and nut tree crops. Many of these diseases have been shown to be caused by, or are at least associated with, bacteria including strains of the *Pseudomonas syringae* species complex, *Xanthomonas* species, and members of the family *Pectobacteriaceae*. Bakhshi ganje et al. (2021) used a plethora of phylogenomic, comparative genomic, and pan-genomic approaches to gain new insights into members of the *Pectobacteriaceae* genus *Brenneria*, which includes several tree-pathogenic species. Among many other findings, the authors determined which pathogenicity and virulence factors separate the genus *Brenneria* from other *Pectobacteriaceae* and which gene families distinguish *Brenneria* species from each other and may thus be involved in the adaptation to different hosts. The emergence of an oomycete tree pathogen was the focus of a study by Tabima et al. (2021). *Phytophthora pluvialis*, an invasive forest pathogen in New Zealand, was only recently also found in the U.S. Pacific Northwest. Previous studies indicated two introductions of the pathogen to New Zealand, while GBS analysis in the current study, based on 1543 SNPs, show that a single introduction can explain the emergence of *Phytophthora pluvialis* in New Zealand, including isolates clustering with isolates from the United States. The genetic structure of the New Zealand and U.S. populations are consistent with clonal expansion even though *Phytophthora pluvialis* is a homothallic species, which is interesting given similar findings for invasive heterothallic *Phytophthora* species even when both mating types are present. *Phytophthora pluvialis* has only been isolated from a single host in the United States, but in New Zealand it has expanded to multiple conifer species since its first identification in 2008. Focusing on another filamentous forest pathogen, Capron et al. (2021) show that populations of the *Dothistroma* needle blight pathogen in western Canada are structured into long-diverged lineages. Their findings increase our knowledge of the biogeography of fungi, and in parallel provide crucial information to prevent human-driven secondary contacts from causing new needle blight epidemics on wild and planted pine trees.

The origins and migration paths of boxwood blight pathogens were investigated by LeBlanc et al. (2021) using whole-genome resequencing data. Comparisons of historical scenarios using approximate Bayesian computations provided evidence for

intercontinental and multiple introductions into the United States of the most widespread of the two boxwood blight pathogens, suggesting that future introductions represent a significant risk that must be considered. These studies reveal the power of large-scale genomics when investigating pathosystems, in particular those that are less amenable to experimentation, such as tree pathogens and their woody hosts.

Historical migration and adaptational routes of new and emerging crop pathogens can also be analyzed using population genomic techniques. Pordel et al. (2021) studied the emergence of blast disease on maize in Iran, caused by *Pyricularia oryzae*. The disease was first documented in Iran in the north of the country in 2012. In 2016, the disease was also detected in the south. *Pyricularia oryzae* was also epidemic on the weed barnyard grass (*Echinochloa* spp.), growing in and surrounding the maize fields. The authors used a combination of traditional simple sequence repeat markers and whole-genome sequencing of multiple isolates, including a historical strain from West Africa, and concluded that strains causing epidemics on maize and on barnyard grass are genetically different. The maize-infecting isolates originate from independent host range expansions from barnyard grass-infecting strains. Findings like these beg the question how widespread and risky such host expansion events are in plant pathogens and whether they can be prevented at all. In another study elucidating pathogen isolates from wild and cultivated hosts, Weldon et al. (2021) show that isolates of the hop powdery mildew pathogen *Podosphaera macularis* from U.S. commercial hop yards belong to one population closely related to isolates from the U.K., while isolates from wild hop in the U.S. group with isolates from Eastern Europe, possibly suggesting trade-associated migration.

The two virology research papers in this Focus Issue also study aspects of pathogen evolution and migration. The above-mentioned study by Zhang et al. (2021b) concluded that the citrus isolates of *Cytorhabdovirus* are probably derived from isolates from other hosts. In a joint paper involving a consortium of laboratories, Fuentes et al. (2021) provide a detailed, global-scale vision of the diversity of potato virus A (PVA). Full-genome characterization using high-throughput sequencing of a large number of isolates allowed the discovery of a previously unknown lineage from the Andean region, the center of origin of the potato crop. The relative dating of the PVA phylogenetic tree compared with that of the biologically similar potato virus Y allowed the authors to propose a tentative evolutionary scenario for PVA spanning centuries and supported by several independent historical coincidences.

In an experimental evolution study, Gambhir et al. (2021) shed light on evolution of fungicide resistance, a pressing problem for agriculture and human health. In the field, fungi can be exposed to a range of doses of active ingredient, including sublethal doses. Interestingly, the stress induced by sublethal doses of fungicide may cause genomic instability. To investigate this phenomenon, the authors quantified mutations associated with sublethal doses of four fungicides against *Sclerotinia sclerotiorum*. They found a significant increase in Indels on one strain and a decrease in transposable element insertions across strains, but more generally they found that mutation patterns depended on the genomic background of the strain. These results show why pathogen evolution is perhaps the most vexing problem in disease management.

In conclusion, the research articles, reviews, and perspectives assembled in this Focus Issue are proof of the power of population genomics and phylogenomics in revealing new insights into plant pathogen biology, evolution, and dissemination. The advances we have witnessed in sequencing technologies and bioinformatics have provided the basis for the described discoveries. Continued dedication of phytopathologists to collecting samples, formulating hypotheses, and critically analyzing the genomic data in the context of real-world plant pathology will make it possible to translate the obtained sequences into healthier plants.

ACKNOWLEDGMENTS

We thank Senior Editor Nian Wang for handling an additional manuscript for this Focus Issue and Editor-in-Chief Harald Scherm for support and insights throughout the whole process of putting the issue together.

LITERATURE CITED

- Ares, A., Tacão, M., Figueira, D., Garcia, E., and Costa, J. 2021. Draft genome resources sequences of six *Pseudomonas syringae* pv. *actinidiae* strains isolated from *Actinidia chinensis* var. *deliciosa* leaves in Portugal. *Phytopathology* 111:237-239.
- Bakhshi ganje, M., Mackay, J., Nicolaisen, M., and Shams-Bakhsh, M. 2021. Comparative genomics, pangenome, and phylogenomic analyses of *Brenneria* spp., and delineation of *Brenneria izadpanahii* sp. nov. *Phytopathology* 111:78-95.
- Bautista-Jalon, L. S., Frenkel, O., Tsrer, L., Malcolm, G. M., Gugino, B., Lebiush, S., Hazanovsky, M., Milgroom, M. G., and Jiménez-Gasco, M. M. 2021. Genetic differentiation of *Verticillium dahliae* populations recovered from symptomatic and asymptomatic hosts. *Phytopathology* 111:149-159.
- Capron, A., Feau, N., Heinzelmann, R., Barnes, I., Benowicz, A., Bradshaw, R. E., Dale, A. L., Lewis, K., Owen, T., Reich, R., Ramsfield, T., Woods, A., and Hamelin, R. 2021. Signatures of post-glacial genetic isolation and human-driven migration in the *Dothistroma* needle blight pathogen in western Canada. *Phytopathology* 111:116-127.
- Everhart, S., Gambhir, N., and Stam, R. 2021. Population genomics of filamentous plant pathogens—A brief overview of research questions, approaches, and pitfalls. *Phytopathology* 111:12-22.
- Fuentes, S., Gibbs, A. J., Adams, I. P., Wilson, C., Botermans, M., Fox, A., Kreuze, J., Boonham, N., Kehoe, M. A., and Jones, R. A. C. 2021. Potato virus A isolates from three continents: Their biological properties, phylogenetics, and prehistory. *Phytopathology* 111:217-226.
- Fumero, M. V., Yue, W., Chiotta, M. L., Chulze, S. N., Leslie, J. F., and Toomajian, C. 2021. Divergence and gene flow between *Fusarium subglutinans* and *F. temperatum* isolated from maize in Argentina. *Phytopathology* 111:170-183.
- Gambhir, N., Kamvar, Z. N., Higgins, R., Amaradasa, B. S., and Everhart, S. E. 2021. Spontaneous and fungicide-induced genomic variation in *Sclerotinia sclerotiorum*. *Phytopathology* 111:160-169.
- Hessenauer, P., Feau, N., Gill, U., Schwessinger, B., Brar, G. S., and Hamelin, R. C. 2021. Evolution and adaptation of forest and crop pathogens in the Anthropocene. *Phytopathology* 111:49-67.
- LeBlanc, N., Cubeta, M. A., and Crouch, J. A. 2021. Population genomics trace clonal diversification and intercontinental migration of an emerging fungal pathogen of boxwood. *Phytopathology* 111:184-193.
- McLeish, M. J., Fraile, A., and Garcia-Arenal, F. 2021. Population genomics of plant viruses: The ecology and evolution of virus emergence. *Phytopathology* 111:32-39.
- Montarry, J., Mimee, B., Danchin, E. G. J., Koutsovoulos, G. D., Ste-Croix, D. T., and Grenier, E. 2021. Recent advances in population genomics of plant-parasitic nematodes. *Phytopathology* 111:40-48.
- Pordel, A., Ravel, S., Charriat, F., Gladieux, P., Cros-Arteil, S., Milazzo, J., Adreit, H., Javan-Nikkhah, M., Mirzadi-Gohari, A., Moumeni, A., and Tharreau, D. 2021. Tracing the origin and evolutionary history of *Pyricularia oryzae* infecting maize and barnyard grass. *Phytopathology* 111:128-136.
- Rasmussen, D. A., and Grünwald, N. J. 2021. Phylogeographic approaches to characterize the emergence of plant pathogens. *Phytopathology* 111:68-77.
- Ste-Croix, D. T., St-Marseille, A.-F. G., Lord, E., Belanger, R. R., Brodeur, J., and Mimee, B. 2021. Genomic profiling of virulence in the soybean cyst nematode using single-nematode sequencing. *Phytopathology* 111:137-148.
- Straub, C., Colombi, E., and McCann, H. C. 2021. Population genomics of bacterial plant pathogens. *Phytopathology* 111:23-31.
- Tabima, J. F., Gonen, L., Gómez-Gallego, M., Panda, P., Grünwald, N. J., Hansen, E. M., McDougal, R., Leboldus, J. M., and Williams, N. M. 2021. Molecular phylogenomics and population structure of *Phytophthora pluvialis*. *Phytopathology* 111:108-115.
- Vogel, G., Gore, M. A., and Smart, C. D. 2021. Genome-wide association study in New York *Phytophthora capsici* isolates reveals loci involved in mating type and mefenoxam sensitivity. *Phytopathology* 111:204-216.
- Weldon, W. A., Knaus, B. J., Grünwald, N. J., Havill, J. S., Block, M. H., Gent, D. H., Cadle-Davidson, L. E., and Gadoury, D. M. 2021. Transcriptome-derived amplicon sequencing markers elucidate the U.S. *Podosphaera macularis* population structure across feral and commercial plantings of *Humulus lupulus*. *Phytopathology* 111:194-203.
- Zhang, F., Chen, H., Zhang, X., Gao, C., Huang, J., Li, L., Shen, D., Wang, L., Huang, C., Ye, W., Zheng, X., Wang, Y., Vossen, J. H., and Dong, S. 2021a. Genome analysis of two newly emerged potato late blight isolates sheds light on pathogen adaptation and provides tools for disease management. *Phytopathology* 111:96-107.
- Zhang, S., Huang, A., Zhou, X., Li, Z., Dietzgen, R. G., Zhou, C., and Cao, M. 2021b. Natural defect of a plant rhabdovirus glycoprotein gene: A case study of virus–plant coevolution. *Phytopathology* 111:227-236.