Environmental DNA Sequencing Offers a Powerful Lens to View Changes in Biodiversity

Next-generation sequencing and eDNA metabarcoding enable researchers to monitor a wide variety of ecosystems accurately.

Introduction

Professor Michael Bunce is a molecular biologist at Curtin University in Perth, Australia. While his research interests are varied, all his projects revolve around extracting, amplifying, and analyzing degraded DNA. His **Tr**ace and **En**vironmental **D**NA (TrEnD) laboratory performs genetic analysis studies of a broad range of samples, including ancient DNA, herbal medicines, wine, wildlife, and seawater.

For the last decade, Dr. Bunce has studied ecosystem biodiversity using next-generation sequencing (NGS) to analyze environmental DNA (eDNA), the DNA that all organisms shed into their immediate settings. His team uses metabarcoding, a combination of DNA barcoding and high-throughput NGS, to analyze eDNA samples and monitor ecosystem changes over time. Recent studies by his team have assessed biodiversity in marine samples and evaluated rehabilitation success in restoring native ecosystems after mining or oil exploration.

iCommunity spoke with Dr. Bunce to discuss his eDNA studies, their value in environmental and biodiversity research, and how he uses the NextSeq™ 550, MiSeq™, and iSeq™ 100 Systems to study samples from various ecosystems.

Q: When did you begin applying molecular biology techniques to areas traditionally regarded as field biology, such as paleobiology and ecology?
Michael Bunce (MB): In 2003, I published a paper with colleagues that showed that DNA could be extracted from prehistoric Siberian permafrost cores and New Zealand cave sediments and sequenced successfully.[1] This was landmark work demonstrating that we don't need fossils to identify the animals that were present in an area many centuries ago. We can detect DNA that these animals left behind and use the data to study how plants and animals have changed over time.

Q: How has NGS impacted your studies?
MB: Analyzing DNA was difficult when we conducted paleobiology research in the early 2000s. We had to extract DNA from the sample, clone it into bacteria, and then sequence the bacterial plasmids using Sanger sequencing, one bacterium at a time. It was time consuming, costly, and didn't delve deep enough into the species present in the sample.

NGS has been pivotal in our ability to analyze environmental samples easily and accurately to determine what species existed there. With bacterial cloning and Sanger sequencing, we saw a limited DNA snapshot of each sample, typically obtaining 50–100 reads per sample. NGS and its massively parallel sequencing enables us to look at tens to hundreds of thousands of reads per sample, giving us a powerful lens into the assemblage within multispecies environmental samples. Over the last decade, we've used NGS-based metabarcoding extensively to look at many different biological substrates. Our work at TrEnD is more about understanding composition and changes in nonbacterial biological communities than sequencing genomes.

Q: What is eDNA and how can it be used to measure biodiversity?
MB: All organisms shed DNA into their environments. These genetic 'breadcrumbs' enable us to assess the different organisms that are or have been present in any given environment. If we concentrate a liter of seawater down to its organic components and extract DNA from it, we can learn about the biodiversity in that sample. For example, we can design PCR assays that act as molecular magnets, latching onto certain sequences of fish DNA to study what fish were in the seawater. We can build a species list and look at the presence or absence of various taxa and use that to build a holistic picture of the biota in a specific environment and how they interact.

eDNA can also help us study organisms that are difficult to spot. For example, seahorses are one of the most elusive fish in the ocean. However, we've been able to find them in South Africa and Australia by detecting their DNA signatures in seawater. We believe it will revolutionize how we study these and other cryptic species that camouflage themselves so well that they are really difficult to survey in a nondestructive way.

Michael Bunce, PhD, is a Professor in the School of Molecular and Life Sciences, and Head of the TrEnD laboratory at Curtin Universtiy in Perth, Western Australia.

In addition, we can use eDNA to look holistically at an entire ecosystem and use it as a barometer of how healthy or interconnected the species are within it. Conversely, when there's a major pollution event, such as an oil spill or intrusion of an invasive species, we can use eDNA to detect what species have been impacted and learn from those experiences.

> "When you compare DNA among species, there are many conserved regions. However, between those highly conserved areas is a highly variable region that we can use as a barcode to discriminate between species and determine the taxa."

Q: What are the common challenges working with eDNA?
MB: The primary challenge is the fact that working with eDNA is a numbers game. Each sample is a large haystack containing only a few needles. If you take a liter of seawater or a gram of dirt, bacteria typically make up more than 99% of the DNA in that sample. We need ways to selectively enrich for specific targets. If we are looking for coral species in seawater and our assays cross-react with phytoplankton, we'll detect mostly phytoplankton and little or no coral. So, we have to design (PCR) assays that are effective and target what we want and actively exclude what we don't want to assay in a given sample.

We also need to optimize sampling approaches to enrich for the DNA that we're studying. If we're assessing fish populations, will we find more fish DNA in sediment samples from the ocean floor, or in seawater? That information enables us to gravitate towards sampling only the biological substrates where eDNA ends up hiding. We just completed a coral study that showed that while there is a significant amount of coral DNA in seawater, there is very little of it in ocean floor sediments.

Another challenge is accessing the DNA. We need to use efficient DNA extraction methods. If we use methods that are suboptimal, we might not be able to detect species present in low abundance.

Q: What NGS methods are you using to study biodiversity in environmental samples?
MB: We use two NGS approaches, shotgun sequencing and eDNA metabarcording.

Shotgun sequencing is a great approach if we've got the numbers on our side. For example, if we're detecting bacteria or smaller eukaryotes, like phytoplankton. However, it's often ineffective for seawater samples because the haystack is too large.

For most environmental samples, we use eDNA metabarcording. Just like supermarket products, every organism has a unique barcode associated with it. In this case, it's a DNA barcode. When you compare DNA among species, there are many conserved regions. However, between those highly conserved areas is a highly variable region that we can use as a barcode to discriminate between species and determine the taxa.

When we want to tell species apart, we'll use metabarcoding and targeted assays. We've used these eDNA metabarcoding approaches on ice cores, dirt from swamps, archaeological sites, seawater, fecal samples, and herbal medicines. The application of eDNA metabarcoding extends across many applications.

Q: What is Tree of Life (ToL) metabarcoding and how have you used the approach for marine ecosystem studies?
MB: The concept of the ToL metabarcoding approach is that no single assay will give us the full diversity of an environmental sample. In a recent study, we obtained one nine-liter water sample and analyzed it with shotgun sequencing, bacterial metagenomic approaches, bacterial 16S sequencing, and 10 different molecular assays.[2] Some of these assays picked up coral, while others picked up fish. We looked at everything we could from a single sample to get a feel of the strengths and weaknesses of different assays and to compare shotgun sequencing with NGS metabarcoding.

We learned very quickly that no single assay – at least for sea water – will give us the breadth of diversity in that entire ecosystem. We need to be inventive with the assays we design and implement, and tailor them depending on the plants or animals that we're trying to study.

> "We need to be inventive with the assays we design and implement, and tailor them depending on the plants or animals that we're trying to study."

Q: How could eDNA metabarcoding be used to study restoration of ecological systems?
MB: There are numerous mine sites in Australia that are being rehabilitated back to their native ecosystems. Usually we plant native trees that used to be at these sites and wait a few years. In the meantime, we don't know if the restoration is on the right trajectory.

We believe that eDNA metabarcoding could assess what used to live there and what should be planted there. It could measure whether the insects, pollinators, and bacteria native to the ecosystem were returning.[3] We could use eDNA metabarcoding to look at what animals and insects are present by analyzing scat, air, and water samples. From this data, we could develop a biological order of the different animals, plants, and insects present in these environments to see if our restoration efforts are effective or not.

One of the main challenges for studies like this is with getting the experimental design right. For example, when we pick up scat

samples, we don't analyze each sample one by one. We put all the scat samples into a large blender, mix it all up, and extract the DNA. The process is much more time and cost efficient.

Q: Are there limitations as to what can be determined with eDNA?
MB: There are questions about whether the eDNA that we're extracting from a sample directly correlates with the number of organisms that are present and could therefore be used to measure abundance. It will never correlate precisely because organisms shed DNA at different rates. However, we can look at the presence and absence of species in different areas, and if we replicate enough samples, we can turn that presence/absence data into a measurement of relative abundance. There are numerous research studies underway focused on extrapolating as much as possible from eDNA data sets.

Another limitation of eDNA analysis is that our ability to assign DNA barcodes to species with fidelity relies on robust databases. Researchers throughout the world are in the process of building those databases.

> "Illumina NGS systems provide excellent sequence fidelity, which is critically important when analyzing environmental samples with metabarcoding."

Q: What advice would you give to other researchers wanting to implement eDNA methodology?
MB: eDNA methodologies might appear to be quite simple as an approach. In actuality, they are difficult to implement well. Over time, contamination of samples can become a real issue. I advise researchers to do their homework about how DNA should be collected, stored, and extracted in a clean-lab environment to maintain the fidelity of the data that is generated. For example, we learned early in our studies that 2-step PCR approaches for producing eDNA libraries generate numerous artifacts and are very susceptible to contamination.

Q: What NGS systems are you using for your current studies?
MB: Illumina NGS systems provide excellent sequence fidelity, which is critically important when analyzing environmental samples with metabarcoding. We need to have confidence that our DNA barcodes have been faithfully copied. In our lab, we use the NextSeq 550 System for all of our shotgun sequencing work (eg, coral and fish mitochondrial genomes). The MiSeq System is used for amplicon sequencing for our metabarcoding workflows and its cost per run is very affordable. The MiSeq System was ideal for our ToL metabarcoding study because it could amplify longer DNA segments, up to 600 base pairs in length.

We recently took delivery of an iSeq 100 System. While it possesses similar capabilities to the MiSeq System, it is smaller and more portable. We're using the iSeq 100 System to simulate how we could perform DNA sequencing outside of a lab environment. The iSeq 100 System will enable us to turn around data quickly while maintaining sequence fidelity that other platforms don't provide currently.

Q: How do you analyze eDNA data?
MB: When we amplify the DNA in metabarcoding workflows, we include indexes, akin to number plates, on the front and back of the DNA barcodes that allow us to assign any DNA read to a particular sample. Before analysis, we remove what we think are errors and artifacts. Bioinformatics tools such as USEARCH and DADA2 help identify these artifacts and remove them from our data set. We actively filter the data to reduce background noise. Once we have a 'clean' dataset, we query all the barcodes against reference databases using alignment algorithms like BLAST. We might use GenBank or custom databases that we have constructed.

As a field, eDNA is still finding its feet regarding data analytics. There are many ways to crunch the data. The key is to distinguish real patterns from background artifacts while still maintaining the sensitivity needed to detect rare taxa.

Q: What other eDNA applications have been enabled with NGS?
MB: The potential applications of eDNA are almost too numerous to list. Currently, they include port monitoring, baseline biodiversity surveys, ballast water testing, soil testing, scat (diet) analysis, fermentation dynamics, and food/medicine ingredients. The list keeps getting longer.

> "There are many ways to crunch the data. The key is to distinguish real patterns from background artifacts while still maintaining sensitivity to detect rare taxa."

Q: You are launching a commercial venture called eDNA Frontiers. How will your eDNA metabarcoding approach be used to assist clients?
MB: There are a few commercial eDNA service labs popping up. Our eDNA Frontiers service offering will launch in 2019 and will be a shopfront model using some of the workflows we've developed over the past decade.[4] Our clients will include resource companies or entities seeking to obtain planning approval in marine or terrestrial systems.

We believe that eDNA data can be used as an insurance policy to monitor biological changes at specific locations over time. For example, some of our clients want to know what their site's ecosystem looks like before they begin operations. eDNA could

be used as a tool to show changes, positive or negative, that occur in the biology of that ecosystem. If their operations have a detrimental effect on the environment, they will know early enough to mitigate and lessen impact.

We're also setting up an eDNA biobank where companies submit samples to us and we retain them for a decade.[5] These samples will serve as a baseline and the data can be compared to more recent samples to show how environments are changing.

Q: How do you envision eDNA being used and studied in the future?
MB: I might be biased, but I think that eDNA could develop into one of the most powerful biomonitoring tools and become even more useful as the field matures. Biomonitoring with eDNA could be a vital tool for environmental protection agencies throughout the world that are charged with preserving the environment. They could offer tests to companies who want to mine or put an oil rig in a certain area. These companies could be required to collect environmental DNA samples before the sites are developed and then submit samples every 6–12 months to a biobank. If there's a major incident, such as an oil spill or the sudden presence of an invasive species, the agency could use these eDNA samples to assess the impact the company had on its environment and hold parties accountable. The data would speak for itself.

> "Biomonitoring with eDNA could be a vital tool for environmental protection agencies throughout the world that are charged with preserving the environment."

Q: What else are you working on?
MB: Our lab is passionate about education and community engagement with the environment. We're about to launch an eDNA community science project called Sequence Our Seas (SoS). We'll visit school classrooms and provide the students with a sample bottle. They'll go down to their local beach or waterway, scoop up a liter of seawater, and filter the biological grunge out of it in their classroom. We'll then use our eDNA workflows to analyze the contents of the filter and create a list of all the fish, birds, and mammals present.

This project provides a conduit where people can engage with the biology in their local marine ecosystems that are under various anthropogenic pressures. It also teaches people genetic literacy, which we feel is extremely important.

Learn more about NGS for eDNA studies:

Environmental DNA Sequencing, www.illumina.com/techniques/sequencing/dna-sequencing/targeted-resequencing/environmental-dna.html

Learn more about the systems mentioned in this article:

NextSeq 550 System, www.illumina.com/systems/sequencing-platforms/nextseq.html

MiSeq System, www.illumina.com/systems/sequencing-platforms/miseq.html

iSeq 100 System, www.illumina.com/systems/sequencing-platforms/iseq.html

References

1. Willerslev E, Hansen AJ, Binladen J, et al. Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science*. 2003; 300: 791–795.
2. Stat M, Huggett MJ, Bernasconi R, et al. Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Sci Rep*. 2017; 7(1):12240. doi: 10. 1038/s41598-017-12501-5.
3. Fernandes, van der Heyde M, Bunce M, et al. DNA metabarcoding–a new approach to fauna monitoring in mine site restoration. *Restor Ecol*. 2018; doi.org/10.1111/rec.12868
4. eDNA frontiers, Environmental DNA, Biomonitoring Solutions. https://scieng.curtin.edu.au/edna-frontiers/. Accessed January 9, 2019.
5. Jarman SN, Berry O, Bunce M. The value of environmental DNA biobanking for long-term biomonitoring. *Nat Ecol Evol*. 2018; 2:1192–1193.

Illumina, Inc. • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com