Machine-Learning for AAV9 Mutant-Capsid Screening for both Production and ALPL-Mediated Transduction Efficiency

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

Voyager Therapeutics Inc., Lexington, MA, USA

1911

INTRODUCTION

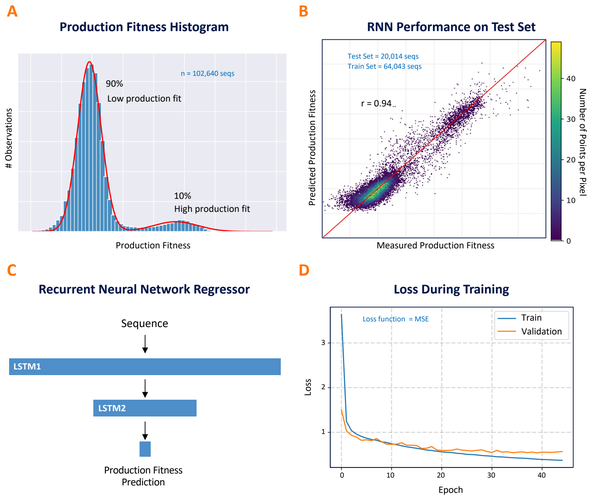
- We recently identified AAV9-capsid variants that show high CNS transduction in rodents and primates.
- They are relying on ALPL as their primary surface receptor for transport across the blood-brain barrier.
- We are also attempting to further engineer our vectors by removing major antigenic sites on the capsid surface. Here we have made mutations in a potentially antigenic region of the capsid.
- However, we have found that most mutations in this region render the capsid nonviable and many reduce its CNS-transduction efficiency.
- Furthermore, there is an inverse relationship between these properties, good producers are often bad transducers and vice versa.
- Thus, building a diverse mutant capsid library in this region — that contains mutants that both produce well and maintain ALPL-mediated brain transduction— has been a challenge.
- To meet this challenge, we have taken a machine-learning-based approach. Specifically, we have used a combination of machine learning models to prescreen mutant-capsid sequences for both production fitness and ALPL-mediated transduction efficiency.
- We show that this approach has allowed us to effectively build the desired library and through it identify novel capsids with enhanced human intravenous immunoglobulin (IVIG) evading properties.

1 DATA FOR ML MODELING

- ~100,000 mutant AAV9 VP1 capsid sequences were generated synthetically, containing random mutations at 8 positions 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 $\log_2$(viral cpm / 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 $\log_2$(cellular cpm / viral input cpm).

2 PRODUCTION FITNESS MODELING

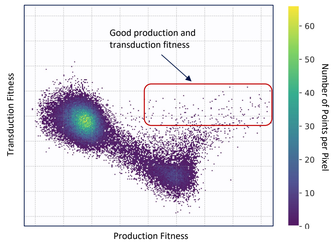
Figure 1. An RNN for Production Fitness Prediction



- (**A**) Only 10% of mutant sequences are high production fit.
- (**B**) A RNN regressor can predict production fitness with a high correlation coefficient between predicted and measured values, Pearson r = 0.94.
- (**C**) Neural network architecture, (**D**) Loss vs. epoch during training.

4 PRODUCTION VS TRANSDUCTION

Figure 3. Production is Inversely Related to Transduction



- There is an inverse relationship between production fitness and ALPL-mediated transduction fitness among mutants.
- The area in red encloses capsids that are both production and transduction fit. It is mutant capsids like these that we would like too screen further for nAb evasion.

3 TRANSDUCTION FITNESS MODELING

Figure2. An RNN for Transduction Fitness Classification

		Predicted	
		Negative	Positive
True	Negative	0.72	0.28
	Positive	0.02	0.98

- (**A**) ALPL-mediated transduction fitness histogram showing low and high transduction peaks.
- (**B-D**) A RNN Classifier that can predict high vs. low transduction fitness with an accuracy of 0.94 and a ROC-AUC of 0.98

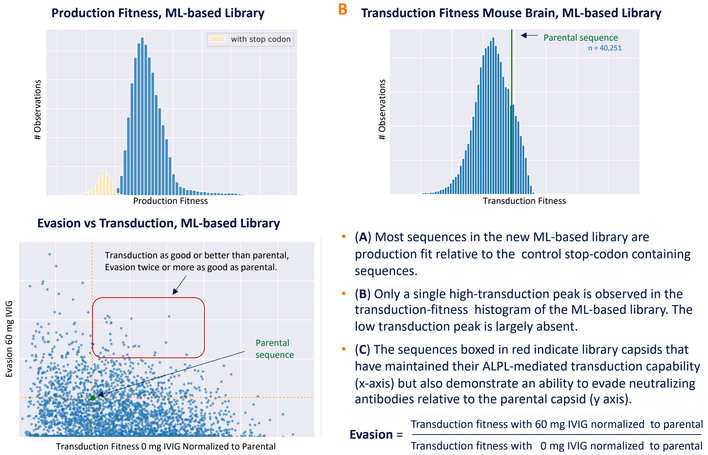
5 An ML-BASED LIBRARY FOR nAb EVASION SCREENING

- 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 containing 3, 4, 5, 6, 7 or 8 mutations.
- The large pool was screened for production fitness with the model shown in Figure 1, and for transduction fitness with the model shown in Figure 2.
- 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.

6 RESULTS FROM THE ML-BASED LIBRARY

- We used the TRACER™[1] 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.

Figure4. Results from the ML-based Library



- (**A**) Most sequences in the new ML-based library are production fit relative to the control stop-codon containing sequences.
- (**B**) Only a single high-transduction peak is observed in the transduction-fitness histogram of the ML-based library. The low transduction peak is largely absent.
- (**C**) 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).

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

CONCLUSIONS

- We have developed a machine-learning model that can make production-fitness predictions for AAV9 capsid variants mutated in a region where 90% of random mutants are nonviable. On our test dataset, measured and predicted values strongly correlate, Pearson r = 0.94.
- We have also developed a machine-learning model that can distinguish between high and low transducing mutants in a cell-based assay of ALPL-mediated viral transduction, accuracy = 0.94.
- We have used these models to prescreen mutants for inclusion in a library to be screened for AAV9 neutralizing antibody evasion.
- We show that this ML-based approach has allowed us to generate a diverse, largely-viable library in this difficult to mutate area.
- Further, it has allowed us to identify mutant capsids which display both enhanced ALPL-mediated brain transduction and AAV9-neutralizing-antibody evasion.

REFERENCES

1. Nonnenmacher, M., et al. (2021). "Rapid evolution of blood-brain-barrier-penetrating AAV capsids by RNA-driven biopanning." *Mol Ther Methods Clin Dev* **20**: 366-378.