SELVENTA™

Analyze. Accelerate. Clarify.

PRISME Forum

Ewan Hunter PhD

Director Technical Support EMEA

May 3-4, 2011

Corporate Overview

- Company founded in 2002; corporate headquarters in Cambridge, MA
- Updated corporate strategy
  - Rebranded end of November 2010
  - Refocusing the company to capture strategic value (Personalized Medicine)
  - 8 commercial partners to date
- Privately held: Series A, Flagship ventures, Pappas ventures
- Experienced leadership team:
  - David de Graaf, Ph.D., President and Chief Executive Officer
  - Louis Latino, Executive Vice President, Sales & Marketing
  - Julian Ray, Ph.D., Senior Vice President, Technology & Engineering
  - David Fryburg, M.D., Chief Medical Officer
  - John Tagliamonte, Senior Vice President, Corporate Development
  - Chris Thomajan, Chief Financial Officer

Selventa Vision and Mission

Vision

Your strategic partner in finding optimal treatments for the right patients

Mission

Apply patient data-driven analytics to increase the value of portfolios

Selventa Is a Personalized Healthcare Company

Devoted to Accelerating Portfolio Optimization Through Patient Stratification

The BEL Language

The Blueprint for Knowledge Sharing in Biomedical Sciences

BEL Represents Scientific Findings with Qualitative Causal Relationships

PubMed ID: 999999999
"We demonstrate that RNA expression of Y is mediated through activation of the X transcription factor"

ta(p(X)) increases r(Y);

"increased transcriptional activity of protein designated by X increased the abundance of RNA designated by Y"

PubMed ID: 999999998
"We demonstrate that phosphorylation of X at T32 results in suppression of the transcriptional activity of the transcription factor X"

p(X, P@T32) directlyDecreases ta(p(X));

"increased abundance of the protein designated by X phosphorylated at threonine 32 directly decreased transcriptional activity of the protein designated by X"

BEL Uses Standard Vocabularies and Ontologies

- Terms are defined by functions that reference external ids
  - **p(EG:207)**
    - "the abundance of the protein designated by EntrezGene id 207"
    - *(AKT1 Human)*
  - **p(UP:P31749)**
    - "the abundance of the protein designated by UniProt id P31749"
    - *(AKT1 Human)*
  - **bp(GO:0006915)**
    - "the biological process designated by the GO id 0006915"
    - *(apoptosis)*
- External ids can include names in well-defined namespaces
  - **p(HUGO:AKT1)**
    - "the abundance of the protein designated by HUGO gene symbol 'AKT1'"
  - **bp(MESH:apoptosis)**
    - "the biological process designated by the MESH heading 'apoptosis'"

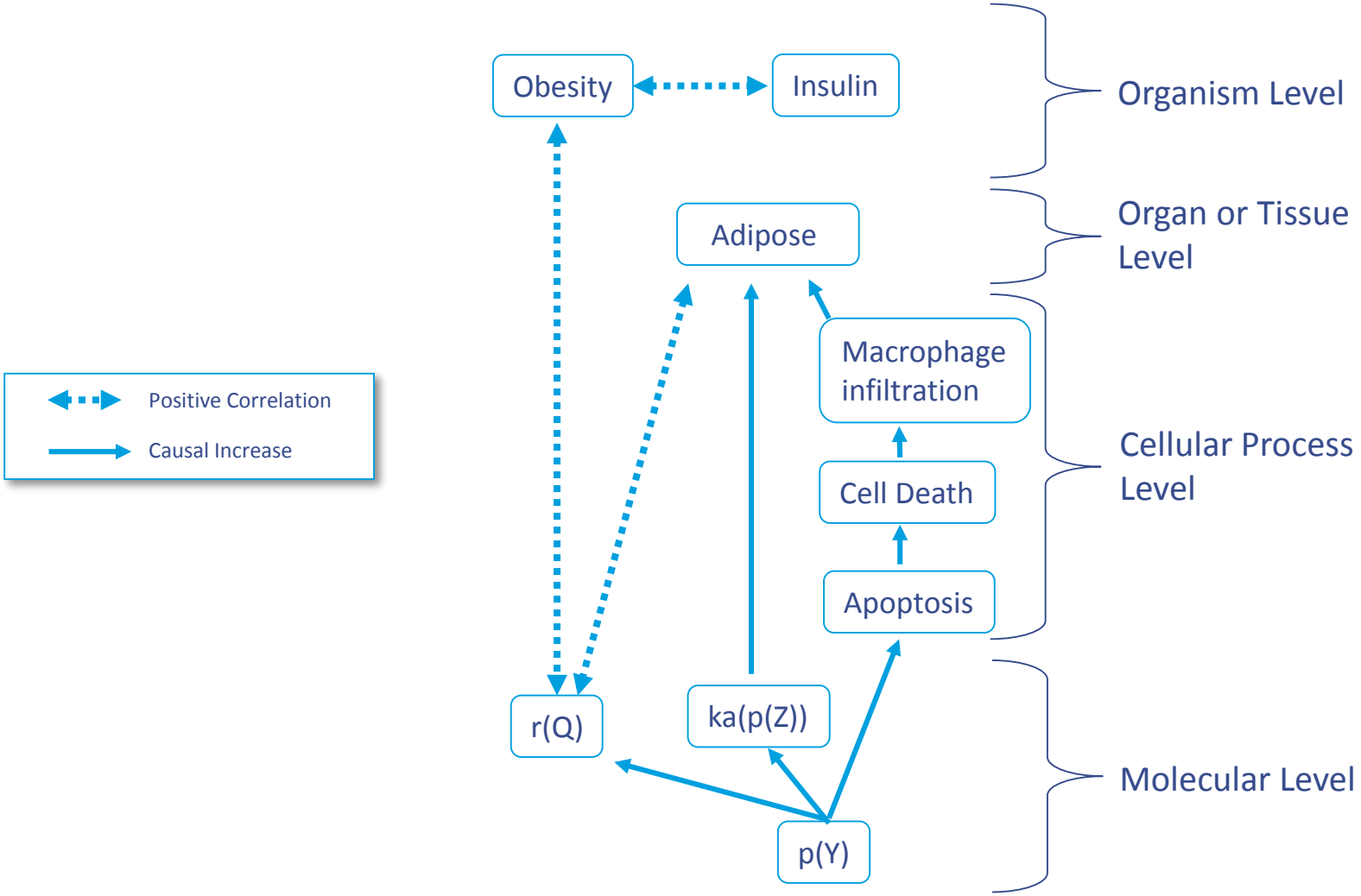
BEL Manages Equivalences Between External IDs

- Original statements in BEL Documents:
  - **ka(p(EG:207)) decreases bp(GO:0006915);**
  - **ka(p(UP:P31749)) directlyDecreases ka(p(UP:P49841));**

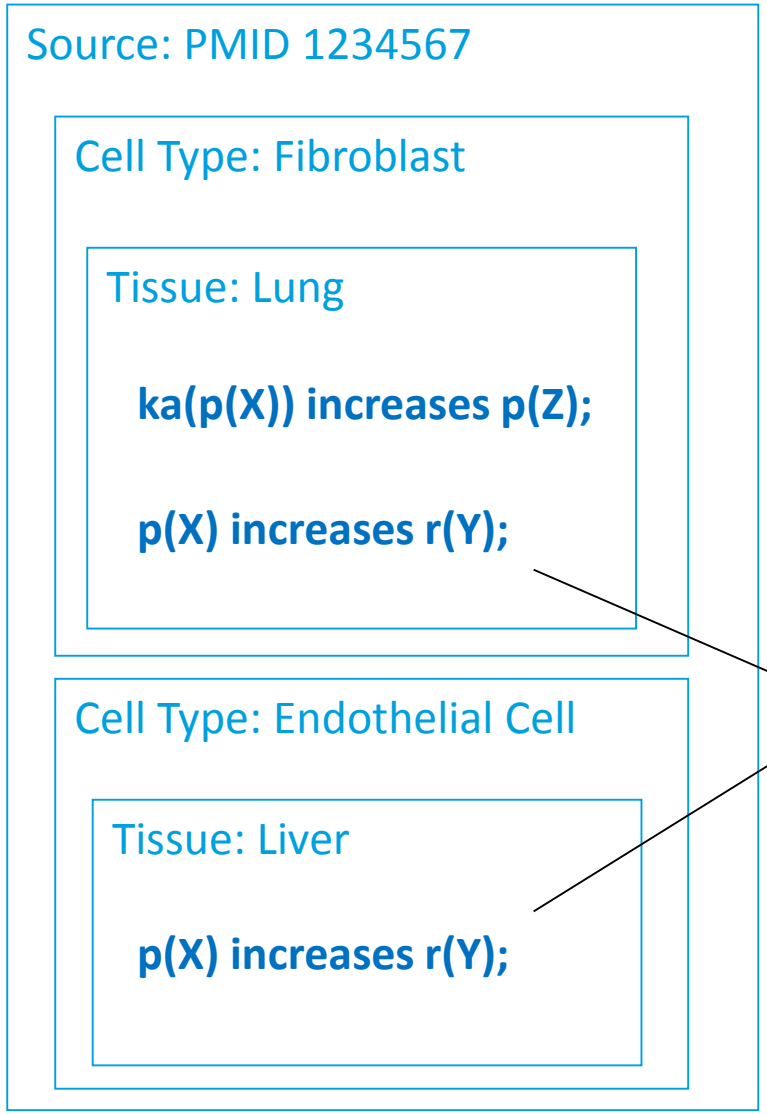
- Equivalence relationships specified to BEL Compiler:
  - **GO:0006915 ➔ MESH:apoptosis**
  - **EG:207 ➔ HUGO:AKT1**
  - **UP:P31749 ➔ HUGO:AKT1**
  - **UP:P49841 ➔ HUGO:GSK3B**

- Processed Statements:
  - **ka(p(HUGO:AKT1)) decreases bp(MESH:apoptosis);**
    - **“kinase activity of AKT1 decreases apoptosis”**
  - **ka(p(HUGO:AKT1)) directlyDecreases ka(p(HUGO:GSK3B));**
    - **“kinase activity of AKT1 directly decreases kinase activity of GSK3B”**

Multiple Representation Levels Coexist



Contexts Annotate Statements



Causal relationships demonstrated in lung fibroblasts, reported in PMID 1234567

Causal relationship demonstrated in liver endothelial cells , reported in PMID 1234567

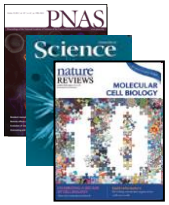
Each Statement is distinct: These Statements have different sets of contexts

BEL Framework

- Suite of **software components** which facilitate the interchange of biological scientific facts between user-communities and between applications

- Supports 3 types of workflows:
  - Knowledge Creation/Management
    - Generating and editing knowledge as BEL Documents
    - Sharing knowledge
    - Moving knowledge between applications or knowledge sources (e.g. NLP methods)
  - Knowledge Publishing
    - Assembling knowledge by combining BEL Documents into KAMs
    - Moving KAMs between BEL-enabled applications
  - Knowledge Use By Applications
    - Using published knowledge in BEL-enabled applications such as the GTP

Using BEL Documents

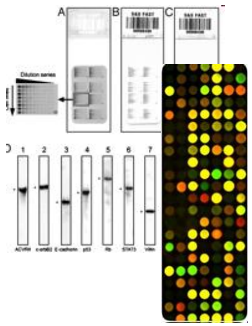


Text Mine Public or Proprietary Sources

Capture the results of text-mining algorithms in a portable format

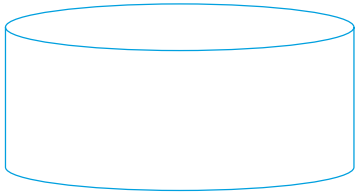
Publish Key Findings in an Area of Biology

Create comprehensive documents that capture the critical findings for an area of research



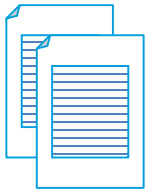
Capture Experimental Results

Manage public or proprietary findings in a form that abstracts biological relationships while preserving experimental context



Transform Database, Pathway or Other Structured Content

Integrate knowledge derived from databases, pathway knowledge bases and other structured sources using BEL



BEL Documents

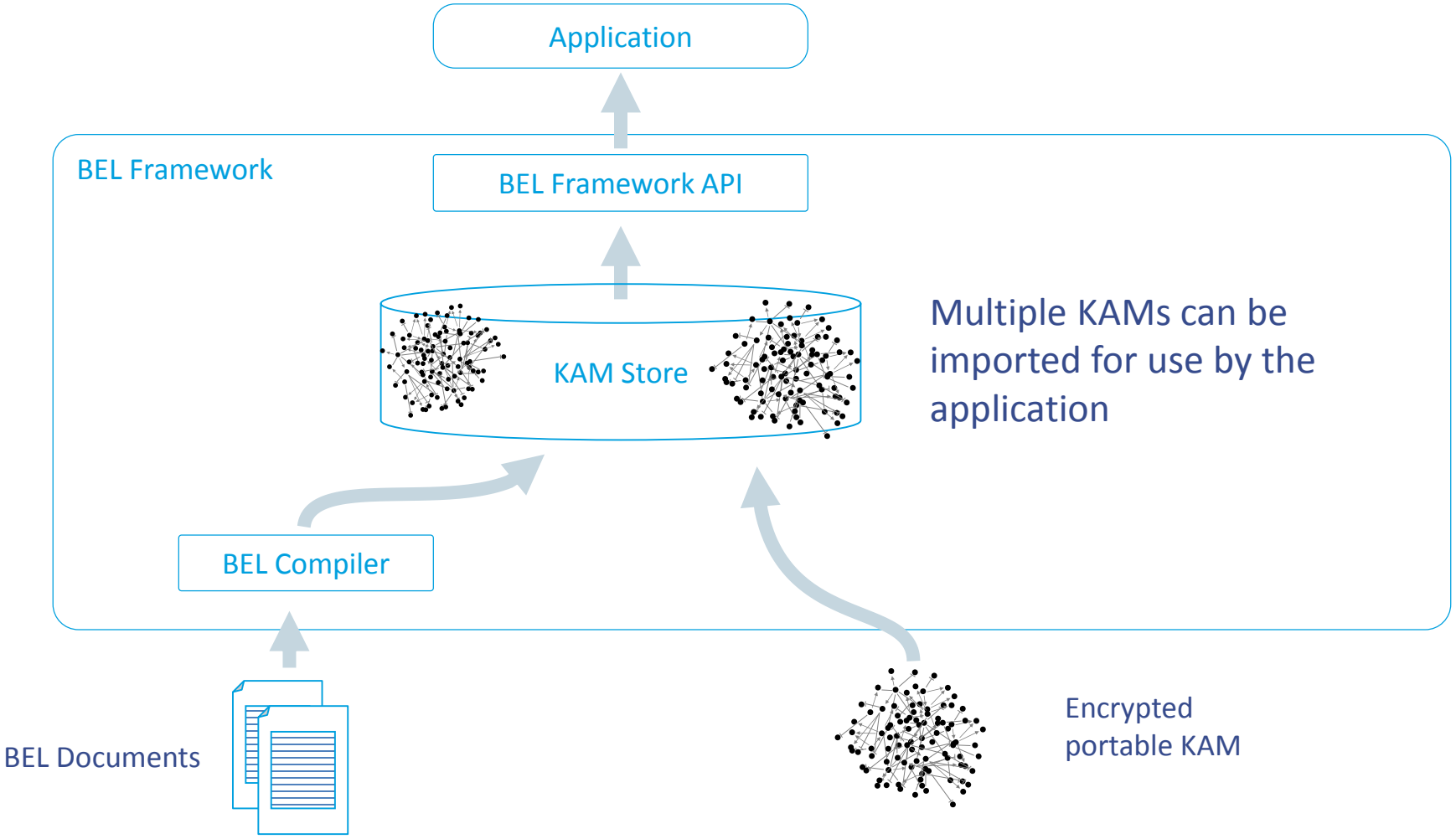
Review, Edit and Organize Knowledge

Edit and organize knowledge published in BEL to create new knowledge resources meeting defined quality standards

Knowledge Assembly Model

- Directed network of biological facts assembled by the BEL Compiler derived from one or more BEL Documents
- Network is composed of KamElements
  - KamVertex – references a biological entity defined in a BEL Document
  - KamEdge – asserts a biological relationship between two biological entities evidenced by one or more BEL Statements defined in a BEL Document
- Each KamEdge references one or more BEL Statements and associated provenance and contexts
- Can be exported into an encrypted, portable format which can be imported into another KAM Store
- KAMs are stored using internal references for biological entities which map to an encrypted symbol table which can only be decrypted by the BEL Framework API

Knowledge User Workflow
BEL Framework and Applications



BEL Framework API

- Java API providing access to KAMs in a KAM Store
- SPARQL API providing access to a RDF representation of the KAM
- Allows users to dynamically filter a KAM by specifying a set of include/exclude filters:
  - Provenance filters
  - Context filters
  - KamVertex filters
  - KamEdge filters
- Allows access to underlying evidence support for KamEdges
  - Statements and contexts
  - Information derived from the BEL Document headers

BEL Framework Web Service API

- Extends the BEL Framework API to the web
- Provides a SOAP-based and RESTful Web Service API to allow non-java based applications to access and query KAMs stored in a KAM Store
- Includes a self-contained web server
- Can be deployed on a server and configured to work with http/https and configurable ports

Comparing BEL with BioPAX

BEL	BioPAX
Represents scientific findings in molecular biology in a form which is both computable and intuitive for life scientists	Represents biological pathways at the molecular and cellular level and to facilitate the exchange of pathway data
Focus: capture of causal relationships to ultimately enable inference by applications	Focus: represent detailed molecular interactions and abstraction to pathways
Can flexibly annotate each represented finding with rich contextual information	Elements in a pathway can be supported with evidence codes from external vocabularies
Facilitates representation of incomplete knowledge, where findings may demonstrate causality but not mechanism	Oriented towards precise representation of interactions and reactions
Designed with a simple structure to empower a broad range of biologists to effectively curate and review knowledge	Designed to provide a data exchange format for pathway data that will represent the key elements of the data models from a wide range of popular pathway databases.
Enables the creation of use-neutral knowledge resources which can subsequently be assembled to create specific models	Pathways are specific models; it would be feasible to assemble BioPAX pathways starting with BEL documents

Comparing BEL with SBML

❖ BEL

Represents scientific findings in molecular biology in a form which is both computable and intuitive for life scientists

Focus: capture of ***qualitative*** causal relationships to ultimately enable inference by applications

Enables the creation of use-neutral knowledge resources which can subsequently be assembled to create specific models

A potential avenue for future BEL extensions would be to enable the capture of additional quantitative information associated with scientific findings, facilitating the assembly of SBML models from information in BEL documents.

• SBML

A machine-readable format for representing biological models.

Focus: describing systems where biological entities are involved in, and modified by, processes that occur over time

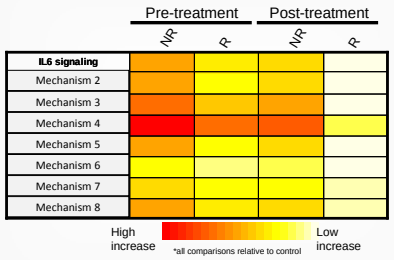
SBML describes specific models.

Example of Portfolio Design through Stratification of Patients

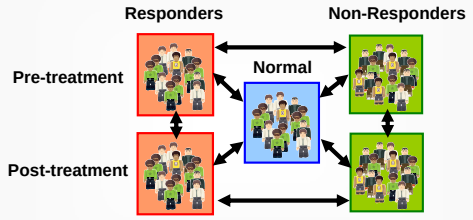
Identifying Disease Driving Mechanisms Resulting in Infliximab Non-response in Ulcerative Colitis

Identification of Optimal Treatments for the Right Patients

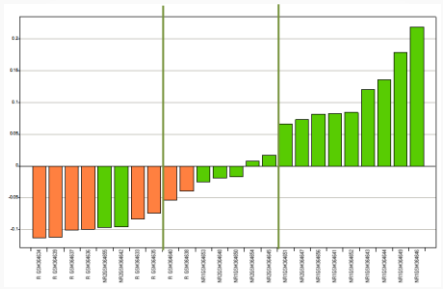
1 Calculate strength of molecular processes in a patient population



2 Identify disease-driving mechanisms in non-responders



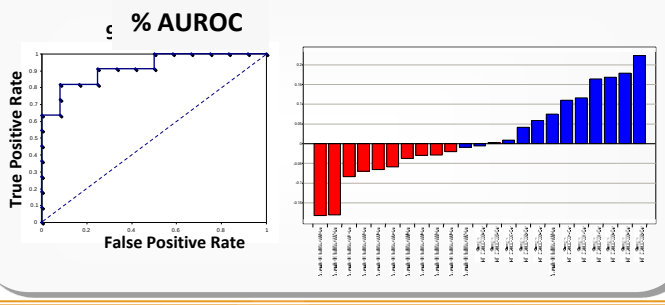
3 Assess Strength of Disease Driving Mechanism in Individual Patients



4 Generate a classifier that differentiates patients with high and low activation of disease-driving mechanism

Gene Classifiers		
IL6	Mechanism 1	Mechanism 3
Gene 1	Gene A	Gene i
Gene 2	Gene B	Gene ii
Gene 3	Gene C	Gene iii
Gene 4	Gene D	Gene iv
Gene 5	Gene E	Gene v
Gene 6	Gene F	Gene vi
Gene 7	Gene G	Gene vii
Gene 8	Gene H	Gene viii
Gene 9	Gene I	Gene ix
Gene 10	Gene J	Gene x
Gene 11	Gene K	Gene xi
Gene 12	Gene L	Gene xii
Gene 13	Gene M	Gene xiii
Gene 14	Gene N	Gene xiv
Gene 15	Gene O	Gene xv
Gene 16	Gene P	Gene xvi
Gene 17	Gene Q	Gene xvii
Gene 18	Gene R	Gene xviii
Gene 19	Gene S	Gene xiv
Gene 20	Gene T	Gene xx

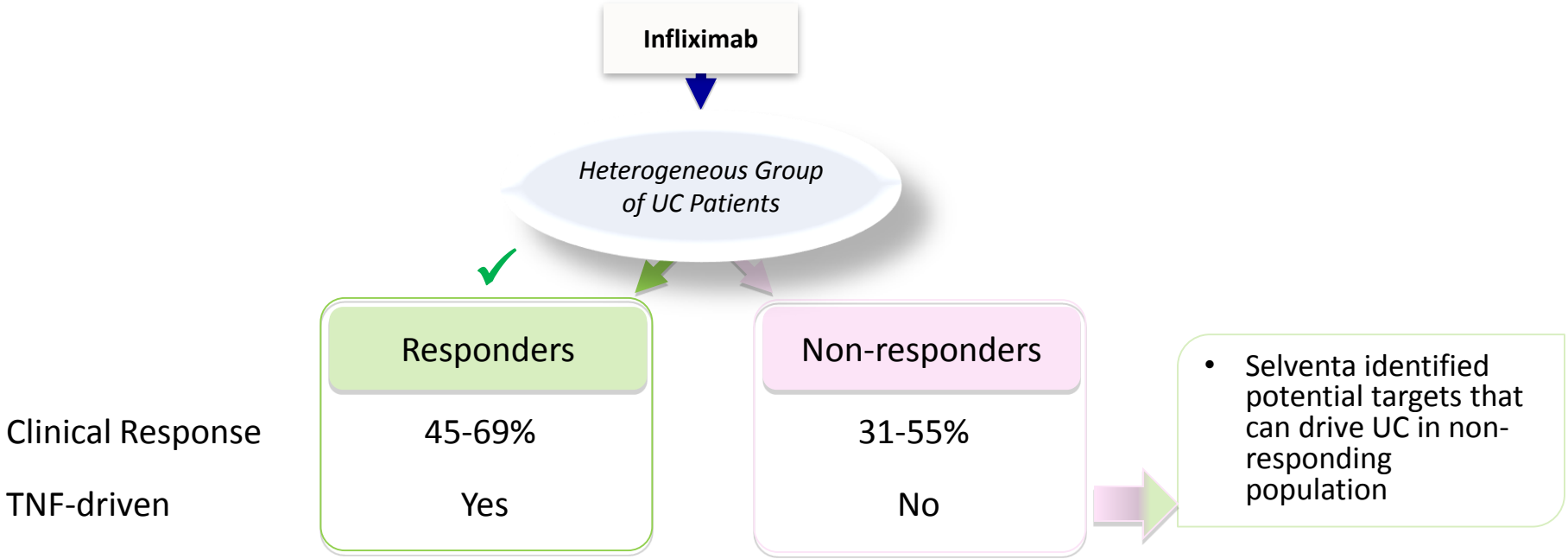
5 Assess classifier performance in an independent test set



6 Portfolio design based on identified mechanisms

Addressing Infliximab Non-Response in Ulcerative Colitis

- Objective:
  - Leverage Selventa's key technology, which is powered by the Genstruct® Technology Platform (GTP), to identify mechanistic differences between responders and non-responders
- An example of a value creation case study:
  - Identification of non-TNF-driven disease mechanisms in ulcerative colitis (UC)
    - PMIDs 19956723, 19700435



Knowledge Encoded in BEL Is a Substrate for RCR and Identifies Mechanistic Causes of the Data

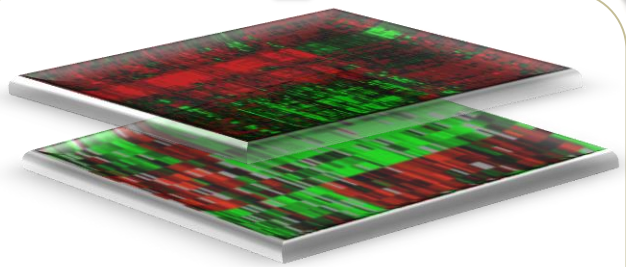
(e.g. *Increase in TNF)*

Identification of mechanistic causes leading to differential gene expression changes

Knowledgebase

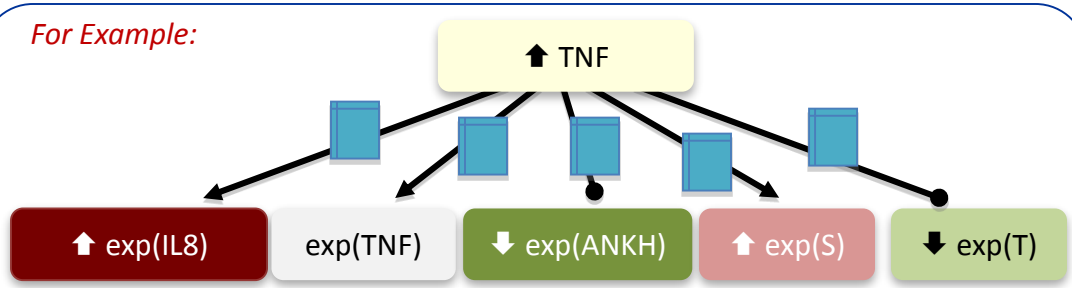
A collection of cause-and-effect relationships

Reverse Causal Reasoning

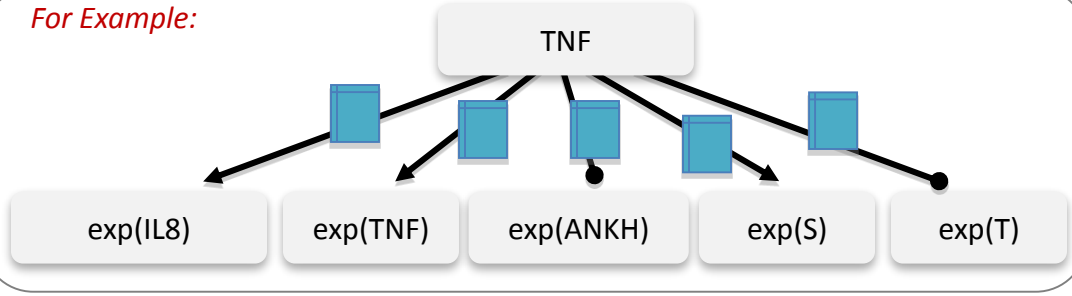


Differentially expressed genes

For Example:

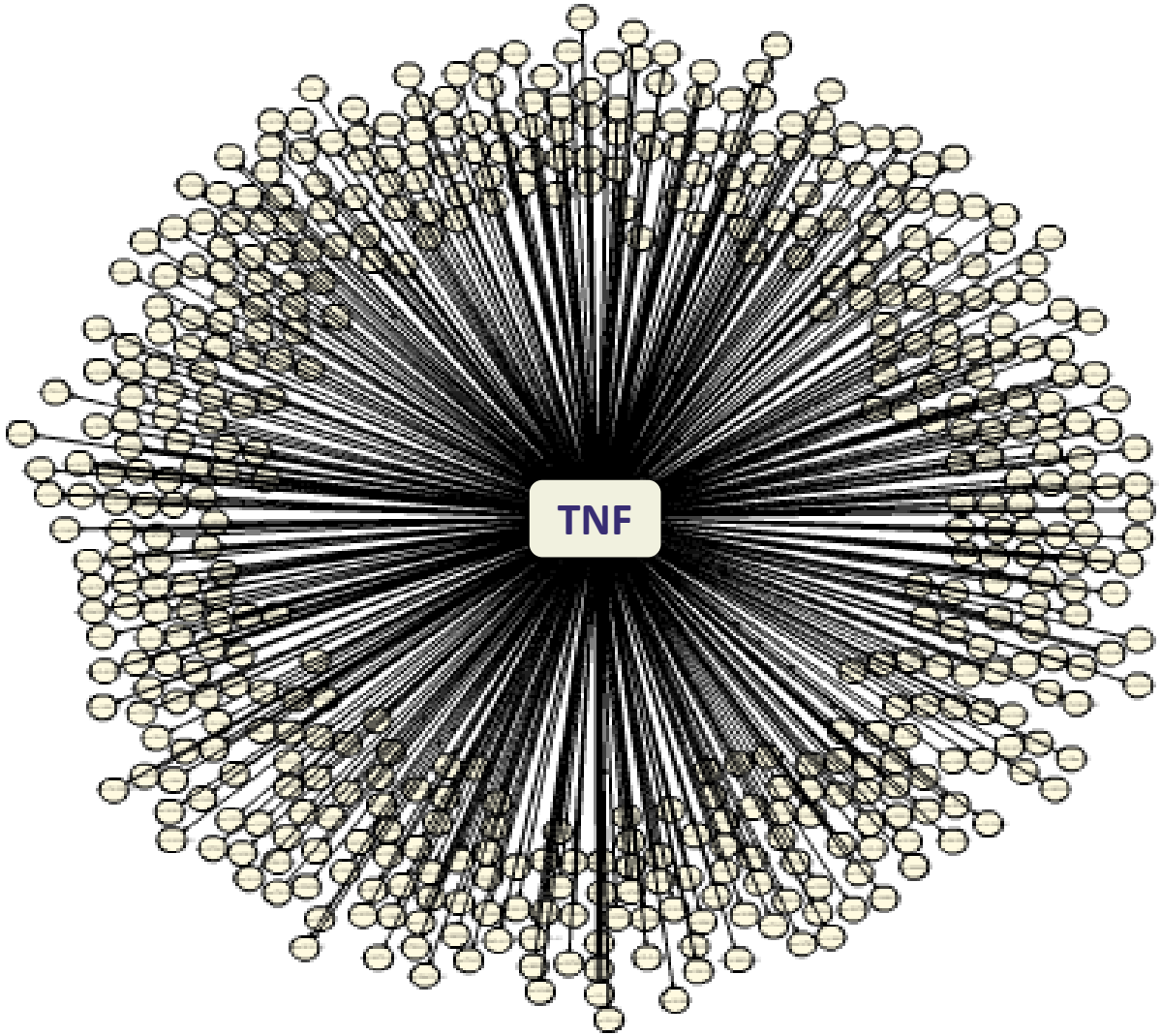


For Example:



For Example:

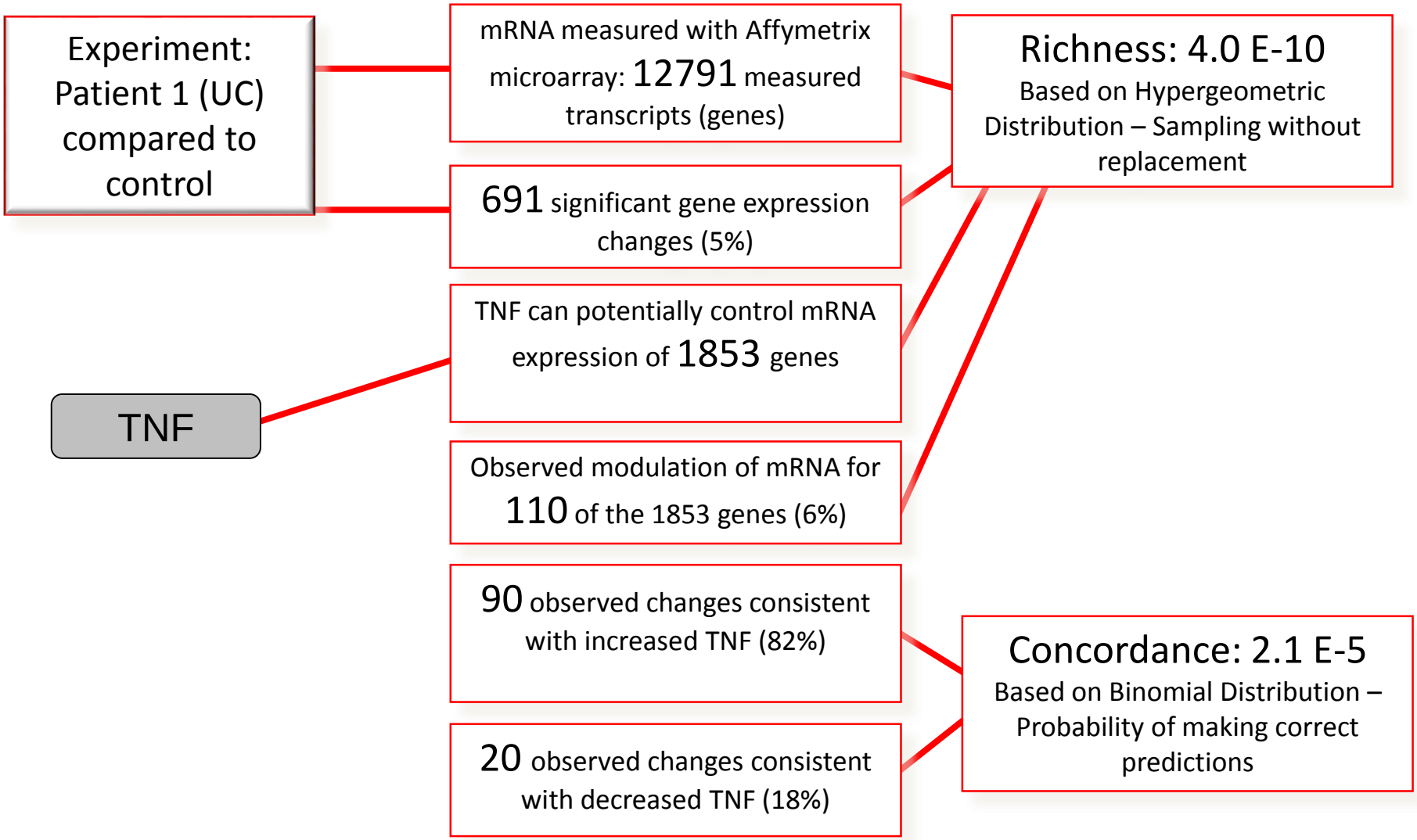
TNF-specific Signature in the Knowledgebase



TNF can potentially regulate 1853 genes based on 691 unique peer-reviewed publications

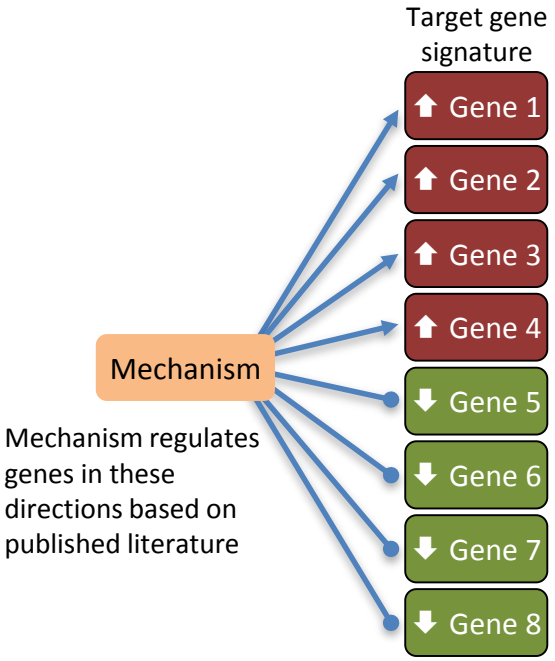
Patient 1 vs. Control has 110 of these significantly changed

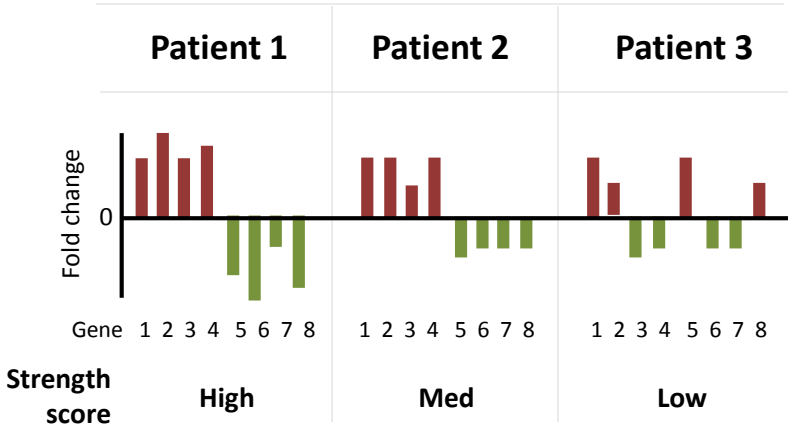
Statistical Significance Is Determined by Calculating Richness and Concordance



Apply a Strength Metric to Stratify IBD Patients for TNF Pathway Activation

- The strength metric is calculated on a target gene signature
  - The strength algorithm calculates the geometric mean of the fold changes in the gene signature
  - A list of gene fold changes can be collapsed into a single number
- A quantitative value is assigned to each patient for their level of signaling specific to the target
  - Assesses the relative signaling strength of a target network in each patient of a population
  - Patients can then be stratified on a continuum of network strength

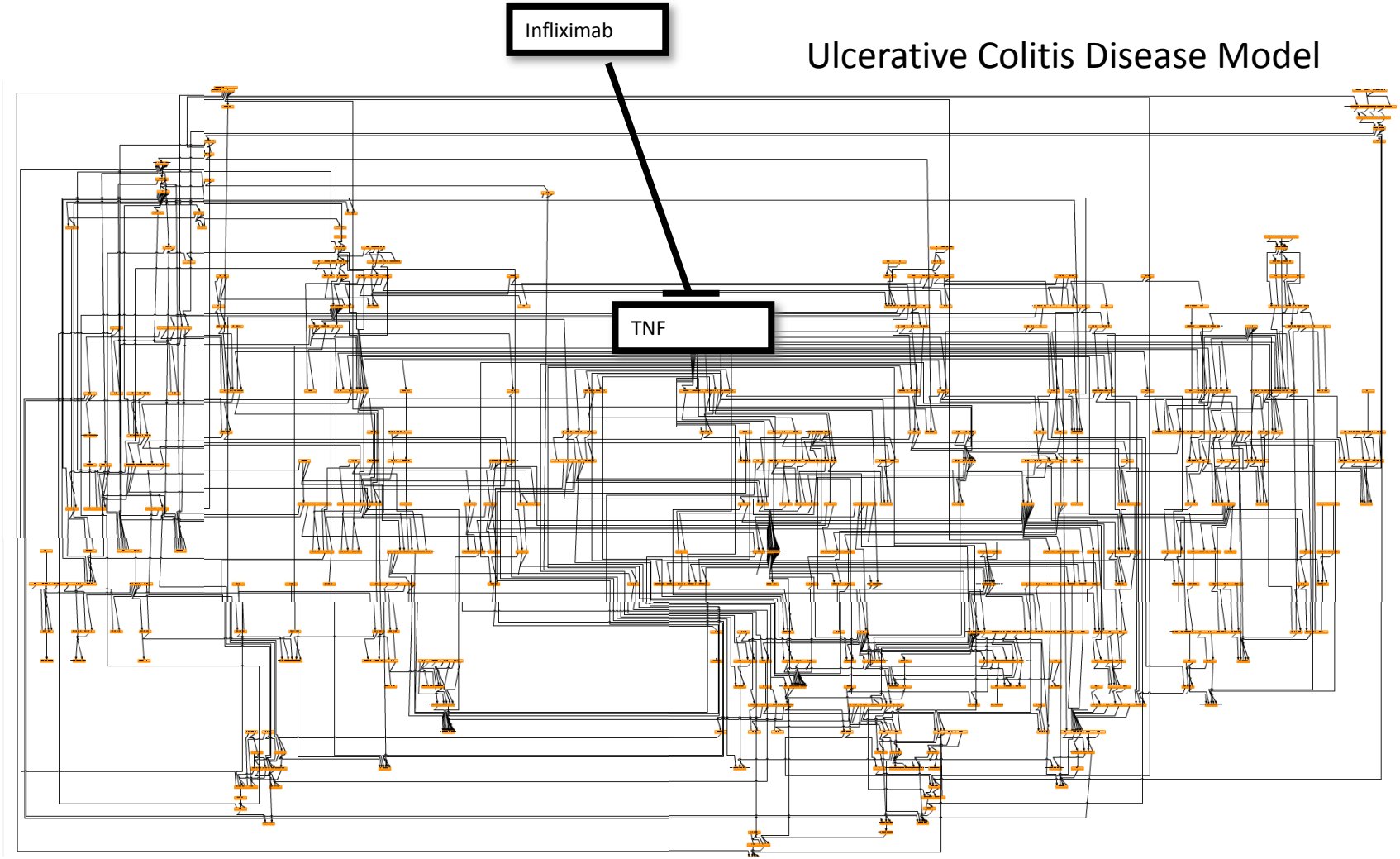




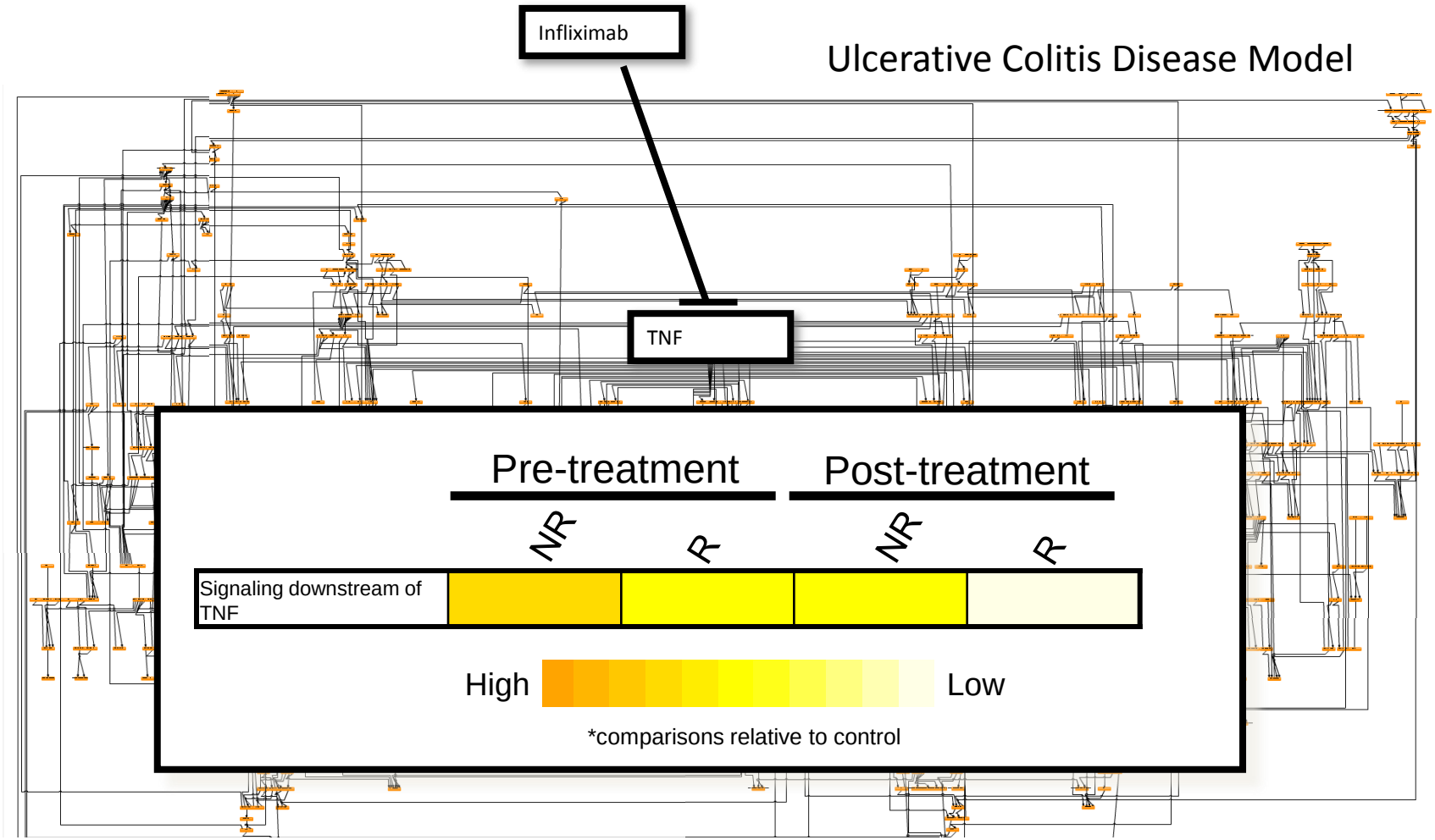
Generation of Strength Scores for Mechanisms

- Selventa's technology assesses the relative strength of hundreds of mechanisms for each individual patient
  - Strength score is calculated for > 2,000 direct mechanisms in the Knowledgebase
- Non-response stems from alternative disease driving mechanisms
  - Reflected in difference in strength of activation of molecular processes between responders and non-responders

Infliximab, a TNF-Inhibitor, Ameliorates UC in ~60% of Patients

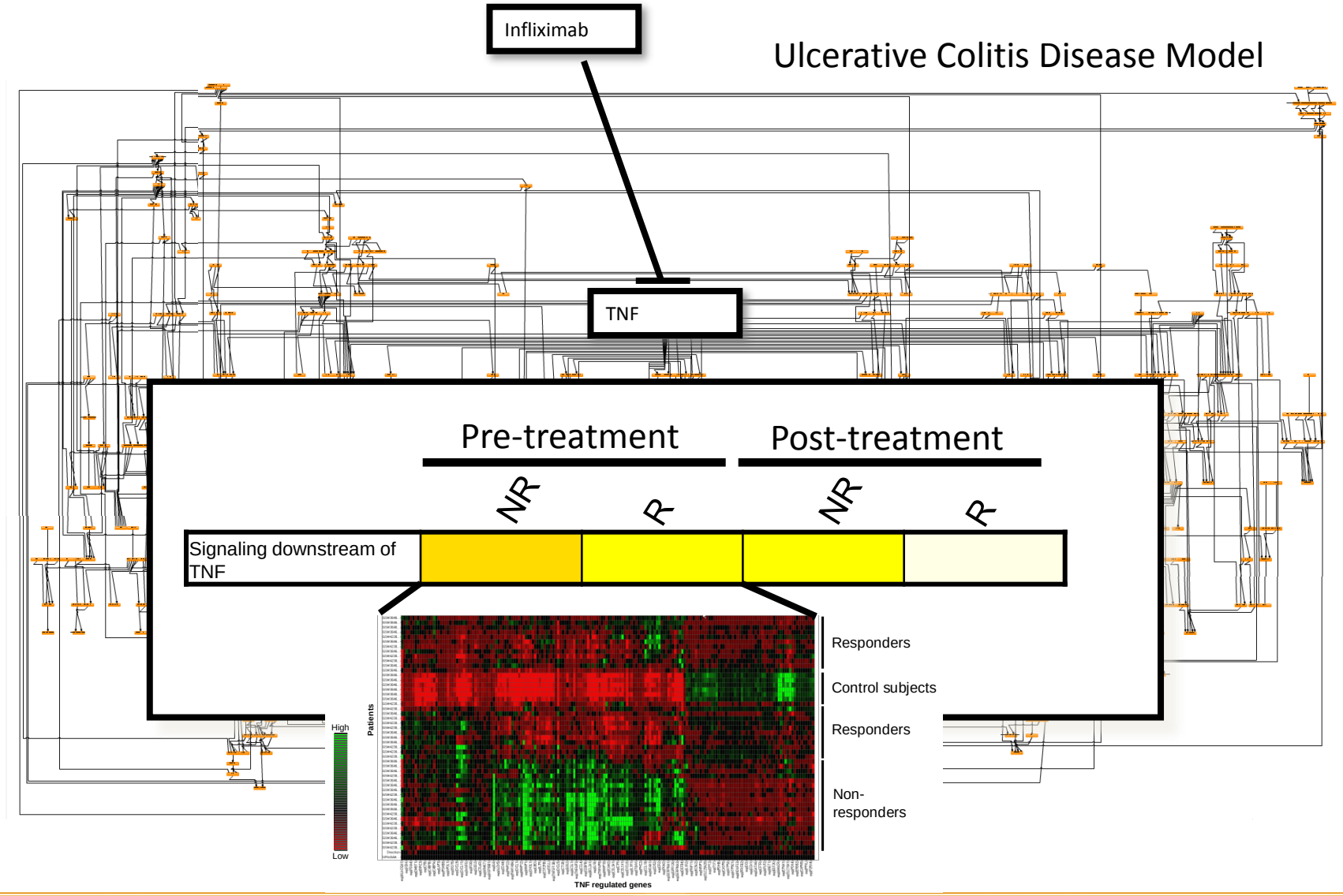


RCR Demonstrates Sustained TNF-like Downstream Signaling in Non-responders after Infliximab



- Antibody is in excess, sustained signaling is likely due to alternative upstream controller

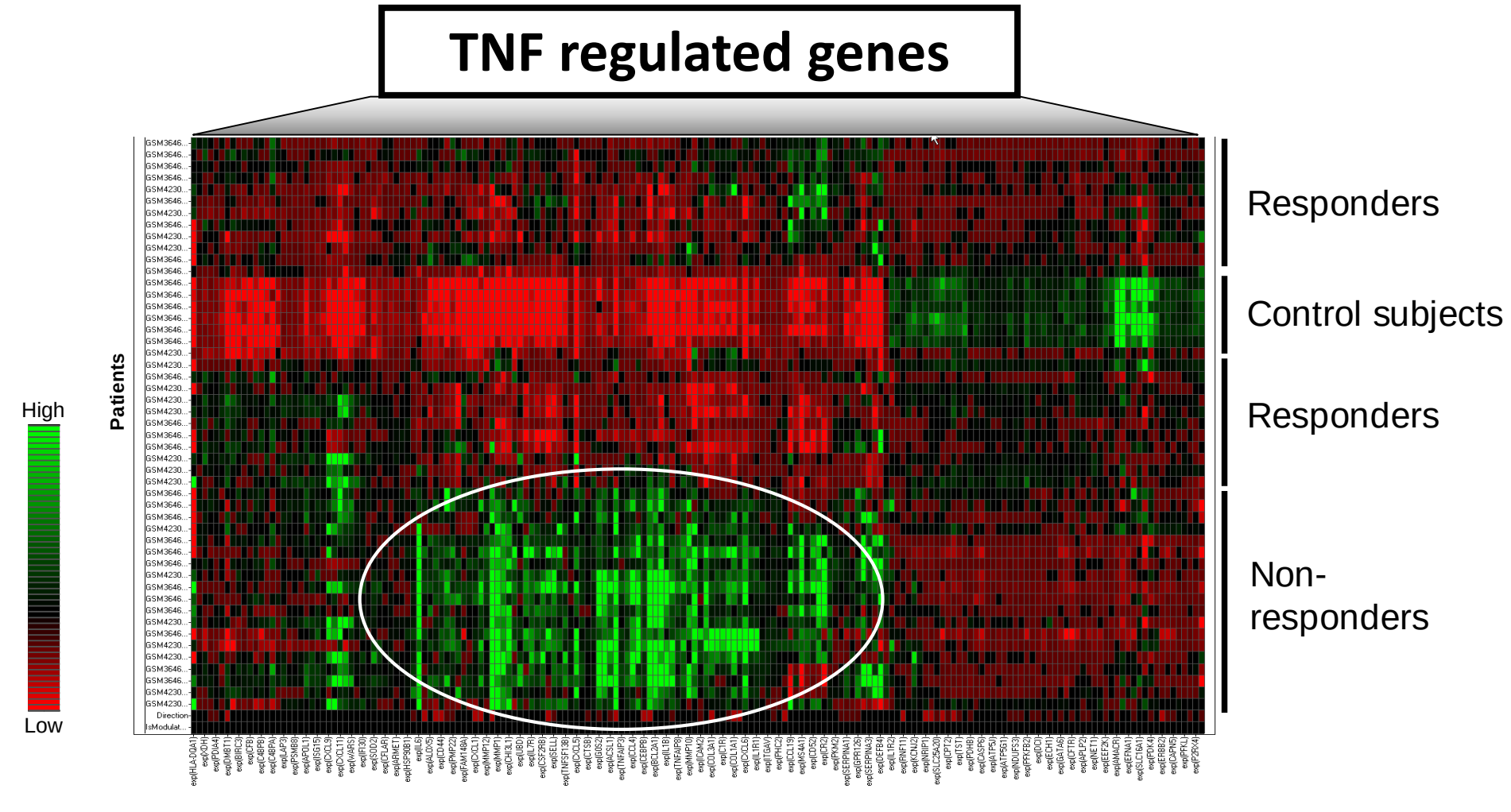
TNF Hypothesis Is Supported in Responders and Non-responders



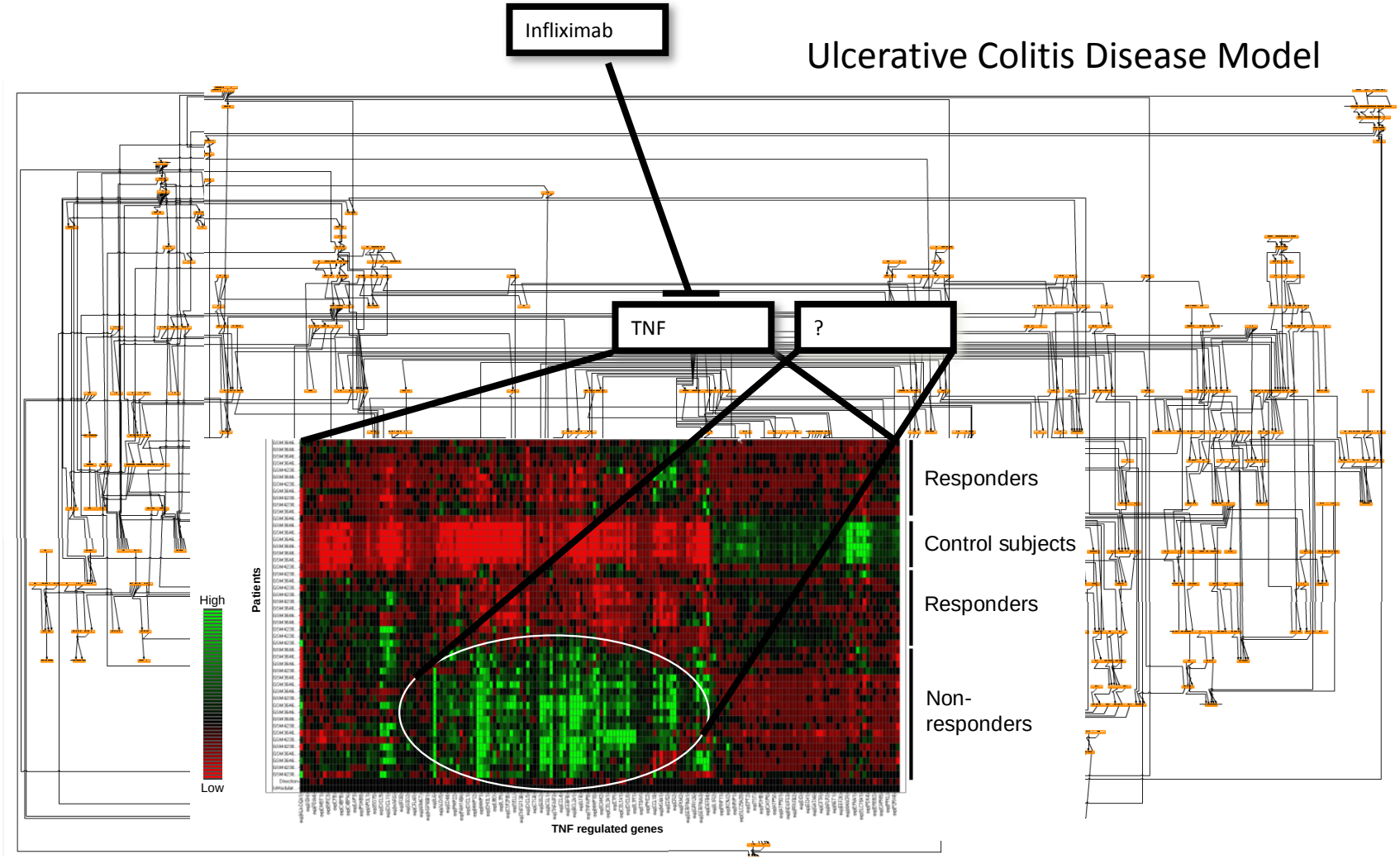
Non-responders and Responders Have Different TNF-like Signatures at Baseline

Identification of Subtle Differences Between Responders and Non-Responders

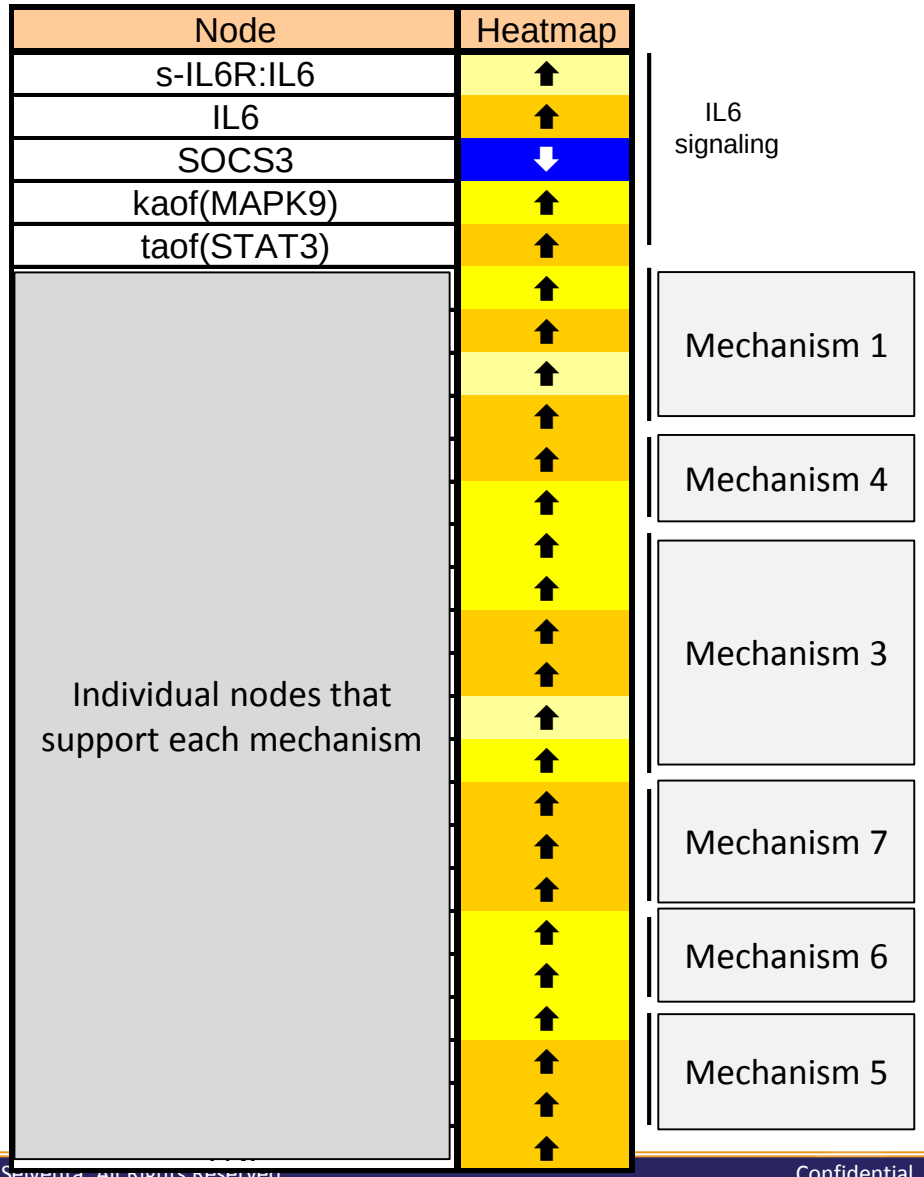
- Activation of alternative upstream controllers for TNF regulated genes in NR?



Alternative Disease-driving Mechanisms in Non-responders



Alternative Upstream Controllers for TNF-Regulated Genes in Non-responders

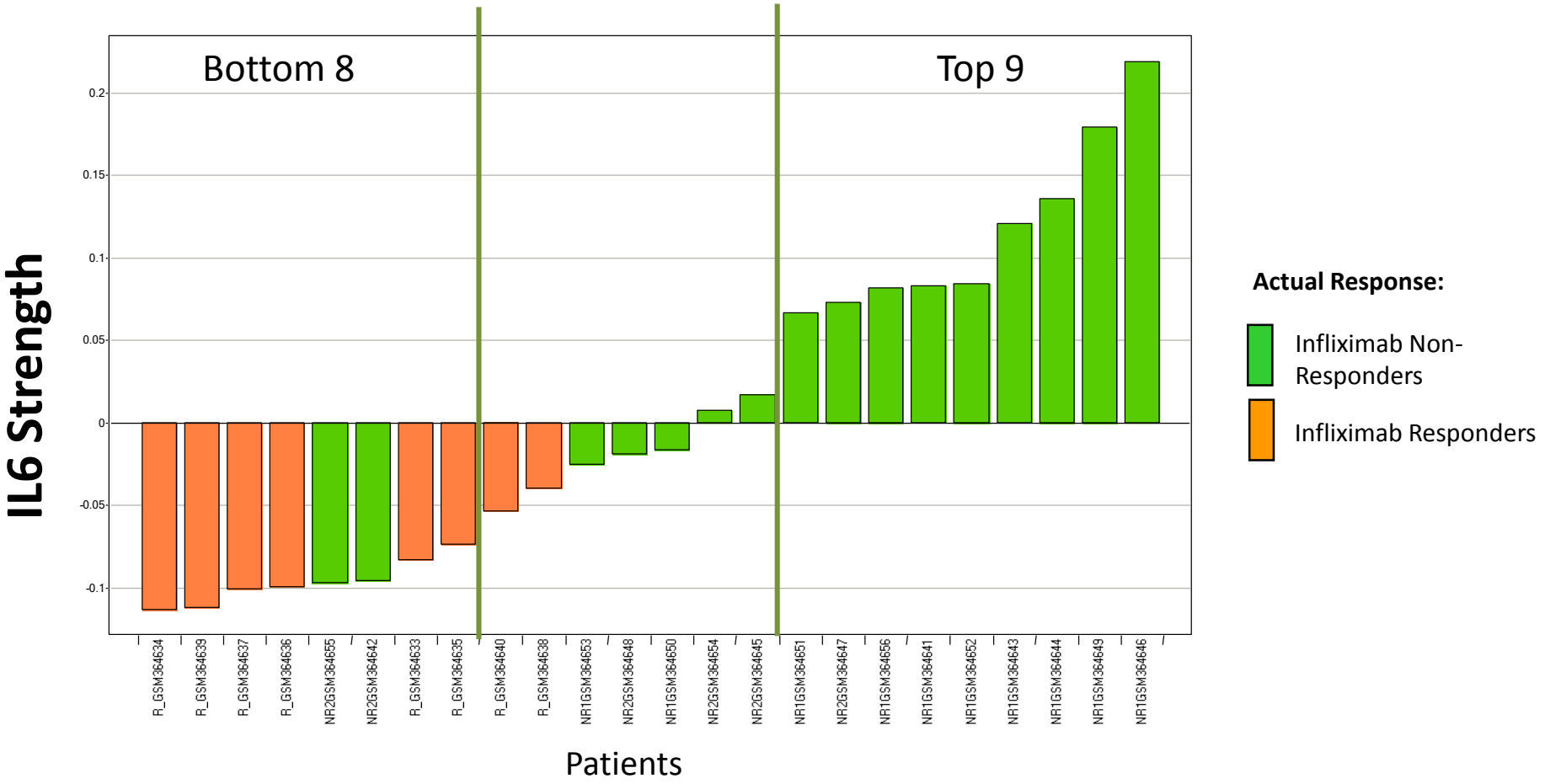


- TNF-regulated genes unique to non-responders implicate TNF-independent disease-driving mechanism in non-responders
  - These mechanisms can be targeted in the non-responder population

A group of causally related hypotheses provides a mechanism which is amenable to therapeutic intervention

Personalized Healthcare

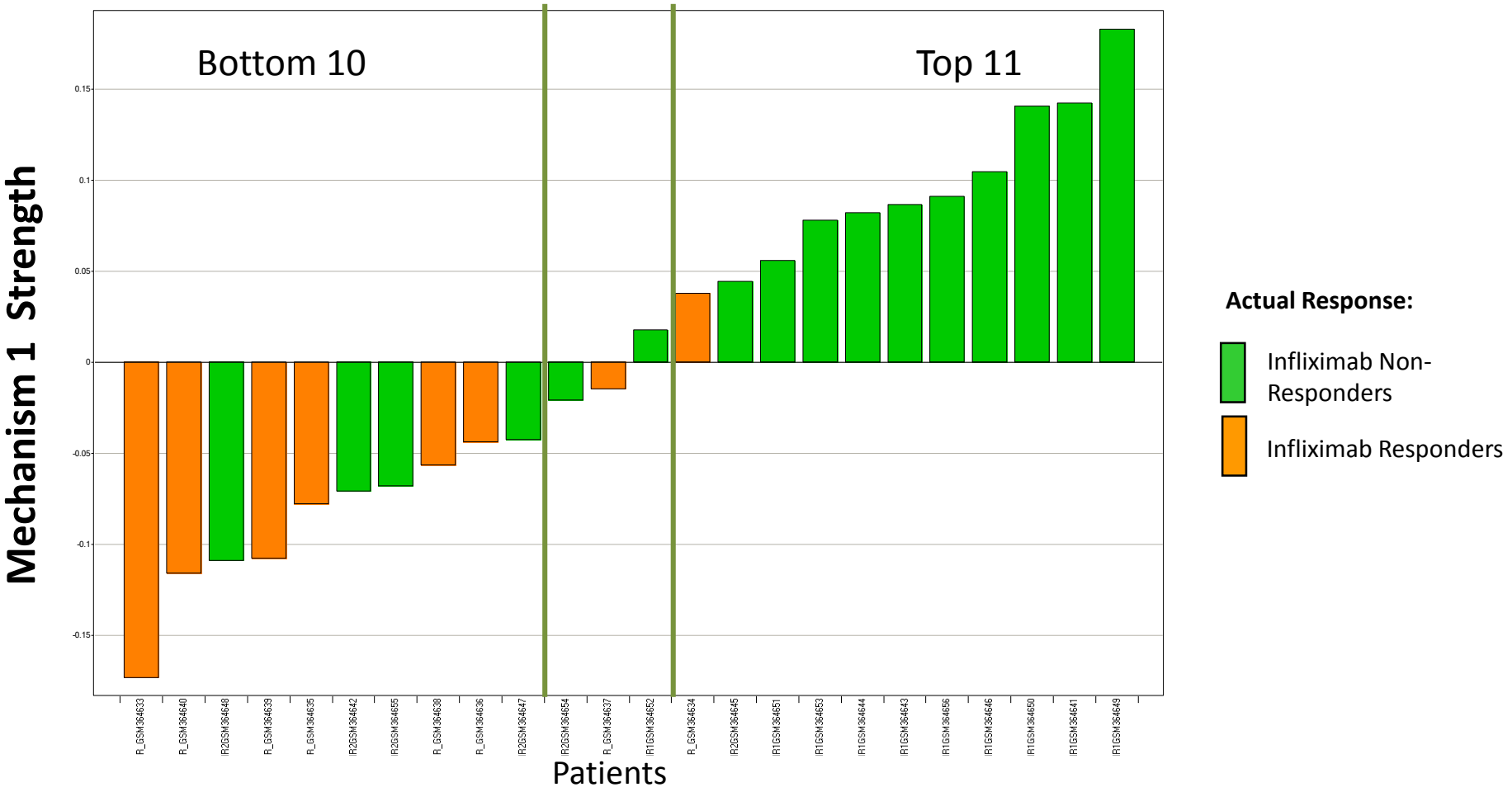
Patients Are Stratified Based on IL6 Strength



- Patients with the highest and lowest levels of IL6 signaling strength were selected for use in gene classifier generation
  - Training set GSE16879

Personalized Healthcare

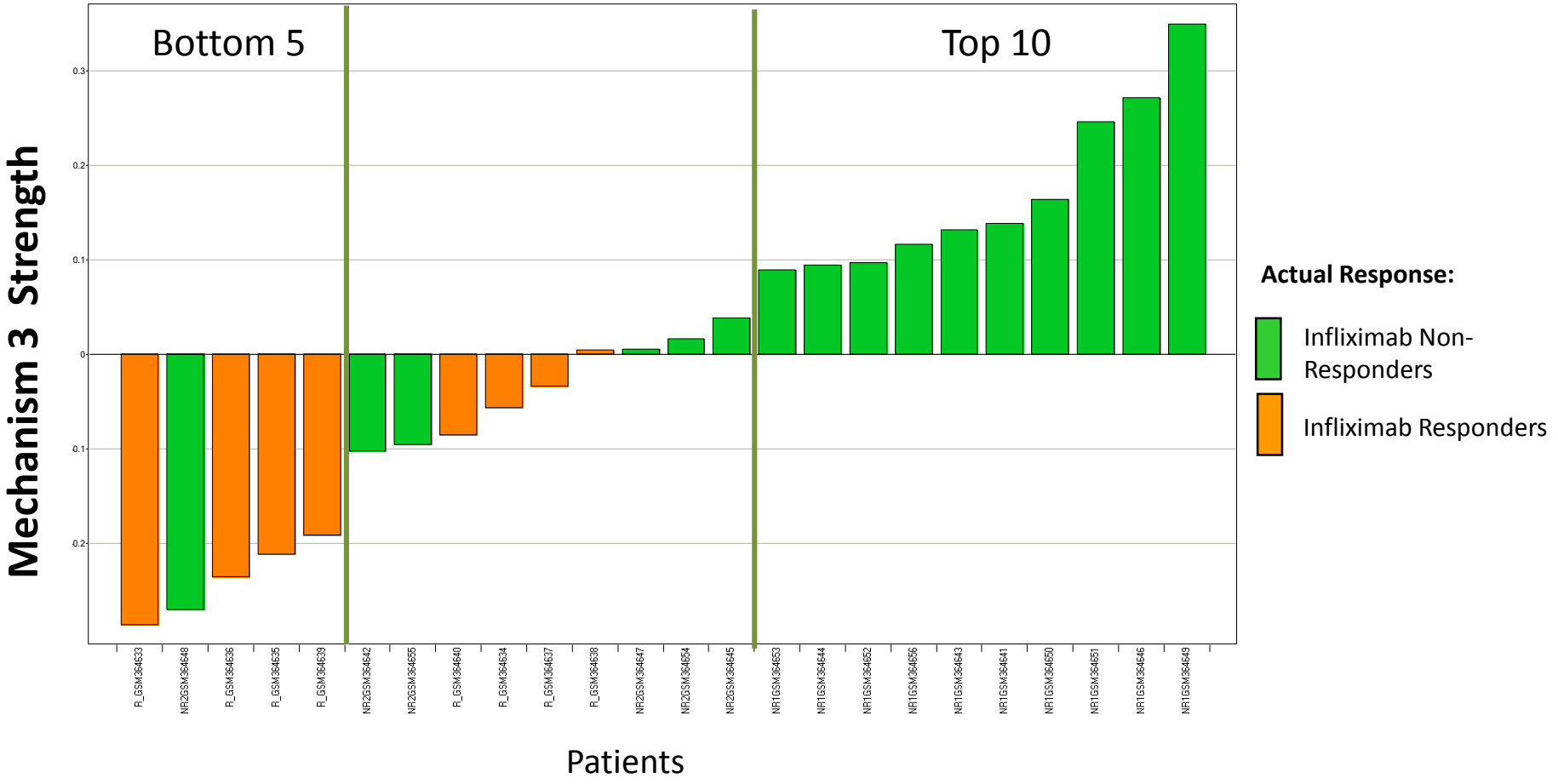
Patients Are Stratified Based on Disease-Driving Mechanism



- Patients with the highest and lowest levels of Mechanism 1 signaling strength were selected for use in gene classifier generation
  - Training set GSE16879

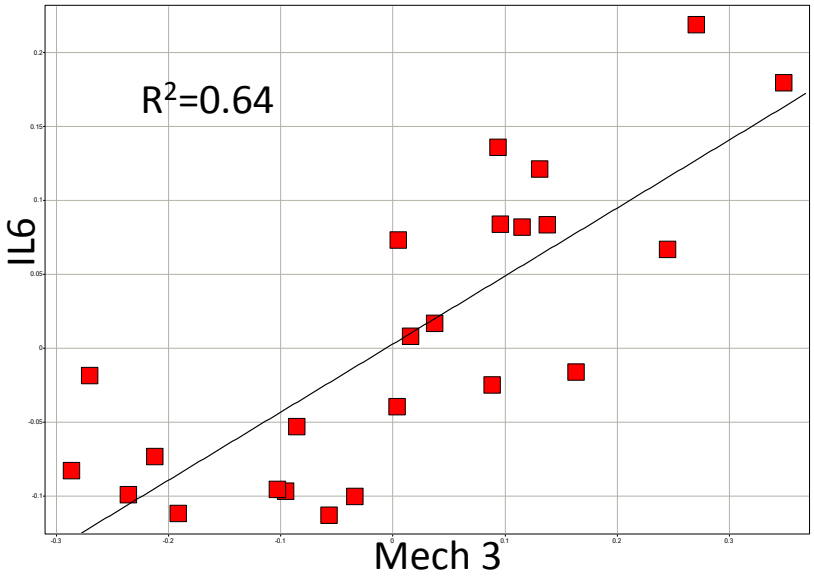
Personalized Healthcare

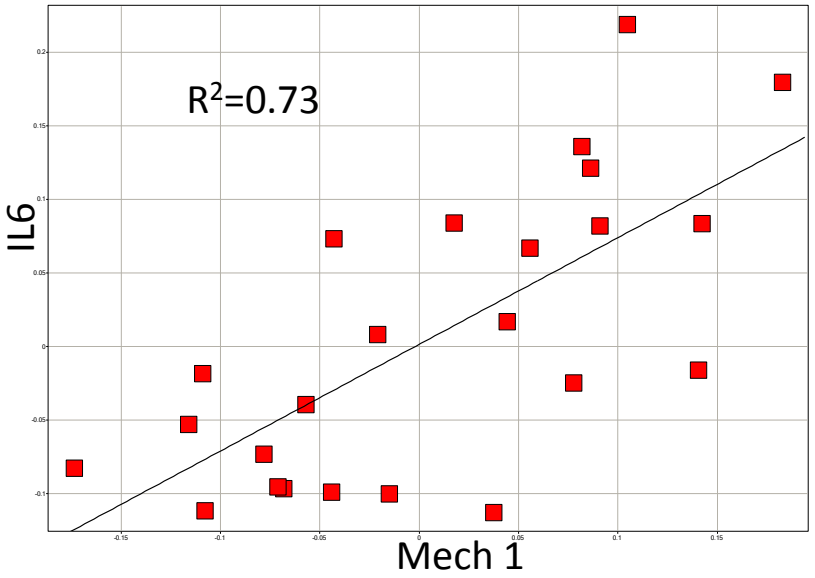
Patients Are Stratified Based on Disease-Driving Mechanism

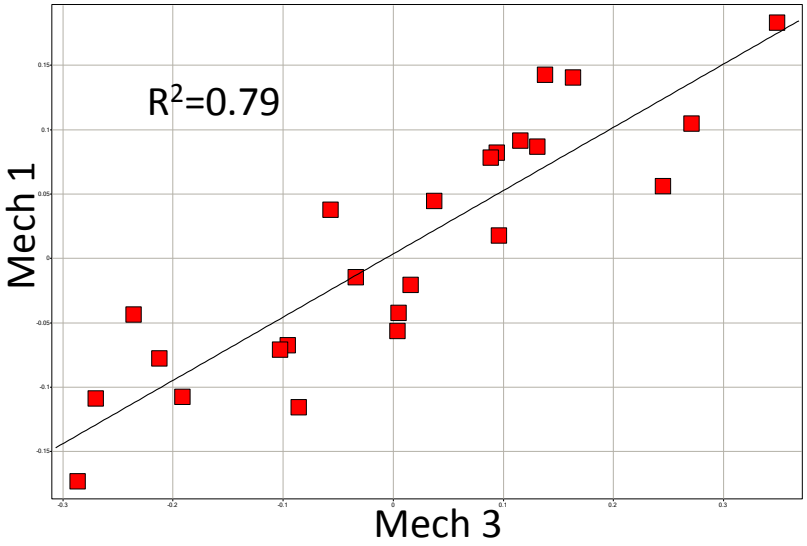


- Patients with the highest and lowest levels of Mechanism 3 signaling strength were selected for use in gene classifier generation
  - Training set GSE16879

Mechanisms Work Together or Separately

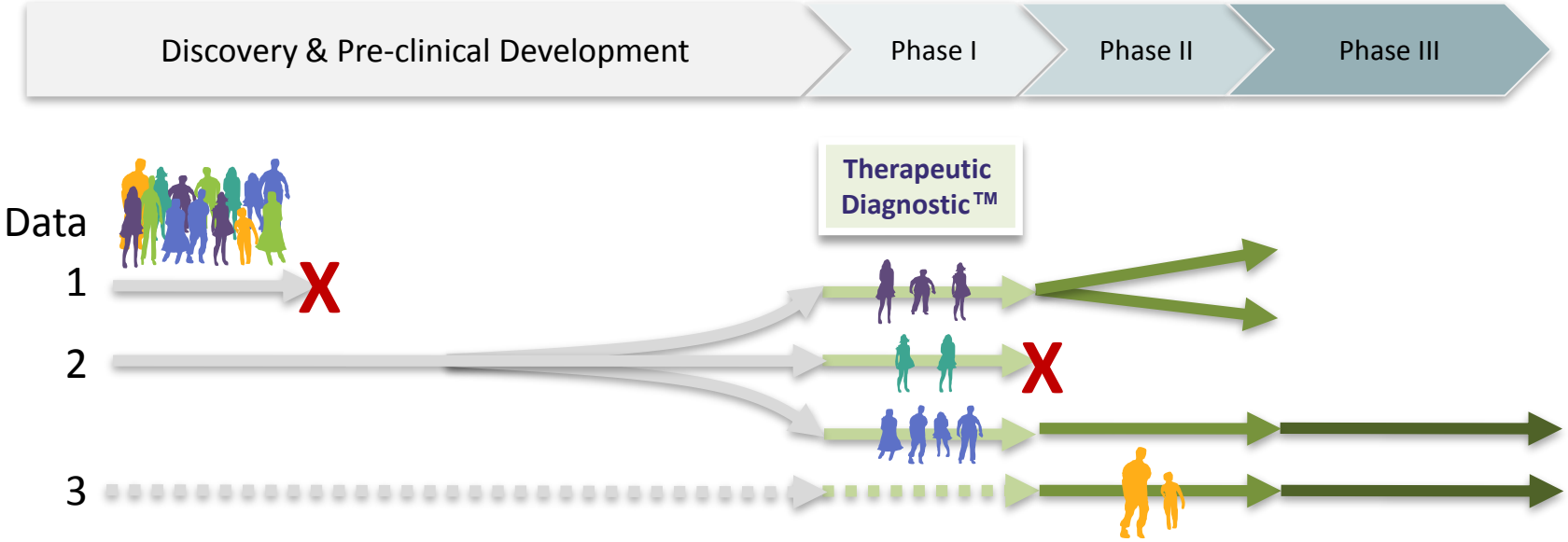






- Mechanism 1 and 3 may overlap, but IL6 is an independent driver

Portfolio Design Based on Mechanisms



Summary

Market Need

Drug Options

Drug A

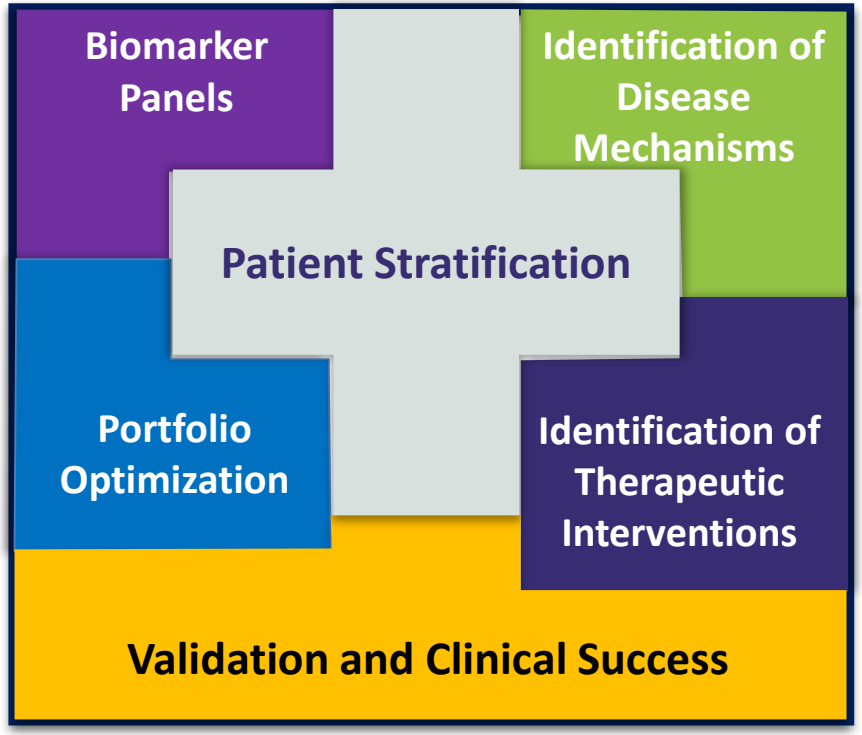
Drug B

Drug C

Drug X

Patients of a Specific Disease

Selventa Solution



Drug A

Optimal Therapy for the Right Patients