Paving the Way for Healthy Ecosystems, One Bee at a Time

Innovative uses of next-generation sequencing data complement the ecologist's toolbox to analyze, understand, and protect natural communities.

Introduction

Bees pollinate up to 30 percent of the food we consume, including sunflowers, rapeseed, apples, and other fruits and berries. Without these buzzing pollinators, our food supply would suffer. Yet, in the past 50 years populations of bee species have declined in droves, including the domesticated honey bee, which has taken a staggering 50 percent hit. Reduced food sources due to habitat loss from extensive agriculture, urban development, deforestation, and diseases are mainly to blame.

Alexander Keller, PhD, at the University of Würzburg in Germany is attempting to understand bee species and other wild pollinators with novel approaches that might provide useful insights for future conservation of these ecosystem services. He and a network of researchers are creatively applying next-generation sequencing (NGS) data to ecological studies. After carefully collecting samples of plants and bees, as well as their associated microbe communities, from the environment, Dr. Keller and his team sequence the DNA with the MiSeq® System and use bioinformatics techniques to determine which species are cavorting with whom in the pollination network. Dr. Keller hopes that this detailed sketch of the intricate environmental interactions between plants, pollinators, and microbes will provide new viewpoints to understand the ecology, current threats, and resilience factors within pollination systems.

iCommunity recently caught up with Dr. Keller to learn more about how these novel NGS-based approaches make things hum in the field of pollination network ecology.

Q: How did you become interested in molecular biodiversity research?
Alexander Keller (AK): Biodiversity research was where my academic career started in the first place, investigating frog communities in Borneo for my Master's thesis. However, after graduating with an MS in Biology with a focus on Ecology, I switched to the Department of Bioinformatics to learn about a new set of tools for my projects. I became very interested in how sequencing data can be used in biodiversity research with phylogenetics—the study of evolutionary relationships among organisms. As a group leader, I'm now combining those two fields of research, using genomic data to address questions in ecological research. That's what motivates me.

Alexander Keller, PhD is Group Leader of the Molecular Biodiversity Group at the University of Würzburg. His dog, Ciya, often accompanies him in the field as he places artificial solitary bee nests and collects samples.

Q: What are the areas of focus in your lab?
AK: We integrate NGS data into ecological studies, especially but not exclusively the interaction triangle between plants, pollinators, and microbes. We're focused on determining the role microbial organisms play in influencing the pollination activities of insects in certain situations and the corresponding significance of microbes in flowers. In collaborative studies with other groups, we also use NGS to investigate pollination networks, where we try to figure out how plants and animal species interact with each other to obtain and maintain pollination services.

Q: What are the overall goals of your studies?
AK: Pollination is an important ecosystem service and many plant species rely essentially on these activities for reproduction. There are many plant and animal species that are very specialized in their pollination links, and others that are more generalized. We try to investigate which species are interacting and what factors lead to specialized pollination patterns. Microbial mutualists might play an important role in bees' decisions to interact with specific plants. This is due to the immense biochemical potential they possess, as well as their ability to transform nutrient sources and toxins, and biocontrol disease-related antagonists. It's very important to understand these characteristics so we can identify potential threats to the health of plants, bees, and other insects, and

maintain ecosystem pollination services. Because pollination also impacts crop yield, we're also interested in the type, effectiveness, and resilience of these interactions.

In addition, if one pollinator agent falls out of the system, because of a disease for example, then other agents might take over the same service for the ecosystem and for agricultural pollination. Nature often finds an alternative strategy to maintain pollination services, and our studies contribute to identifying the resilience factors that natural biodiversity provides.

> "We integrate NGS data into ecological studies, especially but not exclusively the interaction triangle between plants, pollinators, and microbes."

Q: What types of organisms are you studying?
AK: At the moment, we are concentrated on bees; not only the honey bee, but especially several species of wild bees. Some are solitary bee species, such as the mason bee, which is also quite an important pollinator. Further, we are looking at a diversity of plant species and their microbial associates.

For both organism groups, we're primarily sequencing the host-associated microbiomes. We're also studying some very interesting bacterial isolates obtained from those samples. Several provide antimycotic activities that repress fungal development and keep the solitary bee species' nest from getting moldy. Bees mostly use pollens as a source of food for their offspring. There are good indications in honey bees that microbes are involved in pollen predigestion to enhance the nutritive effectiveness. It's currently unknown how solitary bees do it and we're trying to figure that out.

In combination, we're also sequencing the pollens and trying to find out which plants they are collected from and whether there's some importance to having specific microbes for specific types of pollens.

Q: What is your process of sequencing pollen DNA?
AK: When you're using normal sequencing technologies, you need to have one isolate of each species and sequence each of those individually. However, we're simultaneously extracting the DNA for the entire community. Using NGS, we can look for a specific marker with a sequence unique for each species and obtain an overview of all of the species in the sample.

Pollen metabarcoding is very similar to the popular 16S amplicon application for bacteria. By adding some technical adaptations to the workflow and switching the marker region to the ITS2 of the eukaryotic ribosomal cistron, we are now able to look at whole mixed plant samples. With NGS, we can determine the complete bee pollen collections in one step without manually separating individual pollen grains.

Q: What does NGS data enable you to see?
AK: It provides us with a description of all of the species that are within the sample. For pollen studies, this means which plants the foragers visited. It's like going into a field plot and looking for all the plants that are inside by counting them. We do the same thing in our pollen field studies, but our samples are very small. Dependent on the specific origin, our samples might reach high biodiversity containing 200–1,000 bacterial species and 10–50 plant species. We can process the data statistically and determine factors structuring the corresponding communities.

Q: What is a typical day in the field like when you're collecting your samples?
AK: Depending on the project design, we are collecting samples from bees, their nests, and specific tissues of plants, especially flowers through swabs with cotton buds. In specific cases, we need to bring in the organism itself, collecting them from artificial or natural sites, to have them available for bioassays.

When we're collecting samples in the field, we work under as sterile conditions as possible so as not to degenerate the microbiota. We cool the samples down or use specific kits to stabilize the communities. For the samples only processed for pollens, this is not necessary because pollens don't degenerate quickly.

> "With NGS, we can determine the complete bee pollen collections in one step without manually separating individual pollen grains."

Q: How do you extract DNA from the samples?
AK: It's just like you would extract DNA from any other sample. The only difference is that you're not looking at one specific species at a time. We're looking at a whole community of species. That's something we need to consider when we interpret quality assurance steps, such as gel electrophoresis.

Q: What sequencing system are you using?
AK: We started off with other devices, but switched to the MiSeq System a year and a half ago. Several issues contributed to that switch. The main reason is that the MiSeq reads are now long enough to enable DNA barcoding, while providing good sequence quality and high throughput. We

usually do not need all of the capacity of one sequencing run for each sample, and are able to process multiple samples simultaneously on the MiSeq System with one flow cell.

Q: How are you using the MiSeq System for these studies?
AK: In general, it's important for ecologists to have a high sample size because there is so much variability in the environment that clouds the patterns of interest. Ecosystems are also very complex, and in most cases, high sample sizes are needed to get the statistical power required to draw solid conclusions. With the throughput currently generated with MiSeq v2 flow cells, we usually process 384 samples alongside each other for microbial or pollen community assessments, or 10–15 samples for microbial isolate genomes. Thus, the high multiplexing possibilities with the MiSeq allow us to account for the needed sample size cost-effectively.

Technically, we extract DNA from the samples, amplify established DNA barcoding markers, label them with MIDs (unique identification sequences), and add the adapters for the MiSeq System. This all happens in one PCR reaction. Then we normalize the amounts of DNA between samples and generate combined pools that include all samples. Finally, we perform quality control as usual and load the final library on the MiSeq System.

Q: Which Illumina library prep kits are you using?
AK: For microbial genomes, we use Nextera® XT DNA Library Preparation kits, which work quite well. For our metabarcoding approaches with bacteria, we use a protocol that was published in 2013[1], which was very successfully for our samples. We also developed a protocol for mixed pollen samples[2] and are in the process of developing a MiSeq version of it[3].

> "MiSeq reads are now long enough to enable DNA barcoding, while providing good sequence quality and high throughput."

Q: How does NGS compare with the light microscopy methods you used previously to study pollens (palynology)?
AK: The traditional method is to look at pollens with light microscopy. You go out into the field and collect pollens from bees or from airborne samples. You take the pollen, put it under the microscope, and identify the morphological characteristics and colorations to see which plant species they belong. This is a hard, error-prone task. Many types of pollen look the same so you can only identify them down to the family level or to the genus level at maximum. It's time-consuming to look under the microscope and it's very hard to classify the organisms.

You also need a lot of expertise in the botanical situation that you have in your country. An expert in Germany wouldn't be perfectly suited to characterize samples from Australia because the flora is different. This is a limiting factor for palynology studies. That's where using NGS, and now especially the MiSeq System, is beneficial. MiSeq sequencing is objective and the data are comparable between pollen studies worldwide. You don't need to be an expert in palynology to perform NGS-based studies. In most cases, we're able to obtain reliable and deep taxonomic assignments for all the pollen in our samples.

> "For microbial genomes, we use Nextera XT DNA Library Preparation kits, which work quite well."

Q: How have your research studies benefitted from using the MiSeq System?
AK: The sequencing is excellent when compared to other devices. For DNA barcoding, we need long sequences. The MiSeq System allows sequencing up to 500 cycles, which are 500 base pairs. There are only a few devices out there that allow sequencing of such large fragments, and the MiSeq is outcompeting them in terms of sequencing depth. As previously mentioned, we transform this depth by multiplexing, having up to 384 samples processed simultaneously. This eases our laboratory processing effort enormously, enabling us to perform these studies cost-effectively.

The MiSeq System sequence quality is also very good and that's important in mapping the sequences of various species and making assumptions about the biodiversity of an ecosystem. It's also quite flexible. We're performing microbiome and pollination network research, and also using it for shotgun sequencing. We're using several different protocols, but in the end, it's always the MiSeq System as the sequencing device. We have great data from all of these applications and it's really working out well. The MiSeq is perfect for our work.

Q: Will you soon be publishing results with data from the MiSeq System?
AK: Several manuscripts relying on MiSeq data are under review at the moment, so we hope to find the first articles accepted soon. In total, we have generated MiSeq data for approximately 30 different studies of our workgroup and collaborators together. Hopefully, many of them will be published within the next few years.

Q: How will your NGS studies change the field of network ecology?
AK: From the technical side, many ecologists are not aware that they can use molecular techniques to support their data. This is an emerging field and there's much potential to enhance ecological studies with sequencing data. NGS enables us to see, on a different scale and with an objective view, how biodiversity is structured and what interaction networks look like. In microbial ecology, this enables snapshots of microbes associated with a host, in our case with both plants and bees. In combination with other methods, the data sheds new light on how microbes affect pollination ecosystem services.

Our pollen sequencing studies represent a pioneering approach in obtaining detailed views of pollination networks. Studies applying this technique might gain considerable importance in environmental monitoring and agricultural decisions.

> "The MiSeq System sequence quality is also very good and that's important in mapping sequence species and making assumptions about the biodiversity of an ecosystem."

Q: Why is the MiSeq System a great tool for ecologists?
AK: Our studies are good examples that ecologists can gain great information by applying novel methods, especially NGS technologies. The MiSeq is ideal in assessing the biodiversity of mixed samples, and describing and monitoring communities and the changes that occur within them. I expect there will be a wide new array of methods developed for other questions and that they will greatly support the existing ecological toolboxes.

Q: Why is now the time for pollination studies?
AK: Many ecosystems are currently under threat. We're observing a great loss in biodiversity and natural habitats are disappearing. That puts human existence at risk. Pollination is a very central service within most ecosystems, and its resilience and maintenance is important. Every aspect that brings new light into pollination ecology is thus a piece in the puzzle to conserve and protect the stability of this system.

Nature has many ways of providing some means of resilience. Exemplarily, several honey bee populations are currently under threat, so it's important that there are alternative pollinators. Many of those are wild and solitary bees, which can take up the ecosystem services when honey bees are negatively impacted by disease. It's important to protect and conserve pollinator biodiversity, not just a single service species, in order to establish and maintain resilient ecosystems.

> "NGS enables us to see, on a different scale and with an objective view, how biodiversity is structured and what interaction networks look like."

Q: What improvements are needed to make the techniques that your lab is using more useful or universally available?
AK: We are one of the first labs performing NGS sequencing with pollens and developing new techniques in that regard, so this is hard to tell at the moment. However, I don't believe that our sequencing technology, metabarcoding approaches, or lab and field studies need improvement. What we need are improved reference species databases, especially for biogeographic regions that are not yet well investigated.

Q: What are the next steps in your studies?
AK: The next step is linking all of these results. We have quite a lot of information about which bees go to which plants, which microbes are associated with which plants, and which microbes are associated with which bees. We are sequencing the genomes of several very interesting microbes isolated from such hosts that show remarkable biochemical activity. Now we have to get all of those things together and try to figure out how the system works. This is a task that will take several years.

References

1. Kozich JJ, Westcott SL, et al. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol.* 2013; 79(17): 5112–5120.
2. Keller A, Danner N, et al. Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. *Plant Biol (Stuttg).* 2014; doi. 10.1111/plb.12251.
3. Sickel W, Ankenbrand MJ, Grimmer G, et al. Increased efficiency in identifying mixed pollen samples by meta-barcoding with a dual-indexing approach. *BMC Ecol.* 2015;15:20 doi:1186/s12898-015-0051-y.

Illumina • 1.800.809.4566 toll-free (U.S.) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

Pub. No. 770-2014-048 Current as of 04 August 2015

illumina®