Program Booklet

Dear Participants,

On behalf of the Atlas of Variant Effects Alliance and the organizing committee, we would like to welcome you to the **2026 Mutational Scanning Symposium** (MSS26). The symposium aims to propel advancements in Multiplex Assays of Variant Effects (MAVEs) and mutational scanning technologies which are crucial for understanding the human genome.

This year, for the first time in Australia, we are bringing together diverse experts from around the world to discuss high throughput assays of variant impact. We hope this forum provides a stimulating environment for discussion and advancing research in functional genomics. Join us for insights into interpreting the genome and help shape the future of genomics.

MSS26 will feature keynote presentations from

Professor Matthew Hurles,
Director of the Wellcome Sanger Institute

and

Professor Maitreya Dunham,
University of Washington

among many other excellent presentations by genomics researchers and quantitative biologists from across the globe, as well as specialty application workshops on computational modelling and clinical variant interpretation.

From the Atlas of Variant Effects Alliance and the Mutational Scanning Symposium 2026 planning committee:

Irene Gallego Romero, SVI, Melbourne, Australia
Alan Rubin, WEHI, Melbourne, Australia
Rehan Villani, QIMR Berghofer, Brisbane, Australia
Lara Muffley, University of Washington, USA

Atlas of Variant Effects

1.1 Summary program

Wednesday March 25th			
Time	*Speaker*	*Institution*	*Title*
10:00-10:15	**Welcome to Country**		
Session 1 - Chair: Irene Gallego Romero			
10:15-11:00	**Fumitaka Inoue**	**Kyoto University**	***Deciphering human functional genome using Massively Parallel Reporter Assays***
11:00-11:15	Chanodya Pasansi Ranwala	University of Adelaide	Unraveling Regulatory Complexity: Functional Characterisation of Enhancer Variants associated with Common Epilepsy
11:15-11:30	Asfar Lathif Salaudeen	The University of British Columbia	Expanding the gene editing toolkit to decipher endogenous causal variants
11:30-11:45	**Coffee break**		
Session 2 - Chair: Lachlan Jolly			
11:45-12:00	Daniel V Brown	Walter and Eliza Hall institute of Medical Research	Active learning of the sequence-function landscape in a far-red fluorescent protein
12:00-12:15	Maddy Comerford	University of Melbourne	Mapping the gene regulatory landscape of archaic hominin introgression in modern Papuans
12:15:12-30	Justine Shih	Broad Institute of MIT and Harvard	Mapping Genetic Variation to Cellular Phenotypes with Targeted Optical Endogenous Sequencing
12:30-12:45	Bernd Willems	Twist Bioscience	Writing the Future of Mutational Scanning
12:45-13:45	**Lunch + posters on your own**		
Session 3 - Chair: Mandy Spurdle			
13:45-14:15	**Bryony Thompson**	**Royal Melbourne Hospital**	**Lessons Learned from Applying Functional Data in Real-World Clinical Variant Classification**
14:15-14:30	Melissa Gilbert	The Children's Hospital of Philadelphia & University of Pennsylvania	The clinical utility of Multiplexed Assays of Variant Effects (MAVEs) in a pediatric cohort
14:30-14:45	Richard James	Seattle Children's Research Institute	Functional Assessment of Genetic Variants of CARD11 and BCL10 with Saturation Genome Editing
14:45-15:00	Ben Capodanno	Brotman Baty Institute at the University of Washington	MaveMD: A functional data resource for genomic medicine
15:00-15:15	Emmylou C. Nicolas-Martinez	Adelaide University	RNA variants resolved in non-expressed genes by transactivation
15:15-15:45	**Coffee Break**		

Session 4 - Chair: John Christodoulos			
15:45-16:15	**Alex Wagner**	**Nationwide Children's Hospital**	**From Functional Assays to Computable Knowledge: Making MAVE Data Usable in Variant Interpretation**
16:15-16:30	Steven E. Brenner	University of California, Berkeley	Findings from the Critical Assessment of Genome Interpretation (CAGI), seventh edition: a community experiment to evaluate phenotype prediction
16:30-16:45	Ashley P.L. Marsh	Ambry Genetics	Clinical translation of a MUTYH MAVE into a diagnostic laboratory workflow enables large-scale VUS resolution
16:45-17:00	Frederick P. Roth	University of Pittsburgh	Landscapes of human AAT missense-variant effects reveal pathogenic variation and genetic suppressors

Thursday March 26th

Time	*Speaker*	*Institution*	*Title*
Session 5 - Chair: Rehan Villani			
9:00-9:45	**Matthew Hurles**	**Wellcome Sanger Institute**	***Keynote Speaker - Scaling MAVEs: are we at the end of the beginning?***
9:45-10:15	**Chai-Ann Ng**	**Victor Chang Cardiac Research Institute**	**Integrating MAVE and patch clamp data to improve classification and risk prediction in KCNH2-LQTS**
10:15-10:30	Andrew Glazer	Vanderbilt University Medical Center	Three Multimodal Assays of SCN5A Variant Function Inform Arrhythmia Risk Prediction
10:30-10:45	Sujatha Jagannathan	University of Colorado Anschutz Medical Campus	Single-codon resolution mapping of nonsense-mediated mRNA decay via genomic stop codon scanning of LMNA
10:45-11:00	Lea Starita	University of Washington	260,000 variant effect measurements and novel calibration methods drive resolution of VUS
11:00-11:30	**Coffee Break**		
Session 6 - Chair: Sandra Cooper			
11:30-12:00	**Nilah Ioannidis**	**University of California, Santa Cruz**	**Modeling the impact of personal genome variation on molecular phenotypes**
12:00-12:15	Omar Tariq	University of British Columbia	tfGPRA: A High Throughput Platform for Transcription Factor Characterization
12:15-12:30	Vanessa Burns	Wellcome Sanger Institute	Saturation Genome Editing to Clarify Variant Effect within 5’UTRs of Neurodevelopmental Disorder Genes
12:30-12:45	Lorenzo Vaccaro	Telethon Institute of Genetics and Medicine	Genotype-phenotype single-cell transcriptomics for massive parallel assessment of genetic variants
12:45-13:00	Daniel	Ambry Genetics	MAVE Progress Report, How Ambry Genetics Validates

	Zimmerman		and Deploys High-Throughput Functional Assays
13:00-14:00	**Lunch + posters on your own**		
Session 7 - Chair: Matthew Wakefield			
14:00-14:30	**Jian-Rong Yang**	**Sun Yat-sen University, Guangzhou**	**Phenotypic Mutations in Collision: Negative Epistasis Between Transcription Errors and Translation Errors Revealed by DMS**
14:30-14:45	Srivatsan Raman	University of Wisconsin-Madison	From Variant Maps to Functional Design: Using Deep Mutational Scanning and Machine Learning to Engineer Programmable Bacteriophages
14:45-15:00	Guillaume Diss	Friedrich Miescher Institute	The genetic architecture of the human bZIP interaction network
15:00-15:15	Jacob Purcell	Monash University	High-resolution functional assessment of SMAD4 variants
15:15-15:30	Rosa De Santis	Telethon Institute of Genetics and Medicine	From bulk to single cell: benchmarking analytical tools for Deep Mutational Scanning experiments
15:30-16:00	**Coffee Break**		
Session 8 - Chair: Sefi Rosenbluh			
16:00-16:30	**Amelie Stein (v)**	**University of Copenhagen**	**Variant consequences in context**
16:30-16:45	Xinyu Wu	Walter and Eliza Hall institute of Medical Research	Elucidate the active conformation of cytokine receptors using deep mutational scanning
16:45-17:00	Polina V Polunina	University of Freiburg	Modeling Viral Evolution as Sequence Transitions: A Transformer Approach Using Time-Resolved SARS-CoV-2 Spike Data
17:00-17:30	**Selected poster short talks**		
Poster short talks chair: Fritz Roth			
17:30-19:30	**Catered poster session**		
19:45-late	**Social function at the Gertrude Hotel (148 Gertrude St, five minute walk)**		

Friday March 27th

Time	*Speaker*	*Institution*	*Title*
Session 9 - Chair: Alan Rubin			
9:00-9:45	**Maitreya Dunham**	**University of Washington**	***Keynote Speaker: From ARS1 to ZWF1: some adventures in deep mutational scanning***
9:45-10:15	**Mark Dawson**	**Peter MacCallum Cancer Centre**	**Characterising the functional landscape of the human PRC2 complex by base editing at single cell**

			resolution
10:15-10:30	Maximilian Stammnitz (v)	Centre for Genomic Regulation	The genetic architecture of allosteric plant hormone receptors
10:30-10:45	Jun J Yang	St. Jude Children's Research Hospital	MAVE-informed Protein Language Models for Predicting Pharmacogenetic Variant Function at Scale
10:45-11:00	Douglas M. Fowler	University of Washington	Biochemical profiling of ~315,000 MAP kinase pathway protein variants in human cells with LABEL-seq
11:00-11:30	**Coffee Break**		
Session 10 - Chair: Ebony Matotek			
11:30-12:00	**Chris Hahn**	**SA Pathology and University of South Australia**	**Deep mutational scanning to measure the transactivational effect of GATA2 and ERG variants to benefit clinical interpretation**
12:00-12:15	Alistair Dunham	Wellcome Sanger Institute	Standardisation and joint analysis of 2,041 MAVEs
12:15-12:30	Yu-Jen (Jennifer) Lin	University of California, Berkeley	Critical Assessment of Genome Interpretation (CAGI) 7 ARSA missense stability prediction challenge identifies computational advances over state-of-the-art variant impact predictors
12:30-12:45	Linda Wijaya	The Kids Research Institute Australia & University of Western Australia	Unravelling the RASopathy Syndromes using iPSC-derived neural disease modelling
12:45-13:00	Illumina	Illumina Inc	
13:00-14:00	**Lunch (no posters)**		
Session 11 - Chair: Convenors			
14:00-14:30	**Melissa Call**	**Walter and Eliza Hall institute of Medical Research**	**Mutational profiling of SARS-CoV-2 PLpro exposes inhibitor escape routes**
14:30-14:45	Awards and closing		
14:45-15:00	**Coffee Break**		
15:00-17:00	**Workshop**		**Clinical Application Workshop**

1.1 Summary program **1**
1.2 Full information **1**
1.2.1 Date 1
1.2.2 Location 1
1.2.3 The Host 1
1.3 Wifi **1**
1.4 Sponsors **2**
1.5 Talks **3**
Wednesday March 25th 2026 **3**
Deciphering human functional genome using Massively Parallel Reporter Assays 3
Fumitaka Inoue 3
Unraveling Regulatory Complexity: Functional Characterization of Enhancer Variants associated with Common Epilepsy 4
Chanodya Pasansi Ranwala 4
Expanding the gene editing toolkit to decipher endogenous causal variants 5
Asfar Lathif Salaudeen 5
Active learning of the sequence-function landscape in a far-red fluorescent protein 6
Daniel V Brown 6
Mapping the gene regulatory landscape of archaic hominin introgression in modern Papuans 7
Maddy Comerford 7
Mapping Genetic Variation to Cellular Phenotypes with Targeted Optical Endogenous Sequencing 8
Justine Shih 8
Writing the Future of Mutational Scanning 9
Bernd Willems 9
Lessons Learned from Applying Functional Data in Real-World Clinical Variant Classification 10
Bryony Thompson 10
The clinical utility of Multiplexed Assays of Variant Effects (MAVEs) in a pediatric cohort 11
Melissa Gilbert 11
Functional Assessment of Genetic Variants of CARD11 and BCL10 with Saturation Genome Editing 12
Richard James 12
MaveMD: A functional data resource for genomic medicine 13
Ben Capodanno 13
RNA variants resolved in non-expressed genes by transactivation 14
Emmylou C. Nicolas-Martinez 14

From Functional Assays to Computable Knowledge: Making MAVE Data Usable in Variant Interpretation 15
Alex Wagner 15
Findings from the Critical Assessment of Genome Interpretation (CAGI), seventh edition: a community experiment to evaluate phenotype prediction 16
Steven E. Brenner 16
Clinical translation of a MUTYH MAVE into a diagnostic laboratory workflow enables large-scale VUS resolution 17
Ashley P.L. Marsh 17
Landscapes of human AAT missense-variant effects reveal pathogenic variation and genetic suppressors 18
Frederick P. Roth 18
Thursday March 26th 2026 **19**
Scaling MAVEs: are we at the end of the beginning? 19
Matthew Hurles 19
Integrating MAVE and patch clamp data to improve classification and risk prediction in KCNH2-LQTS 20
Chai-Ann Ng 20
Three Multimodal Assays of SCN5A Variant Function Inform Arrhythmia Risk Prediction 21
Andrew Glazer 21
Single-codon resolution mapping of nonsense-mediated mRNA decay via genomic stop codon scanning of LMNA 22
Sujatha Jagannathan 22
260,000 variant effect measurements and novel calibration methods drive resolution of VUS 23
Lea Starita 23
Modeling the impact of personal genome variation on molecular phenotypes 24
Nilah Ioannidis 24
tfGPRA: A High Throughput Platform for Transcription Factor Characterization 25
Omar Tariq 25
Saturation Genome Editing to Clarify Variant Effect within 5’UTRs of Neurodevelopmental Disorder Genes 26
Vanessa Burns 26
Genotype-phenotype single-cell transcriptomics for massive parallel assessment of genetic variants 27
Lorenzo Vaccaro 27
MAVE Progress Report, How Ambry Genetics Validates and Deploys High-Throughput Functional Assays 28
Daniel Zimmerman 28
Phenotypic Mutations in Collision: Negative Epistasis Between Transcription Errors and Translation Errors Revealed by DMS 29
Jian-Rong Yang 29

From Variant Maps to Functional Design: Using Deep Mutational Scanning and Machine Learning to Engineer Programmable Bacteriophages 30
Srivatsan Raman 30
The genetic architecture of the human bZIP interaction network 31
Guillaume Diss 31
High-resolution functional assessment of SMAD4 variants 32
Jacob Purcell 32
From bulk to single cell: benchmarking analytical tools for Deep Mutational Scanning experiments 33
Rosa De Santis 33
Deep mutational scanning to measure the transactivational effect of GATA2 and ERG variants to benefit clinical interpretation 34
Chris Hahn 34
Elucidate the active conformation of cytokine receptors using deep mutational scanning 35
Xinyu Wu 35
Modeling Viral Evolution as Sequence Transitions: A Transformer Approach Using Time-Resolved SARS-CoV-2 Spike Data 36
Polina V Polunina 36
Friday March 27th 2026 37
From ARS1 to ZWF1: some adventures in deep mutational scanning 37
Maitreya Dunham 37
Characterising the functional landscape of the human PRC2 complex by base editing at single cell resolution 38
Mark Dawson 38
The genetic architecture of allosteric plant hormone receptors 39
Maximilian Stammnitz 39
MAVE-informed Protein Language Models for Predicting Pharmacogenetic Variant Function at Scale 40
Jun J Yang 40
Biochemical profiling of ~315,000 MAP kinase pathway protein variants in human cells with LABEL-seq 41
Douglas M. Fowler 41
Variant consequences in context 42
Amelie Stein 42
Standardisation and joint analysis of 2,041 MAVEs 43
Alistair Dunham 43
Critical Assessment of Genome Interpretation (CAGI) 7 ARSA missense stability prediction challenge identifies computational advances over state-of-the-art variant impact predictors 44
Yu-Jen (Jennifer) Lin 44
Unravelling the RASopathy Syndromes using iPSC-derived neural disease modelling 45

Linda Wijaya 45
Illumina Talk title 46
Illumina 46
Mutational profiling of SARS-CoV-2 PLpro exposes inhibitor escape routes 47
Melissa Call 47
1.6 Posters (Alphabetical, number) 48
Predicting changes in protein-protein binding affinity upon mutation with statistical potentials 49
A human cell model system for interrogation of ATAD3 variants, a frequent cause of severe mitochondrial disease 50
High-Throughput DNA Synthesis Enables Large-Scale Mapping of Protein Sequence–Function Relationships 51
Mapping the effects of all mutations on neutralization by antibodies and engineered therapeutics for HIV Envelopes from divergent strains. 52
Rules of the Rod: MYH7 Variant Effect Across the β-MHC Coiled-Coil Domain 53
Deep Characterization of PTEN Variants on Canonical and non-Canonical Signaling Pathways 54
A Bioinformatics Ecosystem for High Throughput MAVE-SGE Experiments 55
Integrating genomic and functional approaches to investigate differences of sex development and SRY regulation 56
It’s Not Just the Variant—It’s the Interface: A Human-Centered Look at MaveDB 57
Functional classification of population variants demonstrates that SLC6A19 loss-of-function is associated with reduced progression to kidney failure 58
Advancing the Laboratory Foundations of a Scalable Saturation Genome Editing Pipeline 59
Scaling DNA Synthesis Enhances Gene Synthesis and DNA Variant Library Generation 60
Importance of higher-order epistasis in protein sequence-function relationships 61
Human CRISPR-oligo recombineering of a ribosomal protein reveals a functional S-acylated cysteine 62
Clinical Impact of Integrating BRCA2 MAVE Data for Variant Reclassification: A Real-World Evaluation 63
DISCOVER, CONNECT, CONTRIBUTE: ENGAGEMENT PATHWAYS WITHIN THE ATLAS OF VARIANT EFFECTS ALLIANCE 64
Understanding the delivery of an ASO therapy for Epidermolysis Bullosa 66
Differentiation of IPS cells into fibroblasts to validate gene targets for Dupuytren’s Contracture 67
Deep insight into protein through Deep Mutational Scanning. 68
The impact of Polygenic Risk Scores on a variable penetrance alleles of BRCA1 69
Investigating the mechanism of a non-coding deletion variant in the promoter of beta-cell repressed gene SLC16A1 causing exercise-induced hyperinsulinism. 70
Allosteric mapping for the rapid discovery of novel therapeutics 71

CountESS, a user-friendly platform for analysing MAVE data 72
Modelling motile cilia for a DNAI2 variant of uncertain significance using CRISPR edited iPSCs and directed differentiation 73
NAXD deficiency: heterogeneous phenotypes, positive response to niacin treatment and drug repurposing studies 74
Variant Effect Predictors Show Systematic Bias in Intrinsically Disordered Regions 75
Unravelling CACNA1A Channelopathies: Precise CRISPR Gene Editing and Transcriptomic Profiling for Rapid Functional Analysis of Rare Genetic Variants 76
Transdifferentiation of Human Dermal Fibroblasts to iNeurons and iAstrocytes for Functional Variant Studies 77
ELIXIR Activities Supporting a FAIR Future for Life Science 78
Implementing MaveDB to identify functionally impactful pharmacogenomic variants in underrepresented Indonesian populations 79
Functional mapping of paralog synthetic lethal domains using base editing: Bridging knockout screens and therapeutic discovery. 80
Development of a CRISPR-oligorecombineering (CORe) and imaging platform for high-throughput profiling of phospho-variants in disease 81
Disease modelling and variant interpretation in the inherited retinal dystrophies using human pluripotent stem cells. 82
Multidimensional deep mutational scanning of the ALS protein SOD1: dissecting the effect of mutations on protein abundance and dimer formation 83
Mapping NTD of Spike protein utilising Deep Mutational Scanning to predict the unfolding evolution of SARS-CoV-2 84
Allosteric regulation of KRAS 85
Community co-designed workshops build confidence in use of functional evidence for variant classification. 86
1.7 Online Posters **87**
How do we define 'truthsets' for clinical validation of functional assays? 87
A systematic approach benign missense variant "truthsets" for validation of functional assays 88
Multiplexed assays of variant effects enhance gene discovery in de novo gene burden testing 90
Saturation genome editing of CTCF resolves variants of uncertain significance and maps domain-specific functional constraint 91
nf-core/deepmutscan: a community-oriented, modular bioinformatics pipeline for deep mutational scanning by shotgun sequencing 92
Precision missense mutagenesis for deciphering the functional impact of disease-associated MED27 variants. 93
1.8 The workshop **94**
1.9 Program committee **94**
1.10 Planning committee **94**

1.2 Full information

1.2.1 Date

25-27th March 2026

1.2.2 Location

MSS2026 will take place at the Aikenhead Centre for Medical Discovery (ACMD) in Melbourne, Australia. ACMD is a new centre for biomedical engineering and translational research that brings together multiple leading national universities and research institutes with a major tertiary hospital. Its unique combination of research and industry tenants means that it is uniquely positioned to tackle some of the world's most challenging health problems, and the building has been purpose-built to foster collaborative research, with multiple floors set aside for industry partners and start-ups.

ACMD is co-located at St Vincent's Hospital Melbourne, at the corner of Nicholson St and Victoria Parade, with fantastic access for local attendees and visiting guests both.

1.2.3 The Host

The Atlas of Variant Effects (AVE) Alliance was founded in 2020 and has grown to include over 700 members from more than 50 countries.

The Mutational Scanning Symposium and Workshop is our flagship annual event that brings together experts in functional genomics, protein science, precision medicine, variant interpretation, and computational genetics. The symposium aims to propel advancements in Multiplex Assays of Variant Effects (MAVEs) and mutational scanning technologies which are crucial for understanding the human genome.

1.3 Wifi

Eduroam is available throughout the ACMD building and the easiest way to connect to the internet. If you do not have access to eduroam, there is a guest network available in the auditorium and foyer. The access details are available in the room.

1.4 Sponsors

SPONSORS

GOLD

SILVER

The Collaborative Centre for
Genomic Cancer Medicine
A joint venture of the University of Melbourne and Peter MacCallum Cancer Centre

BRONZE

VENUE SPONSOR

1.5 The workshop

CLINICAL APPLICATION WORKSHOP

Enabling clinical translation of MAVEs; developing tools and resources

ACMD, March 27th 2026 - 15.00h -16:00h AEST

Multiplex Assays of Variant Effect (MAVE)s provide high-throughput functional data that can help elucidate the biological and clinical impact of variants. To translate this data into clinically actionable insights, the data needs to be evaluated and calibrated so that it can be applied appropriately for clinical use. The workshop is intended for clinicians, researchers and bioinformaticians alike.

Participants will learn about methods for using high throughput assays in clinical diagnostics, and will be introduced to some of the new processes and tools available, but also will have the opportunity to inform some of the field approaches to solving some of the challenges in the field. Through sharing and building resources, this workshop aims to enable those interested to translate experimental data to functional evidence for clinical use.

First, we will discuss current and developing guidance for translating functional data into evidence for variant classification. This will help participants understand how MAVE data is used in a clinical laboratory and how to apply the data in a way that aligns with the ACMG/AMP/ClinGen framework. Next, we will present new tools and resources for sharing, evaluating and using MAVE-derived functional evidence. The workshop will then focus on the new MaveMD, and will include participant consultation, extracting collated feedback to inform MaveMD and future tool development in the field. Participants will work through some demonstrator cases as instructional examples for how these tools might be applied in practice.

Presenters

- **Professor Lea Starita**, Department of Genome Sciences, University of Washington; Brotman Baty Institute
- **Associate Professor Alex Wagner**, Nationwide CHildren's Hospital
- **Jeremy Stone** - MaveDB, Brotman Baty Institute at the University of Washington
- **Sally Grindstaff** - MaveDB development team, Brotman Baty Institute.
- **Estelle Da** - MaveDB development team, University of Melbourne

Facilitator Rehan Villani

Register here - Register for the MSS26 Workshop

1.6 Program committee

Many thanks to our program committee for helping put together this meeting!

Alan Rubin (University of Melbourne)

Ebony Matotek (Australian Functional Genomics Network)

Irene Gallego Romero (St Vincent's Institute of Medical Research)

Lachlan Jolly (The University of Adelaide)

Matthew Wakefield (Walter and Eliza Hall Institute)

Rehan Villani (QIM Berghofer)

Sefi Rosenbluh (Monash University)

Twishi Gulati (Victorian Centre for Functional Genomics, Peter MacCallum Cancer Centre)

Vanessa Fear (The Kids Research Institute Australia)

1.7 Talks

Wednesday March 25th 2026

Deciphering human functional genome using Massively Parallel Reporter Assays

Fumitaka Inoue
Kyoto University

98% of our genome is non-coding, and its function remains largely unknown. The non-coding genome contains a variety of enhancers that play crucial roles in gene regulation. Mutations and variations in enhancers are thought to be a major source of human diseases, individual differences, and evolution.

lentivirus-based massively parallel reporter assay (lentiMPRA), a cutting-edge technology that enables high-throughput functional characterization of enhancers at single-nucleotide resolution by quantifying transcribed barcodes. We have leveraged the lentiMPRA technology to dissect enhancer function in the human genome and understand the molecular basis of gene regulation underlying complex biological phenomena, such as cell differentiation, disease and evolution.

In my presentation, I will introduce the lentiMPRA methodology and various applications, and discuss the future direction of functional genomics.

Unraveling Regulatory Complexity: Functional Characterization of Enhancer Variants associated with Common Epilepsy

Chanodya Pasansi Ranwala

University of Adelaide

Ranwala C [1,2, *], Baer L [1,2], Sadlon T [2,3] , Munro J [4], Bahlo M [4], Barry S [3], Gecz J [1,5], and Jolly L [1,2]

[1] Robinson Research Institute, Adelaide, 5005, SA, Australia

[2] School of Biomedicine, The University of Adelaide, Adelaide, 5005, SA, Australia

[3] Discipline of Paediatrics, Women's and Children's Health Research Institute, The University of Adelaide, Adelaide, 5005, SA, Australia

[4] The Walter and Eliza Hall Institute of Medical Research, Parkville, 3052, Australia

[5] Adelaide Medical School, University of Adelaide, Adelaide, SA, Australia

Epilepsies are a heterogeneous group of neurological disorders characterized by unprovoked seizures, affecting 70 million individuals worldwide. 95% of epilepsies are categorized as 'common epilepsies' and underpinned by a complex polygenic etiology that remains unresolved. Recent common epilepsy GWAS identified an enrichment of 1874 variants located within enhancer regions, but their functional impact remains undetermined. We applied a lentiviral Massively Parallel Reporter Assay (lentiMPRA) in the context of neural progenitor cells (NPCs) to assess the gene regulatory effects of these variants.

A lentiMPRA library consisting of the 1874 enhancer variants in both alternative or reference GWAS variant configurations were cloned upstream of a library of synthetic barcode genes, packaged into lentivirus and delivered to NPCs at high efficiency. Combined short read sequencing of DNA and RNA derived from the transduced cells facilitated functional assessment. A total of 60 enhancer variants were shown to affect enhancer function, with 33 and 27 SNPs demonstrating repressive and activating effect respectively. Notably, 49 of these variants correspond to known eQTLs for key regulators of neuronal excitability, including KCNIP, CDKRAP3, and ZDHHC14. Furthermore, 39 were predicted to disrupt binding of transcription factors such as KLF4, CREB1 and REST which link gene regulatory pathways to neuronal excitation. Collectively, these results link enhancer variants to pathways that may drive epileptogenesis.

This study establishes lentiMPRA in the context of human NPCs as a powerful tool to dissect the complex genetic architecture of common epilepsies and its functional impact on epilepsy gene regulatory networks which are perturbed in pathology.

Expanding the gene editing toolkit to decipher endogenous causal variants

Asfar Lathif Salaudeen

The University of British Columbia

Asfar Lathif Salaudeen1*, Trevor Shyiak1, Nicholas Mateyko1, Carl de Boer2

1 Genome Science and Technology Graduate Program, The University of British Columbia, Vancouver, BC, Canada, V6T 1Z4

2 School of Biomedical Engineering, The University of British Columbia, Vancouver, BC, Canada, V6T 1Z3

Identifying causal variants that drive phenotypic changes remains a major challenge in functional genomics. While genome-wide association and eQTL studies have mapped numerous loci linked to various phenotypes, they cannot disentangle causal mutations from passengers due to linkage disequilibrium. This problem can also be extended to understanding the logic and the impact of mutations within regulatory regions such as promoters and enhancers. Experimentally perturbing the genome in an endogenous context, is therefore essential for establishing causality. We adapted an existing Virus Like Particle (VLP) based CRISPR–Cas9 base editor delivery system to generate a suite of base editors with diverse mutational spectra. Because VLPs transiently deliver the base editor ribonucleoproteins without integrating or continuously expressing the nuclease, they minimize cellular stress and reduce the risk of off-target activity. By using these base editor VLPs either in combination or in a sequential manner, we can efficiently induce multiplexed mutagenesis across diverse genomic targets directly within their endogenous context. This approach allows us to generate libraries of variants at scale without requiring labor-intensive cell engineering. As a proof of concept, we applied this system to the promoter of the PD-L1 gene, a critical player in immune evasion within the tumor microenvironment. By tiling the PD-L1 promoter sequences with random mutations and quantifying the expression changes of PD-L1 via Sort-Seq, we generated a sequence-to-expression dataset that revealed key nucleotides modulating PD-L1 expression. Our base editor VLP system provides a scalable and generalizable approach to decipher genome functions and distinguish causal from passenger variants, thereby enabling systematic insights into genotype to phenotype relationships across various genomic loci.

Active learning of the sequence-function landscape in a far-red fluorescent protein

Daniel V Brown
Walter and Eliza Hall institute of Medical Research

Brown DV 1,2,*, Hill T 1,2, Xu A 1,2, Pranggono R 1,2, Bowden R 1,2

1 - The Walter and Eliza Hall Institute of Medical Research, Melbourne, VIC, Australia.

2 - Department of Medical Biology, University of Melbourne, VIC, Australia

Fluorescent proteins are fundamental tools for cellular imaging. Aequorea victoria–derived proteins require oxygen and are excited by short wavelengths that poorly penetrate tissue. Far-red fluorescent proteins originating from cyanobacteria overcome many of these limitations by enabling excitation with deeper-penetrating light, but their relatively low brightness has hindered widespread use.

To address this challenge, we applied machine-learning-guided directed evolution to enhance the brightness of miRFPnano3. Our workflow combines small-scale experimental cycles, automation and cell-free expression, allowing rapid testing of protein variants. Sequence–function data from each round were used to iteratively update AI models through Bayesian optimisation.

We identified miRFPnano3 variants exhibiting four-fold increased brightness, including mutations that would not have been predicted from structural intuition alone. Compared with traditional pooled mutational scanning, our strategy required dramatically fewer experimental observations and significantly lower cost while accelerating convergence on improved variants.

These results illustrate the efficiency of combining machine learning with minimal laboratory data and demonstrate a generalizable framework for mapping and optimizing sequence–function landscapes in protein engineering.

Mapping the gene regulatory landscape of archaic hominin introgression in modern Papuans

Maddy Comerford

University of Melbourne

Maddy Comerford 1,2,*, Davide M. Vespasiani 2, Navya Shukla 1,2, Laura E. Cook 2, Danat Yermakovich 3,4, Michael Dannemann 3, Matthew Leavesley 5,6,7, Christopher Kinipi 8, François-Xavier Ricaut 4, Nicolas Brucato 4, Murray P. Cox 9,10, Irene Gallego Romero 1,3,11,12

1 - Human Genomics and Evolution, St Vincent's Institute of Medical Research, Melbourne, Australia
2 - School of Biosciences, University of Melbourne, Melbourne, Australia
3 - Center for Genomics, Evolution and Medicine, Institute of Genomics, University of Tartu, Tartu, Estonia
4 - Centre de Recherche sur la Biodiversité et l'Environnement (CRBE), Université de Toulouse, CNRS, IRD, Toulouse INP, Université Toulouse 3 – Paul Sabatier (UT3), Toulouse, France
5 - Strand of Anthropology, Sociology and Archaeology, School of Humanities and Social Sciences, University of Papua New Guinea, Port Moresby, Papua New Guinea
6 - ARC Centre of Excellence for Australian Biodiversity and Heritage, College of Arts, Society and Education, James Cook University, Cairns, Australia
7 - ARC Centre of Excellence for Indigenous and Environmental Histories and Futures, College of Arts, Society and Education, James Cook University, Cairns, Australia
8 - Health Services, University of Papua New Guinea, Port Moresby, Papua New Guinea
9 - Department of Statistics, University of Auckland, Auckland, New Zealand
10 - College of Sciences, Massey University, Palmerston North, New Zealand
11 - Faculty of Health Sciences, Australian Catholic University, Melbourne, Australia
12 - School of Medicine, University of Melbourne, Melbourne, Australia

Interbreeding between anatomically modern humans and archaic hominins has contributed to the genomes of present-day human populations. However, our understanding of the specific gene regulatory consequences of Neanderthal, and particularly, Denisovan introgression is limited. Here, we used a massively parallel reporter assay to investigate the regulatory effects of 25,869 high-confidence introgressed SNPs segregating in present-day individuals of Papuan genetic ancestry in immune cell types. Overall, 8.22% of Denisovan and 8.58\% of Neanderthal sequences showed active regulatory activity, and 9.22% of these displayed differential activity between archaic and modern alleles. We found no association between introgressed allele frequency on activity regardless of introgression source, but introgressed Denisovan alleles at higher frequencies were less likely to be differentially active than expected, suggesting introgression is under some degree of selective constraint. Both activity and differentially activity were associated with distance to the nearest transcription start site, while differential activity was additionally associated with differential transcription factor binding. Genes predicted to be regulated by differentially active sequences included IFIH1 and TNFAIP3, key immune genes and known examples of archaic introgression. Overall, this work provides experimental validation of regulatory activity for thousands of archaic variants in populations with the highest levels of Denisovan ancestry worldwide, revealing how human evolutionary history actively shapes present-day genetic diversity and immune function.

Mapping Genetic Variation to Cellular Phenotypes with Targeted Optical Endogenous Sequencing

Justine Shih
Broad Institute of MIT and Harvard

Justine Shih 1,2*, Russell Walton 1,2, Paul C. Blainey 1,2,3

1 - Broad Institute of MIT and Harvard, Cambridge, MA, USA

2 - Department of Biological Engineering, MIT, Cambridge, MA, USA

3 - Koch Institute for Integrative Cancer Research at MIT, Cambridge, MA, USA

Single cell sequencing has revolutionized our ability to dissect molecular variation in individual cells. However, these methods often lose spatial and morphological context. Optical pooled screening (OPS) bridges that gap by capturing rich, multidimensional phenotypes, such as morphology, signaling dynamics, and subcellular localization patterns, at scale. Yet OPS is limited to indirect readouts rather than sequencing the genomic variation itself. Here, we report an approach for sequencing select mutations at the endogenous locus to directly link genetic variants with their optical cellular phenotypes. Integration of this direct readout with base editing screens allows us to systematically map the effects of diverse genetic variants across millions of single cells. This work establishes a high-throughput, imaged-based framework for measuring the cellular phenotypic effects of genetic variation, enabling comprehensive characterization of genetic variants through multiplexed molecular and morphological readouts.

SPONSOR TALK

Writing the Future of Mutational Scanning

Bernd Willems

Twist Bioscience

Multiplexed assays of variant effect (MAVEs) and deep mutational scanning are transforming our ability to map sequence–function relationships and interpret genetic variation, central themes of the Mutational Scanning Symposium 2026 (MSS2026). This presentation highlights how scalable, high-fidelity DNA synthesis by Twist Bioscience enables the design and construction of variant libraries required for these experiments. We present approaches spanning gene synthesis, pooled DNA, and precision variant libraries that support systematic functional interrogation of regulatory elements and proteins. Case studies illustrate how synthetic DNA technologies accelerate variant effect mapping, enabling reproducible, large-scale experiments that connect genotype to phenotype and advance functional genomics and precision medicine.

Lessons Learned from Applying Functional Data in Real-World Clinical Variant Classification

Bryony Thompson
Royal Melbourne Hospital

This presentation aims to provide the point of view of a stakeholder (a clinical scientist in a diagnostic laboratory) who applies functional assays evidence in day-to-day variant curation and the challenges that can arise from the interpretation of this evidence. This is from the perspective of the pathology department of one of the largest adult hospitals (the Royal Melbourne Hospital) in Australia where the genetic testing covers a variety of conditions/specialties (e.g. nephrology, cardiology, neurology, endocrinology) and thus encountering a diversity of disease associated genes. The challenges that variant curators face when trying to interpret functional assay evidence in the process of variant classification will be highlighted. Particularly the priorities in variant classification in a diagnostic laboratory, such as the benefits of clinically calibrated functional assays and highly weighted pathogenic evidence. Some considerations for researchers that generate mutational scanning data to ensure the best clinical application of this evidence will be discussed from the perspective of a variant curator.

The clinical utility of Multiplexed Assays of Variant Effects (MAVEs) in a pediatric cohort

Melissa Gilbert
The Children's Hospital of Philadelphia & University of Pennsylvania

Gilbert MA1,2*, Bowen BM3 , Hayek TJ1,2, Richardson ME3, Conlin L1,2, Rajagopalan R1,2, Smith ED3, Marsh A3, Murrell J1,2, Spinner NB1,2.

1Division of Genomic Diagnostics, Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA

2 Department of Pathology and Laboratory Medicine, The Perelman School of Medicine at The University of Pennsylvania, Philadelphia, PA, 19104, USA

3Ambry Genetics, Aliso Viejo, CA, 92656, USA

Advances in genome sequencing have outpaced our ability to interpret the functional consequences of many DNA variants identified during clinical diagnostic testing. Consequently, most variants catalogued in ClinVar, a database of DNA variants and their association with disease, are classified as Variants of Uncertain Significance (VUS), limiting their clinical utility. MAVEs, which are high-throughput functional assays that characterize libraries of variants, are revolutionizing our capacity to study variant impact. When properly calibrated, data generated from MAVEs can be directly translated into standardized evidence strengths (very strong, strong, moderate, or supporting) for benign or pathogenic classification, enabling integration into clinical evaluation.

We evaluated the utility of MAVEs datasets for 12 disease-associated genes to reclassify VUSs identified in a pediatric cohort who underwent clinical diagnostic testing at the Children's Hospital of Philadelphia. MAVEs data was obtained from published sources and calibrated to derive evidence weights. Across the genes, 787 variants were detected, of which 276 (35.1%) had available MAVEs data. Preliminary comparison of abnormal vs. normal functional outcomes for 145 variants previously classified as (likely) pathogenic or (likely) benign achieved 80.8% sensitivity (pathogenic variants with abnormal function) and 98.1% specificity (benign variants with normal function). The remaining 131 variants were classified as VUSs. Preliminary analysis demonstrated that 105 (78.9%) of these VUSs had MAVEs data consistent with normal function, supporting benignity, while 16 (12%) demonstrated abnormal function, supporting pathogenicity. Applying calibrated MAVEs functional evidence to these variants has the power to drive systemic reclassification, improving genomic diagnostics and advancing precision medicine.

Functional Assessment of Genetic Variants of CARD11 and BCL10 with Saturation Genome Editing

Richard James
Seattle Children's Research Institute

James RG 1,2*, Camp ND 1, Dawson S 1, Moore N 3, Smith N 2,4, Rubin A 3, Starita LM 2,4

1- Seattle Children's Research Institute

2- University of Washington School of Medicine

3- Walter and Eliza Hall Institute of Medical Research

4- Brotman Baty Institute

The effective use of genomic medicine is often hindered by variants of uncertain significance - genetic variants in disease-associated genes whose impact on disease development remains unknown. The CARD11-BCL10-MALT1 signalosome plays a critical role in B and T cell function, with each gene implicated in various immune-related diseases. Loss-of-function in these genes are linked to immunodeficiency and, in some cases, primary immune regulatory disorders, while gain-of-function is associated with lymphoma. In this study, we used cloning-free saturation genome editing to introduce all possible single nucleotide variants of the coiled-coil and guanylate kinase domains of CARD11, and the CARD domain of BCL10 into human primary T cells and a B cell line. Using a proliferation based assay we were able to functionally score 4,791 genomic variants spanning 502 amino acids of the coiled-coil domain of CARD11 and 2,178 variants covering 234 amino acids of the coding region of BCL10. We were able to show that loss-of-function variants that clinically present as primary immune regulatory disorders, such as those in CADINS Disease (C223T and R235P) or atopic dermatitis (L194P), were scored as loss-of-function in the T-cell assay. Remarkably, gain-of-function variants associated with T-cell lymphoma induced proliferation in both T and B cells, whereas several variants linked to B-cell lymphoma showed little or no phenotype in T cells but were robustly gain-of-function in B cells under BCR-signaling inhibition. These findings reveal previously unrecognized cell-type-specific functional consequences of signalosome variants. Collectively, our study demonstrates that saturation genome editing in primary cells provides a scalable framework for VUS classification and exposes lineage-dependent mechanisms controlling signaling through CARD11 and BCL10.

MaveMD: A functional data resource for genomic medicine

Ben Capodanno

Lead Software Engineer, MaveDB Development Team, Brotman Baty Institute.

Jeremy Stone 1,*, Benjamin J. Capodanno 1, Estelle Y. Da 2, Sally B. Grindstaff 1, David Reinhart 1

1 - Brotman Baty Institute for Precision Medicine, University of Washington, Seattle, WA, USA

2 - Bioinformatics Division, Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia

Variant interpretation remains one of the most significant challenges in clinical genetics. Variants of uncertain significance (VUS) undermine precision medicine implementation because they have an unknown relationship to disease and cannot be used for clinical decision-making. While evidence from multiplexed assays of variant effect (MAVEs) and other functional assays can help classify variants, major barriers prevent routine use in clinical variant classification, including fragmentation across multiple repositories, insufficient data standards, and the need to calibrate assays clinically. Here we address these challenges by presenting a new interface for the MaveDB database called MaveMD (MAVEs for MeDicine) that integrates with external resources such as ClinVar and the ClinGen Allele Registry, displays clinical evidence calibrations, provides intuitive visualizations, and exports structured evidence compatible with ACMG/AMP variant classification guidelines. MaveMD implements automatic mapping of MaveDB datasets to the human reference genome, dramatically simplifying the clinical translation of new MAVE data. We also defined a new metadata model after curating 438,318 variant effect measurements from 74 MAVE datasets spanning 32 disease-associated genes, and created an interface aimed at enabling effective clinical decision-making. Thus, MaveMD makes MAVE data accessible and easily usable for variant classification, and will scale seamlessly with future data generation efforts to empower the use of MAVE evidence in clinical practice.

RNA variants resolved in non-expressed genes by transactivation

Emmylou C. Nicolas-Martinez
Adelaide University

Nicolas-Martinez EC 1,2,*, Pflueger C 3,4, Zhang J 1,2, Gardner A 1,2, Corbett MA 1,2, John T 1,2 , PERSYST Investigator Team 5, Lister R 3,4, Gecz J 1,2, Jolly LA 1,2

1 - Adelaide University, Adelaide, SA, 5005, Australia

2 - The Robinson Research Institute, University of Adelaide, Adelaide, SA, 5005, Australia

3 - Harry Perkins Institute of Medical Research, Nedlands, WA, 6009, Australia

4 - The University of Western Australia, Crawley, WA, 6009, Australia

5 - National Collaboration, Australia

Understanding the impact of suspected disease-causing genetic variants on RNA splicing is crucial for both diagnosis and precision medicine opportunities. This is enabled through RNA transcriptomic and targeted functional studies using RNA isolated from clinically accessible tissues such as the patients' blood or skin. Insufficient disease gene expression in these tissues does however pose a major barrier to RNA-based investigations, which we show is relevant to 1,436 Mendelian disease genes, herein termed 'silent Mendelian genes' (SMGs). We pursued CRISPR-activation-based gene transactivation in patient-derived skin fibroblasts and blood-derived lymphoblastoid cells to overcome this limitation. An initial single-cell Perturb-seq transactivation screen of 40 SMGs in fibroblasts highlighted broad utility (25/40 SMGs transactivated in >100 cells), and an underlying relationship between the level of gRNA expression and transactivation. This stimulated our development of a four-plex gRNA transactivation system culminating in the 6- to 90,000-fold induction of expression of 61/61 (100%) SMGs tested individually in fibroblast and/or lymphoblastoid thus far. RNA-sequencing revealed that the magnitude (1-3000 transcripts per million) and isoform diversity of transactivated SMGs was comparable to the expression of SMGs in their clinically relevant tissues. We applied SMG transactivation combined with short-read and long-read nanopore RNA-sequencing to resolve the splicing impact of 37 variants in SMGs, (e.g. COL2A1, DMD, MED12L, SCN1A, USH2A) using patient-derived fibroblasts and lymphoblastoids, resulting in new diagnosis, disease mechanisms and precision medicine opportunities. Transactivation represents a potential functional genomic solution to investigate the >100K variants of uncertain significance in SMGs captured in ClinVar.

From Functional Assays to Computable Knowledge: Making MAVE Data Usable in Variant Interpretation

Alex Wagner

Nationwide Children's Hospital

The rapid expansion of human genomic sequencing has created a growing gap between the number of observed variants and the evidence available to interpret their functional and clinical significance. Experimental functional assays, including multiplexed assays of variant effect (MAVEs), offer a scalable way to measure the impact of thousands of variants simultaneously. However, results from these experiments are often difficult to integrate into genomic interpretation workflows. Many assays report variants relative to experimental constructs rather than reference genomes, which limits their accessibility for downstream computational analysis.

This presentation describes recent efforts to make MAVE data more accessible and interoperable for human genomics applications. In collaboration with the Atlas of Variant Effects Alliance, methods have been developed to map millions of variants reported in MaveDB to human reference sequences while preserving assay provenance. These mappings allow functional assay results to be represented in machine-readable formats and linked to variants observed in sequencing datasets.

Emerging standards from the Global Alliance for Genomics and Health (GA4GH) Genomic Knowledge Standards workstream, including the Variation Representation Specification (VRS) and related knowledge representation frameworks, provide a foundation for representing and exchanging this information in a consistent and computable form. Using these standards, functional evidence from MAVEs can be linked directly to variants observed in human sequencing data and incorporated into variant interpretation systems. MAVE data contribute functional evidence within a broader federated ecosystem of genomic knowledge resources alongside population and clinical evidence from resources such as gnomAD and ClinVar.

Findings from the Critical Assessment of Genome Interpretation (CAGI), seventh edition: a community experiment to evaluate phenotype prediction

Steven E. Brenner
University of California, Berkeley

Critical Assessment of Genome Interpretation (CAGI) 7 ARSA missense stability prediction challenge identifies computational advances over state-of-the-art variant impact predictors

Lin YJ 1,*, Gelb MH 2, Brenner SE 1

1 - University of California, Berkeley

2 - University of Washington

Metachromatic leukodystrophy (MLD) is an autosomal recessive lysosomal storage disorder caused by deficiency of the lysosomal enzyme arylsulfatase A (ARSA). Early identification of severe ARSA deficiency is critical to newborn treatment decisions.

The CAGI 7 ARSA challenge asked participants to predict the impact of nearly all ARSA missense variants. These were assessed against an unpublished deep mutational scanning dataset measuring ARSA protein stability. We employed several assessment metrics. For example, to relate stability to clinical severity, we assessed prediction against a threshold informed by enzyme stability in patients with late-infantile (most severe) MLD. Across 22 teams (86 models), several submissions (APE, 3BIOBXL, PRESCOTT, CLINTRABI, and Evocoders) outperformed widely used state-of-the-art predictors.

Because missense changes from single-nucleotide variants (SNVs) arise more frequently than multi-nucleotide variants (MNVs), we evaluated methods separately on each. Predictions for SNV-missense variants showed stronger agreement with the experimental measurements, indicating their impact is easier to predict. Because many methods predict pathogenicity rather than stability, we also performed an analysis intended to remove variants known to cause pathogenicity via mechanisms other than stability. After excluding these positions, predictions showed improved agreement with the experimental stability measurements. This highlights the importance of mechanism-aware assessment. Together, the CAGI 7 ARSA challenge assessment reveals advances in variant impact prediction, while highlighting needs and future directions for the field.

Clinical translation of a MUTYH MAVE into a diagnostic laboratory workflow enables large-scale VUS resolution

Ashley P.L. Marsh

Ambry Genetics

Ashley P.L. Marsh 1,*, Shelby L. Hemker 2, Kristin Kohler 1, Daniel Zimmerman 1, Felicia P. Hernandez 1, Jacob O. Kitzman 2,3 and Marcy E. Richardson 1

1. Ambry Genetics, Aliso Viejo, CA 92656, USA

2. Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109, USA

3. Gilbert S. Omenn Department of Computational Medicine & Bioinformatics, University of Michigan Medical School, Ann Arbor, MI 48109, USA

Variant-to-function maps (or MAVEs) offer opportunities for resolving variants of uncertain significance (VUS). However, VUS resolution in medically relevant genes requires calibration and translation into diagnostic workflows to generate clinically actionable results. The recent publication of a MUTYH MAVE offered opportunity for large-scale VUS resolution, which constitute ~90% of missense variants in ClinVar. Biallelic, loss of function variants cause MUTYH-associated polyposis (MAP), an autosomal recessive (AR) condition predisposing to colorectal polyps and cancer.

Population-based estimates from repositories like UK Biobank serve as resources for associating clinical risk with functional categorization. However, for AR conditions like MAP, these repositories may yield insufficient data due to rarity of biallelic pathogenic (LP/P) genotypes and/or phasing requirements. To circumvent this, records were obtained from a diagnostic laboratory. The burden of MAP-associated phenotype for >200 individuals with an assumed or confirmed pathogenic genotype was compared to 155 individuals with one MUTYH pathogenic variant and one functionally abnormal or neutral VUS, revealing that functionally abnormal missense VUS confer similar risk to MAP as pathogenic genotypes and its contrapositive.

Given this association, the OddsPath was calculated, resulting in strong odds of pathogenicity for abnormal (>18.7) and neutral (<0.053) variants. Application of this evidence within a diagnostic laboratory, using an ACMG/AMP framework, resulted in reclassification of 80 VUS to LP/P and 451 VUS to LB/B (~69% of missense VUS), impacting 4,664 individuals. In summary, this study provides insight into the implementation of MAVE data into a diagnostic workflow, leading to large-scale VUS resolution and more accurate clinical care.

Landscapes of human AAT missense-variant effects reveal pathogenic variation and genetic suppressors

Frederick P. Roth
University of Pittsburgh

van Loggerenberg W 1, Gebbia M 2,3,4, Hraha T 5, Rasoulzadeh A 1, Fawcett KM 1, Axakova A 2,3,4, Chawla A 2,3,4, Kishore N 2,3,4, Weile J 2,3,4, Barrera L 5, Banerjee A 5, Packer M 5 and Roth FP 1,2,3,4,*

1 - Department of Computational and Systems Biology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA;

2 - Donnelly Centre, University of Toronto, Toronto, Ontario, Canada;

3 - Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada;

4 - Lunenfeld-Tanenbaum Research Institute, Sinai Health, Toronto, Ontario, Canada;

5 - Beam Therapeutics Inc. Cambridge, Massachusetts, USA

SERPINA1, which encodes the serine protease inhibitor Alpha-1 Antitrypsin (A1AT), harbors variants causing A1AT Deficiency (A1ATD). Associated with reduced A1AT serum levels, A1ATD causes predisposition for inflammatory lung disease. For a subset of variants causing A1AT aggregation in the endoplasmic reticulum (ER), A1ATD also causes liver disease. Despite being a ‘common rare’ disease (affecting ~1 in 2500 individuals in Europe), A1ATD is under- and often mis-diagnosed. Although sequence-based diagnosis can be definitive, one-third of missense variants in A1AT are classified as ""Variants of Uncertain Significance"" (VUS). To address the VUS problem and to reveal sequence-structure-function relationships, we implemented and validated a scalable assay for A1AT variant function, and measured both diminished abundance and ER aggregation for >88% of all possible A1AT amino acid substitutions, thus providing a comprehensive variant effect (VE) map. Given that the majority of individuals with A1ATD express the ER-aggregation-prone Z-allele (p.Glu366Lys), and the potential for direct genome editing of this locus in patients, we also systematically identified variants which might enhance or suppress the Z-allele aggregation phenotype. This new atlas of A1AT VE maps both aligns with biochemical expectations and points to new sequence-structure relationships. Importantly, it provides evidence to discriminate pathogenic from benign variants that is comprehensive and proactive, i.e., providing evidence even for yet-to-be-observed clinical missense variants. Moreover, the atlas reveals a genetic suppressor of aggregation-associated A1AT variants, identifying a potential therapeutic target to treat A1ATD-associated liver disease.

Thursday March 26th 2026

Keynote Speaker

Scaling MAVEs: are we at the end of the beginning?

Matthew Hurles
Wellcome Sanger Institute

BIO:

Matthew Hurles is Director of the Wellcome Sanger Institute and leads a research group focused on deciphering the genetic causes of severe developmental disorders, and understanding how DNA mutates as it is passed from generation to generation.

Integrating MAVE and patch clamp data to improve classification and risk prediction in KCNH2-LQTS

Chai-Ann Ng
Victor Chang Cardiac Research Institute

BIO:

Dr. Chai-Ann Ng completed his PhD at La Trobe University in 2009. His postdoctoral work was undertaken at the Cardiac Electrophysiology Laboratory, Victor Chang Cardiac Research Institute, earning him the Young Biophysicist of the Year Award in 2013. He is now a Senior Staff Scientist at the Institute and a conjoint Senior Lecturer at UNSW. Dr Ng is an expert member on two ClinGen Variant Curation panels (Cardiac Potassium and Sodium/Calcium channels), working on global guidelines for variant interpretation. His research addresses the challenge of variants of uncertain significance in inherited arrhythmia syndromes using Automated Patch Clamp electrophysiology. Dr. Ng has translated research into clinical practice by collaborating with key diagnostic labs nationally and internationally to improve variant classification and diagnoses. His leadership in translating research to improve health outcomes was recognised by his selection as a Finalist for the 2024 Australian Cardiovascular Alliance Game Changer Award.

Three Multimodal Assays of SCN5A Variant Function Inform Arrhythmia Risk Prediction

Andrew Glazer
Vanderbilt University Medical Center

Andrew M. Glazer,1* Matthew J. O'Neill1, Joanne G. Ma2, Jeremy E. Smith1, Joseph F. Solus1, Jamie I. Vandenberg2, Chai-An Ng2, Dan M. Roden1

1 Vanderbilt University Medical Center, Nashville USA

2 Victor Chang Cardiac Research Institute, Darlinghurst Australia

SCN5A encodes the cardiac sodium channel NaV1.5 and is a key gene linked to inherited arrhythmia syndromes, including Brugada Syndrome (BrS), which carries a heightened risk of sudden cardiac death. We describe three high-throughput assays that provide functional insights into SCN5A variants and refine arrhythmia risk prediction.

First, a robotic automated patch clamp (APC) assay was calibrated using 49 benign and pathogenic controls and applied to 252 SCN5A variants from a 3,335-patient BrS cohort. APC stratified variants into normal function, partial loss-of-function, or severe loss-of-function categories, enabling the reclassification of 110 VUS, including 104 as likely pathogenic.

Second, a MAVE assessed 16,602 SCN5A variants, identifying 6,292 with abnormal function, including a network of residues critical for channel inactivation. MAVE scores correlated strongly with APC measurements (Pearson rho=0.76). APC and MAVE loss-of-function variants were enriched in transmembrane domains near the channel pore and correlated strongly with BrS risk (8–25% penetrance in case-control comparisons and biobank analyses).

Third, a multiplexed minigene splicing assay evaluated splicing effects for 224 clinically observed SCN5A variants in induced pluripotent stem cell-derived cardiomyocytes. This assay identified 78 variants that disrupt splicing, including many outside canonical splice sites. Splice disruption measurements correlated well, but not perfectly, with computational predictors (Pearson rho=0.61-0.81).

Together, these multimodal assays provide robust functional evidence for SCN5A variant interpretation, advancing precision medicine and arrhythmia risk stratification. To support clinical adoption, a Sodium and Calcium Channel Arrhythmia Variant Curation Expert Panel is developing SCN5A-specific guidelines for variant classification.

Single-codon resolution mapping of nonsense-mediated mRNA decay via genomic stop codon scanning of LMNA

Sujatha Jagannathan
University of Colorado Anschutz Medical Campus

Cortázar MA 1, Muñoz O 1,2, Campbell AE 1, Coban-Akdemir Z 3, Jagannathan S 1,4,*

1 - Dept. Biochemistry and Molecular Genetics, University of Colorado Anschutz Medical Campus, Aurora CO, USA.

2 - Molecular Biology Graduate Program, University of Colorado Anschutz Medical Campus, Aurora CO, USA.

2 - Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA

3 - RNA Bioscience Initiative, University of Colorado Anschutz Medical Campus, Aurora CO, USA.

Premature termination codons (PTCs) cause a large proportion of inherited genetic diseases. The mRNA surveillance pathway nonsense-mediated mRNA decay (NMD) mitigates the impact of these mutations by detecting and degrading PTC-containing transcripts that could otherwise produce potentially toxic truncated proteins. While variation in NMD efficiency is known to modulate clinical presentation, both the extent and the major determinants of NMD variability on a target mRNA remain elusive. Here, we performed systematic CRISPR/Cas9 genome editing of the clinically relevant LMNA A/C gene (LMNA) to quantify the NMD response at single-codon resolution on the endogenous mRNA in human cells. We integrated all three stop codons (TAA, TAG, TGA) at every codon position across different exon regions and performed targeted sequencing to study 3,321 PTC mRNA variants. By quantifying the transcript abundance with and without NMD inhibition via a small molecule, we calculate the precise NMD activity induced by each PTC across the transcript. We find that the position of the PTC within the open reading frame is the major determinant of NMD variability, followed by the presence of downstream readthrough-promoting sequence contexts and the identity of the PTC at specific exon positions. Our results have direct implications four our understanding of the mechanism of NMD and for the clinical interpretation of nonsense variants in the human population.

260,000 variant effect measurements and novel calibration methods drive resolution of VUS

Lea Starita
University of Washington

Lea Starita* and the IGVF coding variant focus group

Understanding the functional consequences of human genetic variation is essential for realizing the promise of genomics. Functional assays provide fundamental insights into gene and protein biology, reveals the molecular mechanisms by which variants perturb cellular systems, and supplies critical evidence for clinical variant interpretation. However, most variants remain uncharacterized, existing data are fragmented across disparate formats, and translation into clinically actionable evidence has been limited. As part of the NHGRI Impact of Genomic Variation on Function (IGVF) consortium, we deployed production-scale MAVEs to measure the effect of over 60,000 variants across 9 genes on protein abundance and cellular fitness. Moreover, we curated ~200,000 community-generation variant effect measurements in another 30 genes. Using these data, we illuminate variant-driven mechanisms of pathogenicity and provide systematic evidence to support clinical variant classification. We developed new methods for calibrating experimental and predictive data for clinical use that enable us to provide sufficient evidence to reclassify ~76% of ~16,000 ClinVar variants of uncertain significance across 40 genes, with an error rate below 2% on control variants. We explored the relationship between our functional and predictive data and human phenotypes in the All of Us biobank, determining associations between our calibrated data and relevant clinical phenotypes, further validating our approach. Extending this framework, we “pre-classified” more than 100,000 previously unseen variants as likely pathogenic or likely benign, and exposed all calibrated results through MaveMD, a clinician-oriented MaveDB interface. Together, our work establishes a scalable framework for mapping variant function, deepens our understanding of molecular pathomechanisms, and demonstrates a realistic path to eliminating most variants of uncertain significance in genomic medicine.

Modeling the impact of personal genome variation on molecular phenotypes

Nilah Ioannidis
University of California, Santa Cruz

Title: Modeling the impact of personal genome variation on molecular phenotypes

University of California, Berkeley

Abstract: Understanding inter-individual variation in molecular, cellular, and other clinically-relevant phenotypes is an important challenge in precision medicine. Sequence-based genomic deep learning models that predict gene expression and other molecular phenotypes directly from DNA sequence can be applied in silico to sequences containing any combination of rare or common genetic variants, with great potential to predict the genetic contribution to variation in such phenotypes. However, despite success in explaining variation in molecular phenotypes across the genome and across a variety of cell types, we and others recently found that current sequence-based genomic deep learning models have limited ability to explain variation in gene expression across different individuals based on their personal genome sequences. I will discuss our work to characterize the cross-individual performance of such models on gene expression and other molecular phenotypes, with resulting insights into their understanding of regulatory variation. I will also discuss our recent efforts to develop models with improved understanding of variation across individuals using several strategies, such as incorporating personal genome and transcriptome data during model training and using a hierarchical approach to first model more locally-regulated phenotypes such as chromatin accessibility.

tfGPRA: A High Throughput Platform for Transcription Factor Characterization

Omar Tariq
University of British Columbia

Omar Tariq, 1*, William Cheney, 1, Carl de Boer, 1

1 - School of Biomedical Engineering, University of British Columbia

Transcription factors (TFs) are the central regulators of gene expression. TFs recognize and bind short sequence motifs to regulate the expression of nearby target genes. Genome wide sequencing data shows substantial variation in TFs across the human population. The phenotypic impact of these variants remains unclear. Over the years, several methods have been developed for characterizing TF binding genome-wide (ex. ChIP-seq) or against a large pool of synthetic sequences (ex. HT-SELEX). However, these methods are sub-optimal for profiling many TF variants in a single experiment.

We have developed the TF-GPRA (Transcription Factor Gigantically Parallel Reporter Assay) to learn binding for many TF proteins in parallel. The TF-GPRA involves exogenously expressing eukaryotic TFs in yeast, along with a library of randomized promoter sequences controlling the expression of a fluorescent reporter. By sorting cells according to reporter expression, we can develop a tripartite TF-Sequence-Expression dataset for training machine learning models. We show these models can learn motifs for both yeast and exogenously expressed human TFs. This approach could be applied to studying TF variants, isoforms or TFs from a poorly studied genome.

Saturation Genome Editing to Clarify Variant Effect within 5’UTRs of Neurodevelopmental Disorder Genes

Vanessa Burns
Wellcome Sanger Institute

Vanessa Burns[1,2*], Sebastian Gerety[1], Elizabeth Radford[1,2], Matthew Hurles[1]

1 - Wellcome Sanger Institute

2 - University of Cambridge

The 5′ untranslated regions (5′UTRs) of mRNAs play a critical role in ribosome recruitment and contain functional sequences that regulate gene expression through post-transcriptional mechanisms. While some individual 5′UTR motifs are known, how variants affect these sequences and their overall function remains poorly understood. Interpreting 5′UTR variants is especially challenging in neurodevelopmental disorder (NDD) genes, where clinical classification is difficult. To address this gap, we conducted a saturation genome editing (SGE) screen targeting 5′UTRs of selected NDD genes, leveraging multiplexed assays of variant effect (MAVEs) for large-scale functional assessment. We first performed a pilot SGE screen to evaluate ~4,000 unique 5′UTR variants across four NDD-associated genes essential in the Hap1 cell line. This variant set spans all SNVs, indels of varying lengths, insertions of functional 5′UTR elements, and clinically or population-reported variants. Preliminary data reveals distinct impact profiles across variant classes, highlighting individually disruptive variants and broader trends in 5′UTR sensitivity and resilience. Notably, variants creating strong upstream starts or disrupting the endogenous Kozak sequence show increased sensitivity, while most SNVs and deletions are well tolerated. These findings inform the design of a larger-scale 5′UTR SGE study, anticipated to clarify the contribution of 5′UTR variation to NDDs and have immediate large-scale diagnostic utility.

Genotype-phenotype single-cell transcriptomics for massive parallel assessment of genetic variants

Lorenzo Vaccaro

Telethon Institute of Genetics and Medicine

*Lorenzo Vaccaro1,2,$, Rosa De Santis1,$, Francesco Panariello1, Antonio Grimaldi1,2, Gennaro Gandolfo1, Mirella Massa1,2, Tonya Fusco1,2, Marcello Salvi1,3, Sara Riccardo3, Patrizia Annunziata3, Chiara Colantuono3, Giuseppe Martone3, Martina Bamundo4, Gennaro Gambardella5, Riccardo Arnese5,6, Emanuele Agolini7, Claudia Cesario7, Ibrahim M.Abdelrazek8, Hadeer AbdelGhany9, Antonio Novelli7, Beatrice Salvatori3, Marcella Cesana10, Christoph Ziegenhain11, Caterina Missero6,12, Davide Cacchiarelli1,2,13

1Telethon Institute of Genetics and Medicine (TIGEM), Armenise/Harvard Laboratory of Integrative Genomics, Pozzuoli, Italy
2Department of Translational Medicine, University of Naples "Federico II", Naples, Italy
3NEGEDIA srl, Pozzuoli, Naples, Italy
4CEINGE Biotecnologie Avanzate Franco Salvatore, Naples, Italy
5Telethon Institute of Genetics and Medicine (TIGEM)
6Department of Chemical, Materials and Industrial Engineering, University of Naples Federico II, Naples, Italy
7Laboratory of Medical Genetics, Translational Cytogenomics Research Unit, Bambino Gesù Children's Hospital, IRCCS, Rome, Italy
8Department of Human Genetics, Medical Research Institute, Alexandria University, Alexandria, Egypt
9Department of Dermatology, Venereology and Andrology, Faculty of Medicine, Alexandria University, Alexandria, Egypt
10Department of Advanced Biomedical Sciences, University of Naples "Federico II", Naples, Italy
11Division of Medical Systems Bioengineering, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden
12Department of Biology, University of Naples Federico II, Naples, Italy
13Genomics and Experimental Medicine Program, Scuola Superiore Meridionale (SSM, School of Advanced Studies), Naples, Italy
$Co-first author

Predicting the pathogenic effect of rare genetic variants remains a significant challenge, hindered by the limited availability of diagnosed patient cohorts and the difficulty in systematically evaluating the impact of both existing and novel variants across diverse functional contexts. To address this issue, we first leveraged a large-scale protein Multiplexed Assay of Variants Effect (MAVE) to functionally assess ~2,300 missense variants of the TP63 gene, which are strongly associated with autosomal dominant developmental disorders. Variant activity was precisely measured using an optimized fibroblast-to-keratinocyte conversion protocol, providing an in-vivo-like functional readout. This comprehensive MAVE effort yielded a clinically actionable dataset, reclassifying hundreds of common variants and definitively resolving a Variant of Unknown Significance as pathogenic in a patient with ectodermal dysplasia. However, classical MAVE methodologies are fundamentally constrained by the requirement for specific, predefined phenotypic assays. To overcome this bottleneck and enable systematic, future-proof variant screening, we developed SCRAMseq (Single Cell RNAseq Associated with MAVEs by sequencing). This generalizable platform leverages full-length scRNAseq for high-throughput, single-cell genotype-to-phenotype mapping, utilizing the entire transcriptome as the functional readout. Extensive validation across diverse biological contexts, including different genes and distinct disease-driving pathways, confirms SCRAMseq's superior reproducibility and accuracy (also compared with bulk assay based on cell conversion). SCRAMseq is a robust and user-friendly workflow set to accelerate the discovery and dissection of gene variant effects on any potential disease-driving pathway, enabling the application of MAVEs to any possible disease driving gene.

SPONSOR TALK

MAVE Progress Report, How Ambry Genetics Validates and Deploys High-Throughput Functional Assays

Daniel Zimmerman
Ambry Genetics

Daniel Zimmerman 1*, Sarah Dugger 1, Ashley Marsh 1, Dean Hoffer 1, Ana Silverstein 1, John Ranola 1, Heather Zimmermann 1, Sean Korpela 1, Marcy Richardson 1, Katie Yergert 1, Rhonda Lassiter 1, and Tina Pesaran 1

1 - Ambry Genetics, Aliso Viejo, California

As the research community becomes increasingly proficient at high-throughput experiments, clinical laboratories need to adapt their methodologies to accommodate large-scale variant classification. Analyzing variants on a case-by-case basis will inevitably fall behind the rapid pace of research, creating an ever-widening gap between the release of new MAVEs and implementation into clinical settings. At Ambry Genetics, we are developing automated workflows that identify newly released studies, validate the quality of data, and calibrate results to ACMG/AMP evidence strengths.

Ambry collaborates directly with the research community to facilitate variant reclassification by providing clinical data and expertise. New MAVEs are identified via ongoing partnerships with academic institutions, or through AutoLit, an automated search software, and prioritized according to their clinical impact. Data from ClinVar, Ambry, and gnomAD is pulled and filtered to generate a high-quality reference set of pathogenic and benign variants. Logistic regression and local likelihood ratio estimation models are then run to assign variant-specific strengths of evidence. To visualize MAVE data, we developed GeneGlance, an online webserver intended to enable large-scale reclassification, and MAVEtracker to coordinate reclassification efforts. As of November 2025, we have processed 7 MAVEs for genes on our CancerNext panel, enabling personalized care for 18,300 patients via reclassification of 2,161 VUS.

Phenotypic Mutations in Collision: Negative Epistasis Between Transcription Errors and Translation Errors Revealed by DMS

Jian-Rong Yang
Sun Yat-sen University, Guangzhou

Phenotypic mutations consist of non-heritable changes in biomolecular sequences arising from transcription and translation errors. Their significance in evolution and human health is becoming increasingly evident, for instance in cellular aging and the emergence of novel traits in cancer. The interaction between the two types of errors, however, is poorly understood. Using data from humans and four other model eukaryotes, we conducted genome-wide analyses of nucleotide misincorporation during transcription (mistranscription) and amino acid misincorporation during translation (mistranslation). It was found that genes with frequent mistranslations usually have a lower mistranscription rate, a pattern we hypothesized to be due to negative epistasis between the two types of errors. We tested our hypothesis through systematic DMS experiments on within-gene epistasis between mutations. Such epistasis is predominately negative, suggesting that proteins affected simultaneously by both types of error will be significantly more deleterious than expected. More importantly, an in silico simulation of molecular evolution suggests that the extra deleterious effects caused by negative epistasis, when scaled by the error rates, can facilitate selection against mistranscription for genes with frequent mistranslations. Finally, our hypothesis is further supported by the observation that genes with frequent mistranslation purge nonsynonymous mistranscription more efficiently, and transcripts that are more extensively translated tend to have a lower mistranscription rate. Overall, our results indicate that mistranslation suppresses mistranscription in eukaryotes.

From Variant Maps to Functional Design: Using Deep Mutational Scanning and Machine Learning to Engineer Programmable Bacteriophages

Srivatsan Raman
University of Wisconsin-Madison

Srivatsan Raman

Deep mutational scanning (DMS) has enabled high-resolution maps of how protein sequence changes affect function, but using this information to design proteins with new properties remains an open challenge. Here, we present an integrated approach that combines DMS, machine learning, and motif discovery to engineer bacteriophage proteins with programmable host specificity.

We focused on a receptor-binding protein from a model phage and built a comprehensive library of single and combinatorial mutations. By measuring how each variant affected infectivity across multiple bacterial hosts, we created a detailed sequence–function landscape. Using these data, we trained predictive machine learning models that not only forecast functional outcomes for unseen variants, but also helped reveal key sequence motifs and structural features that govern host recognition.

We then used the models to design new protein variants predicted to rewire or refine phage–host interactions. Experimental testing confirmed that many of these designs performed as intended — either expanding, shifting, or restricting host range in predictable ways.

This work demonstrates how combining variant effect maps with computational modeling can unlock new capabilities in phage engineering. More broadly, it offers a general framework for moving from mutational data to functional design across diverse protein systems.

The genetic architecture of the human bZIP interaction network

Guillaume Diss
Friedrich Miescher Institute

Bendel AM 1,2, Eichenberger B 1,2, Kempf G 1, Klein D 1, Shimada K 1, Jaaks-Kraatz S 1, Durdu S 1, Roth G 1, Soneson C 1,3, Schuebeler D 1,2, Stadler M 1,3, Zenke F 1, Diss G 1*

1 - Friedrich Miescher Institute for Biomedical Research (FMI), Basel, Switzerland

2 - University of Basel, Basel, Switzerland

3 - SIB Swiss Institute of Bioinformatics, Basel, Switzerland

Generative biology holds the promise to transform our ability to design and understand living systems by creating novel proteins, pathways, and organisms with tailored functions that address challenges in medicine, sustainability, and technology. However, training generative models requires large quantities of data that captures the genetic architecture of protein function in all its complexity, but these are currently scarce. Here, we systematically mutagenized all 54 human basic-leucine zipper (bZIP) domains and quantified their interactions with each other using bindingPCA, a quantitative deep mutational scanning assay. This resulted in ~2 million interaction measurements, capturing the effect of all single amino acid substitutions at each of the 35 interfacial positions. We found that mutation effects are largely additive in the vicinity of each wild-type bZIPs, but diverge across the family, indicating strong context dependency. A global additive thermodynamic model provided moderate prediction of mutation effects, while individual models per bZIP achieved higher performance, supporting a model of local simplicity and global complexity. Our results therefore suggest that the genetic architecture of protein function is more complex than previously anticipated, which could hinder predictability. However, a convolutional neural network trained on this dataset could accurately predict binding scores from sequence alone. Furthermore, the model enabled the design of synthetic bZIPs with high target specificity, demonstrating practical applicability for bioengineering purposes. Our study shows that capturing family-wide diversity is essential to reveal context dependencies and train accurate quantitative models of protein-protein interactions.

High-resolution functional assessment of SMAD4 variants

Jacob Purcell
Monash University

Purcell J 1*, Liu L 1, Rosenbluh J 1

1 - Biomedicine Discovery Institute, Monash University, Melbourne, Australia

SMAD4 is the central signalling protein of the TGFβ/activin and BMP pathways, and its dysfunction underlies a range of diseases. In the germline, loss-of-function SMAD4 mutations cause Juvenile Polyposis Syndrome and Hereditary Haemorrhagic Telangiectasia, while gain-of-function variants cause Myhre Syndrome. Moreover, somatic SMAD4 mutations are common in gastrointestinal cancers, where they are associated with poor prognoses. Despite this clinical relevance, over half of all SMAD4 variants recorded in ClinVar remain Variants of Uncertain Significance (VUS), including 90% of missense variants. This ambiguity constrains genetic counselling, diagnosis and precision therapies for patients with SMAD4 VUS mutations.

In this study we employed a Multiplexed Assays of Variant Effects (MAVE) screening approach to functionally characterise all possible coding SMAD4 variants. To do so, we developed a GFP-based TGFβ reporter assay in HT29 colorectal cancer cells. This cell line harbours biallelic SMAD4-inactivating mutations and therefore exhibits minimal TGFβ signalling unless complemented with functional SMAD4. Lentiviral SMAD4 variant libraries were delivered into this background, and cells were sorted into GFP-high and GFP-low populations by FACS. Next-generation sequencing of each bin quantified variant abundance, enabling computation of function scores reflecting each variant’s ability to restore signalling.

This data provides a high-resolution functional map of SMAD4 that robustly separates known pathogenic from benign variants and reveals both loss- and gain-of-function effects across the protein. Our dataset provides experimental evidence to support reclassification of many clinically ambiguous SMAD4 VUS.

From bulk to single cell: benchmarking analytical tools for Deep Mutational Scanning experiments

Rosa De Santis
Telethon Institute of Genetics and Medicine

De Santis R 1,2* , Vaccaro L1,2, Grimaldi A 1,2, Gandolfo G 1, Massa M 1,2, Fusco T 1,2, Ziegenhain C 3, Cacchiarelli D1,2

1Telethon Institute of Genetics and Medicine (TIGEM), Armenise/Harvard Laboratory of Integrative Genomics, Pozzuoli, Italy

2 Department of Translational Medicine, University of Naples “Federico II”, Naples, Italy

3 Division of Medical Systems Bioengineering, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden" "Proteins are the primary effectors of cellular function, and genetic variants can modify their structure, stability, or molecular interactions, leading to human disease. Deep mutational scanning (DMS) has become a valuable method for assessing the functional impact of thousands of variants simultaneously. Although numerous DMS studies and analytical tools exist, benchmarks mainly depend on simulated datasets, which complicates evaluating the performance of different strategies across various experimental setups.

In this study, we provide a comprehensive benchmark of DMS tools and analysis pipelines using our validated dataset on TP63, generated via Mutagenesis by Integrated Tiles (MITE) sequencing in both bulk and single-cell RNA-seq setups, combined with or without a functional assay. By utilizing these reference datasets, we examined various analytical approaches, including normalization, statistical modeling, and variant calling, to determine best practices for both bulk and single-cell MAVEs (Multiplexed Assay of Variant Effect) data.

This work represents the first benchmark of DMS methodologies in both bulk and single-cell contexts, highlighting methodological biases and offering parameters to guide the most effective strategies for high-throughput mutagenesis studies and the interpretation of functional variants.

Variant consequences in context

Amelie Stein
University of Copenhagen

BIO:

A/Prof Stein's lab studies the consequences of sequence variants on proteins, focusing on their cellular stability and function. The group performs high-throughput assays on protein variants and build on this data to develop and improve methods for prediction of variant consequences. They then apply these methods to determine whether genomic variants are likely to be pathogenic. Further they aim to integrate effects of multiple mutations for applications in protein engineering.

Elucidate the active conformation of cytokine receptors using deep mutational scanning

Xinyu Wu
Walter and Eliza Hall institute of Medical Research

Xinyu Wu 1,2,*, Margareta Go 1, Wessel Burger 1,2, Matthew Call 1,2, Melissa Call 1,2

1 - Walter and Eliza Hall Institute, Parkville, Australia

2 - Department of Medical Biology, University of Melbourne, Australia

Cytokine receptors convert extracellular cues into intracellular signals, and their dysfunction contributes to many diseases. Clinical mutations in TpoR and its kinase JAK2 are associated with disorders of thrombopoiesis; notably, many activating mutations cluster within the transmembrane domain (TMD), which also serves as the binding site for small-molecule TpoR agonists. Despite this, the structural mechanism of TpoR activation remains poorly understood.

Here, we present a pipeline combining AlphaFold predictions, molecular simulations and deep mutational scanning (DMS) to delineate TpoR conformations. AlphaFold predicts two distinct TMD dimeric arrangements: (1) a triangular configuration contacting at the N-terminus and splaying toward the C-terminus, stabilized by a juxtamembrane helix oriented laterally along the membrane; and (2) a parallel interface with continuous close contacts. DMS analysis revealed that ligand-bound TpoR exhibits a largely featureless TMD mutational landscape suggesting minimal TMD contacts are needed for Tpo-mediated signaling. This was confirmed by poly-valine domain-swapping experiments, showing specific contacts within the TMD of TpoR are not required for Tpo-mediated signalling In contrast, both TMD-targeting agonists, extracellular domain binding romiplostim, and the constitutively active S505N mutation produced specific mutational patterns, indicating defined active conformations.

The distinct TMD signatures elicited by native and non-native stimuli support the relevance of both AlphaFold models: the triangular interface is most consistent for Tpo-dependent signaling and the parallel interface exemplifies non-native agonist-driven activation. These findings reconcile disparate signaling models and illuminate structural principles underlying cytokine receptor activation.

Modeling Viral Evolution as Sequence Transitions: A Transformer Approach Using Time-Resolved SARS-CoV-2 Spike Data

Polina V Polunina

University of Freiburg

Polina V Polunina 1,* , Wolfgang Maier 1 , Alan F Rubin 2,3

1 - Bioinformatics Group, Department of Computer Science, University of Freiburg, Freiburg, Germany.

2 - Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia.

3 - Department of Medical Biology, The University of Melbourne, Parkville, VIC, Australia.

Predicting how viral proteins evolve over time is crucial for understanding variant emergence, immune escape, and functional shifts. Here we present a trajectory-aware transformer model that learns temporal amino-acid sequence transitions in the SARS-CoV-2 Spike receptor binding domain (RBD). Using ~1.9 million high-quality Spike protein sequences from NCBI/Nextclade, we construct ordered ancestor-descendant pairs based on collection dates and clade relationships. A sequence-to-sequence transformer was trained to map each RBD sequence at time t to its observed descendant at time t+Δt. The model achieves high accuracy on short-range Omicron transitions and higher uncertainty on large evolutionary jumps (e.g., non-Omicron to Omicron), consistent with known SARS-CoV-2 dynamics.

Analysis of the model's latent representations, obtained from the encoder embeddings of the ancestor sequences (at time t), shows clear organization by clade and time, suggesting that the model captures meaningful evolutionary structure directly from sequence data. While the current work focuses solely on primary amino-acid sequences, the framework is compatible with integrating additional information - such as structural features or pretrained protein embeddings (e.g., Evo2) - in future extensions. This approach provides a foundation for combining evolutionary trajectory modeling with Deep Mutational Scanning (DMS) datasets to link predicted sequence transitions with functional consequences.

Keynote Speaker

From ARS1 to ZWF1: some adventures in deep mutational scanning

Maitreya Dunham
University of Washington

Saturation mutagenesis coupled with functional assays and deep sequencing is a powerful combination for understanding the impacts of genetic variation, both positive and negative. In my lab, we have used these methods to survey the consequences of large numbers of mutations on diverse genetic sequences in yeast, including replication origins, promoters, and protein-coding sequences. These experiments have led to better understanding of what sequences are crucial for the normal functioning of genetic elements with implications for basic biology and disease risk. They have also let us identify mutations that behave even better than wildtype, leading to insight into adaptive evolution mechanisms and even to improved components for strain engineering. In this talk I'll give a tour of some highlights, from *ARS1* to *ZWF1.*

Characterising the functional landscape of the human PRC2 complex by base editing at single cell resolution

Mark Dawson
Peter MacCallum Cancer Centre

BIO:

Professor Dawson is a Physician-Scientist and an Associate Director of Research at the Peter MacCallum Cancer Centre. He graduated with a medical degree from the University of Melbourne in 1999, and subsequently trained as a haematologist in Melbourne, Australia. After his clinical training, he completed his PhD at the University of Cambridge, followed by a Wellcome Trust Fellowship studying the epigenetic regulation of normal and malignant haematopoiesis. He returned to Melbourne in 2014 to lead a team investigating chromatin regulation in haematopoiesis and cancer. His research has helped define key molecular mechanisms that underpin the initiation, maintenance, and progression of cancer. These insights have led to the development of several first-in-class epigenetic therapies that have been translated into various clinical trials across the world. In recognition of his research achievements, he has been awarded several awards and prizes including as a HHMI International Scholar, the William Dameshek prize from the American Society of Hematology, The Paul Marks Prize for Cancer Research and the Prime Ministers Prize for Science as Life Scientist of the Year 2020. He has been elected to the Australian Academy of Science, Australian Academy of Health and Medical Sciences and the European Molecular Biology Organisation (EMBO).

The genetic architecture of allosteric plant hormone receptors

Maximilian Stammnitz

Centre for Genomic Regulation

Maximilian R. Stammnitz[1,*], Ben Lehner[1,2,3,4]

1 - Centre for Genomic Regulation (CRG), Barcelona, Spain
2 - Wellcome Sanger Institute, Cambridge, UK
3 - Universitat Pompeu Fabra (UPF), Barcelona, Spain
4 - Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Allosteric plant hormone receptors govern critical responses like drought resistance – traits vital for climate resilience and therefore subject to major metabolic engineering campaigns. Yet, it remains elusive how very precise hormone tunability is encoded in the proteins' genetic architecture.

Here, we present a deep mutational scanning strategy to measure >100,000 interactions between receptor protein mutants and their small molecule ligands. Using an inducible protein complementation assay, GluePCA, we created the first complete variant effect maps of the hormone receptors PYL1 and GID1A in response to their natural plant metabolites abscisic acid and gibberellin.

Our data uncovers how mutations can fine-tune receptor properties – including baseline activity, chemical ligand sensitivity, and signal strength – and reveals protein stability as a hidden causal driver. A small fraction of missense mutants confers qualitative phenotypic innovation, producing constitutive, inverted and band-stop dose response patterns. Beyond an improved understanding of hormone receptor biophysics, our study provides a scalable framework to interrogate the function of thousands of engineered environmental sensors in a single deep mutational scanning experiment.

MAVE-informed Protein Language Models for Predicting Pharmacogenetic Variant Function at Scale

Jun J Yang
St. Jude Children's Research Hospital

Li Z 1, Yang W 1, Yang JJ 1

1 - Department of Pharmacy and Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN

Genetics-guided therapy requires accurate functional interpretation of pharmacogenetic variants. While functional consequences of common pharmacogenetic variants are well established (CPIC Level A/B), the majority of variants in these genes are classified as of unknown significance. Mutational scanning via Multiplexed Assays of Variant Effect (MAVE) provides large-scale experimental measurements, although can be challenging to implement uniformly across genes. Artificial intelligence/deep learning-based computational prediction of variant effect offer a scalable solution, and their applicability in pharmacogenetics needs to be assessed systematically.

We evaluated the correlation between computationally predicted scores and MAVE-derived abundance and/or activity scores for six pharmacogenes: NUDT15, TPMT, CYP2C9, CYP2C19, G6PD and VKORC1. Six computational methods were assessed, including conventional tools (SIFT, PolyPhen, CADD, and REVEL) and two deep learning models (ESM-2 and AlphaMissense). Overall, deep learning-based methods consistently demonstrated superior performance, with average Spearman's ρ=0.56 for both ESM-2 and AlphaMissense, compared to 0.47 for other tools.

Next, we developed a MAVE-informed ESM model: fine-tuning ESM-2 for each gene using 85% of the available MAVE-tested variants. As expected, the MAVE-informed models led to substantial improvement in variant effect prediction of the held-out variants (average Spearman's ρ=0.80), with greater gains seen in larger proteins. Even after down-sampling MAVE data to 30% of possible variants, fine-tuning ESM-2 still produced significant improvement.

Together, these findings demonstrate that protein language models can extend MAVE by accurately inferring functional effects for unmeasured variants, enabling more efficient designs of mutagenesis experiments.
Data-informed AI models offer a practical path toward comprehensive functional annotation of pharmacogenetic variation.

Biochemical profiling of ~315,000 MAP kinase pathway protein variants in human cells with LABEL-seq

Douglas M. Fowler
University of Washington

Jessica Simon1,*, Sriram Pendyala2,*, A. Felicia Adebanjo1, Allyssa Vandi2, Daniel Holmes2, Raining Wang2, Melinda Wheelock2, Jordan Opsahl2, Madeline Walsh1, Aman Shihora1, Lea M. Starita2,5, Dustin Maly1,3,#, Douglas Fowler2,4,5,#

1 Department of Chemistry, University of Washington, Seattle, WA USA

2 Department of Genome Sciences, University of Washington, Seattle, WA, USA

3 Department of Biochemistry, University of Washington, Seattle, WA, USA

4 Department of Bioengineering, University of Washington, Seattle, WA, USA

5 Brotman Baty Institute for Precision Medicine, Seattle, WA, USA

*denotes equal contribution

#correspondence to: djmaly@uw.edu, dfowler@uw.edu

Mitogen-activated protein kinase (MAPK) signaling underlies core decisions in cell growth, proliferation, and differentiation. Variants in pathway members, including ErbB, Ras, and Raf proteins, disrupt signaling and drive cancer and developmental disorders, yet mechanistic interpretation of these variants has lagged because multiplexed assays of variant effect have typically focus on single proteins. Here, we overcome this bottleneck, generating ~315,000 variant measurements across 15 proteins in the MAPK pathway by combining an optimized site-saturation mutagenesis pipeline with LABEL-seq, a multiplexed assay that measures variant-driven changes in protein activity, abundance, and interactions directly in human cells. Our pathway-scale dataset exposes trade-offs between abundance and signaling output, reveals dominant-negative mechanisms, and pinpoints functional sites of post-translational modification. Integration with structural modeling uncovers previously unrecognized interaction surfaces, including the K-Ras-LZTR1 interface that governs K-Ras stability. Finally, these multidimensional variant-effect maps distinguish pathogenic from benign variants and clarify cancer-associated mechanisms across the pathway. Thus, we establish a generalizable framework for measuring variant effects at pathway scale and provide the first comprehensive, mechanistic atlas describing how coding variation rewires MAPK signaling.

Deep mutational scanning to measure the transactivational effect of GATA2 and ERG variants to benefit clinical interpretation

Chris Hahn

SA Pathology and University of South Australia

Christopher N Hahn[1,2,3], Wen Teng[1,2], Jiarna R Zerella[1,2], Stuart T Webb[1,3], Claire C Homan[1,2,3], Parvathy Venugopal[1,2,3], Kerry Phillips[4], Nicola Poplawski[4], Devendra Hiwase[5,6], David M Ross[6], Anna L Brown[1,2,3], Catherine Carmichael[7,8], Alan F Rubin[9], Douglas M Fowler[10,11], Matthew Wakefield[12], Hamish S Scott[1,2,3].

[1]Centre for Cancer Biology, SA Pathology and Adelaide University, Adelaide, SA, Australia; [2]College of Health, Adelaide University, Adelaide, SA, Australia; [3]Department of Genetics and Molecular Pathology, SA Pathology, Adelaide, SA, Australia; [4]Adult Genetics Unit, Royal Adelaide Hospital, Adelaide, SA, Australia; [5]South Australian Health & Medical Research Institute, Adelaide, SA, Australia; [6]Royal Adelaide Hospital, Central Adelaide Local Health Network, Adelaide, SA, Australia; [7]Centre for Cancer Research, Hudson Institute of Medical Research, Clayton, VIC, Australia; [8]Department of Molecular and Translational Medicine, Monash University, Clayton, VIC, Australia; [9]Collaborative Centre for Genomic Cancer Medicine, University of Melbourne and Peter Mac, Melbourne, VIC, Australia; [10]Brotman Baty Institute for Precision Medicine, University of Washington, Seattle, WA, USA; [11]Department of Genome Sciences, University of Washington, Seattle, WA, USA; [12]Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia.

Abstract:

The Australian Familial Haematological Conditions Study (AFHCS) has recruited over 280 families from Australia and New Zealand with apparent predisposition to bone marrow failure (BMF) and/or haematological malignancy (HM) since 2004. Our comprehensive genetic and genomic analyses of samples from these families led to the discovery of GATA2 in 2011, after which numerous studies and diagnostic laboratories reported germline GATA2 variants associated not only with BMF/HM, but also other phenotypes. Subsequently, it was found that missense variants in zinc finger 2 (ZF2) were predominantly germline and those in ZF1 mainly somatic. While over 270 different GATA2 variants have been classified as Pathogenic/Likely pathogenic in the literature and ClinVar, over 850 remain as variants of uncertain significance (VUS) where patients and families remain without a clear diagnosis. More recently, we discovered ERG in 2024 leading to BMF/HM and potential cardiovascular defects. It is likely that ERG will follow a similar trajectory to GATA2 as it enters clinical screening panels.

Therefore, we embarked on developing Multiplexed assays of variant effects (MAVEs) for GATA2 and ERG using bespoke transactivation assays that we have developed over the past 16 years. While traditional transactivation assays have employed reporters such as Luciferase, we are adapting MAVEs to digitally count barcodes derived from the 3’UTR of a GFP cDNA driven by biologically appropriate promoters/enhancers in cell lines that faithfully mirror the pathogenicity of clinically relevant variants.

Standardisation and joint analysis of 2,041 MAVEs

Alistair Dunham

Wellcome Sanger Institute

Dunham A 1,2,*, Bourdim M 1, Fang M 1, Fong M 1, Gaunt O 1, Hidari E 1, Zane I 1, Parts L 1,2, Rubin A 3, Sokolov A 2,4, Tsukanov K 2,4, Zakirov A 2,4, Freeburg M 2,4

1 - Wellcome Sanger Institute

2 - OpenTargets

3 - Walter and Eliza Hall Institute of Medical Research

4 - European Bioinformatics Institute

Recent advances in high-throughput experiments have lead to a proliferation of perturbation studies, individually giving us insights into gene and variant properties. Combined analysis of such data has potential to improve the quality of conclusions through joint modelling, learn about more general biology and inform future experiments but is complicated by the wide array of experimental and analysis approaches in use. We are launching the Open Targets Perturbation Catalogue to address this problem, collating a combined catalogue of high-throughput perturbation studies including MAVEs, MPRAs, CRISPR KO and Perturb-Seq into a single annotated and searchable database. The MAVE portion of the data includes 1,188 coding MAVEs from MaveDB, including 2,041 experiments across 1,052 genes in 99 species, in addition to 35 MPRA studies. We have developed a hierarchical Bayesian mixture model to transform all experimental scores into a single standardised score space, with goodness of fit metrics, inferred deleteriousness cutoffs and known controls for quality control. Using our combined dataset we go on to investigate a range of properties of MAVE data, including the power to classify variants, genetic constraint and the relationship to biophysical properties. The combined dataset also allows us to explore the imputation of missing fitness scores at scale. Finally, this standardisation also allows us to generate a combined ML-ready MAVE dataset for benchmarking and training models. We demonstrate the utility of this dataset by training models predicting drug interactions from MAVE data, fitting a simple general protein sequence model and exploring mutational landscapes with a lab-in-the-loop framework.

Critical Assessment of Genome Interpretation (CAGI) 7 ARSA missense stability prediction challenge identifies computational advances over state-of-the-art variant impact predictors

Yu-Jen (Jennifer) Lin
University of California, Berkeley

Lin YJ 1,*, Gelb MH 2, Brenner SE 1

1 - University of California, Berkeley

2 - University of Washington" "Metachromatic leukodystrophy (MLD) is an autosomal recessive lysosomal storage disorder caused by deficiency of the lysosomal enzyme arylsulfatase A (ARSA). Early identification of severe ARSA deficiency is critical to newborn treatment decisions.

The CAGI 7 ARSA challenge asked participants to predict the impact of nearly all ARSA missense variants. These were assessed against an unpublished deep mutational scanning dataset measuring ARSA protein stability. We employed several assessment metrics. For example, to relate stability to clinical severity, we assessed prediction against a threshold informed by enzyme stability in patients with late-infantile (most severe) MLD. Across 22 teams (86 models), several submissions (APE, 3BIOBXL, PRESCOTT, CLINTRABI, and Evocoders) outperformed widely used state-of-the-art predictors.

Because missense changes from single-nucleotide variants (SNVs) arise more frequently than multi-nucleotide variants (MNVs), we evaluated methods separately on each. Predictions for SNV-missense variants showed stronger agreement with the experimental measurements, indicating their impact is easier to predict. Because many methods predict pathogenicity rather than stability, we also performed an analysis intended to remove variants known to cause pathogenicity via mechanisms other than stability. After excluding these positions, predictions showed improved agreement with the experimental stability measurements. This highlights the importance of mechanism-aware assessment. Together, the CAGI 7 ARSA challenge assessment reveals advances in variant impact prediction, while highlighting needs and future directions for the field.

Unravelling the RASopathy Syndromes using iPSC-derived neural disease modelling

Linda Wijaya

The Kids Research Institute Australia & University of Western Australia

Linda K. Wijaya 1,2, Catherine A. Forbes 1, Emma Kuzminski 1,2, Nicole C. Shaw 1,2, Paula Gomez 1, Kevin Chen 1, Mitchell Hedges 1, Gareth Baynam 3, Timo Lassmann 1,2, Vanessa S. Fear 1,2.

RASopathies is a group of developmental disorders caused by germline variants in genes regulating the RAS/MAPK signalling pathway. RASopathy syndromes, such as Noonan syndrome (NS), cardio-facio-cutaneous syndrome (CFC), and others, share overlapping features, including developmental delay, craniofacial dysmorphology, and cardiac malformations, making diagnosis challenging. While pathogenic variants are identified in about half of patients, many remain undiagnosed due to novel variants or variants of uncertain significance (VUS). Each VUS requires functional validation to determine pathogenicity, a process that can take more than 5 years, if achieved at all. There is an urgent need to accelerate variant interpretation to support diagnosis.

To address this, we developed a laboratory platform combining precise CRISPR-Cas9 gene editing in induced pluripotent stem cells (iPSCs) with neural disease modelling to fast-track genetic variant interpretation. In this study, BRAF variants, the most common cause of CFC, and present in a subset of NS cases, were investigated. Three iPSC lines harbouring a CFC-associated pathogenic variant (BRAF p.Leu499Glu), an NS-associated pathogenic variant (BRAF p.Trp532Cys), a VUS (BRAF p.Phe548Ser), and isogenic controls, were generated. These lines were differentiated into neural progenitor cells to assess neurodevelopmental phenotypes. Transcriptomic profiling, protein expression, and cell-based signalling analyses were performed to evaluate biological processes and variant-specific effects on RAS/MAPK pathway regulation.

Our results reveal variant-specific biological processes and differential MAPK activation, providing support data for the classification of the BRAF VUS. This study also demonstrates the utility of CRISPR-Cas9 gene editing and iPSC-based neural models for accelerating the interpretation of RASopathy variants.

SPONSOR TALK

Illumina Talk title

Illumina

TBC

Mutational profiling of SARS-CoV-2 PLpro exposes inhibitor escape routes

Melissa Call

Walter and Eliza Hall institute of Medical Research

Melissa Call[1], Xinyu Wu[1], Margareta Go, Julie Nguyen[1], Matthew Call[1]

1. Structural Biology Division, Walter and Eliza Hall Institute, Parkville, Australia[1]

The SARS-CoV-2 papain-like protease (PLpro) is essential for viral replication and a promising antiviral target. Using deep mutational scanning in mammalian cells, we mapped functional effects of nearly all single-residue substitutions. This analysis confirmed active-site requirements, uncovered an allosteric network connecting distal sites to catalysis, and showed that the blocking loop is mutation-tolerant yet dependent on conformational flexibility. We next assessed inhibitor resistance across three scaffolds: PF-07957472 and Jun12682, both GRL0617 derivatives, and the structurally distinct WEHI-P8. Escape mutations to the GRL0617-like inhibitors overlapped, whereas WEHI-P8 resistance was mediated by unique variants in the S4 pocket and blocking loop. Together, these findings define structural constraints that shape PLpro function and highlight how scaffold diversity is critical for building durable antiviral strategies.

References

[1] Wu, X., Go, M., Nguyen, J.V., Kuchel, N.W., Lu, B.G.C., Zeglinski, K., Lowes, K.N., Calleja, D.J., Mitchell, J.P., Lessene, G., Komander, D., Call, M.E., Call, M.J. (2024). Nat. Commun. 15, 6219.

[2] Wu, X., Devine, S.M., Go, M., Nguyen, J.V., Lu, B.G.C., Loi, K., Kuchel, N.W., Lowes, K.N., Mitchell, J.P., Lessene, G., Komander, K., Call,M.E., Call, M.J. (2025). PLoS Pathog. 21(9): e1013468.

1.8 Posters (Alphabetical, number)

Poster #1

Predicting changes in protein-protein binding affinity upon mutation with statistical potentials

Ciupitu A 1,*, Pucci F 1 (1) Université Libre de Bruxelles

In-silico approaches for predicting variant effects on protein-protein interaction (PPI) binding affinity continue to exhibit critical biases and difficulties generalizing beyond their training datasets. We propose a novel physics-based approach, termed "interface potentials," for the description of PPI interfaces and the prediction of the effect of mutations. These statistical potentials are derived from datasets of known PPI structures using the Boltzmann Law and describe the interfaces in terms of distances, torsion angles, and solvent accessibility of residues.

Poster #2

A human cell model system for interrogation of ATAD3 variants, a frequent cause of severe mitochondrial disease

Ann E. Frazier 1,2,*, Serene Regi John 3,4, Linden Muellner-Wong 1,3,4, Alison Compton 1,2,5, David Thorburn 1,2,5, David Stroud 1,3,4,5

1 – Murdoch Children's Research Institute, Melbourne, Australia

2 – Department of Paediatrics, University of Melbourne, Melbourne, Australia

3 – Department of Biochemistry and Pharmacology, University of Melbourne, Australia

4 – Bio21 Molecular and Biotechnology Institute, Melbourne, Australia

5 – Victorian Clinical Genetics Services, Melbourne, Australia

ATAD3 is an essential mitochondrial membrane-associated AAA+ ATPase implicated in diverse cellular processes, including mitochondrial dynamics, cholesterol metabolism, and nucleoid organization, although its precise molecular function is unclear. Notably, in hominids, the ATAD3 locus is a tandem repeat region encoding 3 genes: ATAD3A, ATAD3B and pseudogene ATAD3C. Variants in the highly conserved ATAD3A have been linked to mitochondrial disorders with a range of severities and symptoms. We have developed a human cell model system of ATAD3 deficiency, using genome editing to remove the entire ATAD3 locus while exogenously expressing ATAD3A under a doxycycline promoter. Removal of doxycycline from the cells results in rapid turnover of ATAD3A and halts cell growth within an 8-day period. A retroviral system enables constitutive expression of an additional copy of ATAD3A under an alternate promoter, establishing a system for testing of variants of uncertain significance (VUS). We have established a library of >15 cell lines expressing ATAD3A with a range of known pathogenic and VUS, across all domains of the protein. While expression of WT ATAD3A rescues cell growth upon doxycycline removal, preliminary data from several VUS demonstrates protein instability and reduced cell growth rates, confirming their pathogenicity. Further bespoke functional assessments (e.g. immunoblotting, proteomics, imaging) on select variants is in progress to determine their impact on ATAD3 function. Our strategy will not only advance mechanistic insights into ATAD3 biology but also informs the interpretation of pathogenic variants identified in patients and planned site-saturated mutational scanning approaches to interrogate poorly defined protein domains.

nf-core/deepmutscan: a community-oriented, modular bioinformatics pipeline for deep mutational scanning by shotgun sequencing

Poster #3

Mapping the effects of all mutations on neutralization by antibodies and engineered therapeutics for HIV Envelopes from divergent strains.

Caelan Radford 1,*, Jesse Bloom 1

1 - Basic Sciences and Computational Biology Program, Fred Hutch Cancer Center

Potential vaccines and therapeutics that target the HIV Envelope protein (Env) are challenged by its incredible diversity. This challenge is compounded by the fact that mutations can have vastly different effects on Env function or escape from neutralization in divergent strains of Env. We have developed a lentiviral vector deep mutational scanning platform that uses pooled libraries of barcoded genotype-phenotype linked pseudoviruses to measure the effects of all possible Env mutations on entry into cells and neutralization by antibodies and sera. We have used this platform to comprehensively measure the effects of mutations on escape from neutralization by antibodies for Envs from divergent strains, and found the extent to which measurements of mutation effects on neutralization can be generalized across Envs varies among antibodies. For some antibodies, mutations that cause escape from neutralization differ greatly between Envs. Even for antibodies for which mutations have similar effects on escape between Envs, mutation effects can have subtle differences. We are interested in continuing to use this platform to characterize newly identified antibodies and engineered therapeutics that target Env and compare them to therapeutics currently being developed. We are also interested in further studying human anti-HIV sera from individuals living with HIV or elicited through vaccination to better understand the neutralizing activity of broad and potent anti-HIV sera and how it could be elicited through vaccination.

Poster #4

Rules of the Rod: MYH7 Variant Effect Across the β-MHC Coiled-Coil Domain

Clayton E. Friedman 1,2,3,*, Shawn Fayer 4, Sriram Pendyala 4, Wei-Ming Chien 1,2,3,5, Charles E. Murry 1,2,3,6,7, Lea M. Starita 4,8, Douglas M. Fowler 4,7,8, Kai-Chun Yang 1,2,3,5

1 - Institute for Stem Cell and Regenerative Medicine, University of Washington, School of Medicine, Seattle, WA 98109, USA.

2 - Center for Cardiovascular Biology, University of Washington, Seattle, WA 98109, USA.

3 - Department of Medicine/Cardiology, University of Washington, Seattle, WA 98109, USA.

4 - Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA.

5 - Cardiology/Hospital Specialty Medicine, VA Puget Sound HCS, Seattle, WA 98108, USA.

6 - Department of Pathology, University of Washington, Seattle, WA 98109, USA.

7 - Department of Bioengineering, University of Washington, Seattle, WA 98195, USA.

8 - Brotman Baty Institute for Precision Medicine, Seattle, WA 98195, USA.

Hypertrophic cardiomyopathy (HCM) is an inheritable condition characterized by left ventricular heart thickening leading to heart failure and sudden cardiac death. Missense variants in MYH7 (encoding β-MHC) cause HCM, however, most MYH7 variants are of uncertain significance (VUS), limiting the utility of genetic testing for pre-clinical HCM screening. We established a high-throughput phenotyping pipeline and predictive model to prospectively determine MYH7 variant effect at scale. We used CRISPR-Cas9 to knock a MYH7 variant library spanning 45 amino acids of the coiled-coil domain into the endogenous MYH7 locus of human induced pluripotent stem cells. We differentiated this pooled library to cardiomyocytes and used FACS with next-generation sequencing to measure β-MHC-meGFP abundance for each variant. Aberrant β-MHC abundance accurately segregated known pathogenic (n=25) and benign (n=18) MYH7 variants (sensitivity 0.83; specificity 0.81) and reached Moderate strength OddsPath and OddsBenign scores, supporting classification of 15 VUS to likely pathogenic and 14 VUS to likely benign. We observe that β-MHC abundance is dependent on where amino acid substitutions occur within a repeating heptad motif. We built a structurally informed model to predict the effect of 17306 MYH7 variants spanning 1047 amino acids of the coiled-coil domain, which displayed similar accuracy to direct measurements and supports classification of >1000 VUS. We also find destruction of post-translational modification sites correlates with the ‘increased’ but not ‘decreased’ β-MHC abundance phenotype. We show disrupted β-MHC dosage largely explains MYH7 variant pathogenicity and provide an atlas of MYH7 variant effects to bridge genetic testing with life-saving interventions for HCM.

Poster #5

Deep Characterization of PTEN Variants on Canonical and non-Canonical Signaling Pathways

Corbin Glufka* 1, Mahir Taher 1, WC Sin 1, Jerry Tong 1, Warren Meyers 1, Kurt Haas 1

1 - University Of British Columbia

PTEN is an essential dual-specific protein and lipid phosphatase that tightly regulates various cellular processes. It is one of the most frequently mutated genes in cancer and is directly linked to PTEN hamartoma tumor syndrome (PHTS), a grouping of disorders that result in benign tumors throughout the body and significantly increase cancer risk. PTEN mutations have also been associated with developmental delays (DD) and autism spectrum disorder (ASD). Despite the widespread impact of PTEN-related conditions, treatment effectiveness varies greatly between individuals, largely due to unique genetic variations. The most common disease-associated PTEN mutations are single-nucleotide missense variants with largely unknown effects on protein function. Hundreds of PTEN mutations have been found throughout its structure, across disease states, making it difficult to link individual mutations to specific outcomes. Traditionally, studies of variant pathophysiology have focused on PTEN's canonical role in suppressing the growth-promoting Akt/PI3K pathway. However, PTEN has been linked to various non-canonical signaling through protein-protein interactions, its protein-phosphatase activity, and nuclear functions including DNA repair, Wnt, integrin, EGF, and insulin signaling. Here, six PTEN variants were functionally characterized relative to RefSeq across 18 phospho-markers associated with both canonical and non-canonical functions in three human cell lines using spectral flow cytometry. Unique variant- and cell-specific signaling profiles were identified across canonical and non-canonical pathways, highlighting the diverse functional consequences of PTEN mutations. These findings underscore the need for comprehensive variant-specific functional analysis to advance precision medicine and improve interpretation of PTEN mutations' clinical significance.

Poster #6

Integrating genomic and functional approaches to investigate differences of sex development and SRY regulation

Firman P. Idris[1,2*], Jocelyn van den Bergen[1], Gorjana Robevska[1], Lucas G. A. Ferreira[1,3], Hennie T. Bruggenwirth[4], Karen R. Ferreira[3], Marina M. L. Kizys[3], Yolande van Bever[4], Tara Hussein Tayib[5], Magnus R. Dias da Silva[3], Andrew H. Sinclair[1,2], Katie L. Ayers[1,2]

[1]The Murdoch Children's Research Institute, Melbourne, Australia.

[2]Department of Paediatrics, The University of Melbourne, Melbourne, Australia.

[3]Laboratory of Molecular and Translational Endocrinology, Endocrinology Division, Department of Medicine, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, Brazil

[4]Department of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands

[5]College of Medicine, Sulaimani University, Sulaimani, Iraq

Differences of Sex Development (DSD) are congenital conditions where chromosomal, gonadal, or anatomical sex development is atypical. Despite advances in genomic technologies, fewer than 50% of DSD cases receive a genetic diagnosis. To improve diagnostic yield and uncover the underlying biological mechanisms, it is crucial to explore both established and understudied contributors to sex development.

We analysed a cohort of 21 individuals with DSD and their families using chromosomal microarray and whole exome sequencing. Functional characterisation was performed for novel (p.Arg75Lys and p.Tyr198Cysfs*18) and previously reported (p.Asp58Glu, p.Met85Thr, and p.Arg86Term) SRY variants, including subcellular localisation assays and transcriptional activity assays using dual-luciferase reporter systems. We also examined non-coding variant in the SRY enhancer region (NC_000024.10:g.2792834C>T) responsive to SF1 using dual-luciferase reporter systems. Beyond direct genetic changes, regulatory factors influencing SRY function was aexamined. This included investigating how phosphorylation of SRY-regulating factors may affect its expression.

We identified both reported and novel variants in DSD-associated genes, including two novel frameshift variants in NR5A1 (p.Leu437Cysfs*59 and p.Val369Alafs*12), a key regulator of gonadal development. Functional analyses demonstrated that some SRY variants impair nuclear localisation and transcriptional activity, supporting their pathogenicity. Additionally, a variant in an enhancer region upstream of SRY had no impact on SRY activation and phosphorylation affect the activity of SF1 and GATA4.

Together, our findings underscore the complex interplay between coding variants, regulatory elements, and protein-level regulation in DSD, and emphasize the importance of combining genomic and functional approaches to improve diagnostic outcomes and understand the mechanisms of sex development.

Poster #7

It’s Not Just the Variant—It’s the Interface: A Human-Centered Look at MaveDB

Nguyen H 1, Starita L1,2, Stone J 2, Fowler D 1,2

1- Department of Genome Sciences, University of Washington, Seattle, WA, USA

2- Brotman Baty Institute for Precision Medicine, University of Washington, Seattle, WA, USA

Accurately interpreting genetic variants poses a major challenge in clinical genomics. In particular, a substantial proportion of variants discovered via genetic testing are of uncertain significance (VUS) and thus cannot inform clinical decisions. Although multiplexed assays of variant effect (MAVEs) and other functional assays offer evidence capable of resolving VUS at scale, fragmented data resources, inconsistent metadata structures, and substantial usability barriers in existing interfaces limit the integration of these data into clinical workflows. To address these challenges, we are evaluating MaveDB, an open-sourced database for MAVE datasets, to assess how effectively the platform supports researchers, computational analysts, and clinical users. This evaluation deploys a user-centered research framework to investigate the usability, accessibility, and workflow alignment of MaveDB. Our work includes a heuristic evaluation of the existing interface, a systematic review of user pain points, and the ongoing development of a clinical usability study. We are investigating how users search for and interpret MAVE data, as well as how they act upon that data to identify features, visualization, or interaction patterns that impact the use of functional evidence in both clinical and research settings. The goal of our work is to inform design improvements to MaveDB that support intuition in navigation, data context, and better integration with clinical variant classification paradigms. The motivation of the study is to enable the variant effect data in MaveDB to reach their full potential in precision medicine by maximizing usability, accessibility, and effectiveness in supporting clinical decision-making workflows.

Poster #8

Functional classification of population variants demonstrates that SLC6A19 loss-of-function is associated with reduced progression to kidney failure

Hannah N. De Jong 1,*, Sahar V. Mozaffari 1, Nora I. Scherer 2, Elena Butz 2, Nathan Fastman 1, Julie Ullman 1, Maarten Hoek 1, Robert R. Graham 1, Nathan A. Sallee 1, Anna Köttgen 2, Karol Estrada 1

1- Maze Therapeutics, South San Francisco, CA, USA

2- University of Freiburg, Freiburg im Breisgau, Germany

SLC6A19 is expressed in the kidney, and recaptures free amino acids from the diet before excretion in the urine. Protein-truncating variants in SLC6A19 are strongly associated with improved estimated glomerular filtration rate (eGFR), a metric for kidney function. However, the impact of naturally occurring missense variants in SLC6A19 is not well described.

We therefore compiled 365 SLC6A19 variants for assay, including all those present at an allele count >1 among UK Biobank (UKBB), All of Us (AOU), and German Chronic Kidney Disease (GCKD). Expressing each variant in HEK293 cells, we quantified SLC6A19 protein abundance by flow cytometry and (for a small subset) radiolabeled amino acid uptake activity. Of these, 222 variants were classified as experimental loss-of-function ("xLOF") by low protein abundance, and an additional 7 variants were classified as xLOF by reduced uptake activity.

Using these xLOF variants, we performed a burden analysis to elucidate the role of SLC6A19 in kidney disease. The xLOF variants combined are significantly associated with higher eGFR in UKBB (p=2.48e-14) and AOU (p=0.005). Further, we demonstrate that xLOF variants, in combination with previously identified LOF variants, are significantly associated with decreased incidence of composite kidney endpoint (CKE: 40% eGFR decline, eGFR <15 mL/min/1.73m2, dialysis, or transplant) in both a general cohort (UKBB, HR = 0.63, p=0.016) and meta-analysis of chronic kidney disease cohorts (UKBB stage 2 and GCKD, HR= 0.51, p = 0.001). These data support the scientific rationale for studying SLC6A19 inhibition as a potential therapeutic strategy for CKD.

Poster #9

Advancing the Laboratory Foundations of a Scalable Saturation Genome Editing Pipeline

Ian Lobb 1,*, Dijue Sun 1, Stuart Fawke 1, Leah Wells 1, Jamie Billington 1, Adam Hunter 1, Sanger MAVE Project Team 1, Sebastian Gerety 1, Andrew Waters 1, David Adams 1, Matthew Hurles 1, Sonia Goncalves 1, Charlotte Beaver 1

1 – Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SA, United Kingdom

Multiplexed assays of variant effect (MAVEs), particularly Saturation Genome Editing (SGE), have been widely adopted for interrogating the functional consequences of genetic variation They provide a powerful framework for systematically assessing the functional impact of genetic variation within its native genomic context. To realise this potential at scale, we have focused on strengthening the laboratory components of our SGE pipeline to deliver high editing precision, reproducibility, and throughput.

As part of this optimisation, we conducted a genome-wide CRISPR knockout screen in our HAP1 LIG4 KO cells to identify genes displaying appropriate essentiality profiles for inclusion in SGE experiments. This data-driven approach enables informed gene selection, giving us the best possible chance of success and minimising failure rates. In parallel, we introduced an improved DNA quantification method to more accurately measure and normalise donor and guide components, enabling precise stoichiometry at transfection. Refinements to transfection parameters — including reagent ratios, timing, and cell handling — further increased editing efficiency and cell viability.

Together, these improvements have led to homology-directed repair (HDR) rates routinely exceeding 50%, significantly enhancing the reliability and consistency of experimental outcomes. These optimisations form a robust experimental foundation for high-quality SGE data generation and integration with informatic analysis pipelines.

By embedding systematic testing and iterative refinement into each stage of the workflow, we have established a scalable and reproducible laboratory platform that supports comprehensive functional genomics studies across diverse gene targets.

Poster #10

Scaling DNA Synthesis Enhances Gene Synthesis and DNA Variant Library Generation

Ju Zheng, Ming Dong, Xiandi Zhang, Xiang’er Jiang, Yue Shen

GCATbio UK

Conventional high-throughput DNA synthesis technologies use intricate chip and microfluidic systems to produce large-scale synthetic oligonucleotides, which depend on chip manufacturing and complex microfluidic systems, but with low concentration and limited compatibility for long DNA assembly. Here we highlight the microchip-based massive in parallel synthesis (mMPS) system, with an “identification–sorting–synthesis–recycling” iteration mechanism applied to microchips for high-throughput DNA synthesis. This achieves not only high-throughput but also offers higher yield and enhanced quality stability compared to conventional methods. We further explore its broad impact on downstream applications, particularly in high-throughput gene synthesis and DNA variant library generation.

Poster #11

Importance of higher-order epistasis in protein sequence-function relationships

Sethi, Palash 1, Zhou J 1, 2, *

1 - Department of Biology, University of Florida

2 - Genetics Institute, University of Florida

Protein sequence–function relationships are inherently complex, as amino acids at different positions can interact in highly unpredictable ways. A key question for protein evolution and engineering is how often epistasis extends beyond pairwise interactions to involve three or more positions. Although experimental data has accumulated rapidly in recent years, addressing this question remains challenging, as the number of possible interactions is typically enormous even for proteins of moderate size. Here, we introduce an interpretable machine learning framework for studying higher-order epistasis scalable to full-length proteins. Our model builds on the transformer architecture, with key modifications allowing us to assess the importance of higher-order interactions by fitting a series of models with increasing complexity. Applying our method to 10 large protein sequence-function datasets, we found that while additive effects explain the majority of the variance, within the epistatic component, the contribution of higher-order epistasis ranges from negligible to up to 60%. We also found higher-order epistasis is particularly important for generalizing locally sampled fitness data to distant regions of sequence space and for modeling an additional multi-peak fitness landscape. Our findings suggest that higher-order epistasis can play important roles in protein sequence-function relationships, and thus should be properly considered in protein engineering and evolutionary data analysis.

Poster #12

Human CRISPR-oligo recombineering of a ribosomal protein reveals a functional S-acylated cysteine

Kevin W. Yeung 1,2,*, Sarah Hassan 2,3, Ione C. Goodwin 2, Louise J. Walport 1,4, Mauricio Barahona 5, Sophia N. Yaliraki 1, Edward W. Tate 1,4, Matthew A. Child 2

1 - Department of Chemistry, Imperial College London, London, UK

2 - Department of Life Sciences, Imperial College London, London, UK

3 - London Biofoundry, Innovation Hub, White City Campus, Imperial College London, London, UK

4 - The Francis Crick Institute, UK

5 - Department of Mathematics, Imperial College London, London, UK

Post-translational modifications (PTMs) tune the function and localization of proteins following their biosynthesis, altering the biochemistry of the target residue. Consequently, PTM dysfunction has been associated with a wide range of diseases. While chemical proteomics has pioneered the identification and characterization of PTMs, their functional annotation has been hindered by their dynamic nature and strong dependence on local structural and regulatory context. Here, we apply human CRISPR-oligo recombineering (huCORe) to a highly conserved region of a ribosomal protein to assess the functionality of an S-acylated cysteine. We find distinct patterns of mutational sensitivity at this cysteine consistent with disruption of PTM-dependent regulation. In contrast, despite sequence conservation, most mutations proximal to this cysteine are well-tolerated, and in some instances, lead to a gain of function. Follow-up biochemical analysis confirms that S-acylation is abrogated in the cysteine mutant. Our approach demonstrates that biochemical mutational tolerance analysis can effectively resolve the functional roles of individual PTMs, enabling systematic interrogation of protein regulation at scale.

Poster #13

Clinical Impact of Integrating BRCA2 MAVE Data for Variant Reclassification: A Real-World Evaluation

Kristin Kohler 1,*, Ashley P. Marsh 1, Colin Young 1, Jamie Weyandt 1, Sarah Dugger 1, Daniel Zimmerman 1, Magan Trottier 1, Steven Harrison 1, Marcy E. Richardson 1 1-Ambry Genetics, Aliso Viejo, CA 92656, USA

MAVE studies provide high-throughput functional data that can refine the interpretation of genetic variants, especially for cancer risk genes like BRCA2. However, the clinical application of MAVE data to reclassify variants of uncertain significance (VUS) has not been thoroughly evaluated in a real-world setting. This study investigates the clinical utility of integrating MAVE data into variant reclassification workflows and quantifies the downstream patient impact.

Following ACMG/AMP guidelines, we incorporated functional data from two MAVEs to reclassify BRCA2 VUS previously classified in a clinical laboratory. Known pathogenic/likely pathogenic (P/LP) and benign/likely benign (B/LB) variants were used to calibrate these assays. 1,430 variants functionally assessed in the MAVE assays had a prior classification-including 981 VUS. MAVE data were highly concordant with prior clinical interpretations, with only one variant showing discordance across both assays. Of 981 VUS, 54% (534) were reclassified, with 97% (518) reclassified as LB/B and 3% (16) as P/LP. Reclassifications impacted 4,362 patients, with 4,290 reports reclassified as VUS to LB/B, and 72 reports reclassified as VUS to P/LP. Variants reclassified as P/LP included canonical splice (3/16), missense (10/16), and intronic (3/16), whereas reclassified LB/B variants included missense (517/518) and synonymous (1/518). Cascade testing of 24 individuals across 13 families resulted in 14 additional positive reports, highlighting the broader impact of MAVE data on family-based risk evaluation.

This study provides one of the first large-scale, real-world applications of MAVE data integration into a clinical workflow, offering a valuable model for resolving VUS and improving precision in hereditary cancer risk assessment.

Poster #14

DISCOVER, CONNECT, CONTRIBUTE: ENGAGEMENT PATHWAYS WITHIN THE ATLAS OF VARIANT EFFECTS ALLIANCE

Lara Muffley1,2, Lea M Starita1,2 Alan F Rubin4,5, Atlas of Variant Effects Alliance Research group, Douglas M Fowler1,2

1Department of Genome Sciences, University of Washington, Seattle, WA, USA
2 Brotman Baty Institute for Precision Medicine, Seattle, WA, USA
4 Bioinformatics Division, WEHI, Parkville, VIC, Australia
5 Department of Medical Biology, University of Melbourne, Parkville, VIC, Australia

The Atlas of Variant Effects (AVE) Alliance is a global community of over 800 members across 57 countries, dedicated to accelerating discoveries in functional genomics with direct applications in precision medicine. At its core are Multiplexed Assays of Variant Effect (MAVEs), which enable high-throughput measurement of thousands of genetic variants in a single experiment.

The Alliance fosters a collaborative and inclusive community committed to the open exchange of MAVE data, protocols, and methodologies. Outputs from Alliance activities include shared resources that accelerate progress in variant interpretation.

Participation in the Alliance is open to all, and members are actively encouraged to engage through a variety of pathways:

-Propose a topical interest group to explore emerging scientific themes

-Join workstreams and committees during open calls

-Share experimental protocols via the AVE protocols.io repository

-Attend or present at the monthly Variant Effects Seminar Series

-Participate in the annual Mutational Scanning Symposium

-Tune into the Variants and Us podcast for community insights

-Explore educational content on AVE’s YouTube channel

-Engage with the community on LinkedIn and BlueSky

-Contribute to or utilize the MAVE resource collection

-Register your project on MaveRegistry

-Deposit a variant effect map into MaveDB

-Utilize MaveMD for clinical insights

Through our community of practice, we have now amassed over 7 million variant effect measurements. These measurements span over 700 distinct human genes as well as genes in model organisms, all publicly available through MaveDB, our open-source database built to enable FAIR (Findable, Accessible, Interoperable, and Reusable) data sharing and reuse.

Through dedicated workstreams and collaborations, AVE has produced a rich community-driven body of work which include over a dozen peer-reviewed publications that advance MAVE tools, set standards, inform experimental design, and clinical interpretation of the data.

Here we showcase these collaborative achievements and outline specific pathways for participation in this growing international effort to map and apply variant effects at scale.

Keywords: Atlas of Variant Effects, Multiplex Assays of Variant Effect, Community Outreach, Education, International Coordination

Email address of first author: muffley@uw.edu

Poster #15

Differentiation of IPS cells into fibroblasts to validate gene targets for Dupuytren's Contracture

Worth L 1,2,3, Pearce R 1,4, O'Brien A 1,5, Barrett L 1, Kuzminski E 3, Hedges M 3, Forbes C 3, Shaw N 3, Fear V 3, Fear M 1, Stevenson A 1

1 Burn Injury Research Unit, University of Western Australia

2 Paediatric Burns Care, The Kids Research Institute Australia

3 Translational Genetics Team, The Kids Research Institute Australia

4 The University of Notre Dame Australia, 5Royal College of Surgeons in Ireland

Dupuytren's contracture (DD) is a painful fibroproliferative disease of the hand palmar fascia, and has a strong genetic risk component, although the exact mutation/s are unknown. Reducing movement of the fourth and fifth fingers, they become painful to straighten, strongly impacting quality of life. DD has a prevalence of 22% in people above the age of 60 years in Australia, and current treatment options for DD include physiotherapy and surgical intervention which only offer a temporary respite. This study tested the significance of a potential SNV for DD while developing a novel fibroblast differentiation protocol with the hope to take a step towards better treatment options for patients. Whole exome sequencing of a Western Australian family with DD identified a series of single nucleotide variants (SNVs) of interest. To validate one of the SNVs as a genetic cause of DD, CRISPR Cas9 was used to introduce a heterozygous SNV seen in the LRSAM1 gene of DD patients into IPS cells. The mutated IPS cells underwent single cell cloning for SNV or wildtype genotype, and populations were validated by targeted amplicon sequencing before fibroblast differentiation. Differentiated cells were compared phenotypically to DD patient cells, using transcriptomics to validate whether this SNV in LRSAM1 was pathogenetic. Further testing was done through fibrosis assays, immunocytochemistry, and immunoblotting, demonstrating that the SNV clones exhibit a more fibrotic phenotype then their wildtype pairs. By identifying and validating causative SNVs for DD, more effective treatment options can be designed resulting in a better quality of life.

Poster #16

Understanding the delivery of an ASO therapy for Epidermolysis Bullosa

Worth L 1*, Ham K 2, Joesoef E 1, Aung-Htut M 2, Norret M 3, Evans C 3, Iyer K 3, Wood F 1, Barrett L 1, Fear M 1, Stevenson A 1

1 Burn Injury Research Unit, University of Western Australia

2 Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University

3 Bionano, University of Western Australia

Recessive dystrophic Epidermolysis Bullosa (RDEB) is the most severe form of a painful, rare genetic disorder resulting in easily blistered and broken skin, which has no cure. RDEB is caused by mutations in COL7A1, causing a collagen VII deficit between the epidermis and dermis in the skin, allowing these layers to easily separate and form open wounds. Current treatments are limited and treat symptoms rather than the underlying cause of the disease.

A promising treatment option for RDEB involves targeting COL7A1 with antisense oligonucleotides (ASOs). An ASO has been designed to restore functional collagen VII production in RDEB patients with mutations in exon 73 by inducing exon-skipping. This study investigated delivery of this ASO to both skin cells in vitro and a simulated ex vivo delivery method. Using high resolution confocal microscopy, the ASO in dermal fibroblasts and epidermal keratinocytes was observed to enter the cell within 15 min and the nucleus by 30 min. To develop a new method of ASO delivery to patients, a novel protocol of lipid transfection in a commercially available Autologous Skin Cell Suspension (ASCS) product was tested. This ASCS is used intraoperatively to treat wounds, and has potential in treating genetic skin disorders. High efficiency of ASO transfection in the ASCS has been achieved in surgically relevant time points, allowing potential clinical application. By understanding ASO delivery to RDEB patients, development of a treatment that addresses the cause of RDEB becomes one step closer, allowing patients' better quality of life.

Poster #17

Deep insight into protein through Deep Mutational Scanning.

Margareta Go 1,*, Xinyu Wu1,2, Melissa Call 1,2

1. Walter and Eliza Hall Institute of Medicl Research

2. Department of Medical Biology. University of Melbourne

Studying protein function, behaviour and structure can be both costly and time- consuming.

Traditional individual mutagenesis approaches often provide limited insights and may lead to incomplete conclusion. Nowadays, advances in cloning techniques and Next Generation Sequencing platform have transformed protein research, enabling more detailed analysis in less time and at lower cost.

The Multiplexed Assay Technology Hub (MATH) at Walter and Elisa Institute offers a Deep Mutational Scanning (DMS) service designed to optimise both time and resources.

Our services include support for experimental design, library preparation, selection steps, preparation of next-generation sequencing samples, and data processing.

All services are customised and delivered in a time-efficient and cost-effective manner.

Our aim is to assist researchers in obtaining meaningful insights into the functional impact of protein variants and achieving a better understanding of their proteins.

Poster #18

A Bioinformatics Ecosystem for High Throughput MAVE-SGE Experiments

Matthew Forbes 1,*, Anna Shtanova 1, Elvira Hernández Gutiérrez 1, Shriniwas Kulkarni 1, Swee Hoe Ong 1, Miguel Rada 1, Zakiya Salemahomed 1, Helen Schuilenburg 1, Priyanka Surana 1, Leah Wells 1, Jamie Billington 1, Adam Hunter 1, Ian Lobb 1, Rahul Sadanandan Nair 1, Wendy Thorpe 1, Sanger MAVE Project Team 1, Sonia Gonçalves 1, David Parry-Smith 1

1 – Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SA, United Kingdom

Multiplexed assays of variant effect (MAVEs) have become a widely adopted approach for systematically assessing the functional consequences of genetic variation, with saturation genome editing (SGE) emerging as a powerful CRISPR-based MAVE technique that enables variant impact to be measured in its native genomic context. A key challenge to scaling these experiments lies in the lack of integrated bioinformatics workflows supporting the entire MAVE-SGE process — from target design to laboratory execution, through to the analysis and interpretation of next-generation sequencing data.

In this poster, we present the conceptual framework underpinning the Sanger Institute's MAVE Programme for building scalable informatics solutions. Central to this approach is the concept of the "targeton": a consistently sized segment of the target sequence that provides a standardised unit for designing, executing, and analysing MAVE-SGE experiments.

We outline how this concept informs the design of robust workflows for targeton selection, supports laboratory processes, and enables high-throughput analysis and interpretation of functional data at scale. We also describe the technical capabilities and technology stack on which these solutions will be built. This work aims to enable a scalable, end-to-end MAVE-SGE pipeline that produces high-quality data and expands the potential impact across a broad range of target genes.

Poster #19

The impact of Polygenic Risk Scores on a variable penetrance alleles of BRCA1

Wakefield MJ 1,2,*, Beard S 1, James P 2, Spurdle A 3, Rubin AF 2, Starita L 4, AusMAVE Consortium

1 - WEHI, Parkville, Australia

2 - University of Melbourne, Parkville, Australia

3 - QIMR Berghofer, Brisbane, Australia

4 - University of Washington, Seattle, USA

The BRCA1 exon 17 allele R1699Q confers an intermediate risk for breast and ovarian cancer. We set out to determine if saturation genome editing (SGE) can detect this intermediate phenotype and determine if individual differences and polygenic risk scores (PRS) influence the SGE assay estimate of this variant effect. We established lymphoblastoid cell lines from individuals with known PRS scores and used dual CRISPR editing with nucleofected riboprotein to generate cell lines hemizygous for the BRCA1 locus. We successfully established hemizygous cell lines from three high risk, one intermediate risk and one low risk PRS genotype.

These cell lines have been used for SGE of BRCA1 exon 17 by nuclofection of riboprotein with an oligo homologous repair library.

A pilot experiment with a limited allele set suggests that variation in intermediate allele impact exists between individuals and that this variation is independent of the polygenic risk score, and that the individual variation also differs between PARP inhibitor selection and growth advantage in the absence of inhibitor.

We will present results from the full replicated exon 17 SGE and discuss potential drivers of independent variant effect variability in cell lines and PRS scores for BRCA1 alleles.

Poster #20

Investigating the mechanism of a non-coding deletion variant in the promoter of beta-cell repressed gene SLC16A1 causing exercise-induced hyperinsulinism.

Vasiliadis M 1*, Bennett J 1, Hopkinson J 1, Owens NDL 1, Flanagan S. 1-

Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter, UK

We recently identified a novel heterozygous 94 base pair deletion variant in the promoter region of SLC16A1 causing a rare congenital form of insulin-secretion disorder known as exercise-induced hyperinsulinism. This variant was found in 31 individuals from 11 families with hyperinsulinaemic hypoglycaemia, characterised by low blood glucose levels due to inappropriate insulin secretion. SLC16A1 is expressed in most cell types and encodes monocarboxylate transporter-1, which facilitates the import of pyruvate and other monocarboxylates into cells. In pancreatic beta-cells, SLC16A1 is silenced to prevent entry of pyruvate or lactate that would trigger glucose-independent insulin secretion. Immunohistochemistry and functional studies of the variant demonstrated aberrant SLC16A1 expression in carriers' beta-cells, strongly suggesting this as the causal disease mechanism. The mechanism through which the deletion variant ultimately leads to loss of repression of SLC16A1 in beta-cells, however, remains unclear.

Using publicly available chromatin accessibility data across 222 human cell-types, we identified a peak of accessibility overlapping the promoter variant in beta-cells, despite its repressed state. Comparably, cell types which express SLC16A1 have a substantially larger peak of accessibility over the gene's promoter, in line with the presence of transcriptional machinery. Analysis of chromatin conformation data reveals that cells expressing SLC16A1 exhibit a chromatin loop to 74bp upstream, and we hypothesise that as-yet unknown factors binding in beta-cells prevent the formation of this loop, maintaining SLC16A1 repression. Ongoing work aims to characterise the factors bound in beta-cells creating this peak of accessibility and possibly being disrupted by the variant.

Poster #21

CountESS, a user-friendly platform for analysing MAVE data

Moore N 1*

1 - Mnemote Pty Ltd

We present CountESS (Count-based Experiment Scoring and Statistics), a visual, modular and extensible open-source program for processing experimental data from MAVEs and other functional genomics experiments. CountESS supports a variety of assays and methodologies including saturation genome editing and bin-based experiments (e.g. VAMP-seq) and provides a flexible framework for interactive or pipeline-driven data analysis. At each step, the graphical user interface shows tabular data previews to enable prototyping and exploration, allowing customised filtering and quality control to be added and immediately reviewed throughout the workflow.

Modules are included for reading common bioinformatics file formats and for performing standard data analysis tasks, such as associating barcodes to full-length sequences and calling variants with CountESS's own simple variant caller or using a minimap2 wrapper module. The software also has several pre-built methods for transforming counts into scores, as well as robust support for implementing and benchmarking custom approaches. Building a pipeline interactively from pre-existing modules allows for fast, flexible and reproducible data analysis with minimal programming requirements. Tools for automatically creating ready-to-run workflows based on a template are also provided.

CountESS uses Python, Apache Arrow and DuckDB to easily handle larger-than-memory data sets. The cross-platform graphical user interface can run locally or be forwarded from a server over SSH. Workflows can also be executed from the command line locally or on an HPC cluster. The software is fully open source and features extensive developer documentation for researchers who wish to contribute to the project or create and share their own modules independently.

Poster #22

Modelling motile cilia for a DNAI2 variant of uncertain significance using CRISPR edited iPSCs and directed differentiation

Shaw N 1*, Forbes C 1, Wijaya L 1, Wright K 1,2 Fear V 1

1 - The Kids Research Institute Australia, The University of Western Australia, Nedlands, WA, Australia.

2- School of Biomedical Sciences, The University Western Australia, Nedlands, WA, Australia.

Dynein Axonemal Intermediate Chain 2 (DNAI2) is one of approximately 45 genes that have been associated with Primary Ciliary Dyskinesia (PCD) and is essential for dynein complex assembly in motile cilia. In this study, we investigate a variant of unknown significance (VUS) in DNAI2 (c.565C>T, p.R189W) identified by prenatal genetic screening of a foetus with ventriculomegaly. We hypothesised that abnormal cilia function due to the compound heterozygous DNAI2 mutation is responsible for the patient's phenotype. CRISPR/Cas9 gene-editing was used to introduce the DNAI2 VUS into iPSCs. Two wild-type and two compound heterozygous DNAI2 VUS iPSC clones with corresponding knock-out variants, c.571_580del, p.(P191*) and c.571_579delinsACAGG, p.(P191Tfs*3), were derived. iPSCs were differentiated into ciliated airway epithelial cells via intermediate enrichment steps using Magnetic-activated Cell Sorting (MACS) for lung progenitors (CPM+) and airway basal cells (NGFR+) prior to seeding on Transwell permeable supports for culture at the air-liquid interface. Mature airway epithelial cells were assessed for the presence and function of cilia by immunofluorescence, cilia ultrastructure by transmission electron microscopy (TEM) and cilia beat frequency (CBF) using high-speed video-microscopy (HSVM). Ciliated airway epithelial cells were successfully differentiated from all iPSC clones with positive acetylated α-tubulin immunostaining and beating cilia. Cilia developed in airway epithelial cells with the DNAI2 c.565C>T VUS had normal ciliary axoneme ultrastructure and similar CBF as healthy airway cells. SEM structural analysis, and cilia beat pattern were investigated. The motile cilia disease model provides a robust system for the investigation of cilia dysfunction for a compound heterozygous DNAI2 mutation.

Poster #23

NAXD deficiency: heterogeneous phenotypes, positive response to niacin treatment and drug repurposing studies

Nicole Van Bergen241, Najmesadat Seyedkatouli2, Liana Semcesen3, Lucia Gallucci2, Jean-François Conrotte2, Mei Du1, Tim Sikora1, Marat Kasakin2, Gezime Seferi2, Licia Corona2, Martin Jakubec2, Brunda Nijagal4, Sajel Lala5, Rebecca D Ganetzky6, Ana Maria Rodriguez5, Marina Szlago7, Melanie Wong8, Margit Shah8,9, James Nurse10,11, Nicola Foulds10, Shankar Sadagopan10, Thu Ha Nguyen12, Chi Dung Vu12, Ngoc Khanh Nguyen12, Michelle De Silva13, Fernando Rossello3, Mirana Ramialison3, David R. Thorburn1,13,14, Matthew Lynch15, Pauline McGrath15, Henry Beetham1, Alejandro Hidalgo-Gonzalez1, David A. Stroud1,3,14, John Christodoulou1,13,14, Carole L Linster2.

1 Murdoch Children's Research Institute, Royal Children's Hospital, Parkville, VIC 3052, Australia.
2 Enzymology and Metabolism Group, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 6 Avenue du Swing L-4362 Esch-sur-Alzette, Luxembourg.
3 Department of Biochemistry and Pharmacology, Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Parkville, VIC 3010, Australia.
4 Metabolomics Australia, Bio21 Institute, The University of Melbourne, Melbourne, VIC, Australia.
5 Division of Clinical Genetics, Nickelaus Children's Health System, 3100 SW 62(nd) Avenue, Coral Gables, USA.
6 Mitochondrial Medicine Frontier Program, Division of Human Genetics, Department of Pediatrics, Children's Hospital of Philadelphia.
7 Hospital de Niños Dr. Ricardo Gutiérrez, City of Buenos Aires, Argentina.
8 The Children's Hospital at Westmead, Westmead, Australia.
9 Faculty of Medicine and Health, University of Sydney, Sydney, NSW, Australia.
10 University Hospital Southampton NHS Foundation Trust, Southampton, UK.
11 Faculty of Medicine, University of Southampton, Southampton, UK.
12 Center for Endocrinology, Metabolism, Genetics/Genomics and Molecular Therapy, Vietnam National Children's Hospital
13 Victorian Clinical Genetics Services, Murdoch Children's Research Institute, Melbourne, Victoria, Australia.
14 Department of Paediatrics, University of Melbourne, Melbourne, Victoria, Australia.
15 Neurosciences Unit, Queensland Children's Hospital, Brisbane, Australia.

Encephalopathy, progressive, early-onset, with brain edema and/or leukoencephalopathy, 2 (PEBEL2) features rapidly progressive neurodegeneration following febrile illness or infection in infancy. PEBEL2 is caused by biallelic pathogenic variants in NAXD. Evidence suggests that location of the pathogenic variant may correlate with clinical presentation. Here we describe nine new cases of PEBEL2. Four showed neurological deterioration triggered by periods of fever and/or illness. These children had either a homozygous missense variant in the catalytic domain, or a missense variant and frameshift variant in trans. Four cases presented with a cardiac phenotype, triggered after illness. These children had pathogenic variants in the mitochondrial targeting sequence (MTS). One individual was prematurely stillborn with severe in utero neurodegeneration and had biallelic missense and frameshift variants.

We will present the study of four missense variants, all of which affected NAXD protein solubility, NADHX dehydratase activity and/or thermostability. Fibroblasts from three cases had elevated concentrations of the damaged NAD(P)HX cofactors and significantly reduced NAXD protein levels. Individual PEBEL2 cases have shown benefit in response to high-dose vitamin B3. Four of the PEBEL2 cases have been treated with niacin, and all treated patients are still alive and have survived repeated febrile illness events. In ongoing work we are identifying novel therapeutic interventions using high-throughput screening (HTS). We have screened a 4222 FDA-approved compound/drug library and identified 11 hit compounds, currently undergoing further testing in our cell and animal models. With early diagnosis and prompt introduction of these precision therapies, we envisage that the clinical course of this devastating disorder could be dramatically improved.

Poster #24

Variant Effect Predictors Show Systematic Bias in Intrinsically Disordered Regions

Norbert Deutsch 1,*, Zsuzsanna Dosztanyi, 1

1 - Eötvös Loránd University, Budapest

VEP benchmarking is crucial but remains structurally incomplete, focusing on folded domains while overlooking Intrinsically Disordered Regions (IDRs).This is a critical gap, as IDRs mediate key functions from signaling and liquid-liquid phase separation to providing interaction sites via short linear motifs (SLiMs). These diverse functions evolve under different constraints, rendering uniform metrics like sequence conservation unreliable.

Our systematic analysis shows VEPs have significantly reduced accuracy in IDRs, a structural bias evident both in classifying ClinVar variants and in reduced correlation with deep mutational scanning data from the ProteinGym benchmark. We find this bias is compounded by a dataset-specific bias, where high inter-predictor correlation on ClinVar breaks down proteome-wide, indicating overfitting. This general underperformance, however, masks a critical distinction: predictors that perform relatively better are precisely those whose score distributions show significant separation, assigning higher average pathogenicity to known functional elements versus other regions. This ability correlates with their capacity to capture distinct, IDR-specific functional patterns, such as ""island-like"" pathogenicity of SLiMs, highlighting the importance of local functional context.

Improving IDR variant assessment requires a new approach centered on this functional context. This highlights an urgent need for new benchmarks that incorporate these IDR-specific functional annotations, and for integrating IDP functional expertise with AVE assessment expertise. Our work provides the quantitative basis to initiate this integration.

Poster #25

Unravelling CACNA1A Channelopathies: Precise CRISPR Gene Editing and Transcriptomic Profiling for Rapid Functional Analysis of Rare Genetic Variants

Wright O K 1,2,*, Forbes C 1, Gomez P 1,2, Morris K 1, Smith M 1, Wijaya L K 1, Shaw N 1, Fear V S 1

1 - Translational Genetics Team, The Kids Research Institute Australia

2 - School of Biomedical Science, University of Western Australia

Pathogenic variants in the neural calcium channel gene CACNA1A are associated with a spectrum of rare disease characterised by paroxysmal movement disorders, cerebellar dysfunction, and neurodevelopmental phenotypes. Genomic sequencing of patients with suspected CACNA1A channelopathies often reveals novel genetic variants, or variants of uncertain significance (VUS). For these patients, diagnosis is delayed until a functional assessment of variant effect is performed. This study investigates the use of CRISPR gene editing, stem cell neural disease modelling, and transcriptomic profiling to rapidly assess the functional impact of two CACNA1A patient VUS (CACNA1A p.Ala1959Asp, and CACNA1A p.Lys1438del). Precise CRISPR-Cas9 HDR gene editing is used to introduce patient VUS into induced pluripotent stem cells (iPSCs). Paired VUS and WT iPSCs are differentiated into neural progenitor cells (NPCs) and further matured into synaptically active neurons. Expression of stem markers (OCT3/4 and NANOG) and neural markers (SOX1, PAX6, and Nestin) is measured through flow cytometry to assess NPC differentiation. Neural maturation is assessed through immunofluorescence staining for neural markers TUBB3, MAP2, NeuN, Syanapsin and FOXG1, and glial marker GFAP. Expression and localisation of CACNA1A is also confirmed by immunofluorescence staining. The impact of each CACNA1A VUS on cellular function is assessed through whole-cell patch-clamp electrophysiology, calcium flux assay, and transcriptomic profiling. Beyond providing evidence toward the classification of two CACNA1A patient VUS, this work assesses the capacity of transcriptomic profiling to provide rapid, high throughput functional analysis of CACNA1A channelopathy variants, decreasing the time to diagnosis for many patients.

Poster #26

Transdifferentiation of Human Dermal Fibroblasts to iNeurons and iAstrocytes for Functional Variant Studies

Olivia Robinson 1,2*, Alison Gardner 1,3, Emmylou Nicolas-Martinez 1,2, Jamie V Zhang 1,2, Tarin C John 1,2, Tracy Dudding 4, Jean-Vianney Barnier 5, Jozef Gecz 1,3, Lachlan Jolly 1,2

1. The Robinson Research Institute, The University of Adelaide, Adelaide, SA, Australia
2. School of Biomedicine, The University of Adelaide, Adelaide, SA, Australia
3. School of Medicine, The University of Adelaide, Adelaide, SA, Australia
4. NSW Genetics of Learning Disability (GOLD) Service, Hunter Genetics, Newcastle, NSW, Australia
5. Institut des Neurosciences Paris-Saclay, Université Paris-Saclay, Saclay, France

We defined a list of 516 silent neurological disease genes (SNGs) based on limited or no expression in clinically accessible tissues of blood and skin forgoing functional investigations of genetic variants using patient derived cells. We investigated whether transdifferentiation of patient derived human dermal fibroblasts directly into induced Neurons (iNeurons) or iAstrocytes could induce the endogenous expression of SNGs providing an approach to overcome this challenge.

Overexpression of cell fate directing transcription factors ASCL1 & NEUROG2 or SOX9 & NFIB, combined with small molecule cocktails generated iNeurons and iAstrocytes respectively over 3-6 weeks as confirmed by short- and long-read RNAseq profiling. Transcriptome profiles correlated with respective cell types derived from human pluripotent stem cells. iNeurons and iAstrocytes expressed 193 and 153 SNGs respectively, with a combined 264/516 (51%) SNGs expressed. Expressed SNGs recapitulated expression levels and isoform diversity consistent with adult human cortex.

Patient derived iNeurons were used to investigate disease associated variants in SNGs. Three nonsense PAK3 variants were proven to be regulated by nonsense-mediated mRNA decay (NMD) confirming a loss-of-function (LOF) disease mechanism. An intronic variant in DLG3 (c.840+4A>G) resulted in a 31bp insertion of intronic sequence leading to a premature stop codon, NMD and a 3.8-fold reduction in expression confirming a LOF disease mechanism. The impacts of both PAK3 and DLG3 variants were confirmed using orthologous CRISPRa gene transactivation approaches.

These data support use of transdifferentiation as a robust and systematic method to functionally investigate the mechanisms of disease associated variants impacting SNGs.

Poster #27

ELIXIR Activities Supporting a FAIR Future for Life Science

Freiburg Galaxy Team 1,*

1 – Bioinformatics Group, Department of Computer Science, University of

Freiburg, Freiburg, Germany

ELIXIR, a pan-European life-science infrastructure, coordinates interoperable biological data services and connects national nodes to support research, medicine, and industry. Its vision is to enable FAIR data, robust software, and reusable workflows across the life sciences. Several ELIXIR Communities directly support functional genomics, protein science, computational genomics, and precision medicine, including Single-Cell Omics, Federated Human Data, Rare Diseases, Cancer Data, Proteomics, and 3D-BioInfo. These groups advance standards, tools, and workflows that underpin genomics interpretation, protein function analysis, spatial omics, and translational research. Parallel Galaxy Communities - such as Single-Cell, Microbiology, and Imaging - provide complementary workflows and trainings widely adopted across ELIXIR.

As members of ELIXIR Germany, we, the Freiburg Galaxy team, contribute to multiple ELIXIR initiatives and highlight here a subset of ongoing activities. Within the ELIXIR BioHackathon programme, we developed a suite of FAIR bioimage-analysis training materials and Galaxy workflows. These resources support reproducible extraction of quantitative image-derived phenotypes and support integrative analyses linking spatial context and cellular morphology to functional genomics readouts.

ELIXIR Research Software Ecosystem initiative develops shared design principles, metadata conventions, and community-driven tooling that enhance software quality, sustainability, and discoverability across the ELIXIR and Galaxy ecosystems. These efforts also align with international partners, including the Australian BioCommons, to strengthen globally interoperable life-science data infrastructures.

Poster #28

Implementing MaveDB to identify functionally impactful pharmacogenomic variants in underrepresented Indonesian populations

Limardi PC 1*, Malik SG 2, Sudoyo H 2, Gallego-Romero I 1

1 - St. Vincent's Institute of Medical Research, Melbourne, Australia

2 - Mochtar Riady Institute for Nanotechnology, Tangerang, Banten, Indonesia

Most genetic reference datasets focus on European populations, leaving functionally important variants in underrepresented groups poorly characterised. Deep mutational scanning (DMS) enables large-scale experimental measurement of variant effects, providing a powerful approach to identify such variants.

In this study, we integrated pharmacogene variants from Indonesian cohorts (n ≈ 3,000) with functional scores from MaveDB to explore their potential impact. All identified single-nucelotide polymorphisms (SNPs) were annotated using Ensembl Variant Effect Predictor (VEP) and cross-referenced with MaveDB datasets. Variants were then mapped to functional category (nonsense/nonsense-like, decreased, wildtype/wildtype-like, increased abundance or activity) based on DMS effect scores.

We identified several rare variants unique to Indonesian populations that fell into nonsense, decreased or increased functional scores, suggesting potential functional consequences not reported in current pharmacogenomic databases. Preliminary comparisons with computational predictors (SIFT, PolyPhen-2, REVEL, ClinPred, CADD Score) revealed some discrepancies with DMS results, highlighting the usefulness of experimental data in underrepresented populations.

These findings demonstrate the potential of MaveDB to support functional interpretation of variants in underrepresented populations and emphasise the importance to expand variant effect mapping efforts to diverse global populations.

Poster #29

Functional mapping of paralog synthetic lethal domains using base editing: Bridging knockout screens and therapeutic discovery.

Olvera-León R 1,2,*, Belattar Y 1, Dennler O 3, Heride C 4, Ryan C 3, Tzelepis K 2, Coelho M 1, Adams D 1, Mehta I 1

1 - Wellcome Sanger Institute, Cambridge, UK

2 - Cambridge Stem Cell Institute, Cambridge, UK

3 - Conway Institute, University College Dublin, Dublin, Ireland

4 - Cancer Research Horizons Therapeutic Innovation, Cambridge, UK

Around 70% of human genes have at least one identifiable paralog, reflecting widespread functional redundancy that buffers essential cellular processes. While this redundancy can mask loss-of-function phenotypes, it also presents an opportunity for therapeutic exploitation: as cancers accumulate deletions or transcriptional silencing of specific genes, they may become synthetically lethal (SL) when the remaining paralog is inhibited.

Using multiplexed CRISPR knockout screening based on a novel tRNA dual-guide expression system, we identified more than fifty reproducible SL paralog interactions across diverse cancer contexts and cellular lineages. Each represents a potential dependency in which loss of one paralog sensitises cells to inhibition of its partner. From these, we prioritised pairs such as CDS1/CDS2 and FAM50A/FAM50B, which display strong SL effects and clear clinical relevance, with TCGA data revealing recurrent loss of one paralog at the genomic or transcriptomic level across multiple tumour types.

We are now applying systematic base-editing mutational scanning to map the residues and domains that mediate paralog-specific dependency in isogenic models lacking the biomarker paralog. By inducing precise nucleotide substitutions rather than complete knockouts, base editing enables residue-level dissection of functional architecture and identification of motifs driving SL sensitivity.

Integration of mutational fitness maps with evolutionary conservation, AlphaFold-predicted structures, and ligand annotations will pinpoint regions mediating paralog dependency and highlight selective, druggable sites for therapeutic development. This framework positions base editing as a powerful link between genome-scale knockout screens and compound discovery, translating synthetic lethal interactions into actionable molecular targets.

Poster #30

Community co-designed workshops build confidence in use of functional evidence for variant classification.

Rehan Marie Villani 1, Bronwyn Terrill 2, 3, 4, Emma Tudini 1, Ebony Matotek 5, Tessa Mattiske 5, Corrina C. Cliffe 6, Bryony A. Thompson 7, Laveniya Satgunaseelan 8, Christopher N. Hahn 9, Ben Lundie 10, 11, John Christodoulou 12, 13, Alan F. Rubin 14, 15, Amanda B. Spurdle 1, 11

Affiliations: 1 - QIMR Berghofer, 2 - Australian Genomics, Murdoch Children's Research Institute, Melbourne, 3 - Clinical Translation and Engagement Platform, Garvan Institute of Medical Research, Sydney, 4 - Faculty of Medicine and Health, UNSW Sydney, 5 - Australian Functional Genomics Network, Murdoch Children's Research Institute, 6 - Douglass Hanly Moir Pathology, 7 - Department of Pathology, The Royal Melbourne Hospital, 8 - Department of Neuropathology, Royal Prince Alfred Hospital, 9 - Department of Genetics and Molecular Pathology, SA Pathology, 10- Pathology Queensland, Queensland Health, 11 - University of Queensland, 12 - Genomic Medicine Theme, Murdoch Children's Research Institute, 13 - Department of Paediatrics, University of Melbourne, 14 - The Walter and Eliza Hall Institute of Medical Research, 15 - Department of Medical Biology, University of Melbourne

Experimental data can be used as evidence for variant classification, however the complexity of evaluating assays for application as evidence creates challenges. Our previous results showed that uncertainty around practice recommendations leads to a lack of confidence in evidence evaluation.

We proposed that capability may be improved though improving confidence of those performing functional data evaluation. We used a community of practice discussion and consultation workshop approach to develop and evaluate a solution to developing confidence that builds capability while supporting practice development. A series of workshops were delivered, each with learning content centred on current expert recommendations while surveying participants in their experimental evaluation practice and opinions around evaluating experimental data.

Workshop participants were highly engaged and indicated a desire for further training and resources. Consultation suggested that ambiguity in the current recommendations underpinned differences in interpretation and ultimately resulted in practice variation. Collaborative, case-based discussion was identified as the overall preferred method of building confidence. We delivered case-based community of practice discussion, after which participants reported improved comfort in the evaluation practice.

Through the consultative workshop approach, workshop participants have built confidence but also contributed to informing field recommendation development. We propose that this approach supports capability development for functional evidence evaluation in variant classification, and recommend the overall approach as a means of capability development for evolving fields.

Poster #31

Development of a CRISPR-oligorecombineering (CORe) and imaging platform for high-throughput profiling of phospho-variants in disease

Hassan S.[1,2,*], Das A.[2], Reid K.[1], Storch M.[2], Tate E.W.[1], Child M.A.[1]

[1] Department of Life Sciences, Imperial College London, London, UK

[2] London Biofoundry, Innovation Hub, Imperial College London, London, UK

Phosphorylation is a dynamic and pervasive post-translational modification (PTM) that regulates protein localization, stability, and signaling. Disruptions to phosphosite integrity through missense variants are increasingly linked to disease, yet their mechanistic and phenotypic consequences remain poorly annotated. Here, we describe the development of a high-throughput and modular CRISPR-oligorecombineering (CORe) platform designed to functionally annotate phosphosite-associated variants in human cells. Using CORe, we introduce defined phospho-status controls across 11 well-characterized sites implicated in signaling and organelle organization. These serve as positive controls to benchmark phenotypic profiles associated with site-specific phosphorylation changes. We integrate high-content imaging via CellPaint to capture multi-parametric morphological and organelle-level perturbations induced by variant introduction, enabling quantification of subtle phenotypic signatures beyond cellular fitness. By coupling this imaging-based phenotyping with phospho-enrichment proteomics, we aim to define the morphological and biochemical fingerprints of disrupted phosphorylation networks. This pipeline establishes a scalable framework for systematic phosphosite function mapping and provides a foundation for high-throughput variant interpretation in the context of dynamic PTM regulation.

Poster #32

Disease modelling and variant interpretation in the inherited retinal dystrophies using human pluripotent stem cells.

Loi T 1,*, Nash BN 1,2,3, Aryamanesh N 4, Grigg JR 1,5, Yang P 6, Gonzalez-Cordero A 7, Jamieson RV 1,3,8

1 - Eye Genetics Research Unit, Children's Medical Research Institute, Save Sight Institute, Sydney Children's Hospitals Network, University of Sydney, New South Wales, Australia
2 - Sydney Genome Diagnostics, Western Sydney Genetics Program, Sydney Children's Hospitals Network, New South Wales, Australia
3 - Specialty of Genomic Medicine, Faculty of Medicine and Health, University of Sydney, New South Wales, Australia
4 - Bioinformatics Facility, Children's Medical Research Institute, University of Sydney, Westmead, New South Wales, Australia.
5 - Department of Ophthalmology, Sydney Children's Hospitals Network, New South Wales, Australia
6 - Computational Systems Biology Unit, Children's Medical Research Institute, The University of Sydney, New South Wales, Australia
7 - Stem Cell Medicine Group and Stem Cell and Organoid Facility, Children's Medical Research Institute, University of Sydney, New South Wales, Australia
8 - Department of Clinical Genetics, Western Sydney Genetics Program, Sydney Children's Hospitals Network, New South Wales, Australia

Inherited retinal diseases (IRDs) affect 1:1000 people, causing aberrant photoreceptor function and vision loss due to variants in >250 disease genes. Genomic technologies provide genetic diagnosis in ~50-60% of cases but also identify variants of uncertain significance (VUS) in 20-30% of cases, hindering appropriate disease management and access to specific genetic therapies. Here, we aim to examine the utility of human induced pluripotent stem cells (iPSCs) differentiated to retinal organoids for VUS pathogenicity determination.

We investigated the IRD gene RPGRIP1, impacted by >1000 VUS. We created patient-derived iPSC lines with known pathogenic variants and normal iPSCs with CRISPR/Cas9 genomic introduction of novel VUS for testing. All iPSC lines were then differentiated to retinal organoids and examined for biomarkers by immunohistochemistry and transcriptomics. For validation, pathogenic retinal organoids were transduced with AAV containing normal copies of RPGRIP1.

In both pathogenic and VUS retinal organoids, we saw diminished or mislocalised expression of RPGRIP1 along with abnormal photoreceptor cell staining and localisation compared to controls. Transcriptomic analysis found expression of photoreceptor related genes lower in the variant organoids than controls. Furthermore, gene signatures associated with disease were similarly enriched in both forms of variant organoids providing further evidence of VUS pathogenicity. Post AAV therapy, pathogenic variant organoids displayed rescued biomarker features including improved expression of RPGRIP1 and photoreceptor morphology.

Overall, the use of human iPSCs is highly beneficial for disease modelling and establishing key biomarkers for RPGRIP1-IRD to further examine other RPGRIP1 variants at scale.

Poster #33

High-Throughput DNA Synthesis Enables Large-Scale Mapping of Protein Sequence–Function Relationships

Xiang'er Jiang

I will be presenting my primary work published in 2025: https://www.nature.com/articles/s41586-024-08370-4 https://www.nature.com/articles/s41587-025-02844-0 High-throughput DNA synthesis and comprehensive protein variant libraries are essential for systematically mapping sequence–function relationships. We have developed a massively parallel DNA synthesis platform that significantly increases oligonucleotide yield and simplifies downstream gene assembly, enabling the rapid construction of large, high-complexity DNA libraries. Leveraging this platform, we generated deep mutational scanning libraries covering over 120 million protein variants across hundreds of human protein domains. These libraries allow precise quantification of the effects of individual amino acid substitutions on protein stability and function, providing a systematic resource for protein engineering and functional genomics. Our approach integrates optimized synthesis chemistry, automated library assembly, and quality control measures to ensure reproducibility and scalability. By enabling large-scale, high-resolution mapping of protein sequence–function space, these technologies support applications ranging from variant interpretation in clinical genetics to the design of novel proteins with tailored properties. Herein, discussion of methodological advances, practical challenges in library construction, and strategies for integrating high-throughput synthesis with functional screening will be presented.

Poster #34

Mapping NTD of Spike protein utilising Deep Mutational Scanning to predict the unfolding evolution of SARS-CoV-2

Yastika Banerjee 1, Rhys Parry 1, Helen Mostafavi 1, Alexander Khromykh 1

1School of Chemistry and Molecular Biosciences, The University of Queensland, St Lucia, Queensland, Australia

This study investigates the N-terminal domain (NTD) of the SARS-CoV-2 Spike protein, which has evolved to evade humoral immunity. Using Deep Mutational Scanning (DMS), we aim to construct a comprehensive viral library on the ancestral SARS-CoV-2QLD02 strain isolated in Queensland, containing mutations across the NTD. A diverse mutagenized plasmid library (DMS-NTD) was generated using tiling primers with NNK/MNN nucleotides by mutating each NTD codon. The plasmid library was validated through Sanger sequencing of individual clones, Nanopore, and Illumina sequencing. Sanger analysis shows an average of 2.75 mutations per clone with a potential amino acid representation of 88.5%. We successfully recovered both SARS-CoV-2QLD02 and SARS-CoV-2QLD02-DMS-NTD harbouring the DMS-NTD library using Circular Polymerase Extension Reaction (CPER). Future work entails screening the mutant library against a panel of sera from vaccinated and infected individuals to identify immune escape mutations, followed by deep sequencing to provide insights into early viral evolution and inform next-generation vaccine antigen design.

Poster #35

Allosteric regulation of KRAS

Yujing Ma1,2*, Qile Jin1,2, Ben Lehner3,4,5,6, Chenchun Weng1,2

1 - Department of Pathology, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui, 230036, China

2 - Intelligent Pathology Institute, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui, 230036, China

3 - Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Barcelona, Spain

4 - Universitat Pompeu Fabra (UPF), Barcelona, Spain

5 - Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

6 - Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, United Kingdom

Allosteric communication in proteins underlies biological regulation, variant pathogenicity, and the efficacy of many drugs. Maps of allosteric regulation have been quantified for individual protein functions. However, most proteins are multi-functional, for example binding multiple interaction partners through different binding interfaces. It is not clear how allostery operates for spatially distinct binding sites in the same protein. Here we use the KRAS oncoprotein as a model system to address this question. By quantifying the energetic effects of mutations on binding to eight proteins — a total of over 30,000 free energy measurements — we find that allosteric control of structurally distinct interfaces is largely orthogonal. Each interface has a unique allosteric network defined by its structural location. The protein therefore has a modular allosteric architecture, with different core and surface residues controlling binding through different interfaces. Only where the distance-dependent allosteric networks intersect is regulation of both interfaces possible. We believe the existence of this modular allosteric architecture has important implications for biological regulation, evolution, and drug discovery.

1.9 Online Posters

Poster #Online 1

How do we define 'truthsets' for clinical validation of functional assays?

Allen S 1,*, Rowlands C F 1, Turnbull C 1,2

1 - Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK

2 - The Royal Marsden NHS Foundation Trust, London, UK

To improve consistency in clinical assay validation, in 2020 the ClinGen Sequence Variant Interpretation group published a framework for assessing assay robustness, which included using a 'truthset' of previously classified pathogenic and benign variants. However, there remains wide variation in approaches for defining such truthsets, and in the method by which evidence strength is assigned (either at an assay-level, where all variants are assigned the same evidence strength, or at a variant-level, where each variant attains a specific evidence strength using mixture-model based methods).

We have explored the variety in approaches to clinical validation for several large-scale functional assays, identifying up to 172 plausible approaches to defining variant truthsets using both amalgamated ClinVar classifications and systematically-classified variants. For each approach, we have examined the influence on evidence strength for assay-level scoring using the Brnich et al 2020 approach, and for variant-level scoring using Gaussian Mixture Modelling.

Through these examples, we demonstrate the impact on applicable evidence strength when using a truthset of nonsense/synonymous variants versus missense variants only, and the trade-off in power when curating more stringent ClinVar truthsets (such as restricting to classifications of 2 stars or above). We also explore feasibility of variant-level evidence assignment without the use of author-defined function thresholds, and show that variants may attain evidence beyond a single assay-level strength. We anticipate that these results will inform future guidance and approaches for standardisation of clinical assay validation under the remit of the ACMG v4.0 framework.
MAVE Progress Report, How Ambry Genetics Validates and Deploys High-Throughput Functional Assays

Poster #Online 2

A systematic approach benign missense variant "truthsets" for validation of functional assays

Charlie F. Rowlands 1*, Sophie Allen 1, Alice Garrett 1,2, Miranda Durkie 3, George J. Burghel 4,5, Rachel Robinson 6, Alison Callaway 7, Joanne Field 8, Bethan Frugtniet 2, Sheila. Palmer-Smith 9, Jonathan Grant 10, Judith Pagan 11, Trudi McDevitt 12, Katie Snape 2, Helen Hanson 13,14, Terri McVeigh 15, Clare Turnbull 1,15 and CanVIG-UK

1 - Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK.
2 - St George's University Hospitals NHS Foundation Trust, Tooting, London, UK
3 - Sheffield Diagnostic Genetics Service, NEY Genomic Laboratory Hub, Sheffield Children's NHS Foundation Trust, Sheffield, UK
4 - Manchester Centre for Genomic Medicine and NW Laboratory Genetics Hub, Manchester University Hospitals NHS Foundation Trust, Manchester, UK
5 - Division of Cancer Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK
6 - Yorkshire Regional Genetics Service, Leeds Teaching Hospitals NHS Trust, Leeds, UK
7 - Central and South Genomics Laboratory Hub, Wessex Genomics Laboratory Service, University Hospital Southampton NHS Foundation Trust, Salisbury, UK
8 - Genomics and Molecular Medicine Service, Nottingham University Hospitals NHS Trust, Nottingham, UK
9 - Wales Genomic Health Centre, Cardiff and Vale University Health Board, Cardiff, UK
10 - Laboratory Genetics, Queen Elizabeth University Hospital, NHS Greater Glasgow and Clyde, Glasgow, UK
11 - South East Scotland Clinical Genetics, Western General Hospital, Edinburgh, UK.
12 - Department of Clinical Genetics, CHI at Crumlin, Dublin, Ireland
13 - Peninsula Regional Genetics Service, Royal Devon University Healthcare NHS Foundation Trust, Exeter, UK
14 - Department of Clinical and Biomedical Sciences, University of Exeter Medical School, Exeter, United Kingdom
15 - The Royal Marsden NHS Foundation Trust, Fulham Road, London, UK

BACKGROUND

Effective adoption of MAVEs and other functional assays requires robust quantification of their applicable evidence strength. Bias, circularity and the limited availability of truthset variants remain substantive obstacles to quantification of MAVE strength for many genes.

AIMS

We aimed to quantify the extent to which available truthsets of missense variants would enable robust clinical validation of forthcoming MAVES, and sought to develop new approaches to improve the lack of available clinically classified benign missense variants against which to validate MAVE strength.

RESULTS

We present quantitative analyses showing that scarcity of available high-confidence ClinVar classifications for 116 cancer susceptibility genes will limit our ability to clinically validate new MAVES for most of these genes.

We present an approach for systematic generation of benign missense variant sets for validation of assays. We demonstrate that, for eight hereditary breast and ovarian cancer genes, validation variant sets generated via our approach would allow greater strength of evidence for pathogenicity (PS3) to be applied compared to using existing sets of ClinVar classifications.

CONCLUSIONS

We demonstrate an unmet need regarding availability of missense variants for MAVE validation. We present a systematic approach to generation of benign missense variants allowing improvement in the applicable evidence strength for MAVEs, even in genes with scarce existing classifications. Improving truthsets for clinical validation of MAVEs is key to unlocking the value of these resources to improve our ability to classify variants as benign or pathogenic (and out of the VUS status assigned to so many rare missense variants).

Saturation genome editing of CTCF resolves variants of uncertain significance and maps domain-specific functional constraint

Poster #Online 3

Multiplexed assays of variant effects enhance gene discovery in de novo gene burden testing

Daniel A. Jaramillo-Calle[1,2], Elizabeth Radford[1,3], Sarah Lindsay[1], David Adams[1], Andrew Waters[1], Sebastian Gerety[1], Matthew Hurles[1]

1 Wellcome Sanger Institute, Hinxton, UK

2 School of Biological Sciences, University of Cambridge, Cambridge, UK

3 Department of Paediatrics, University of Cambridge, Cambridge, UK

Large-scale sequencing studies use de novo gene burden testing (dnGBT) to discover disease genes. However, associations remain undetected despite adequate sample sizes, partly due to limitations of computational variant effect predictors for variant prioritisation. Experimental measurements of variant effects can outperform computational predictors, but their sparse availability has restricted use to variants observed in case–control cohorts. Multiplexed assays of variant effect (MAVEs) now provide near-saturation functional maps, allowing variant-level weights for almost all variants in a gene, as required for dnGBT to generate unbiased null distributions and expected mutation counts.

We present W_MAVE, a functional weighting scheme for dnGBT that integrates saturation genome editing (SGE) data. Using BAP1, BRCA1, and DDX3X datasets, we transformed experimental scores into probabilistic weights via logistic regression and evaluated statistical power against DeNovoWEST, BayesDel, and AlphaMissense.

W_MAVE outperformed other methods for missense variants, uniquely identifying a BAP1–neurodevelopmental disorder association in 31,058 trios that competing approaches missed and reducing sample size requirements for detecting known associations. BayesDel performed better when protein-truncating variants were included, reflecting uniformly high predictions compared with more variable experimental scores. Genome-wide type I error remained conservative.

By leveraging functional maps, W_MAVE enables functional weighting in dnGBT to uncover missed associations in existing cohorts when the assay readout captures the relevant disease mechanism. Because genes are often pleiotropic, W_MAVE can repurpose maps generated in one disease context (e.g., cancer) to power gene discovery in others, such as neurodevelopmental disorders.

Poster #Online 4

Saturation genome editing of CTCF resolves variants of uncertain significance and maps domain-specific functional constraint

Daniel A. Jaramillo-Calle[1,2], Elizabeth Radford[1,3], Malin H. L. Andersson[1], Sarah Lindsay[1], Petr Daněček[1], Hong Kee Tan[1], Andrew Waters[1], David Adams[1], Silvia Casadei[4], Nahum T Smith[4], Sebastian Gerety[1], Lea Starita[4], Matthew Hurles[1]

1 Wellcome Sanger Institute, Hinxton, UK

2 School of Biological Sciences, University of Cambridge, Cambridge, UK

3 Department of Paediatrics, University of Cambridge, Cambridge, UK

4 Department of Genome Sciences and Brotman Baty Institute for Precision Medicine, University of Washington, Seattle, WA, USA

CTCF is a 11-zinc finger transcription factor and architectural protein that organises 3D chromatin and regulates gene expression. Germline variants cause an autosomal dominant neurodevelopmental disorder and somatic mutations are recurrently observed across diverse cancers. However, most missense variants in clinical and tumour sequencing are classified as variants of uncertain significance (VUS), limiting diagnosis, risk estimation, and counselling.

We performed saturation genome editing (SGE) in HAP1-A5 cells with LIG4 knockout and stable Cas9 expression, targeting CTCF exons 3–12 and ~15 bp of flanking intronic sequence. This yielded functional scores for >16,000 variants with near-complete single-nucleotide substitution coverage, efficient editing, high replicate concordance and clear separation of nonsense and synonymous variants.

Despite haploinsufficiency, over two-thirds of variants were functionally normal, with the remainder showing strong or intermediate depletion. CTCF showed marked regional heterogeneity in mutational tolerance: residue-level analyses showed most functionally abnormal variants clustered within zinc fingers and the YDF motif, whereas N- and C-terminal regions were more tolerant. Homologous-residue analyses based on zinc finger multiple sequence alignments revealed intolerance at DNA-contact and zinc-coordinating positions. Functional scores showed high concordance with low-throughput assays and perfect specificity and ~85% sensitivity for clinical classifications.

Applying ACMG/AMP guidelines with calibrated PS3/BS3 evidence, SGE data enabled proactive classification of variants and definitive reclassification of over half of CTCF VUS, providing immediate clinical utility.

Poster #Online 5

nf-core/deepmutscan: a community-oriented, modular bioinformatics pipeline for deep mutational scanning by shotgun sequencing

Benjamin Wehnert[1,*], Taylor L. Mighell[1], Fei Sang[2], Ben Lehner[1,2,3,4], Maximilian R. Stammnitz[1]

1 - Centre for Genomic Regulation (CRG), Barcelona, Spain

2 - Wellcome Sanger Institute, Cambridge, UK

3 - Universitat Pompeu Fabra (UPF), Barcelona, Spain

4 - Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Deep mutational scanning of long open reading frames (ORFs) remains technically challenging, despite major advances in DNA synthesis and sequencing. For long ORF targets, most studies still rely on molecular barcoding or tiled libraries of programmed mutants – approaches that are labor-intensive, error-prone, and difficult to scale. An attractive but underused alternative is to quantify variants directly from unbiased fragment (“shotgun”) short-read sequencing. Shotgun DNA sequencing can be used to greatly speed up the inference of long ORF mutant fitness landscapes, theoretically at no expense in accuracy.

We present nf-core/deepmutscan, a Nextflow-based computational workflow that unlocks shotgun-sequencing based deep mutational scanning of long ORFs. The pipeline is fully containerized, easy to deploy and highly modular. It currently implements (1) raw shotgun sequencing QC, (2) read mapping and filtering, (3) variant quantification, (4) mutant library coverage assessments, (5) count error correction, and (6) fitness estimation. All components are integrated into the nf-core bioinformatics ecosystem, thereby enabling continuous, community-driven maintenance and extension of the workflow architecture.

By providing a standardized, reproducible, and extensible framework for shotgun-based analysis, we believe that nf-core/deepmutscan lowers the barrier to long-ORF deep mutational scanning and promotes the broader adoption of best-practices genomics workflows in our community (https://nf-co.re/deepmutscan/).

Poster #Online 6

Precision missense mutagenesis for deciphering the functional impact of disease-associated MED27 variants.

Ione C. Goodwin 1*, Kimberly M. Reid 1, Matthew A. Child 1.

1 - Imperial College London

Recent advances in genomics and sequencing technologies have increased the diagnostic yield within the field of rare neurodevelopmental disorders. Despite this, a large proportion of patients remain undiagnosed, and diseases lack defined molecular mechanism due to the paucity of functional investigation of VUSs. One example of this is MED27-related disease – a complex neurodevelopmental disorder caused by variants in the MED27 gene. MED27 encodes subunit 27 of the Mediator (MED) complex – a large protein complex that facilitates RNA polymerase II-mediated DNA transcription. Despite multiple variants being reported in 57 known patients, the impact of these variants on MED27 function remains undetermined, leaving gaps in our understanding of disease mechanism, and patients far from therapeutic intervention.

We are functionally interrogating variant impact on MED27 function using our CRISPR-based oligo recombineering (CORe) platform. CORe uses an HDR-mediated targeted missense mutagenesis strategy to introduce variants associated with clinically relevant mutations. A total of 11 likely-pathogenic and 12 VUS missense variants were identified from publications or ClinVar, respectively. Three population missense variants were selected from Gnomad as likely-benign 'negative' controls. Variants were introduced into HAP1 cells and their impact on overall fitness was assessed. Further biochemical assays were performed in HEK293 cells to investigate the impact of triaged variants on the MED interactome, and a MED27 knock-down was generated in the physiologically relevant SH-SY5Y neuronal cell line for further validation studies. Combined, these data provide mechanistic evidence for the function of disease-associated MED27 variants and permits VUS prioritisation for future orthogonal investigations into disease-relevant molecular phenotypes.

Poster #Online 7

Allosteric mapping for the rapid discovery of novel therapeutics

Mireia Seuma*, Sol Adaglio, Pablo Baeza, Grzegorz Chojnowski, Júlia Domingo, André J Faure, Lukas Lim, Cristina Lorca, Youssef Mansour, Andrei Papkou

Authors are listed in alphabetical order and are all affiliated with ALLOX

ALLOX is a biotech spin-off from the Lehner Lab at the Centre for Genomic Regulation (CRG). We've developed a proprietary platform that combines systematic mutagenesis, high-throughput phenotyping and biophysical modelling to systematically identify allosteric switches and modulatory sites at scale. ALLOX is advancing towards its goal to rapidly discover novel medicines targeting therapeutically relevant protein-protein interactions (PPI).

We will present the latest discoveries and improvements of our hybrid experimental and computational platform, highlighting its growing ability to address challenging and previously inaccessible protein systems. In particular, we will share results from our first complete allosteric map of the transcription factor TEAD1 and its disordered co-activator YAP1. This work expands the scope of our technology beyond globular-globular interactions and demonstrates its power to reveal both orthosteric and allosteric druggable pockets in complex protein interfaces.

In addition, we will showcase progress in our internal drug discovery pipeline, where deep mutational scanning (DMS) data guides multiple hit identification strategies and informs the design of novel molecules targeting clinically relevant proteins.

Together, these advances underscore the potential of DMS-based technologies like ALLOX's to transform how molecules are designed, identified, and optimised for therapeutic benefit.

Poster #Online 8

Multidimensional deep mutational scanning of the ALS protein SOD1: dissecting the effect of mutations on protein abundance and dimer formation

Quiroga T 1*, Martin M 1, Ashcroft L 2, Rice L 2, Boratav D 3, Thompson B 3, McAlary L 2, Bolognesi B 1

1 - Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Techonology (BIST), Spain

2 - Molecular Horizons and School of Science, Faculty of Science, Medicine and Health, Univerisy of Wollongong, Australia

3 - Department of Pathology, Royal Melbourne Hospital, Australia

Mutations in SOD1 cause Amyotrophic Lateral Sclerosis (ALS) through diverse mechanisms, including monomer destabilization, dimer dissociation, aggregation, and neurotoxicity. However, quantitative evidence linking each phenotype to pathogenicity is available only for few mutants. To systematically evaluate mutational impacts in SOD1, we performed a multidimensional deep mutational scan measuring the effect of ~6000 mutations on protein abundance and dimer formation. Our dual-phenotype atlas includes amino acid substitutions, insertions, and deletions, thus reporting both side-chain and backbone perturbations. Our results show that amino acid substitutions decreasing abundance are enriched in buried β-strands residues, while insertions and deletions are overall disruptive in β-strands. Shorter loops are generally more affected, whereas longer metal-binding loops are more tolerant, except at zinc-binding sites. Abundance and binding scores are strongly correlated, revealing interdependence between monomer abundance and dimerization. Outlier variants deviating from this trend mainly map to the dimer interface, indicating specific disruption of dimer formation without broadly affecting monomer abundance. Our dataset includes quantitative scores for all known pathogenic and uncertain significance variants (VUS): 88% of pathogenic, 75% of VUS, and 77% of unreported variants reduce protein abundance. Integrating insertions and deletions provides structural information about backbone flexibility that, combined with the dual readout of abundance and binding, reveals mechanistic features. This multidimensional phenotyping also identifies allosteric sites that affect dimerization while maintaining monomer abundance, revealing potential new druggable regions to stabilize the SOD1 dimer and prevent monomerization and aggregation. Overall, our work provides a comprehensive mechanistic atlas of SOD1 mutational effects that advances our understanding of SOD1-ALS pathogenesis, offering a potential source for variant clinical variant interpretation and drug discovery in ALS.

1.10 Planning committee

Contact Information:

Mutational Scanning Symposium 2026 planning committee

A/Prof Irene Gallego Romero; e irene.gallego@svi.edu.au

Dr Alan Rubin; e alan.rubin@unimelb.edu.au

Dr Rehan Villani; e rehan.villani@qimrb.edu.au

Lara Muffley; e muffley@uw.edu