A HIGH-THROUGHPUT METAGENOMIC APPROACH FOR IDENTIFYING ENVIRONMENTS RICH IN NPRS AND PKS CONTAINING MICROORGANISMS

Jason N Woodhouse, Rocky Chau, Lu Fan, Mark V Brown, Torsten Thomas, Brett A Neilan
School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney 2052, New South Wales, Australia. j.woodhouse@unsw.edu.au

Key words: Pyrosequencing, Non-ribosomal peptide, Polyketide

Introduction. Non-ribosomal peptides (NRP) and polyketides (PK) are the cornerstone of many modern pharmaceuticals (1). An increase in the understanding of the molecular mechanisms underpinning the biosynthesis of these molecules (2) has provided researchers with targets for the identification of these genes present in the environment. Traditional culture-dependent (3) and culture-independent approaches (4) do not provide access to the full depth of diversity present in the environment.
Here we present an approach for evaluating the true genetic diversity of NRP and PK biosynthesis simultaneously from multiple environments.

Methods. NRP and PK biosynthetic genes were amplified from the environment and sequenced using targeted degenerate primers. Amplicons were then sequenced using the *454* FLX sequencing platform to generate thousands of sequences from each sample. These sequences were subjected to a stringent pipeline that revealed the functional and taxonomic composition of these genes within the environment. The approach presented also provided sufficient information for the targeted recovery of entire biosynthetic pathways suitable for heterologous expression in a suitable host.

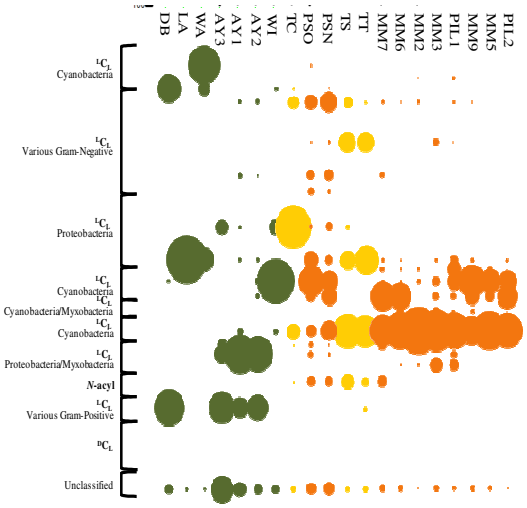

Fig.1 Taxonomic and functional composition of NRP biosynthesis domains from numerous environments.

Results. The generation of 21565 NRP and 1895 KS sequences from 20 environments resulted in the identification 3326 novel biosynthetic domains. Accounting for modular redundancy, we predict that this represents at least 300 novel biosynthetic pathways available for exploitation. While taxonomic classification suggested the origin of these domains was fairly limited, the functional diversity observed, indicates the production of a complex mixture of bioactive small molecules in each environment.

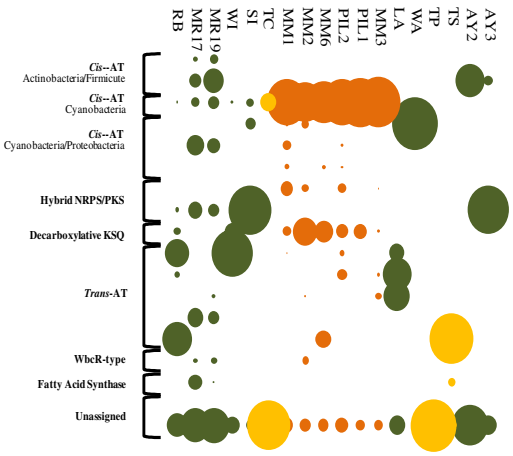

Fig.2 Taxonomic and functional composition of PK biosynthesis domains from numerous environments.

Conclusions. This method provides a novel strategy for the effective identification and evaluation of NRP and PK biosynthesis within the environment. Application of this method across multiple environments allows those containing higher metabolic diversity to be prioritized for downstream recovery of producing organisms or biosynthetic pathways in genomic libraries.

Acknowledgements. This work is supported by the Australian Research Council.

References.

1. Newman DJ, Cragg GM. (2012). *J Nat Prod.* 75(3):311-335.
2. Stachelhaus T, Marahiel M. (2006). *FEMS Microbiol Lett.* 123(1):3-14.
3. Strobel G, Daisy B. (2003). *Micobiol Mol Biol Rev.* 67(4):491-502.
4. Fisch KM, Gurgui C, Heycke N, van der Sar SA, Anderson SA, *et al.* (2009). *Nat Chem Biol.* 5:494-501.