Resource Article

Machine Learning Identifies Signatures of Macrophage Reactivity and Tolerance that Predict Disease Outcomes

Authors: Pradipta Ghosh[1-3]*†, Saptarshi Sinha[1,4], Gajanan D. Katkar[1], Daniella Vo[4], Sahar Taheri[4], Dharanidhar Dang[4], Soumita Das[3,5†] and Debashis Sahoo [3,4,6]*†

Affiliations:
[1]Department of Cellular and Molecular Medicine, University of California San Diego.
[2]Department of Medicine, University of California San Diego.
[3]Moores Cancer Center, University of California San Diego.
[4]Department of Pediatrics, University of California San Diego.
[5]Department of Pathology, University of California San Diego.
[6]Department of Computer Science and Engineering, Jacob's School of Engineering, University of California San Diego.

One Sentence Summary: *S*ignatures of *ma*crophage *r*eactivity and *t*olerance (*SMaRT*) predict disease outcomes

Key Words:
- Artificial Intelligence/Machine Learning
- Boolean Equivalent Clusters
- Macrophage
- Reactive
- Tolerant
- Innate immune response
- Outcome prediction

*Contributed Equally

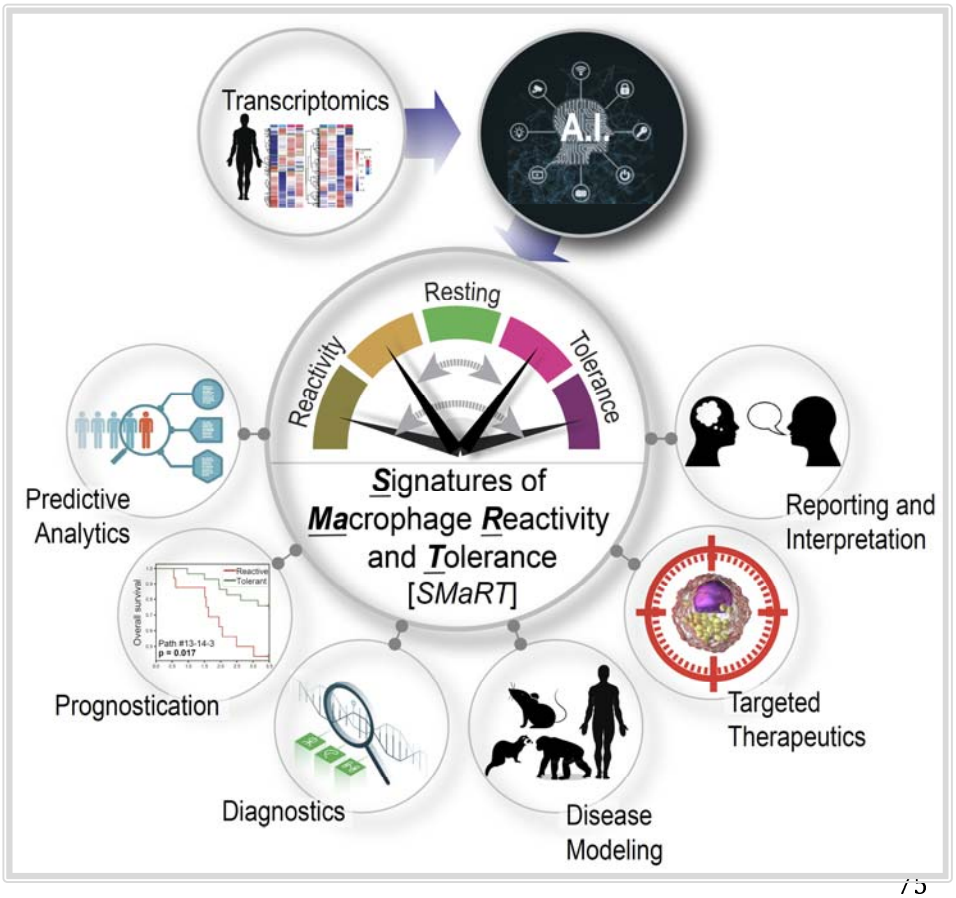
†**Corresponding author:**

Debashis Sahoo, Ph.D.; Associate Professor, Department of Pediatrics, University of California San Diego; 9500 Gilman Drive, MC 0703, Leichtag Building 132; La Jolla, CA 92093-0831. **Phone:** 858-246-1803: **Fax:** 858-246-0019: **Email:** dsahoo@ucsd.edu

Soumita Das, Ph.D.; Associate Professor, Department of Pathology, University of California San Diego; 9500 Gilman Drive, George E. Palade Bldg, Rm 256; La Jolla, CA 92093. **Phone**: 858-246-2062: **Email:** sodas@ucsd.edu

Pradipta Ghosh, M.D.; Professor, Departments of Medicine and Cell and Molecular Medicine, University of California San Diego; 9500 Gilman Drive (MC 0651), George E. Palade Bldg, Rm 232; La Jolla, CA 92093. **Phone**: 858-822-7633: **Email:** prghosh@ucsd.edu

GRAPHIC ABSTRACT



In Brief: The authors use machine learning approaches to identify universally relevant definition of macrophage polarization states and create a predictive framework for developing macrophage-targeted precision diagnostics and therapeutics.

Highlights:

- *S*ignatures of *Ma*crophage *R*eactivity and *T*olerance were identified using AI
- *SMaRT* genes are relevant across tissues, organs, species and immune cells
- *SMaRT* genes identify physiologic and pathologic macrophage states
- *SMaRT* genes can prognosticate outcome and serve as therapeutic targets

Summary/Abstract (n = 170)

A continuum of macrophage polarization states is essential for the initiation, maintenance, and resolution of inflammation. We built a network using pooled human macrophage transcriptomic datasets and used machine learning algorithms to identify the path of such continuum states. One path, comprised of 338 genes emerged as the best; it accurately identified both physiologic and pathologic spectra of "*reactivity*" and "*tolerance*", and remained relevant across tissues, organs, species and immune cells (> 12,500 diverse datasets). This 338-gene signature identified macrophage polarization states in physiology and across diverse human diseases and objectively analyzed the appropriateness of mice as pre-clinical models for such diseases. The signature consistently outperformed conventional signatures in the degree of transcriptome-proteome overlap and in prognosticating outcomes across diverse acute and chronic diseases, e.g., sepsis, liver fibrosis, aging and cancers. Crowd-sourced genetic and pharmacologic studies confirmed that model-rationalized interventions trigger predictable macrophage fates. These findings provide a formal and universally relevant definition of macrophage states and a predictive framework for the scientific community to develop macrophage-targeted precision diagnostics and therapeutics.

INTRODUCTION

Macrophages are complex; as sentinel cells of the innate immune system, they are found in various organs and their dysregulated activation can directly impact organ functions and the outcome of all diseases (Murray and Wynn 2011; Pollard 2009). Macrophages were initially classified as M1 (the classically activated macrophages) and M2 (the alternatively activated macrophages) based on their functions at the extremes of polarization states (Mills et al. 2000). However, the M1/M2 nomenclature is considered as too simplistic; it fails to describe the diverse, polyfunctional and plastic cells, and the myriad of continuum states that they adopt in the tissue at steady-state and during disease (Amit et al. 2016; Ginhoux et al. 2016; Glass and Natoli 2016; Okabe and Medzhitov 2016). To cope with this degree of diversity and plasticity, several definitions of macrophage subtypes have emerged, each representing specialized contexts, e.g., TAMs, tumor-associated macrophages (Qian and Pollard 2010); LAMs, lipid-associated macrophages in atherosclerosis (Jaitin et al. 2019); DAMs, disease-associated microglia in neurodegenerative disorders (Keren-Shaul et al. 2017); SAMs, scar-associated macrophages in liver fibrosis (Duffield et al. 2005; MacParland et al. 2018; Ramachandran et al. 2019). These definitions were geared to identify divergent markers, spatial localization, origin, and functional pathways associated with macrophages during disease; however, they fall short in predictive or prognostic abilities.

We sought to create and validate a comprehensive model of macrophage processes for defining, tracking, and even predicting macrophage fate after perturbation (see **Fig 1a**, **S1A** for workflow outline). We hypothesized that such a model might inspire *formal* definitions for macrophage polarization states that are reflective of fundamental processes and maintain relevance across tissues, organs, diseases and species. In addition, it may also rationalize diagnostics and therapeutics to detect and reset, respectively, deranged macrophage states in disease. We show that such formal definition(s) of macrophage states is not only possible, but also provide evidence for their usefulness in prediction and prognostication.

RESULTS AND DISCUSSION

A computational model of continuum states in macrophage processes

We chose a Boolean approach to build transcriptomic network (Sahoo et al. 2008); this approach has been used to create maps of evolving cellular states along any disease continuum and identify cellular states in diverse tissues and contexts with high degrees of precision (see detailed *Methods*). The Boolean approach relies on invariant relationships that are conserved despite heterogeneity in the samples used for the analysis. Invariant relationships among pairs of genes that are conserved across samples representative of maximum possible diversity, i.e., irrespective of their origin (normal or disease), laboratories/cohorts, different perturbations, are assumed to be fundamentally important for any given process.

For model training and development, we used a pooled all-human microarray dataset that included 197 manually annotated heterogeneous macrophage datasets from GEO (GSE134312 (Dang et al. 2020); **Fig 1a-c**; **Fig S1A-B;** see **Supplemental Information 1** for catalog of datasets). These datasets contained primary tissue-derived macrophages (both healthy and diseased tissues) and cultured macrophage cell lines (e.g., THP1), either untreated or treated with diverse sets of ligands that are known to induce either M1 (n = 13) or M2 (n = 8) polarized states (see **Table S1**).

A graph (**Fig 1d**; **Fig S1C; Fig S2A**) is built, comprised of gene clusters (nodes) connected to each other using Boolean implication relationships (edges). The network displayed scale-free properties, as expected (**Fig S1D**). We oriented ourselves to the resultant network by querying and locating the known 'M1/M2' samples; the 'M1' samples segregated towards one end, and 'M2' samples on the other, implying that the paths of connected clusters within the resultant network represent a continuum of cellular states in macrophages within the immunologic spectrum (**Fig S1E-G**). Reactome pathway analyses (Fabregat et al. 2018) of each cluster along the top continuum paths revealed a multitude of cellular processes that are impacted during macrophage polarization (**Fig 1e;** Gene clusters and reactome pathways can be queried at: http://hegemon.ucsd.edu/SMaRT/).

Identification of signatures of macrophage 'reactivity' and 'tolerance' (SMaRT)

Using machine learning approaches, various interconnected gene clusters (i.e., Boolean paths) were assessed for their ability to accurately classify the samples (based on the genes in the clusters and computing a weighted average of gene expression values outlined in **Fig S2B**) (**Fig 1f**). A multivariate analysis of the top five Boolean paths revealed that the path connecting clusters(C)#13→14→3 is the best ($p < 0.001$) at discriminating M1 (ROC-AUC 0.98) and M2 (ROC-AUC 0.99) (**Fig 1f; Fig S2C**). Path #13→14→3 was subsequently validated in five other independent datasets (**Fig 1g**). A comparative analysis of #13→14→3 path *vs*. other traditional approaches, e.g., Differential Expression (Becker et al. 2015), Correlation Network (Becker et al. 2015), Hierarchical Clustering (Coates et al. 2008) and Differential and interactome analyzes (Martinez et al. 2006) showed the superiority of the *BoNE*-derived path in separating M0-M1-M2 states. The Boolean path matched differential expression in its ability to distinguish M1 state, while exceeding the remaining traditional approaches (**Fig 1h**). A heatmap of the pattern of gene expression in each cluster in M0-M1-M2 states is shown in **Fig 1i**.

Furthermore, C#13 predicted M1 perfectly (ROC-AUC = 1.0); the path #14→3 predicted M2 close to perfection (ROC-AUC = ranging from 0.80 to 1.00) in all cohorts tested (**Fig 1j**). This indicates that while the path #13→14→3 is the most accurate path across all human macrophage-derived datasets collected and analyzed, C#13 and the path #14→3 carry relevant information on macrophage states independently of each other. C#13 is associated with M1-like state and expression of these genes is predicted to reflect the extent of "immunoreactivity" of macrophages. Path #14→3 is associated with a M2-like state and expression of these genes is predicted to reflect the extent of "immunotolerance". We define the two distinct macrophage polarization states in physiology as "reactive" and "tolerant" based on basal C#13 and #14→3 scores, respectively (**Fig. 1k**). Four additional macrophage states could also exist, presumably in disease states, i.e., hyperreactive (high C#13), hypertolerant (high #14→3), hyporeactive (low C#13), and hypotolerant (low #14→3) (**Fig. 1k**). Henceforth, we refer to these genes as signatures of *ma*crophage *r*eactivity and *t*olerance, abbreviated as '*SMaRT*' (See http://hegemon.ucsd.edu/SMaRT/ and **Supplemental Information 2** for the list of genes).

We tested the independence of M1/M2 signatures using single cell RNASeq dataset GSE150708 (human), GSE117176 (mouse), where we artificially created pseudobulk samples using various mixtures of M1 and M2 cells (**Fig S2D-H**). M1 and M2 cells were categorized as both tolerant and reactive using C#13 and path #14-3 signatures (**Fig S2D**). Mouse single cell dataset GSE117176 showed the same patterns after revising the genes from path #14-3 as many genes were not captured by scRNASeq (**Fig S2E-H**).

SMaRT genes are relevant across tissues, organs, species and immune cells

We found that the path #13→14→3 successfully identified M1/M2-polarization states in diverse tissue-resident macrophages (brain-resident microglia, the Langerhan's cells in the skin, intestinal and lung alveolar macrophages, etc.), in both humans and mice (**Fig 2a-b**). See **Supplemental Information 1** for the degree of heterogeneity represented in these datasets. Surprisingly, the path could also separate reactive and tolerant states of other immune cells, including lymphocytes (B/T and NK-T), natural killer (NK) cells, neutrophils, dendritic, basophils, eosinophils and mast cells (**Fig 2c**). Together, these findings indicate that the *SMaRT*-based definitions of 'reactivity' and 'tolerance' remain relevant in the context of tissue-resident macrophages despite their adaptation to the tissue and/or organ-specific microenvironment for their identity (Gordon and Pluddemann 2017; Lavin et al. 2014; Stout and Suttles 2004). These definitions also maintain relevance in mice, whose immune system is different from ours (Mestas and Hughes 2004). Findings suggest that the *SMaRT*-based definitions may reflect the *fundamental* immune-reactive and tolerant gene regulatory mechanisms that are shared among diverse cells in our immune system, regardless of whether they are derived from the myeloid or lymphoid lineage (**Fig 2d**).

The network captures physiologic macrophage states and functions

We found that diverse macrophage subtypes are represented within our model of macrophage processes (**Fig 2e**). The classical M1 subtype was represented in C#1 and #13 on the reactive end of the model, alongside TCR+ macrophages in C#1 and #12; the latter is known to release CCL2 and have high phagocytic abilities (Chavez-Galan et al. 2015). On the tolerant end of the model we found the TAMs in

C#2, #5, #6 and the *CD169*+ macrophages in C#2, #3 and #7; both subtypes have been implicated in immunological tolerance (Liu and Cao 2015; Ravishankar et al. 2014; Saunderson et al. 2014). As one would anticipate, the tissue-resident macrophages (M2a-d) that are known for their plasticity of polarization states were more centrally placed in C#2 and #5. Finally, gene signatures of scar-associated non-inflammatory (ni) macrophages that restrict inflammation in liver cirrhosis (SAM B (Ramachandran et al. 2019) and SAM ni (MacParland et al. 2018), **Fig 2f**) and damage-associated microglia (DAMs (Keren-Shaul et al. 2017); **Fig 2f**) that restrict the progression of neurodegeneration significantly overlapped with the tolerant clusters C#14 and #3. A gene signature that was recently shown to be induced in monocytes and macrophages in all viral pandemics (ViP), past and present, overlapped with the reactive C#13 as expected (**Supplemental Information 2** lists all gene signatures in **Fig 2f**).

Members of the family of pattern recognition receptors (PRRs; **Table S2**), *via* which macrophages 'sense' its surroundings (Zhou et al. 2015), were distributed in various nodes within the model, overlapping with each other (**Fig 2g**). PRRs that sense pathogens or apoptotic cells to stimulate phagocytosis and mediate inflammation, e.g., toll-like (TLRs), nucleotide oligomerization domain (NODs) and receptor for advanced glycation end products (RAGE) were found on the 'reactive' side of the model. The TLRs, scavengers and C-type lectins also overlapped with path#13→14→3, but only on the tolerant end (cluster #3) of the spectrum.

The circadian genes were distributed within clusters along a path (#1→2→3→4) (**Fig 2h**), intersecting at the tolerant end of the path#13→14→3, i.e., C#3. The daytime circadian genes were in the reactive end of the model and showed an inverse high=>low Boolean relationship with night-time circadian genes; the latter were mostly in the tolerant end of the model (**Fig S3A-C**). This finding is consistent with the current belief that macrophages 'kill' (reactive) during the day and 'heal' (tolerant) during the night (Early and Curtis 2016). We also show that the performance of the tolerant signature (C#14-3) in diseases that have an intricate relationship with circadian rhythms, such as metabolic syndrome (Eckel-Mahan and Sassone-Corsi 2013), can be further improved by normalization based on a clock gene or clock gene signature (**Fig S4**).

SMaRT genes identify pathologic polarization states in diseases

We next asked how the Boolean network-derived *formal* definitions perform in disease states. A plethora of disease conditions and tissues were analyzed (**Fig 3a-n; Supplemental Information 1**). We computed a composite immune response score derived from C#13 alone or C#14 and #3, which quantitatively estimates the degree of "reactivity" and "tolerance", respectively, and tested it in diverse conditions. An analysis of full-thickness colon tissues representing the 2 major subtypes of inflammatory bowel disease (IBD), ulcerative colitis (UC) and Crohn's disease (CD) (**Fig S5A**) revealed that reactivity is a common feature in both UC and CD (**Fig 3a**, *top*; **Fig S5B**-*left*). However, tolerance was enhanced only in CD (**Fig S5B**-*right*), which is consistent with the notion that 'alternatively' activated tolerant macrophages may drive the transmural nature of the inflammation, ineffective bacterial clearance, and accompanying tissue remodeling (fibrosis, stricture, fistula), all features that are observed uniquely in CD (Cho 2008), but not UC.

Reactivity alone could prognosticate outcome (i.e., segregate responder *vs* non-responder) regardless of the heterogeneity of the UC cohorts and the diverse treatment modalities (**Fig S5C-D**), consistent with the widely-accepted notion that hyperinflammatory macrophages are drivers (Steinbach and Plevy 2014) of the disease and key targets for therapeutics (Peters et al. 2017). Insufficient datasets precluded similar analyses in the case of CD.

We also found that "reactivity" and "tolerance" differs along the length of the colon crypt—the surface is more reactive, whereas the stem-cell niche at the bottom is more "tolerant" (**Fig 3b**; **Fig S5E-F**). We also found that "hypo-reactivity" [low C#13] and "complete tolerance" [high #14→3] are two states that are progressively accentuated during colorectal carcinoma (CRC) initiation and the emergence of chemoresistance (**3c**; **Fig S5G-H**). Consistent with the fact that most of the CRCs are found located in the left (distal) colon and microbe-driven risk is high in that segment (Drewes et al. 2016), we found that segment to be more tolerant than the right (proximal) segment (**3d**).

We detected altered macrophage states during the initiation and progression of several human other diseases, ranging from arthritis, through neurodegenerative diseases to viral pandemics (see **Fig 3e-n; Fig S6A-N, Fig S7A-E**). Our definitions for "reactivity" and "tolerance" could accurately identify the underlying pathologic macrophage states implicated in each condition. Together, these results show that the *BoNE*-derived signature can detect different subsets of macrophages are essential to the pathogenesis of many diseases. Findings also agree with the notion that disease chronicity is invariably associated with mixed polarization states (whose detection has largely been enabled by scSeq studies) where each state plays an opposing (balanced) role (Duffield et al. 2005; Jaitin et al. 2019; Keren-Shaul et al. 2017; MacParland et al. 2018; Murray and Wynn 2011; Qian and Pollard 2010; Ramachandran et al. 2019).

SMaRT genes rationalize the choice of mouse models

Although mice are the preferred model species for research (Rosenthal and Brown 2007), most agree that their innate immune systems differ (Mestas and Hughes 2004). C57BL/6J and Balb/c mice are two most commonly used mouse strains that differ in their immune responses, giving rise to distinct disease outcomes, which in turn rationalizes their use as pre-clinical models for human diseases (**Fig 3o**). Our signature successfully classified the macrophages from these two strains in three independent cohorts (Howes et al. 2016; Link et al. 2018) (**Fig 3p**); C57BL/6 emerged as more reactive and Balb/c as more tolerant (**Fig 3q**). These findings are consistent with the observation that BALB/c mice are more susceptible to a variety of pathogens (Mainou-Fowler et al. 1988; Sacks and Noben-Trauth 2002; Schluter et al. 1999), and are useful for modeling tumor initiation and progression and for making antibodies. By contrast, C57BL/6 mice are resistant to infections and are the most common strain used for modeling inflammatory diseases, e.g., arthritis, metabolic disorders [NASH, atherosclerosis, etc. (Champy et al. 2008; Ishida et al. 1991; Toye et al. 2005)]. We conclude that the model-derived definitions for "reactivity" and "tolerance" —(i) capture the contrasting immunophenotypes of these two murine strains previously reported by Mills et al., (Mills et al. 2000) and (ii) rationalize the choice of each strain as preferred models for modeling a unique

set of human diseases. Findings also suggest that the model-derived signatures could serve as an objective guide for assessing the appropriateness of any species/strains/sub-strains as pre-clinical models.

SMaRT genes carry diagnostic value

Next we compared head-to-head the diagnostic and prognostic potential of the newly defined polarization states against four traditional definitions: differential expression analysis (Becker et al. 2015) (DExp), correlation network (Bell et al. 2016) (CorN), hierarchical clustering + fold change (Coates et al. 2008) (HiClu), and differential + interactome analysis (Martinez et al. 2006) (Dif+Int). A composite immune response score derived from C#13 alone, which quantitatively estimates the degree of "reactivity" was tested on multiple datasets generated from tissues derived from patients with known clinically relevant diagnoses. A hyper-reactive state was invariably associated with graft rejection in transplanted hearts, livers and kidneys (**Fig 3r**). A 'hyper-reactive' state also classified IBD-afflicted children from those with non-IBD indications (8-18 y age) with reasonable accuracy in a prospective study where the blood samples were drawn at the time of diagnostic colonoscopy (**Fig 3r**). Among the critically ill patients in the ICU, a hyper-reactive state was associated with better 28-day survival for those with ARDS on ventilators (**Fig 3r**) and improved survival without the need for liver transplantation in those diagnosed with Tylenol-induced acute liver failure (**Fig 3r**). While some of the four other traditional methodologies fared similar to the new definitions in some cohorts, none performed as well, and/or as consistently. Findings suggest that the *BoNE*-derived signatures may capture fundamental aspects of macrophage polarization that drive disease states.

SMaRT genes can prognosticate outcome

We next computed a composite immune response score based on either the path #13-14-3 or C#13 alone. When used as a composite score, a low score value represents "reactive" and high score value represent "tolerant" states. This signature was tested on all transcriptomic datasets found on the NCBI GEO database (as of *04/2022*) originating from prospective studies, regardless of disease. Prospective studies were chosen because they rarely have selection bias from enrollment procedures because the outcomes have not yet occurred at the time of enrollment. In the context of cancers, "reactive" tumors carried a worse prognosis than "tolerant" ones across a variety of solid tumor subtypes, e.g., colorectal (n = 555; **Fig 4a**), breast, pancreas, prostate, glioblastoma and bladder cancers (**Fig S7F**). Undetectable by any of the traditional methodologies, these findings are consistent with the well-recognized role of inflammatory cells in the tumor microenvironment (Coussens and Werb 2002).

In a cohort of 216 patients with HCV-related liver fibrosis, overall survival was reduced among patients with a "reactive" signature on their liver biopsies compared to those with a "tolerant" signature (**Fig 4b**). Again, undetectable by any of the traditional methodologies, these findings are consistent with the known role of activated macrophages in chronic liver injury, inflammation and fibrosis (Ehling et al. 2014; Heinrichs et al. 2011; Kazankov et al. 2014; Sunami et al. 2012).

In a cohort of 802 patients with sepsis, 28-day mortality was worse among those with a "tolerant" signature compared to those with a "reactive" signature (**Fig 4c**). This finding is consistent with the notion

that "endotoxin tolerance" during sepsis carries poor outcome (Pena et al. 2014). Two of the four traditional models, correlation network and hierarchical clustering, also performed reasonably well in sepsis, which is not surprising because 30% of the reactivity signature within C#13 overlaps with the 'M1'-state definition in some of the traditional methods.

In a cohort of 114 patients with idiopathic pulmonary fibrosis (IPF), an incurable disease that is characterized by progressive fibrosis requiring lung transplantation (George et al. 2011), a "reactive" signature was associated with shorter transplant-free survival (**Fig 4d**). Two of the four traditional models, correlation network and differential + interactome analysis, also performed reasonably well. Results are in keeping with the widely accepted notion that proinflammatory pulmonary macrophages are known to drive inflammation and fibrosis in the lung (Byrne et al. 2016).

Among 517 recipients of kidney transplants, a "reactive" signature was associated with increased graft loss in two independent cohorts (**Fig 4e-f**). Findings are in keeping with prior body of work implicating inflammatory macrophages (both number and extent of activation) as culprits in both acute and chronic allograft rejection and graft loss (Azad et al. 2018; Bergler et al. 2016; Liu et al. 2016). Two of the four traditional models, differential expression and correlation network approaches, performed reasonably well in one cohort (**Fig 4e**), but none reached significance in the other (**Fig 4f**).

Finally, among 151 nonagenarians in the Vitality 90+ study (Nosraty et al. 2015), a "reactive" signature was associated with higher mortality in men (**Fig 4g**-*left*), but not women (**Fig 4g**-*right*). Results are in keeping with the fact that the plasma levels of the 'classical' marker of inflammaging, i.e., interleukin-6 (IL-6) and a pro-inflammatory gene signature in PBMCs were correlated in men, whereas no correlations were observed in women (Nevalainen et al. 2015). None of the traditional methodologies could detect this gender-specific difference, nor did they prognosticate survival.

These findings demonstrate a degree of robustness and consistency in the prognostic ability of the newly defined signatures of macrophage polarization across diverse diseases and independent datasets.

SMaRT genes are significantly enriched in the macrophage proteome

We used Tandem Mass Tag (TMT) proteomics datasets from THP1-derived macrophages (M0, PMA) that were polarized to M1-M2 states (see workflow **Fig 5a**) and asked if the *BoNE*-derived gene clusters are translated to proteins. We found that the *BoNE*-derived *SMaRT* genes were induced significantly in the THP1 proteome (**Supplemental Information 3**). Consistent with our hypothesis that C#13 and path #14→3 carry independent information regarding "reactivity" and "tolerance", we found that LPS and IFNγ-induced M1 polarization was associated with significant differential translation of genes in C#13 (**Fig 5b**-*top*), whereas IL4-induced polarization was associated with significant differential translation of genes in C#14 and C#3 (**Fig 5b**-*bottom*). Such differential protein translation continued to take place over 24 h (**Fig 5b**).

Comparative analyses showed that while the "reactivity" signatures identified by two other conventional methodologies--Differential Expression and Correlation Network-- also reached significance; **Fig 5b**-*top*), "tolerance" signatures derived by all other conventional approaches did not (**Fig 5b**-*bottom*).

Heatmaps show the dynamic and opposing nature of the proteins translated by the genes within the *BoNE*-derived gene signatures during polarization (**Fig 5c-d**).

Findings demonstrate that the gene signatures of 'reactivity' and 'tolerance' identified here are significantly represented also in the translated proteome.

Perturbation of SMaRT genes results in predictable outcomes

We next asked if network-rationalized interventions result in predictable outcomes upon perturbation, e.g., gene depletion (CRISPR, shRNA, KO mice) or overexpression, expression of functionally defective mutants, or chemical agonists/inhibitors. To this end, we carried out real-world crowdsourcing experiments on macrophage datasets in which interventions were conducted by different groups using diverse manipulations (**Fig 6a**). Depletion or pharmacologic inhibition of any gene in C#13 was predicted to suppress reactivity and enhance tolerance, whereas overexpression or pharmacologic stimulation of the same should have an opposite impact, i.e., enhance reactivity and suppress tolerance. Similarly, depletion/inhibition of any gene in C#14 was predicted to enhance reactivity and suppress tolerance (**Fig 6b**, *left*; **Table S3**). The depletion of genes in C#3 is predicted to not have a robust impact the network because of the Low=>Low relationship with C#14.

We began with the ENCODE portal (Davis et al. 2018), a resource that was born out of the larger initiative called the ENCODE integrative analysis (Consortium 2012); it is an encyclopedia of large, unbiased shRNA library screen on the human K562 chronic myeloid leukemia cell line. This dataset contained 4 of the 137 genes in C#14 and none from C#13 (Davis et al. 2018). In all 4 cases, the depletion of genes in C#14 resulted in the predicted outcome of enhanced reactivity and hypotolerance (**Fig 6b**, *right*). A systematic search of the NCBI GEO database also revealed 16 other independent datasets reporting the impact of interventions on genes in C#13 (9 datasets) and C#14 (7 datasets) (**Table S3**). Regardless of the heterogeneous nature of the interventions and lab-to-lab variations in the type of cells/tissues used, predictions matched the observed outcomes in each instance. At least in one instance (i.e., STAT3), we could confirm the alignment of phenotypes between gene deletion and pharmacologic inhibition, implying that both approaches must have converged on the same biology. Because such alignment and/or convergence is seen in many instances (Weiss et al. 2007), findings suggest that the current model can accurately guide outcome-driven pharmacologic interventions.

Together, these crowd-sourced studies rigorously and independently validate the definitions of macrophage polarization states; the fundamental nature of these definitions appear to remain relevant despite the thunderous heterogeneity of models and methods used by so many.

CONCLUSIONS

The lack of consensus on how to define macrophage activation has impeded progress in multiple ways; despite a panoply of existing descriptors, most remain contentious and/or confusing. AI-guided gene expression signatures presented here, *SMaRT*, offers a set of standardized definitions of macrophage polarization that encompasses four principles: (i) they are comprised of an unbiased collection of markers of macrophage activation that are represented in both the transcriptome and the proteome; (ii) they remain meaningful and relevant regardless of the source of macrophages (i.e., bone marrow, circulation, tissue-resident); (iii) they perform well across diverse activators, both in vitro and in vivo (i.e., recombinant ligands and cytokines, microbes, or multifactorial, as in the setting of complex disease states), and (iv) they provide a predictive framework that can be exploited for diagnostic purposes and for outcome-rationalized therapeutic interventions. These principles unify experimental standards for diverse experimental scenarios and interpretations across diverse tissues and diseases.

Finaly, these *SMaRT* genes provide a common framework for macrophage activation nomenclature, which should enable laboratories to detect and report a given immunophenotype of macrophage in a standardized way. Standardization is expected to spur the development of robust strategies to address the multitude of macrophage-related disorders. It also serves as a starting point for the development of new diagnostics and immunomodulatory therapies.

Acknowledgements

We thank Gordon Gill and Christopher K. Glass for critiques and suggestions during manuscript preparation. This work was supported by the National Institutes for Health (NIH) grant R01-AI155696 (to PG, DS and SD). Other sources of support include: R01-GM138385 (to DS), R01-AI141630, CA100768 and CA160911 (to P.G), R01-DK107585 (to S.D) and UG3TR002968 (to D.S., S.D and P.G). D.S was also supported by two Padres Pedal the Cause awards (Padres Pedal the Cause/RADY #PTC2017 and San Diego NCI Cancer Centers Council (C3) #PTC2017). D.S, P.G and S.D were also supported by the Leona M. and Harry B. Helmsley Charitable Trust.

Author Contributions:

Conceptualization: D.S, P.G, S.D

Methodology: D.S, S.S, D.V., S.T., D.D.

Investigation: D.S, S.S, P.G

Visualization: D.S, P.G, S.S, G.D.K, D.V.

Funding acquisition: D.S, S.D, P.G

Project administration: D.S, S.D, P.G

Supervision: D.S, P.G

Writing – original draft: D.S, P.G

Writing – review & editing: D.S, P.G., S.D, G.D.K, S.S

Competing interests:

Authors declare that they have no competing interests.

Data and materials availability:

All data are available in the main text or the supplementary materials. A website (http://hegemon.ucsd.edu/SMaRT/) of the macrophage network is built to support interactive query.

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Pooled human macrophage array	NCBI GEO (The National Center for Biotechnology Information- Gene expression omnibus)	GSE134312
Ccdc88a KO peritoneal macrophages		GSE203423
Proteomics dataset, reanalyzed from PMID: 34731634	MassIVE repository	MSV000084672
Experimental models: Organisms/strains		
Ccdc88a fl/fl LysMCre/- mice	PMID: 33055214	
Software and algorithms		
Numpy	Python	https://numpy.org
Scipy	Python	https://scipy.org
Seaborn	Python	https://seaborn.pydata.org
Matplotlib	Python	https://matplotlib.org
Hierarchical Exploration of Gene Expression Microarrays Online (Hegemon)	HTML, JavaScript, Python, PHP	https://github.com/sahoo00/Hegemon
Boolean Network Explorer (BoNE)	Python	https://github.com/sahoo00/BoNE
Other		
Interactive website	This paper	http://hegemon.ucsd.edu/SMaRT/

Materials and Methods

Data Collection and Annotation

Publicly available microarray and RNASeq databases were downloaded from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) website (Barrett et al. 2005; Barrett et al. 2013; Edgar et al. 2002). Gene expression summarization was performed by normalizing Affymetrix platforms by RMA (Robust Multichip Average) (Irizarry et al. 2003a; Irizarry et al. 2003b) and RNASeq platforms by computing TPM (Transcripts Per Millions) (Li and Dewey 2011; Pachter 2011) values whenever normalized data were not available in GEO. We used log2(TPM) if TPM > 1 and (TPM – 1) if TPM < 1 as the final gene expression value for analyses. We also used log2(TPM + 1) in some datasets. We also used publicly data normalized using RPKM (Mortazavi et al. 2008), FPKM (Trapnell et al. 2009; Trapnell et al. 2010), TPM (Li et al. 2010; Wagner et al. 2012), and CPM (Law et al. 2016; Robinson et al. 2010). In the context of Affymetrix microarray data we believe that RMA works better than MAS 5.0 (Pandey and Sahoo 2019).

Macrophage datasets used for network analysis

Previously published pooled macrophage dataset from GEO (GSE134312, n = 197) assayed on the Human U133 Plus 2.0 (GPL570), Human U133A 2.0 (GPL571) and Human U133A (GPL96) platforms were used to perform macrophage network analysis. This dataset was manually annotated with M0, M1 or M2 phenotypes. Accession numbers for the M0, M1 and M2 phenotypes are presented in table S4. Five validation datasets are used to test the macrophage gene signature: GSE35449 (7 M0, 7 M1, 7 M2), GSE46903 (64 M0, 29 M1, 40 M2), GSE61298 (6 M0, 6 M1, 6 M2), GSE55536 human peripheral blood mononuclear cell-derived macrophage (6 M0, 6 M1, 6 M2), GSE55536 iPSC derived macrophages (3 M0, 3 M1, 3 M2). See **Supplementary Information 1** for all datasets analyzed in this work.

Computational Approaches

StepMiner Analysis

StepMiner is a computational tool that identifies step-wise transitions in a time-series data (Sahoo et al. 2007). StepMiner performs an adaptive regression scheme to identify the best possible step up or down based on sum-of-square errors. The steps are placed between time points at the sharpest change between low expression and high expression levels, which gives insight into the timing of the gene expression-switching event. To fit a step function, the algorithm evaluates all possible step positions, and for each position, it computes the average of the values on both sides of the step for the constant segments. An adaptive regression scheme is used that chooses the step positions that minimize the square error with the fitted data. Finally, a regression test statistic is computed as follows:

$$F\ stat = \frac{\sum_{i=1}^{n}(\hat{X}_i - \bar{X})^2/(m-1)}{\sum_{i=1}^{n}(X_i - \hat{X}_i)^2/(n-m)}$$

Where X_i for $i = 1$ to n are the values, $\widehat{X_i}$ for $i = 1$ to n are fitted values. m is the degrees of freedom used for the adaptive regression analysis. $\bar{X}$ is the average of all the values: $\bar{X} = \frac{1}{n} * \sum_{j=1}^{n} X_j$. For a step position at k, the fitted values $\widehat{X_l}$ are computed by using $\frac{1}{k} * \sum_{j=1}^{n} X_j$ for $i = 1$ to k and $\frac{1}{(n-k)} * \sum_{j=k+1}^{n} X_j$ for $i = k + 1$ to n.

Boolean Analysis

Boolean logic is a simple mathematic relationship of two values, i.e., high/low, 1/0, or positive/negative. The Boolean analysis of gene expression data requires the conversion of expression levels into two possible values. The ***StepMiner*** algorithm is reused to perform Boolean analysis of gene expression data (Sahoo et al. 2008). **The Boolean analysis** is a statistical approach which creates binary logical inferences that explain the relationships between phenomena. Boolean analysis is performed to determine the relationship between the expression levels of pairs of genes. The ***StepMiner*** algorithm is applied to gene expression levels to convert them into Boolean values (high and low). In this algorithm, first the expression values are sorted from low to high and a rising step function is fitted to the series to identify the threshold. Middle of the step is used as the StepMiner threshold. This threshold is used to convert gene expression values into Boolean values. A noise margin of 2-fold change is applied around the threshold to determine intermediate values, and these values are ignored during Boolean analysis. In a scatter plot, there are four possible quadrants based on Boolean values: (low, low), (low, high), (high, low), (high, high). A Boolean implication relationship is observed if any one of the four possible quadrants or two diagonally opposite quadrants are sparsely populated. Based on this rule, there are six kinds of Boolean implication relationships. Two of them are symmetric: equivalent (corresponding to the positively correlated genes), opposite (corresponding to the highly negatively correlated genes). Four of the Boolean relationships are asymmetric, and each corresponds to one sparse quadrant: (low => low), (high => low), (low => high), (high => high). BooleanNet statistics (**Fig. 2a**) is used to assess the sparsity of a quadrant and the significance of the Boolean implication relationships (Sahoo et al. 2008; Sahoo et al. 2010). Given a pair of genes A and B, four quadrants are identified by using the StepMiner thresholds on A and B by ignoring the Intermediate values defined by the noise margin of 2 fold change (+/- 0.5 around StepMiner threshold). Number of samples in each quadrant are defined as a_{00}, a_{01}, a_{10}, and a_{11} (Figure 1A) which is different from X in the previous equation of F stat. Total number of samples where gene expression values for A and B are low is computed using the following equations.

$$nA_{low} = (a_{00} + a_{01}), nB_{low} = (a_{00} + a_{10}),$$

Total number of samples considered is computed using following equation.

$$total = a_{00} + a_{01} + a_{10} + a_{11}$$

Expected number of samples in each quadrant is computed by assuming independence between A and B. For example, expected number of samples in the bottom left quadrant $e_{00} = \hat{n}$ is computed as probability of

A low ((a_{00} + a_{01})/total) multiplied by probability of B low ((a_{00} + a_{10})/total) multiplied by total number of samples. Following equation is used to compute the expected number of samples.

$$n = a_{ij}, \hat{n} = (nA_{low}/total * nB_{low}/total) * total$$

To check whether a quadrant is sparse, a statistical test for ($e_{00} > a_{00}$) or ($\hat{n} > n$) is performed by computing S_{00} and p_{00} using following equations. A quadrant is considered sparse if S_{00} is high ($\hat{n} > n$) and p_{00} is small.

$$S_{ij} = \frac{\hat{n} - n}{\sqrt{\hat{n}}}$$

$$p_{00} = \frac{1}{2}\left(\frac{a_{00}}{(a_{00} + a_{01})} + \frac{a_{00}}{(a_{00} + a_{10})}\right)$$

A suitable threshold is chosen for S_{00} > sThr and p_{00} < pThr to check sparse quadrant. A Boolean implication relationship is identified when a sparse quadrant is discovered using following equation.

Boolean Implication = (S_{ij} > sThr, p_{ij} < pThr)

A relationship is called Boolean equivalent if top-left and bottom-right quadrants are sparse.

Equivalent = $(S_{01} > sThr,\ P_{01} < pThr, S_{10} > sThr, P_{10} < pThr)$

Boolean opposite relationships have sparse top-right (a_{11}) and bottom-left (a_{00}) quadrants.

Opposite = $(S_{00} > sThr,\ P_{00} < pThr, S_{11} > sThr, P_{11} < pThr)$

Boolean equivalent and opposite are symmetric relationship because the relationship from A to B is same as from B to A. Asymmetric relationship forms when there is only one quadrant sparse (A low => B low: top-left; A low => B high: bottom-left; A high=> B high: bottom-right; A high => B low: top-right). These relationships are asymmetric because the relationship from A to B is different from B to A. For example, A low => B low and B low => A low are two different relationships.

A low => B high is discovered if the bottom-left (a_{00}) quadrant is sparse and this relationship satisfies following conditions.

A low => B high = $(S_{00} > sThr,\ P_{00} < pThr)$

Similarly, A low => B low is identified if the top-left (a_{01}) quadrant is sparse.

A low => B low = $(S_{01} > sThr,\ P_{01} < pThr)$

A high => B high Boolean implication is established if the bottom-right (a_{10}) quadrant is sparse as described below.

A high => B high = $(S_{10} > sThr,\ P_{10} < pThr)$

Boolean implication A high => B low is found if the top-right (a_{11}) quadrant is sparse using following equation.

A high => B low = $(S_{11} > sThr,\ P_{11} < pThr)$

For each quadrant a statistic S_{ij} and an error rate p_{ij} is computed. S_{ij} > sThr and p_{ij} < pThr are the thresholds used on the BooleanNet statistics to identify Boolean implication relationships.

Boolean analyses in the test dataset GSE134312 uses a threshold of sThr = 3 and pThr = 0.1. These thresholds are exactly same as the previously used thresholds sThr = 3 and pThr = 0.1 for BooleanNet (Dabydeen et al. 2019; Pandey and Sahoo 2019; Sahoo et al. 2008).

Boolean Network Explorer (BoNE)

Boolean network explorer (BoNE) provides an integrated platform for the construction, visualization and querying of a network of progressive changes underlying a disease or a biological process in three steps (**Fig S1A**): First, the expression levels of all genes in these datasets were converted to binary values (high or low) using the StepMiner algorithm. Second, gene expression relationships between pairs of genes were classified into one-of-six possible Boolean Implication Relationships (BIRs), two symmetric and four asymmetric, and expressed as Boolean implication statements. This offers a distinct advantage from conventional computational methods (Bayesian, Differential, etc.) that rely exclusively on symmetric linear relationships in networks. The other advantage of using BIRs is that they are robust to the noise of sample heterogeneity (i.e., healthy, diseased, genotypic, phenotypic, ethnic, interventions, disease severity) and every sample follows the same mathematical equation, and hence is likely to be reproducible in independent validation datasets. Third, genes with similar expression architectures, determined by sharing at least half of the equivalences among gene pairs, were grouped into clusters and organized into a network by determining the overwhelming Boolean relationships observed between any two clusters. In the resultant Boolean implication network, clusters of genes are the nodes, and the BIR between the clusters are the directed edges; BoNE enables their discovery in an unsupervised way while remaining agnostic to the sample type.

Statistical Analyses

Gene signature is used to classify sample categories and the performance of the multi-class classification is measured by ROC-AUC (Receiver Operating Characteristics Area Under The Curve) values. A color-coded bar plot is combined with a density or violin+swarm plot to visualize the gene signature-based classification. All statistical tests were performed using R version 3.2.3 (2015-12-10). Standard t-tests were performed using python scipy.stats.ttest_ind package (version 0.19.0) with Welch's Two Sample t-test (unpaired, unequal variance (equal_var=False), and unequal sample size) parameters. Multiple hypothesis corrections were performed by adjusting *p* values with statsmodels.stats.multitest.multipletests (fdr_bh: Benjamini/Hochberg principles). The results were independently validated with R statistical software (R version 3.6.1; 2019-07-05). Pathway analysis of gene lists were carried out via the Reactome database and algorithm (Fabregat et al. 2018). Reactome identifies signaling and metabolic molecules and organizes their relations into biological pathways and processes. Kaplan-Meier analysis is performed using lifelines python package version 0.14.6.

Boolean implication network construction

A Boolean implication network (BIN) is created by identifying all significant pairwise Boolean implication relationships (BIRs) for GSE134312 datasets (**Fig S1A**). The Boolean implication network contains the six possible Boolean relationships between genes in the form of a directed graph with nodes as genes and edges as the Boolean relationship between the genes. The nodes in the BIN are genes and the edges correspond to BIRs. Equivalent and Opposite relationships are denoted by undirected edges and the other four types (low => low; high => low; low => high; high => high) of BIRs are denoted by having a directed edge between them. The network of equivalences seems to follow a scale-free trend; however, other asymmetric relations in the network do not follow scale-free properties. BIR is strong and robust when the sample sizes are usually more than 200. However, it is also possible to build BIN for smaller dataset such as the selected macrophage GSE134312 dataset (n = 197). The macrophage dataset GSE134312 was prepared for Boolean analysis by filtering genes that had a reasonable dynamic range of expression values. When the dynamic range of expression values was small, it was difficult to distinguish if the values were all low or all high or there were some high and some low values. Thus, it was determined to be best to ignore them during Boolean analysis. The filtering step was performed by analyzing the fraction of high and low values identified by the StepMiner algorithm (Sahoo et al. 2007). Any probe set or genes which contained less than 5% of high or low values were dropped from the analysis.

Clustered Boolean Implication network

Clustering was performed in the Boolean implication network to dramatically reduce the complexity of the network (**Fig S1C**). A clustered Boolean implication network (CBIN) was created by clustering nodes in the original BIN by following the equivalent BIRs. One approach is to build connected components in a undirected graph of Boolean equivalences. However, because of noise the connected components become internally inconsistent e.g. two genes opposite to each other becomes part of the same connected component. In order to avoid such situation, we need to break the component by removing the weak links. To identify the weakest links, we first computed a minimum spanning tree for the graph and computed Jaccard similarity coefficient for every edge in this tree. Ideally if two members are part of the same cluster they should share as many connections as possible. If they share less than half of their total individual connections (Jaccard similarity coefficient less than 0.5) the edges are dropped from further analysis. Thus, many weak equivalences were dropped using the above algorithm leaving the clusters internally consistent. We removed all edges that have Jaccard similarity coefficient less than 0.5 and built the connected components with the rest. The connected components were used to cluster the BIN which is converted to the nodes of the CBIN. The distribution of cluster sizes was plotted in a log-log scale to observe the characteristic of the Boolean network (**Fig S1C**). The clusters sizes were distributed along a straight line in a log-log plot suggesting scale-free properties (**Fig S1D**). A new graph was built that connected the individual clusters to each other using Boolean relationships. Link between two clusters (A, B) was established by using the top representative node from A that was connected to most of the member of A

and sampling 6 nodes from cluster B and identifying the overwhelming majority of BIRs between the nodes from each cluster.

A CBIN was created using the selected GSE134312 datasets. Each cluster was associated with healthy or disease samples based on where these gene clusters were highly expressed. The edges between the clusters represented the Boolean relationships that are color-coded as follows: orange for low => high, dark blue for low => low, green for high => high, red for high => low, light blue for equivalent and black for opposite.

Boolean paths

The asymmetric BIRs provide a unique dimension to the network that is fundamentally different from any other gene expression networks in the literature. Traversing a set of nodes in a directed graph of the Boolean network constitutes a Boolean path that can be interpreted as follows. A simple Boolean path involves two nodes and the directed edge between them. This simple Boolean path can be interpreted as shown in the supplementary figure (**Fig S1E**). For the nodes X and Y with X low => Y low only quadrant #1 is sparse; the other quadrants #0, #2, and #3 are filled with samples (**Fig S1E**). Assuming monotonicity in X and Y, the quadrants can be ordered in two possible ways: 0-2-3 and 3-2-0. The path corresponds to 0-2-3 begins with X low and Y low. This is interpreted as X turns on first and then Y turns on along a hypothetical biological path defined by the sample order. Similarly, Y turns off first and then X turns off in the path 3-2-0. A complex path in the Boolean network involves more than one Boolean implication relationship (**Fig S1F**). Three Boolean implication relationships can be used to group samples into five bins and the bins can be ordered in two possible ways (**Fig S1F**, forward, reverse). Another example of a path is illustrated in supplementary figure (**Fig S1G**).

Discovery of Paths in Clustered Boolean Implication network

Discovery of paths start with a node that represents the biggest cluster in the CBIN. Since a path of high=>high, high=>low, and low=>low can be used to order samples as shown in **Fig S1G**, we try to identify paths of this type that intersects the big clusters in the network. We developed a simple, intuitive algorithm that traverses the nodes of the CBIN starting with the biggest cluster and greedily chooses next big cluster connected to the nodes visited in sequence. The emphasis on cluster sizes comes from the fundamental assumption that size determines importance and relevance. Therefore, we start from a big cluster (A1) and identify other clusters that form a chain of low => low. Further, we identify other clusters that are either opposite to A1 or they have high=>low relationship with A1, and the biggest cluster (A2) among these clusters were chosen. In addition, a chain of low=> low relationship from A2 is identified. In each subsequent step, again the biggest cluster among the different choices was greedily chosen. Finally equivalence relationship from each cluster is used to gather more genes in each cluster and the whole path is clustered based on equivalence relationships. Depth-first traversal (DFS) was used to follow the path of low => low where bigger clusters are visited first. The search was performed until a cluster was reached for which there is no low => low relationships. For example, starting with cluster S, the search will return S low => A1 low, A1 low => A2 low, and A2 low => A3 low if A3 doesn't have any low => low relationships.

Similarly, a new starting point is considered S2 such that S2 is the biggest cluster X that has either S high => X low or S Opposite X. From cluster S2 another DFS was performed to retrieve the longest possible path of low => low. The search may return S2 low => B1 low, B1 low => B2 low if B2 doesn't have any low => low relationships. In summary, the most prominent Boolean path was discovered by starting with the largest cluster and then exploring edges that connected to the next largest cluster in a greedy manner. This process was repeated to explore paths that connect the big clusters in the network.

Scoring Boolean path for sample order

A score was computed for a specified Boolean path that can be used to order the sample which was consistent with the logical order. To compute the final score, first the genes present in each cluster were normalized and averaged. Gene expression values were normalized according to a modified Z-score approach centered around StepMiner threshold (formula = (expr - SThr)/3*stddev; **Fig S2B**). Weighted linear combination of the averages from the clusters of a Boolean path was used to create a score for each sample. The weights along the path either monotonically increased or decreased to make the sample order consistent with the logical order based on BIR. The samples were ordered based on the final weighted and linearly combined score (**Fig S2C**). The direction of the path was derived from the connection from a reactive cluster to a tolerant cluster. The sample order is visualized by a color-coded bar plot and a violin+swarm plot (**Fig S2C**).

Summary of genes in the clusters

Reactome pathway analysis of each cluster along the top continuum paths was performed to identify the enriched pathways (Fabregat et al. 2018). The pathway description was used to summarize at a high-level what kind of biological processes are enriched in a particular cluster.

Signatures of macrophage reactivity and tolerance (S-Ma-R-T) computation

BoNE uses Boolean implication network on macrophage dataset to build a signature of macrophage polarization. Selected clusters by size connected by high => high (green arrow), high => low (red arrows) and low => low (blue arrows) Boolean implication relationships. Reactome analysis of each clusters shows the biological processes the genes are involved in (**Fig S2A**). A path is selected in the network that is used to test M1/M2 states classification. This process is demonstrated by using a path #13-14-3 on GSE134312 (**Fig S2B-C**).

Normalization of gene expression based on circadian rhythm

Since the state of macrophage swings from reactive to tolerant from day to night, it is important to control for this variation during analysis of macrophage polarization. To start the normalization process, clock genes (such as DBP, ARNTL, etc.) or gene signatures that capture circadian rhythm is used to adjust the BoNE score (**Fig S4**). First, both the BoNE score (**Fig S4B**) and the clock gene expression are scaled for each sample type based on their dynamic range of expression values (min – max). For example, the dataset GSE98895 contains two sample types: C (Control), and MetS (Metabolic Syndrome). Let's take

one sample from the MetS group (x, y) where x is the clock gene expression value and y is the original BoNE score (**Fig S4C**). Bounding box for the MetS group demonstrates the range of values for both the BoNE score (S1) and the clock gene expression (S2). Average of BoNE scores and the clock gene expression is shown using an orange diamond. The distance of (x, y) from the orange diamond (S3, S4) is used to scale both values: (x – S3 * (S2 + 1) / (S1 + 1) , y + S4 * (S1 + 1) /(S2 + 1)). This process is repeated using control (C) samples using the green diamond. Linear regression is used to compute the trend between the transformed BoNE score and clock gene expression (y = mx + c; **Fig S4D**). The trend is subtracted from the transformed BoNE score to compute the final normalized BoNE score (y – mx - c). Samples are now rank ordered based on the final normalized BoNE score to visualize the effect of normalization process.

Proteomics analysis

A multiplexed TMT (tandem mass tags) quantitative proteomics dataset has been obtained from He, L. et al (He et al. 2021) (see *Key Resource Table*). To generate this dataset, authors had differentiated human THP-1 cells with phorbol myristate acetate (PMA) for 24 h into macrophages (M0 state). The M0 cells were subsequently treated with IL4 for M2 polarization and with LPS and IFNγ for M1 polarization over a 24-h time-period. Samples were processed for quantitative mass spectrometry at 1 h, 4 h, 8 h and 24 h. Ratio of raw intensity values has been compared between M1 and M2 states to obtain the list of induced proteins at various time points (see **Supplemental Information 3**). To obtain the list of proteins induced in M1 state, the cut-off used for induction of proteins when comparing the raw intensity ratio for LPS/IFNγ over IL4 stimulation for all time points was >=2. To obtain the list of proteins induced in M2 state, the cut-off used for induction of proteins when comparing the raw intensity ratio for IL4 over LPS/IFNγ stimulation for all time points was >=1.5.

To assess the differential enrichment of proteins across different signatures for both M1 and M2 polarization states at various time points, we used the following equation to calculate the z-test of proportions,

$$z= p1−p2p(1−p)(1n1+ 1n2)−−−−−−−−−−−−−−−√z= p1−p2p(1−p)(1n1+ 1n2)$$

Here, p1 is sample proportion (x1/n1) of proteins translated from the "reactive" signature that were induced >=2 fold upon LPS stimulation. And p2 is the sample proportion (x2/n2) of proteins translated from the "tolerance" signature that were induced >=1.5 fold upon IL4 stimulation. Here, p = (x1+x2)/(n1+n2).

Supplementary materials:

- Figures S1 to S7
- Table S1 to S3

Supplemental Information (excel datasheets uploaded separately) 1 to 3.

- **Supplemental Information 1**: Excel datasheet with an inventory of all publicly available gene expression datasets analyzed in this work.
- **Supplemental Information 2**: Excel datasheet with an inventory of all previously published gene signatures cited in this work.
- **Supplemental Information 3**: Excel datasheet with list of proteins translated by genes in clusters #13 and 14+3 at various time points after ligand stimulation of THP1 cells.

Figures and legends

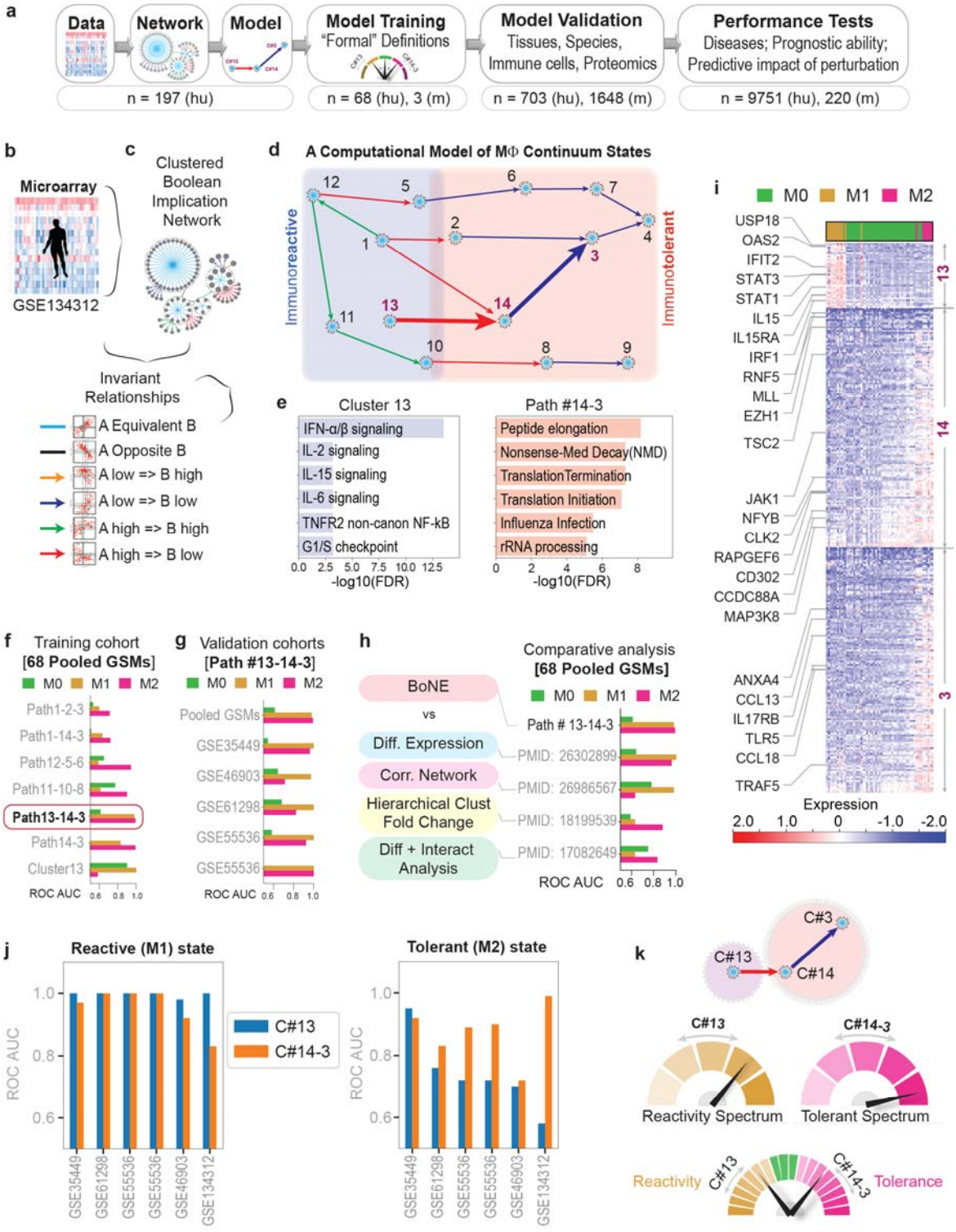


Figure 1: ***BoNE*-assisted formulation of formal definitions of macrophage polarization.**

a) Overview of workflow and approach used in this work.

b-c) A pooled dataset of diverse human transcriptomes (**b**; n = 197) was used to build a Boolean implication network (**c**-*top*) and visualized as gene clusters (nodes, comprised of genes that are equivalent

to each other) that are interconnected based on one of the six overwhelming Boolean implication relationship between the clusters (directed edges; **c**-*bottom*).

d) Display of the major Boolean paths within the network prioritized based on the cluster size. Annotations of "immunoreactive" and "immunotolerant" ends of the spectrum are based on the expression profile of the gene clusters in 68 samples within the pooled dataset that were stimulated *in vitro* as M1 and M2, respectively.

e) Reactome pathway analysis of each cluster along the top continuum paths was performed to identify the enriched pathways (for other clusters see http://hegemon.ucsd.edu/SMaRT/).

f-g) Training (**f**) was performed on the 68 pooled samples using machine-learning approaches; the best-performing Boolean path, #13-14-3 was then validated (**g**) in multiple independent human macrophage datasets. For a list of datasets used see **Table S1**. The performance was measured by computing *ROC AUC* for a logistic regression model.

h) Comparative analysis of performance of the *BoNE*-derived *versus* other traditional approaches in segregating M0/M1/M2 polarization states.

i) Heatmap displaying the pattern of gene expression in C#13, 14 and 3. Selective genes are labelled.

j) Validation studies assessing the ability of the genes in either C#13 alone or C#14-3 alone to classify M0/M1/M2 polarization states in multiple human macrophage datasets.

k) *Top*: Schematic summarizing the model-derived formal definitions of macrophage polarization states based on the levels of expression of genes in C#13 (hypo to hyper- "reactivity" spectrum) and those in C#14+3 (hypo to hyper- "tolerant" spectrum). *Bottom*: A composite score of the entire range of physiologic and pathologic response can be assessed via the *BoNE*-derived path #13→14→3.

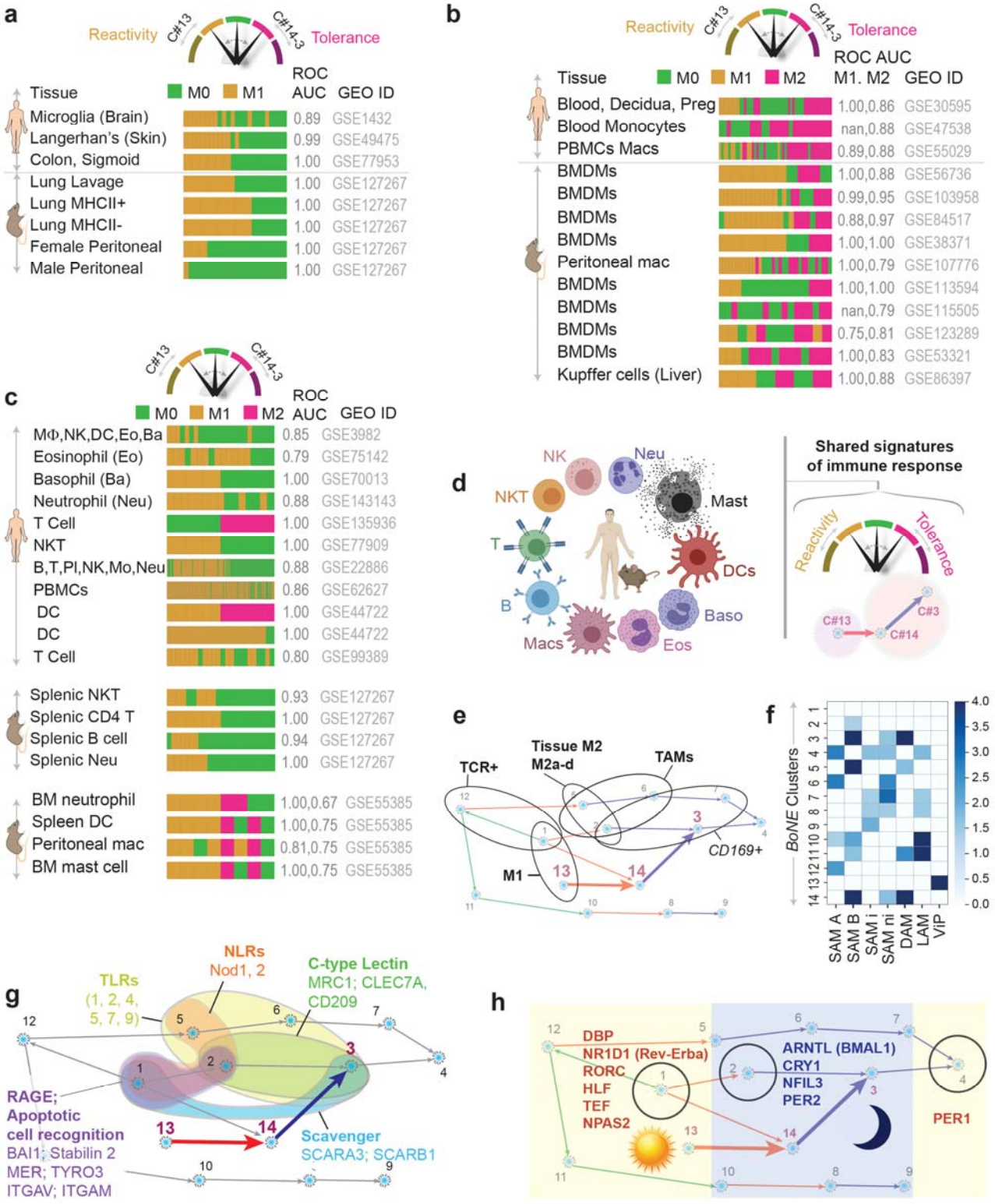


Figure 2. **Definitions of "reactivity" and "tolerance" are conserved across tissues, organs, species and diverse immune cell types.**

a-b) Validation studies assessing the ability of *SMaRT* genes to classify diverse tissue-resident macrophage datasets from both humans and mice. Performance is measured by computing *ROC AUC* for a logistic regression model.

c-d) Validation studies (c) assessing the ability of *SMaRT* genes to classify active vs inactive states of diverse immune cell types in both humans and mice. The schematic (d) summarizes findings in c.

e) Known macrophage subtypes, as defined by marker genes, are projected on the Boolean map of macrophage processes.

f) Published disease-associated macrophage gene signatures (see **Supplemental Information 2**) are are analyzed for significant overlaps with various gene clusters in the Boolean map of macrophage processes. Results are displayed as heatmaps of -Log10(*p)* values as determined by a hypergeometric test.

g) The distribution of pattern recognition receptors (PRRs) [see **Table S2**] within various gene clusters of the Boolean map of macrophage processes is displayed.

h) The positions of key circadian genes that are present in the network are shown on the Boolean map of macrophage processes. See also **Fig S4**.

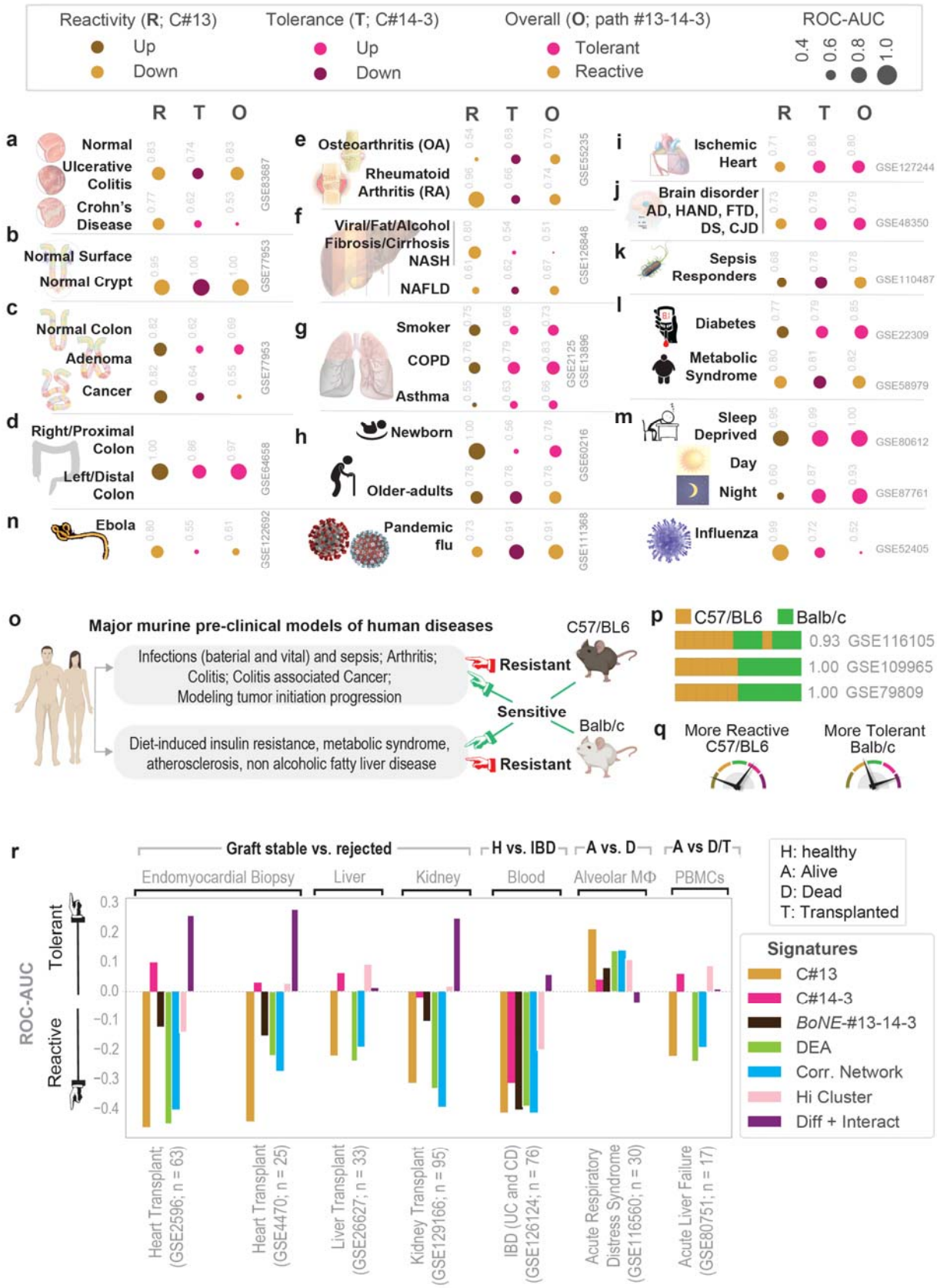


Figure 3. Definitions of "reactivity" and "tolerance" detects pathologic macrophage states in disease.

Tissue immune microenvironment is visualized (in panels a-n) as bubble plots of ROC-AUC values (radii of circles are based on the ROC-AUC; Key on top) demonstrating the direction of gene regulation (Up vs

Down; Key on top) for the classification of samples using BoNE-derived gene signatures of either reactive (R; C#13) or tolerant (T; C#14-3) or overall (O; path #13→14→3) in columns. The ROC-AUC values are provided next to the bubble. Sample diversity and sizes are as follows:

a) IBD; GSE83687, n = 134; 60 Normal, 32 Ulcerative Colitis, 42 Crohn's Disease.

b) Colon crypt; GSE77953, 6 Normal Surface *vs*. 7 Normal Crypt base. **c**; Colon cancer: Pooled colon dataset from NCBI GEO; n=170 Normal, 68 Adenomas, 1662 CRCs.

d) Colon anatomy: Proximal (right) vs distal (left) normal colon from mouse (GSE64423, n = 6) and human (GSE20881, n = 75). See **Figure S5** for violin plots.

e) Arthritis; GSE55235, GSE55457 and GSE55584, n = 79; 20 Normal, 33 Rheumatoid Arthritis, 26 Osteoarthritis.

f) Hepatitis: GSE89632, n=63; 20 fatty liver, 19 Non-alcoholic steatohepatitis (NASH) and 24 healthy, alcoholic liver disease (GSE94417, GSE94397 and GSE94399, n = 195; 109 Healthy, 13 Alcoholic Hepatitis, 6 Alcoholic fatty liver (AFL), 67 Alcoholic cirrhosis (AC) and viral hepatitis (GSE70779, n=18; 9 Pre-treatment, 9 Post-treatment with direct-acting anti-virals).

g) Chronic lung disease; GSE2125 and GSE13896, n = 115; 39 Non-smoker, 49 Smoker, 15 Asthma, 12, Chronic Obstructive Pulmonary Disease (COPD).

h) Aging process; GSE60216, n = 9; 3 Newborn babies, 3 Adults, 3 Old-adults.

i) Cardiomyopathy (CM), ischemic and non-ischemic (I/NI); GSE104423, n = 25 human samples; 14 NICM, 11 ICM; GSE127244, n = 24 mouse samples, 16 NICM, 8 ICM.

j) Neurodegenerative brain disorders; GSE118553 (n = 401) and GSE48350 (n = 253), Alzheimer's disease (AD); GSE35864, HIV-associated neurocognitive disorder (HAND; n = 72); GSE13162, frontotemporal dementia (FTD; n = 56); GSE59630, Down's Syndrome (DS; n = 116); GSE124571, Creutzfeldt-Jakob Disease (CJD; n = 21).

k) Systemic inflammatory response syndrome (SIRS) and sepsis; GSE63042 (n = 129); GSE110487 (n = 31).

l) Type 2 diabetes and metabolic syndrome; GSE22309 (n = 110), Pre- and post- insulin treatment muscle biopsies from 20 insulin sensitive, 20 insulin resistant, 15 T2DM; GSE98895 (n = 40), PBMCs from 20 control, 20 metabolic syndrome.

m) Sleep deprivation and circadian rhythm; GSE9444, n = 131 mouse brain and liver samples; GSE80612, twin, n = 22 human peripheral blood leukocytes; GSE98582, n = 555 human blood samples; GSE104674, n = 48, 24 healthy and 24 T2DM.

n) Viral pandemics, such as SARS, MERS, Ebola and others [**N**; numerous datasets, see **Figure S7E**].

See **Figure S6** and **S7** for violin plots relevant to panel **e-n**.

o-q) Schematic (o) summarizes the use of two major mouse strains (C57/B6 and Balb/c) commonly used for modeling two broad categories of human diseases. Bar plots (p) showing sample classification of genetically diverse macrophage datasets based on expression levels of genes in C#13. Schematic (q) summarizes findings.

r) The diagnostic potential of various indicated gene signatures were tested on multiple datasets generated from tissues derived from patients with the known clinically relevant outcome, as indicated. In each case, BoNE-derived signatures were compared against four traditional approaches.

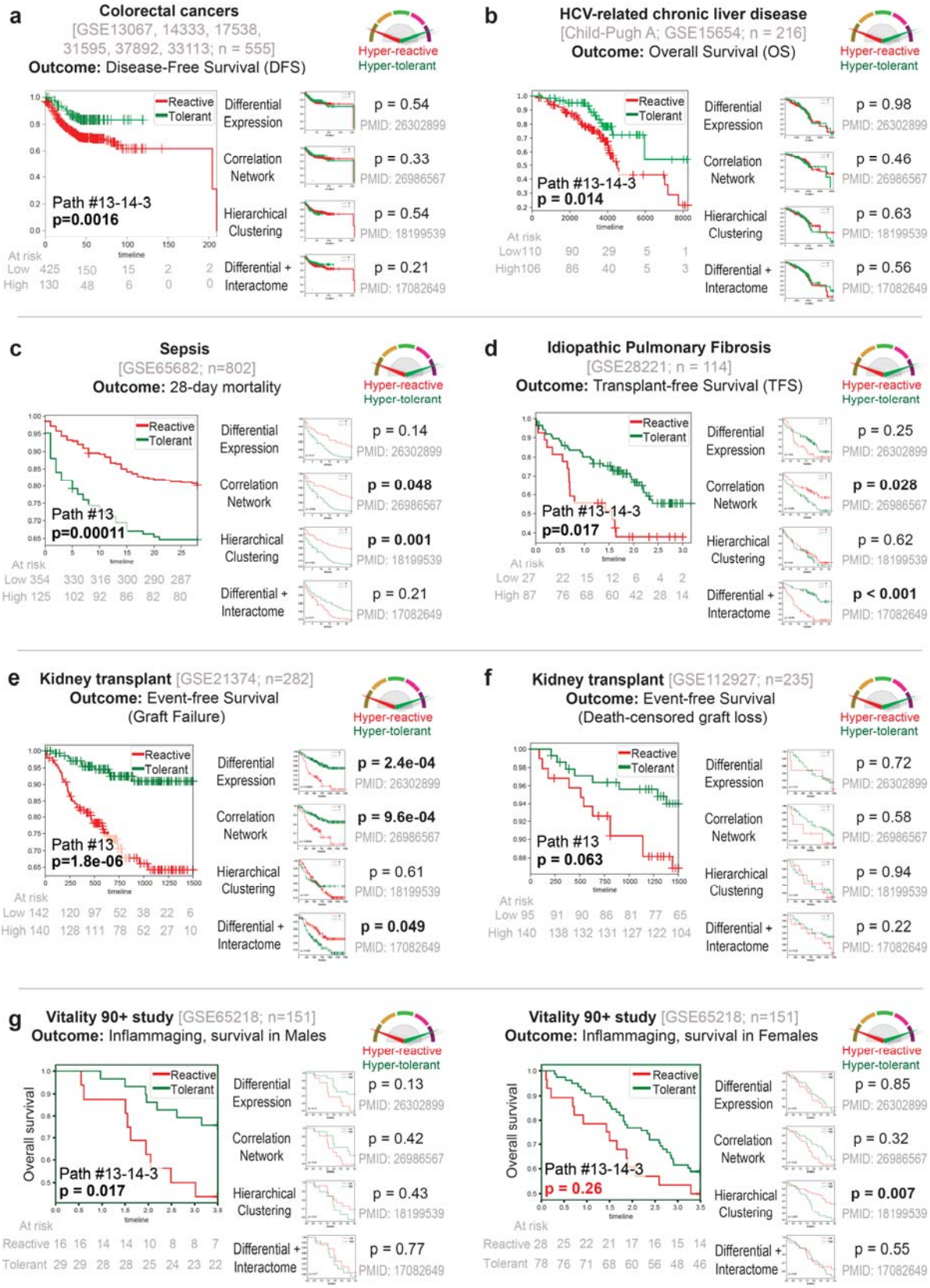


Figure 4. Comparison of the prognostic and diagnostic potentials of *BoNE vs*. other traditional approaches.

a-g) The prognostic performance of the *BoNE*-derived *SMaRT* genes is compared head-to-head with signatures derived from four other traditional approaches across diverse disease conditions (colon cancer,

a; liver fibrosis, b; sepsis, c; idiopathic pulmonary fibrosis, d; kidney transplantation, e-f; inflammaging, g). Results are displayed as Kaplan Meier (KM) curves with significance (*p* values) as assessed by log-rank-test. A composite immune response score is computed using Boolean path #13→14→3 or C#13 alone, as indicated within each KM plot. Low score = “reactive”; high score = “tolerant”. A threshold is computed using StepMiner on the immune score to separate these two states. See also **Fig S7F** for other cancers.

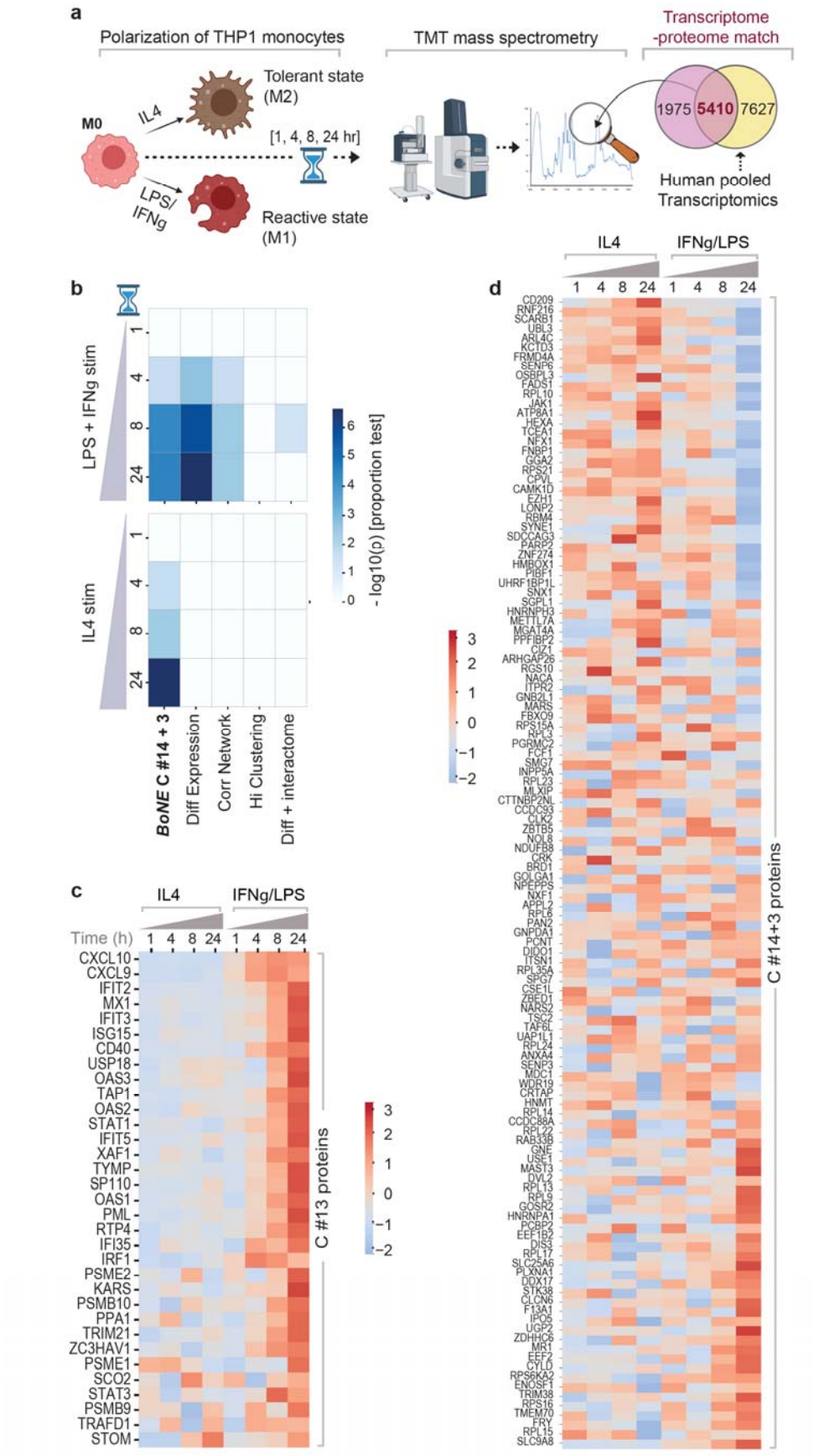


Figure 5. *SMaRT* genes are differentially translated in polarized macrophages

a) Overview of the experimental design. PMA-treated human THP-1 cell lines (M0) are polarized to M1 (with LPS and IGNg) or M2 (with IL4), followed by multiplexed mass spectrometry at indicated time points. The fraction of the global macrophage transcriptome (from the pooled 197 macrophage datasets) that is represented in the global macrophage proteome is subsequently assessed for induction (or not) of proteins that are translated by various gene signatures.

b) Selectivity of induction of proteins upon LPS and IFNg (top) or IL4 (bottom) stimulation at various timepoints was assessed across different signatures using z-test of proportions and -log(10)p values are displayed as heatmaps.

c-d) z normalized Log of intensities of proteins (**Supplemental Information 3**) translated at different time points by genes in C#13 (c) and C#14+3 (d) are displayed as heatmaps.

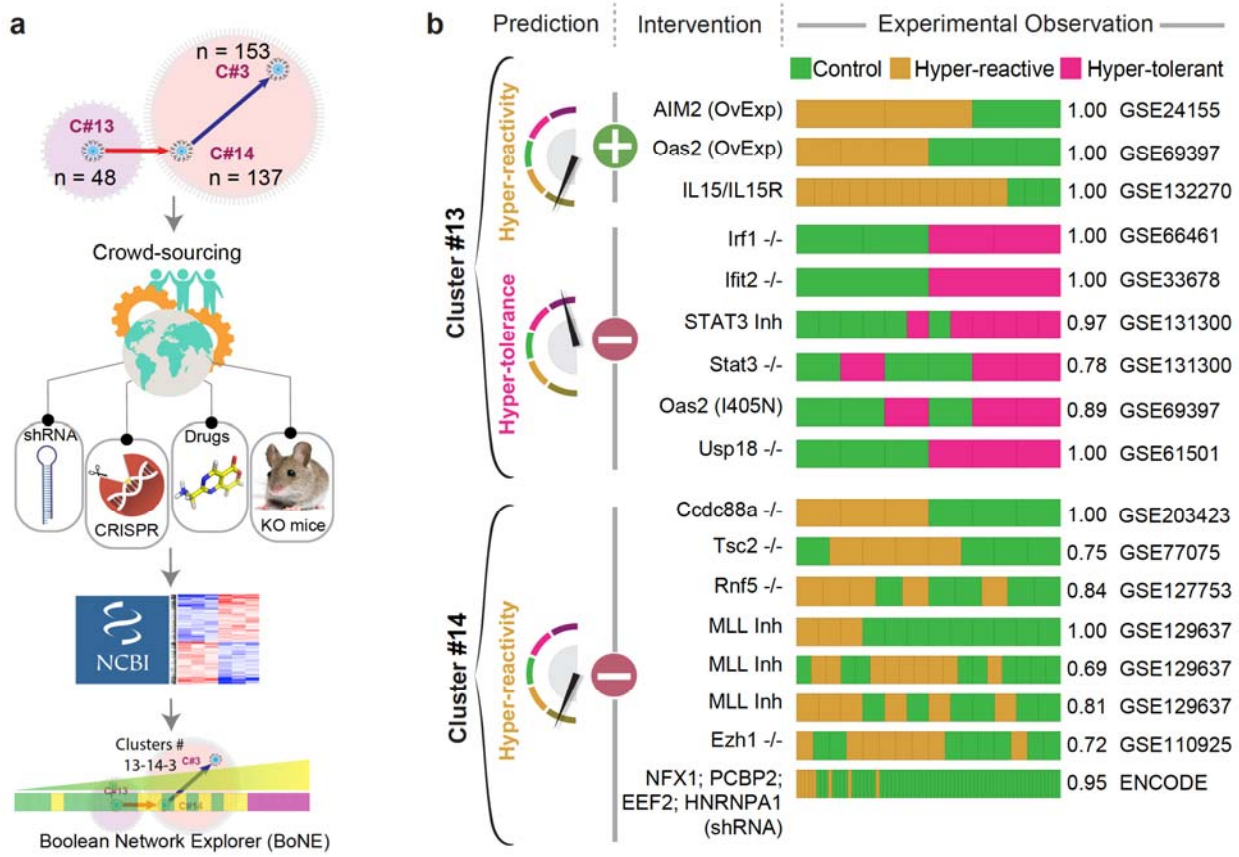


Figure 6. **Crowd-sourced assessment of the predictive potential of the *SMaRT* genes.**

a) Overview of our workflow and approach for crowd-sourced validation. Publicly available transcriptomic datasets reporting the outcome of intervention studies (genetic or pharmacologic manipulations) on macrophages/monocytes targeting any of the 185 genes in C#13 and C#14 were analyzed using the *BoNE* platform for macrophage states.

b) Predicted impact of positive (+, either overexpression [OvExp] or agonist stimulations) or negative (-; genetic -/- models, shRNA or chemical inhibitors) interventions and observed macrophage polarization states are shown. Performance is measured by computing *ROC AUC* for a logistic regression model. See **Table S3**.

References

Amit, I., Winter, D.R., and Jung, S. (2016). The role of the local environment and epigenetics in shaping macrophage identity and their effect on tissue homeostasis. Nat Immunol 17, 18-25.

Azad, T.D., Donato, M., Heylen, L., Liu, A.B., Shen-Orr, S.S., Sweeney, T.E., Maltzman, J.S., Naesens, M., and Khatri, P. (2018). Inflammatory macrophage-associated 3-gene signature predicts subclinical allograft injury and graft survival. JCI Insight 3.

Barrett, T., Suzek, T.O., Troup, D.B., Wilhite, S.E., Ngau, W.C., Ledoux, P., Rudnev, D., Lash, A.E., Fujibuchi, W., and Edgar, R. (2005). Ncbi geo: Mining millions of expression profiles--database and tools. Nucleic Acids Res 33, D562-566.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., *et al.* (2013). Ncbi geo: Archive for functional genomics data sets--update. Nucleic Acids Res 41, D991-995.

Becker, M., De Bastiani, M.A., Parisi, M.M., Guma, F.T., Markoski, M.M., Castro, M.A., Kaplan, M.H., Barbe-Tuana, F.M., and Klamt, F. (2015). Integrated transcriptomics establish macrophage polarization signatures and have potential applications for clinical health and disease. Sci Rep 5, 13351.

Bell, L.C., Pollara, G., Pascoe, M., Tomlinson, G.S., Lehloenya, R.J., Roe, J., Meldau, R., Miller, R.F., Ramsay, A., Chain, B.M., *et al.* (2016). In vivo molecular dissection of the effects of hiv-1 in active tuberculosis. PLoS Pathog 12, e1005469.

Bergler, T., Jung, B., Bourier, F., Kuhne, L., Banas, M.C., Rummele, P., Wurm, S., and Banas, B. (2016). Infiltration of macrophages correlates with severity of allograft rejection and outcome in human kidney transplantation. PLoS One 11, e0156900.

Byrne, A.J., Maher, T.M., and Lloyd, C.M. (2016). Pulmonary macrophages: A new therapeutic pathway in fibrosing lung disease? Trends Mol Med 22, 303-316.

Champy, M.F., Selloum, M., Zeitler, V., Caradec, C., Jung, B., Rousseau, S., Pouilly, L., Sorg, T., and Auwerx, J. (2008). Genetic background determines metabolic phenotypes in the mouse. Mamm Genome 19, 318-331.

Chavez-Galan, L., Olleros, M.L., Vesin, D., and Garcia, I. (2015). Much more than m1 and m2 macrophages, there are also cd169(+) and tcr(+) macrophages. Front Immunol 6, 263.

Cho, J.H. (2008). The genetics and immunopathogenesis of inflammatory bowel disease. Nat Rev Immunol 8, 458-466.

Coates, P.J., Rundle, J.K., Lorimore, S.A., and Wright, E.G. (2008). Indirect macrophage responses to ionizing radiation: Implications for genotype-dependent bystander signaling. Cancer Res 68, 450-456.

Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57-74.

Coussens, L.M., and Werb, Z. (2002). Inflammation and cancer. Nature 420, 860-867.

Dabydeen, S.A., Desai, A., and Sahoo, D. (2019). Unbiased boolean analysis of public gene expression data for cell cycle gene identification. Mol Biol Cell 30, 1770-1779.

Dang, D., Taheri, S., Das, S., Ghosh, P., Prince, L.S., and Sahoo, D. (2020). Computational approach to identifying universal macrophage biomarkers. Frontiers in Physiology 11.

Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K., *et al.* (2018). The encyclopedia of DNA elements (encode): Data portal update. Nucleic Acids Res 46, D794-D801.

Drewes, J.L., Housseau, F., and Sears, C.L. (2016). Sporadic colorectal cancer: Microbial contributors to disease prevention, development and therapy. Br J Cancer 115, 273-280.

Duffield, J.S., Forbes, S.J., Constandinou, C.M., Clay, S., Partolina, M., Vuthoori, S., Wu, S., Lang, R., and Iredale, J.P. (2005). Selective depletion of macrophages reveals distinct, opposing roles during liver injury and repair. J Clin Invest 115, 56-65.

Early, J.O., and Curtis, A.M. (2016). Immunometabolism: Is it under the eye of the clock? Semin Immunol 28, 478-490.

Eckel-Mahan, K., and Sassone-Corsi, P. (2013). Metabolism and the circadian clock converge. Physiol Rev 93, 107-135.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene expression omnibus: Ncbi gene expression and hybridization array data repository. Nucleic Acids Res 30, 207-210.

Ehling, J., Bartneck, M., Wei, X., Gremse, F., Fech, V., Mockel, D., Baeck, C., Hittatiya, K., Eulberg, D., Luedde, T.*, et al.* (2014). Ccl2-dependent infiltrating macrophages promote angiogenesis in progressive liver fibrosis. Gut 63, 1960-1971.

Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., Haw, R., Jassal, B., Korninger, F., May, B.*, et al.* (2018). The reactome pathway knowledgebase. Nucleic Acids Res 46, D649-D655.

George, T.J., Arnaoutakis, G.J., and Shah, A.S. (2011). Lung transplant in idiopathic pulmonary fibrosis. Arch Surg 146, 1204-1209.

Ginhoux, F., Schultze, J.L., Murray, P.J., Ochando, J., and Biswas, S.K. (2016). New insights into the multidimensional concept of macrophage ontogeny, activation and function. Nat Immunol 17, 34-40.

Glass, C.K., and Natoli, G. (2016). Molecular control of activation and priming in macrophages. Nat Immunol 17, 26-33.

Gordon, S., and Pluddemann, A. (2017). Tissue macrophages: Heterogeneity and functions. BMC Biol 15, 53.

He, L., Jhong, J.H., Chen, Q., Huang, K.Y., Strittmatter, K., Kreuzer, J., DeRan, M., Wu, X., Lee, T.Y., Slavov, N.*, et al.* (2021). Global characterization of macrophage polarization mechanisms and identification of m2-type polarization inhibitors. Cell Rep 37, 109955.

Heinrichs, D., Knauel, M., Offermanns, C., Berres, M.L., Nellen, A., Leng, L., Schmitz, P., Bucala, R., Trautwein, C., Weber, C.*, et al.* (2011). Macrophage migration inhibitory factor (mif) exerts antifibrotic effects in experimental liver fibrosis via cd74. Proceedings of the National Academy of Sciences of the United States of America 108, 17444-17449.

Howes, A., Taubert, C., Blankley, S., Spink, N., Wu, X., Graham, C.M., Zhao, J., Saraiva, M., Ricciardi-Castagnoli, P., Bancroft, G.J.*, et al.* (2016). Differential production of type i ifn determines the reciprocal levels of il-10 and proinflammatory cytokines produced by c57bl/6 and balb/c macrophages. J Immunol 197, 2838-2853.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., and Speed, T.P. (2003a). Summaries of affymetrix genechip probe level data. Nucleic Acids Res 31, e15.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003b). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4, 249-264.

Ishida, B.Y., Blanche, P.J., Nichols, A.V., Yashar, M., and Paigen, B. (1991). Effects of atherogenic diet consumption on lipoproteins in mouse strains c57bl/6 and c3h. J Lipid Res 32, 559-568.

Jaitin, D.A., Adlung, L., Thaiss, C.A., Weiner, A., Li, B., Descamps, H., Lundgren, P., Bleriot, C., Liu, Z., Deczkowska, A.*, et al.* (2019). Lipid-associated macrophages control metabolic homeostasis in a trem2-dependent manner. Cell 178, 686-698.e614.

Kazankov, K., Barrera, F., Moller, H.J., Bibby, B.M., Vilstrup, H., George, J., and Gronbaek, H. (2014). Soluble cd163, a macrophage activation marker, is independently associated with fibrosis in patients with chronic viral hepatitis b and c. Hepatology 60, 521-530.

Keren-Shaul, H., Spinrad, A., Weiner, A., Matcovitch-Natan, O., Dvir-Szternfeld, R., Ulland, T.K., David, E., Baruch, K., Lara-Astaiso, D., Toth, B.*, et al.* (2017). A unique microglia type associated with restricting development of alzheimer's disease. Cell 169, 1276-1290.e1217.

Lavin, Y., Winter, D., Blecher-Gonen, R., David, E., Keren-Shaul, H., Merad, M., Jung, S., and Amit, I. (2014). Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. Cell 159, 1312-1326.

Law, C.W., Alhamdoosh, M., Su, S., Dong, X., Tian, L., Smyth, G.K., and Ritchie, M.E. (2016). Rna-seq analysis is easy as 1-2-3 with limma, glimma and edger. F1000Res 5.

Li, B., and Dewey, C.N. (2011). Rsem: Accurate transcript quantification from rna-seq data with or without a reference genome. BMC Bioinformatics 12, 323.

Li, B., Ruotti, V., Stewart, R.M., Thomson, J.A., and Dewey, C.N. (2010). Rna-seq gene expression estimation with read mapping uncertainty. Bioinformatics 26, 493-500.

Link, V.M., Duttke, S.H., Chun, H.B., Holtman, I.R., Westin, E., Hoeksema, M.A., Abe, Y., Skola, D., Romanoski, C.E., Tao, J.*, et al.* (2018). Analysis of genetically diverse macrophages reveals local and domain-wide mechanisms that control transcription factor binding and function. Cell 173, 1796-1809 e1717.

Liu, Y., and Cao, X. (2015). The origin and function of tumor-associated macrophages. Cell Mol Immunol 12, 1-4.

Liu, Y., Kloc, M., and Li, X.C. (2016). Macrophages as effectors of acute and chronic allograft injury. Curr Transplant Rep 3, 303-312.

MacParland, S.A., Liu, J.C., Ma, X.Z., Innes, B.T., Bartczak, A.M., Gage, B.K., Manuel, J., Khuu, N., Echeverri, J., Linares, I.*, et al.* (2018). Single cell rna sequencing of human liver reveals distinct intrahepatic macrophage populations. Nat Commun 9, 4383.

Mainou-Fowler, T., MacGowan, A.P., and Postlethwaite, R. (1988). Virulence of listeria spp.: Course of infection in resistant and susceptible mice. J Med Microbiol 27, 131-140.

Martinez, F.O., Gordon, S., Locati, M., and Mantovani, A. (2006). Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: New molecules and patterns of gene expression. J Immunol 177, 7303-7311.

Mestas, J., and Hughes, C.C. (2004). Of mice and not men: Differences between mouse and human immunology. J Immunol 172, 2731-2738.

Mills, C.D., Kincaid, K., Alt, J.M., Heilman, M.J., and Hill, A.M. (2000). M-1/m-2 macrophages and the th1/th2 paradigm. J Immunol 164, 6166-6173.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by rna-seq. Nat Methods 5, 621-628.

Murray, P.J., and Wynn, T.A. (2011). Protective and pathogenic functions of macrophage subsets. Nat Rev Immunol 11, 723-737.

Nevalainen, T., Kananen, L., Marttila, S., Jylha, M., Hervonen, A., Hurme, M., and Jylhava, J. (2015). Transcriptomic and epigenetic analyses reveal a gender difference in aging-associated inflammation: The vitality 90+ study. Age (Dordr) 37, 9814.

Nosraty, L., Enroth, L., Raitanen, J., Hervonen, A., and Jylha, M. (2015). Do successful agers live longer? The vitality 90+ study. J Aging Health 27, 35-53.

Okabe, Y., and Medzhitov, R. (2016). Tissue biology perspective on macrophages. Nat Immunol 17, 9-17.

Pachter, L. (2011). Models for transcript quantification from rna-seq. In arXiv e-prints.

Pandey, S., and Sahoo, D. (2019). Identification of gene expression logical invariants in arabidopsis. Plant Direct 3, e00123.

Pena, O.M., Hancock, D.G., Lyle, N.H., Linder, A., Russell, J.A., Xia, J., Fjell, C.D., Boyd, J.H., and Hancock, R.E. (2014). An endotoxin tolerance signature predicts sepsis and organ dysfunction at initial clinical presentation. EBioMedicine 1, 64-71.

Peters, L.A., Perrigoue, J., Mortha, A., Iuga, A., Song, W.M., Neiman, E.M., Llewellyn, S.R., Di Narzo, A., Kidd, B.A., Telesco, S.E.*, et al.* (2017). A functional genomics predictive network model identifies regulators of inflammatory bowel disease. Nat Genet 49, 1437-1449.

Pollard, J.W. (2009). Trophic macrophages in development and disease. Nat Rev Immunol 9, 259-270.

Qian, B.Z., and Pollard, J.W. (2010). Macrophage diversity enhances tumor progression and metastasis. Cell 141, 39-51.

Ramachandran, P., Dobie, R., Wilson-Kanamori, J.R., Dora, E.F., Henderson, B.E.P., Luu, N.T., Portman, J.R., Matchett, K.P., Brice, M., Marwick, J.A.*, et al.* (2019). Resolving the fibrotic niche of human liver cirrhosis at single-cell level. Nature 575, 512-518.

Ravishankar, B., Shinde, R., Liu, H., Chaudhary, K., Bradley, J., Lemos, H.P., Chandler, P., Tanaka, M., Munn, D.H., Mellor, A.L.*, et al.* (2014). Marginal zone cd169+ macrophages coordinate apoptotic cell-driven cellular recruitment and tolerance. Proceedings of the National Academy of Sciences of the United States of America 111, 4215-4220.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). Edger: A bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139-140.

Rosenthal, N., and Brown, S. (2007). The mouse ascending: Perspectives for human-disease models. Nat Cell Biol 9, 993-999.

Sacks, D., and Noben-Trauth, N. (2002). The immunology of susceptibility and resistance to leishmania major in mice. Nat Rev Immunol 2, 845-858.

Sahoo, D., Dill, D.L., Gentles, A.J., Tibshirani, R., and Plevritis, S.K. (2008). Boolean implication networks derived from large scale, whole genome microarray datasets. Genome Biol 9, R157.

Sahoo, D., Dill, D.L., Tibshirani, R., and Plevritis, S.K. (2007). Extracting binary signals from microarray time-course data. Nucleic Acids Res 35, 3705-3712.

Sahoo, D., Seita, J., Bhattacharya, D., Inlay, M.A., Weissman, I.L., Plevritis, S.K., and Dill, D.L. (2010). Midreg: A method of mining developmentally regulated genes using boolean implications. Proc Natl Acad Sci U S A 107, 5732-5737.

Saunderson, S.C., Dunn, A.C., Crocker, P.R., and McLellan, A.D. (2014). Cd169 mediates the capture of exosomes in spleen and lymph node. Blood 123, 208-216.

Schluter, D., Deckert-Schluter, M., Lorenz, E., Meyer, T., Rollinghoff, M., and Bogdan, C. (1999). Inhibition of inducible nitric oxide synthase exacerbates chronic cerebral toxoplasmosis in toxoplasma gondii-susceptible c57bl/6 mice but does not reactivate the latent disease in t. Gondii-resistant balb/c mice. J Immunol 162, 3512-3518.

Steinbach, E.C., and Plevy, S.E. (2014). The role of macrophages and dendritic cells in the initiation of inflammation in ibd. Inflamm Bowel Dis 20, 166-175.

Stout, R.D., and Suttles, J. (2004). Functional plasticity of macrophages: Reversible adaptation to changing microenvironments. J Leukoc Biol 76, 509-513.

Sunami, Y., Leithauser, F., Gul, S., Fiedler, K., Guldiken, N., Espenlaub, S., Holzmann, K.H., Hipp, N., Sindrilaru, A., Luedde, T.*, et al.* (2012). Hepatic activation of ikk/nfkappab signaling induces liver fibrosis via macrophage-mediated chronic inflammation. Hepatology 56, 1117-1128.

Toye, A.A., Lippiat, J.D., Proks, P., Shimomura, K., Bentley, L., Hugill, A., Mijat, V., Goldsworthy, M., Moir, L., Haynes, A.*, et al.* (2005). A genetic and physiological study of impaired glucose homeostasis control in c57bl/6j mice. Diabetologia 48, 675-686.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). Tophat: Discovering splice junctions with rna-seq. Bioinformatics 25, 1105-1111.

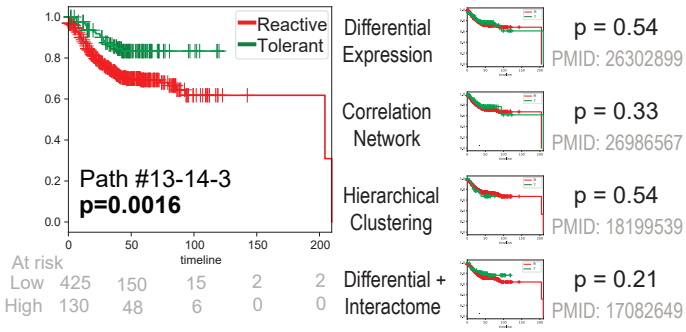
Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by rna-seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28, 511-515.

Wagner, G.P., Kin, K., and Lynch, V.J. (2012). Measurement of mrna abundance using rna-seq data: Rpkm measure is inconsistent among samples. Theory Biosci 131, 281-285.

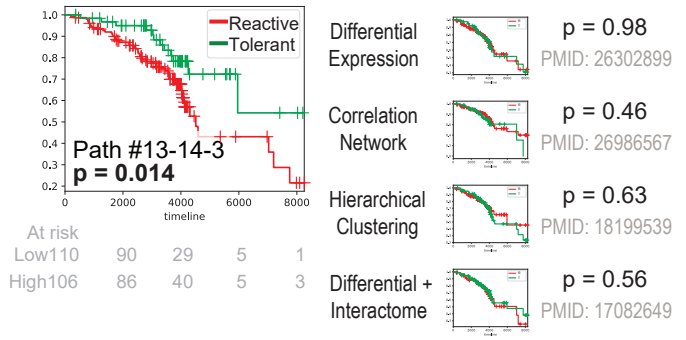
Weiss, W.A., Taylor, S.S., and Shokat, K.M. (2007). Recognizing and exploiting differences between rnai and small-molecule inhibitors. Nat Chem Biol 3, 739-744.

Zhou, L., Cao, X., Fang, J., Li, Y., and Fan, M. (2015). Macrophages polarization is mediated by the combination of prr ligands and distinct inflammatory cytokines. Int J Clin Exp Pathol 8, 10964-10974.

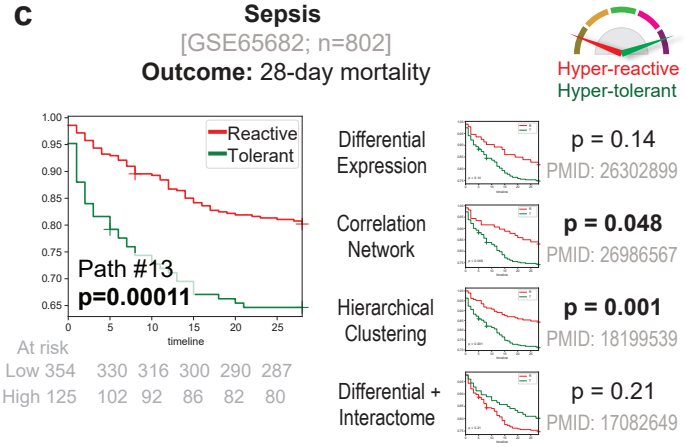
a **Colorectal cancers**
[GSE13067, 14333, 17538, 31595, 37892, 33113; n = 555]
Outcome: Disease-Free Survival (DFS)



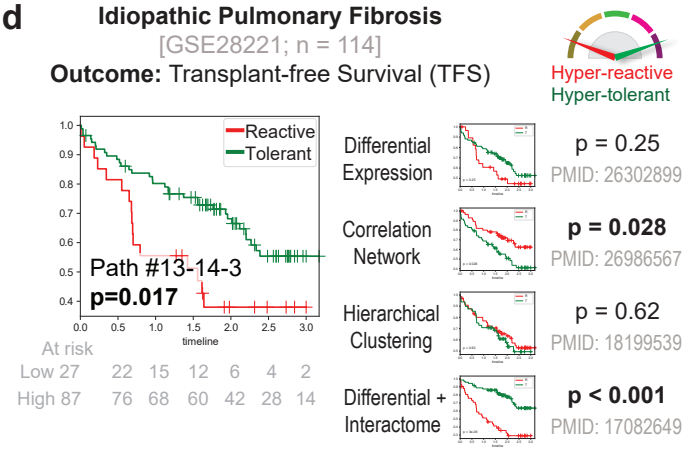
b **HCV-related chronic liver disease**
[Child-Pugh A; GSE15654; n = 216]
Outcome: Overall Survival (OS)



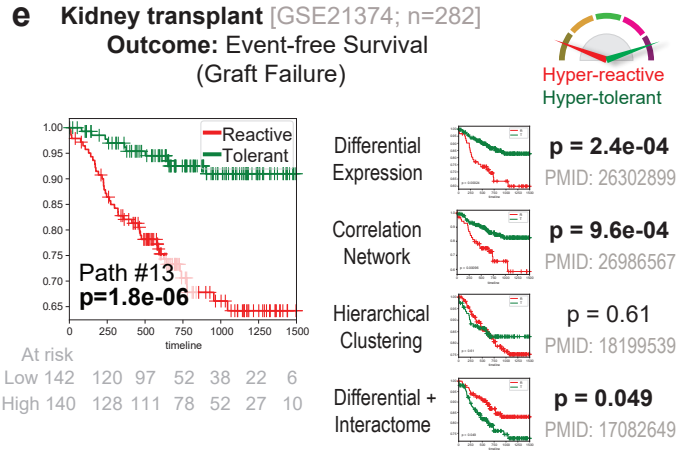
c **Sepsis**
[GSE65682; n=802]
Outcome: 28-day mortality



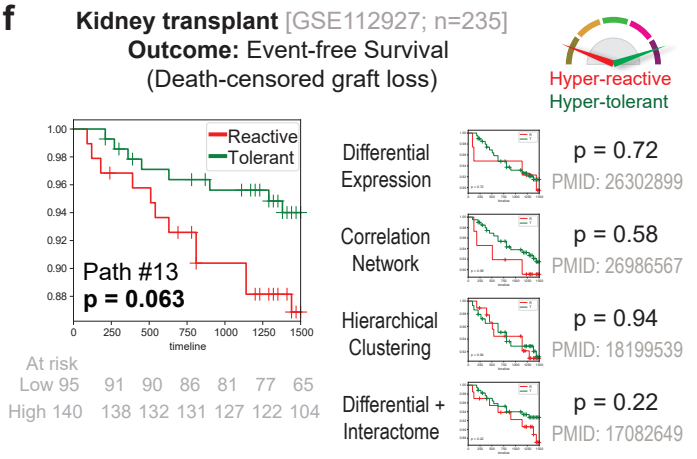
d **Idiopathic Pulmonary Fibrosis**
[GSE28221; n = 114]
Outcome: Transplant-free Survival (TFS)



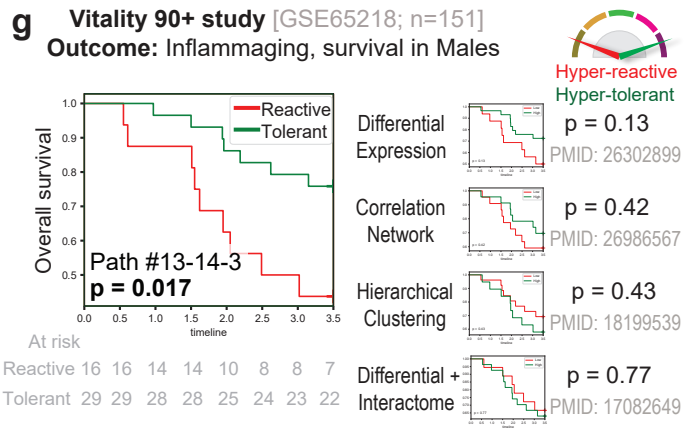
e **Kidney transplant** [GSE21374; n=282]
Outcome: Event-free Survival (Graft Failure)



f **Kidney transplant** [GSE112927; n=235]
Outcome: Event-free Survival (Death-censored graft loss)



g **Vitality 90+ study** [GSE65218; n=151]
Outcome: Inflammaging, survival in Males



Vitality 90+ study [GSE65218; n=151]
Outcome: Inflammaging, survival in Females

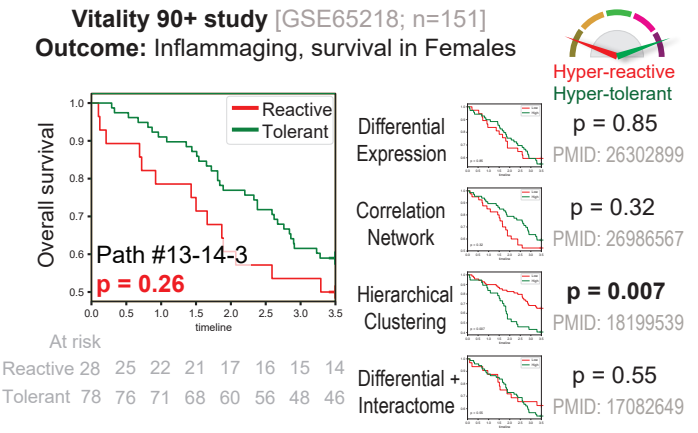
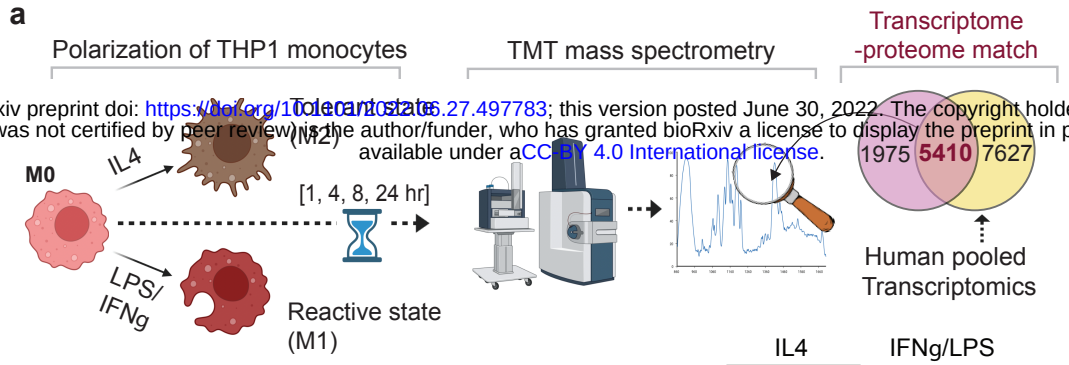
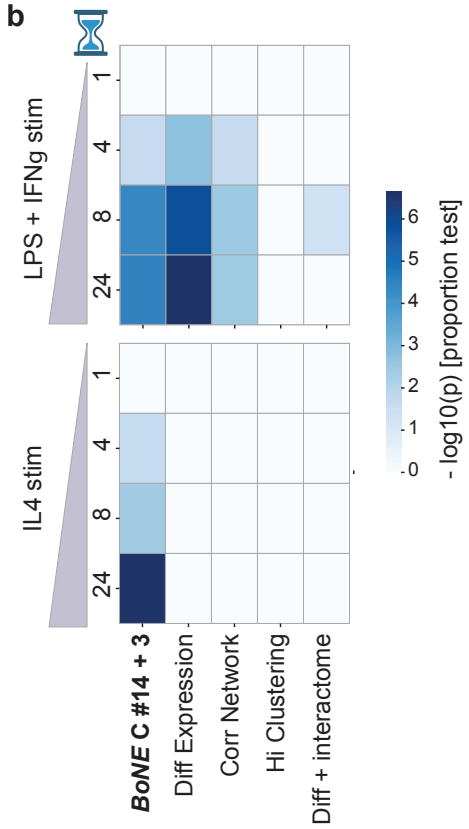
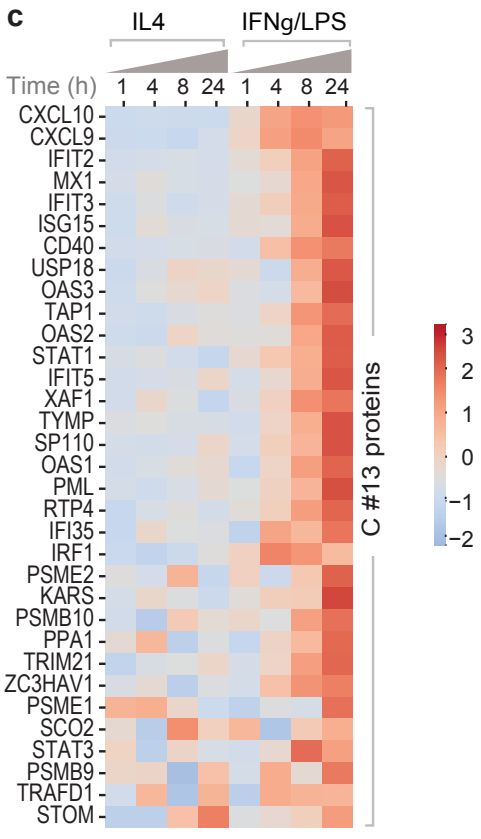
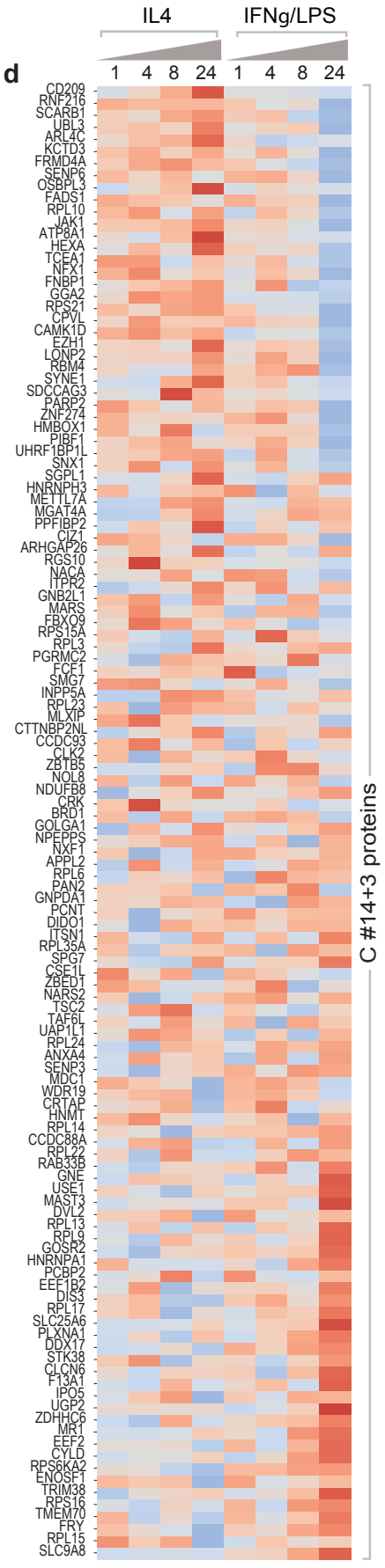


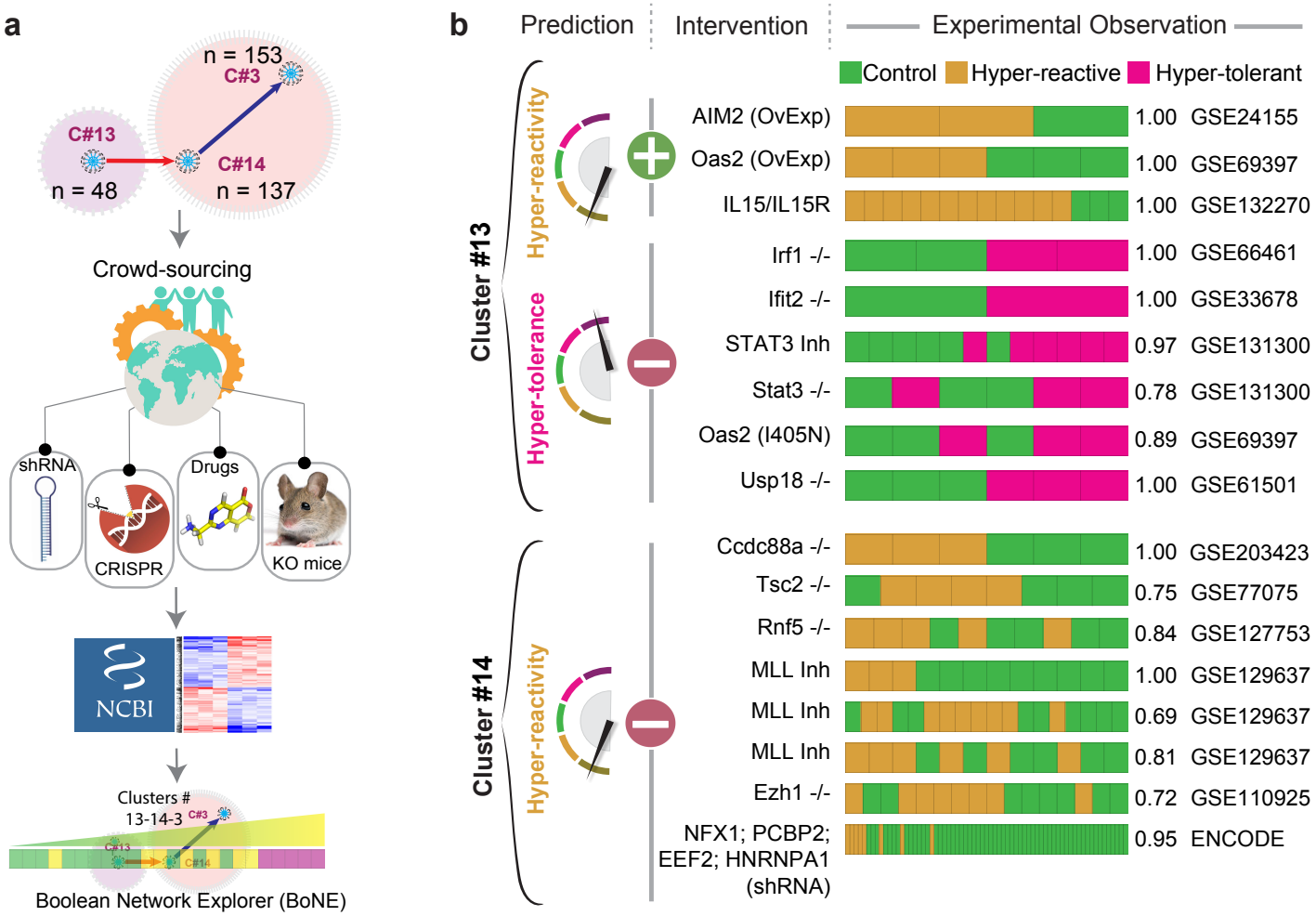
Figure 4

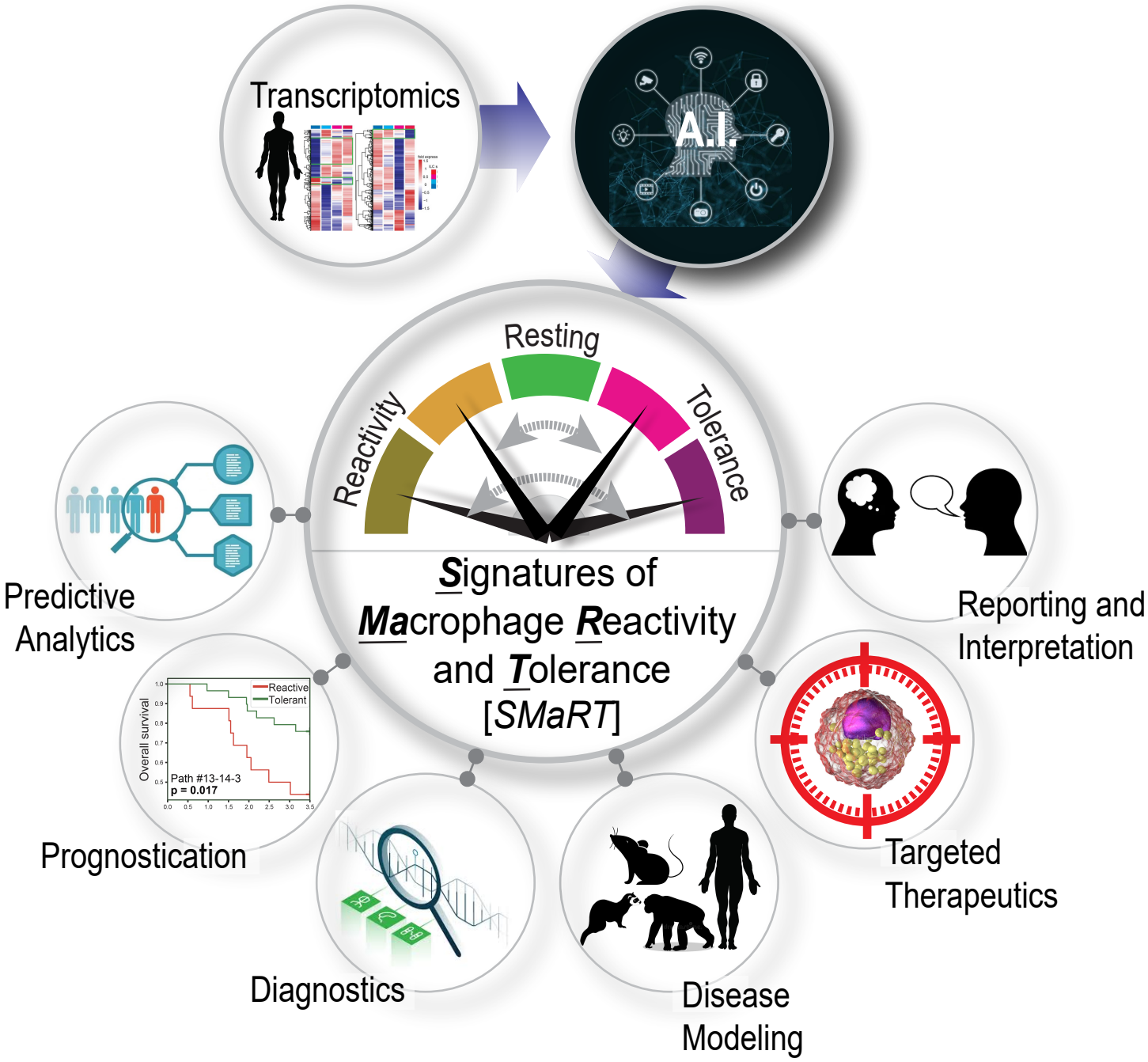




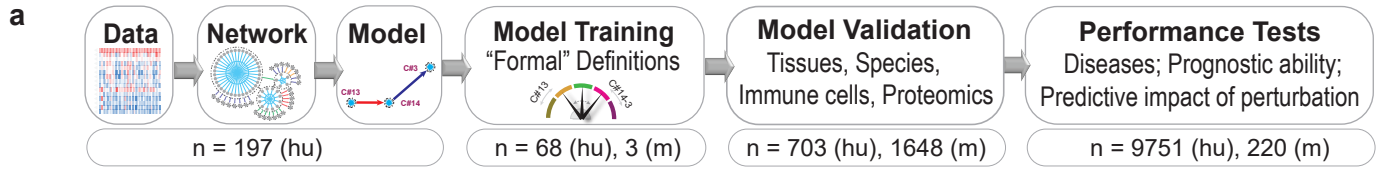


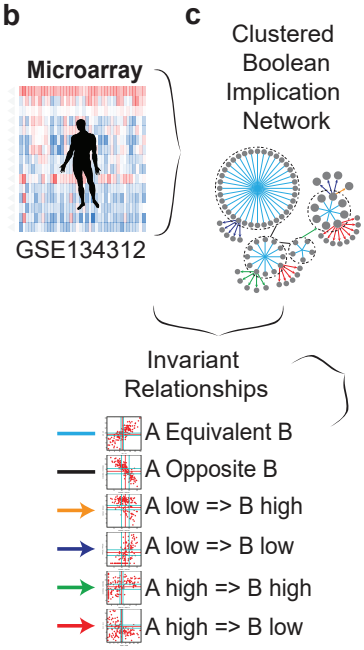


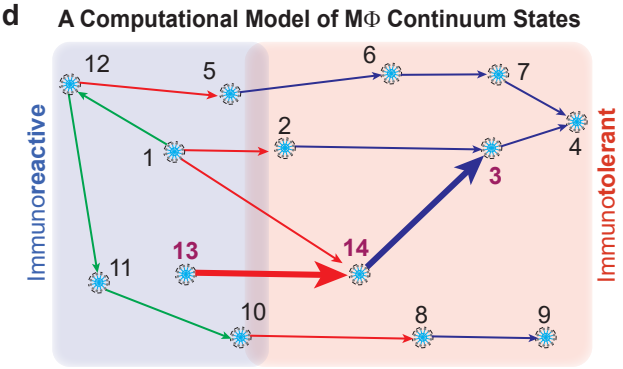


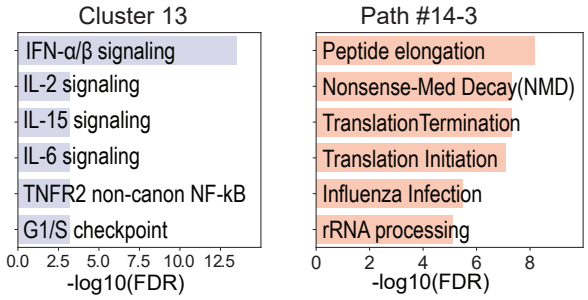


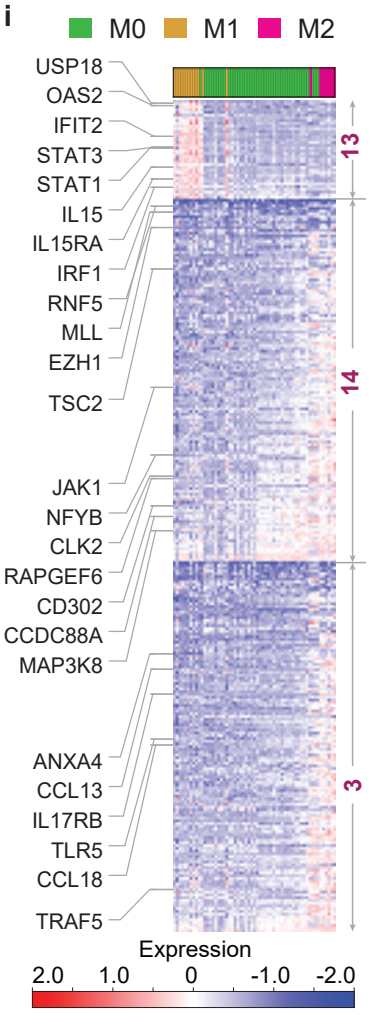
Graphic Abstract

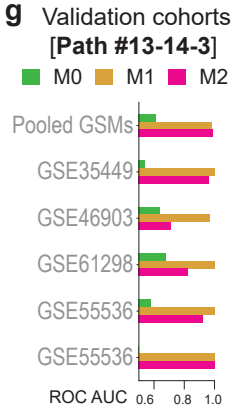


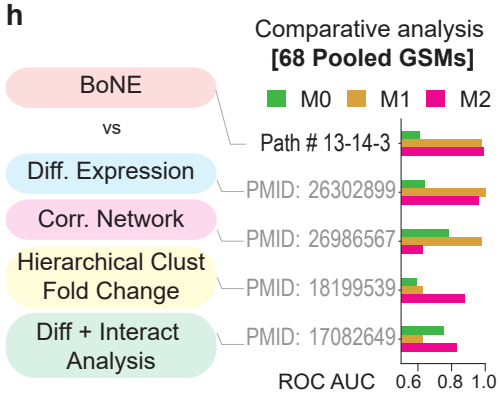


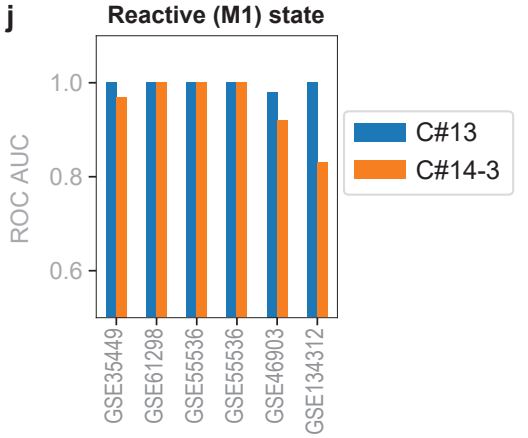












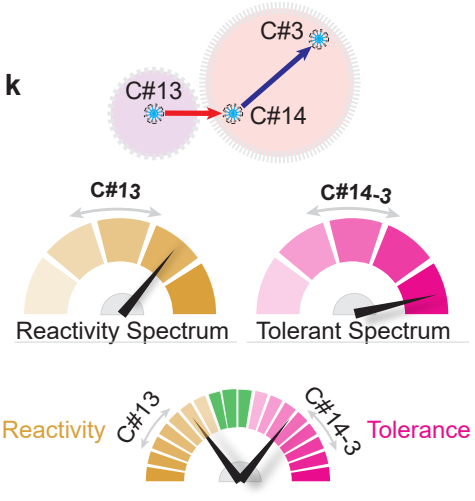


Figure 1

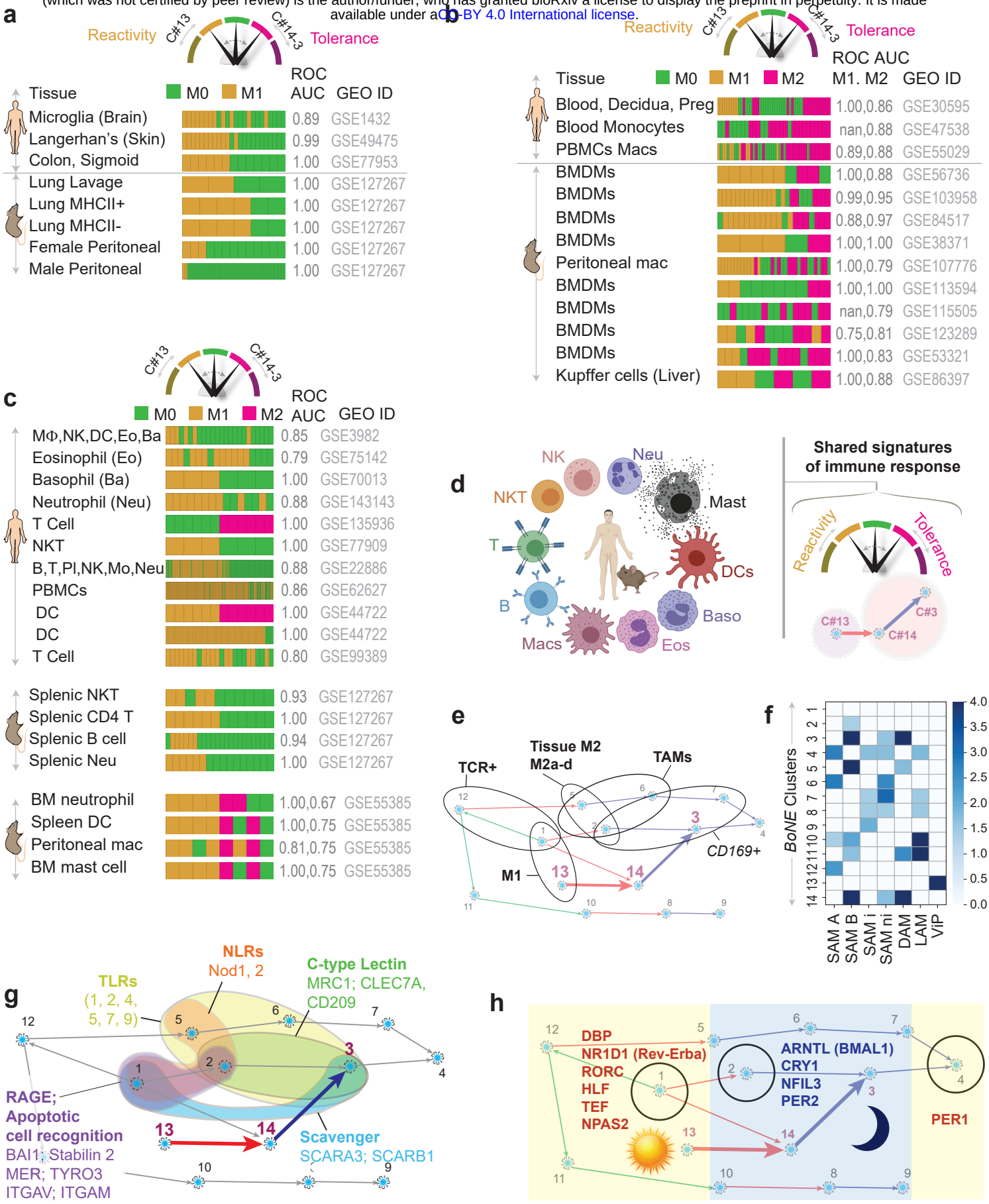


Figure 2

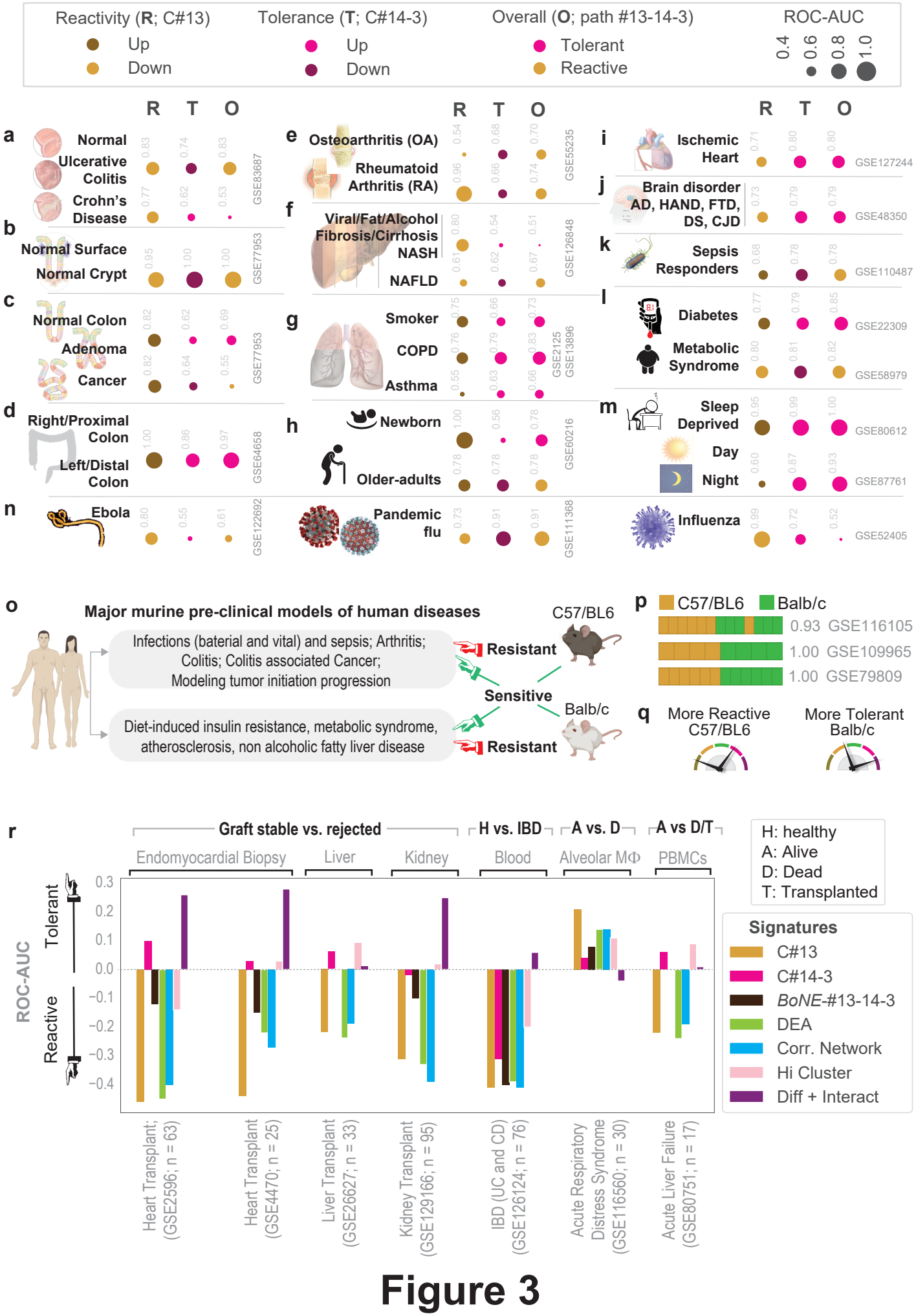


Figure 3