# WhiteLab Genomics: Key Applications

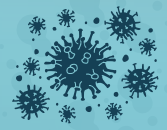
WHITELAB GENOMICS

## Target Discovery

- Target expression profiling
- Cellular biomarkers screening
- Epigenetic signatures identification
- Protein structures and interactions simulations
- Target related biomedical knowledge

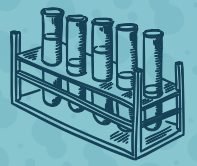
## Payload Design

- Promoter/enhancer choice
- Transgene expression modulation



## Vector Design

- Viral capsid & envelope design
- Non viral vectors design
- Transduction optimization
- Biodistribution prediction
- Immunoreactivity modulation

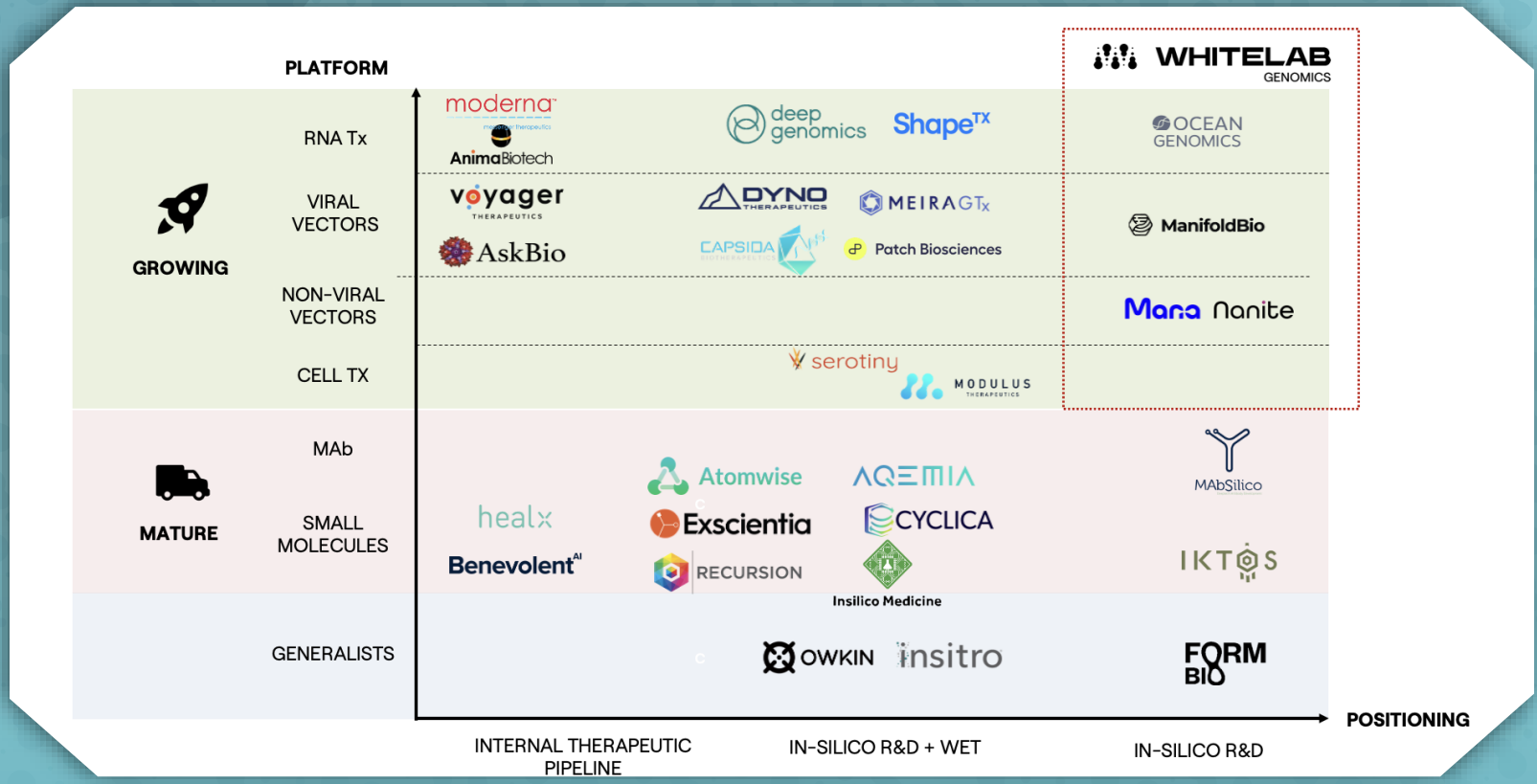


## Bioproduction

- Capsid viability
- Helper function optimization
- Bioproduction optimization (% Yield)
- Full/empty ratio

# What makes WhiteLab Genomics different ?

WHITELAB GENOMICS




WHITELAB GENOMICS

**WhiteLab Genomics is the only in-silico company using data science and AI to foster and de-risk research and non-clinical development exclusively for genomic therapies**

# WhiteLab's Platform - WKNOW™ AI

WHITELAB GENOMICS

## DATA

Text Based Data
Publications, Clinical Trials, Patents, Regulatory

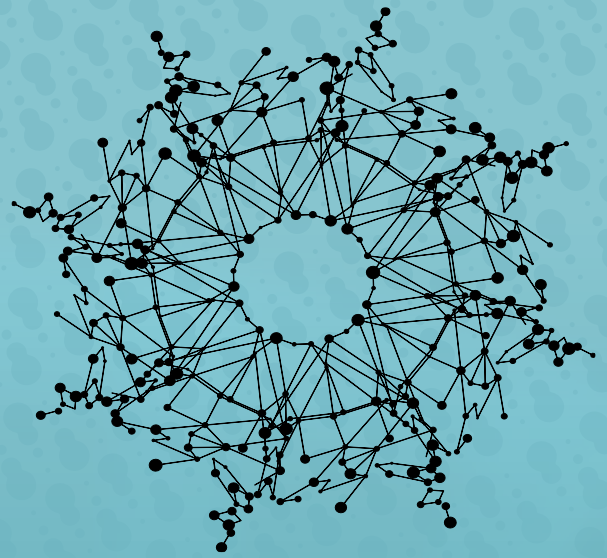
Biological Data,
Omics Tissue/Organs/Cells, Vectors, Plasmid

Experimental Data
Expression in tissues / organs /cells
Genotoxicity
Vector production

## KNOWLEDGE AI

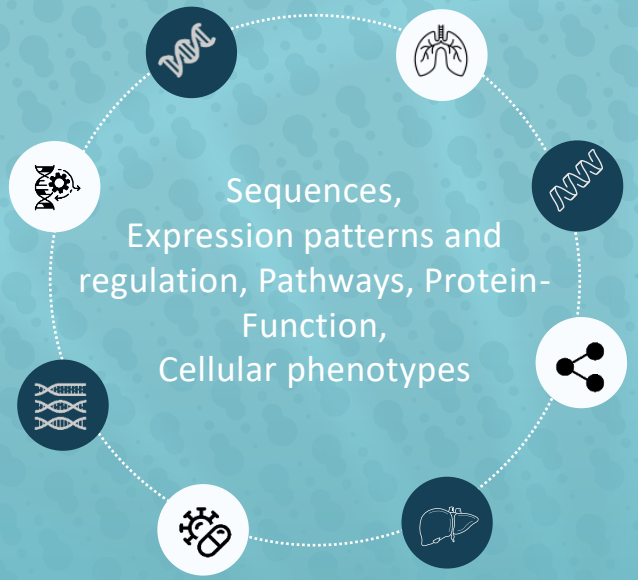
**Data Curation**

**Ontologies**



**Data Science Tools**

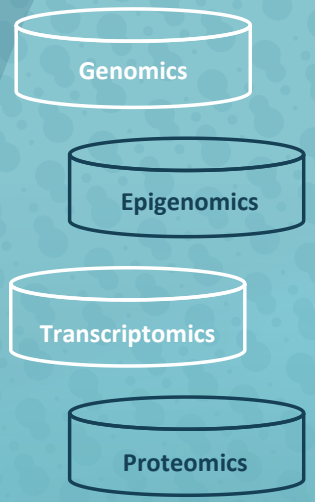
## PREDICTIVE MODELING



# WhiteLab's Platform – WMAP™ Atlas

WHITELAB GENOMICS

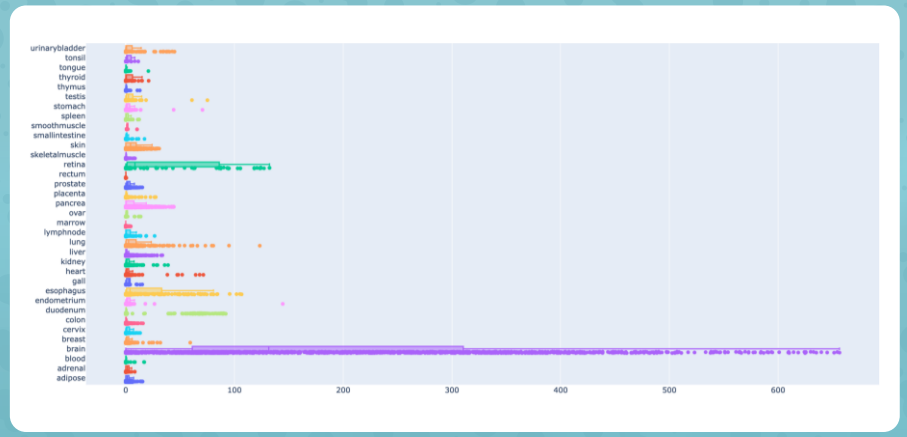
## DATABASES




- Experimental data sets - public and private data
- New genomic therapy sequences
- Biological public data sets (genes, proteins, taxonomies, etc.)

## BIOMARKER CELLULAR ATLAS

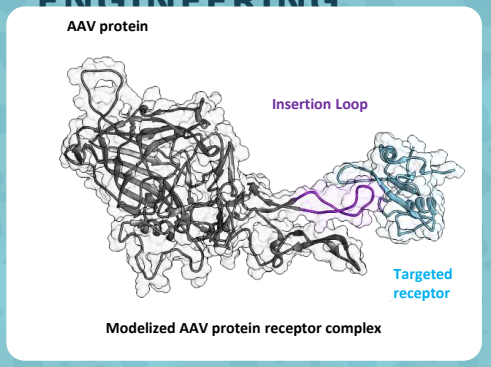
**A catalogue of 850+ biomarkers across the body**



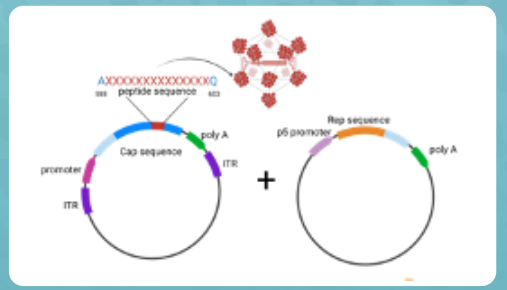

*A specific gene expression pattern in brain tissue identified*

- Specific biomarkers per tissue or cell type
- Analysis performed by 14 machine learning models
- Validation by experimental datasets

## VECTOR/PAYLOAD ENGINEERING




We build a new generation of rationally guided cell type specific vectors through a structurally-AI based peptide / ligand insertion design




We develop new approaches to optimize expression cassette and CAR design

Poster presented at ASGCT 2023

# Novel machine learning protocol for the phenotypic prediction of AAV variants library production

**Dylan SERILLON**[1], Tiziana LA BELLA[2], Anastasia BARKOVA[1], Jean Philippe BUFFET[1], David Del Bourgo[1], Ante MIHALJEVIC[2], Giuseppe RONZITTI[2], Julien COTTINEAU[1]

[1]WhiteLab Genomics, Future4Care, 8 Rue Jean Antoine de Baïf, 75013 Paris, France
[2]INTEGRARE, Genethon, INSERM, Université Evry, Université Paris-Saclay, 91002 Evry, France

## ABSTRACT

Adeno-associated virus (AAV)-based capsid libraries are becoming increasingly popular as a candidate selection tool for gene therapy vectors with a recent advance in the study of the specific VR-VIII region.

However, the amount of available data regarding variants carrying multiple mutations outside this specific region remains poorly available and functionally unexplored

Shuffling of natural adeno-associated virus (AAV) allowed to create a specific library composed by 272 capsids, mutated in 43 different positions. We decided to build a machine learning models based on two different protocols:

- *(i)* Learning from the sequence, we generated numerical descriptors that describe the physico-chemical properties of the mutated residues.
- *(ii)* Based on the sequence, the AAV2 variants were partially reconstructed as a 3D object, and a machine learning algorithm was trained on the geometric data.

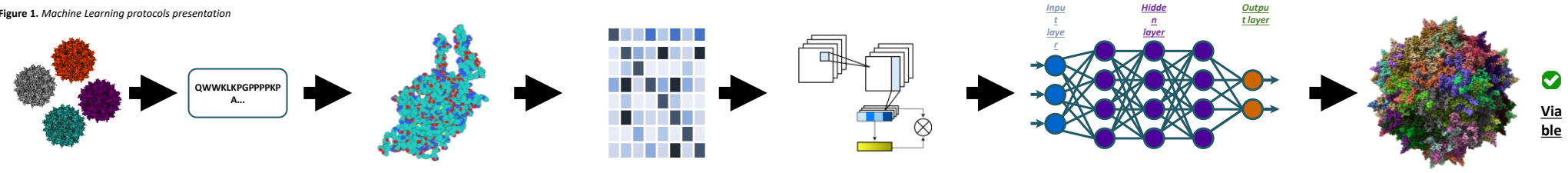
## CONTACT

Dr. Dylan SERILLON
*(PharmD, PhD)*
dserillon@whitelabgx.com

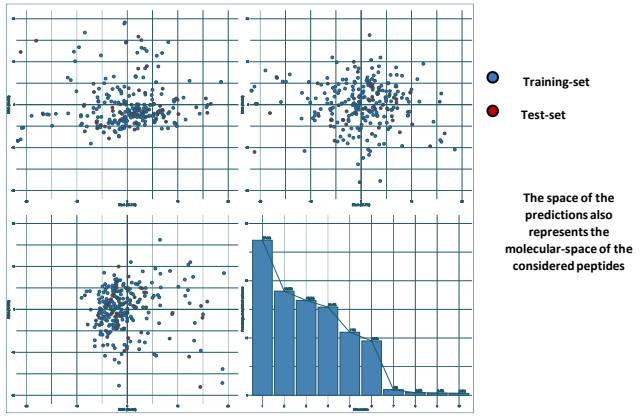
WhiteLab Genomics
https://whitelabgx.com/

## MATERIAL & METHODS

**Figure 1.** *Machine Learning protocols presentation*




## RESULTS




The space of the predictions also represents the molecular-space of the considered peptides

**Figure 2.** *Vizualization of the « **space** » of the predictions*

Controlling the machine learning model space is a challenging process, which allows for a little more rationality behind the classic 'black box' of artificial intelligence models. Our protocol is able not only to predict the capsid viability, but also to structure conformational changes to analyze physico-chemical variations.

## STATISTICAL ANALYSIS

**Table 1.** *Performances of the machine learning models*

| | Training-set | Test-set |
|---|---|---|
| | **Sequential approach** | |
| **Accuracy** | 100 % | 100 % |
| **Sensitivity** | 100 % | 100 % |
| **Specificity** | 100 % | 100 % |
| **Prevalence** | 17.1 % | 17.3 % |
| | **Geometric approach** | |
| **Accuracy** | 98.5 % | 100 % |
| **Sensitivity** | 91.7 % | 100 % |
| **Specificity** | 100 % | 100 % |
| **Prevalence** | 17.1 % | 17.3 % |

- **Prevalence** : percentage of non-viable AAV variants in the datasets.
- **Accuracy** : global performances.
- **Sensitivity** : represents the capability for a model to predict non-viable variants.
- **Specificity** : capability to predict viable variants.

In total, more than a million models were generated using the variation of the tuning parameters, the cross-validation protocol, and different data partitions.

Both models have different way to predict the same output and achieve powerful performances: respectively (*i*) 100% accuracy, (*ii*) and 98% accuracy with associated specificity (100%) and sensitivity (92%).

## DISCUSSION

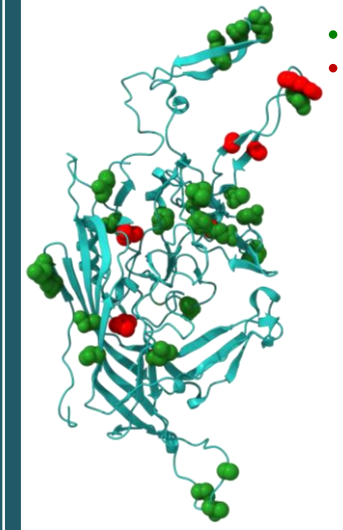


**Figure 3.** *AAV protein with highlighted mutated residues*

- Leading mutated residue
- Mutated residue

The models lead us to the identification of several positions which are predicted as fundamental for the stability of the viral capsid.

**It is interesting to note that some of the residues that have been predicted to be very important for the model are located outside the well-known VR-VIII region**

**As a perspective, a list of interesting variants, which were mutated on the 43 positions identified as interesting by our AI protocols** can be generated using rational approaches

# They talk about us

WHITELAB GENOMICS

*"This collaboration with WhiteLab Genomics will allow us to integrate a large body of information on critical parameters and their interconnections in complex processes so that we can very rapidly adapt our experiments; a new way for us to work, to accelerate our strategic projects and to innovate "*

**Anne Galy, Head of Inserm's Genomic Therapy Technology Research Accelerator (ART-TG)**

*"The tools developed by WhiteLab will make it possible for us to review thousands of sequences and devise new and innovative combinations. We aim to develop a new generation of more specific AAV vectors, contributing to the emergence of original treatments for neuromuscular disorders"*

**Giuseppe Ronzitti Group Leader, Genethon**

*"The value of using this AI-based platform is the potential acceleration of pre-clinical, translational stage, helping drug research companies quickly design payloads and vectors and identify the best experimental protocols for in vitro and in vivo tests"*

**Hamzeh Abdul Hadi, Investment Director at Debiopharm Innovation Fund**

## Fundraising, 2022 (10m€ series A)

Y Combinator
477 780 abonnés
6 mois •

French biotech company WhiteLab Genomics (YC W22) has raised $10 million in funding for an AI platform designed to aid the discovery and development of genomic therapies.

Thousands of diseases, including cystic fibrosis, Parkinson's, and Alzheimer's stem from flaws in an individual's DNA, and emerging research in gene and cell therapies may eventually treat such conditions at their source, supplanting the need for drugs or surgery. However, such therapies are typically costly to develop with no guarantee that they'll work.

Founded out of Paris by David Del Bourgo and Julien Cottineau in 2019, WhiteLab Genomics provides gene and cell therapy companies with predictive software simulations to expedite the design of gene and cell therapies. The company is currently working on projects including DNA-based therapies for metabolic conditions such as lysosomal diseases, as well as cell stem therapies for blood diseases such as sickle cell disease and immuno-gene therapies to treat cancers.

Congrats David, Julien and team WhiteLab on the round!

MNES

Debiopharm