

Leveraging Artificial Intelligence to Design AAV Mutant Capsids Optimized for Antibody Evasion

Daniel H. Cox, Damien Maura, Weitong Chen, Mathieu Nonnenmacher

Voyager Therapeutics Inc., Lexington, MA, USA

ABSTRACT

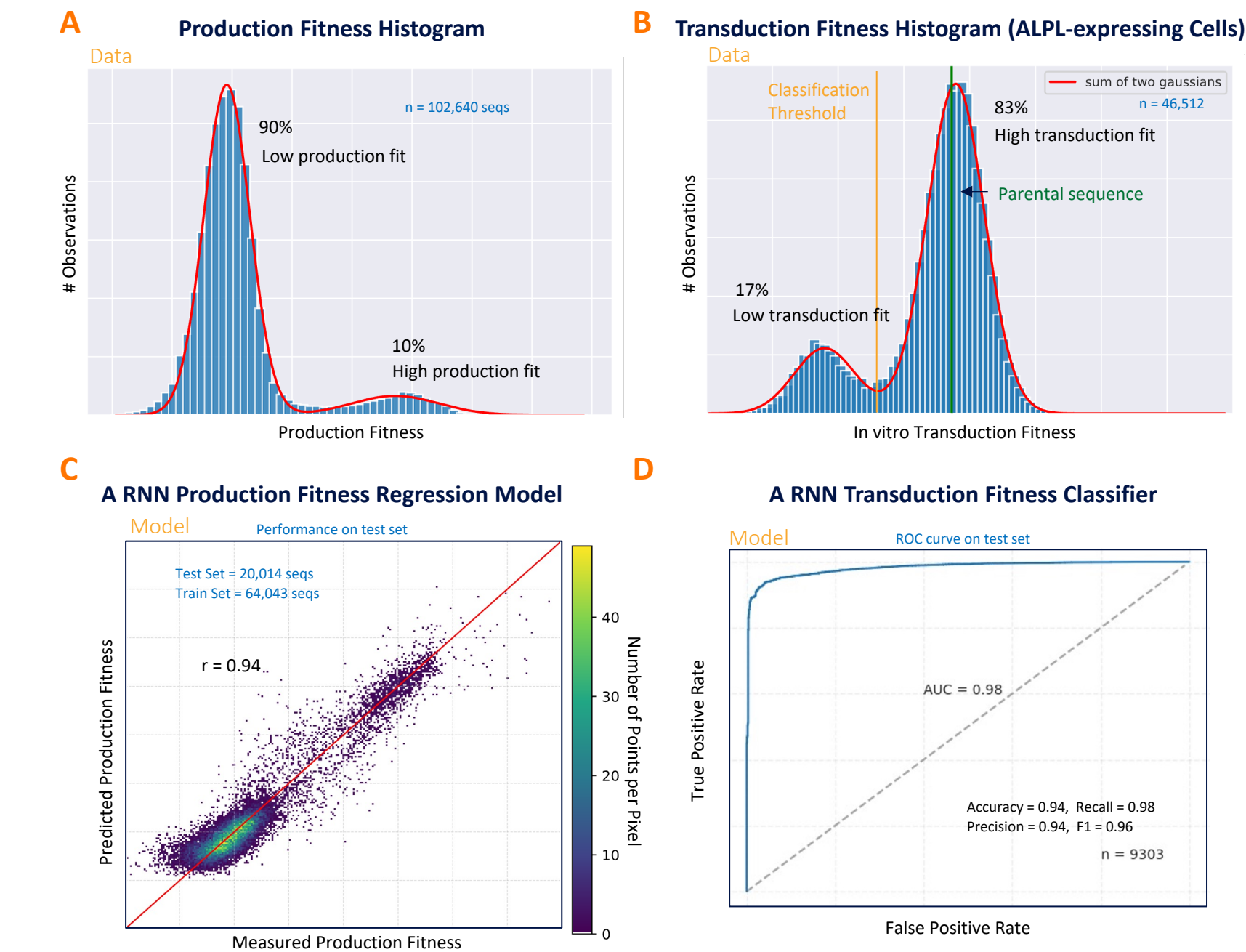
Using our AAV-display TRACER™ technology, we have identified AAV-capsid variants that effectively transduce the CNS in both rodents and primates. These variants may be valuable vehicles for the delivery of gene therapies to the brain. To make these vectors more universally applicable, we are also engineering them to evade pre-existing AAV-neutralizing antibodies (NABs). These NABs are present in a significant fraction of prospective patients and are believed to reduce the efficacy of intravenous gene therapy treatments, leading to the exclusion of seropositive patients from clinical trials. We initially mutated several capsid regions hypothesized to be immunogenic and tested the resulting capsids for NAb-evading properties. Capsid mutants were screened for three different properties: maintenance of capsid assembly and efficient viral production, maintenance of effective CNS transduction, and measurable NAb evasion. In some capsid regions, however, we have found that such mutations are rare and searching for them by random chance is impractical. In one region of interest, for example, less than 10% of randomly generated mutants were viable, and of these only a small fraction maintained the full CNS-transducing phenotype. In this region, therefore, we have used machine-learning to computationally optimize capsid mutant pools by preemptively eliminating defective mutants and maximizing viable sequence representation. Specifically, based on data from two rounds of mutation and empirical laboratory screening, we have developed regression models using recurrent neural networks and transformers to predict — from amino-acid sequence — capsid-assembly fitness, transduction fitness, and NAb evasion. Here we show the performance of these models and how we have used them to generate novel AAV variants that show large improvements in antibody evasion while maintaining their favorable manufacturing and CNS-transduction properties.

A DATASET FOR ML MODELING

- ~100,000 mutant AAV9 VP1 capsid sequences were generated synthetically, containing random mutations in region B.
- These mutants were constructed in an ALPL-binding parental sequence previously shown to transduce mouse and NHP brain well.
- The sequences were packaged into virus, and both plasmid and viral DNA sequences were quantified by next generation sequencing (NGS).
- Production fitness was calculated as $\log_2(\text{viral cpm} / \text{plasmid cpm})$ for each VP1 mutant sequence.
- Viral particles were exposed to HEK293 cells expressing ALPL.
- Cell transduction was measured by NGS after capsid RNA isolation.
- Transduction fitness was calculated as $\log_2(\text{cellular cpm} / \text{viral input cpm})$.

ML MODELING ROUND 1

Figure 1. RNNs for Production and Transduction Fitness Prediction



- (A) Only 10% of mutant sequences are high-production fit.
- (B) 83% of mutant sequences can transduce ALPL-expressing-cells.
- (C) An RNN regressor can predict production fitness with a high correlation coefficient between measured and predicted.
- (D) An RNN classifier can identify transduction-fit sequences with high accuracy.

Confusion Matrix		Predicted	
	Negative	Positive	
True Negative	0.72	0.28	
True Positive	0.02	0.98	

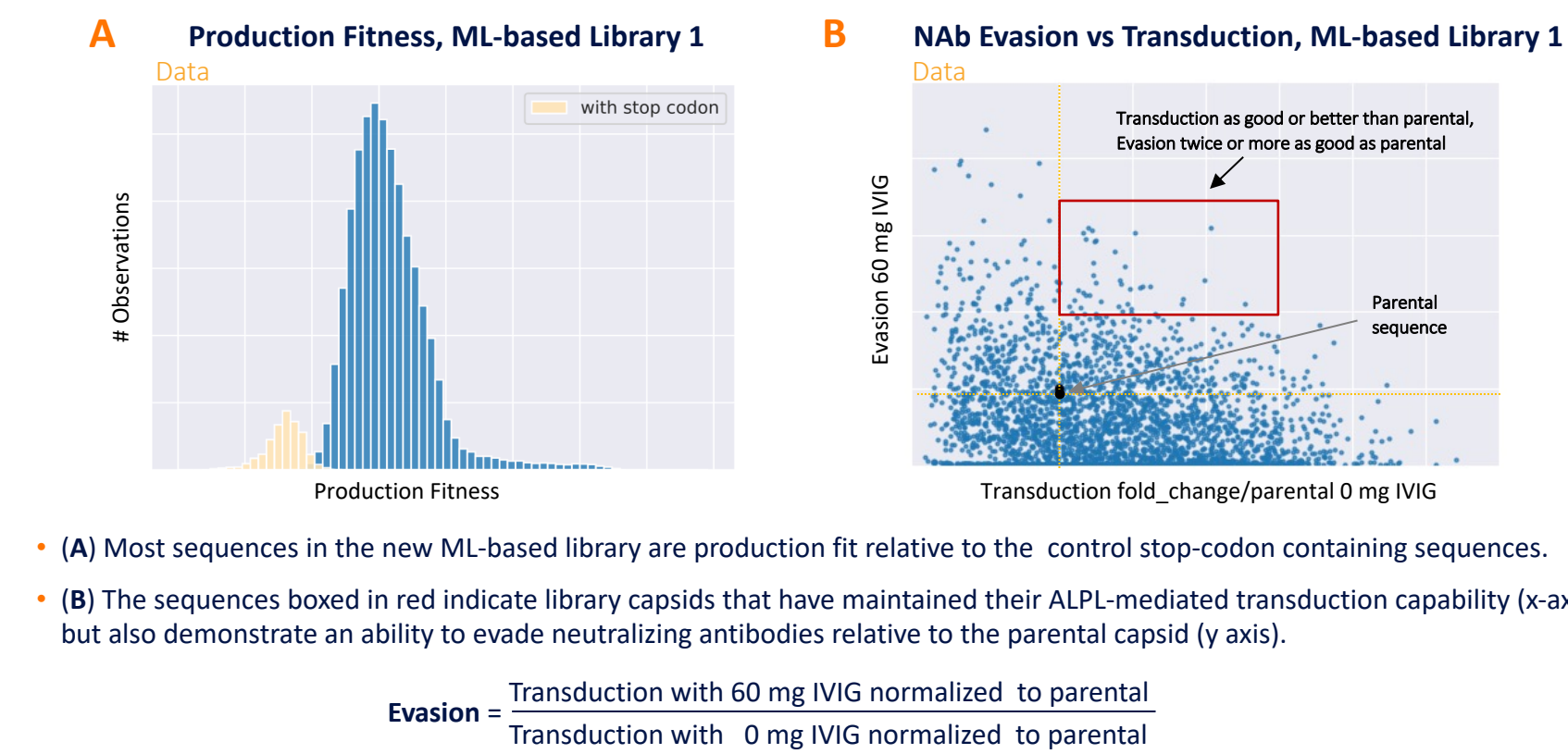
ML-BASED LIBRARY 1 FOR NAB EVASION SCREENING

- 100,000 mutant sequences were selected for a new library to be screened for AAV9 neutralizing-antibody evasion.
- These mutants were made in an ALPL-binding parental sequence.
- Each sequence passed an in silico double screen of a larger pool of ~25 million randomly generated sequences.
- The large pool was screened for production fitness with the model shown in Figure 1C, and for transduction fitness with the model shown in Figure 1D.
- 3,000 stop-codon-containing sequences were also included as negative controls.
- To sample uniformly across the sequence space a minimum hamming distance constraint was applied.

RESULTS FROM ML-BASED LIBRARY 1

- We used the TRACER™ system to quantify the ability of each mutant in the new ML-based library to transduce mouse brain in the absence or presence of human AAV9 neutralizing antibodies.
- Mice were injected with the new library at 2.5e13 VG/kg with either 0 or 60 mg human intravenous immunoglobulin (IVIG).
- After 14 days the mice were sacrificed, and NGS was used to quantify RNA expression in the brain from each mutant capsid. The results from six mice were averaged, and a CV filter of 1.0 was applied.

Figure 2. Results from ML-based Library 1

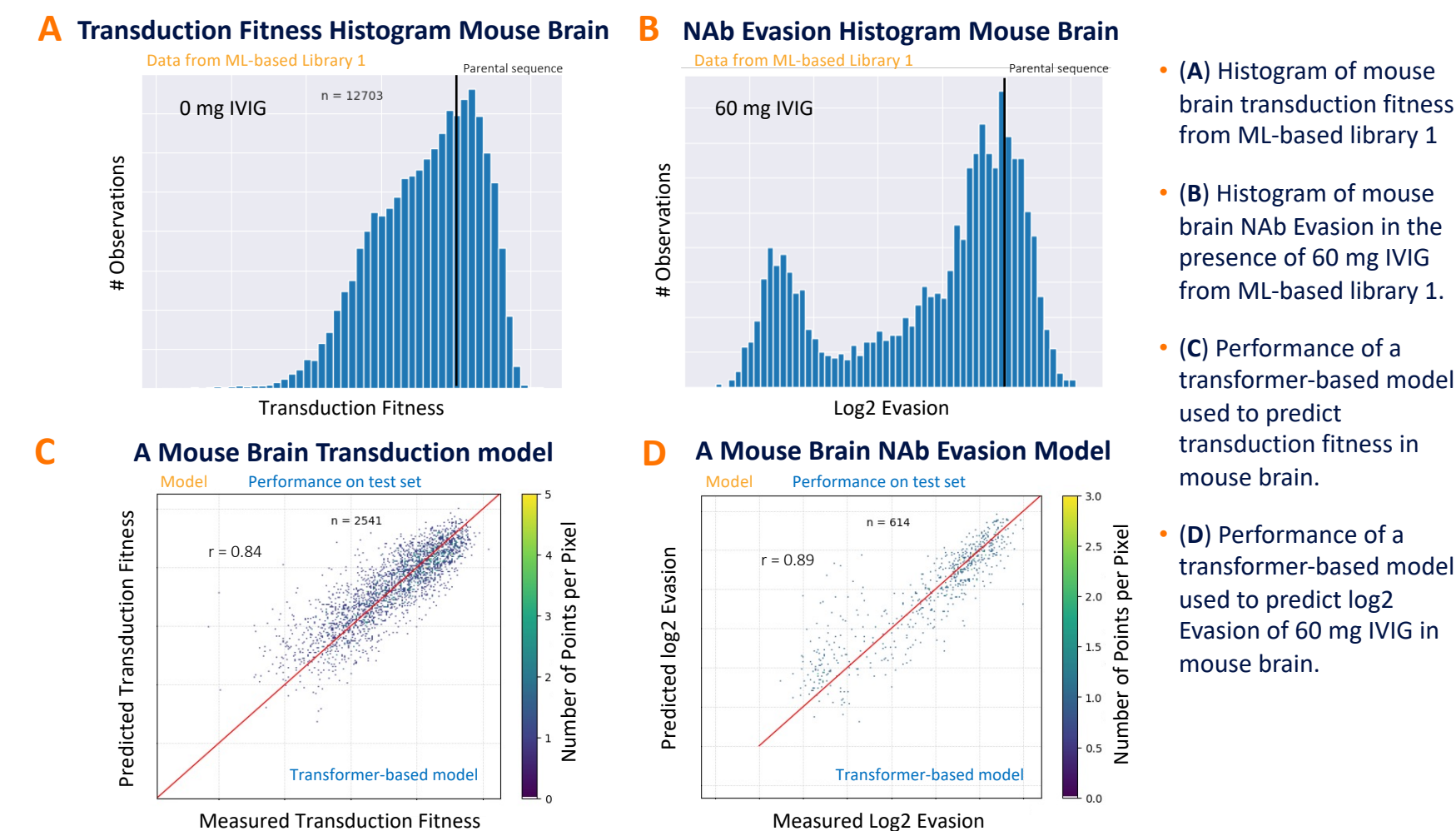


- (A) Most sequences in the new ML-based library are production fit relative to the control stop-codon containing sequences.
- (B) The sequences boxed in red indicate library capsids that have maintained their ALPL-mediated transduction capability (x-axis) but also demonstrate an ability to evade neutralizing antibodies relative to the parental capsid (y axis).

$$\text{Evasion} = \frac{\text{Transduction with 60 mg IVIG normalized to parental}}{\text{Transduction with 0 mg IVIG normalized to parental}}$$

ML MODELING ROUND 2

Figure 3. Transformers for Transduction and NAb Evasion Prediction

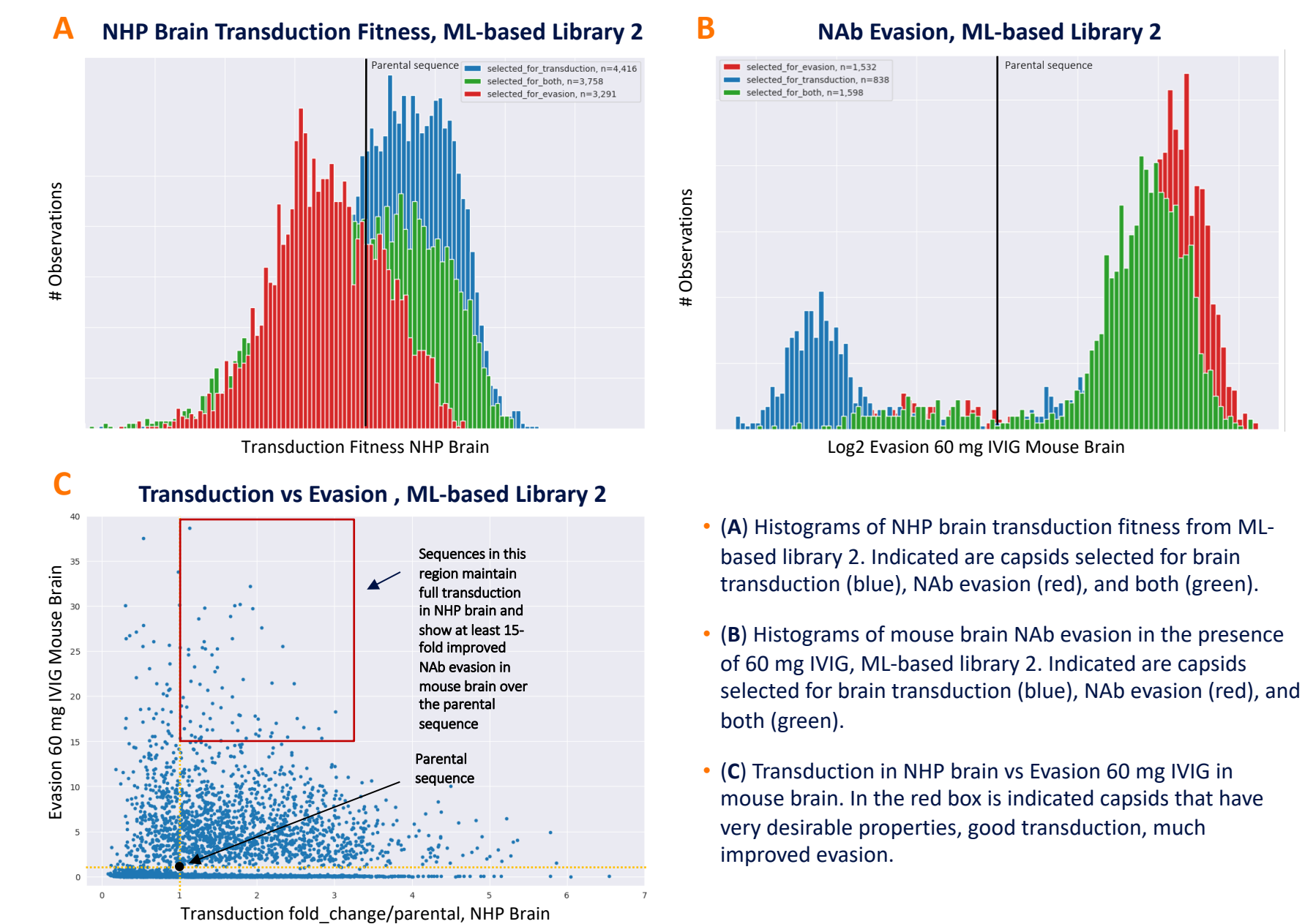


- (A) Histogram of mouse brain transduction fitness from ML-based library 1
- (B) Histogram of mouse brain NAb Evasion in the presence of 60 mg IVIG from ML-based library 1.
- (C) Performance of a transformer-based model used to predict transduction fitness in mouse brain.
- (D) Performance of a transformer-based model used to predict log2 Evasion of 60 mg IVIG in mouse brain.

ML-BASED LIBRARY 2

- 15,000 mutant capsid sequences were selected from a larger pool of 1.7 billion sequences for a new library to be screened for brain transduction in NHP and AAV9-neutralizing-antibody evasion in mouse.
- These mutants were made in an ALPL-binding parental sequence that already had very good brain-transduction properties.
- Sequences were screened in silico with the models in Fig. 1C, Fig. 3C, Fig. 3D, for production fitness, brain transduction, and NAb evasion respectively.
- For laboratory experiments, 5,000 sequences were chosen that passed in silico screening only for brain transduction (blue). 5,000 sequences were chosen that passed in silico screening only for NAb evasion (red), and 5,000 sequences were chosen that passed screening for both (green).

Figure 4. Results from ML-based Library 2



- (A) Histograms of NHP brain transduction fitness from ML-based library 2. Indicated are capsids selected for brain transduction (blue), NAb evasion (red), and both (green).
- (B) Histograms of mouse brain NAb evasion in the presence of 60 mg IVIG, ML-based library 2. Indicated are capsids selected for brain transduction (blue), NAb evasion (red), and both (green).
- (C) Transduction in NHP brain vs Evasion 60 mg IVIG in mouse brain. In the red box is indicated capsids that have very desirable properties, good transduction, much improved evasion.

CONCLUSIONS

- With a machine learning model for production fitness, we have been able to screen for viable AAV9-capsid mutants in a very difficult to mutate area where only 10% of random mutants are viable.
- Through two rounds of laboratory experimentation and ML-modeling of brain transduction and NAb evasion, we have been able to identify many capsid mutants that retain very good brain transduction and demonstrate greatly improved NAb evasion.
- Such capsids may unlock gene therapies for a larger percentage of potential patients than would have previously been possible.