ADVANCING CELL AND GENE THERAPY RESEARCH WITH BIOINFORMATICS

June 2024

fios GENOMICS

fios GENOMICS

CONTENTS

INTRODUCTION

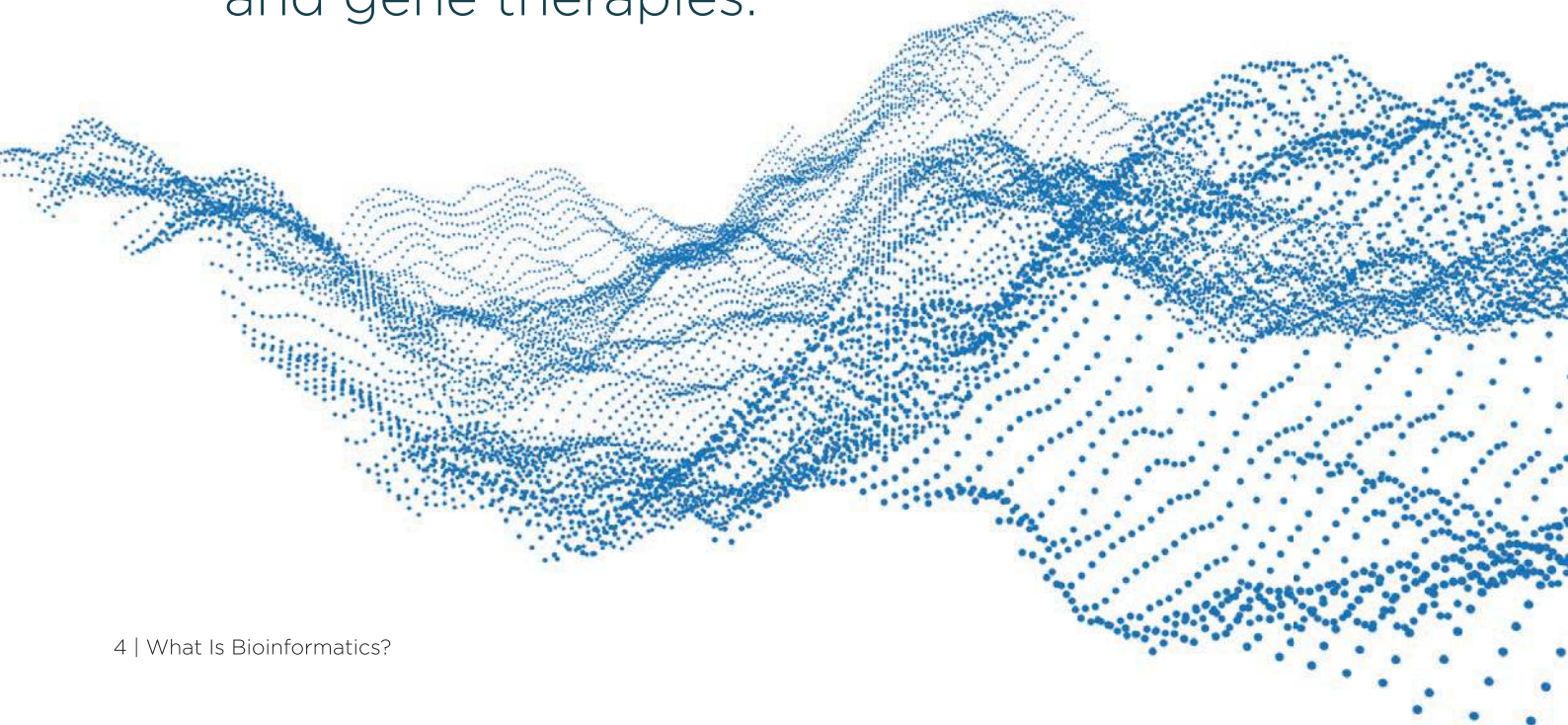
Cell and Gene therapies are cutting-edge approaches to treating and preventing disease. Cell therapies work by transferring certain cells or cell types into a person to treat or prevent a specific disease. Meanwhile, gene therapies do this by placing (or replacing) genetic material (a particular gene for example) into a person's cells to "over-ride" a defective process.

The research and development of cell and gene therapies requires the use of large-scale biological data. Vast quantities of biological data are analysed to further understanding of biological processes, identify potential therapeutic targets, optimise treatment strategies, and improve patient outcomes.

In order to analyse large-scale biological data for these insights, bioinformatics approaches are required. This whitepaper will detail the key ways bioinformatics advance cell and gene therapy research throughout the discovery and development process.

WHAT IS BIOINFORMATICS?

Bioinformatics refers to the use of computational and statistical analyses to extract meaningful information from biological data. It is an interdisciplinary field comprised of biology, computer science and statistics. In cell and gene therapy research, bioinformatics approaches are used to pinpoint the relevant information from large-scale biological data sets that will inform the development of cell and gene therapies.



APPLICATIONS OF BIOINFORMATICS IN CELL AND GENE THERAPY

Data Mining

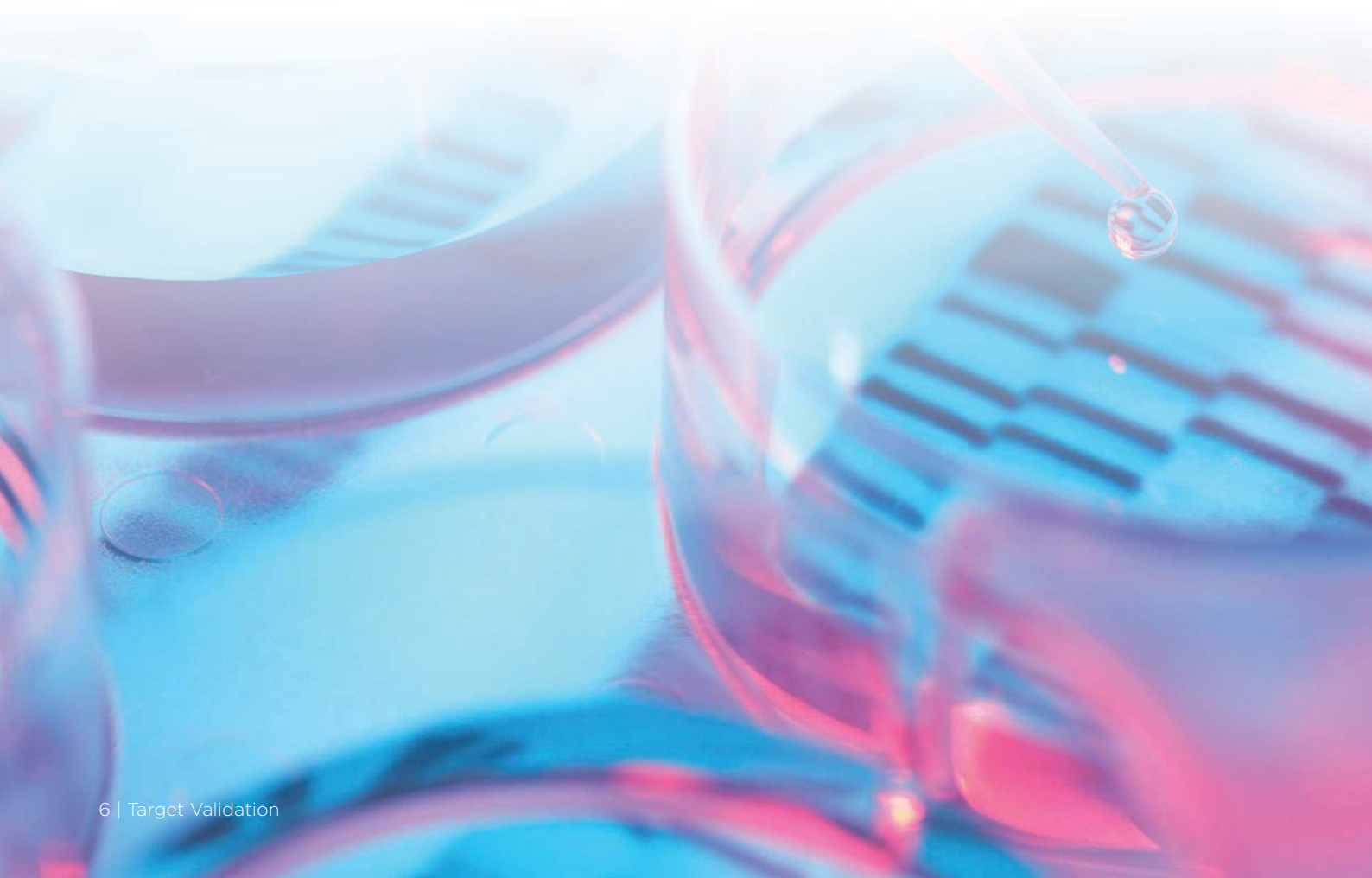
Bioinformatics approaches can be used to mine large-scale biological datasets for key insights which can be used to further cell and gene therapy research. Datasets can be mined for a variety of purposes including, to identify drug targets, discover biomarkers associated with response and reveal biological pathways of interest.

Target Identification

Target identification involves identifying a molecule in the body which is associated with the disease to be treated by a cell or gene therapy. Bioinformatics supports target identification by analysing biological data (such as that gathered from high-throughput screens, preclinical experiments on animal models, diseased human tissue and cell lines) to reveal potential drug targets. It is important that downstream biological pathway analysis is included as part of this to ensure the target is relevant in the pathology of the disease in question.

Target Validation

Once a drug target is identified, it needs to be investigated to ensure that engaging it will have a potential therapeutic outcome either by alleviating symptoms or impacting disease progression. This is target validation. Bioinformatics helps validate drug targets for cell and gene therapies by analysing various data sets (including those from CRISPR screens, knock-out models and gene expression studies) to reveal a potential therapeutic benefit from engaging the target, achieving target validation. Additionally, bioinformatics approaches can further validate targets by enabling in silico comparisons of the analysis results with publicly available data.



Variant Identification

Cell and gene therapies often involve modifying the genetic material of cells to achieve a therapeutic effect. To do this, it is first necessary to identify genetic variations or mutations that are associated with the disease being targeted. Bioinformatics supports variant identification in a number of ways:

Analysis of sequencing data to identify gene variants present in the diseased population of interest

Sequencing technologies are used to 'read' the DNA of an individual's cells. This results in the creation of large-scale biological data sets. Bioinformatics can be applied to these data sets to identify the genetic variant or variants that are associated with a particular disease or phenotype. These variants may be unique to an individual.

Bioinformatics analysis of DNA sequencing data can identify variants such as single nucleotide polymorphisms (SNPs), insertions or deletions of DNA and structural variations in the DNA. This is known as variant calling.

Quality control

Bioinformatics tools also help with the quality control of DNA sequencing data by filtering out low-quality variants as well as artifacts introduced during the sequencing process. Such quality control measures ensure the accuracy and reliability of variant calls.

Annotation and interpretation of gene variants

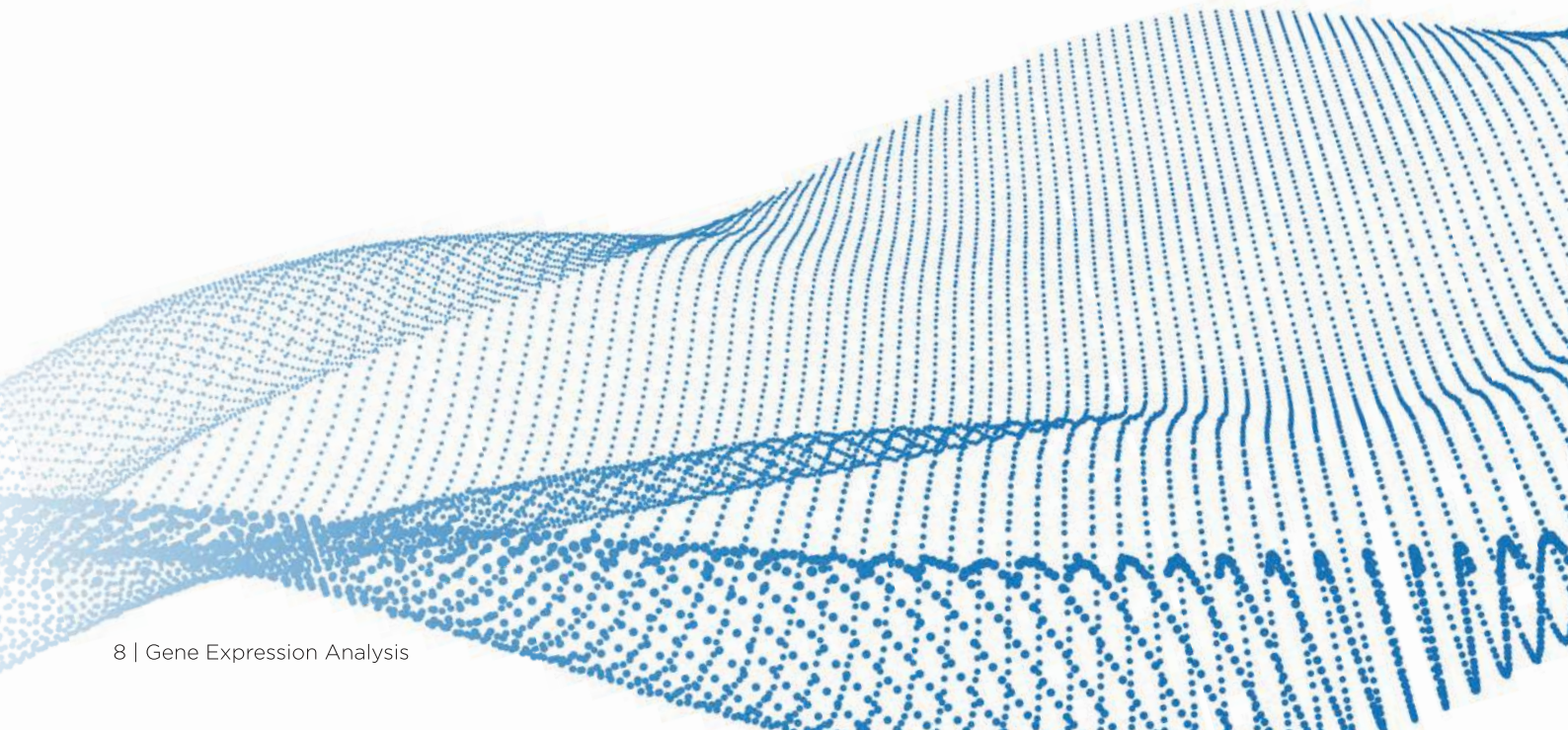
After gene variants are identified, bioinformatics tools can provide annotations that describe their potential consequences by investigating the location and associated biology of the gene the variant is present in. This information is essential for understanding the impact of genetic changes on the therapeutic properties of any modified cells.

Gene Expression Analysis

Scientists researching cell and gene therapies need information about the patterns and levels of gene activity in cells, which can be found in gene expression data. They need this information to help develop the therapies, as well as to evaluate their impact. Bioinformatics analysis of gene expression data is essential to enable it to be used for these purposes.

Bioinformatics approaches assist with:

- Processing raw gene expression data
- Assessing the quality of the data, by identifying and removing artifacts or errors introduced during the experimental phase or during sample processing
- Normalisation of the data for technical variations (including batch correction)
- Identifying differentially expressed genes via comparison of gene expression profiles between different conditions, such as treated vs untreated cells, to identify significantly upregulated or downregulated genes
- Statistical analyses to investigate the significance of observed expression changes and control for false positives
- Gene ontology (GO) analysis to provide insights into the biological processes, molecular functions, and cellular components affected by the gene expression changes



- Pathway analysis to identify any biological pathways overrepresented among the differentially expressed genes, which helps with understanding broader functional implications of changes in gene expression

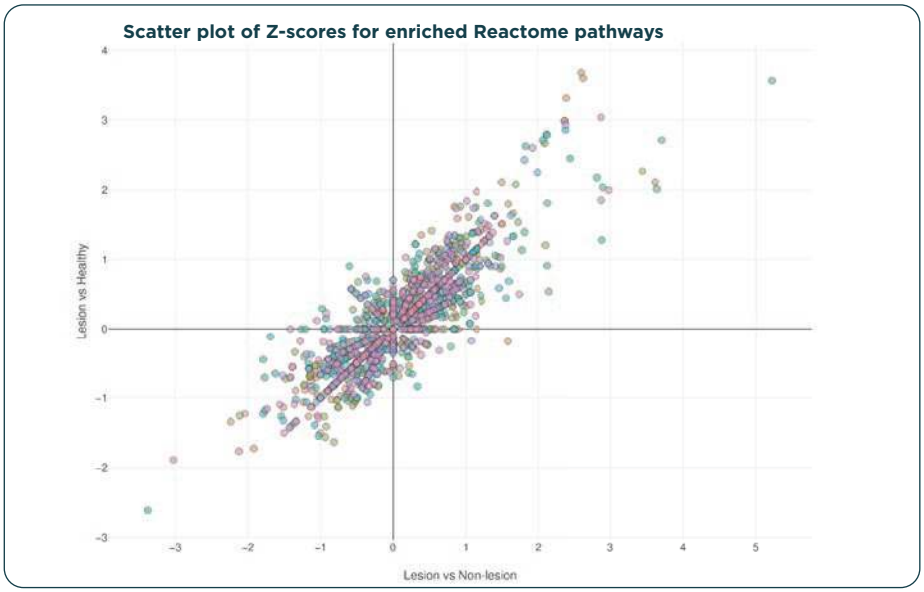
- Cluster analysis to group together genes or samples that have similar expression patterns. This helps to identify co-regulated genes as well as subgroups within the gene expression data

- Classifying modified cells by their gene expression profiles to predict their therapeutic efficacy, using machine learning

- Creating visualisations of gene expression data to aid the interpretation of complex patterns within it. Using, for example, scatter plots, pathway diagrams and heatmaps

- Integrating gene expression data with genomic data to identify correlations between gene expression changes and genetic variants in modified cells

Figure 2:
Pairwise Z-scores are plotted for significantly enriched Reactome pathways.



Source: Fios Genomics' data analysis reports (public) - Analysis of RNA sequencing data generated from psoriatic and healthy skin. This report profiles candidate antisense oligonucleotides (ASOs) for 7 transcripts originating from the LGALS1 gene locus.

Sequence Design

A key component of cell and gene therapies is sequence design. This refers to designing the nucleotide sequence of therapeutic genetic materials, such as genes, with the aim of making the therapy as safe and effective as possible.

Bioinformatics assists with sequence design via:

- Codon optimisation, via analysis of the target cell or organism's codon usage preferences so that the gene sequence used in the therapy aligns with the host's codon preferences
- Predicting the secondary structure of the gene sequence. This helps to identify and minimise RNA structural elements that could affect translation or stability, preventing them from impeding gene expression
- Analysing the GC content of gene sequences to optimise them for efficient transcription and stability
- Facilitating the avoidance of CpG islands. Bioinformatics tools help in designing gene sequences that have reduced CpG content, to enhance safety and minimise immunogenicity
- Predicting RNA splicing via analysis of potential splice sites and splicing events. This enables the design of gene sequences aligned with desired splicing patterns, to ensure proper processing and translation of mRNAs
- Assisting with promoter selection by analysing genomic data to find and characterise regulatory elements and assess promoter strength and specificity
- Analysing epigenetic marks and chromatin structure in target cells. This assists in designing gene sequences that are compatible with the epigenetic landscape, enhancing the chances of successful gene expression
- Enabling *in silico* cloning by simulating the cloning process. This can predict potential issues, (related to sequence compatibility or restriction site availability for example) to ensure gene expression vectors are constructed as efficiently as possible
- Predicting potential immunogenic epitopes within the gene sequence to help minimise immune responses, contributing to the safety of gene therapy
- Minimising non-specific binding potential, to reduce the chance of off-target effects *in vivo*

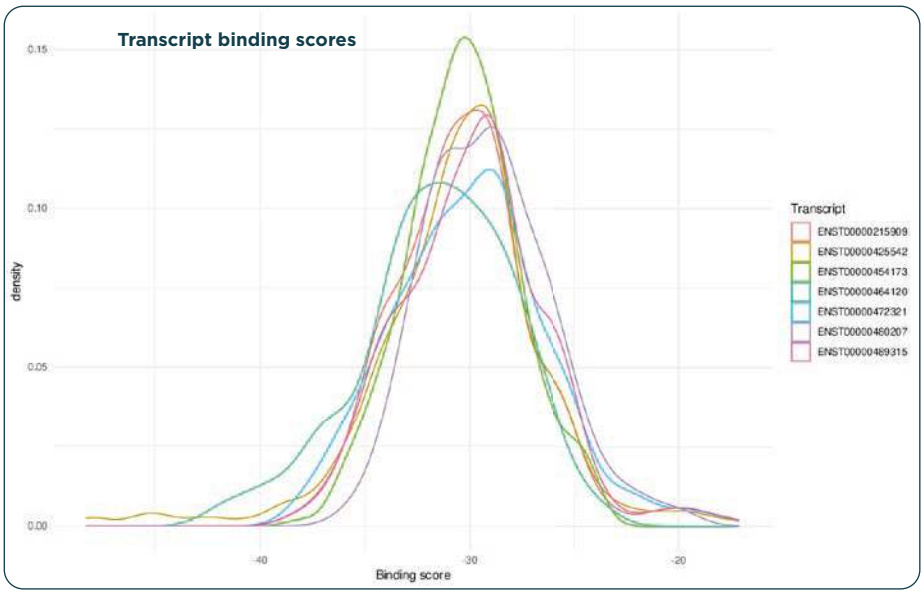
Antisense Oligonucleotide (ASO) Research

ASOs can be used and cell in gene therapy research to both down- and up-regulate the expression of genes associated with diseases, as well modulate splicing to correct associated genetic abnormalities.

Bioinformatics supports ASO research by:

- Assisting with ASO target identification via the analysis of genomic and transcriptomic data
- Predicting suitable candidate ASOs via target sequence design, optimising ASO length, melting temperature, GC content, potential toxicity, non-specific binding potential, and target binding affinity.

Figure 1:
Scores were calculated as the thermodynamic binding affinity of an ASO for its target region. This was quantified using IntaRNA to incorporate both the ASO-target interaction as well as the target region's accessibility. A more negative score can be interpreted as an increase in binding affinity. Binding score is presented individually for each transcript.



Source: Fios Genomics' data analysis reports (public) - Antisense oligonucleotide design
This report profiles candidate antisense oligonucleotides (ASOs) for 7 transcripts originating from the LGALS1 gene locus.

Antigen Transduction

Antigen transduction in cell and gene therapy involves introducing genetic material, that carries instructions for making specific antigens, into target cells. The aim is to stimulate an immune response against the encoded antigens, potentially enhancing the immune system's ability to recognise and eliminate targeted cells or pathogens.

Bioinformatics supports antigen transduction via:

- Analysis of genomic and proteomic data to find potential antigens associated with diseases, which helps with antigen selection
- The use of computation methods to predict epitopes within antigens that are likely to elicit immune responses
- Optimisation of codon usage within DNA or RNA sequences to enhance the efficiency of antigen expression in target cells
- Providing computational tools that assist in the optimisation of vector genomes for efficient antigen transduction
- Analysis of gene expression patterns in target cells to help ensure the introduced genetic material will be properly translated and transcribed. In turn, this enables proper antigen production
- Analysis of off-target effects to ensure the transduction process specifically targets intended cells, minimising risks to non-target tissues
- The use of computational methods to study the interaction between antigens and the immune system. This assists the design of constructs that produce robust and targeted immune responses

T Cell Expansion

T cell expansion is used in cell and gene therapy research to boost immune response for therapeutic purposes.

Bioinformatics supports T cell expansion by:

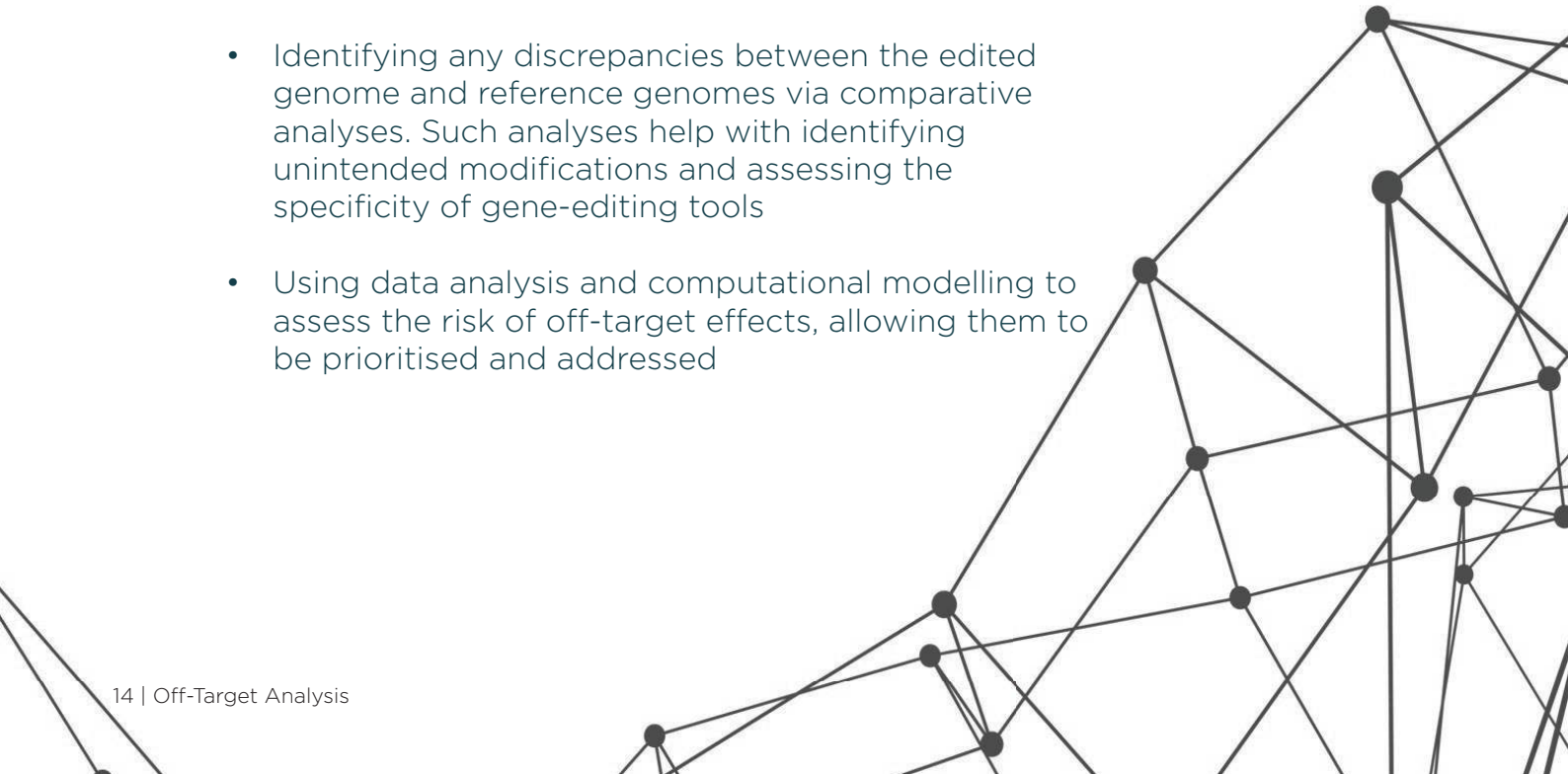
- Analysing genomic and proteomic data in order to prioritise antigens for T cell recognition, to ensure the expanded T cells target specific antigens associated with the target disease
- Predicting epitopes within antigens that will likely induce strong T cell responses. In turn, this helps with designing T cell expansion protocols that focus on immunogenic regions
- Analysing T Cell Receptor sequences to understand the diversity and specificity of T cell clones. This information guides the selection of T cells with desired antigen specificity for expansion
- Evaluating the clonality of expanded T cell populations to ensure a diverse collection of T cells with varied antigen specificities
- Monitoring the quality of expanded T cells (via computation analyses) by assessing factors such as cell viability, functionality, and potential off-target effects
- Combining genomics, transcriptomics, and proteomics data to provide a thorough understanding of T cell behaviour during expansion, which helps with optimising the expansion process for therapeutic efficacy
- Contributing to personalised T cell expansion strategies by assessing individual genetic and molecular profiles, to tailor the therapy to individual patients
- Simulating T cell expansion dynamics (via computational modelling) to predict the optimal conditions and factors influencing expansion outcomes
- Assisting in the study of T cell interactions with antigens, predicting potential immunogenic epitopes, and optimising T cell receptor (TCR) engagement
- Integrating different data types to identify potential biomarkers associated with successful T cell expansion

Off-Target Analysis

Off-target analysis is an essential component of cell and gene therapy research. Its role is to identify and address any unintended genetic modifications a potential therapy could cause. Therefore, it plays a vital role in ensuring the safety, precision, and overall effectiveness of cell and gene therapies.

Bioinformatics facilitates off-target analysis in various ways, such as:

- Assisting with target prediction. Bioinformatics tools can be used to compare the target sequence with the entire genome to identify any genomic regions that are similar to the intended target. This highlights potential off-target locations
- Analysing genomic data from gene-editing experiments to identify any off-target effects
- Variant calling. Bioinformatics can be leveraged to call and annotate genetic variants, which helps distinguish between intended modifications and off-target changes. This helps pinpoint genomic alterations and assess their potential impact on cellular function
- Identifying any discrepancies between the edited genome and reference genomes via comparative analyses. Such analyses help with identifying unintended modifications and assessing the specificity of gene-editing tools
- Using data analysis and computational modelling to assess the risk of off-target effects, allowing them to be prioritised and addressed



Pathway Analysis

Pathway analysis plays a valuable role in cell and gene therapy research as it helps in understanding the interactions and signalling pathways within cells and provides a comprehensive overview of the genes or proteins that are working together to regulate biological processes. This furthers cell and gene therapy research by assisting with target identification, and elucidating mechanisms of action. It also aids biomarker discovery, helps optimise therapeutic strategies and assists with ensuring the safety of therapies.

Bioinformatics supports pathway analysis by:

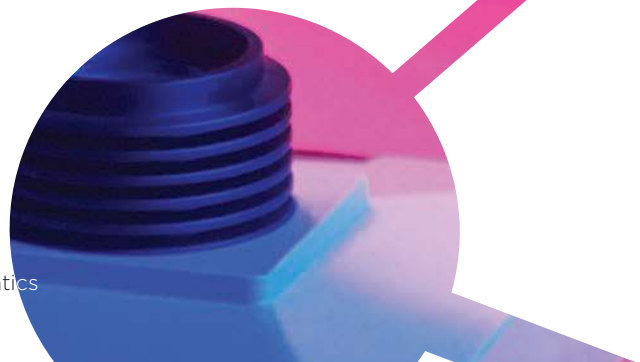
- Enabling the integration of omics data sets, such as genomics, transcriptomics, and proteomics data, to provide a detailed picture of the molecular landscape of cells and tissues
- Providing network analysis of molecular interaction networks to help understand the interactions between genes, proteins, and other molecular entities within pathways
- Identifying significant pathway enrichment or depletion in experimental datasets to focus research on biologically relevant pathways
- Enabling the annotation of genes and proteins with functional information. In turn, this helps with interpreting the roles they play within specific pathways and cellular processes
- Providing comparative analysis of pathway activities across different treatment groups, genotypes, and experimental conditions. This helps with the identification of key regulatory pathways
- Facilitating the development of computational models that can predict the impact of genetic modifications on cellular pathways, to assist with designing and optimising therapeutic strategies
- Enabling biomarker discovery to identify potential biomarkers associated with specific pathways. This assists with monitoring the response to cell and gene therapies

Structural Bioinformatics

Structural bioinformatics is a particular subset of bioinformatics that plays an important role in cell and gene therapy research. This is because structural bioinformatics provides insights into the three-dimensional structures of biological molecules, such as proteins and nucleic acids, which ultimately helps to create more targeted and effective cell and gene therapies.

It does this by:

- Assisting with design and optimisation of therapeutic molecules. Understanding the structure of a molecule allows researchers to modify and enhance its properties for improved efficacy and specificity
- Analysing the three-dimensional structures of molecules to assist with target identification and validation
- Providing insights into the interactions between therapeutic agents and their target molecules. This aids in the development of therapies that interact specifically and selectively with intended targets, minimising off-target effects
- Analysing and predicting protein-protein interactions to assist in the development of therapies to modulate specific cellular processes
- Analysing the molecular elements of antigen-antibody interactions to help develop safer and more efficient gene delivery systems
- Analysing structural information to provide mechanistic insights into the functioning of therapeutic agents at the molecular level. This helps with optimising the mechanisms of action in cell and gene therapies



ADVANCING YOUR CELL AND GENE THERAPY RESEARCH WITH BIOINFORMATICS

As you will now be aware, bioinformatics analyses can support cell and gene therapy research in a multitude of ways, assisting with everything from target identification to optimising therapeutic efficacy. This is why expert bioinformatics support can significantly reduce the time required to develop safe and effective cell and gene therapies.

Fios Genomics has over 15 years of experience providing bioinformatics support for drug discovery and development projects. In recent years, we have seen a significant increase in clients developing cell and gene therapies and we have updated our service offering to align with this. Our analyses supporting Antisense Oligonucleotide (ASO) research have been particularly popular. However, we provide bioinformatics services to support cell and gene therapy research throughout the discovery, development, and manufacturing processes.

Areas of cell and gene therapy research we support with bioinformatics include:

- Target identification
- Target validation
- Copy number and clonal analysis
- Gene expression analysis
- ASO research
- T cell expansion & Antigen Transduction research

If you would like to learn more about how we can support cell and gene therapy research, or if you would like to discuss support for your own research project, contact us and we will be happy to help!

CONTACT US

fios GENOMICS

Fios Genomics is a bioinformatic analysis provider **helping clients to gain more insight from their research data.**

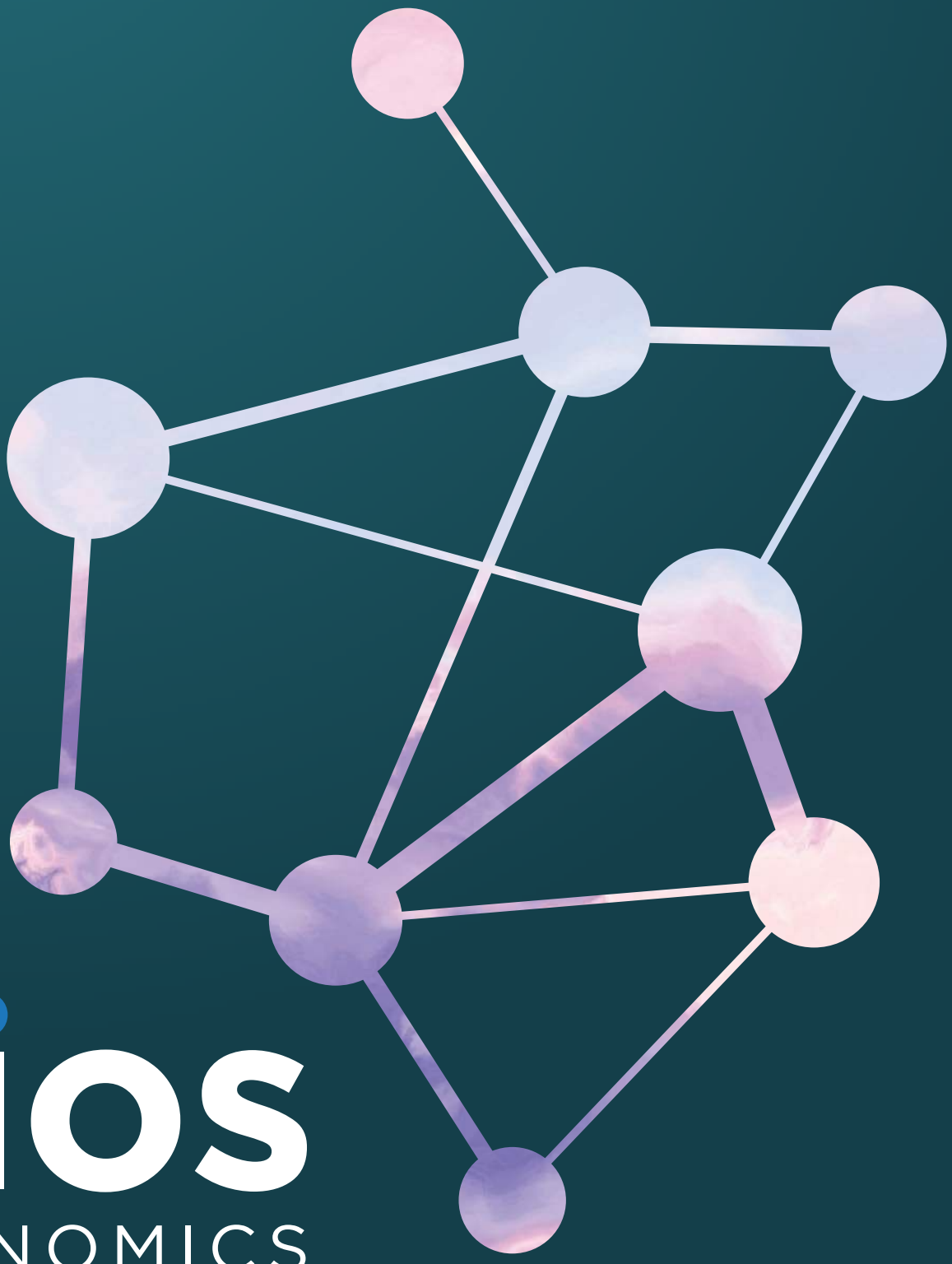
OVERVIEW

With over 15 years of experience in supporting scientists, researchers and bioinformaticians in data analysis, Fios have extensive experience in handling all types of data sets for drug discovery & development, diagnostics, agricultural research, veterinary medicine and applied research across all species. Our specialised team of bioinformaticians, statisticians, and biologists are able to analyse and interpret any genomic, transcriptomic, proteomic & metabolomic data, independent of the platform used.

LEARN MORE

"We have utilized the Bioinformatics team at Fios Genomics for many of our drug discovery projects, as they provide expertise in the analysis of complex bioinformatic datasets. We have been consistently impressed with the rigor of Fios Genomics' work, their communication throughout the projects, and the rapid speed at which they complete their analyses."

- Dr Scott Ribich, Vice President of Biology at Accent Therapeutics



fios
GENOMICS