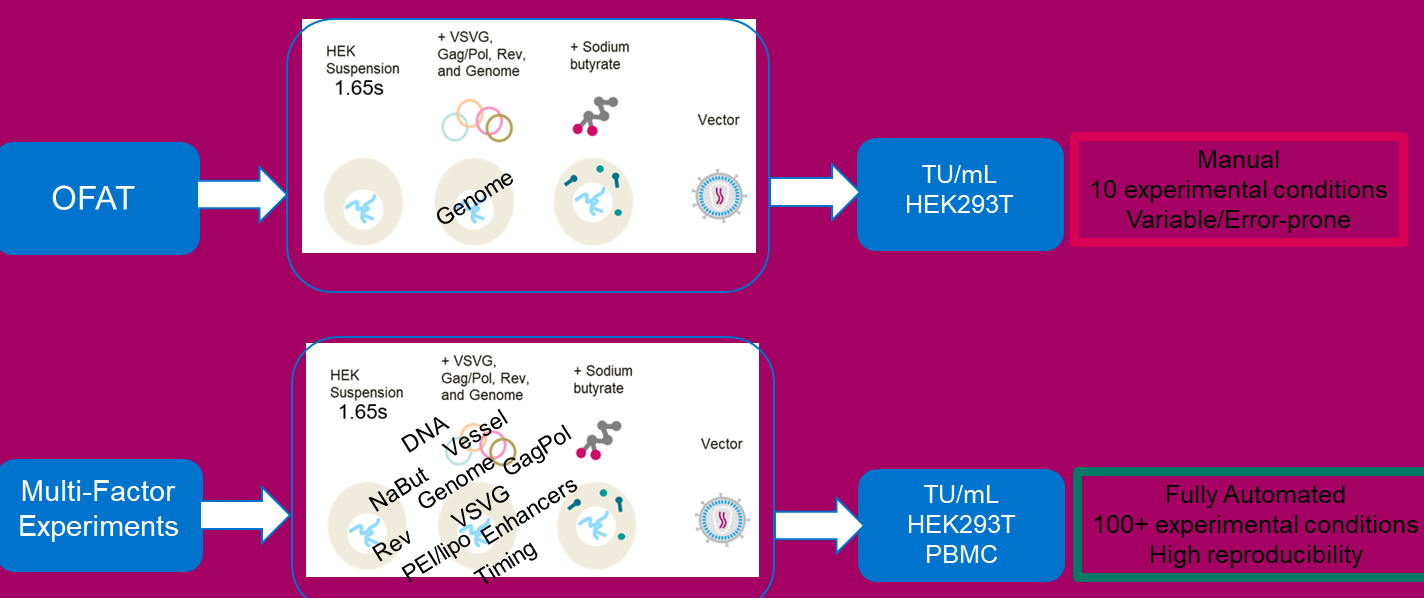


# Rapid Optimisations for Cell & Gene Therapy Applications

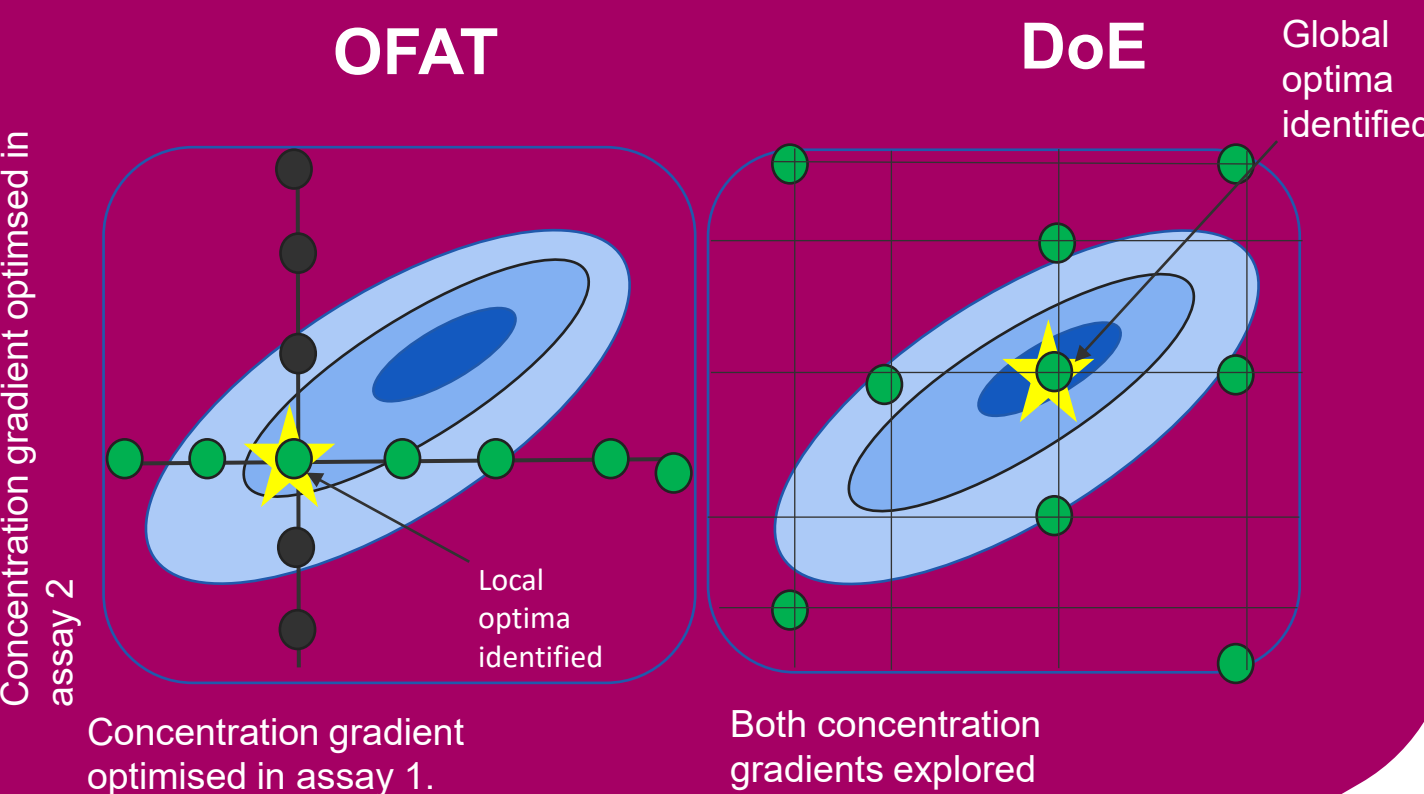
Thomas Evans<sup>1</sup>, Charles Moore-Kelly<sup>1</sup>, Anurag Kukarni<sup>1</sup>, John O'Driscoll<sup>1</sup>, Kyriacos Mitrophanous<sup>1</sup>, Nicholas Clarkson<sup>1</sup> and R. André Raposo<sup>1</sup>

## Design of Experiments (DoE)

A DoE is a multiple variable statistical method used to characterise and optimise systems quickly and efficiently. It has many advantages over one factor at a time approaches (OFAT). However, executed manually DoE methods can be time consuming and error prone.

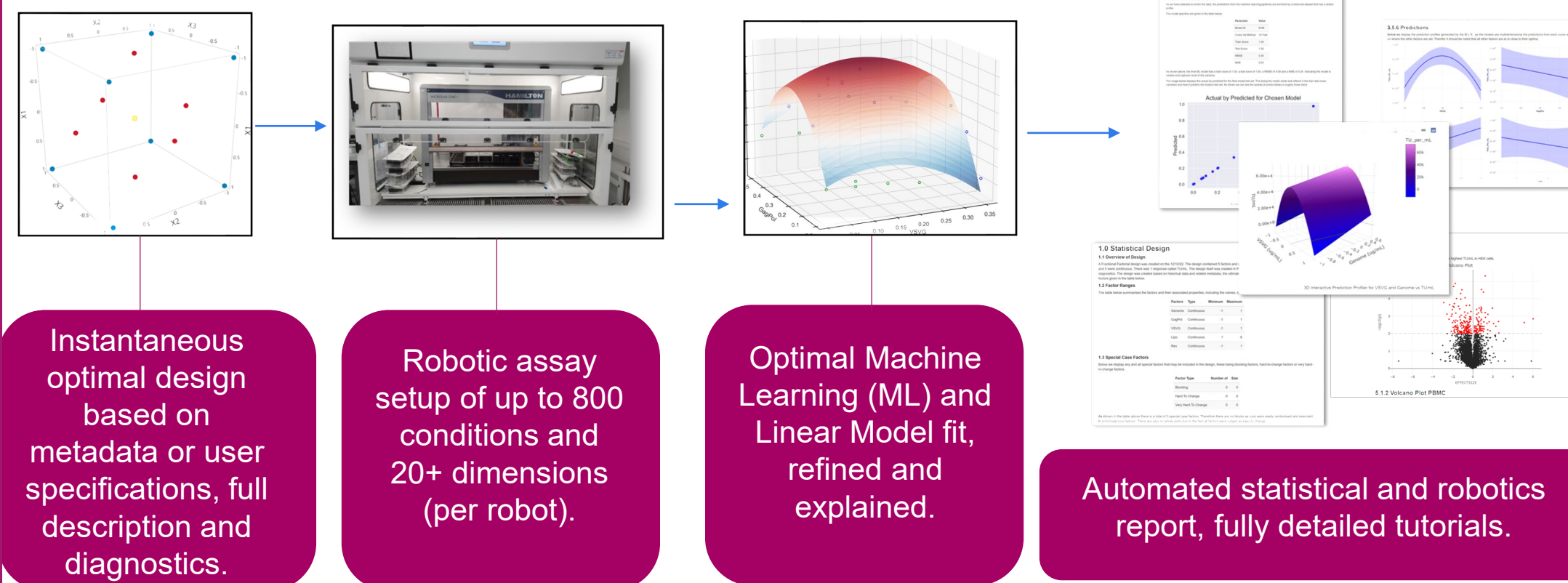


DoE methods have also been shown to identify global optimal conditions where OFAT might miss.



## Overview of System

State-of-the-art automated optimisation platform with applications in automated DoE, synthetic DNA optimisations, small molecule optimisation, AQBd and high-throughput screening.



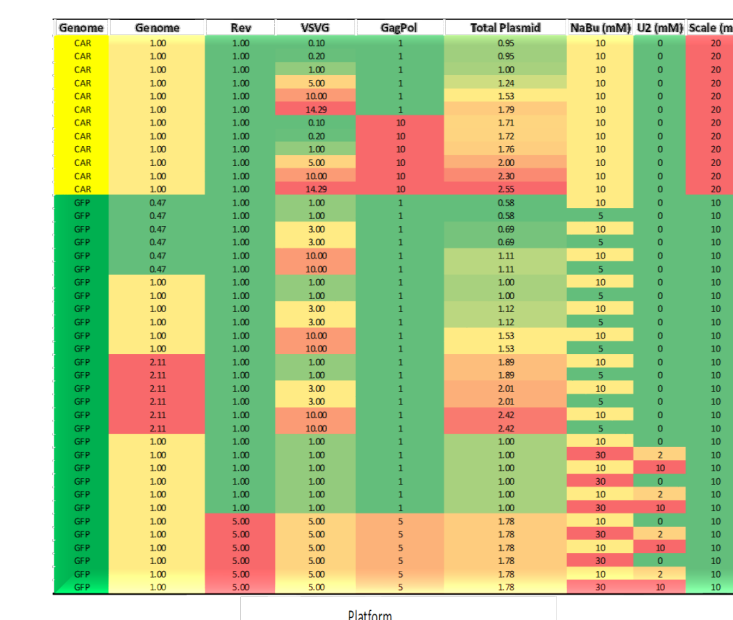
## Case Study: Side-by-Side Optimisation of HEK293T Titre and Primary T Cell Transductions Using Optimal Plasmid Ratios and Proteomics

Optimal plasmid ratios for HEK293T titre may not represent the optimal ratios for target cell transductions, in this case primary T cell transductions.

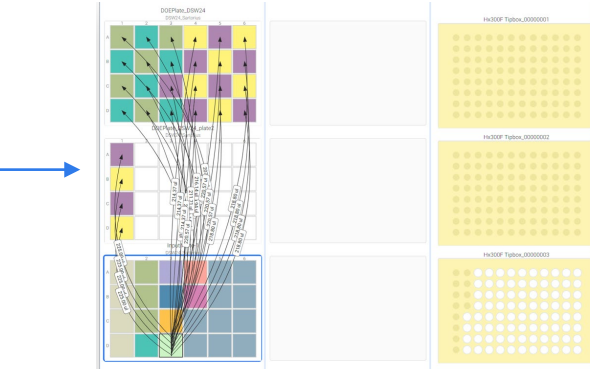
DoE parsed into Synthace Cloud platform software and executed using a Hamilton STARLet liquid handler.

Data from multiple sources joined, modelled and reported

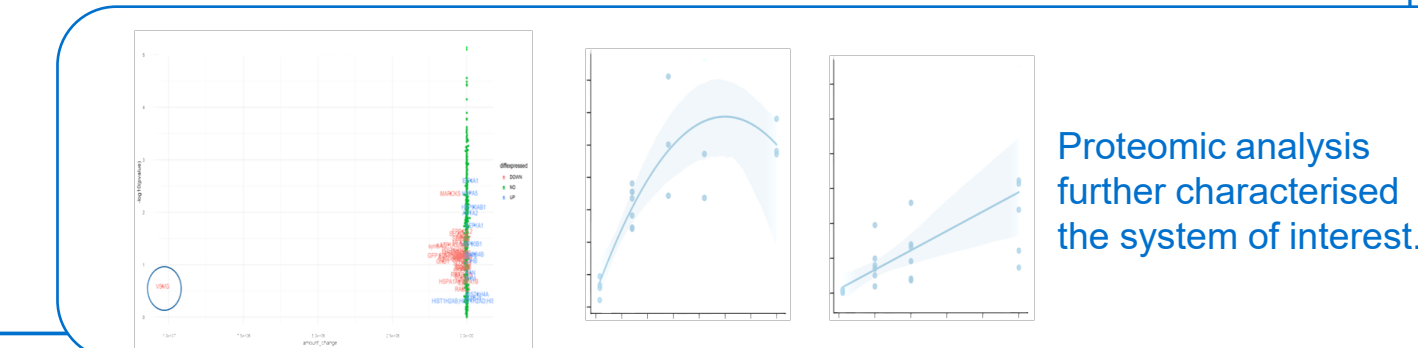
Exploratory analysis carried out to describe general trends.



Ratios over Platform concentrations



Linear modelling and ML used to characterise all optima.

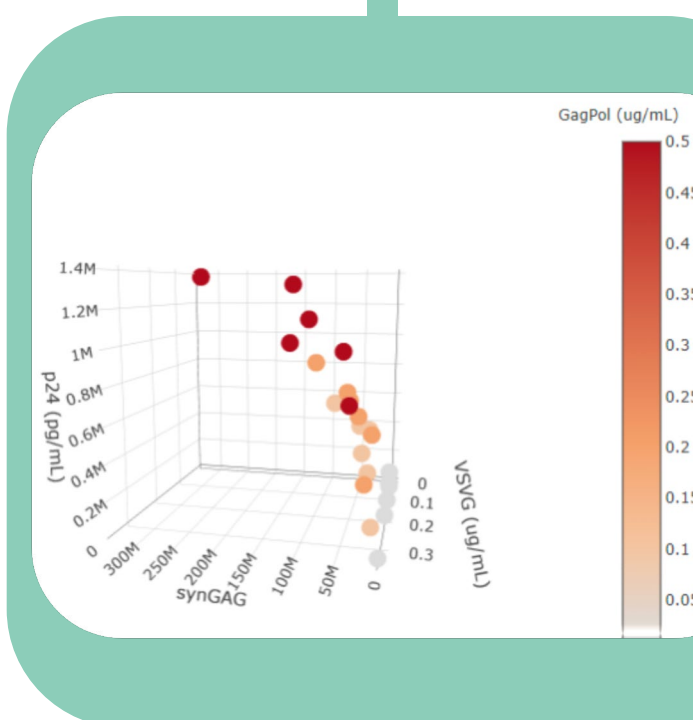
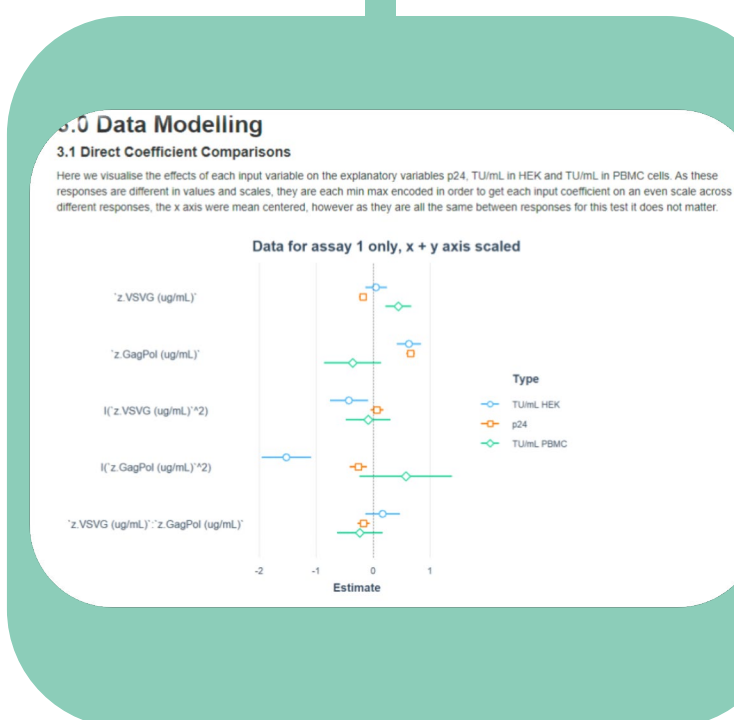


Proteomic analysis further characterised the system of interest.

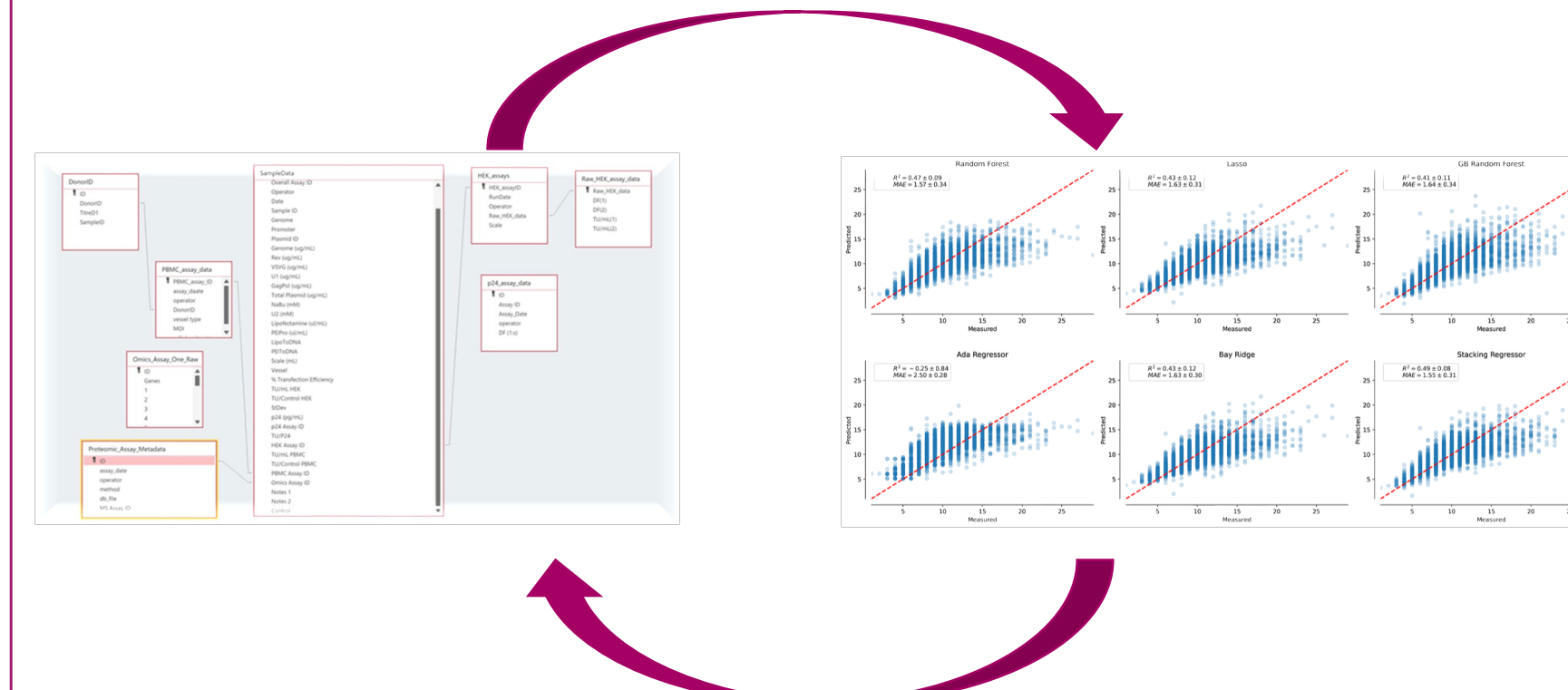
## Interactive Multi-Response Optimisation Including Proteomic Augmented Models

Simultaneous optimisation of multiple responses i.e. Primary T Cell Transduction and HEK293T TU/mL

Exploratory proteomic enriched models allowing plasmid – protein – Titre interactions

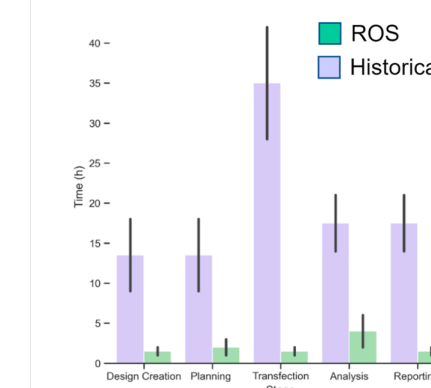


## Machine Learning and Data Enrichment



- Explore non-linear trends with up to ten machine learning kernels and a flexible neural network architecture, automatically run, refined and fully reported.
- Draw upon high quality curated historical data and associated metadata to add power to your model, fill in blank design space, and extrapolate out of the design space.

## Time, Accuracy & Precision



- Significant time saving in design, planning, analysis and reporting

- Full DoE repeated 2 times.
- Paired T test and REML run.
- Anonymized distributions shown on the right.

Test	Value	Result
T Test	0.9844	No evidence of difference
REML	0.68	High evidence of equivalence
Upper Bound P	<0.0001	>99.99% Probability that values differ by less than +/- 10%
Lower Bound P	<0.0001	

## Conclusions

- Enhanced High Throughput Optimisations**
  - 4000+ Runs (Flasks / Wells)
  - 100+ Input Variables
  - 10+ Response Variables
  - Model responses as functions of other responses

- Model Scale and Transgene Properties**
  - Multi Scale Systems
  - Multi Transgene Optimisations

- Flexibility**
  - Ability to characterise and optimise other measurable biological systems

- Time and Reproducibility**
  - Reduction in wet lab work as well as analysis by over 80%
  - Highly reproducible assays

- Data Analysis**
  - State of the art linear and non-linear modelling
  - Clear explanations and conclusions

