Intrinsic transcriptional heterogeneity in B cells controls early class switching to IgE

Yee Ling Wu,1 Michael J.T. Stubbington,2 Maria Daly,1 Sarah A. Teichmann,2 and Cristina Rada1

1Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 0QH, England, UK
2The Wellcome Trust Sanger Institute, Cambridge CB10 1SA, England, UK

Noncoding transcripts originating upstream of the immunoglobulin constant region (I transcripts) are required to direct activation-induced deaminase to initiate class switching in B cells. Differential regulation of Ic and Iγ1 transcription in response to interleukin 4 (IL-4), hence class switching to IgE and IgG1, is not fully understood. In this study, we combine novel mouse reporters and single-cell RNA sequencing to reveal the heterogeneity in IL-4–induced I transcription. We identify an early population of cells expressing Ic but not Iγ1 and demonstrate that early Ic transcription leads to switching to IgE and occurs at lower activation levels than Iγ1. Our results reveal how probabilistic transcription with a lower activation threshold for Ic directs the early choice of IgE versus IgG1, a key physiological response against parasitic infestations and a mediator of allergy and asthma.

INTRODUCTION
Heterogeneity of allele activation in single cells is emerging as a general property of transcriptional programs (Trapnell et al., 2014). In many tissues and in organisms as diverse as worms, flies, and humans, single-cell analysis has revealed the prevalence of monoallelic and probabilistic expression of many genes. At the population level, this heterogeneity in the expression pattern of individual cells does not necessarily have functional consequences, as the overall phenotype reflects the average pattern of gene expression for the whole tissue (Little et al., 2013). Nonetheless, this transcriptional noise can be crucial in specific cases and has been implicated as a mechanism that facilitates cell fate choice, dosage compensation, stem cell differentiation, and functional plasticity (Chang et al., 2008; Paul et al., 2015; Reinius and Sandberg, 2015). Although it is still unclear how the heterogeneity is established (Ravarani et al., 2016), its general prevalence has been interpreted as a reflection of the basic molecular processes that govern transcription, an emerging intrinsic property of transcriptional networks (Li and Xie, 2011). Accordingly, genetically identical cells at the same developmental stage are not necessarily functionally equivalent, a property that enables cells to respond differently to the same external cues (Kærn et al., 2005).

An example where diversity in the response is of essential functional importance is class-switch recombination (CSR) at the Ig-constant region loci. CSR generates diverse antibody isotypes with the same specificity and affinity to antigens but crucially with different effector functions (Stavnezer and Schrader, 2014). Among the isotypes, IgE is a powerful mediator for type 2 immune responses, and although protective in helminth and other parasitic infections, IgE can also mediate pathological conditions such as asthma and allergies (Wu and Zarrin, 2014). In contrast to B cells directed toward switching to other isotypes, IgE B cells rarely contribute to the memory compartment or to the long-lived plasma cell pool (Yang et al., 2012). This explains the low levels of circulating IgE found in most individuals in contrast to the high levels of IgG1 in mice (IgG4 in humans) that arise in response to the same T helper 2 cell (Th2 cell) type of stimuli (Gould and Ramadani, 2015). CSR is thus critical in determining the development and terminal differentiation of B cells.

Ig class switching to IgE is a highly regulated process that relies on signals from Th2-type immune responses including the cytokines IL-4 and IL-13, as well as direct interaction with Th cells, leading to the intracellular activation of the NF-κB and STAT6 signaling pathways in the responding B cell (Geha et al., 2003; Xiong et al., 2012b). It also depends on the specific recruitment of activation-induced deaminase (AID) to the DNA-switch region adjacent to the constant ε region (Xue et al., 2006). AID recruitment is linked to the transcription of specific noncoding RNAs (ncRNAs, also called germline transcripts) that originate at promoters upstream of the constant regions of each antibody isotype (I promoters) and proceed through repetitive G:C-rich switch regions (Matthews et al., 2014). Transcription of ncRNAs is critical to allow AID access to DNA (Pefanis et al., 2014) and is mechanistically linked to its targeting, both by the cytoplasmic–dependent selective activation of the I promoters and by the association of the transcription machinery with AID.
targeting (Pavri et al., 2010; Willmann et al., 2012). However, type 2 cytokines induce both IgY1 and Igε ncRNAs in B cells, raising the question as to how the choice between IgG1 and IgE is implemented.

Class switching to IgE is an irreversible differentiation event because it involves deletion of the genes encoding the Cα-, Cδ-, and Cγ-constant regions as well as the Igε promoter. Molecularly, switching to IgE can proceed directly from IgM to IgE or sequentially from IgM to IgG1 and then to IgE (Siebenkotten et al., 1992). The molecular path to IgE switching depends on intrinsic properties of the switch region, such as size and locus architecture (Misaghi et al., 2013), but it is directly linked to the developmental regulation of transcription of the I promoters (Wesemann et al., 2011). In vivo studies in mice have suggested that sequentially switched IgE molecules are of higher affinities and potentially are pathogenic, with the ability to elicit hypersensitive responses. In contrast, directly switched IgE antibodies are of lower affinities and less likely to promote adverse allergic reactions (Xiong et al., 2012a). Low-affinity IgE is produced by extracellular B cells differentiating into plasma cells and early differentiating germinal center (GC) IgE plasma cells (typically before or at 10 d after challenge with parasitic worms such as Nippostrongylus brasiliensis) that peak before the full-blown GCs and IgG1-producing cells that dominate the later phase of the response (He et al., 2013). Thus, the timing or choice of molecular path to IgE switching has important biological consequences.

In vitro activation of mouse B cells via the TLR4 agonist LPS in the presence of IL-4 strongly activates both IgY1 and Igε promoters and induces robust switching to IgG1 and IgE (Mandler et al., 1993). To understand the intrinsic molecular basis for the activation of IgY1 and Igε promoters and the choice between class switching to IgG1 and IgE, we generated knock-in mouse lines with enhanced GFP (eGFP) and tdTomato (tdTom) controlled by the endogenous IgY1 and Igε promoters combined with single-cell transcriptomics and quantitative PCR analyses to follow the temporal activation of IgY1 and Igε in single cells. Our study reveals that heterogeneity in promoter activation is a characteristic of transcription for multiple gene networks even in genetically identical, highly homogeneous, and synchronously activated B cells. In individual B cells, this intrinsic heterogeneity of promoter activation is crucial for control of AID targeting and patterns of IgE switching.

RESULTS
I promoter allelic activation is stochastic, asynchronous, and independent of allelic exclusion

To monitor the dynamics of I transcript regulation in single cells, we introduced a fluorescent reporter gene into the first exon of the γ-1 ncRNA (Iγ1-eGFP line; see the Mice section of Materials and methods and Fig. S1). The resulting transcripts terminate after the eGFP gene and, unlike the endogenous Iγ1 ncRNAs, are not spliced but are efficiently translated (Fig. 1, A and B). The emergence of reporter-positive cells closely mimicked the accumulation of Iγ1 transcripts in control mice, preceded the appearance of surface IgG1-positive cells in LPS–IL-4–stimulated B cells (Fig. 1 B; Fig. S2, A and B; Fig. S3 A), and declined after 72 h as a result of class switching and deletion of the Iγ1-eGFP switch region. B cells homozygous for the modified Iγ1 allele were severely impaired in their ability to switch to IgG1 because of transcription termination before the switch repeat region. As indicated by the kinetics of IgG1 switching in the heterozygous B cells, the presence of a modified allele had no effect on the kinetics of IgG1 switching of the unaltered allele, suggesting there is no cross talk between the two Iγ1 allelic promoters.

To establish the relationship between the activation of the two Iγ1 alleles, we generated a second mouse line, where a tdTom fluorescent protein gene was inserted into the Iγ1 ncRNA (Iγ1-tdTom line) in a C57BL/6j background and crossed with the Iγ1-eGFP BALB/c line (see the Mice section of Materials and methods, Fig. 1 C, Fig. S1 B, Fig. S2, and Fig. S3 B). After 24 h of cytokine induction, ~30% of B cells expressed fluorescence, but only one third of these expressed both reporters (Fig. 1 C), indicating that the activation of the two allelic promoters within the same cell is independent and asynchronous, with most cells showing stable monoallelic transcription. The choice of initial allele activation was stable over several hours and even through cell division, and only a fraction of cells activated the second allele after a further 24 h of culture (Fig. S2, F and G). This suggested that once switched on, promoters are more likely to remain active and that the subsequent activation of the allelic promoter follows a similar probability of activation; thus allelic promoter activation appears to be independent: transcription of one allele is not always accompanied by transcription of the second allele in the same cell.

To determine whether transcription interference from the upstream VDJ promoter or the previous epigenetic history of the allele, such as allelic exclusion (Mostoslavsky et al., 2004; Cedar and Bergman, 2008), could influence the choice of early activation of the Iγ1 promoters, we selected B cells before activation on the basis of the rearrangement status of the locus and followed the emergence of cells expressing the Iγ1 reporters (Fig. 1 D). Heterozygous B cells harboring both the IgMε (C57BL/6j-Iγ1-tdTom) allotype and the IgMε (BALB/c-Iγ1-eGFP) allotype activated the Iγ1 reporters with similar frequencies (Fig. 1 D and Fig. S2). This demonstrated that allelic expression is unrelated to the previous history of the locus because the Iγ1 promoter in the allelically excluded chromosome had the same probability of activation as the productive one.

Paradox in the molecular paths to IgE switching: early direct versus late sequential

In vitro switching to IgE proceeds in most cases via an intermediate that deletes first the Cγ1 region and then joins the Sm region with the Se region (sequential switching;
Siebenkotten et al., 1992; Mandler et al., 1993). As indicated in Fig. 1 B, \(I\gamma 1\) transcripts rapidly accumulate in B cells upon cytokine stimulation, in contrast to the slow accumulation of \(I\varepsilon\) transcripts or the delay in the accumulation of Aicda mRNA (Fig. 2 A). Therefore, it was surprising that early time points after stimulation were enriched for IgE-switched cells with a direct-switch molecular signature (Fig. 2 B). To investigate the kinetics of both \(I\gamma 1\) and \(I\varepsilon\) transcript initiation, we generated a third mouse line with a tdTom fluorescent protein gene inserted into the \(I\varepsilon\) transcript (\(I\varepsilon\)-tdTom line; see the Mice section of Materials and methods, Fig. S1 C, and Fig. S2). Breeding with the \(I\gamma 1\)-eGFP line allowed us to monitor the cytokine-induced activation of both promoters at single-cell resolution (Fig. 2 C and Fig. S3 D). More than 10% of B cells expressed \(I\gamma 1\)-eGFP as early as 24 h after activation. A small percentage of cells expressed both \(I\gamma 1\)-eGFP and \(I\varepsilon\)-tdTom (<1%), and this increased over time as more cells activated the \(I\varepsilon\) promoter. A small population (~1%) expressed only \(I\varepsilon\)-tdTom at the earliest time point, and this single-positive population also grew over time. The emergence of IgE-switched cells followed the activation of the \(I\varepsilon\) promoter (Fig. 2 C and Fig. S2 C). This suggested that the activation of the \(I\varepsilon\) promoters also behaves in a probabilistic fixed way over time, albeit with a lower probability than that of the \(I\gamma 1\) promoters.

The accumulation of \(I\gamma 1\)-eGFP fluorescent cells declined after 72 h, whereas cells expressing the \(I\varepsilon\) reporter or both \(I\varepsilon\) and \(I\gamma 1\) continued to accumulate (Fig. 2 C). This reflected the emergence of cells that had deleted the \(I\gamma 1\)-eGFP region by switching to \(I\varepsilon\) but also demonstrated that the emergence of double-positive fluorescent cells is limited by the fixed frequency/probability of \(I\varepsilon\) activation.

Given the presence of single tdTom-positive cells early after activation, we tested the possibility that early direct switching to IgE could simply reflect the absence of \(I\gamma 1\) transcription, rather than preferential recruitment of AID to the \(I\varepsilon\) locus, by performing single-cell RNA sequencing (scRNAseq) on tdTom-expressing cells 24 h after cytokine activation. After stringently mapping the different Ig transcripts (see the Single-cell analyses section of Materials and methods), we...
determined the frequency of cells coexpressing Iγ1 and Iε transcripts (Fig. 2 D). Consistent with the strong induction of the Iγ1 promoter observed in the Iγ1-eGFP/Iε-tdTom cells, most tdTom-positive cells also expressed Iγ1 transcripts (25/31). Nonetheless, we detected a few tdTom-positive cells without Iγ1 or Iε transcripts and a further few expressing Iε transcripts and no Iγ1 (Fig. 2 D). Only one of the Iγ1-negative/eGFP-negative cells had Iμ–Cy1 transcripts (diagnostic of IgG1 switching). These results suggested that activation of the individual Iε alleles occurs independently of each other and also of Iγ1 and follows a predetermined probabilistic rule. Furthermore, our data strongly suggested that transcription alone is sufficient to explain differential isotype switching in this system.

**Transcription in activated B cells is intrinsically noisy**

Our observations revealed the stochastic nature of the transcription activation of the I promoters in B cells upon cytokine induction. To ascertain whether transcriptional heterogeneity contributed to the choice of isotype switching by determining the early activation of the Iγ1 or Iε ncRNAs in single cells, we used the fluorescent reporters in the Iγ1-eGFP × Iε-tdTom line to identify cells that had entered the class-switching differentiation pathway. The transcription profile of all 157 single cells, sorted on the basis of increased size (blasting), was determined, and tdTom or eGFP fluorescence (see the Single-cell methods section of Materials and methods and Fig. S4) confirmed their activation and high concordance of computationally assigned cell cycle status (Table S3; Scialdone et al., 2015). The transcription profile was consistent with B cell identity (expression of Cd40 or Cd19), cytokine activation (induction of the low-affinity IgE receptor Fcer2a [Acharya et al., 2010] or the transcription regulator nuclear factor IL-3–regulated Nfil3 [Kashiwada et al., 2010]), and the G1 stage of the cell cycle (Fig. 3 A).

General transcriptional heterogeneity results from differences in the cell cycle stage, which directly or indirectly affect the levels of many transcripts. In addition to standard strategies to monitor technical variability (spike-ins), we focused on a highly synchronous cell population. Indeed, analysis of transcriptional variance accounted for by cell
cycle (Buettner et al., 2015) indicated that only a small proportion of the variance (between 1.9 and 2.8%) was caused by cell cycle differences. Despite homogeneity in the cell cycle status of the single cells, we observed additional transcriptional heterogeneity that is unlikely to be attributable to cell cycle differences and indeed indicated general cell-to-cell variation in addition to that observed in the specific case of the I-ncRNA alleles or γ1/ε fluorescent reporters.

To determine the degree of heterogeneity in the expression of individual genes in individual cells, we applied Bayesian Analysis of Single-Cell Sequencing statistics (BASiCs; algorithm as in Vallejos et al., 2015) along with quantification of spike-ins to evaluate the heterogeneity attributable to biological variability (Sigma-Aldrich) in the data (see the Single-cell analyses section of Materials and methods). Of the 13,840 expressed genes, 6,470 were classified as highly variable genes (HVGs), whereas at the opposite end, 887 genes were classified as low variability genes (LVGs; Fig. 3B and Table S5).

Unsurprisingly, the LVG list included highly expressed and housekeeping genes, such as β2-microglobulin (B2m) and β-actin (Actb), highly induced genes such as eElf1a and Hsp90, and some of the B cell–specific highly expressed genes such as the Ig light chain Igκ and the B cell receptor–associated protein Ig-β (Cd79b). The highly variable subset included genes with low expression levels, where biological variability is less readily discernible from the technical noise. However, it also included genes with transcription levels that likely represent functional heterogeneity such as the secretory leukocyte protease inhibitor Slpi and the chemokine receptor Ccr7.

Essentially, each B cell has a unique complement of expressed V genes because of the rearrangement of its antigen receptor. Single-cell transcriptomes revealed the biological heterogeneity exemplified by the IgV genes, which were among the highly variable transcripts identified by the BASiCs algorithm (Fig. 3C). Even highly synchronous and developmentally equivalent B cells showed a high degree of heterogeneity in the expression of master regulator genes, notably transcription factors, which were enriched in the HVG subset (Fig. 3C and Table S4). Again, cell cycle status was not correlated with the expression of the top HVGs or key B cell regulators such as Myc, Bcl6, or Aicda (Fig. S5 A and Fig. 3C). Thus, transcriptional heterogeneity is a general feature of many gene networks in activated B cells.
The $I_e$ promoters have different activation thresholds: $I_e$ responds at lower levels of activation

Given the transcriptional heterogeneity among cells, we used hierarchical clustering to evaluate the differences and identified two groups of cells (Fig. 4 A). Cluster 1 was enriched for reporter-negative cells (Fig. 4 A, gray bars) but showed no differences regarding the cell cycle. Based on the expression of genes known to become up-regulated in B cells 24 h after in vitro cytokine stimulation (Shi et al., 2015), we assigned an activation score to each cell (for details, see the Single-cell analyses section of Materials and methods, Tables S3 and S4, and Fig. S5 B). This revealed that the two clades identified by the cluster analysis reflected differences in the activation scores of the cells (Fig. S5 C). Higher activation scores identified cells with active $I_e$ reporters as well as blastic (increase in size), a direct physical measure of activation (Fig. 4 B and Fig. S5 D). Both $I_{y1}$-eGFP–positive and $I_e$-tdTom–positive cells had a significantly higher mean activation score (0.761, $P < 0.0001$; and 0.666, $P = 0.0007$, respectively; Student’s t test) compared with reporter-negative cells (0.574; Fig. 4 B). This suggested that different thresholds of activation are correlated with the switching on of the $I_y1$ and $I_e$ promoters. Furthermore, the activation score for the $I_e$-tdTom–positive cells was also significantly lower than that of the $I_{y1}$-eGFP–positive cells.

The responsiveness of the individual $I$ promoters was reflected in the activation score of the cells, with a stepwise increase in the mean activation score with each additional activated allele (Fig. 4 C). Even when limiting the analysis to cells with a single active promoter, the differential thresholds of the $I_e$ versus $I_{y1}$ were maintained. Equally, among cells with two promoters on, a significantly lower activation score was observed for cells with an active $I_e$ (Le-tdTom or $I_e$) plus an active $I_{y1}$ compared with cells with both the $I_{y1}$-eGFP and $I_{y1}$ on (Fig. 4 C). Such a difference in activation threshold is corroborated by the subtle differences in the proportion of cells expressing both alleles in the pool of cells that have at least one active allele, which is higher for $I_e$ than for $I_{y1}$ (Fig. 1 B, and Fig. S2, B and C).

To test the differential responsiveness of the I promoters to IL-4, the main driver of IgE switching, we monitored the activation of the reporter at limiting levels of IL-4, whereas the TLR4 signal driving general activation and mitogenic response remained constant. The kinetics of the $I_e$ reporter were unaffected by a 10-fold reduction in the levels of IL-4, whereas the number of $I_{y1}$-eGFP+ cells was significantly reduced (Fig. 4 D, left). Similarly, early after activation (24 h), the number of $I_e$ cells responsive to low levels of IL-4 was not significantly changed, whereas the number of $I_{y1}$-eGFP cells was almost halved. Whereas maximum activation of $I_e$ was achieved at low levels of IL-4, activation of $I_{y1}$ only reached saturation at high IL-4 concentrations (Fig. 4 D, right). Altogether, these data confirm the differential activation thresholds for the $I_e$ and $I_{y1}$ promoters inferred from the scRNAsq analyses.
Early transcription of the I-ncRNAs accelerates the emergence of switched cells

As a population, IgE-switched cells appeared later than IgG1 cells, despite the lower activation score associated with active IgE promoter cells. Therefore, we asked whether the early presence of IgE transcripts had functional consequences. switching was generally enhanced in cells selected for early expression of tdTom or eGFP reporters (Fig. 5, A and B). Furthermore, enrichment of cells with early expression of IgE increased the number of cells in the culture that underwent switching to IgE (greater than twofold increase, normalized to overall switching) compared with unselected cultures. As expected, however, the same did not apply to cells selected on the basis of early IgF expression in relation to IgG1 switching (Fig. 5 C), as most activated cells already express If1 transcripts and switch to IgG1 (Fig. 5 C). The enhancement in overall switching (47–51% compared with 32% in the unsorted cultures) matched the correlation we observed in the single cells between I transcript expression and activation status (Fig. 4, C and D). Thus, early transcription of I-ncRNAs had a functional outcome, favoring the emergence of IgE-producing cells.

Based on the frequency of cells activating the fluorescent reporters (Fig. 1 B, Fig. S2 C, and Fig. S3, A and C), we estimated that under normal in vitro stimulation conditions, 20% of B cells transcribed If1 and between 1–2% were IgE positive, whereas If1–Cy1–positive cells were four times more frequent than IgE–Cy1 positive in the tdTom cells (Fig. 2 D). These ratios underlie the relative emergence of the transcripts measured at the population level (Fig. 2 A) but do not reflect the responsiveness of each I promoter to the main Th2 mediator IL-4 (Mao and Stavnezer, 2001). The single-cell transcriptome analyses also showed that, even at this early time point, at least 6% of the cells expressed Aicda at levels (transcripts per million [TPM] of >50) equivalent to those of key B cell genes such as Myc (median TPM 52.3), Cd74 (median TPM 89.2), or Cd19 (median TPM 90.9; Fig. 3, A and C; and Tables S1 and S2). Indeed, the mean TPM of low-expressed but essential genes such as the transcription regulator Supt5 (TPM > 1 in 73/157 cells; median 4.47) is comparable with the levels detected in cells for Aicda (94/157 cells; median 3.66).

The presence of transcripts that can only arise as a result of the physical deletion of the intervening region, the Iμ-Cε or Iμ-Cγ1 transcripts, although rare (Fig. 5 D), allowed us to monitor the frequency of functional switch outcomes. Data from the two independent scRNAseq analyses, using stringent mapping limited to the unique junctions of the hybrid switch transcripts, identified at least 2/213 cells with switched transcripts (Fig. S5 C). Significantly, these switched cells also had the highest TPM counts for Aicda transcripts (TPM > 100) and reduced counts for Iμ-Cγ1. Stringent detection of switched transcripts is likely to underestimate the true number of switch events, so we also considered cells with detectable TPM counts for the switched transcripts. Here, we found 28/213 cells with Iμ-Cε-Cγ1–switched transcripts (Fig. 5 E). Thus, we concluded that the heterogeneity we observed in the activation of the Iε and If1 together with Aicda expression has functional outcomes at the single-cell level: the emergence of either Iμ-Cε– or Iμ-Cγ1–switched transcripts and the switched Ig.

The number of cells expressing Aicda (45%) but not Ie or If1 supported the independence of Aicda transcription activation, as did the absence of correlation between activation scores and TPM levels in the reporter-positive cells (Fig. 5 F). Although higher levels of Aicda transcripts were not perfectly correlated with the presence of switch transcripts, switched cells were more frequently associated with higher activation scores, as expected from the association between activation score and I promoter activation (Fig. 5 F). This suggested that the activation thresholds for the I promoters are not shared by the Aicda promoter, and therefore, at early time points, the activation threshold of the I-ncRNAs is the main determinant of isotype choice.

**DISCUSSION**

In this study, we attempt to understand how individual B cells integrate external signals to become terminally differentiated as effector IgE-secreting cells. This cell fate is ultimately fixed by the recruitment of programmed DNA damage adjacent to the constant region of Ig genes leading to a genetic switch, with the deletion of the genetic material encoding for other potential isotypes. This provides a robust system to explore how gene activation in response to external stimuli contributes to differentiation at the single-cell level.

As documented in other systems (Shalek et al., 2014), we detected heterogeneity in the transcriptional output of cells upon integration of identical external stimuli, even in highly homogeneous quiescent splenic B cells responding to strong activation signals. Although the activation signal was not limiting, there were clear differences among individual cells, including the number of cells that progressed into mitosis. Even within the population of cells that proceed in a synchronous way toward S phase, the level of activation within
each cell was variable, suggesting that global transcription (the readout that we used as a proxy for activation) behaves in a probabilistic manner, with activation of individual promoters behaving in a stochastic way (Little et al., 2013). This was particularly evident when directly comparing the two alleles of the cytokine-inducible ncRNAs that are associated with class switching to IgG1. In most cells, one allele was stably transcribed for several hours before the second allele also became active. This digital behavior of the promoters is an intrinsic property in the sense that, even at later time points, the number of cells that activated a second allele was fixed. The same activation probability applied to the allelic promoter after cells had undergone cell division and had reached activation levels past the threshold required for IgG1 expression.

Our data also suggest that cells can retain memory of which allele becomes active first. Allelic epigenetic memory controls the early stages of VDJ rearrangement and allelic exclusion (Cedar and Bergman, 2008) and even somatic hypermutation allelic choice in mature B cells (Fraenkel et al., 2007). However, the early activation of the Iγ1 promoters is not affected by the rearrangement history of the cis IgH locus but is rather stochastic (Deng et al., 2014; Reinius and Sandberg, 2015).

Our data show that apart from the stochasticity in the activation of allelic promoters, defined probabilities do operate, as indicated by the reproducible frequencies at which individual cells activate a particular promoter in response to the same stimulus over time (~20% for Iγ1 versus ~1% for Iε). Probabilistic regulation of cytokine production in differentiating Th2 cells has been invoked as a mechanism to fine tune local production of IL-4 and IgE class switching (Guo et al., 2005). It underpins monoallelic expression of cytokines in T cell subsets as a mechanism to control protein levels and contributes to the functional diversity of T cells in response to stimuli (Holländer et al., 1998; Kelly and Locksley, 2000; Calado et al., 2006). Unlike conventional master transcription factors that regulate lineage choice (Rothenberg, 2014) like PAX5 or BCL11B, the choice of pathway in mature B cells can be decided by expression of the I-ncRNAs. At its
The simplest, the choice of isotype is determined by which I promoter becomes active in the presence of functional levels of AID (Fig. 5, D–F). Even at early time points, when Aicda transcription was barely detectable at a population level, we observed cells that had functional levels of AID as well as Iε but no Iγ1 transcripts and preferentially switched to IgE. This explains the higher prevalence of early direct switching to IgE, as cells with active Iε but not Iγ1 promoters are only able to recruit AID to the ε switch region. The presence of detectable direct switching to IgE at late time points suggests that the recruitment of AID is not particularly selective once most cells express both Iγ1 and Iε. Our data also suggest an alternative explanation, in that even late after activation, the probabilistic nature of the I promoter's activation because of the low probability of Iγ1 promoters, reveals by the I reporters.

In fact, the ratio of IgG1 versus IgE cells at all time points could be a direct reflection of the relative frequency of cells with only Iγ1 or Iε promoters active.

Class switching is tightly linked to cell division. However, our single-cell analysis confirms that cell division is not per se a requirement for AID activity (Schrader et al., 2007). The presence of switched transcripts in cells in the G1 phase demonstrates that class switching takes place before full entry into S phase, consistent with the use of the nonhomologous end-joining pathway to resolve the double-stranded DNA breaks generated as a result of AID-induced lesions (Boboila et al., 2010).

Although we could not identify individual genes associated with preferential early activation of either I promoter, we observed a robust difference in activation scores for the three groups of cells selected on the basis of the I promoter activity. Our data clearly demonstrated that, in vitro, the Iε promoter is switched on at a lower activation level (Fig. 4, B and C). This observation corroborates the well-documented higher affinity of the Iε promoter for phospho-Stat6 compared with the Iγ1 promoter (Mao and Stavnezer, 2001), suggesting that Iε promoters are more responsive to lower overall induction levels. This presumably underlies the increased levels of direct switching to IgE observed in immature and transitional B cells (Wesemann et al., 2011). Our results also highlight the probabilistic nature of the I promoter's activation because most cells with active Iγ1 promoters have not induced the Iε promoter despite achieving higher overall activation levels.

Direct evidence for a lower activation threshold for the Iε promoter has important mechanistic implications in vivo, where stimuli are most likely limiting, in particular at the initial stages of an immune response. A lower activation threshold that still allows functional levels of AID would favor the early emergence of IgE-producing plasma cells, where at later stages the amplification of the Th2 response would ensure the production of IgG1 cells that can also populate the GC reaction, become selected, and contribute to memory responses (Fig. 6). Significantly, we also observe heterogeneity in the expression of multiple transcription factors, in particular Bcl6 and Prdm1 (BLIMP-1), which are associated with divergent differentiation pathways (GC vs. plasma cell; Reljic et al., 2000; Shaffer et al., 2002). As in uncommitted precursor cells (Paul et al., 2015), the expression of these factors would presumably, as in the case of early Iε transcripts, contribute to the functional commitment of the cells (Harris et al., 1999; Huang et al., 2013).

Our results reveal for the first time, how in vivo limiting IL-4 signaling would favor the emergence of an early wave of IgE-producing cells (via direct switching) caused by the lower activation requirements of Iε-ncRNAs. They also explain how the number of IgE-producing cells is restricted because of the low probability of Iε promoter activation, which remains constant over time, even at later stages when signaling levels become nonlimiting. Thus, the probabilistic nature of the transcription activation of the I promoters, an intrinsic property of transcription networks, affords individual B cells...
the temporal flexibility to respond to subtle differences in Th2 signals, which is critical in balancing physiological responses and pathological conditions including allergies and asthma.

**MATERIALS AND METHODS**

**Mice**

Iγ1-eGFP and Iε-tdTom mice were generated on a BALB/c background. Iγ1-tdTom mice were generated on a C57BL/6 background.

**Iγ1-eGFP mice.** The Iγ1-eGFP-loxp-Neo-loxp targeting vector was generated using recombineering (Warming et al., 2005). In brief, a 6.4-kb DNA fragment upstream of the Ig heavy chain γ1 switch region encompassing the Iγ promoter was amplified by PCR (primer sequences are listed in Table S6) and cloned into pL2XR, containing the thymidine kinase gene (a gift from A. McKenzie’s group, Medical Research Council, Cambridge, England, UK). The pL2XR-γ1 (BALB/c) plasmid was transformed into the recombineering bacterial strain SW106 by electroporation. To generate the recombineering cassette, the eGFP-loxp-Neo-loxp gene was amplified from plasmid pL452-eGFP (obtained from A. McKenzie’s group; Liu et al., 2003) to add 60-bp and 50-bp mini homology arms corresponding to the sequence immediately downstream of the Iγ exon but before the splice donor. The resulting purified linear fragment Iγ1-eGFP-loxp-Neo-loxp was electroporated into SW106 containing pL2XR-γ1 (BALB/c). Successfully recombined kanamycin resistance colonies were confirmed by sequencing. The pL2XR-Iγ1-eGFP-loxp-Neo-loxp targeting vector was linearized and transfected into an embryonic stem cell (ES cell) line (BALB/c origin). Targeted clones were confirmed by Southern blotting using an 866-bp probe (PCR fragment generated from RP24-258E20 BAC DNA; Fig S1 A). Mice were genotyped using PCR and bred to a Tg(CMV-cre)1Cgn cre-expressing mouse line on a BALB/c background (Schwenk et al., 1995), and deletion of the neo resistance cassette NeoR was confirmed by PCR (Fig. S1 C).

Homozygous Iγ1-eGFP and Iγ1-tdTom or Iε-tdTom mice were bred to produce Iγ1-eGFP × Iγ1-tdTom or Iγ1-eGFP × Iε-tdTom F1 mice, respectively.

All animal work was performed under Medical Research Council guidelines and with the approval of the UK Home Office (PPL 707571).

**Isolation of splenic B cells and in vitro class-switching assay**

Splenic B cells were isolated from mice age 9–16 wk by depletion of CD43+ cells using magnetic microbeads (nos. 130-049–801; Miltenyi Biotec). Purity of the isolated B cells was assessed using anti-B220 antibody (clone RA3-6B2; no. 103212; BioLegend), and 20 ng/ml recombinant mouse IL-4 (no. 404-ML; R&D Systems), 40 μg/ml LPS (L4391 or L3129; Sigma-Aldrich), and 2 ng/ml recombinant mouse B cell–activating factor (no. 2106-BF; R&D Systems). Cells were harvested at 24-h intervals for flow cytometry analysis. Long-term cultures (up to 7 d) were split on day 4 after stimulation.

**Flow cytometry analysis and sorting**

eGFP- and tdtom-positive cells were harvested and stained with anti-B220 antibody (clone RA3-6B2; no. 103212; BioLegend), and dead cells staining positive for Sytox blue stain with anti-B220 antibody (clone RA3-6B2; no. 563097; BD). Then, cells were washed, fixed, and permeabilized with a Fixation/Permeabilization Solution kit (no. 554714; BD). Permeabilized cells were intracellularly stained with BV510-conjugated anti-IgE (clone R35-72; no. 553413; BD). Cells were analyzed on a flow cytometer (LSR II Fortessa; BD). The percentages of cell populations were quantified using FlowJo software (Tree Star). Gating strategies are illustrated in Fig. S2 A.

Cell sorting of eGFP- and tdtom-positive cells was performed on a cell sorter (iCyt Synergy; Sony) with dead cells excluded using Sytox blue stain. For sorting IgM+ and IgMβ cells, freshly isolated splenic B cells were stained with
levels of Gapdh or Pol2a as per the manufacturer's instructions normalizing to the expression of Aicda, I transcripts (Iγ1-Cγ1 and Iε–Cε), and the switched circle transcripts (Iε–Cγ1 and Iε–Cμ) was estimated using Power SYBR green (no. 4367659; Thermo Fisher Scientific) and the absolute quantification method normalizing to the expression levels of Gapdh or Pol2a as per the manufacturer's instructions. Amplicons were cloned into the pCR-blunt vector (no. 44-0302; Thermo Fisher Scientific) to construct standard curves, and the relative abundance of each standard was further calibrated using an internal amplicon for the pCR-blunt vector. All samples were done in duplicate. All primers are listed in Table S6.

Single-cell methods

Index sorting, preparation of cDNA libraries, and sequencing. Cultured B cells were stimulated with IL-4 plus LPS for 24 h, and single cells were deposited into 96-well PCR plates containing lysis buffer using an iCyt Synergy cell sorter (see gating strategy for index sorting in Fig. S4, A and B). The generation of the single-cell libraries followed the Smart-seq2 protocol (Picelli et al., 2014) and then switched to the Fluidigm C1 protocol for tagmentation and purification of the pooled libraries (PN 100-7168; Fluidigm). Pooled libraries of 84 and 180 single cells from two independent cell-sorting experiments were sequenced on a Hi-seq 2500 sequencer (Illumina) at the Cancer Institute facility in Cambridge, England, UK.

Single-cell analyses. A reference transcriptome file was constructed by appending the following sequences to the mouse GR.Cm38 genome reference transcripts: External RNA Controls Consortium (ERCC) spike-in sequences, Iγ1-GFP and Iε-Tom transcripts; the ncRNAs from the Iγ1, Iε, Iκ, Iκ2a, Iκ2b, Iγ3 promoters; and the Iκ–Cμ transcript and the switched transcripts Iγ1-Cγ1, Iε–Cε, Iκ–Cκ, Iκ2a-Cγ1a, Iγ1-Cγ2b, and Iγ3-Cγ3. An index for Bowtie (v1.1.2; Langmead et al., 2009) was created using the bowtie-build indexer. Reads were mapped against this index using Bowtie (v1.1.2) and the options—best -m 1 -X 2000 to ensure that only the best uniquely mapping reads were reported as aligned. Unaligned reads were removed from the resulting output SAM file using samtools view with the following options: -h -f 3. Only reads uniquely mapped to each transcript sequence were counted using a custom Python script. Plots indicating mapping coverage of these reads were generated using bedtools genomc cov to count read depth at each base within the transcripts followed by visualization of these depths using a custom Python script. This stringent mapping restricts detection to reads that map uniquely to one of the transcripts. It is likely that this lowers the detection sensitivity and hence underestimates the numbers of transcript-positive cells, but it ensures high confidence in those transcripts that are detected.

Markers of B cell activation were determined by analyzing existing transcriptomic data (Shi et al., 2015), which provides gene expression values (as fragments per kilobase million) for one sample of resting B cells and one of uninduced B cells 24 h after activation. 626 genes indicative of activation were chosen as those with at least a fourfold increase in expression in the activated sample and with a minimum fragments per kilobase million of 2 in that sample. The activation score for each cell in this study was calculated as the sum of TPM values for the activation marker genes normalized to the maximum activation score from each batch (see Tables S3 and S4).

Cell cycle assignment was performed using Cyclone (Tables S3 and S4; Scialdone et al., 2015). Gene expression values measured in TPM were used as input. This generated G1, G2M, and S-phase scores for each cell. Cells were assigned to a particular cell cycle stage as previously described (Scialdone et al., 2015) such that if either the G1 or G2M score was above 0.5, the cell was assigned to the stage with the highest score. If neither of these scores was above 0.5, the cell was assigned to S phase.

For stringent mapping, a reference transcriptome was constructed containing solely the Iγ1-GFP and Iε-Tom transcripts; the ncRNAs from the Iγ1, Iε, Iκ, Iκ2a, Iκ2b, and Iγ3 promoters; and the Iκ–Cμ transcript and the switched transcripts Iγ1-Cγ1, Iε–Cε, Iκ–Cκ, Iκ2a-Cγ1a, Iγ1-Cγ2b, and Iγ3-Cγ3. An index for Bowtie (v1.1.2; Langmead et al., 2009) was created using the bowtie-build indexer. Reads were mapped against this index using Bowtie (v1.1.2) and the options—best -m 1 -X 2000 to ensure that only the best uniquely mapping reads were reported as aligned. Unaligned reads were removed from the resulting output SAM file using samtools view with the following options: -h -f 3. Only reads uniquely mapped to each transcript sequence were counted using a custom Python script. Plots indicating mapping coverage of these reads were generated using bedtools genomc cov to count read depth at each base within the transcripts followed by visualization of these depths using a custom Python script. This stringent mapping restricts detection to reads that map uniquely to one of the transcripts. It is likely that this lowers the detection sensitivity and hence underestimates the numbers of transcript-positive cells, but it ensures high confidence in those transcripts that are detected.

Markers of B cell activation were determined by analyzing existing transcriptomic data (Shi et al., 2015), which provides gene expression values (as fragments per kilobase million) for one sample of resting B cells and one of uninduced B cells 24 h after activation. 626 genes indicative of activation were chosen as those with at least a fourfold increase in expression in the activated sample and with a minimum fragments per kilobase million of 2 in that sample. The activation score for each cell in this study was calculated as the sum of TPM values for the activation marker genes normalized to the maximum activation score from each batch (see Tables S3 and S4).

Cell cycle assignment was performed using Cyclone (Tables S3 and S4; Scialdone et al., 2015). Gene expression values measured in TPM were used as input. This generated G1, G2M, and S-phase scores for each cell. Cells were assigned to a particular cell cycle stage as previously described (Scialdone et al., 2015) such that if either the G1 or G2M score was above 0.5, the cell was assigned to the stage with the highest score. If neither of these scores was above 0.5, the cell was assigned to S phase.
HVGs and LVGs were determined using BASiCs (data as described in Vallejos et al., 2015). Estimated read counts from Salmon (rounded to integer values) were used as input. HVGs were detected at a variance contribution threshold of 89% leading to an optimized evidence threshold of 0.5145 and giving an estimated false discovery rate of 5.71% and an estimated false negative rate of 5.7%. LVGs were detected at a variance contribution threshold of 40% with an optimized evidence threshold of 0.8115 and estimated false discovery rate and estimated false negative rate both of 2.28%. The value sigma indicates the proportion of the total variability that is due to biological heterogeneity.

Hierarchical clustering of cells was performed and visualized using the clustermap function provided by the Seaborn Python package with the Ward method of clustering. Genes were selected for use in hierarchical clustering if they were expressed (TPM ≥ 1) in at least one cell. Validation of the clusters was performed using the clusterboot function from the fpc R package. The data were resampled by bootstrapping, and Jaccard similarities were calculated between the original clusters and the most similar clusters in the resampled data. Mean Jaccard similarities were calculated from 5,000 iterations. Sequence data described in this manuscript can be accessed at Experiment ArrayExpress (accession no. E-MTAB-4825).

Online supplemental material

Fig. S1 shows a schematic representation of the targeting strategies used to generate I transcript fluorescent reporters in mouse ES cell lines. Fig. S2 shows gating strategies and accumulation kinetics of switched and fluorescent cells in cytokine-stimulated B cell cultures from I promoter reporter mouse lines. Fig. S3 shows examples of individual experiments monitoring the emergence of fluorescent cells in cytokine-stimulated B cell cultures from I promoter reporter mouse lines and the accumulation of class switched IgG1- and IgE-expressing cells. Fig. S4 shows single-cell isolation and scRNAseq quality control. Fig. S5 shows parameters determining transcription heterogeneity in single cells. Tables S1–S6 are available as Excel files. Table S1 shows lists of TPM values for 56 cells from batch 1. Table S2 shows lists of TPM values for 157 cells from batch 2. Table S3 shows cell metadata. Table S4 shows gene lists used for analyses. Table S5 shows heterogeneity analyses using the BASiCs algorithm. Table S6 is a list of primers.

ACKNOWLEDGMENTS

We are grateful to Richard Pannell for help with generating transgenic lines and Ricardo Miraglia, Nick Barry, Fan Zhang, Veronika Romashova, Andrew McKenzie, and members of the McKenzie group for technical advice, critical comments, and helpful suggestions. We thank the staff at the Laboratory of Molecular Biology Biological Services for their excellent work in animal handling and husbandry. Work in the authors’ laboratories is supported by a grant from the Medical Research Council (MC_U105178806 to Y.L. Wu and C. Rada), the Ruth L. Kirschstein National Research Service Award from the National Institute of Allergy and Infectious Diseases, National Institutes of Health (F32AI091311 to Y.L. Wu), and the European Research Council ThNSWITCH grant (260507 to M.J.T. Stubbington and S.A. Teichmann).

REFERENCES

Acharya, M., G. Borland, A.L. Edkins, L.M. Macellan, J. Matheson, B.W. Oznanne, and W. Cushley. 2010. CD23/FcERII: molecular multi-tasking. Clin. Exp. Immunol. 162:12–23. http://dx.doi.org/10.1111/j.1365-2249.2010.04210.x

Bobola, C., C. Yan, D.R. Wesemann, M. Jankovic, J.H. Wang, J. Manis, A. Nussenzwieg, M. Nussenzwieg, and F.W. Alt. 2010. Alternative end-joining catalyzes class switch recombination in the absence of both Ku70 and DNA ligase 4. J. Exp. Med. 207:417–427. http://dx.doi.org/10.1084/jem.20092449

Buettner, F., K.N. Nataraian, F.P. Casale, V. Proserpio, F.J. Theis, S.A. Teichmann, J.C. Marioni, and O. Stegle. 2015. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. Nat. Biotechnol. 33:155–160. http://dx.doi.org/10.1038/nbt.3102

Calado, D.P., T. Panzão, D. Holmberg, and M. Haury. 2006. Stochastic monoallelic expression of IL-10 in T cells. J. Immunol. 177:5358–5364. http://dx.doi.org/10.4049/jimmunol.177.8.5358

Cedar, H., and Y. Bergman. 2008. Choreography of Ig allelic exclusion. Curr. Opin. Immunol. 20:308–317. http://dx.doi.org/10.1016/j.coi.2008.02.002

Chang, H.H., M. Hemberg, M. Barahona, D.E. Ingber, and S. Huang. 2008. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature. 453:544–547. http://dx.doi.org/10.1038/nature06965

Collins, J.T., and W.A. Dumnick. 1993. Germline transcripts of the murine immunoglobulin γ2a gene: structure and induction by IFN-γ. Int. Immunol. 5:885–891. http://dx.doi.org/10.1093/intimm/5.8.885

Deng, Q., D. Ramsköld, B. Reinius, and R. Sandberg. 2014. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science. 343:193–196. http://dx.doi.org/10.1126/science.1245316

Fraenkel, S., R. Mostoslavsky, T.I. Novobrantseva, R. Pelanda, J. Chaudhuri, G. Esposito, S. Jung, F.W. Alt, K. Rajewsky, H. Cedar, and Y. Bergman. 2007. Allelic ‘choice’ governs somatic hypermutation in vivo at the immunoglobulin κ-chain locus. Nat. Immunol. 8:715–722. http://dx.doi.org/10.1038/ni.1476

Gelà, R.S., H.H. Jahara, and S.R. Brodeur. 2003. The regulation of immunoglobulin E class-switch recombination. Nat. Rev. Immunol. 3:721–732. http://dx.doi.org/10.1038/nri1181

Gerondakis, S., 1990. Structure and expression of murine germ-line immunoglobulin epsilon heavy chain transcripts induced by interleukin 4. Proc. Natl. Acad. Sci. U.S.A. 87:1581–1585. http://dx.doi.org/10.1073/pnas.87.4.1581

Gerondakis, S., C. Gaff, D.J. Goodman, and R.J. Grunmot. 1991. Structure and expression of mouse germline immunoglobulin γ3 heavy chain transcripts induced by the mitogen lipopolysaccharide. Immunogenetics. 34:392–400. http://dx.doi.org/10.1007/BF01787490
Shaffer, A.L., K.I. Lin, T.C. Kuo, X. Yu, E.M. Hurt, A. Rosenwald, J.M. Giltinan, L. Yang, H. Zhao, K. Calame, and L.M. Staudt. 2002. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. Immunity. 17:51–62. http://dx.doi.org/10.1016/S1074-7613(02)00335-7

Shalek, A.K., R. Satija, J. Shuga, J.J. Trombetta, D. Gennert, D. Lu, P. Chen, R.S. Gertner, J.T. Gaublomme, N. Yosef, et al. 2014. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature. 510:363–369. http://dx.doi.org/10.1038/nature13437

Shi, W., Y. Liao, S.N. Willis, N. Taubenheim, M. Inouye, D.M. Tarlinton, G.K. Smyth, P.D. Hodgkin, S.L. Nutt, and L.M. Corcoran. 2015. Transcriptional profiling of mouse B cell terminal differentiation defines a signature for antibody-secreting plasma cells. Nat. Immunol. 16:663–673. http://dx.doi.org/10.1038/ni.3154

Siebenkotten, G., C. Esser, M. Wabl, and A. Radbruch. 1992. The murine IgG1/IgE class switch program. Eur. J. Immunol. 22:1827–1834. http://dx.doi.org/10.1002/eji.1830220723

Stavnezer, J., and C.E. Schrader. 2014. IgH chain class switch recombination: mechanism and regulation. J. Immunol. 193:5370–5378. http://dx.doi.org/10.4049/jimmunol.1401849

Vallejos, C.A., J.C. Marioni, and S. Richardson. 2015. BASiCS: Bayesian analysis of single-cell sequencing data. PLoS Comput. Biol. 11:e1004333. http://dx.doi.org/10.1371/journal.pcbi.1004333

Warming, S., N. Costantino, D.L. Court, N.A. Jenkins, and N.G. Copeland. 2005. Simple and highly efficient BAC recombineering using galK selection. Nucleic Acids Res. 33:e36. http://dx.doi.org/10.1093/nar/gni035

Xue, K., C. Rada, and M.S. Neuberger. 2006. The in vivo pattern of AID targeting to immunoglobulin switch regions deduced from mutation spectra in msh2−/−ung−/− mice. J. Exp. Med. 203:2085–2094. http://dx.doi.org/10.1084/jem.20061067

Yang, Z., B.M. Sullivan, and C.D.C. Allen. 2012. Fluorescent in vivo detection reveals that IgE+ B cells are restrained by an intrinsic cell fate predisposition. Immunity. 36:857–872. http://dx.doi.org/10.1016/j.immuni.2012.02.009