A personalized network framework reveals predictive axis of anti-TNF response across diseases

Shiran Gerassy-Vainberg[1,2], Elina Starosvetsky[#1], Renaud Gaujoux[#1,4], Alexandra Blatt[#2], Naama Maimon[1,2], Yuri Gorelik[2], Sigal Pressman[2], Ayelet Alpert[1], Haggai Bar-Yoseph[1,2], Tania Dubovik[1], Benny Perets[1], Adir Katz[4], Neta Milman[1], Yehuda Chowers[*1,2, 3], Shai S. Shen-Orr[*1]

on behalf of the Israeli IBD research Network (IIRN)

#, * Contributed equally to the study

Author affiliations

1. Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel
2. Department of Gastroenterology, Rambam Health Care Campus, Haifa, Israel
3. Clinical Research Institute, Rambam Health Care Campus, Haifa, Israel
4. CytoReason, Tel Aviv 67012, Israel

Corresponding author contact information:

Shai Shen-Orr, Faculty of Medicine, Technion-Israel Institute of Technology, Haifa 32000, Israel; shenorr@technion.ac.il and Yehuda Chowers, Department of Gastroenterology, Rambam Health Care Campus, Haifa, Israel; y_chowers@rambam.health.gov.il.

Abstract

Personalized treatment of complex diseases has been mostly predicated on biomarker identification of one drug-disease combination at a time. Here, we used a novel computational approach termed Disruption Networks to generate a new data type, contextualized by cell-centered individual-level networks, that captures biology otherwise overlooked when performing standard statistics. The new data-type extends beyond the 'feature level space', to the 'relations space', by quantifying individual-level breaking or rewiring of cross-feature relations. Applying disruption network to dissect high-dimensional blood data, we discover and validate that the RAC1-PAK1 axis is predictive of anti-TNF response in inflammatory bowel disease. Intermediate monocytes, which correlate with the inflammatory state, play a key role in the RAC1-PAK1 responses, supporting their modulation as a therapeutic target. This axis also predicts response in rheumatoid arthritis, validated in three public cohorts. Our findings support blood-based drug response diagnostics across immune-mediated diseases, implicating common mechanisms of non-response.

Keywords

Precision medicine, Individual-level network analysis, Drug response, Anti-TNF antibodies, Infliximab, Immune-mediated diseases, Inflammatory bowel disease, Rheumatoid arthritis, Pan-disease drug response diagnostics.

Introduction

Biologic therapies are used in a broad range of therapeutic areas including immune-mediated diseases, oncology, and hematology and have demonstrated effectiveness by improving disease clinical course, morbidity and patient quality of life. However, a subset of patients do not respond to therapy and therefore are exposed to the consequences of uncontrolled disease activity, unwanted side effects and increasing care costs. Therefore, the development of biomarkers for response prediction is an unmet medical need, necessary for achieving a favorable therapeutic index, cost/benefit ratio and overall improved patient care. Although biologics' targets are highly specific (e.g. PD1, TNFα) and target particular molecular processes across diseases (e.g. CD8 T-cell exhaustion, or TNF induced inflammation), the presence of these pathways in an individual patient is necessary but not sufficient to predict response to therapy, implying a more nuanced therapeutic mechanism which may be disease specific[1,2].

One of the most frequently used biologic drug classes are anti-TNFα antibodies, with sales of over $US 25 billion per year[3]. Anti-TNF agents are thought to exert their effects through several mechanisms, including TNFα neutralization, induction of cell and complement cytotoxicity through the FC drug fragment and cytokine suppression via reverse signaling or apoptosis[4]. Similar to other drugs and across inflammatory diseases including inflammatory bowel disease (IBD) and rheumatoid arthritis (RA), a sizable proportion of 20-40% of the treated patients, will primarily not-respond to treatment[5,6].

Previous studies used systematic screening of in-house and meta-analysis data for the identification of biomarkers associated with anti-TNFα treatment failure. Different markers were identified in different disease contexts[7]. Among these, in IBD, Oncostatin M (OSM) was identified as a potent mucosal biomarker[8]. This gene correlated closely with Triggering Receptor Expressed On Myeloid Cells 1 (TREM1), a biomarker found by us, which was predictive of response in biopsy and importantly also in blood, albeit in an inverted ratio[9]. In RA, myeloid related sICAM1 and CXCL13, and type I IFN activity were associated with anti-TNF response[10]. The identification of these markers suggests that biomarkers of pretreatment

immune status may be useful for patient screening. However, little is known regarding molecular dynamics of anti-TNF response and resistance, and whether drug biomarkers are disease dependent, or represent a patient-specific property which can be generalized across diseases.

The availability of high-resolution molecular data provides opportunities for achieving improved modeling of the complex therapeutic landscape using systems biology and network-based approaches. Yet, most of the statistical methods used are based on population averages, which do not suffice to fully investigate these complex diseases. Although several personalized approaches were recently suggested for exploring sample-level network information[11,12], these studies were not cell-centered, and did not decouple cell frequency and cell regulatory program changes. Network structure was used to identify individual alterations in cross-feature relationships between groups, however, these were validated only in the unicellular level. The same is true for the identification of individual-level time series analysis. Thus, immunologic as well as time-dependent qualifiers, within and across patients, must be accounted for when attempting to predict and reassess response to immunotherapy over the course of therapy and in context to standard methods of clinical response assessment.

We therefore employed a longitudinal cell-centered systems analysis, combining high-dimensional data of whole blood from anti-TNF responding and non-responding IBD patients at baseline and following two and fourteen weeks post first treatment. We focused on immune responses in blood, because although presenting an analytical challenge due to high background noise, blood-biomarkers have a clear advantage of accessibility, standardization, and cost-effectiveness. To understand individual variation in drug resistance, we devised a single sample-based network approach, termed 'Disruption Networks', which generates a new data type providing individual information of cross-feature relations, indicating changes in regulation. Using the new data-type information, we inferred patient-specific hypotheses for lack of response with respect to a global response network. We demonstrate that the monocytic expression of the RAC1-PAK1 axis, which is a final common pathway of multiple immune-receptor signaling cascades, is predictive of anti-TNF response in IBD as well as for the same treatment in RA, providing validation for the signature's predictivity and supporting common baseline elements that contribute to response across infliximab (IFX) treated immune-mediated diseases.

Results

Treatment response is associated with forward movement along an inflammatory axis, whereas non-responders regress.

To understand the cellular and molecular changes associated with IFX response and non-response, we performed longitudinal deep immunophenotyping of peripheral blood in Crohn's disease (CD) patients who received first-time therapy with IFX during standard clinical care (Fig. 1a, left, hereon IFX cohort). Patients were profiled by gene expression, CyTOF and Luminex, a total of three times: pre-treatment (day 0), week 2 (W2) and week 14 (W14) post-treatment initiation. At W14, 15 patients showed clinical response whereas 9 were classified as non-responders at the study end (Supp. Table 1 for clinical demographics; see Methods for response classification).

To define an individual-specific unbiased expectation of peripheral blood immune dynamics during disease course, we used a public gene expression dataset of whole blood samples from healthy individuals and 75 IBD patients in varying disease states treated with standard of care therapies (Fig. 1a, right; see Methods). We constructed an external data-driven reference IBD axis (Fig. 1b, left), which describes in a

104 dimensionality-reduced Principal Component Analysis (PCA) space the molecular transition from active-
105 through inactive disease to healthy- state, based on differentially expressed genes (hereon 'inflammatory
106 axis', see Methods). Next, we projected the position of our in-house IFX cohort on the PCA (Fig. 1b, right)
107 and calculated the distance each patient traversed on the axis over time, providing continuous molecular
108 information to characterize a patient's immune state shift (Fig. 1c). Analyzing the distance between paired
109 sample time-points, we observed that responders progressed on the inflammatory axis (*i.e.,* a positive
110 shift on the axis towards the centroid of healthy reference samples), while non-responders regressed on
111 it (Figure 1c, $P<0.05$, one-sided permutation test). Breaking up these dynamics by time point, we observed
112 that responders exhibited increased progress along the inflammatory axis following first drug treatment,
113 and reduced progress in the following period (Figure 1c). The negative correlation between progress along
114 the axis between baseline-W2 and progress in the following segment W2-W14 suggests that patients
115 progressing to 'response' early, slow down during subsequent timepoints whereas those showing a slow
116 progress initially, progress more thereafter (Fig. 1d). In fact, temporal patterns in axis progression provide
117 statistically significant context to the rate of response to therapies, that depends on immune contexture.
118 Importantly our results suggest that clinical non-responders are immunologically affected by treatment
119 as well, with an overall opposite direction from responders' progress. Collectively, our inflammatory axis,
120 captures blood molecular changes which are clinically relevant for treatment response.

121 **Early IFX response reduces expression of innate immune pathways attributed mainly to monocyte**
122 **function.**
123 To identify cellular changes following treatment in each response group, we characterized major immune
124 cell compositional changes in 16 canonical immune populations (Fig. 2, Supp. Table 2-3 for CyTOF panel
125 and Citrus clusters annotation). Then, to compare how cellular peripheral blood state differs as a function
126 of treatment response, we computed a PCA on the fold change of patients' cell phenotyping profiles (Fig.
127 2a, left). We observed significant difference in cell abundance changes between responders and non-
128 responders for W2 and W14 changes relative to baseline (P=0.005, NPMANOVA).

129 Multiple cell subset changes in responders were already apparent at W2 including reduced abundance of
130 monocytes, granulocytes, Tregs, naïve CD4+ T cells, CD4+ central memory T cells and increased abundance
131 of CD4+ and CD8+ effector memory T cells and B cells (FDR≤0.15, Paired Wilcoxon test; Supp. Fig. 1a).
132 Based on the PCA loadings we deduced that monocytes and Tregs were the primary drivers of changes
133 following treatment (Supp. Fig. 1b), evidence for which was also supported by the univariate comparison
134 showing that monocytes were significantly reduced in responders throughout both W2 and W14, whereas
135 in non-responders monocyte frequency was unchanged in W2 and elevated at W14 (P=0.0015 and
136 P=0.048 in responders, as opposed to P= 0.64 and P=0.016 in non-responders at W2 and W14 respectively,
137 Paired Wilcoxon test). Moreover, monocyte frequency was also correlated with changes in CRP
138 (Spearman's r = 0.4, P=0.01), suggesting their relevance to treatment response (Fig. 2a center, right and
139 Supp. Fig. 1c for correlation of CRP with other cell-types). Taken together, our results demonstrate
140 significant differential cell composition following IFX treatment as a function of response, with monocytes
141 likely playing a major role.

142 Given the observed cell composition alterations, we performed a cell-centered analysis to identify
143 changes in transcriptional programs following treatment in each response group, by adjusting the gene
144 expression for variation in major cell-type proportions. This procedure places focus on detection of
145 differences between conditions of the gene regulatory programs the cells are undergoing rather than
146 those differences detected due to cell compositional differences, and has been shown to unmask

147 additional signal (i.e. false-negative of direct bulk analysis) while decreasing false-positives (Fig. 2b, see
148 Methods)[9]. In this analysis, we identified 1400 (5.99%) and 589 (2.52%) differential features in responders
149 (FDR<0.15, permutation test; Supp. Tables 7) at W2 and W14 compared to baseline respectively,
150 suggesting enhanced response at W2 followed by reduced dynamics in W14. Compared to responders,
151 non-responders showed attenuated dynamics in the parallel treatment periods, with only 542 (2.32%,
152 Supp. Table 7) differential features at W2 compared to baseline, and no significantly detected dynamics
153 at W14. To ensure the differences in dynamics between the two response groups were not due to sample
154 size, we subsampled responders to match the non-responder group size and observed that responding
155 patients exhibit more dynamic changes compared to non-responders (Supp. Fig. 2). Furthermore,
156 comparing the two response groups, we observed only a minor overlap in the post treatment dynamic
157 features (23 features, 1.2% at W2). In line with the 'inflammatory axis', these results suggest that there
158 are increased early dynamics in responders compared to non-responders and that responders and non-
159 responders presented different alterations following treatment.

160 To understand the relationship during IFX response between gene regulatory programs in a biological
161 context, we constructed a cell-centered co-expression network, which was expanded by known
162 interacting genes, followed by functional enrichment analysis (see Methods, Supp. Tables 8 for network
163 edges and Supp. Fig. 3b for functional enriched pathways respectively). Interestingly, despite this being a
164 blood-based network, we noted genes which were previously associated with anti-TNF response in IBD
165 biopsies such as TREM1 and OSM[8,9], suggesting that relevant signals originally detected in tissue, are also
166 reflected in blood. We identified potential mediating pathways, i.e. pathways possessing higher
167 connectivity to other nodes in the response network, using degree and betweenness centrality
168 measurements (Fig. 2c).

169 We observed that most central pathways associated with the W2 early response were related to the
170 innate immune system (Supp. Fig. 3b). At the pathway level, consistent with the 'inflammatory axis' and
171 feature level analysis, we found augmented response at W2, which was attenuated in the following period
172 (151 vs. 88 enriched dynamic pathways in responders at W2 and W14 respectively; Supp. Fig. 3a-b). As
173 expected, among the innate related altered functions, we observed pathways related to downregulation
174 of NF-kB and TNF signaling via NF-kB (Fig. 2c, *FDR*<0.005 for W2 vs. baseline pathway score comparison,
175 by Wilcoxon test; *FDR*<0.01 for enrichment in network by GSEA). Pathways with high network centrality
176 included downregulation of FC receptor signaling and phagocytosis, cytoskeleton organization, Toll-like
177 receptors (TLRs) and vascular endothelial growth factor (VEGF) signaling responses (Fig. 2c; top 25th
178 percentile for both degree and betweenness; *FDR*<0.005 for W2 vs. baseline, by Wilcoxon test; *FDR*<0.1
179 for enrichment by GSEA). These pathways also correlated with CRP measured in the clinical setting
180 (Spearman's *r* *FDR*<0.05 and Supp. Fig. 3d). Of note, FCYR is known to be regulated by TNFα[13] and
181 mediates a number of responses, including the phagocytosis of IgG-coated particles, accompanied by
182 cytoskeleton rearrangements and phagosome formation, central pathways that were downregulated in
183 responders (Fig. 2c and Supp. Fig. 3b, *FDR*<0.001 for W2 vs. baseline, by Wilcoxon test; *FDR*<0.15 for
184 enrichment by GSEA). We also observed the downregulation of reactive oxygen species (ROS) pathway,
185 which is crucial for the digestion of engulfed materials in phagosomes (*FDR*<0.001 for W2 vs. baseline, by
186 Wilcoxon test; *FDR*<0.05 for enrichment by GSEA). This pathway was also correlated with CRP (Spearman's
187 *r* 0.43, *FDR*<0.005, Supp. Fig. 3b and Supp. Fig. 3d). To identify the most likely cell expressing these
188 pathways, we regressed the unadjusted fold change gene expression on major blood immune cell
189 abundance changes (see Methods). We observed that monocytes and granulocytes were the major
190 contributors associated with the dynamic pathways (Supp. Fig. 3c). This further supports the considerable
191 contribution of monocytes to treatment response, on top of their significant frequency alteration and
192 their frequency correlation with CRP.

'Disruption Networks' as a framework to understand individual variation in non-responders' dynamics. Whether non-responders' transcriptional profile reflects fundamental routes of IFX resistance, is essential for tailoring treatment. To elucidate molecular mechanisms of individual-specific pathways of treatment non-response, we devised a systematic framework we term 'Disruption Networks' which generates a new data-type to provide individual-level information of cell-centered changes in cross-feature relations. The generation of the new data-type relies on studying relations between features across a predefined reference population of individuals (i.e., a population level reference network), and then inferring how these relations differ (i.e., are disrupted) at the single sample level. The new data-type can serve as an input to multiple analyses including integration, differential signal detection, patient stratification based on disruption profile, assessment of disruption in functional modules and evaluation of individual's molecular network behavior under specific perturbation effects or biological conditions (Fig. 3a).

To identify how non-responding-individuals differ with respect to the IFX response dynamics, we iteratively added a single non-responding patient to the response reference network we had studied and calculated the disruption in the correlation structure in each edge for that patient (hereon 'dropout'). This procedure was performed separately for each non-responder. We considered only negative dropouts, that is, events in which the relation (i.e., correlation) between two features was weakened once the non-responder data was spiked into the responders' group, indicating deviance from treatment response (Supp. Fig. 4a, for an example). To evaluate non-responders' dropout significance, we generated empirical null distribution of dropouts ('normal response' dropouts) by iterative addition of each responder's sample to the other responders' samples. We calculated P-values as a left-tail percentile, within the null distribution of the normal dropouts, which were further corrected for multiple testing (Fig. 3a; see Methods). By applying the 'Disruption Networks' framework, we considerably expanded the detected differential signal between response groups as compared to standard differential analysis (one feature by Wilcoxon test (*FDR*<0.1) vs. 180 features by mean drop intensity, including the single feature identified by Wilcoxon test (*FDR*<0.1 for dropout significance and 10th top percentile of mean drop intensity); Fig. 3b and Supp. Fig. 4b for mean drop intensity, disrupted edge ratio parameters and the agreement of both respectively).

To understand disruption in the functional context, we aggregated the dropouts to calculate a pathway-level personalized disruption (Fig. 3c for mean drop intensity and Supp. Fig. 4c for disrupted edge and node ratio parameters; see Methods). We found that the major disrupted dynamics at W2 was related to the cytoskeleton/fiber organization and VEGFR signaling which were central functions during normal treatment dynamics. Interestingly, nodes related to these disrupted pathways exhibited high centrality (*P*<9.999e-05 and *P*=0.034 for degree and betweenness correspondingly by permutation test; Fig. 3d). On the meta-pathway level, monocytes were the most central cell-type associated with the disrupted pathways (Fig. 3e, left, top 5th percentile for degree and betweenness centrality). The disrupted meta-pathway included the core genes consisting of the HCK-RAC1-PAK1 signaling cascade, which presented high combined degree and betweenness centrality (*P*=0.017, n=1000 random triple node subsampling). This core perturbed axis is a final common pathway involving signaling through several proximal immune-receptors by a range of inflammatory ligands including chemokines, growth factors such as VEGFR, and FC receptor ligands which induce FC-mediated phagocytosis involving coordinated process of cytoskeleton rearrangement[14]. Indeed, these pathways were functionally enriched in the disrupted meta-pathway (q-value<0.05, hypergeometric test; Fig. 3e, right). The latter are also linked to ROS and NADPH oxidase activation through the regulation of RAC1[15].Of note, suppression of RAC1-PAK1 signaling, predominately in innate immune cells was shown to mediate remission in CD[16]. Taken together, these observations showcase the power of 'Disruption Networks' to identify masked, individual level, signal and suggest that the RAC1-PAK1 signaling cascade, is significantly disrupted in non-responders, during treatment.

239 **RAC1-PAK1 signaling is elevated in responders' peripheral monocytes pre-treatment.**

240 We next asked whether cellular programs found to be disrupted during treatment dynamics can be
241 identified pre-treatment, since direct differential analysis in the feature expression space did not yield
242 significant signal. Looking at the feature level, we found that most of the pre-treatment differentially
243 expressed genes were increased in responders, including genes involved in the RAC1-PAK1 axis (*FDR*<0.1,
244 Wilcoxon test, Supp. Fig. 5a). On the pathway level we observed that the fiber organization pathway,
245 presented pre-treatment disparity between the two response groups (*FDR*<0.1, NPMANOVA) and
246 correlated with clinical CRP (Spearman's *r* =0.4, *P*=0.06), in addition to its high centrality in the response
247 network (Fig. 4a). The relative pathway score of the cytoskeleton-organization pathway was higher in
248 responders pre-treatment compared to non-responders (*P*<0.0006, one-tailed Wilcoxon test), and was
249 downregulated following efficient treatment (*P*<0.001 and *P*<0.05 for W2 and W14 compared to baseline,
250 one-tailed Wilcoxon test; Fig. 4b). This was in contrast to non-responders which showed insignificant
251 dynamics at W2 and even an opposite trend in W14 (*P*=0.52 and *P*=0.041 for W2 and W14 compared to
252 baseline, one-tailed Wilcoxon test; Fig. 4b).

253 The fiber organization pathway associated with treatment dynamics and response already at pre-
254 treatment state, represents distinctive differences in cellular transcriptional states between response
255 groups, rather than differences reflecting cellular composition alterations, as our analyses accounted for
256 cell proportions. Therefore, we next aimed to dissect the cellular origin of the fiber organization related
257 core genes. First, we tested the correlation between the canonical cellular frequencies as obtained by
258 CyTOF, and the bulk unadjusted expression of the fiber organization genes (Supp. Fig. 5b). We observed
259 that the majority of the genes in the target pathway were positively associated with monocytes
260 abundance. To further validate the cellular origin and the fiber organization related transcriptional cell
261 state in the two response groups, we performed single-cell RNA sequencing (scRNA-seq) using peripheral
262 blood mononuclear cells (PBMCs) from pre-treatment samples of a representative responder and non-
263 responder (Fig. 4c; see Methods). Assessment of the fiber organization related expression in the cellular
264 level, confirmed that monocytes were highly associated with the distinctive pathway expression (*P*<2.2e-
265 16, for expression in monocytes compared to the other cell types, Wilcoxon test, Fig. 4c, right and Supp.
266 Fig. 6).

267 To understand the molecular events associated with the fiber organization pathway in the relevant cell
268 and subset specific context, we expanded the fiber organization differential genes through intersection of
269 knowledge- and data-driven based networks (see Methods). Then, we assessed the pathway related
270 expression in monocyte subsets, which were previously shown to exhibit distinct phenotypes and
271 functions in health, and immune-mediated disease states[17]. The results indicated that intermediate
272 monocytes contributed most to the fiber organization distinctive expression between the response
273 groups, pre-treatment (|FC|=2.13, *P*<2.2e-16 in intermediate monocytes vs. |FC|=1.3, *P*<2.2e-16 and
274 |FC|=1.1, *P*<0.05 in classical and non-classical monocytes respectively by Wilcoxon test, Fig. 4d).
275 Interestingly, we detected significantly increased membrane TNF (mTNF) on intermediate monocytes
276 compared to the other subsets, by CyTOF (*P*<5e-07, one-tailed Wilcoxon test, Fig. 4e), suggesting these
277 cells serve as drug targets, thereby explaining their tight linkage to drug response.

278 **Pre-treatment RAC1-PAK1 axis is predictive for IFX response across immune mediated diseases.**

279 We next tested whether the pre-treatment fiber organization pathway could predict treatment response
280 (see Methods). We observed that the pathway score of a set of 6 core genes (RAC1, PAK1, LYN, ICAM1,
281 IL1B and FCGR3A) could discriminate responders from non-responders at a mean AUC of 0.90 (95CI 0.74,
282 1; P=0.0001 by Permutation test), supporting a common mechanism of non-response to treatment (Supp.
283 Fig. 5c). By applying targeted network analysis of the predictive fiber organization pathway in
284 intermediate monocytes, we found that the FCYR signaling and functionally related pathways including
285 phagocytosis and ROS metabolism were highly enriched in the co-expression network effectively
286 differentiating between response groups at baseline (Supp. Fig. 7).

287 To further validate our findings, we tested an additional independent validation cohort of 29 CD patients,
288 which were naive to biological treatment and were treated with thiopurines or steroids only as a co-
289 therapy (Supp. Table 9 for clinical demographics). The results indicated that the pre-treatment RAC1-PAK1
290 axis, was differentially expressed between response groups in the validation cohort (P<0.01, Wilcoxon
291 test) as well, supporting the primary findings and thereby demonstrating that reduced pre-treatment
292 expression of the RAC1-PAK1 axis is associated with non-response (AUC=0.78; Fig. 5a).

293 To assess whether the predictive RAC1-PAK1 axis is disease dependent or whether it could be generalized
294 across diseases, we tested public datasets of blood samples from RA patients, pre-IFX treatment
295 (GSE20690[18], GSE33377[19], GSE42296[20]). Gene expression was adjusted to major cell type contributions
296 which was evaluated by deconvolution (see Methods). The results confirmed the increased pre-treatment
297 expression of the axis genes in RA responders, (representative cohort GEO20690, Fig. 5b). Application of
298 fiber organization predictive signature to multiple pre-treatment RA cohorts separated IFX response
299 groups effectively (Meta ROC AUC=0.72, Fig. 5c). These findings expand the predictive value of the RAC1-
300 PAK1 axis to other IFX-treated related diseases such as RA. Taken together, these observations
301 demonstrate that the baseline RAC1-PAK1 axis expression in monocytes differentiates response groups
302 and ultimately impacts response potential across immune-mediated diseases.

303 Discussion

304 Despite substantial inter-individual heterogeneity and our growing ability to measure it, commonly used
305 statistical frameworks for analyzing high-dimensional data describe changes happening on average
306 between conditions or groups. This is especially true in the case of networks which form a natural way of
307 describing the possible interactions occurring between measured biological species, yet are population-
308 based, and thus limited in their ability to monitor individual variation from those interactions and the
309 ensuing emergent phenomena these interactions yield. Here we studied the dynamics of IFX response in
310 IBD, in a small cohort, over time. To address this challenge, we devised the 'Disruption Networks'
311 approach, a cell-centered personalized statistical framework which unmasks differences between
312 individuals. The approach enables a systematic dissection of IFX effect on response dynamics from blood,
313 by generating a new data-type which quantifies individual-level breaking or rewiring of cross-feature
314 relations. The generated data-type is cell-centered considering both cellular composition changes and
315 changes in cellular regulatory programs, allowing us to identify robust functional pathways deviating from
316 normal response in non-responders, and robustly associate these with drug resistance in both IBD and
317 RA.

318 Although TNF is a pleiotropic cytokine, functioning in both the innate and adaptive immune system[21], we
319 found that the early response alterations following IFX treatment were mostly related to innate pathways
320 of which monocytes were the major driver. Evidence supporting this has been previously implicated by

321 the decreased frequency of monocytes during treatment in anti-TNF treated IBD[22] and RA[23] patients.
322 Furthermore, the anti-proliferative and cell-activation suppressive effect of IFX was shown to depend on
323 FC-expressing monocytes in a mixed lymphocyte reaction[24]. In addition, the regained long term response
324 following granulocyte/monocyte adsorption treatment following loss of response during IFX treatment
325 further corroborates our findings[25]. Taken together, these results support the potential for subset specific
326 targeted therapy to augment IFX treatment.

327 By applying the 'Disruption Networks' framework, we identified RAC1-PAK1 signaling, as a central
328 pathway associated with IFX response. This pathway exhibited disrupted dynamics in non-responders and
329 was predictive of treatment response at baseline. Although abnormal RAC1 signaling was linked to
330 immune-mediated diseases pathogenesis[26], its direct relation to anti-TNF response has not been
331 demonstrated. The RAC1-PAK1 axis is a final common pathway shared by several proximal immune
332 receptors, controlling actin cytoskeletal movement, activation of the respiratory burst and phagocytic
333 activity in innate cells. RAC1 was identified as a susceptibility gene for IBD[27], and TNF was shown to
334 stimulate RAC1-GTP loading[16], supporting efficacy of antagonizing this effect by anti-TNF. In line with our
335 findings demonstrating IFX suppressive effect on the RAC1-PAK1 axis during treatment, thiopurines,
336 another effective IBD treatment were also shown to inhibit RAC1 activity[28]. The superior effect of anti-
337 TNF -thiopurines combination over monotherapy[29] suggests that the enhanced therapeutic effect is
338 mediated not only by controlling anti-drug antibody (ADA) levels, but conceivably also by the induction of
339 a mutual additive effect on RAC1 suppression. Interestingly, the TREM adaptor (TYROBP/DAP12), which
340 we previously found to be predictive for anti-TNF response by meta-analysis[9], was detected in the
341 differential RAC1-PAK1 signature, exhibiting significant correlation with the RAC1-PAK1 axis in monocytes,
342 and is also functionally related through shared signaling[30] .

343 The monocytes single-cell based RAC1-PAK1 co-expression network demonstrated pre-treatment
344 differential expression, primarily in intermediate monocytes, related to FcγR dependent phagocytosis and
345 interferon signaling. This is consistent with prior reports showing that FcγR affinity affects anti-TNF
346 therapeutic response[31–33]. Interestingly, the RAC1-PAK1 axis was predictive of IFX responsiveness also in
347 RA, an observation which provides additional validation for the signature predictivity and supports
348 common baseline elements contributing to response across IFX-treated immune-mediated diseases.
349 Similarly to IBD, also in RA, the RAC1-PAK1 upstream activator FcγR was linked to disease
350 susceptibility[34,35]. The FcγR3A, which is a part of the predictive signature, is known as a key receptor for
351 monocytes effector response including antibody-dependent cellular cytotoxicity (ADCC), immune IgG-
352 containing complexes clearance and phagocytosis[36,37]. These further corroborate the common element of
353 enhanced RAC1-PAK1 signaling through increased expression or affinity for FcγR3A expressed on
354 monocytes that may enhance the efficacy of IFX in IBD and RA. These results extend the relevance of
355 molecular commonalities for disease activity[38] and pan-pathology[39], also to interconnected pathways of
356 drug responsiveness across immune-mediated diseases.

357 Whether the RAC1-PAK1 axis and the upstream FcγR are applicable to IFX response in other immune-
358 related diseases or other anti-TNF therapeutic antibodies remains to be determined. While we identified
359 the RAC1-PAK1 axis as predictive for therapy response in IFX-naive patients, our results do not yet provide
360 an understanding of how this axis is expressed in previously-treated patients. Considering the backwards
361 immune shift in non-responders along the 'inflammatory axis' we identified, analysis of previously-treated
362 patients should be addressed separately. The 'inflammatory axis' further provides a potential explanation
363 for the inferior response rates to subsequent treatments in treatment-experienced compared to naïve

364 patients treated with the same agents[40]. Of note, our real-life cohorts consisted of clinically comparable
365 responding and non-responding groups, in terms of demographics and concurrent therapies, except for
366 lower drug levels in non-responders at W14 in the primary cohort. The disrupted axis was identified at
367 the early W2 response period in which drug levels were comparable and thus response is not expected to
368 be affected by the subsequent difference. In this context, the lower drug levels are likely a consequence
369 rather than a cause of non-response, maybe due to "inflammatory sink" drug consumption, or drug loss
370 through a "leaky gut"[41,42].

371 Blood-based pre-treatment biomarkers are highly important for precision medicine, since when identified
372 across diseases and drugs as performed here, they offer the vision of data-driven choices for physician
373 treatment and personalized care. Our results suggest that the road to this vision may be shorter than
374 anticipated, as at least for immunotherapies, blood is a relevant tissue for signal detection and non-
375 response mechanisms appear to be conserved across immune-mediated diseases. We note that this pan-
376 disease drug response conserved pattern may not necessarily hold in biopsies from the site of disease,
377 which being different tissues, may present different cells playing a role. Our combined experimental-
378 computational approach, where small time series experiments are combined with an individual-level
379 analytical framework, can be generalized to other diseases and conditions including mechanisms of drug
380 mode of action, drug non-response, comparison of drug effects and disease courses. These will ultimately
381 allow to make sense of blood and accelerate an era of immune-based precision diagnostics.

382

383 # Methods

384 **Patients and study design**

385 **Primary real-life IBD cohort**

386 A primary real-life cohort consisting of 24 Crohn's disease (CD) patients who received IFX treatment at the
387 gastroenterology department of the Rambam Health Care Campus (RHCC). All patients met the study
388 inclusion criteria as follows: 1) Adequately documented active luminal CD, as phenotyped by a
389 gastroenterologist with expertise in IBD. 2) Documented decision to initiate full IFX induction regimen
390 with 5 mg/kg induction dosing (i.e., at weeks 0, 2, 6). Patients that had past exposure to Infliximab,
391 Adalimumab or Vedolizumab, or patients who had active infection including febrile diseases or intra-
392 abdominal or perianal abscess were excluded. The study was approved by the institutional review board
393 (0052-17-RMB), and patients provided written informed consent. Demographic and clinical characteristics
394 of the patients are shown in Supp. table 1.

395 Patient samples were obtained at three time points: at baseline, before IFX treatment, and two and
396 fourteen weeks post first treatment and assayed for gene expression microarray data, high-resolution
397 granulocytes and lymphocytes subtype frequencies and functional markers by CyTOF, and a panel of 51
398 cytokines and chemokines by Luminex. CyTOF panel including Clone, vendor, and conjugation
399 information, and Luminex panel are detailed in Supp. table 2 and 3 respectively.

400 Patient response classification was defined by decision algorithm, which we used and described previously
401 [9]. Briefly, patients were classified as responders based on clinical remission, which was defined as
402 cessation of diarrhea and abdominal cramping or, in the cases of patients with fistulas, cessation of fistula
403 drainage and complete closure of all draining fistulas at W14, coupled with a decision of the treating

404 physician to continue IFX therapy at the current dosing and schedule. In patients that were initially
405 clinically defined as partial responders, classification was determined by a decision algorithm that
406 included the following hierarchical rules: 1) steroid dependency at week fourteen; 2) biomarker dynamics
407 (calprotectin and CRP) and 3) response according to clinical state at week 26. Applying the decision
408 algorithm and exclusion criteria, yielded a final study cohort of 15 and 9 responding and non-responding
409 patients respectively.

410 As shown in Supp. table 1, responders significantly reduced CRP, already at W2 post first treatment while
411 non-responders presented a trend of reduced CRP at W2, but their CRP level following 14 weeks was
412 elevated and significantly higher than CRP level in responders. No significant difference was found in
413 target TNFα levels, neither in responders or non-responders, as measured by either serum cytokine level
414 using Luminex or by adjusted gene expression. As expected, IFX drug levels were shown to be significantly
415 reduced, in both responders and non-responders at W14 compared to W2, due to the transition from
416 induction to maintenance therapy. Drug levels of responders were significantly higher compared to non-
417 responders at W14. However, at W2, no significant difference in drug levels was measured. Responders
418 also showed improved albumin levels along treatment, with significantly higher levels compared to non-
419 responders at W14. All other parameters were comparable between the two response groups.

420 **Validation real life IBD cohort**

421 The validation cohort consisted of 29 CD patients from the RHCC, which were classified to 20 and 9 clinical
422 responding and non-responding respectively patients according to the above-described decision
423 algorithm (Supp. table 9).

424 **CyTOF sample processing and analysis**

425 A total of 2×10^6 cells of each sample were stained (1 h; room temperature) with a mixture of metal-
426 tagged antibodies (complete list of antibodies and their catalog numbers is provided in Supp. table 2). This
427 mix contained antibodies against phenotyping markers of the main immune populations and some central
428 cytokine and chemokine receptors. All antibodies were validated by the manufacturers for flow
429 application (as indicated on the manufacturer's datasheet, available online) and were conjugated by using
430 the MAXPAR reagent (Fluidigm Inc.). Iridium intercalators were used to identify live and dead cells. The
431 cells were fixed in 1.6% formaldehyde (Sigma-Aldrich) at 4°C until they were subjected to CyTOF mass
432 cytometry analysis on a CyTOF I machine (Fluidigm Inc.). Cell events were acquired at approximately 500
433 events/s. To overcome potential differences in machine sensitivity and a decline of marker intensity over
434 time, we spiked each sample with internal metal-isotope bead standards for sample normalization by
435 CyTOF software (Fluidigm Inc.) as previously described[43].

436 For data preprocessing, the acquired data were uploaded to the Cytobank web server (Cytobank Inc.) to
437 exclude dead cells and bead standards. The processed data were analyzed using Citrus algorithm, which
438 performs hierarchical clustering of single cell-events by a set of cell-type defining markers and then assigns
439 per sample, per cluster its relative abundance in each sample as well as the median marker expression for
440 each functional marker per cluster[44]. Citrus analysis was applied separately on PBMCs and Granulocytes
441 population in each sample using the following parameters: minimum cluster size percentage of 0.01 and
442 0.02 for PBMCs and Granulocytes respectively, subsampling of 15,000 events per sample and arcsin
443 hyperbolic transform cofactor of 5. The gating for the classification of the clusters is detailed in Supp. table
444 3.

445 **Blood transcriptome analysis**

446 Whole blood was maintained in PAXgene Blood RNA tubes (PreAnalytiX). RNA was extracted and assayed
447 using Affymetrix Clariom S chips (Thermo Fisher Scientific). The microarray data are available at the Gene
448 Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/). The raw gene array data were
449 processed to obtain a log2 expression value for each gene probe set using the RMA (robust multichip
450 average) method available in the affy R package. Probe set annotation was performed using affycoretools
451 and clariomshumantranscriptcluster.db packages in R. Data were further adjusted for batch effect using
452 empirical Bayes framework applied by the Combat R package.

453 Gene expression data were further adjusted for variations in frequency of major cell types across samples
454 as measured by CyTOF, including CD4+ T cells, CD8+ T cells, CD19+ B cells, NK cells, monocytes and
455 granulocytes, to allow detection of differential biological signals that do not stem from cell proportion
456 differences, which might be otherwise masked in unadjusted gene expression data. Adjustment was
457 performed using the CellMix R package.

458 **Cytokines and chemokines measurement using Luminex bead-based multiplex assay**

459 Serum was separated from whole blood specimens and stored at -80°C until used for cytokine
460 determination. Samples were assayed in duplicate according to the manufacturers' specifications
461 (ProcartaPlex™ Immunoassay, EPX450-12171-901, eBioscience, Cytokine/Chemokine/Growth Factor 45-
462 Plex Human Panel 1, Supp. table 4).

463 Data were collected on a Luminex 200 instrument and analyzed using Analyst 5.1 software (Millipore) and
464 NFI (Median Fluorescence Intensity) values were used for further data processing. A pre-filtering was
465 applied as follows: samples with low mean bead count, below 50 were excluded from analysis. In addition,
466 duplicates with high CV values (Coefficient of variation) above 40% were omitted. NFI values with low
467 bead count, below 20 were filtered out, but in cases which one replicate had acceptable bead count and
468 the CV values for both replicates were less than 25%, NFI values were retained.

469 Finally, net MFI values were calculated by blank reduction followed by log2 transformation. Data were
470 further adjusted for batch effect using the empirical Bayes framework applied by the Combat R package.

471 **Characterization of IFX responders and non-responders' dynamics through integrative molecular**
472 **response axis combining external and in-house data**

473 An integrative molecular response axis was constructed to recapitulate the complex nature of anti-TNFα
474 response progression dynamics which enables to track individual immune dynamics of both responding
475 and non-responding patients. This methodology was assessed using an external data-based axis.

476 For unbiased definition of the 'inflammatory axis' and validation of our own data we used public gene
477 expression data of whole blood from 25 UC patients and 50 CD patients in active or inactive disease states,
478 available in Gene Expression Omnibus (GSE94648). The patients in this external cohort were treated with
479 different medications including 5-ASAs, Immunosuppressants, anti-TNF agents, steroids and combinations
480 of these therapies, as previously described[45], representative of a relatively large portion of the treated
481 IBD patient population. The analysis was performed in several steps: (1) Differential expression analysis
482 between active disease and healthy states for UC and CD separately (Supp. Table 5), using the limma R

483 package, followed by PCA (Principal Component Analysis). (2) Ordinal lasso was used to select the principal
484 components that best describe the desired directionality from active through inactive to healthy state,
485 based on optimal absolute coefficient values and percentage of variance explained parameters (Supp.
486 Table 6). (3) The 'inflammatory axis' coordinates were defined based on initial and terminal points
487 determined as the mean of the two end-point coordinates of active and healthy states. (4) Applying vector
488 multiplication (dot product) for the calculation of the projection of sample vector from our in-house
489 cohort in the direction of the external 'inflammatory axis', to estimate sample position on the axis. (5)
490 Evaluation of the distance of patient samples between two time points based on sample axis location.

491 **Multi-omics network of anti-TNF blood response dynamics**

492 **Core co-expression response network**

493 To identify features that change over time in responders, a linear mixed-effects model was used, in which
494 time was treated as a fixed effect and individuals were treated as a random effect (lmer R package) to
495 allow testing differential expression by time while accounting for between-subject variations. P-values
496 were calculated empirically through a permutation test (n perm=1000). In each permutation, feature
497 measurements were shuffled between visits for each responding patient. Permutation based p-values
498 were obtained by comparing the absolute value of the non-permuted β coefficient for each feature to the
499 null distribution of permuted β coefficients for the same feature. In order to calculate FDR based on the
500 permutation results, permuted p-value was determined for each permuted β coefficient, by comparing
501 the tested permuted β coefficient to the distribution of the other permuted β coefficients for each feature.
502 Then FDR was estimated by comparing the non-permuted p-values to the null distribution of the
503 permuted *p-values*. A similar calculation was performed for non-responders (max n perm =512).
504 In addition to the determination of dynamic features in the full responders' sample data, a random
505 subsampling of samples from the responders group, without replacement, was applied to achieve equal
506 sample size between responders and non-responders. Two-hundred subsamples were generated and
507 tested using linear mixed-effects models. In this part, for the comparison of equally sized responders and
508 non-responders' groups, p-values were calculated based on the t-statistic using the Satterthwaite
509 approximation, implemented in the lmerTest R package, followed by multiple hypotheses correction
510 using the Benjamini-Hochberg procedure.
511 Co-expression network based on V1-V2 fold-change expression values of the significantly altered features
512 (FDR<0.15) was constructed, based on pairwise Spearman's rank correlation using the psych R package.
513 Filtering was applied to remove feature-pairs with insignificant correlation with a cutoff of FDR<0.1.

514 **Network propagation**

515 Network propagation procedure was applied to enhance the biological signal of the obtained networks as
516 previously described [46] with slight modifications. Briefly, for each node in the network, protein interactors
517 with a combined score above 700 were extracted based on STRING database (functional protein
518 association networks; https://string-db.org/cgi/download.pl) using STRINGdb R package. A node
519 interactor was added as a linker gene to the network if its own interactors (hubs) were significantly
520 enriched in the core network features. Enrichment was calculated using the hypergeometric test in the
521 stats R package. Calculated p-values were adjusted for multiple hypotheses using the Benjamini-Hochberg
522 procedure. A cutoff of FDR<0.05 was selected for significant enrichment of the tested interactor hubs in
523 the immune network.

524 **Functional enrichment assessment for the response network**

525 To assess dynamics in the functional level, genes were grouped to functional sets by using a semi-
526 supervised approach combining both network structure and known gene set annotations from Hallmark,
527 Kegg, Reactome, Biocarta, PID and BP Go terms. Each edge in the network was classified to a specific
528 pathway if its two linked nodes were annotated in the same biological group. Pathways with less than 5
529 mapped edges were filtered out. This was followed by a global gene set enrichment analysis using fGSEA
530 (FDR<0.15, nperm=1000, minSize=10, maxSize=400).

531 The dynamic enriched pathway structures were further tested for significance by comparing the density
532 (graph density score) of each pathway associated sub-network to a parallel sub-network density obtained
533 from 100 random networks with a matched size according to the Erdos-Renyi model which assigns equal
534 probability to all graphs with identical edge count (igraph R package). P-value was evaluated as the
535 proportion of random module density scores that were higher than the real module density score.
536 Additional filtering was applied according to the number of connected components in a pathway sub-
537 graph (igraph R package). Only highly connected pathways (percentage of largest connected
538 component>50%, size of the connected component>10) were included.
539 The dynamic pathways list was further condensed by filtering out high overlapping pathways using Jaccard
540 index. Accordingly, in overlapping pathways pairs that presented a Jaccard index above 0.5 the smaller
541 module was omitted.
542 To further associate the assigned pathways with treatment response, the Wilcoxon test was used to
543 compare V1 to V2 and V1 to V3 relative pathway scores in responders and non-responders. p-values were
544 adjusted for multiple hypotheses using the Benjamini-Hochberg procedure (FDR<0.05). Relative pathway
545 scores were calculated for each sample as previously described [38,47] (see Relative pathway score
546 evaluation). To assess cellular contributions for each pathway, the non-adjusted expression of each gene
547 in the dynamic pathways was regressed over the major peripheral cell type frequencies as determined by
548 CyTOF including granulocytes, CD4 and CD8 T cells, B cells, NK cells and monocytes. The cell-specific
549 contribution to each pathway was determined as the mean of the coefficients of the tested cell type across
550 all genes in the module. The centrality of each pathway in the response network was also evaluated by
551 calculating the pathway based mean betweenness and degree across all gene members of the pathway
552 (igraph R package). To further assess the clinical relevance of the dynamic pathways to the treatment
553 response, the calculated pathway score at all tested time points was correlated with CRP using Spearman's
554 rank correlation test.

555 **Relative pathway score evaluation**

556 The expression of each gene in the pathway was standardized by the z-score transformation, to enable
557 comparable contribution of each gene member to the pathway score, followed by mean value calculation
558 across the transformed genes in the pathway for each sample.

559 **'Disruption Networks' framework**

560 To understand individual variation in non-response dynamics, we developed an approach termed
561 'Disruption Networks' in which individual non-responders are iteratively added to the obtained normal
562 IFX response network, and the disruption in the correlation structures is assessed for each edge in the
563 reference response network. The disruption is evaluated in the node (gene/cell) or the module level to
564 determine biological mechanisms that may explain patterns of the non-response.

565 More specifically, consider a feature matrix $F_{n\times m}$ where n is the number of samples for a given condition,
566 in our case, n is the number of samples of responding patients and *m* is the number of features, where

f(i,j) refers to a fold change measured value at a given time point relative to baseline, of the j-th feature in the i-th sample. Let matrix $R_{m\times m}$ be the feature pairwise Spearman's rank correlation matrix based on F which represents the global response network, where r(j,k)=cor(j,k) for genes j and k. Insignificant correlation values according to FDR thresholds, as described above, were presented as NAs in the matrix.

The 'Disruption Networks' construction was assessed individually for each non-responder as follows: a new $F'_{(n+1)\times m}$ matrix was generated by the addition of the tested non-responder to the responders' samples. Based on F', a new pairwise Spearman's rank correlation matrix was calculated to obtain $R'_{m\times m}$, in which r'(j,k) is the correlation between j and k genes when including the non-responder in the responders' samples.

For correlation coefficients comparison, correlation coefficient values were transformed using Fisher z-transformation by the following formula:

$z(r) = 0.5 * ln(\frac{1+r}{1-r})$ and a standard error of $SEz(r) = \frac{1}{\sqrt{n-3}}$ where n is the number of samples. We define a 'disruption' term as the drop in the Fisher z transformed values between two genes as a result of the non-responder addition using the statistical z score which is defined as:

$$disruption(j,k) = z\ score = \frac{z(r\prime)-Z(r)}{Pooled.SEz} = \frac{z(r\prime)-Z(r)}{\sqrt{\frac{1}{(n+1)-3}+\frac{1}{n-3}}}.$$

Only negative values of $sign(r * (z(r') - Z(r)))$, which indicate weakening of the original correlation obtained in responders were included, while positive values were set to zeros. Drop degree of confidence for non-responders was assessed empirically for each drop value in each edge, based on the non-responder drop value percentile in the responders' normal drop distribution. This was further corrected for multiple testing using the Benjamini-Hochberg procedure. Edges with drop adjusted percentile <0.1 were considered as significantly disrupted. Insignificant drop values were set to zeros. Analysis of disruption parameters in the feature level, revealed a considerably expansion of the detected differential signal between response groups, compared to standard differential analysis by Wilcoxon test. While using the Wilcoxon test we detected only one feature (0.06%), with significant differential dynamics between response groups at W2, we identified this feature together with 179 additional features (10%) when using disruption parameter of top mean drop intensity (FDR<0.1 by Wilcoxon test, FDR<0.1 for significant dropout and top 0.1 percentile of mean drop intensity, Figure 3b). We observed similar results for the disrupted edge ratio (0.06% Vs. 14.4% significant features identified by Wilcoxon test (FDR<0.1) and top disrupted edge ratio parameter (FDR<0.1 for significant dropout and top 0.1th percentile of node disrupted edges) respectively, Supp. figure 4a). Testing the agreement of both disruption parameters, we identified 9.4% dynamics differential features including the single feature identified by Wilcoxon test (Supp. figure 4b).

Disruption was also measured in the pathway level for each individual using three different measurements: (1) Pathway specific mean drop intensity in which a mean drop intensity was calculated across the relevant edges in the module, for a specific individual. (2) Pathway specific percentage of disrupted edges which determines the percentage of edges in the pathway that the specific individual is

603 significantly disrupted in. (3) Pathway specific percentage of disrupted nodes which evaluate the
604 percentage of disrupted nodes for a specific individual out of all module nodes.

605 For binary classification of disrupted pathways, we quantify the disruption measure across a range of
606 percentile values in each parameter. For each parameter, in each percentile, the selected positive
607 disrupted modules were those that were disrupted in at least 50% of the non-responding patients and in
608 less than 20% of the responders, or in cases where the difference between the percentage of disrupted
609 non-responders to responders is higher than 50%. The top significantly positive disrupted modules were
610 defined as those with a complete agreement of all three parameters in the highest percentile with shared
611 selected pathways across all parameters, which in our case was determined as the 0.8 percentile.

612 **Single cell RNA sequencing**

613 **Peripheral blood mononuclear cells (PBMCs) cryopreservation and thawing**

614 Blood samples were drawn before IFX first infusion. PBMCs were isolated using density gradient
615 centrifugation by spinning blood over UNI-SEPmaxi+ tubes (Novamed Ltd.) following the manufacturer's
616 protocol. Isolated cells were resuspended in 1 ml freezing solution, containing 10% DMSO and 90% FCS.
617 The samples were kept in Nalgene Mr. Frost® Cryo 1°C Freezing Container (ThermoFisher scientific) with
618 Isopropyl alcohol at -80°C over-night, and immediately after placed in a liquid nitrogen container for long-
619 term storage.
620 For thawing, frozen PBMCs were immediately transferred to a water bath at 37°C for 2-3 min, until a
621 tiny ice crystal was remained. Thawed cells were transferred into 50 mL centrifuge tubes and rinsed with
622 1 mL of warm (37 °C) RPMI 1640 supplemented with 10% of FCS which was added dropwise to the
623 DMSO containing fraction while gently shaking the cells. Next, the cells were sequentially diluted by first
624 adding 2 mL of medium followed by another 4, 8 and 16 mL respectively with 1 min wait between the
625 four dilution steps. The diluted cell suspension was centrifuged for 5 min at 300 g. Most of the
626 supernatant was discarded leaving ~1 ml, and the cells were resuspended in 9 ml of medium followed by
627 additional centrifugation for 5 min at 300 g and resuspended with the same media to reach the desired
628 cell concentration.

629 **Single cell RNA sequencing in 10X genomics platform**

630 PBMCs from responder and non-responder patients pre-treatment (N=2) were prepared for scRNA-seq
631 according to the 10x Genomics Single Cell protocols for fresh frozen human peripheral blood mononuclear
632 cells (see above for cell preservation and thawing). The cells were adjusted to a final cell concentration of
633 1000 cells/Ul and placed on ice until loading into the 10x Genomics Chromium system. The scRNA
634 sequencing was performed in the genomic center of the biomedical core facility in the Rappaport faculty
635 of medicine at the Technion - Israel Institute of Technology. Libraries were prepared using 10x Genomics
636 Library Kits (Chromium Next GEM Single Cell 3' Library & Gel Bead Kit v3.1, PN-1000121) using 20,000
637 input cells per sample. Single cell separation was performed using the Chromium Next GEM Chip G Single
638 Cell Kit (PN-1000120). The RNAseq data was generated on Illumina NextSeq500, high-output mode
639 (Illumina, FC-404-2005), 75 bp paired-end reads (Read1- 28 bp, Read2- 56 bp, Index- 8 bp).

640 **Single cell data analysis**

641 Cell Ranger single cell software suite was used for sample de-multiplexing, alignment to human reference
642 genome (GRCh38-3.0.0), cell barcode processing and single cell UMI counting following default settings.
643 The UMI count matrix was further processed using the Seurat R package (version 3.1.4). First, as a QC

644 step, cells that had a unique feature count of less than 200 were filtered out. Additional filtering was
645 applied to remove features detected in less than 3 cells. we further filtered cells based on mitochondrial
646 gene content above 0.25%. After this step, 19275 single cells and 20673 genes in total were retained and
647 included in downstream analyses. This was followed by Global-scaling library size normalization. Genes
648 were scaled in comparison to all other cells and regressed out the effects of unwanted sources of variation
649 including UMI counts and percentage of mitochondrial genes for the remaining cells. At the next step, we
650 performed linear dimensionality reduction on the scaled data of the top 2000 highly variable genes.
651 Resampling test based on the jackstraw procedure and Elbow plot were performed to identify the first 30
652 significance principal components that were used for downstream visualization by t-SNE plot.
653 SingleR was used to annotate cell types based on correlation profiles with two different resolutions of
654 cell classification using the Blueprint-Encode[48] and the Monaco Immune Cell[49] reference datasets of pure
655 cell types. Differential expression analysis between responders and non-responders was performed for
656 each cell population using a Wilcoxon Rank Sum test implemented in the FindAllMarkers function in the
657 Seurat package.
658 Relative pathway score based on the expended fiber-organization baseline differential genes was
659 calculated for each single cell and compared between cell subsets and response groups using Wilcoxon
660 test (for the expended fiber organization differential genes assessment see below description for selection
661 and evaluation of predictive model for IFX treatment response; see the above description for relative
662 pathway score calculation).
663 To identify cell specific enriched pathways that are associated with the predictive fiber-organization
664 related signature, we constructed a co-expression network based on the pre-treatment expression of the
665 predictive genes: RAC1, PAK1, ICAM1, LYN, FCGR3A and IL-1β, in intermediate monocyte subset in each
666 response group using the MTGOsc R package (Spearman's correlation, thinning net by 0.1 top percentile).
667 Functional enrichment analysis was performed based on the co-expressed network nodes, by a
668 hypergeometric test based on the Reactome database using the Clusterprofiler R package (P-adjust<0.05).
669 Wilcoxon test was assessed to identify significant differences in pathway scores between response groups
670 for each enriched pathway in each monocyte subset. P-values were further adjusted for multiple testing
671 using the Benjamini-Hochberg procedure.

672 **Predictive model for IFX treatment response**

673 Given the significant linkage between monocytes and the differential fiber organization pathway, in order
674 to build a cell specific pre-treatment classifier, we expanded the fiber organization adjusted-bulk based
675 differential genes through intersection of knowledge based- (combined score>900,
676 9606.protein.links.detailed.v11.0 from the STRING protein interaction database: http://string-db.org/ and
677 data-driven networks (Monocytes single-cell based co-expression from a representative responder and
678 non-responder patients at baseline , Spearman's r, thinning percentile: 0.05, MTGOsc R package). This
679 yielded a combined network of 42 edges containing 23 nodes. To build a predictive signature, we used
680 elastic net regularized logistic regression for predictors selection, which has the advantage of including all
681 correlated predictors sharing transcriptional signal (grouping effect), rather than selecting one variable
682 from a group of correlated predictors while ignoring the others[50]. We used the glmnet R package
683 implemented within the caret R package for model fitting by tuning over both alpha (ranging from 0.5-1,
684 n=6) and lambda (ranging from 0.0001-1, n=20) parameters with 100 repeated 2-fold cross-validation.
685 The optimized model was chosen based on the best performance value using the Receiver operating
686 characteristic (ROC) metric (alpha=0.5, lambda=0.26).

687 After variable selection, we calculated AUC based on relative pathway score combining the selected genes
688 using the pROC R package.
689 Internal validation was performed by bootstrapping (n=1000 bootstrap samples) for the AUC by randomly
690 drawing subjects with the same sample size from the original cohort (with replacement).
691 A permutation test was used for estimating one-tailed P-value (n=10000 permutations) by shuffling the
692 subject labels between the response groups and the expression of the selected signature genes. Then we
693 tested the null hypothesis that the observed AUC was drawn from this null distribution.

694 **External validation of the predictive signature using additional independent real-life IBD cohort**

695 For independent validation of the predictive signature, we used an independent IBD cohort of 29 patients
696 (see Patient in the validation real life cohort). RNA was then extracted using RNeasy mini kit (QIAGEN)
697 according to the manufacturer's instruction (for preservation and thawing of PBMCs see Peripheral blood
698 mononuclear cells (PBMCs) cryopreservation). Complementary DNA was synthesized using Maxima first
699 strand cDNA synthesis kit with dsDNase (Thermo Scientific). qPCR was performed using 7300 Real-Time
700 PCR System (AB Applied Biosystems). Relative cytokine expression was calculated following normalization
701 to glyceraldehyde-3 phosphate dehydrogenase (GAPDH) expression (Supp. table 10 for the PCR primer
702 sets). Primers were purchased from Sigma Aldrich. The expression of the genes in the predictive signature
703 was calculated relative to CD14 expression, to measure monocytes' centered differential expression
704 between response groups pre-treatment. Relative pathway score was used to assess prediction
705 performance (see Relative pathway score evaluation).

706 **Assessment of the predictive signature performance in RA**

707 The prediction performance of the RAC1-PAK1 signature in RA public expression datasets was evaluated
708 using the following datasets: GSE20690 (n=68 of which 43 and 25 are responders and non-responders
709 respectively), GSE33377 (n=42 of which 18 and 24 are responders and non-responders respectively) and
710 GSE42296 (n=19 of which 13 and 6 are responders and non-responders respectively).
711 Gene expression was adjusted to major cell type contributions (see Blood transcriptome analysis), which
712 were evaluated by deconvolution using a linear regression framework in which individual samples were
713 regressed based on a characteristic expression of marker genes expressed in 17 cell-types (CellMix R
714 package). This was followed by performance prediction calculation for each study based on the relative
715 signature score based on the adjusted gene expression. Due to differences in expression platforms
716 between studies, there were genes in the signature which were not present in a specific dataset, therefore
717 those genes were not used in the calculation of the relative signature score for the prediction of the
718 specific study. To combine prediction performance from these independent studies we constructed a
719 summary ROC curve (meta-ROC) using the nsROC R package which performs a simple linear interpolation
720 between pairs of points of each individual ROC.
721

Acknowledgments

This work was supported by funding of the Helmsley Charitable Trust to Y.C and S.S.S-O. We thank T.Shvedov for contribution to patient enrollment and clinical data collection. S.Pollok, L.Pinzur, N.Molshatzki and Y.Benita for fruitful discussions and advice on the computational methodology. V.Barsan for his insightful comments in reviewing the manuscript.

Author contributions

S.S.S.-O, Y.C. conceived the idea; S.S.S-O, Y.C, S.G.V, E.S and R.G designed the analyses, S.S.S-O, Y.C. and S.G.V performed the interpretation; S.G.V and R.G performed the design and development of the computational pipeline and validation; A.K, B.P, Y.G and A.A performed development of the computational methodology; N.Ma, A.B, S.P and E.S counseled regarding the biological interpretation; E.S performed the experimental design of the collected cohort and E.S, N.Ma, A.A and T.D performed the data generation; A.B and N.Mi performed the experimental validation; A.B, S.P performed the sample collection; Y.C, H.B.Y and Y.G performed patient enrollment and clinical characterization; S.S.S-O, Y.C and S.G.V wrote the manuscript.

Competing interests

These authors disclose the following: Y.C received consulting fees from AbbVie, Janssen, Takeda, Pfizer and CytoReason; speaker fees from AbbVie, Janssen, and Takeda; and grants from AbbVie, Takeda and Janssen. S.S.S-O received grant fees from Takeda, S.S.S.-O, E.S. and R.G declares CytoReason equity and advisory fees. N. Ma and A.K are employees at CytoReason. S.G.V declares CytoReason advisory fees. The remaining authors disclose no conflicts.

743 ## References

744

745 1. Jiang, P. *et al.* Signatures of T cell dysfunction and exclusion predict cancer immunotherapy
746 response. *Nat. Med.* (2018). doi:10.1038/s41591-018-0136-1
747 2. Murthy, S. K. *et al.* Introduction of anti-TNF therapy has not yielded expected declines in
748 hospitalisation and intestinal resection rates in inflammatory bowel diseases: A population-based
749 interrupted time series study. *Gut* (2020). doi:10.1136/gutjnl-2019-318440
750 3. Berns, M. & Hommes, D. W. Anti-TNF-α therapies for the treatment of Crohn's disease: the past,
751 present and future. *Expert Opin. Investig. Drugs* (2015). doi:10.1517/13543784.2016.1126247
752 4. Mitoma, H., Horiuchi, T., Tsukamoto, H. & Ueda, N. Molecular mechanisms of action of anti-TNF-
753 α agents – Comparison among therapeutic TNF-α antagonists. *Cytokine* (2018).
754 doi:10.1016/j.cyto.2016.08.014
755 5. Roda, G., Jharap, B., Neeraj, N. & Colombel, J. F. Loss of Response to Anti-TNFs: Definition,
756 Epidemiology, and Management. *Clinical and Translational Gastroenterology* (2016).
757 doi:10.1038/ctg.2015.63
758 6. Seymour, H. E., Worsley, A., Smith, J. M. & Thomas, S. H. L. Anti-TNF agents for rheumatoid
759 arthritis. *Br. J. Clin. Pharmacol.* (2001). doi:10.1046/j.1365-2125.2001.00321.x
760 7. Digby-Bell, J. L., Atreya, R., Monteleone, G. & Powell, N. Interrogating host immunity to predict
761 treatment response in inflammatory bowel disease. *Nature Reviews Gastroenterology and*
762 *Hepatology* (2020). doi:10.1038/s41575-019-0228-5
763 8. West, N. R. *et al.* Oncostatin M drives intestinal inflammation and predicts response to tumor
764 necrosis factor–neutralizing therapy in patients with inflammatory bowel disease. *Nat. Med.* **23,**
765 579–589 (2017).
766 9. Gaujoux, R. *et al.* Cell-centred meta-Analysis reveals baseline predictors of anti-TNFα non-
767 response in biopsy and blood of patients with IBD. *Gut* **68,** 604–614 (2019).
768 10. Mulhearn, B., Barton, A. & Viatte, S. Using the immunophenotype to predict response to biologic
769 drugs in rheumatoid arthritis. *Journal of Personalized Medicine* (2019). doi:10.3390/jpm9040046
770 11. Kuijjer, M. L., Tung, M. G., Yuan, G. C., Quackenbush, J. & Glass, K. Estimating Sample-Specific
771 Regulatory Networks. *iScience* (2019). doi:10.1016/j.isci.2019.03.021
772 12. Liu, X., Wang, Y., Ji, H., Aihara, K. & Chen, L. Personalized characterization of diseases using
773 sample-specific networks. *Nucleic Acids Res.* (2016). doi:10.1093/nar/gkw772
774 13. Liu, Y. *et al.* Cytokine-mediated regulation of activating and inhibitory Fcγ receptors in human
775 monocytes. *J. Leukoc. Biol.* (2005). doi:10.1189/jlb.0904532
776 14. Turner, M. & Billadeau, D. D. VAV proteins as signal integrators for multi-subunit immune-
777 recognition receptors. *Nature Reviews Immunology* (2002). doi:10.1038/nri840
778 15. Zhao, X., Carnevale, K. A. & Cathcart, M. K. Human Monocytes Use Rac1, Not Rac2, in the NADPH
779 Oxidase Complex. *J. Biol. Chem.* (2003). doi:10.1074/jbc.M302208200
780 16. Parikh, K. *et al.* Suppression of p21Rac signaling and increased innate immunity mediate
781 remission in Crohn's disease. *Sci. Transl. Med.* (2014). doi:10.1126/scitranslmed.3006763
782 17. Kapellos, T. S. *et al.* Human monocyte subsets and phenotypes in major chronic inflammatory
783 diseases. *Frontiers in Immunology* (2019). doi:10.3389/fimmu.2019.02035
784 18. Tanino, M. *et al.* Prediction of efficacy of anti-TNF biologic agent, infliximab, for rheumatoid
785 arthritis patients using a comprehensive transcriptome analysis of white blood cells. *Biochem.*
786 *Biophys. Res. Commun.* (2009). doi:10.1016/j.bbrc.2009.06.149
787 19. Toonen, E. J. M. *et al.* Validation study of existing gene expression signatures for anti-TNF
788 treatment in patients with rheumatoid arthritis. *PLoS One* (2012).
789 doi:10.1371/journal.pone.0033199
790 20. Mesko, B. *et al.* Peripheral blood derived gene panels predict response to infliximab in
791 rheumatoid arthritis and Crohn's disease. *Genome Med.* (2013). doi:10.1186/gm463
792 21. Kalliolias, G. D. & Ivashkiv, L. B. TNF biology, pathogenic mechanisms and emerging therapeutic
793 strategies. *Nat. Rev. Rheumatol.* (2016). doi:10.1038/nrrheum.2015.169

794 22. Lügering, A. *et al.* Infliximab induces apoptosis in monocytes from patients with chronic active
795 Crohn's disease by using a caspase-dependent pathway. *Gastroenterology* (2001).
796 doi:10.1053/gast.2001.28702
797 23. Batko, B., Schramm-Luc, A., Skiba, D. S., Mikolajczyk, T. P. & Siedlinski, M. TNF-α inhibitors
798 decrease classical CD14 hi cd16− monocyte subsets in highly active, conventional treatment
799 refractory rheumatoid arthritis and ankylosing spondylitis. *Int. J. Mol. Sci.* (2019).
800 doi:10.3390/ijms20020291
801 24. Vos, A. C. W. *et al.* AntiTumor necrosis factor-α antibodies induce regulatory macrophages in an
802 Fc region-dependent manner. *Gastroenterology* (2011). doi:10.1053/j.gastro.2010.10.008
803 25. Yokoyama, Y. *et al.* Efficacy of granulocyte and monocyte adsorptive apheresis in patients with
804 inflammatory bowel disease showing lost response to infliximab. *J. Crohn's Colitis* (2020).
805 doi:10.1093/ecco-jcc/jjaa051
806 26. Marei, H. & Malliri, A. Rac1 in human diseases: The therapeutic potential of targeting Rac1
807 signaling regulatory mechanisms. *Small GTPases* (2017). doi:10.1080/21541248.2016.1211398
808 27. Muise, A. M. *et al.* Single nucleotide polymorphisms that increase expression of the guanosine
809 triphosphatase RAC1 are associated with ulcerative colitis. *Gastroenterology* (2011).
810 doi:10.1053/j.gastro.2011.04.057
811 28. Tiede, I. *et al.* CD28-dependent Rac1 activation is the molecular target of azathioprine in primary
812 human CD4+ T lymphocytes. *J. Clin. Invest.* (2003). doi:10.1172/JCI16432
813 29. Lim, S. Z. & Chua, E. W. Revisiting the role of thiopurines in inflammatory bowel disease through
814 pharmacogenomics and use of novel methods for therapeutic drug monitoring. *Frontiers in*
815 *Pharmacology* (2018). doi:10.3389/fphar.2018.01107
816 30. Hamerman, J. A., Ni, M., Killebrew, J. R., Chu, C. L. & Lowell, C. A. The expanding roles of ITAM
817 adapters FcRγ and DAP12 in myeloid cells. *Immunological Reviews* (2009). doi:10.1111/j.1600-
818 065X.2009.00841.x
819 31. Louis, E. *et al.* Association between polymorphism in IgG Fc receptor IIIa coding gene and
820 biological response to infliximab in Crohn's disease. *Aliment. Pharmacol. Ther.* (2004).
821 doi:10.1111/j.1365-2036.2004.01871.x
822 32. Moroi, R. *et al.* FCGR3A-158 polymorphism influences the biological response to infliximab in
823 Crohn's disease through affecting the ADCC activity. *Immunogenetics* (2013).
824 doi:10.1007/s00251-013-0679-8
825 33. McRae, B. L. *et al.* Fc receptor-mediated effector function contributes to the therapeutic
826 response of anti-TNF monoclonal antibodies in a mouse model of inflammatory bowel disease. *J.*
827 *Crohn's Colitis* (2016). doi:10.1093/ecco-jcc/jjv179
828 34. Chen, Y., Li, H., Lai, L., Feng, Q. & Shen, J. Identification of Common Differentially Expressed
829 Genes and Potential Therapeutic Targets in Ulcerative Colitis and Rheumatoid Arthritis. *Front.*
830 *Genet.* (2020). doi:10.3389/fgene.2020.572194
831 35. Chen, J. Y. *et al.* Association of FCGR3A and FCGR3B copy number variations with systemic lupus
832 erythematosus and rheumatoid arthritis in Taiwanese patients. *Arthritis Rheumatol.* (2014).
833 doi:10.1002/art.38813
834 36. Roberts, J. T., Patel, K. R. & Barb, A. W. Site-specific N-glycan Analysis of Antibody-binding Fc γ
835 Receptors from Primary Human Monocytes. *Mol. Cell. Proteomics* (2020).
836 doi:10.1074/mcp.RA119.001733
837 37. Herter, S. *et al.* Glycoengineering of Therapeutic Antibodies Enhances Monocyte/Macrophage-
838 Mediated Phagocytosis and Cytotoxicity. *J. Immunol.* (2014). doi:10.4049/jimmunol.1301249
839 38. Kotliarov, Y. *et al.* Broad immune activation underlies shared set point signatures for vaccine
840 responsiveness in healthy individuals and disease activity in patients with lupus. *Nat. Med.*
841 (2020). doi:10.1038/s41591-020-0769-8
842 39. Goh, K. Il *et al.* The human disease network. *Proc. Natl. Acad. Sci. U. S. A.* (2007).
843 doi:10.1073/pnas.0701361104
844 40. Singh, S., George, J., Boland, B. S., Casteele, N. Vande & Sandborn, W. J. Primary non-response to

845 tumor necrosis factor antagonists is associated with inferior response to second-line biologics in
846 patients with Inflammatory bowel diseases: A systematic review and meta-analysis. *J. Crohn's*
847 *Colitis* (2018). doi:10.1093/ecco-jcc/jjy004
848 41. Rosen, M. J., Minar, P. & Vinks, A. A. Review article: Applying pharmacokinetics to optimise
849 dosing of anti-TNF biologics in acute severe ulcerative colitis. *Alimentary Pharmacology and*
850 *Therapeutics* (2015). doi:10.1111/apt.13175
851 42. Sarikaya, I. *et al.* Tc-99m dextran and Tc-99m HIG findings in patients with ulcerative colitis. *Clin.*
852 *Nucl. Med.* (1999). doi:10.1097/00003072-199904000-00005
853 43. Finck, R. *et al.* Normalization of mass cytometry data with bead standards. *Cytom. Part A* (2013).
854 doi:10.1002/cyto.a.22271
855 44. Bruggner, R. V., Bodenmiller, B., Dill, D. L., Tibshirani, R. J. & Nolan, G. P. Automated
856 identification of stratifying signatures in cellular subpopulations. *Proc. Natl. Acad. Sci. U. S. A.*
857 (2014). doi:10.1073/pnas.1408792111
858 45. N., P. *et al.* Usefulness of transcriptional blood biomarkers as a non-invasive surrogate marker of
859 mucosal healing and endoscopic response in ulcerative colitis. *J. Crohn's Colitis* (2017).
860 doi:10.1093/ecco-jcc/jjx091 LK -
861 http://sfx.cineca.it:9003/unito?sid=EMBASE&sid=EMBASE&issn=18764479&id=doi:10.1093%2Fe
862 cco-jcc%2Fjjx091&atitle=Usefulness+of+transcriptional+blood+biomarkers+as+a+non-
863 invasive+surrogate+marker+of+mucosal+healing+and+endoscopic+response+in+ulcerative+coliti
864 s&stitle=J.+Crohn%27s+Colitis&title=Journal+of+Crohn%27s+and+Colitis&volume=11&issue=11&
865 spage=1335&epage=1346&aulast=Planell&aufirst=N%C3%BAria&auinit=N.&aufull=Planell+N.&co
866 den=&isbn=&pages=1335-1346&date=2017&auin
867 46. Li, S. *et al.* Molecular signatures of antibody responses derived from a systems biology study of
868 five human vaccines. *Nat. Immunol.* **15,** 195–204 (2014).
869 47. Tsang, J. S. *et al.* Global analyses of human immune variation reveal baseline predictors of
870 postvaccination responses. *Cell* (2014). doi:10.1016/j.cell.2014.03.031
871 48. Fernández, J. M. *et al.* The BLUEPRINT Data Analysis Portal. *Cell Syst.* (2016).
872 doi:10.1016/j.cels.2016.10.021
873 49. Xu, W. *et al.* Mapping of γ/δ T cells reveals Vδ2+ T cells resistance to senescence. *EBioMedicine*
874 (2019). doi:10.1016/j.ebiom.2018.11.053
875 50. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B*
876 *Stat. Methodol.* (2005). doi:10.1111/j.1467-9868.2005.00503.x
877

878 ## Figure titles and legends

879 **Fig 1| External data-driven disease specific molecular response metric, termed 'inflammatory axis',**
880 **indicated that responders exhibit a trajectory of treatment-induced immune dynamics while non-**
881 **responders exhibit an overall opposite direction. a**, Overview of the 'inflammatory axis' analysis. **b**,
882 'Inflammatory axis' assessment. Left panel, external public (GSE94648) based 'inflammatory axis' which
883 defines a transition from IBD active disease through inactive disease to healthy state by PCA based
884 differential expressed genes between disease/health states. Right panel, the projection distance of
885 responding and non-responding patients' samples from our real-life cohort on the 'inflammatory axis' at
886 W2 compared to baseline. **c**, Boxplots comparing responders' and non-responders' projection dynamics
887 on the 'inflammatory axis' at each treatment interval (One-tailed permutation P-values shown, n=10000).
888 **d**, Scatterplot of the relationship between progress on the 'inflammatory axis' between W2 to baseline
889 and between W2 to W14 (n=23, Spearman's r=-0.44, P<0.1).

890 **Fig 2| Normal infliximab dynamics correlated with changes in monocytes and reduced expression of**
891 **innate immune related pathways. a**, Cell frequency alterations following IFX treatment. Left panel, PCA
892 presenting immune cell frequency changes following treatment based on 16 canonical immune
893 populations determined by CyTOF. Arrow tail and head indicate the early W2 and later W14 relative to
894 baseline compositional changes correspondingly. Ellipses represent the Euclidean distance from the
895 center. Center panel, boxplots showing change in monocytes abundance following treatment relative to
896 baseline in responders and non-responders (paired-Wilcoxon P-values shown). Right panel, scatterplot
897 showing the relationship between changes in monocytes abundance (log transformed fold change relative
898 to baseline) and changes in CRP (fold change relative to baseline) (n=23, Spearman correlation=0.4,
899 P=0.01). **b**, Venn diagram showing dynamic features which significantly changed over time at 2 weeks and
900 14 weeks post treatment compared with baseline for each response group using linear mixed-effects
901 models (FDR<0.15, n=1000 & n=519 permutations for responders and non-responders respectively). **c**,
902 Scatterplot presenting the normal response network centrality of significantly enriched dynamic pathways
903 at the early response period (GSEA, FDR<0.25, n perm=1000). Colors indicate pathway median fold change
904 expression at the early response period relative to baseline in responders (colored dots denote significant
905 change in relative pathway score by Wilcoxon test, FDR<0.05).

906 **Fig 3| 'Disruption Networks' as a framework to perform sample level inferences to identify individual**
907 **variation in drug response**. **a**, 'Disruption Networks' concept and applications. Bulk gene expression
908 constitutes both effects of cell composition and cell-specific regulatory programs. 'Disruption Networks'
909 initially decouples cell composition and cell-specific regulatory programs from bulk gene expression
910 providing a cell-centered regulatory network of genes and cells. Then, 'Disruption Networks' learns
911 individual-level breaking or rewiring of cross-feature relations, and by that forms a new data-type
912 providing complementary biological information which increase signal detection. The new data-type can
913 be used for diverse downstream analyses including data integration that accounts for both dimensions of
914 feature expression and relation levels, disruption assessment in functional modules, stratification of
915 patients by disruption profile, assessment of perturbation effects by measuring disruption level
916 throughout the network. **b**, Feature specific differential signal between responders and non-responders
917 dynamics at the early response period using disruption measurement of top mean drop intensity (x axis)
918 and standard statistics by Wilcoxon test (y axis). **c**, 'Disruption Networks' statistic was aggregated across
919 pathways to estimate sample specific disruption in the functional level, according to mean drop intensity,
920 a representative disruption parameter out of three different defined parameters. The heatmap represents
921 the disrupted dynamics for each pathway and sample at W2 compared to baseline. Top significantly
922 disrupted pathways are presented, defined as those with a complete agreement of all three parameters
923 in the 0.8 percentile. Line graphs describe the percentage of disrupted patients in each response group.
924 **d**, Distribution of degree and betweenness centrality for nodes belonging to the top disrupted pathways
925 compared to other nodes in the network. Significance was determined using permutation test (n

926 perm=10000). **e**, Meta disrupted pathway. Left panel, response network subgraph consist of nodes from
927 the baseline differential disrupted pathways (FDR<0.1). Diamond shape and orange color represent cell
928 frequency; circle shape represent cell centered expression; Red circles indicate the fiber organization
929 pathway related central axis. Right panel, enrichment analysis of the disrupted pathways by
930 hypergeometric test.

931 **Fig. 4| Fiber-organization signaling, highly expressed in monocytes, predicts infliximab response at**
932 **baseline. a**, Baseline expression differences in the disrupted pathways between response groups
933 (NPMANOVA; bottom primary axis). Colors denote response network betweenness. The line graph
934 represent correlation of changes in pathway score with changes in CRP (top secondary axis). **b**, The fiber
935 organization differential nodes dynamics assessed by mean relative score across visits for each response
936 group (Wilcoxon one-tailed P-values shown). **c**, Analysis of the cellular origin of the baseline differential
937 fiber organization pathway using scRNA-seq analysis of PBMCs collected from representative responder
938 and non-responder pre-treatment. Left panel, tSNE plot representing cell types identities annotated using
939 singleR based on correlation profiles based on two reference datasets: the Blueprint-Encode and the
940 Monaco Immune Cell datasets. Right panel, tSNE plot colored by the expended fiber organization scaled
941 expression. The fiber organization baseline differential genes were expended through intersecting
942 knowledge based (stringDB) and data-driven based (Monocyte single cell data) networks. **d**, The expended
943 fiber organization scaled expression in the different monocyte subsets (Wilcoxon P-values shown). **e**,
944 Mean mTNF expression in the different monocyte subsets as measured by CyTOF (Wilcoxon one-tailed P-
945 values shown).

946 **Fig. 5| Validation of the fiber organization predictive signature in an independent IBD cohort and three**
947 **public RA cohorts pre IFX treatment. a**, Validation of the pre-treatment predictive fiber organization
948 signature in an additional independent cohort of 20 and 9 responders and non-responders respectively
949 by qPCR. Gene values were normalized to CD14 expression for cell-centered values. Left panel, bar graph
950 of the pre-treatment normalized expression of the signature genes and signature pathway score in each
951 response group (Wilcoxon one-tailed P-values shown). Right panel, ROC based on the predictive signature
952 relative score. **b**, Prediction performance of fiber organization signaling signature in RA public expression
953 datasets. Left panel, boxplots comparing the fiber organization signature related genes and the pathway
954 score between IFX RA responders (n=43) and non-responders (n=25) in a representative public dataset
955 GSE20690 (Wilcoxon one-tailed P-values shown). Right panel, ROC based on the predictive signature
956 relative score of the relevant cohort. **c**, Meta-ROC presenting the predictive performance of three
957 independent public RA cohorts.
958

959 ## Supplemental Information titles and legends

960 **Supp. Fig 1| CyTOF reveals multiple cell subset changes in responders following treatment and**
961 **differences between response groups. a**, Loading plot of PC2 based on major canonical cell composition
962 changes at W2 and W14 compared to baseline. **b**, Cell-type specific alteration in cellular relative
963 abundance during IFX treatment in responders and non-responders (paired-Wilcoxon P-values shown). **c**,
964 Correlation of cell abundance changes at W2 and W14 relative to baseline, with changes in CRP
965 (Spearman's correlation coefficients are shown, P-values are calculated by two tailed probability of the t-
966 statistic, P<0.05 for significant p-values).

967 **Supp. Fig 2| The cumulative number of discovered dynamic features, at a range of target FDR values by**
968 **data-type for each response group.** Top and bottom panels represent significant changes at W2 and W14
969 relative to baseline respectively. FDR was calculated using the Benjamini-Hochberg procedure.
970 Responders were subsampled (n=200) to match the non-responder group size. For responders, mean±
971 SEM values are shown.

972 **Supp. Fig 3| Functional pathways associated with IFX response. a,** Scatterplot of p-values obtained by a
973 comparison of pathway scores between W2 and baseline against those obtained by comparing W14 to
974 baseline (-log10 of paired-Wilcoxon P-values shown). Only globally enriched and network connected
975 pathways were included. **b**, Pathway score related dynamics between W2 and W14 relative to baseline.
976 Top 70 pathways are shown. Pathways are ordered by fold change effect size. P-values for pathway score
977 differences between time points were calculated by paired-Wilcoxon test. Significance was determined
978 by FDR<0.05 (Benjamini-Hochberg procedure). **c**, Heatmap representing a cell-specific contribution for
979 the change in the dynamic pathways. The contribution was determined for each gene in the pathway by
980 regressing its unadjusted fold change expression over the major peripheral cell type frequencies. The
981 reported values represent the mean of the coefficients across all genes in the pathway. **d**, Correlation of
982 pathway score expression with CRP. All time point and response groups are included. (Spearman's
983 correlation coefficients are shown, P-values are calculated by two tailed probability of the t-statistic,
984 Pathway which significantly correlated with CRP (FDR<0.05, Benjamini-Hochberg procedure) are colored.

985 **Supp. Fig 4| Additional disruption parameters and comparison of the differential signal between**
986 **response groups dynamics as obtained by the 'Disruption Networks' framework and standard statistics.**
987 **a,** Representative highly disrupted edge demonstrating significant dropout values for non-responders. **b,**
988 Feature-specific differential signal between responders and non-responders' dynamics at W2 relative to
989 baseline, based on the disruption parameters and standard statistics. Left panel, top disrupted edge ratio
990 (x axis, FDR<0.1 for dropout significance and 10th top percentile of disrupted edge ratio) and standard
991 statistics by Wilcoxon test (y axis, FDR<0.1); Right panel, Scatterplot showing feature specific disruption
992 parameters of mean drop intensity against disrupted edge ratio. Points are colored by quartile thresholds
993 (FDR<0.1 for dropout significance and 10th top percentile of the specific disruption parameter). The
994 feature which agreed with the disruption parameters and standard Wilcoxon test is marked with black
995 border. **c,** Aggregation of 'Disruption Networks' statistic across pathways to estimate sample specific
996 disruption in the functional level, according to percentage of disrupted edges and percentage of disrupted
997 nodes. Heatmaps represent the disrupted dynamics in each parameter for each pathway and sample at
998 W2 compared to baseline. Top significantly disrupted pathways are presented, defined as those with a
999 complete agreement of all three parameters in the 0.8 percentile. Line graphs describe the percentage of
1000 disrupted patients in each response group.

1001 **Supp. Fig 5| Baseline differences of the significantly dynamics disrupted pathways**. **a**, Heatmap
1002 representing the feature-level baseline differences among genes in the dynamics meta-disrupted pathway
1003 (FDR<0.1, Wilcoxon test). **b**, Correlation between the canonical cellular frequencies as obtained by CyTOF,
1004 and the bulk unadjusted expression of the fiber organization related genes in responders (Spearman's
1005 correlation coefficients are shown, P-values are calculated by two tailed probability of the t-statistic). Only

significant correlation values are shown (P<0.05 and |r|≥0.5). **c,** Baseline prediction of IFX response in the primary IFX cohort based on the expended fiber organization predictive signature score, in the cell adjusted space. Left panel, receiver operating characteristic (ROC) plots of 200-bootsraps. The predictive signature was determined using elastic net (a=0.5, lambda=0.26, 100 repeated 2-fold CV) based on the adjusted baseline differential fiber organization related genes. Significance was determined by permutation test (n perm=10000). Right panel, boxplots of the fiber organization predictive signature score pre-treatment, in the different response groups in the cell-centered bulk expression

Supp. Fig 6| scRNA-seq based comparison of the baseline fiber organization related expression between the main cell-types and response groups. The fiber organization scaled score based on its baseline differential genes was compared between PBMCs major cell types, and between response groups for monocytes (Wilcoxon P-values shown).

Supp. Fig 7| Intermediate monocytes functional pathways associated with the predictive fiber organization signature. Heatmap representing the top 20 intermediate-monocytes specific enriched pathways associated with the predictive fiber-organization related signature is shown. Pathways were determined by co-expression network based on the pre-treatment expression of the signature predictive genes in intermediate monocyte based on the scRNA-seq data in each response group followed by enrichment analysis (Spearman's correlation, thinning net by 0.1 top percentile, P-adjust<0.05 for functional enrichment significance by hypergeometric test). Pathways displaying significant differences between response groups in each cell subset are colored (FDR<0.05 by Wilcoxon test).

Supplementary Table titles

ST1: Clinical and demographic characteristics of patients included in the primary real life CD cohort
ST2: CyTOF Panel
ST3: Cell type unsupervised clustering using Citrus algorithm
ST4: Luminex Panel. List of analytes tested in the Luminex assay
ST5: Differentially expressed features between CD and UC active patients, and healthy controls for the construction of an external reference 'inflammatory axis'
ST6: Selection of highly informative PCs to best describe an inflammatory axis directionality from active, through inactive disease states to healthy state using ordinal lasso
ST7: Dynamic features at W2 and W14 relative to baseline in responders and non-responders using linear mixed-effects models
ST8: Normal anti-TNF response dynamics network at the early W2 response period
ST9: Clinical and demographic characteristics of patients included in the validation real-life CD cohort
ST10: qPCR primers used in the IBD validation cohort for measuring expression of the fiber organization predictive signature

Fig. 1

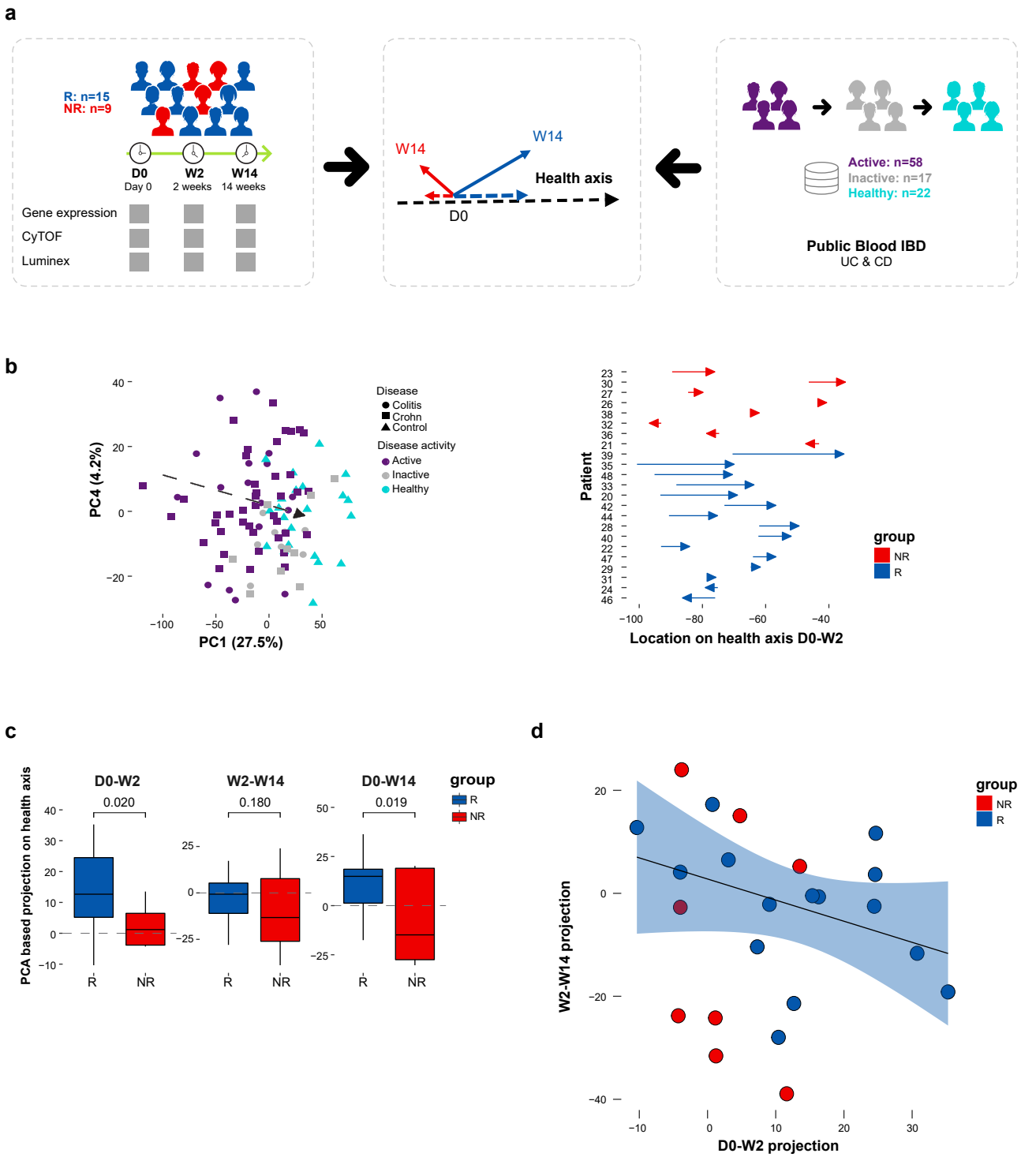

a
R: n=15
NR: n=9
D0
Day 0
W2
2 weeks
W14
14 weeks
Gene expression
CyTOF
Luminex
W14
W14
Health axis
D0
Active: n=58
Inactive: n=17
Healthy: n=22
Public Blood IBD
UC & CD
b
PC4 (4.2%)
PC1 (27.5%)
Disease
Colitis
Crohn
Control
Disease activity
Active
Inactive
Healthy
Patient
Location on health axis D0-W2
group
NR
R
c
PCA based projection on health axis
D0-W2
0.020
W2-W14
0.180
D0-W14
0.019
group
R
NR
d
W2-W14 projection
D0-W2 projection
group
NR
R

Fig. 2

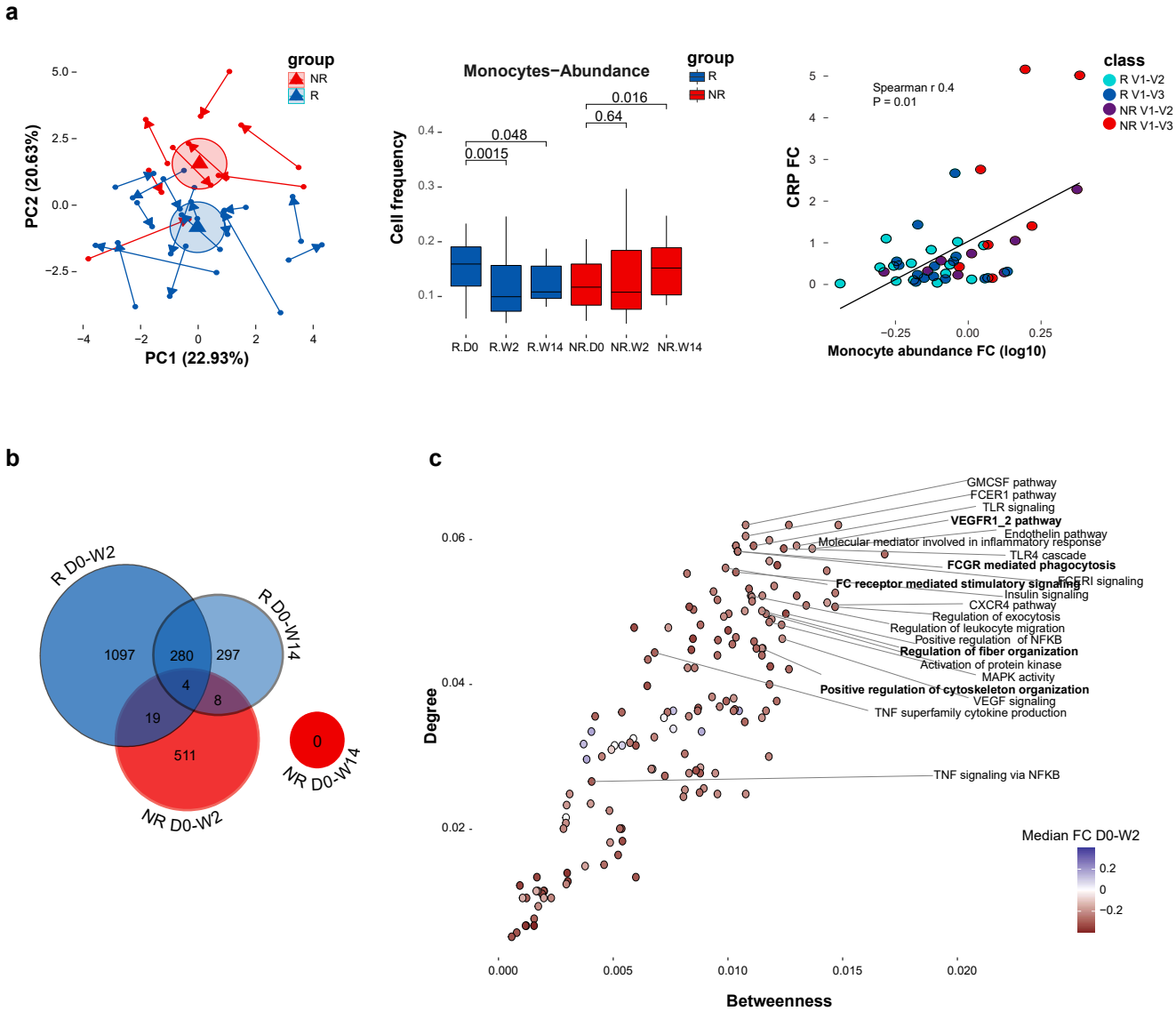

a
group
NR
R
PC2 (20.63%)
PC1 (22.93%)
Monocytes−Abundance
group
R
NR
0.048
0.0015
0.016
0.64
Cell frequency
R.D0
R.W2
R.W14
NR.D0
NR.W2
NR.W14
class
R V1-V2
R V1-V3
NR V1-V2
NR V1-V3
Spearman r 0.4
P = 0.01
CRP FC
Monocyte abundance FC (log10)
b
R D0-W2
R D0-W14
NR D0-W2
NR D0-W14
1097
280
297
4
8
19
511
0
c
GMCSF pathway
FCER1 pathway
TLR signaling
VEGFR1_2 pathway
Endothelin pathway
Molecular mediator involved in inflammatory response
TLR4 cascade
FCGR mediated phagocytosis
FCERI signaling
FC receptor mediated stimulatory signaling
Insulin signaling
CXCR4 pathway
Regulation of exocytosis
Regulation of leukocyte migration
Positive regulation of NFKB
Regulation of fiber organization
Activation of protein kinase
MAPK activity
Positive regulation of cytoskeleton organization
VEGF signaling
TNF superfamily cytokine production
TNF signaling via NFKB
Degree
Betweenness
Median FC D0-W2

Fig. 3

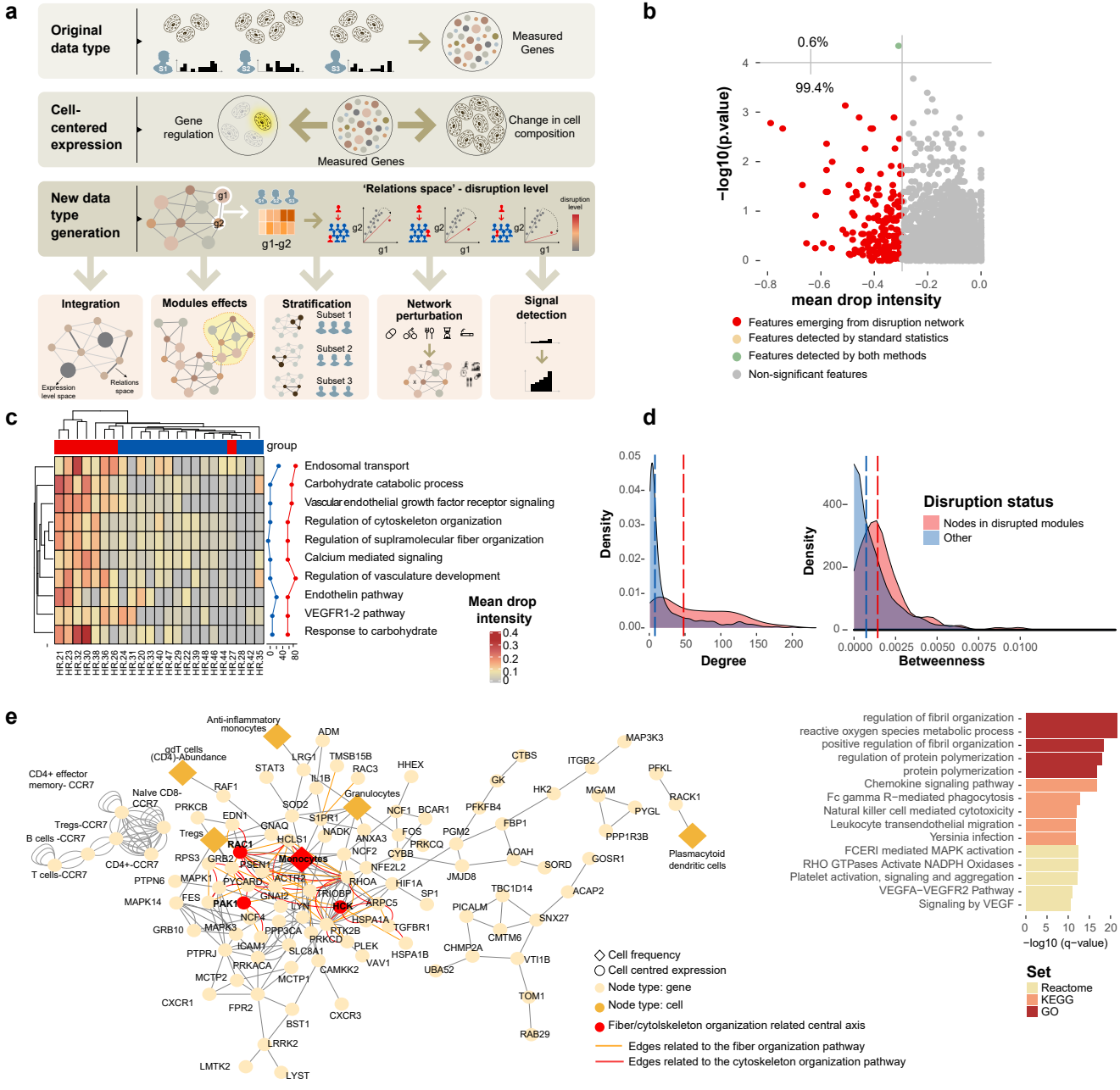

a
Original data type
Measured Genes
Cell-centered expression
Gene regulation
Measured Genes
Change in cell composition
New data type generation
'Relations space' - disruption level
g1-g2
disruption level
Integration
Expression level space
Relations space
Modules effects
Stratification
Subset 1
Subset 2
Subset 3
Network perturbation
Signal detection
b
0.6%
99.4%
−log10(p.value)
mean drop intensity
Features emerging from disruption network
Features detected by standard statistics
Features detected by both methods
Non-significant features
c
group
Endosomal transport
Carbohydrate catabolic process
Vascular endothelial growth factor receptor signaling
Regulation of cytoskeleton organization
Regulation of suplramolecular fiber organization
Calcium mediated signaling
Regulation of vasculature development
Endothelin pathway
VEGFR1-2 pathway
Response to carbohydrate
Mean drop intensity
d
Density
Degree
Disruption status
Nodes in disrupted modules
Other
Betweenness
e
Anti-inflammatory monocytes
gdT cells (CD4)-Abundance
CD4+ effector memory- CCR7
Nalve CD8- CCR7
Tregs-CCR7
B cells -CCR7
T cells-CCR7
CD4+-CCR7
Tregs
Granulocytes
Monocytes
Plasmacytoid dendritic cells
Cell frequency
Cell centred expression
Node type: gene
Node type: cell
Fiber/cytolskeleton organization related central axis
Edges related to the fiber organization pathway
Edges related to the cytoskeleton organization pathway
regulation of fibril organization
reactive oxygen species metabolic process
positive regulation of fibril organization
regulation of protein polymerization
protein polymerization
Chemokine signaling pathway
Fc gamma R−mediated phagocytosis
Natural killer cell mediated cytotoxicity
Leukocyte transendothelial migration
Yersinia infection
FCERI mediated MAPK activation
RHO GTPases Activate NADPH Oxidases
Platelet activation, signaling and aggregation
VEGFA−VEGFR2 Pathway
Signaling by VEGF
−log10 (q−value)
Set
Reactome
KEGG
GO

Fig. 4

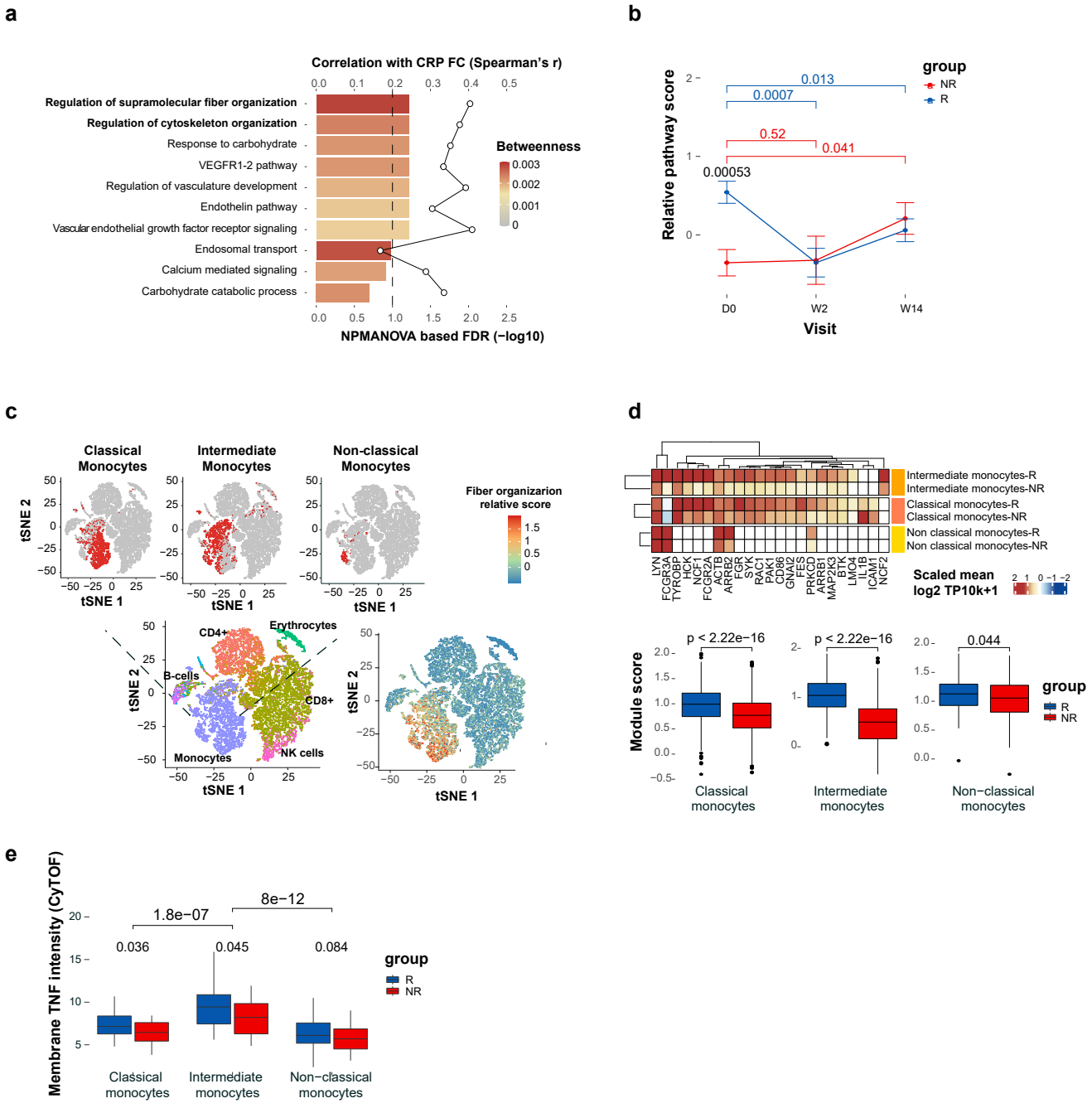
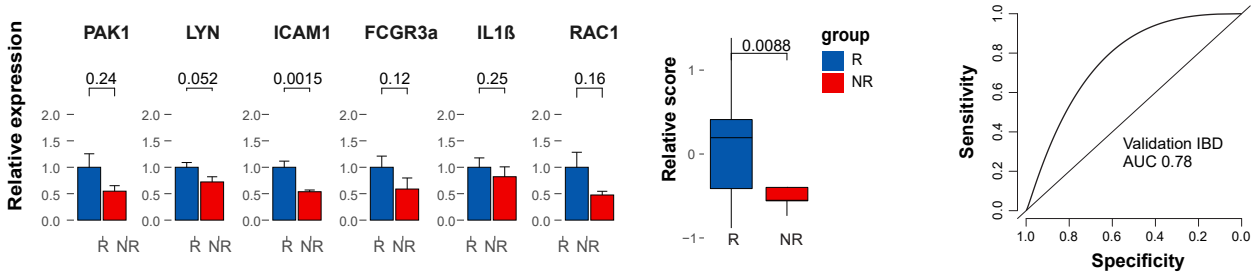
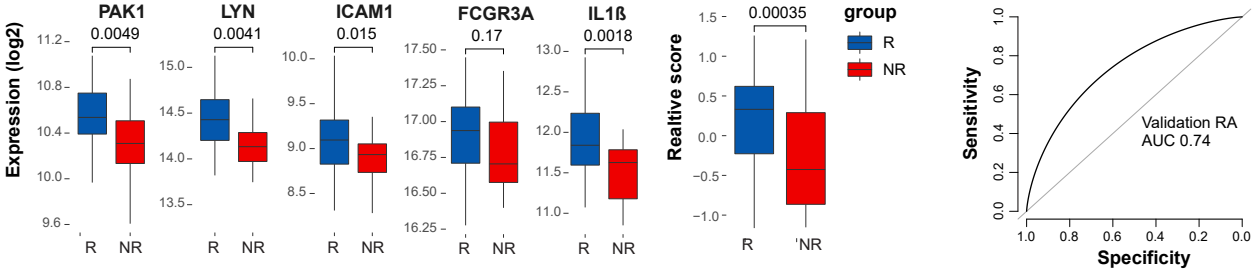

a
Correlation with CRP FC (Spearman's r)
Regulation of supramolecular fiber organization
Regulation of cytoskeleton organization
Response to carbohydrate
VEGFR1-2 pathway
Regulation of vasculature development
Endothelin pathway
Vascular endothelial growth factor receptor signaling
Endosomal transport
Calcium mediated signaling
Carbohydrate catabolic process
Betweenness
NPMANOVA based FDR (−log10)
b
Relative pathway score
group
NR
R
0.013
0.0007
0.52
0.041
0.00053
D0
W2
W14
Visit
c
Classical Monocytes
Intermediate Monocytes
Non-classical Monocytes
Fiber organizarion relative score
tSNE 1
tSNE 2
CD4+
Erythrocytes
B-cells
CD8+
Monocytes
NK cells
d
Intermediate monocytes-R
Intermediate monocytes-NR
Classical monocytes-R
Classical monocytes-NR
Non classical monocytes-R
Non classical monocytes-NR
LYN
FCGR3A
TYROBP
HCK
NCF1
FCGR2A
ACTB
ARRB2
FGR
SYK
RAC1
PAK1
CD86
GNAI2
FES
PRKCD
ARRB1
MAP2K3
BTK
LMO4
IL1B
ICAM1
NCF2
Scaled mean log2 TP10k+1
Module score
p < 2.22e−16
p < 2.22e−16
0.044
Classical monocytes
Intermediate monocytes
Non−classical monocytes
e
Membrane TNF intensity (CyTOF)
1.8e−07
8e−12
0.036
0.045
0.084
group
R
NR
Classical monocytes
Intermediate monocytes
Non−classical monocytes

Fig. 5

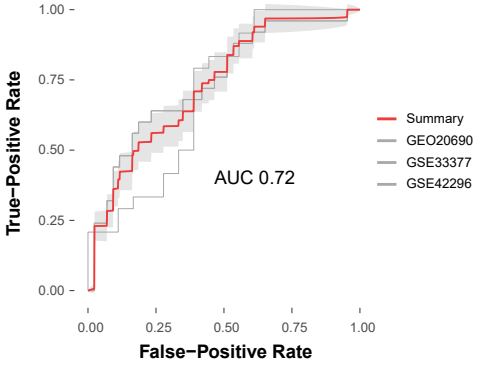
a


Relative expression
PAK1
LYN
ICAM1
FCGR3a
IL1ß
RAC1
0.24
0.052
0.0015
0.12
0.25
0.16
R NR
Relative score
0.0088
group
R
NR
Sensitivity
Specificity
Validation IBD
AUC 0.78

b


Expression (log2)
PAK1
0.0049
LYN
0.0041
ICAM1
0.015
FCGR3A
0.17
IL1ß
0.0018
R NR
Realtive score
0.00035
group
R
NR
Sensitivity
Specificity
Validation RA
AUC 0.74

c


True−Positive Rate
False−Positive Rate
AUC 0.72
Summary
GEO20690
GSE33377
GSE42296