

## Systematic Functional Characterization of the AAV Capsid Fitness Landscape

Eric D. Kelsic<sup>1</sup>, Pierce J. Ogden<sup>1,3</sup>, Sam Sinai<sup>1,4,5</sup>, and George M. Church<sup>1,2</sup>

<sup>1</sup>Wyss Institute, Harvard Medical School, Boston, MA, USA; <sup>2</sup>Genetics Department, Harvard Medical School, Boston, MA, USA; <sup>3</sup>Department of Biological and Biomedical Sciences, Harvard University, Cambridge, MA; <sup>4</sup>Department of Organismal and Evolutionary Biology, Harvard University, Cambridge, MA; <sup>5</sup>Program for Evolutionary Dynamics, Harvard University; Cambridge, MA, USA

Understanding the basic biology of AAV has enabled therapeutic applications and continues to facilitate the engineering of AAV capsids for enhanced *in vivo* gene delivery. In a complementary manner, data generated while engineering AAV can improve our understanding of viral function and suggest new directions for researching the virology of AAV and closely related parvoviruses. As an example of this second approach, we describe how our recent high-throughput characterization of an AAV fitness landscapes provides new insight into basic AAV biology. To better understand capsid sequence-function relationships and facilitate capsid engineering, we systematically created a barcoded library of all possible single codon substitutions, insertions and deletions across the entire AAV2 capsid (94,080 mutations in total). This library contains all possible first-order changes to the protein, enabling us to create high-resolution data-rich maps of the AAV capsid fitness landscape for functions such as capsid assembly, thermostability, heparin binding, infection *in vitro* and biodistribution *in vivo*. Our results reveal surprising new functions of the VP1 region and reveal empirical design rules for engineering tissue tropism. In addition to providing new insights into the function AAV, our work highlights the benefits of using next-gen DNA synthesis and DNA sequencing for building and screening AAV libraries, while demonstrating the potential of machine learning and data-driven protein engineering. Leveraging these emerging technologies will enable the field to design better AAV capsids that are safe, have low prevalence of pre-existing immunity, are efficiently produced, and are highly efficient and specific for target tissues. We expect that the experimental characterization of fitness landscapes will become an essential tool for AAV capsid engineering and for elucidating the basic biology of AAV and other parvoviruses.