ANALYZING GENETIC VARIATION: WHAT CAN POPULATION GENOMICS TELL US ABOUT THE BIOLOGY AND MANAGEMENT OF OUR FAVORITE FOREST PATHOGENS?

Patrick I. Bennett[1,2*] and Jared M. LeBoldus[2,3]

The central dogma of molecular biology states that deoxyribonucleic acid (DNA) is transcribed to a form of ribonucleic acid known as messenger RNA (mRNA). These mRNA molecules are then translated to a string of amino acids known as a protein. These processes form the foundation of our understanding of the relationships between genotype (the sequence of nucleotides at a given locus) and phenotype (the observable characteristics resulting from the expression of a genotype in a given environment). A DNA molecule consists of four chemical residues (adenine, thymine, cytosine, and guanine) known as nucleotides. All of the variation we observe in nature reflects differences in the composition, structure, and expression of the nucleotide sequences present in DNA. Population genomics provides invaluable tools for investigating the evolution of plant pathogens by analyzing the variation in nucleotide sequences. This enables studies of phenotypic characteristics that may be relevant to managing plant diseases such as pathogenicity, virulence, and host specialization (Grünwald et al. 2016). These tools also allow for the study of basic biological and evolutionary characteristics of plant pathogens including reproduction, migration, and natural selection (Grünwald et al. 2016). The population genomics framework has particular utility for investigating the epidemiology of invasive pathogen populations including introduction event(s), dispersal, colonization history, centers of origin, and novel adaptations in new environments (Barnes et al. 2014, Garbelotto et al. 2013, Goss et al. 2009, Kamvar et al. 2015). This information is essential for developing efficient and effective plant disease management strategies.

The methodology and experimental design for population studies generally follow a framework that includes collecting samples of infected plant tissues, isolating the pathogen (if possible), extracting and purifying DNA, sequencing genomes or molecular markers, and analyzing genetic variation. With recent advancements in DNA sequencing technologies, the molecular markers used for population genomics have evolved from individual gene sequences to whole-genome data. Single nucleotide polymorphisms (SNPs) represent variation in a single nucleotide at a single locus. In population genomic studies that utilize SNPs, each isolate or individual will have a genotype consisting of all SNP loci identified across the genome. These SNP loci can only be identified by comparing individual genomes to a reference genome from a type specimen. Genetic diversity is estimated by analyzing the variation in these SNP genotypes among individuals. The spatial or geographic distribution of this genetic variation within or between pathogen populations is known as population structure.

The invasive fungal pathogen *Cronartium ribicola* causes white pine blister rust, a devastating disease of five-needle pines (*Pinus* spp.) This pathogen is a particularly good case study for demonstrating the utility of population genomics in describing the invasion and colonization history of a non-native forest pathogen. Brar et al. (2015) using SNPs studied *C. ribicola* populations in North America in order to evaluate the impacts of the human-mediated introduction and subsequent colonization on its population genetic structure. The authors sampled 76 sites in the northern United States and Canada and genotyped 1,336

*[1]USDA Forest Service, Forest Health Protection, Northern and Intermountain Regions, Missoula, Montana, USA, *patrick.bennett@usda.gov. [2]Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR, USA. [3]Department of Forest Engineering, Resources, and Management (FERM), Oregon State University, Corvallis, OR, USA*

individual isolates with 31 SNP loci. Their analyses identified two distinct genetic clusters of *C. ribicola* with one cluster represented by isolates collected in northeastern North America and the midwestern United States, and the other represented by isolates collected in the western U.S. and western Canada (Brar et al. 2015). There were differences in diversity and population structure between eastern and western subpopulations that reflected the colonization histories of these regions. For example, there was greater genetic diversity in the eastern subpopulations where multiple introductions of *C. ribicola* are known to have occurred, and less diversity in the western U.S. and Canada subpopulations where one or a few introductions of the pathogen occurred (Brar et al. 2015). They also hypothesized that the absence of host species in the central U.S. and Canada could be a barrier to gene flow between these eastern and western subpopulations (Brar et al. 2015). However, their analyses revealed that some gene flow has occurred possibly due to long distance aerial spore dispersal or anthropogenic movement of infected plant material (Brar et al. 2015). The results of analyses such as these can be valuable for describing basic pathogen biology and epidemiology and can inform land management decisions that rely on knowledge of pathogen dispersal and migration (Brar et al. 2015).

Analyses of variation in SNP genotypes has also been used recently to investigate the dispersal capabilities and epidemiology of the Douglas-fir black stain root disease pathogen, *Leptographium wageneri* var. *pseudotsugae*, which is vectored by insects. The preliminary results from this study are described in Bennett et al. (*these proceedings*) and further demonstrate the utility of population genomics when studying the epidemiology of forest pathogens.

Literature Cited

Barnes, I., Wingfield, M. J., Carbone, I., Kirisits, T., Wingfield, B. D. (2014). Population structure and diversity of an invasive pine needle pathogen reflects anthropogenic activity. *Ecology and Evolution, 4(18), 3642–3661. https://doi.org/10.1002/ece3.1200*

Brar, S., Tsui, C.K.M., Dhillon, B., Bergeron, M.J., Joly, D.L., et al. (2015) Colonization history, host distribution, anthropogenic influence and landscape features shape opulations of white pine blister rust, an invasive alien tree pathogen. *PLOS ONE 10(5): e0127916. https://doi.org/10.1371/journal.pone.0127916*

Garbelotto, M., Guglielmo, F., Mascheretti, S., Croucher, P.J.P., & Gonthier, P. (2013). Population genetic analyses provide insights on the introduction pathway and spread patterns of the North American forest pathogen Heterobasidion irregulare in Italy. *Molecular Ecology, 22(19), 4855–4869. https://doi.org/10.1111/mec.12452*

Goss, E.M., Larsen, M., Chastagner, G.A., Givens, D.R., Grünwald, N.J. (2009). Population genetic analysis infers migration pathways of *Phytophthora ramorum* in US nurseries. *PLoS Pathogens, 5(9), e1000583. https://doi.org/10.1371/journal.ppat.1000583*

Grünwald, N.J., McDonald, B.A., Milgroom, M.G. (2016). Population genomics of fungal and oomycete pathogens. *Annual Review of Phytopathology, 54(1), 323–346. https://doi.org/10.1146/annurev-phyto-080614-115913*

Kamvar, Z.N., Larsen, M.M., Kanaskie, A.M., Hansen, E.M., Grünwald, N.J. (2015). Spatial and temporal analysis of populations of the sudden oak death pathogen in Oregon forests. *Phytopathology, 105(7), 982–989. https://doi.org/10.1094/PHYTO-12-14-0350-FI*