A narrow repertoire of transcriptional modules responsive to pyogenic bacteria is impaired in patients carrying loss-of-function mutations in MYD88 or IRAK4

Laia Alsina1,2, Elisabeth Israelsson3, Matthew C Altman3, Kristen K Dang3, Pegah Ghandi4,5, Laura Israel4,5, Horst von Bernuth4–6, Nicole Baldwin1, Huanying Qin1, Zongbo Jin1, Romain Banchereau1, Esperanza Anguiano1, Alexei Ionan1, Laurent Abel4,5, Anne Puel4,5,7, Capucine Picard4,5,7–9, Virginia Pascual1, Jean Laurent Casanova4,5,7,9,10 & Damien Chaussabel1,3,11

Loss of function of the kinase IRAK4 or the adaptor MyD88 in humans interrupts a pathway critical for pathogen sensing and ignition of inflammation. However, patients with loss-of-function mutations in the genes encoding these factors are, unexpectedly, susceptible to only a limited range of pathogens. We employed a systems approach to investigate transcriptome responses following in vitro exposure of patients’ blood to agonists of Toll-like receptors (TLRs) and receptors for interleukin 1 (IL-1Rs) and to whole pathogens. Responses to purified agonists were globally abolished, but variable residual responses were present following exposure to whole pathogens. Further delineation of the latter responses identified a narrow repertoire of transcriptional programs affected by loss of MYD88 function or IRAK4 function. Our work introduces the use of a systems approach for the global assessment of innate immune responses and the characterization of human primary immunodeficiencies.

Studies of children with recurrent invasive pneumococcal infection have led to the discovery of two primary immunodeficiencies (PIDs) caused by autosomal recessive mutations in the gene encoding the kinase IRAK4 or the gene encoding the adaptor MyD88, which impair signaling via Toll-like receptors (TLRs) and the family of receptors for interleukin 1 (IL-1Rs) (collectively referred to as ‘TIRs’). TLRs act as sensors of the innate immune system by recognizing specific components conserved among a variety of microorganisms that invade the host. The ligation of TLRs by their agonists induces an inflammatory response to control the infection. In humans, ten functional members of the TLR family have been identified (TLR1–TLR10). MyD88 is a key downstream adaptor for most IL-1Rs and all TLRs except TLR3, which signals independently of MyD88, and TLR4, which can also signal via another adaptor (TIRAP); IRAK4 is selectively recruited to TLRs and IL-1Rs by MyD88 (refs. 4–6). Thus, loss of the function of IRAK4 or MyD88 interrupts a pathway critical for pathogen sensing (TLR) and inflammation (IL-1R), known as ‘TIR deficiency’.

Indeed, animal models indicate that deficiency in either of these two molecules results in heightened susceptibility to a wide range of pathogens. However, unexpectedly, patients deficient in IRAK4 or MyD88 display a narrow and transient infectious phenotype mostly limited to pyogenic Gram-positive bacteria ( Streptococcus pneumoniae and Staphylococcus aureus) and Gram-negative bacteria ( Pseudomonas aeruginosa) in particular, and mostly from birth until adolescence, when the condition improves.

The contrast between the broad and profound immunological phenotype and the narrow and transient infectious phenotype of such patients is notable. In assessments of the immunocompetency of patients with PIDs, the choice of the panel of analytes used as a ‘readout’ is crucial. Indeed, a normal response does not exclude the possibility of a defect that affects a pathway not covered by the panel. Conversely, a defective response detected by a given panel does not exclude the possibility of the presence of a conserved, potentially redundant, response. Adopting a systems-scale profiling approach as the ‘readout’ can address such limitations. A commonly used systems approach is ex vivo profiling of the blood transcriptome. This approach has led to the identification of novel therapeutic targets and the development of biomarker signatures. Subsequently, the same approach has been adopted to investigate responses to vaccines.

We have developed an unbiased approach for evaluating the transcriptional profiles of blood cells from patients deficient in IRAK4 or MyD88 in response to a broad array of agonists of TLRs and IL-1Rs, as...
well as in response to whole pathogens (bacteria, viruses and fungi). Our hypothesis was that immunological mechanisms dependent on and independent of MyD88 and IRAK4 that are used for the recognition of pathogens would be revealed in vitro, which might contribute to understanding of the in vivo protection from most microbes in such patients. This approach allowed assessment of the range of responses that could be elicited in the absence of functional IRAK4 or MyD88 and provided insight into mechanisms essential for the maintenance of immunity to pathogens.

**RESULTS**

*In vitro* transcriptome responses to purified TIR agonists

Before evaluating the effect of deficiency in IRAK4 or MyD88 on the function of the innate immune system, we set out to establish in a set of control subjects the baseline response elicited by the engagement of purified TIR agonists. We measured transcriptional responses to TIR agonists in control subjects (*r* = 14) on a genome-wide scale. We stimulated whole blood in *vitro* for 2 h with an array of agonists spanning all TLRs (the synthetic lipopeptide Pam3CSK4 (PAM3), an agonist of TLR1 and TLR2; the synthetic lipopeptide Pam2CSK4 (PAM2), an agonist of TLR2 and TLR6; the synthetic RNA duplex poly(I:C), an agonist of TLR3; lipopolysaccharide (LPS), an agonist of TLR4; flagellin, an agonist of TLR5; the imidazoquinoline 3M13, an agonist of TLR8; the synthetic lipopeptide Pam 2CSK4 (PAM2), an agonist of TLR1 and TLR2; the synthetic lipopeptide Pam 3CSK4 (PAM3), an agonist of TLR3 for which signaling is MyD88 independent; for 2 h with an array of agonists spanning all TLRs (the synthetic lipopeptide Pam3CSK4 (PAM3), an agonist of TLR1 and TLR2; the synthetic lipopeptide Pam2CSK4 (PAM2), an agonist of TLR2 and TLR6; the synthetic RNA duplex poly(I:C), an agonist of TLR3; lipopolysaccharide (LPS), an agonist of TLR4; flagellin, an agonist of TLR5; the imidazoquinoline 3M13, an agonist of TLR8; the synthetic imidazoquinoline resiquimod (R-848), an agonist of TLR7 and TLR8; and the dinucleotides Cpg-D19 and Cpg-C, agonists of TLR9) and agonists of IL-1Rs (IL-1β and TLR8; and the dinucleotides CpG-D19 and CpG-C, agonists of TLR9) and agonists of IL-1Rs (IL-1β and TLR8; and the synthetic imidazoquinoline resiquimod (R-848), an agonist of TLR7 and TLR8; and the dinucleotides Cpg-D19 and Cpg-C, agonists of TLR9) and agonists of IL-1Rs (IL-1β, IL-18 and IL-33), along with two positive controls, tumor-necrosis factor (TNF) and the combination of the phosphor ester PMA plus ionomycin. We selected transcripts displaying consistent differences in stimulated expression in samples from healthy control subjects as described (Online Methods). This filter identified sets of transcripts responding to each stimulation. These included transcripts encoding various chemokines, cytokines, costimulatory molecules, antibacterial peptides and transcriptional regulators known to be involved in the TIR signaling pathway. Pathway analysis (with Ingenuity Pathway Analysis software) confirmed that genes encoding molecules involved in trafficking of cells of the immune system and inflammatory response were significantly over-represented in these gene lists (*P* < 0.0001), as were genes encoding molecules involved in signaling via type I interferons after stimulation with LPS, poly(I:C), 3M13, 3M2 or R-848 (*P* < 0.001). The magnitude of transcriptional changes varied for each stimulus: LPS, R-848 and TNF induced stronger responses (713 probes, 535 probes and 550 probes, respectively), and Pam3, Pam2, 3M13, IL-1β and IL-18 induced weaker responses (101 probes, 183 probes, 80 probes, 99 probes and 119 probes, respectively). Several transcripts showed overlap among stimulation conditions. For example, 70 annotated probes (59 genes) were induced by all TLR agonists except poly(I:C) (a nonspecific ligand of TLR3 for which signaling is MyD88 independent) (Supplementary Fig. 1). CpG-D19, CPG-C and IL-33 were weak inducers at this early time point, and culture conditions including these agonists were not used in further analyses. This first step established the *in vitro* transcriptional response to TIR agonists in blood from healthy people.

**Characterization of responses in patients with TIR deficiency**

Next we assessed the ability of IRAK4- or MyD88-deficient patients to respond following engagement of TIRs by purified agonists. Our initial cohort included four patients (P1–P4) with IRAK4 deficiency and four with MyD88 deficiency (Table 1). All the mutations of IRAK4

| Table 1 Patient information | Batch 1 | Batch 2 | Batch 3 |
|-----------------------------|---------|---------|---------|
| IRAK4−/− patients | P1 (refs. 1, 9): male, 12 years of age Mutation: 1189+5020A>G/1189+1G>T (splice IVIG P2 (refs. 1, 9): male, 12 years of age Mutation: E402X/E402X (nonsense) Cotrimoxazole P3 (refs. 1, 9): male, 18 years of age Mutation: Q293X/BAC210N13del (nonsense; large deletion) Cotrimoxazole P4 (ref. 9): female, 14 years of age Mutation: Q293X/Q293X (nonsense) Azthyromycin and IVIG |
| MYD88−/− patients | P1a,b (refs. 2, 9): female, 18 years of age Mutation: R209C/R209C (missense) Cotrimoxazole and Peni V P2a (refs. 2, 9): male, 11 years of age Mutation: R209C/R209C (missense) Peni V and IVIG P3 (ref. 9): female, 1 year of age Mutation: E65del/E65del (deletion) Cotrimoxazole, Peni V and IVIG P4a (refs. 2, 9): female, 5 years of age Mutation: L106P/R209C (missense) Cotrimoxazole and IVIG |
| Control subjects | n = 14 (6 for bacterial stimulation) |
| Stimulation | 9 TIR agonists + TNF + PMA-iono + 4 HK (IRAK4−/−, P3 and P4; MYD88−/−, P3 and P4) 2 TIR agonists + TNF + PMA-iono + 4 HK (all patients) 4 TIR agonists + PMA-iono + 4 HK + BCG + HSV + HK.CA (all patients) |

Patient characteristics (including sex and age; mutation present; and treatment), as well as conditions used for stimulation of samples plus patient samples used (bottom row). IVIG, intravenous immunoglobulin; Peni V, penicillin V; 9 TIR agonists: Pam3, Pam2, poly(I:C), flagellin, LPS, R-848, Cpg, IL-1β and IL-18; 2 TIR agonists: Pam2 and LPS; 4 TIR agonists: Pam2, Pam3, poly(I:C), and R-848; PMA-iono: PMA-iono plus ionomycin; 4 HK: three strains of heat killed S. pneumoniae and S. aureus; HK.CA: heat killed C. albicans. Patients who are siblings. Patients included in more than one batch for reproduction of findings and investigation of additional pathogen stimulation.
or MYD88 in these patients result in loss of function of the protein product\(^1,2,9\). All patients were asymptomatic and had no active infectious process at the time of sample collection. They were 1–18 years of age. The responses to PMA plus ionomycin (positive control) and to poly(I:C) were conserved: 92.4% of transcripts responsive in healthy control subjects were also responsive in these patients (called ‘92.4% of the healthy response’ here). The response to LPS (partially MyD88 dependent) was reduced but was not abolished (40.2% of the healthy response).

However, we observed considerably fewer responsive transcripts in both IRAK4-deficient patients and MyD88-deficient patients than in the healthy control subjects in response to PAM3 (8.8% of the healthy response), PAM2 (6% of the healthy response), flagellin (19.7% of the healthy response), 3M13 (13.8% of the healthy response), 3M2 (16.3% of the healthy response), R-848 (3.9% of the healthy response), IL-1β (19.5% of the healthy response) and IL-18 (23.4% of the healthy response). The extent of the defect for each stimulus was consistent across the patients, both quantitatively and qualitatively (Fig. 1a and Supplementary Fig. 2). When we considered all stimuli together, clustering according to level of responsiveness across subjects resulted in a distinct separation between MyD88-dependent signals and MyD88-independent signals (Fig. 1b). Conversely, clustering of subjects across all stimuli resulted, as expected, in a distinct separation of patients versus control subjects (Fig. 1b). We observed no difference between MyD88-deficient patients and IRAK4-deficient patients in their responsiveness (Fig. 1b). The residual response...
Modular repertoire mapping of TIR transcriptome responses

We next sought to characterize the transcriptional programs affected by loss of MyD88- or IRAK4-dependent signaling. We used a data-mining strategy that consists of mapping relationships among group of genes on the basis of similarities in expression patterns across a wide range of conditions. The approach devised for the construction of such modular transcriptional repertoires has been described in detail elsewhere. For this study, we constructed a large gene co-clustering network using responses to each stimulus across all subjects as input data sets. Network analysis identified gene sets that clustered together (‘transcriptional modules’). The resulting collection of modules served as a framework for downstream data analysis and interpretation. This data-driven process identified a repertoire consisting of 66 modules composed of 1,088 transcripts. We obtained functional annotations for each module (Supplementary Table 1). In addition, we broadly categorized modules on the basis of their patterns of response, in control subjects, to the various TIR agonists and bacteria (Fig. 3). Thus, we grouped all 66 modules into seven clusters (0–6).

Marginal impairment of responses to whole bacteria

To use a stimulus more akin to that experienced by patients during infection, we exposed samples from MyD88-deficient patients (n = 5) and IRAK4-deficient patients (n = 3) to whole heat-killed bacteria (Table 1). The bacteria we used in this assay were those to which IRAK4- or MyD88-deficient patients are most susceptible: S. pneumoniae and S. aureus. For S. pneumoniae, we used three different strains, including the less-virulent non-encapsulated R6 strain.

In control subjects, the magnitude of transcriptional change varied according to the bacteria: S. aureus Cowan strain (SAC) induced the lowest response (101 transcripts), lower than the response to S. pneumoniae (33 transcripts, 282 transcripts and 264 transcripts for the R11470 strain, R8450 strain and R6 strain, respectively) (Fig. 2). Patients displayed substantial levels of transcriptional activity in response to heat-killed bacteria. The response in MyD88-deficient patients was consistently lower than that in IRAK4-deficient patient: R11470, 63% versus 84%; R8450, 45% versus 99%; R6, 46% versus 89%; and SAC, 47% versus 75% (all '% of the healthy response'). We did not observe this difference earlier when we used purified TIR agonists. Thus, this evaluation of transcriptional response to bacterial pathogens showed that, in contrast to the results obtained with purified TIR agonists, patients deficient in MyD88 or IRAK4 retained the ability to respond to S. pneumoniae and S. aureus at the transcriptional level (>50% of the healthy response (average; range, 27–102%) through MyD88- and IRAK4-independent mechanisms; this highlighted the redundancy of the microbe-sensing system in blood leukocytes. Furthermore, our findings suggested that patients with MyD88 deficiency might differ from those with IRAK4 deficiency in their ability to respond to Gram-positive bacteria (an average of 50% (MyD88 deficiency) versus 87% (IRAK4 deficiency) of the healthy response for all four bacterial stimulations).

Modular repertoire mapping of TIR transcriptome responses

We next sought to characterize the transcriptional programs affected by loss of MyD88- or IRAK4-dependent signaling. We used a data-mining strategy that consists of mapping relationships among group of genes on the basis of similarities in expression patterns across a wide range of conditions. The approach devised for the construction of such modular transcriptional repertoires has been described in detail elsewhere. For this study, we constructed a large gene co-clustering network using responses to each stimulus across all subjects as input data sets. Network analysis identified gene sets that clustered together (‘transcriptional modules’). The resulting collection of modules served as a framework for downstream data analysis and interpretation. This data-driven process identified a repertoire consisting of 66 modules composed of 1,088 transcripts. We obtained functional annotations for each module (Supplementary Table 1). In addition, we broadly categorized modules on the basis of their patterns of response, in control subjects, to the various TIR agonists and bacteria (Fig. 3). Thus, we grouped all 66 modules into seven clusters (0–6). Clusters 0, 1, and 6 consisted of modules of genes responsive to either TIR agonists or bacteria or to both. Cluster 6 was composed mainly of inflammation-related modules, including cytokines (IL-1β, IL-6 and IL-8), chemokines, the transcription factor NF-κB, acute-phase response elements and neutrophil function and phagocytosis (Fig. 4 and Table 2); the genes encoding these were upregulated by both TIR agonists and bacteria. Cluster 4 was composed of modules consisting of genes responsive to TIR agonists. For example, the genes in modules M5.4 and M7.2 were ‘preferentially’ induced by LPS, 3M2 or R-848 and, to a lesser extent, poly(I:C), and corresponded to an interferon-related inflammatory response not induced by Gram-positive bacteria. Bacterial stimulation triggered genes in a set of modules not induced by stimulation of individual TLRs with pathogen-associated molecular patterns and thus could be considered specific to whole...
bacteria (cluster 5) (Fig. 4 and Table 2). Annotation of cluster 5 suggested a role in cell signaling for the molecules encoded by those genes (Fig. 4 and Table 2). This suggested that in blood cultures exposed to whole bacteria, receptors other than TIRs were engaged (possibly receptors associated with immunoreceptor tyrosine-based activation motifs and the transmembrane adaptor DAP12 (ref. 19), such as CD300, integrins, Fc receptors, C-type lectins and complement) and for delineating the transcriptional programs affected by deficiency.

We next used the modular repertoire reported above as a framework for delineating the transcriptional programs affected by deficiency in MyD88 or IRAK4. We plotted the data in circular heat maps to represent residual patient responses to each TIR stimulation across the seven module clusters, in which each ‘ring’ of the circle represented a patient and each ‘spoke’ represented a module (Fig. 5 and Supplementary Fig. 3). As expected on the basis of our findings reported above, patients were unable to mount a response to purified MyD88-dependent TIR agonists for any of the modules described above, including those in clusters 4 and 6, which were associated with inflammation (residual responses to Pam2, Supplementary Fig. 3); all others, Supplementary Fig. 3). We observed residual responses to stimulation with LPS (partially MyD88 dependent) for modules in cluster 6 (inflammation) displaying overall well-conserved responses, in contrast to modules in cluster 4 (interferon), which displayed poor residual responses (Fig. 5). Next we analyzed in a similar fashion the responses to heat-killed bacteria in patients deficient in IRAK4 or MyD88. The overall responses to all four heat-killed bacteria were relatively well preserved (Fig. 5). The most-preserved modular responses to Gram-positive bacteria in both MyD88-deficient patients and IRAK4-deficient patients were as follows: module M2.11, which includes IL1A, IL1B, TNF, CCL20, CCL3L1 and CXCL1; module M5.5, which includes IL8, IRAK2, CCL3L3, CXCL2, IL1RN, PLAU and PTGS2; module M6.6, which includes CCR2, CYP4B1, NLRP3, OSM, PTGS2 and TAGAP; and module M8.4, which includes CCL3,
CCL3L1, CCL4L2, NFKBIA, NLRP3 and PLAUR. The genes in these four modules, all from cluster 6, displayed substantial upregulation for all donors (>90% for each control subject and IRAK4-deficient patient, and >50% for each MyD88-deficient patient) relative to the average of healthy subjects (Fig. 6). This indicated that patients were able to induce a proinflammatory program in response to activation with S. pneumoniae and S. aureus, probably through the participation of other sensors, such as activators of the receptor NLRP3 in the form of inflammasome components, which were present in two of the four most-preserved modules.

**Impairment of a narrow modular repertoire in response to bacteria**

Further examination of modular patterns of responsiveness to whole bacteria conversely revealed impairment of specific transcriptional programs. Indeed, the residual responsiveness among the modules constituting cluster 6 diverged markedly (Figs. 5 and 6). We observed this heterogeneity both across modules and across bacterial species or strains (Fig. 6 and Supplementary Fig. 4). However, three modules (M4.3, M4.7 and M6.3) did present consistently low residual responsiveness across these conditions. Of the three modules, responsiveness was most consistently impaired for module M4.3. This module includes genes encoding molecules related to the activation of NF-κB (NFKB1, NFKB2 and IRAK3), regulation (SRC and TNIP1), apoptosis (BIRC3, CFLAR, DENND5A (which encodes an activator of Rab39 and IER5), IL-1β and the inflammasome (CARD16 and P2RX7). The responses of genes in module M4.3 to the R6 strain of S. pneumoniae were abolished in all but one IRAK4-deficient patient (Fig. 6a); samples from this patient, however, displayed a blunted response when exposed to S. aureus (Fig. 6b). This suggested species- or strain-specific fluctuations in susceptibility to pathogens (responses to strains R8450 and R11470 of S. pneumoniae, Supplementary Fig. 4). Overall responses were also similarly blunted for module M6.3, which includes TNAIP8, IRG1 and genes encoding molecules related to cell metabolism (ACSL1, PDE4B and RIN2). Of interest, IRG1 enhances macrophage bactericidal activity21, and ACSL1 has been found to serve a role as an inflammatory mediator in LPS-stimulated macrophages22. M4.7 was another module with overall decreased responsiveness. Notably, mutation of NBN (encoding nibrin), which is one of the nine genes of module M4.7, causes Nijmegen breakage syndrome, a DNA-repair PID characterized by combined cellular and humoral immunodeficiency with severe recurrent sinopulmonary infections that cause substantial mortality in this patient population23. A single-nucleotide polymorphism in NFKBIE, another gene of module M4.7, has been associated with susceptibility to invasive pneumococcal diseases24.
Finally, a third gene of module M4.7, CLIC4, has been found to encode a molecule with an important role in mediating innate resistance to bacterial infection in an animal model. We observed these module-specific defects again in a third set of samples (batch 3, Table 1) that included additional blood samples drawn from two patients who had been included in cohorts studied before (IRAK4−/− patient P5 and MYD88−/− patient P4) and from new two patients (IRAK4−/− patient P6 and MYD88−/− patient P7). This new set of samples showed the same degree of preserved transcriptional activity in response to heat-killed bacteria as that of samples from the previously studied set of patients (Supplementary Fig. 5), and when we delineated this response, we again found that all showed low residual responsiveness for modules M4.3, M4.7 and M6.3 (Supplementary Fig. 6).

**Residual responses to viral, fungal and bacterial pathogens**

Mice lacking IRAK4 or MyD88 are susceptible to a wide range of pathogens, in contrast to the narrow range of susceptibility observed in humans with inborn errors in IRAK4 or MyD88 (ref. 8). Thus, we subsequently investigated patterns of transcriptional response in these patients upon stimulation of whole blood with pathogens to which deficient mice are susceptible but humans are not. We measured responses to heat-killed *Candida albicans*, *Mycobacterium bovis* bacillus Calmette-Guérin (BCG) or herpes simplex virus type 1 (HSV) in blood from two MyD88-deficient patients and two IRAK4-deficient patients, as well as eight control subjects (batch 3, Table 1). Patients displayed substantial transcriptional activity in response to heat-killed *C. albicans*, with nearly normal residual responsiveness in three of the four patients (Fig. 7a). In response to stimulation with BCG, MyD88-deficient patients showed markedly deficient responses (<30% of the healthy response, globally) that affected all the inflammatory modules (Fig. 7a), while an IRAK4-deficient patient had preserved responsiveness both globally (90% of the healthy response) (Fig. 7a) and at the module level (Fig. 7b, middle; data could not be obtained for the second IRAK4-deficient patient). These observations could not be attributed to prior status of vaccination with BCG, as this varied among patients (MyD88-deficient patient P1, no vaccination; MyD88-deficient patient P2, prior vaccination; IRAK4-deficient patient P2, prior vaccination). Finally, both MyD88-deficient patients and IRAK4-deficient patients showed globally defective transcriptional responses to the stimulation of whole blood with HSV (<20% of the healthy response; Fig. 7a). At the module level, one single inflammatory module, M7.4, was selectively preserved upon stimulation with HSV (Fig. 7b). This module includes genes encoding molecules such as NFIL3 (E4BP4), related to the differentiation and function of macrophages, dendritic cells and natural killer cells, that are required for antiviral immunity, and c-Myc and c-Rel, related to signaling via NF-kB. This finding suggested that while signaling via TLRs in response to HSV was for the most part abolished, alternative pathways involving signaling via NF-kB might have been preserved and contributed to maintaining resistance to this pathogen in patients deficient in MyD88 or IRAK4.

In summary, our use of a systems approach revealed a multifaceted pattern of responsiveness of the innate immune system in patients with deficiency in MyD88 or IRAK4. Our results demonstrated a profound defect in the ability of blood leukocytes from patients deficient in MyD88 or IRAK4 to respond to pathogen- and host-derived agonists of TIRs. Responses to specific TIR agonists were completely abrogated, while responses to whole organisms ranged from being relatively normal (heat-killed *C. albicans*) to considerably diminished (HSV). We also observed differences between patients deficient in MyD88 and those deficient in IRAK4 in their responsiveness in several conditions. Transcriptional programs elicited by pyogenic bacteria to which both groups are clinically susceptible were not associated with a substantial reduction in overall responsiveness but instead were associated with impairment of specific inflammatory transcriptional programs.
DISCUSSION

We devised a systems approach to assess TIRs and antibacterial immunity in patients with enhanced susceptibility to infection. The distinct benefit of using a genome-wide approach is that it provides an unbiased means with which to investigate responses to stimuli of the innate immune system, in contrast to more traditional approaches that require reliance on a priori knowledge for the selection of a small panel of analytes as a ‘readout’. Here we assessed the global effect of defects in IRAK4 or MyD88 on transcriptional programs of the innate immune system. Patients’ responses to TLR2 agonists were less than 10% of those of healthy subjects. TLR2 is known to be crucial for the recognition of Gram-positive bacteria (S. pneumoniae and S. aureus), which are pathogens to which patients deficient in IRAK4 or MyD88 are most susceptible.

We confirmed the close interdependence of IRAK4 and MyD88 in the signaling pathway downstream of TIRs, concordant with the functional structure they form (the ‘Myddosome’). However, in contrast to stimulation with purified TIR agonists, stimulation with whole pathogens revealed differences in the ability of patients deficient in IRAK4 or MyD88 to mount transcriptional responses. This might be explained by the involvement of MyD88 in signaling via the GTPase Ras and MAPKs through direct interaction with the kinase Erk without the participation of IRAKs. Nevertheless, such differences do not ‘translate’ into cytokine production, which has been shown to be equally impaired in the same cultures assessed with different ‘readouts’, or into the infectious clinical phenotype, which is indistinguishable in patients deficient in IRAK4 versus those deficient in MyD88.

Despite their profound defects in the common TIR signaling pathway, patients deficient in IRAK4 or MyD88 were able to upregulate the expression of genes in the major inflammatory modules (modules M2.11, M5.5, M8.4 and M6.6) when their blood was exposed to S. pneumoniae or S. aureus. This indicated that a MyD88-independent but NF-kB- and MAPK-dependent pathway is responsible for this proinflammatory response, since the same inflammatory program is observed after stimulation of specific TIRs. S. aureus and S. pneumoniae can be recognized by the receptors Nod2 (ref. 32) and TNFR1 (ref. 33), both of which are expressed in leukocytes, activate NF-kB and MAPKs and initiate the same inflammatory processes initiated by TLRs and do not signal via MyD88. There is consistent evidence that both TLRs and Nod2 act in synergy to induce combined responses. The lack of synergism with TLR2 might explain the suboptimal response observed (lack of upregulation of genes in the modules M4.3, M4.7, M6.3). Those results are consistent with published data obtained at the protein level: the secretion of TNF in whole blood from IRAK4-deficient patients was undetectable 24 h after stimulation of TLRs but was detectable after stimulation with heat-killed S. aureus, although to a lower extent than in blood from healthy control subjects. The residual induction of cytokines via MyD88- or IRAK4-independent signaling pathways might account for both the resistance to other infections and the fever and inflammation that can be observed at late stages of infection.

Such results raise the following question: can the susceptibility to bacteria in patients with TIR deficiencies be ascribed to this partially defective inflammatory response, or do other mechanisms contribute to the phenotype? Our assay evaluated responses only in whole blood, not in mucosal or skin. IL-1Rs are crucial in maintaining protective immunity to invasive staphylococcal skin infection. Indeed, neutrophils have a key role in clearance of bacteria from epithelial sites, and IL-1R-mediated signaling by skin-resident cells would be essential for the adequate recruitment of neutrophils to the site of infection. A published study of skin-derived fibroblasts from both IRAK4-deficient patients and MyD88-deficient patients showed no response to activation of IL-1R. Also, IL-1Rs have proven crucial in other infections with S. aureus. It might be that patients deficient in IRAK4 or MyD88 are able to generate an initial proinflammatory response after an initial encounter with Gram-positive bacteria not only in whole blood, as we have shown here, but presumably at epithelial sites as well (Nod2 expression in keratinocytes and the lungs has been reported). TNFR1 is widely distributed on the airway epithelium. Nevertheless, defective signaling via IL-1Rs could result in an impaired systemic amplification of this initial response, ultimately increasing the risk of not only epithelial infections but also bacteremia starting from the skin and mucosa, which is characteristic of these patients. This would also explain why patients deficient in IRAK4 or MyD88 are clinical phenocopies despite their different degrees of response to whole bacteria in blood, since they displayed similar defects in response to stimulation with IL-1B.

MyD88-deficient patients showed marked impairment in their transcriptional responses to the activation of whole blood with BCG, despite prior vaccination with BCG. Such results are consistent with the strong TLR2-agonist activity of BCG in mice. However, MyD88-deficient patients are not susceptible to mycobacteria or to BCG. In humans, resistance to BCG is known to be dependent on interferon-γ, and, production of interferon-γ is probably preserved through MyD88- and IRAK4-independent pathways. Similarly, while the abnormal responses to HSV observed in both MyD88-deficient patients and IRAK4-deficient patients attest to the role of TLR2 and TLR9 in the initiation of immune responses to HSV, they do not result in enhanced susceptibility to this pathogen. Published work has demonstrated that control of HSV in the central nervous system is selectively dependent on a functional TLR3 pathway in neurons and oligodendrocytes but not in blood leukocytes.

Technologies available for profiling transcript abundance on a genome-wide scale have become robust and cost-effective. The application of such technologies in whole blood has been used extensively for the investigation of disease pathogenesis, identification of biomarkers and assessment of responses to vaccines. Molecular phenotypes may not always be apparent in blood profiled ex vivo in the steady state, as is the case for patients with deficiency in MyD88 or IRAK4 (data not shown). We have shown here that using global profiling as a ‘readout’ in an in vitro functional assay can identify comprehensive phenotypes of the innate immune system in patients with PID. Possible clinical applications for a targeted assay derived from the modular transcriptional framework constructed in this study can be foreseen beyond PID; for example, it could be used in clinical settings in which defects in innate immunity are suspected, such as recurrent pyogenic infections or aging, and also for the prediction of responses to vaccines. The choice of conditions for in vitro stimulation is obviously critical; here we selected a panel tailored for inborn errors in TIR pathway. This can be adapted depending on the population being screened.

In conclusion, our work has demonstrated the use of systems approaches as a robust means for the global assessment of competence of the innate immune system. Applied to the study of patients with inborn defects of TIR signaling, this strategy revealed that patients with deficiency in MyD88 or IRAK4 suffered a profound loss of responsiveness to soluble TIR agonists and that their ability to respond to whole bacteria was selectively affected. Our findings are consistent with the delayed clinical and biological signs of inflammation observed in such patients during the course of infection. More broadly, by leveraging high-throughput profiling technology,
together with an effective analytical framework, our work also opens new avenues for the widespread use of systems approaches as a global ‘readout’ in functional immunological assays.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. GEO: microarray data, GSE25742.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank I. Marodi, J.C. Rodriguez Gallego, G. Davies, Y. Camcioglu, A. Bousfiha, J. Vasconcelos and J.I. Aróstegui for blood from the patients with deficiency in IRAK4 or MyD88; B. Lemoine and P. Nguyen for sample processing; E. Baron (Assistance Publique-Hôpitaux de Paris) for bacterial strains; F.J. Barrat (Dynavax Technologies) for CpG reagents; M. Chrabieh for technical help; and the patients and their families for help. Supported by the US National Institutes of Health (1R01AR050770-02), the Dana Foundation, La Fundació La Caixa (L.A.), 2008–2010 and the Instituto de Salud Carlos III (PIF12/01990 (also financed by the European Regional Development Fund) to L.A., 2012 to the present.

AUTHOR CONTRIBUTIONS

L.A., acquisition of data, analysis and interpretation of data and drafting of the manuscript; E.I., M.C.A. and K.K.D., figures prepared with the Circos visualization tool; Z.J., acquisition of data; P.G., H.v.B., A.P., L.I. and C.P. sample collection and whole-blood stimulation; A.P. and C.P., critical revision of the manuscript; C.P., sending of patients’ blood; N.B., modular analysis; H.Q., R.B. and A.I., statistical analysis of microarray data; E.A., sample collection and data acquisition; and L.A., V.P., J.L.C. and D.C., study conception and design and critical revision of the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Picard, C. et al. Pyogenic bacterial infections in humans with IRAK-4 deficiency. Science 299, 2076–2079 (2003).
2. von Bernuth, H. et al. Pyogenic bacterial infections in humans with MyD88 deficiency. Science 321, 691–696 (2008).
3. Akira, S., Uematsu, S. & Takeuchi, O. Pathogen recognition and innate immunity. Nat. Rev. Immunol. 11, 37–49 (2011).
4. Lee, C.C., Avalos, A.M. & Ploegh, H.L. Accessory molecules for Toll-like receptors and proteasome function. Curr. Opin. Immunol. 24, 465–470 (2012).
5. Kawai, T. & Akira, S. The role of pattern-recognition receptors in innate immunity: new avenues for the widespread use of systems approaches as a global ‘readout’ in functional immunological assays. Curr. Opin. Immunol. 24, 465–468 (2012).
6. Grass, C. et al. Mechanisms of immunosenescence. Immun. Ageing 9, 36–49 (2012).
7. Schmitt, G. et al. Pyogenic bacterial infections in humans with IRAK-4 deficiency. Cell 124, 783–801 (2006).
8. Kuwana, T. & Akira, S. The role of pattern recognition receptors in innate immunity: new avenues for the widespread use of systems approaches as a global ‘readout’ in functional immunological assays. Nat. Rev. Immunol. 11, 37–49 (2011).
9. Lee, C.C., Avalos, A.M. & Ploegh, H.L. Accessory molecules for Toll-like receptors and their function. Nat. Rev. Immunol. 12, 168–179 (2012).
10. Lin, S.C., Lo, Y.C. & Wu, H. Helical assembly in the MyD88-IRAK4-IRAK2 complex in Toll/LR-1IR signalling. Nature 465, 885–890 (2010).
11. Casanova, J.L., Abel, L. & Quintana-Murci, L. Human TLRs and IL-1Rs in host defense: new insights from evolutionary, epidemiological, and clinical genetics. Annu. Rev. Immunol. 29, 447–491 (2011).
12. von Bernuth, H., Picard, C., Puel, A. & Casanova, J.L. Experimental and natural infections in MyD88- and IRAK-4-deficient mice and humans. Eur. J. Immunol. 42, 3126–3135 (2012).
13. Picard, C. et al. Clinical features and outcomes of patients with IRAK-4 and MyD88 deficiency. Medicine 89, 403–425 (2010).
14. Picard, C., Casanova, J.L. & Puel, A. Infectious diseases in patients with IRAK-4, MyD88, NEMO, or IκBα deficiency. Clin. Microbiol. Rev. 24, 490–497 (2011).
15. Aliantat, F., Chauvat, D., Banchereau, J. & Pascual, V. Microarray-based identification of novel biomarkers in IL-1-mediated diseases. Curr. Opin. Immunol. 21, 623–632 (2009).
16. Berry, M.P. et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. Nature 466, 973–977 (2010).
17. Pascual, V., Chauvat, D. & Banchereau, J. A genomic approach to human autoimmune diseases. Annu. Rev. Immunol. 28, 535–571 (2010).
18. Ramilo, O. et al. Gene expression patterns in blood leukocytes discriminate patients with acute infections. Blood 109, 2066–2077 (2007).
19. Banchereau, R. et al. Host immune transcriptional profiles reflect the variability in clinical disease manifestations in patients with Staphylococcus aureus infections. PLoS ONE 7, e34390 (2012).

ONLINE METHODS

Patient information and sample collection. A total of 40 blood samples were collected from three groups of subjects: 7 patients with complete MyD88 deficiency; 6 patients with complete IRAK4 deficiency; and 27 healthy donors (Table 1). IRAK4 mutations in patients P1, P2, and P3 are null (no IRAK4 protein detected\(^2\)). Mutations in IRAK4-deficient patients P4, P5 and P6 (ref 9) and in MyD88-deficient patients 1–7 are loss-of-function mutations. MyD88-deficient patients P1 and P2 have normal levels of nonfunctional MyD88 protein; MyD88-deficient patients P3–P7 and IRAK4-deficient patients P4 and P5 have small amounts of nonfunctional protein\(^2\). Subjects were recruited in four different sets. The first set included 4 healthy control subjects, MyD88-deficient patients P1 and P2, and IRAK4-deficient patients P1 and P2. The second set included 10 healthy control subjects (different from those in set 1), MyD88-deficient patients P3 and P4, and IRAK4-deficient patients P3 and P4. The first and second sets (batch 1) were analyzed to evaluate the responses of patient samples to thirteen different TLR agonists (ten TLR agonists and three IL-1R agonists). A third set (batch 2) was included afterward to increase the number of MyD88- or IRAK4-deficient patients for bacterial stimulation (three strains of heat killed S. pneumoniae and S. aureus); batch 2 also included five healthy control subjects (different from those in batch 1), MyD88-deficient patients P5 and P6, additional samples from MyD88-deficient patient P1, and IRAK4-deficient patient P5. Four patients from batch 1 and four from batch 2 were analyzed to evaluate the responses to bacteria of samples from MyD88- or IRAK4-deficient patients. The fourth set (batch 3) was included to assess stimulation with alternative pathogen (IRAK4-deficient patient P6, MyD88-deficient patient P7, additional samples from IRAK4-deficient patient P5 and MyD88-deficient patient P4, and eight healthy control subjects (different from those in previous batches). Information about antimicrobial prophylaxis at the time blood was drawn is provided in Table 1.

Subjects were considered to be ‘healthy’ on the basis of past medical history (no recurrent infections) and current health status. All healthy control subjects were adult subjects, except for batch 5, which included two healthy young children 3 and 7 years of age. TLR function is mainly age dependent for the newborn\(^49\), and there was an evident limitation in obtaining the requisite large volume of blood from healthy children to use as control subjects.

Blood samples were obtained at Necker Hospital (Paris, France) with approval of the local Research Ethics Committee. All participants over 18 years of age or their legal guardians (participants under 18 years of age) provided written informed consent.

Blood culture. Peripheral blood was drawn into sodium heparin vacuum tubes at the clinic site (Necker Hospital). Immediately, 500 µl of whole blood was diluted with an equal volume of RPMI medium before the addition of any stimulus; replicates were used for each experimental condition. Diluted whole blood was left unstimulated for 2 h or was activated for 2 h with 10 µg/ml of polymyxin B (to clear contamination with LPS) plus the following agonists: Pam3 (100 ng/ml; InvivoGen), Pam2 (100 ng/ml; InvivoGen), poly(I:C) (25 µg/ml; InvivoGen), LPS (100 ng/ml; Sigma), flagellin (100 ng/ml; InvivoGen), M3i3 (3 µg/ml; 3M Pharmaceuticals), M3M2 (3 µg/ml; 3M Pharmaceuticals), R-848 (3 µg/ml; InvivoGen), CpG-D19 and CpG-C (3 µg/ml; provided by F.J. Barrat), IL-1β (20 ng/ml; R&D Systems), IL-18 (50 ng/ml; R&D Systems), IL-33 (50 ng/ml; R&D Systems), TNF (TNFR; 20ng/ml; R&D systems), PMA (phorbol 12-myristate 13-acetate; 25 ng/ml; Sigma) + Ionomycin (1 µg/ml; Sigma, heated-killled (65 °C for 15 min) S. pneumoniae (three strains: R6 (5 × 10⁶ particles per ml), R8450 (1 × 10⁶ particles per ml) and R11470 (5 × 10⁶ particles per ml); provided by E. Varon), heat-killed S. aureus (1 × 10⁶ particles per ml; InvivoGen); HSV (strain KOS-1; multiplicity of infection, 1), BCG (live Pasteur strain at a multiplicity of infection of 20 per leukocyte) and heat-killed C. albicans (1 × 10⁶ particles per ml; InvivoGen). For stabilization of RNA after stimulation, whole blood was lysed with Tempus solution (from Tempus tubes; Applied Biosystems), at a ratio of 1:3. lysates were stored at ~80 °C until extraction of mRNA.

RNA extraction and processing for microarray analysis. Total RNA was isolated from whole-blood lysates with an Tempus MagMAX-96 Blood RNA isolation kit (Applied Biosystems/Ambion). The quality and quantity of RNA were assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies) and a NanoDrop 1000 (NanoDrop Products, Thermo Fisher Scientific). Samples with RNA integrity values of >6 were retained for further processing. Samples were depleted of globin-encoding mRNA with GLOBINclear (Applied Biosystems/Ambion). RNA that had undergone such depletion was amplified and labeled with an Illumina TotalPrep-96 RNA Amplification Kit (Applied Biosystems/Ambion). Three different batches (Table 1) were treated to this process. Biotin-labeled cRNA was hybridized overnight to Human HT-12 V4 BeadChip arrays (Illumina), which contains 47,231 probes, and arrays were scanned on an Illumina BeadStation 500 (batch 1), iScan (batch 2), or HiScan (batch 3) to generate signal intensity values.

Microarray data analysis. To diminish the potential ‘batch effect’ between sets, background was subtracted from raw signal values, which were extracted with Illumina Beadstudio (version 2) and were processed with the ComBat method (Bayes frameworks for adjusting data)\(^30\). The data were then quantile-normalized and the minimum intensity was set to 10. Only probes ‘called’ as present in at least 10% of the samples (P < 0.01) were retained for downstream analysis (n = 19,152). Transcripts regulated differently upon stimulation were defined on the basis of a minimum change in expression of 1.5-fold (upregulation or downregulation) and a minimum absolute raw intensity difference of 150 relative to expression in the corresponding unstimulated sample. Probes that passed these ‘filters’ in at least 75% of samples from healthy control subjects were considered substantially affected by stimulation and were used in the relevant figures (Figs. 1, 2, 5 and 7). No explicit statistical analysis of differences in expression was performed. The ‘percent of the healthy response’ quantifies the number of probes that were different in the patients than in healthy control subjects. It is calculated as the number of probes affected by stimulation divided by the average number of probes affected by stimulation in samples from healthy control subject, for the same stimulus. Red and white plots (Figs. 1b and 2b) represent the normalized probe counts (number of probes that passed the cutoffs for samples divided by number of probes that passed the cutoffs for 75% of control subjects) for each subject-stimulation pair.

Module construction. Clustering of probes into coexpressed modules. We generated sets of coexpressed probes as described\(^41\), with variations as follows. The data used for clustering included samples from ten healthy control subjects and four patients (two MyD88-deficient patients and two IRAK4-deficient patients) for 19 stimulations (all TLR + IL-1R, TNF and four bacteria) (GEO accession code GSE25742). Background was subtracted from the raw expression data and the results were average-normalized before clustering. Only probes with difference in expression of at least a 1.5-fold and ±100 relative to their expression in an unstimulated sample for all stimulations were clustered. To improve clustering efficiency, the expression values of the selected probes were first converted into ‘trinary’ values (for example, –1, 0, 1) as follows: for cases for which the absolute difference was <0 or the absolute value of the difference in expression (log, ‘fold’ value) was <0.85, the signal was considered to be unchanged and was set as 0. Otherwise, the signal was set as 1 if it was greater than the baseline and as –1 if it was smaller than the baseline. Next, among the samples for each stimulus, a ‘group reduction’ was performed such that for each probe, each group of samples (from healthy subjects and MYD88ΔΔ or IRAK4ΔΔ patients) was replaced by a single value that indicated whether the trinary signal was consistent within 75% of the samples of the group. This yielded three trinary values (one per group) per probe. We clustered the ‘group-reduced’ trinary probe values separately for each stimulus by grouping probes with the same pattern of values, creating as many clusters as necessary for all probes to be included in a cluster. The clusters were then used as input for our module extraction algorithm. The extraction required that the ‘maximal clique’ (set of genes in a network in which every gene has an link to every other gene in a given network) used as input for our module extraction algorithm. The extraction required that the ‘maximal clique’ (set of genes in a network in which every gene has an link to every other gene) in a given network) used as the seed of a module contained at least ten probes. We did not use ‘paralogue’ to expand the module ‘clique’ is too restrictive when noise in experimental data can lead to missing links that should be present; a paralogue consists of the ‘maximum clique’ and all genes with at least some proportion of links to the ‘maximum clique’; a cutoff is set to determine the number of links that must be present for a gene to be included in a ‘paralogue’). When completed,
14 rounds of modules were produced, with the number corresponding to the number of input clustered data sets. A total of 66 modules were obtained, composed of 1,088 transcripts.

**Analysis of module-level data.** Module activity was calculated as the difference between the percent of upregulated probes and that of downregulated probes (Fig. 3). To compare the module response of samples from patients with that of samples from healthy control subjects, we calculated a ‘residual response’ value, which is the per-patient module score divided by the mean of the module score for healthy control subjects for each stimulus. Module scores for any stimulus with an absolute value mean score of <10% for samples from healthy control subjects were considered ‘undetected’, so residual responses were not computed for these modules (gray, Fig. 5 and Supplementary Fig. 3). For cases in which the mean scores for healthy control subjects and an individual patient were of opposite signs, the residual response was set to 0. Circular heat map plots were generated with the Circos visualization tool.

**Functional annotation of modules.** The Acumenta Literature Lab literature-mining process was used to associate each probe in a particular module to terms in PubMed abstracts. Association scores reflecting the strength of the associations were used in heat maps and to calculate ‘cumulative LitLab Scores.’ The terms that showed a strong association with a module were used to create the functional annotation. Gene networks for each module were created with the ‘direct interactions’ algorithm of the GeneGo tool of the MetaCore software suite (version 6.10). Modules were annotated with the network processes that were indicated for each module. The final annotations are a summary of Acumenta Literature Lab and GeneGo MetaCore. In cases for which no terms had a strong association with a module and the gene networks were inconclusive, the modules were left without annotation (14 of 66 modules) (Table 2 and Supplementary Table 1). The association scores obtained from Acumenta Literature Lab were plotted as heat maps (Fig. 4).

49. Levy, O. et al. Selective impairment of TLR-mediated innate immunity in human newborns: neonatal blood plasma reduces monocyte TNF-α induction by bacterial lipopeptides, lipopolysaccharide, and imiquimod, but preserves the response to R-848. *J. Immunol.* **173**, 4627–4634 (2004).

50. Johnson, W.E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).