Precision Medicine for Amyotrophic Lateral Sclerosis: Combinatorial Analysis of Patient Genomes

A. Malinowski, K. Taylor, S. Das, C. Stubberfield
PrecisionLife Ltd, Oxford, United Kingdom

Scan to find out more

Introduction

PrecisionLife (PL) is a pioneering techbio company, with a unique approach that finds more signal in complex disease patient data than standard methods. Our high-resolution patient stratification identifies subgroups of patients with similar disease drivers and treatment responses, to make precision medicine possible in chronic diseases.

Amyotrophic lateral sclerosis (ALS) is characterized by a high degree of heterogeneity across the patient population, reflected in several subtypes with different disease progression rates, severity, age of onset and death. GWAS have identified several disease-associated genes, but these findings have shown limited translation into therapies, especially for sporadic disease. This likely reflects the limitations of GWAS in only identifying single variants, while the key to understanding complex diseases that are influenced by multiple genetic loci is to find combinations of variants that distinguish one patient subgroup from another.

Methods

DATASET:
PrecisionLife analysed genetic data from ALS patients found in the Project MinE dataset. These patients were split into two distinct cohorts based on the single SNP array used for genotyping, and analysed against healthy matched controls. The dataset did not contain SNPs related to known ALS genes such as *SOD1*. The patient information included additional clinical and phenotypical data.

Cohort 1:
610 ALS Cases, 1,046 Controls
Cohort 2:
736 ALS Cases, 1,472 Controls

COMBINATORIAL ANALYSIS:
The dataset was analysed in the PrecisionLife platform to identify combinations of SNP genotypes that when observed together in a patient are strongly associated with ALS. The phenotypic and clinical data for each of the patients was used to provide additional insights into the results generated.

SNP combinations that have high odds ratios, low *p*-values and high prevalence in cases are prioritized. This process utilises 1,000 cycles of fully randomized permutations and combinations must meet a specified FDR threshold.

SNPs are scored using a Random Forest algorithm in a 5-fold cross-validation framework and prioritized based on their ability to differentiate cases and controls.

The highest scoring SNPs are then mapped to genes and clustered by the patients they co-occur in to generate a disease architecture.

Results

Figure 1. Combinatorial analysis of genomic data vs GWAS

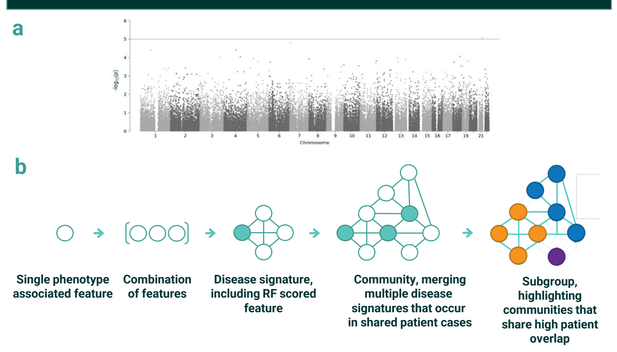

Figure 1. (a) Manhattan plot of genome-wide *p*-values of association for the AD UK-only Project MinE Cohort 2. The dashed line represents the genome-wide significance threshold at p<1e-05. (b) Conceptual representation of features, combinations, disease signatures and communities used to build up the disease architecture in the PrecisionLife combinatorial methodology.

GWAS	Combinatorial Analysis
Single SNP associations must be significant across large groups of patients	Specific combinations of variants associated with each patient subgroup serve as a genetic stratification biomarker
Limited insights unless disease is likely to be caused by a small number of rare variants with large effect sizes (often in gene coding regions affecting protein 3D structure)	Patient subgroups with different causes of disease or even incorrect diagnoses can be distinguished (stratified) by different mechanistic etiology
Does not account for the effects of interactions between SNPs, genes and metabolic networks	Captures epistatic and non-linear additive effects of all interactions between SNPs, genes, environmental factors and metabolic networks

Table 1. Key results from PrecisionLife's ALS study

	Cohort 1	Cohort 2
Validated disease signatures (SNP combinations)	201	74
Significant RF-scored SNPs associated with ALS	48	10
Significant genes associated with ALS	18	6
Genes targeted by at least one drug in clinical development*	4	0
Patient communities identified	9	8

*source: DrugBank and ChEMBL

Table 1. Summary of the results of PrecisionLife's combinatorial analysis of the two cohorts from the Project MinE patient dataset.

Figure 2. ALS disease architecture reveals different stratified genetic communities of patients

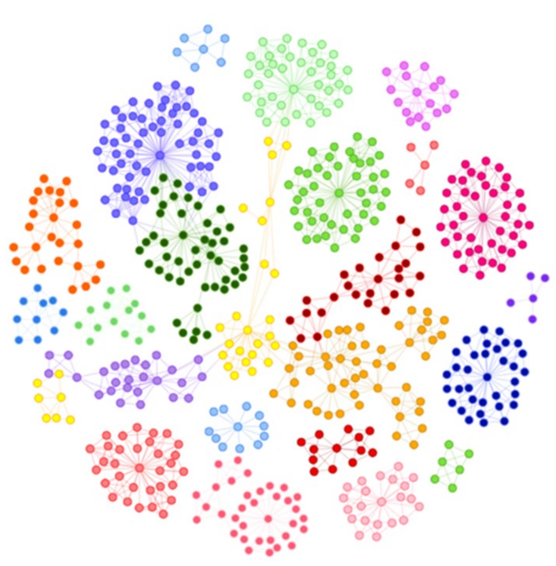

Figure 2. ALS disease architecture diagram demonstrating distinct communities of SNPs comprising the structure of the patient subpopulations generated by the PrecisionLife platform from combined cohort 1 and 2. Each circle represents a disease-associated SNP genotype, edges represent co-association in patients, and each color a distinct community of SNPs.

Figure 3. Genetic stratification of three clinical patient subgroups

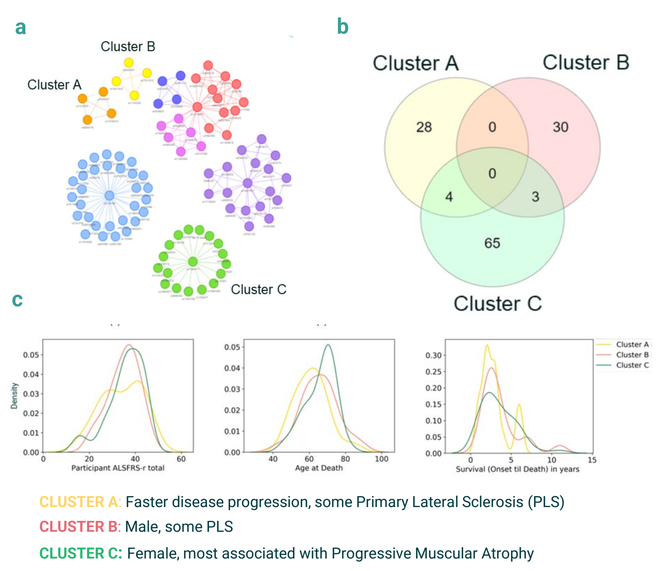

Figure 3. (a) Disease architectures of the patient populations generated by the PrecisionLife platform for ALS Cohort 2. Each circle represents a disease-associated SNP genotype; edges represent co-association in patients and colors represent distinct patient subpopulations. (b) Venn diagram showing the overlap of patients who are found in the clusters A, B and C. (c) Comparison of the distribution of three clinical features; ALSFRS-R, age at death and survival from disease onset until death between the three clusters.

Figure 4. Identification and prioritisation of ALS targets

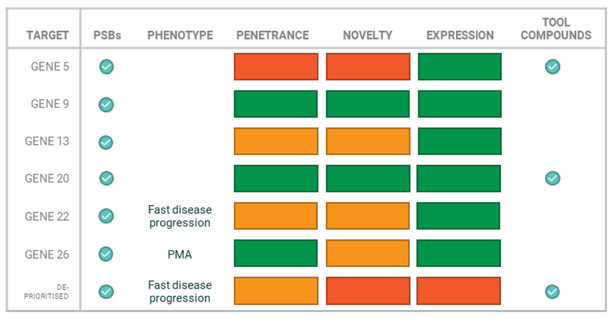

Figure 4. Identified ALS targets are evaluated against tractability criteria to facilitate target prioritisation. Unique insights from the combinatorial analysis include estimation of case penetrance and identification of patient stratification biomarkers for each target.

Figure 5. ALS-relevant target validation approaches

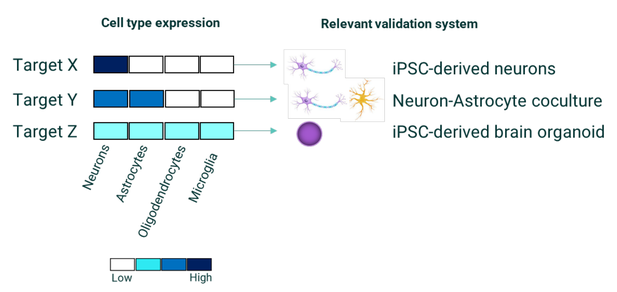

Figure 5. Target expression is evaluated in ALS-relevant cell types and for each target a specific human-centric validation model is selected based on the expression pattern and mechanism-of-action hypothesis. Validation models include simple iPSC-derived monocultures, co-cultures and complex organoids, combined with ALS-relevant readouts.

Figure 6. Systematic Indication Extension in ALS

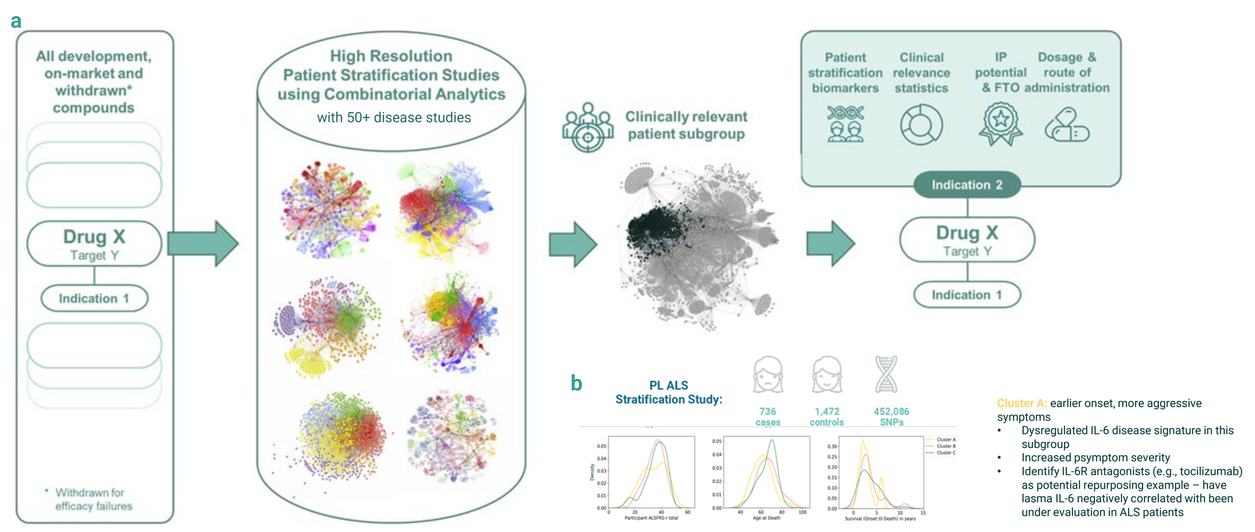

Figure 6. (a) Systematic drug indication extension approach based on high-resolution patient stratification insights generated by combinatorial analytics in over 50 disease studies. (b) IL-6R antagonist repositioning opportunity identified in patient subgroup with early onset and more aggressive form of the disease (Cohort 2 analysis). Adapted from https://doi.org/10.1016/j.patter.2022.100496.

Next Steps

LifeArc Partnership

PrecisionLife and LifeArc announced a strategic R&D collaboration to accelerate the discovery and development of precision medicine treatments for ALS. LifeArc will select multiple novel targets identified by PrecisionLife with supporting patient stratification biomarkers for collaborative validation and development.

Analysis of Project MinE Whole Genome Sequencing (WGS) Dataset

PrecisionLife has accessed Project MinE dataset with patient WGS and clinical information curated by Professor Ammar Al-Chalabi's Group based at King's College London. The dataset is currently being analysed on the PL platform as part of the LifeArc collaboration, in order to confirm and extend the results of previous SNP array-based studies.

Biological Validation of Novel Targets

LifeArc and PrecisionLife will work together to select and validate multiple novel targets together with accompanying patient stratification biomarkers. LifeArc researchers will lead and coordinate target validation projects using ALS-relevant *in vitro* and *in vivo* models.

Drug Repurposing

Among the genes identified within ALS subgroups, several are targeted by drugs in clinical development in other indications. We have developed a pipeline to systematically evaluate the potential of repurposing these to accelerate implementation of safe and effective therapies for ALS patients.

Conclusion

The results demonstrate that the PrecisionLife combinatorial analysis is uniquely able to mechanistically stratify heterogenous patient populations within ALS. We can use these insights to identify more effective therapeutic strategies and accompanying biomarker sets to match them to the patient subgroups that are most likely to demonstrate benefit in downstream clinical trials.

Acknowledgements

Research described in this study has been conducted using data from the Project MinE

precisionlife
improving health | **for everyone**

For more information, visit:
www.precisionlife.com