# Selection of sponsor and institution

Dr. Justin Kinney is an expert in using massively parallel assays and modeling to interrogate the mechanisms of biochemical processes. He was one of the original inventors of the massively parallel reporter assay and showed that concepts from information theory allow the inference of quantitative biophysical models from these experiments. Currently, his lab focuses on applying massively parallel assays of variant effect, mathematical theory, and machine learning to the study of bacterial transcription and human mRNA splicing. Further, the Kinney lab develops software for the analysis of massively parallel experiments and inferring models from their results. This combination of expertise makes the Kinney lab well suited to this project. The lab has experience performing massively parallel assays of splicing activity like the one proposed in this project while also having a depth of experience training computational models to extract patterns from this sort of genomics dataset. By acquiring training in this lab environment, I can learn from Justin’s expertise on running a joint wet and dry lab effectively. This will be of great use to me as a training opportunity, because my previous lab experience has been in the context of labs that are run primarily as wet labs, so this will provide a very useful opportunity to learn about how run an effective mixed lab. Further, while Dr. Kinney’s expertise on its own would make his lab ideal for this project, the suitability of the lab is further improved by his established collaborations with specific subject matter experts.

These collaborations include an active collaboration with Dr. Adrian Krainer to study the mechanisms of splicing. Dr. Krainer has deep expertise in the field of splicing. His focus includes not only mechanistic studies but also translational work including the invention of Spinraza a splice switching antisense oligonucleotide therapeutic that became the first clinically approved therapy for spinal muscular atrophy in 2016. The location of the Krainer lab in the same building as the Kinney lab and the established collaboration ensure that I can easily access the splicing expertise of the Krainer lab. Easy access to this expertise will allow me to gain direct mentorship on the topic of splicing. A major training goal for me in this project is to learn how to apply my expertise in massively parallel assays and computational methods to RNA splicing. Joining such a close collaboration with the Krainer lab and attending their lab meetings will be an integral part of my efforts towards this goal. Further, this mentorship will greatly aid the execution of the project by providing advice and mentorship for the splicing related aspects of the project. This close collaboration and sharing of expertise makes the Kinney lab a perfect location to study splicing both from a scientific perspective and as a trainee.

In addition to collaboration with the Krainer lab to study splicing, Dr. Kinney also has an established collaboration with Dr. Peter Koo to develop methods for applying deep neural networks to genomics and for interpreting such networks. Dr. Koo in an expert in these fields, and leads a lab dedicated to these techniques. His lab has pioneered techniques that range from architectural changes to improve neural net interpretability to post hoc model interpretability analysis. In addition to collaborating closely the Kinney and Koo labs have adjacent dry labs and hold lab meeting and journal clubs jointly. This collaboration will also provide mentorship that will directly support my scientific and training goals. One of my major training goals is to gain more experience with cutting edge deep learning methods and interpretability approaches. Through this collaboration, I can gain direct mentorship on these techniques, and by attending lab meetings and journal clubs with the group, I will broaden my knowledge of this field even outside of my direct area of focus. Further, this collaboration with furnish mentorship and advice for the deep learning and model interpretability aspects of the project. This close relationship further strengthens the Kinney lab’s suitability for performing the parts of this project that rely on training and interpreting deep neural networks.

The particular expertise of Dr. Kinney for massively parallel assays, splicing, and computational modeling in addition to his close collaborations with experts in the fields of splicing, deep learning for genomics, and neural network interpretability make him an ideal choice as the sponsor of this project.

Cold Spring Harbor Laboratory (CSHL) is a world-class laboratory known for its pioneering science and strong academic commitment to the next generation of scientists. CSHL has shaped contemporary biomedical research and education with programs in cancer, neuroscience, plant biology and quantitative biology. Altogether, CSHL employs 1,000 people including 600 scientists, students and technicians. The Simons Center for Quantitative Biology (SCQB) at CSHL, where I will be doing my research, is highly vibrant, interdisciplinary, and interactive. Members of SCQB are actively involved in the department’s weekly QB seminar series and several journal clubs that focus on topics in deep learning, sequence-function relationships, and evolutionary and computational genomics. CSHL is also the host of numerous conferences that invite scientists from all over the world to present and share their ideas and research. CSHL has a wide range of resources that enable both computational and experimental research. This includes a high-performance computing cluster as well as a variety of core facilities which are described elsewhere. This highly collaborative and rigorous environment will provide the resources and opportunities I need for both my scientific and training objectives. By being a member of this community, I will gain exposure to a wide variety of techniques, disciplines, and ideas that will help me broaden the suite of tools that I can bring to bear on scientific questions. Additionally, it will introduce me to a wide variety of scientists who can provide mentorship or be potential collaborators. While the meetings and courses offer these opportunities in a more structured manner, there are also many more informal opportunities for this sort of training and mentorship that come from being in a location frequented by so many scientists coming to share knowledge and learn. In addition to these directly scientific opportunities, CSHL has ample opportunities for training in skills like grant writing and mentorship through professional development workshops. CSHL also offers opportunities to sharpen my teaching skills through TA-ing. By taking advantage of these opportunities, I will strengthen the skills I will need as an independent researcher.