

# Biological Insights into TCR $\gamma\delta^+$ and TCR $\alpha\beta^+$ Intraepithelial Lymphocytes Provided by Serial Analysis of Gene Expression (SAGE)

John Shires,<sup>1,2</sup> Efsthios Theodoridis,<sup>1</sup>  
and Adrian C. Hayday<sup>1,2,3</sup>

<sup>1</sup>Peter Gorer Department of Immunobiology  
Guy's, King's, and St. Thomas' Medical School  
King's College  
University of London  
London Bridge  
London SE1 9RT  
United Kingdom

<sup>2</sup>Department of Molecular, Cellular,  
and Developmental Biology  
Yale University  
New Haven, Connecticut 06520

## Summary

Intraepithelial lymphocytes (IELs) are abundant, evolutionarily conserved T cells, commonly enriched in T cell receptor (TCR)  $\gamma\delta$  expression. However, their primary functional potential and constitutive activation state are incompletely understood. To address this, serial analysis of gene expression (SAGE) was applied to murine TCR $\gamma\delta^+$  and TCR $\alpha\beta^+$  intestinal IELs directly *ex vivo*, identifying 15,574 unique transcripts that collectively portray an “activated yet resting,” Th1-skewed, cytolytic, and immunoregulatory phenotype applicable to multiple subsets of gut IELs. Expression of granzymes, Fas ligand, RANTES, prothymosin  $\beta$ 4, junB, RGS1, Btg1, and related molecules is high, whereas expression of conventional cytokines and high-affinity cytokine receptors is low. Differentially expressed genes readily identify heterogeneity among TCR $\alpha\beta^+$  IELs, whereas differences between resident TCR $\gamma\delta^+$  IELs and TCR $\alpha\beta^+$  IELs are less obvious.

## Introduction

T cells are conventionally viewed as thymus-derived cells that recirculate through lymph nodes and spleen via blood and lymph, entering tissues only following activation. By contrast, large numbers of T cells comprise intraepithelial lymphocytes (IELs) that are constitutively associated with epithelia, such as the gut. In mice, and most likely in humans, IELs are the first T cells to develop (Allison and Havran, 1991; McVay et al., 1998). Despite this, they are poorly understood.

IELs display large granules and are cytolytic directly *ex vivo*, provoking the hypothesis that they provide a first line of defense against infected or transformed epithelial cells (Janeway et al., 1988; Hayday and Viney, 2000). Consistent with this, TCR $\alpha\beta^+$  IELs reportedly lyse CMV-infected targets (Muller et al., 2000), while TCR $\gamma\delta^+$  IELs can target heat-stressed epithelial cells (Havran et al., 1991). Nonetheless, to assess the veracity of this hypothesis *in vivo*, much use has been made of the TCR $\delta^{-/-}$  mouse, since many IEL compartments are

highly enriched in  $\gamma\delta$  T cells. Taken collectively, the results of such studies have been difficult to interpret. While some results support an immunoprotective role of TCR $\gamma\delta^+$  IELs (Lepage et al., 1998; E. Ramsburg and A.C.H., unpublished data), others indicate a variety of functions including: anti-inflammatory effects on  $\alpha\beta$  T cells infiltrating the skin or gut (Shiohara et al., 1996; Roberts et al., 1996; M. Girardi, A.C.H., and R. Tigelaar, submitted); pro-inflammatory recruitment of neutrophils into infected or hypoxic lungs (King et al., 1999); the regulation of lumenal IgA (Fujihashi et al., 1996); the promotion of oral tolerance (Ke et al., 1997); and the maintenance of epithelial cell growth (Boismenu and Havran, 1994) and homeostasis (Komano et al., 1995). These various functions are supported by claims that IELs produce a spectrum of relevant effectors including cytokines (Taguchi et al., 1991), chemokines (Mazzucchelli et al., 1996; Boismenu et al., 1996), and keratinocyte growth factor (KGF/ FGFVII; Boismenu and Havran, 1994).

There are at least two, nonmutually exclusive explanations for these various findings. First, IELs may express effector functions in a hierarchical fashion, being primarily cells of a certain phenotype, with secondary effector potentials manifest at lower levels or under particular circumstances. This hierarchy could be defined by a better understanding of the hierarchy of genes that IELs express. Second, different effector potentials might reflect the selective expression of particular molecules by distinct subsets of IELs. A better understanding of the genes that IELs express would permit a better definition of IEL subset heterogeneity.

IELs are primarily subdivided into TCR $\gamma\delta^+$  and TCR $\alpha\beta^+$  cells, and it has been claimed that TCR $\gamma\delta^+$  IELs, but neither TCR $\alpha\beta^+$  IELs nor systemic T cells (either TCR $\gamma\delta^+$  or TCR $\alpha\beta^+$ ), express FGFVII (Boismenu and Havran, 1994). Nonetheless, the functional relatedness of TCR $\gamma\delta^+$  and TCR $\alpha\beta^+$  cells remains unresolved, in part because there has never been a clear “head-to-head” comparison of the genes expressed by the two cell types from the same anatomical location. The expression of CD8 $\alpha$ , CD8 $\beta$ , and CD4 permits the additional subdivision of IELs into the four most abundant subsets: TCR $\gamma\delta^+$  CD4 $^-$  CD8 $\alpha^-$  CD8 $\beta^-$  ( $\gamma\delta$  DN); TCR $\gamma\delta^+$  CD4 $^-$  CD8 $\alpha^+$  CD8 $\beta^-$  ( $\gamma\delta$  CD8 $\alpha\alpha$ ); TCR $\alpha\beta^+$  CD4 $^-$  CD8 $\alpha^+$  CD8 $\beta^-$  ( $\alpha\beta$  CD8 $\alpha\alpha$ ); and TCR $\alpha\beta^+$  CD4 $^-$  CD8 $\alpha^+$  CD8 $\beta^+$  ( $\alpha\beta$  CD8 $\alpha\beta$ ). The functional differences among these subsets are uncertain, although some results indicate that  $\alpha\beta$  CD8 $\alpha\beta$  cells may enter the epithelium following systemic activation, whereas the other three subsets may develop extrathymically and associate with the epithelium directly (Rocha et al., 1994). Indeed, mice lacking classical class I molecules lack  $\alpha\beta$  CD8 $\alpha\beta$  intestinal IELs but not  $\alpha\beta$  CD8 $\alpha\alpha$  cells (Das et al., 2000). Nonetheless, there is uncertainty over whether IELs take up residence as naïve or as antigen-experienced “memory-type” cells (Lin et al., 1999). Again, a better understanding of the genes that IELs express may provide insight into this.

Information on IEL gene expression could be obtained

<sup>3</sup>Correspondence: [adrian.hayday@kcl.ac.uk](mailto:adrian.hayday@kcl.ac.uk)

from microarrays, but because IELs are poorly understood, many highly relevant genes may be underrepresented on arrays. Useful comparisons of TCR $\gamma\delta^+$  IELs and TCR $\alpha\beta^+$  IELs could be obtained from representational difference analysis, although the sparse characterization of IELs would leave differentially expressed sequences without a clear molecular context. Instead, to characterize the genes that “define” IELs and to compare TCR $\gamma\delta^+$  and TCR $\alpha\beta^+$  IELs, serial analysis of gene expression (SAGE) has been applied (Velculescu et al., 1995). By compiling large libraries of expressed sequences de novo, SAGE identifies genes expressed by IELs with no bias toward previously described sequences, and also creates databases that can be interrogated for the expression of any known sequence, including new genes, as they are characterized. Additionally, the development of such large databases facilitates future comparisons of gene expression with other cell types, such as systemic T cells, as those databases are developed.

## Results and Discussion

### SAGE Libraries of TCR $\alpha\beta^+$ and TCR $\gamma\delta^+$ IELs

From outwardly healthy mice, established procedures were used to extract the small intestinal IEL population, which, at the age taken, comprised  $\sim 50\%$  TCR $\alpha\beta^+$  and  $\sim 50\%$  TCR $\gamma\delta^+$  cells. To reduce the chances of cell activation ex vivo, IELs were antibody stained on ice and sorted by high-performance flow cytometry, which combines rapid manipulation with high purity (Figure 1A). RNA was extracted and subjected to SAGE. In essence, cDNA libraries were prepared from each cell type and sequenced, in order to reveal qualitative information on the identity of genes expressed and quantitative information from how many times the same sequence reiterated. To reduce the time and effort required to obtain this information, only short segments (tags) of each cDNA are sequenced, which are demarcated by the Nla3 site closest to the 3' end of each mRNA. Approximately 99% of tags uniquely identify a single gene.

For SAGE to be applicable to  $\leq 5 \times 10^6$  primary cells, several steps were revised including mRNA isolation, modification of oligo-dT biotin with phosphorothioate linkages, cDNA synthesis protocols, and the development of economical, high-throughput methods for cycle sequencing (J.S., E.T., and A.C.H., unpublished data). The result was a collection of  $\sim 3$  Mb of DNA sequence composing two libraries of  $\sim 91,000$  identified tags for each of the TCR $\alpha\beta^+$  and TCR $\gamma\delta^+$  IEL subsets (Figure 2). Collectively, the two libraries identify 15,574 unique sequence tags expressed in IELs. To compensate for sequencing errors, no tag was registered as unique until it occurred in duplicate. This complete database is available to download on a supplementary web site, <http://www.immunobiology.umds.ac.uk/SAGE>, subsequent reference to which will be abbreviated in the text as “sws.”

From the rate of discovery of novel tags as a function of tags recorded, a polynomial algorithm allowed a prediction of 19,614 as the number of unique transcripts expressed by unactivated IELs. For several reasons, this will be an underestimate: (1) a small number of different genes share the same tag; (2) several transcripts from

a single locus will have a single 3' end tag, but will be differentially spliced to encode distinct products; (3) some genes will not register in the SAGE analysis, primarily because they lack an Nla 3 site; and (4) some biologically relevant genes may be expressed at vanishingly low levels. This notwithstanding, the current analysis suggests that the bulk of the constitutive IEL transcriptome is composed by  $\sim 20,000$  genes.

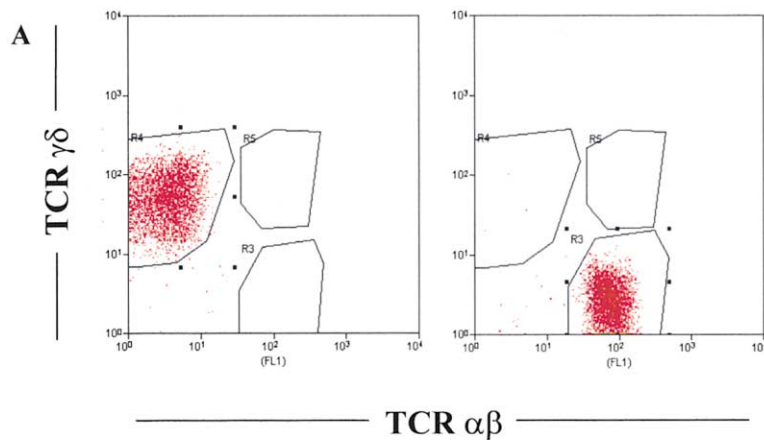
### Gene Abundance Class Correlations

As expected, gene tags fell into abundance classes that operationally were termed: “very high” ( $>800$  tags collectively in the two subsets); “high” ( $800 \geq n > 80$  tags); “moderate” ( $80 \geq n > 20$  tags); “low” ( $20 \geq n \geq 5$  tags); and “rare” ( $n < 5$  tags) (Figure 2). By undertaking 30-cycle RT-PCR using primers for genes from different abundance classes and examining the products on ethidium bromide-stained gels (Figure 1B), it was possible to establish a general correlation between abundance class and conventional measures of gene expression. Additionally, the demonstration that “rare” tags include ones difficult to detect by RT-PCR (Figure 1B) indicates that SAGE has high sensitivity.

Because the library contains  $\leq 75\%$  of expressed transcripts, it is premature to conclude that genes such as *IL10*, for which tag representation = 2 and which are clearly expressed, differ significantly from *IL21*, for which tag representation = 0. Nonetheless, it can be concluded that by comparison to other genes, neither *IL10* nor *IL21* are well expressed in IELs directly ex vivo. Only further expansion of the databases, coupled with RT-PCR validation, will distinguish which of the “zero-tag” genes are rarely expressed and which are unexpressed. Therefore, to avoid the implication that we can make biologically relevant distinctions between genes for which tag numbers are 0–5, all “rare” tags are listed as  $< 5$  (Figure 3).

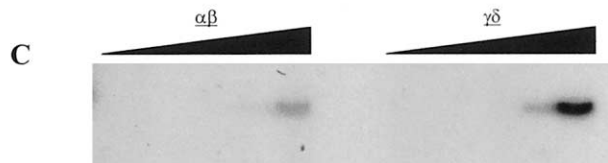
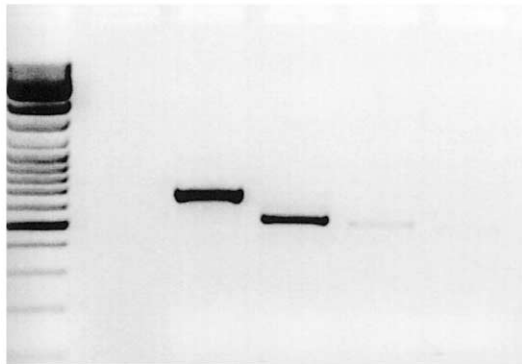
While there will be exceptions, a general relationship between abundance class and protein expression is evident in the common clustering of tags for related protein functions. For example,  $\beta 2$  microglobulin representation (422) was within 1.4-fold of that for MHC class I H2K/D (583; Figure 2); integrins  $\alpha_E$  and  $\beta 7$  that form a heterodimer are expressed at 154 and 106 tags, respectively (Figure 4); 35 genes encoding large and small ribosomal subunit components are clustered between the 23<sup>rd</sup> and 96<sup>th</sup> most abundant mRNAs within 2.5-fold of each other (sws); and tags for 10 of the 11 cyclins and CDKs detected are closely clustered as “low” abundance (sws). Tags also show strong tissue specificity; for example, muscle cofilin is “rare” (3), while nonmuscle cofilin is “high” (210).

Despite the expected representation of essential cellular functions in the “high” abundance class, the most abundant genes include tissue-specific effector molecules (Figure 2). Approximately 3% of all IEL mRNA is devoted to two granzymes and one chemokine. By this criterion, TCR $\gamma\delta^+$  IELs and TCR $\alpha\beta^+$  IELs constitutively resident within the intestine of unchallenged mice are effector T cells, just as plasma cells expressing abundant immunoglobulin RNA are effector cells in the B lineage. Furthermore, the “very high”, “high”, and “moderate” representation of certain effectors, compared



**B**

Gene:	CD8 $\alpha$	LAT	Tyk2	NKG2D
TAG frequency:	185	34	16	2



**Figure 1. IEL Purification and Gene Expression**

(A) Total IELs were isolated from 6-week-old C57BL/6 female mice. Cells were stained with PE- and FITC-conjugated monoclonal antibodies to TCR $\gamma\delta$  and  $\alpha\beta$ , respectively. Cell sorting was done on the MoFlo (Cytomation). Lymphocytes were pregated on light scatter characteristics to eliminate dead cells and pulse width to eliminate multicell aggregates. Sorted cells were reanalyzed to assess purities, which were  $\geq 99.5\%$  in all cases.

(B) Independently generated IEL cDNA was used in an RT-PCR for 30 cycles. Equal volumes of the PCR reactions were run out on an agarose gel containing ethidium bromide and using a 100 bp ladder as a size standard. The gel was then exposed to UV and photographed (negative image is shown). The target gene and the number of times it was detected in the combined SAGE libraries are shown above each lane.

(C) Independently generated IEL cDNAs were normalized against  $\beta$ -actin expression and used in a cycle time course RT-PCR to check for perforin expression. Individual reactions were cycled 18, 20, 22, 24, 26, and 28 cycles and the entire reactions were run out on an agarose gel which was fixed, dried, and exposed to autoradiography film overnight.

with the “low” and “rare” expression of others, establishes a hierarchy of effector potential that portrays IELs as primarily cytolytic cells, with additional potential for highly selective regulation of other cells (Figures 2 and 3).

#### Cytolytic Effector Potential

The highest abundance tag in both subsets is granzyme A (GzmA), while GzmB is the eighth most abundant (Figures 2 and 3). Granzymes are secreted by cytolytic T cells (CTLs), and on entry into target cells, induce apoptosis (Beresford et al., 1999; Shresta et al., 1999). IEL expression of granzymes is selective: GzmC and F were “low” abundance, and GzmD, E, G, and M were “rare”. Granzymes may gain access to target cells via a receptor or by membrane pores created by perforin, the tag for which is “low” abundance ( $\alpha\beta$  3;  $\gamma\delta$  6). Nonetheless, quantitative RT-PCR confirmed perforin expression in TCR $\gamma\delta^+$  and TCR $\alpha\beta^+$  IELs (Figure 1C). A further demonstration of cytolytic potential is the expression of FasL

(35 tags; Figure 3), which induces apoptosis following engagement of Fas on target cells (Nagata and Golstein, 1995).

The tag frequencies for granzymes and FasL were similar irrespective of TCR usage (Figure 3). However, to determine that they are generally representative of IELs, rather than reflecting subset-specific expression, RT-PCR was carried out in the nonsaturating range on a series of dilutions of cDNA made from purified populations of each of the four main IEL subsets:  $\gamma\delta$  DN;  $\gamma\delta$  CD8 $\alpha\alpha$ ;  $\alpha\beta$  CD8 $\alpha\alpha$ ; and  $\alpha\beta$  CD8 $\alpha\beta$  (Figures 5A and 5B). At the same time, gene expression was compared with that in resting and activated splenic CD8 $^+$  T cells that are the systemic cells most closely related to IELs. Examples of the data (Figures 5B and 5C) reveal several points. First, GzmB and FasL expression are representative of all subsets. Second, GzmB and FasL expression combine to distinguish IELs from both resting and activated splenic CD8 $^+$  cells. Third, GzmB expression is

A

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	ATTGTGCAG	1117	1994	3111	Granzyme A
	GTGGCTCAC	940	873	1813	Unknown
	AAGATCTCTG	582	933	1515	RANTES
	GCTGCCCTCC	650	623	1273	Mitochondrial cytochrome c oxidase
	TTGGTGAAGG	521	603	1124	Prothymosin beta 4
	GATACTTGA	482	421	903	Unknown
	AGGCAGACAG	428	410	838	Elongation factor 1-alpha
	CTCTGCTTC	335	477	812	Granzyme B
	CCCTGAGTCC	417	374	791	Beta-actin
	CTAGTCTTTG	408	375	783	Ribosomal protein S29
	TGGTGTCT	411	371	782	Translationally regulated transcript 21 kDa
	AGTCGGGTG	391	369	760	Ribosomal protein L13a
	CACAAACGGT	369	367	736	Ribosomal protein S27
	GGCTCCGGTC	339	335	674	Ribosomal phosphoprotein P1
	GATTGAGAAT	326	257	583	MHC class I
	GGATTGGCT	286	283	569	Ribosomal protein P2 acidic
	AAGAGGCAAG	282	269	551	Ribosomal protein S15a
	ATACTGACAT	273	259	532	Unknown
	GCCTTATGA	246	253	499	Ribosomal protein S24
	GAGCGTTTGG	216	255	471	Peptidylprolyl isomerase A
	TGACCCCGGG	243	216	459	Ubiquitin/60S ribosomal fusion protein
	CATCCCAAC	147	307	454	CD7
	TGGATCAGTC	239	201	440	Ribosomal protein L19
	TTTTCAAAA	187	235	422	Beta-2 microglobulin
	TATGTCAAGC	214	206	420	Ribosomal protein S12
	GATTCGGTA	208	205	413	Ribosomal protein L37
	GCCCCGGAAT	236	176	412	Hexokinase
	CTGTAGGTGA	231	177	408	Ribosomal protein S23
	TCCCTATTAA	166	226	392	Unknown
	GTTGCTGAGA	182	206	388	Ribosomal protein L10
	AAGATCAAGA	197	189	386	Junctional adhesion molecule (JAM)
	ATTCTCCAGT	209	167	376	Ribosomal protein L23
	AAGATTTTT	121	226	347	Rgs1
	AAGACCCCA	82	263	345	TCRg C1, C2
	TAAAGAGCC	179	159	338	Ribosomal protein S26
	TGCTTGTTT	148	189	337	Unknown
	CTAATAAAGC	182	155	337	MNSF beta
	AGCAGTCCCC	163	172	335	Unknown
	CGCTGGTTC	182	150	332	Replication factor C, large subunit
	ATACTGAAGC	166	164	330	Ribosomal protein L13
	ACATCATAGA	172	151	323	Ribosomal protein L12
	GCCAAGTGA	165	157	322	Elongation factor 2
	AAGTGGGAAG	148	172	320	Ribosomal protein L18a
	GGCAAGCCCC	158	159	317	Ribosomal protein L10A
	GCCCCCTTCC	151	161	312	Jun B
	TCCGTACAT	150	159	309	Unknown
	TGGCCCAAT	153	155	308	Ribosomal protein S16
	ATCCGAAAAG	162	146	308	Ribosomal protein L18
	AACAATTTGG	172	136	308	Ribosomal protein L9
	TGGCTCGGTC	150	150	300	Gamma actin
	GCCTTCCAAT	143	153	296	p68 RNA helicase
	CTGCTATCCG	166	130	296	Ribosomal protein L5
	CCTGATCTTT	159	128	287	Laminin receptor
	CAAGGTGACA	162	123	285	Ribosomal protein S2
	GTCTGCTGAT	150	133	283	Guanine binding protein beta-2 rs1
	CAGAACCCAC	146	132	278	Ribosomal protein S18
	GCCTCCAAGG	166	109	275	GAPDH
	AAGAAAATAG	141	132	273	Unknown
	GGGAAGGCGG	129	143	272	Ribosomal protein S3a
	TGTGCCAAGT	140	131	271	Ribosomal protein L34
	ACAGTGCCTA	163	106	269	Actin binding protein 1A (coronin/Taco)
	CCCCAGCCAG	138	122	260	Ribosomal protein S3
	CCTACCAAGA	143	113	256	Ribosomal protein S20
	GATGACACCA	112	139	251	Ribosomal protein S28
	AATCCTGTGG	132	114	246	Ribosomal protein L8
	CCGAACAGA	133	112	245	Ribosomal protein L30
	CCTTAAATCC	127	109	236	Unknown
	CTGAACATCT	133	99	232	Ribosomal acidic phosphoprotein PO
	TCTGTGCACC	113	118	231	Ribosomal protein S11
	GGAAGCCACT	115	111	226	Unknown
	TTGGTGCCTC	103	116	219	Ribosomal protein S14
	TATATTGATT	116	103	219	B-cell translocation gene 1 (anti-proliferative)

Transcript Classes

- effectors/potential effectors
- translational machinery
- housekeeping/metabolic
- MHC/surface receptors
- transcription factors/signalling
- unidentified genes

B

**TAGS Sequenced**

$\alpha\beta$  IELs 90,852

$\gamma\delta$  IELs 91,055

Total Tags 181,907

Unique Tags 15,574

**Classes of Transcript Abundance**

Category	Tag Abundance	Unique Tags in Category	% of Library
very high	800-3111	8	7.4
high	80-799	197	26.9
moderate	20-79	982	22.1
low	5-19	4818	27.6
rare	<5	9569	16.0

Figure 2. The SAGE Database

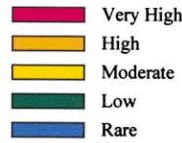
(A) The 75 most abundant tags are shown in descending order. The tags are sorted by the combined abundance within both SAGE libraries (NET). The tag abundance within each library is shown under respective headings. The color coding demonstrates the clustering of certain classes of molecules, particularly the immunological effectors at the highest abundance. The blank cells are unknown genes.

(B) Only tags detected a minimum of two times were considered to represent unique sequences. Unique sequences were operationally broken down into five abundance categories. Tag abundance here is defined as the number of times a given tag was represented in the combined libraries. "% of library" is defined for each abundance class as the (total number of specific tags/total overall tag number)  $\times$  100. This is a close approximation to the percentage of the mRNA that each tag accounts for. Thus, the eight tags in the Very "high" abundance category account for ~7.4% of the total mRNA.

**A**

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	ATTTGTGCAG	1117	1994	3111	Granzyme A
	AAGATCTCTG	582	933	1515	RANTES
	TTGGTGAAGG	521	603	1124	Prothymosin beta 4
	CTCTGCTTCC	335	477	812	Granzyme B
	CTAATAAAGC	182	155	337	MNSF beta
	CATTTTCATA	39	38	77	MIP-1 beta (ScyA4)
	AGCAGATTCT	39	36	75	Cystatin F
	TGCTGTGCAT	32	16	48	Fgf inducible 14
	AACTGACAAA	18	27	45	MIP-1 alpha (ScyA3)
	CTGAACTGG	13	23	36	Fit-3 ligand
	CATACCTCCA	21	14	35	Fas ligand
	GTGGACTACA	22	8	30	Lymphotoxin beta
	GCTGAAACTG	15	14	29	Lymphotoxin (Scyc1)
	AACGCTGCCA	16	12	28	MIF
	GAGACCATCG	5	13	18	IL 17B
	GTCGGTGTGA	8	7	15	Kit ligand
	TCACAGTGT	5	7	12	Lens derived growth factor
	TGTGGGAGCC	7	4	11	GRO1 (GRO alpha)(Scyb1)
	ATTTGCGCT	5	6	11	TGF beta 1
	AGTACTGAGG	5	3	8	Defensin-like cryptdin peptide1,3
	CGTTTCTGGAG	3	4	7	Granzyme F
	AATGAATAAA	3	3	6	Beta Defensin
	TGAAGCCAGG	2	4	6	Granzyme C
	GATGGCATCG	2	4	6	Hepatoma-derived growth factor
	CCCAAGACTC	2	3	5	Insulin-like growth factor 1
	GAACAAGGAA	<5	<5	<5	Fgf-7 (KGF)
	AGTATTGCCA	<5	<5	<5	IFN gamma
	GCACATTCTG	<5	<5	<5	IL 1 beta
	GGTCTTGGGA	<5	<5	<5	IL 10
	AGCCTATTTA	<5	<5	<5	IL 12p35
	AAGAAGCTGT	<5	<5	<5	IL 12p40
	AAATAACTTA	<5	<5	<5	IL 13
	AAATCTTGGG	<5	<5	<5	IL 15
	GCACATCAGG	<5	<5	<5	IL 16
	ATTTTCAGAT	<5	<5	<5	IL 17
	ATTTATAGAG	<5	<5	<5	IL 18
	TAACATCAA	<5	<5	<5	IL 1alpha
	TGTAGGTAGA	<5	<5	<5	IL 2
	GCTTCTTATT	<5	<5	<5	IL 21
	TTCAATTTGTC	<5	<5	<5	IL 3
	CTTTAACTTA	<5	<5	<5	IL 4
	CCAACAGAAA	<5	<5	<5	IL 5
	AAGAACAACCT	<5	<5	<5	IL 6
	TCAACTGCAA	<5	<5	<5	IL 7
	AAGACAGATG	<5	<5	<5	IL 9
	TGATAAGCTA	<5	<5	<5	ScyA1 (TCA3)
	CCAAGCCAGA	<5	<5	<5	ScyA11 (Eotaxin)
	TCTAGGCTGA	<5	<5	<5	ScyA12 (MCP5)
	GTTTCAACAT	<5	<5	<5	ScyA17 (TARC)
	GCCCCCGTGT	<5	<5	<5	ScyA19
	AAACTAAAAA	<5	<5	<5	ScyA2 (MCP1)
	GGGGCGCTTA	<5	<5	<5	ScyA20 (LARC)
	GGCCCTCCC	<5	<5	<5	ScyA21a (Exodus 2)
	GCGCCCTTCC	<5	<5	<5	ScyA21b
	TCCTGAAAGT	<5	<5	<5	ScyA22
	GAAGTGCAGG	<5	<5	<5	ScyA27(Skinkine)
	TATAGCCCTG	<5	<5	<5	ScyA6
	TCACITTAAC	<5	<5	<5	ScyA7 (MCP3, Interkrine)
	GAGATCCTTG	<5	<5	<5	ScyA8 (MCP2)
	ACCATCTTGA	<5	<5	<5	ScyA9 (MIP 1 gamma)
	TGCACACTCC	<5	<5	<5	ScyB11
	CTGTAAACCT	<5	<5	<5	sdf1
	TTCTTTCTAA	<5	<5	<5	Scyb14 (BRAK, MIP2 gamma)
	AGGCAAAGAG	<5	<5	<5	Scyb15 (lungkine)
	TATTTTTACA	<5	<5	<5	Scyb2 (MIP2, GRObeta)
	TTTGAATATG	<5	<5	<5	Scyb5 (LIX)
	TTTTTCCAAT	<5	<5	<5	CXCL7
	GGATTTCAAT	<5	<5	<5	Scyb9 (MIG)
	CTCCACAGAA	<5	<5	<5	ScyD1 (Fractalkine)
	TGAAATGGTC	<5	<5	<5	TECK (ScyA25)
	TTTTATTGGT	<5	<5	<5	TGF beta (GDF1)
	GGGTGTATTA	<5	<5	<5	TGF beta 2
	TATAGATACT	<5	<5	<5	TGF beta 4 (lefty)
	GATGACTTCA	<5	<5	<5	TGF beta 3
	TTTTCTGTGA	<5	<5	<5	TNF alpha
	CTACTTTGTA	<5	<5	<5	TSLP

Abundance Classes



**B**

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	GCCCCCTTCC	151	161	312	Jun B
	TCTGCAAAGG	71	54	125	HMG2
	AGGACAAATA	53	66	119	TAFII250
	TTCTTTCCCG	48	34	82	Muscleblind-like
	TGGAAAGTGA	34	35	69	c-fos
	TGTAAGGAG	26	23	49	ATF-4
	TGTTGGACCT	28	20	48	Interferon regulatory factor 1
	GGGTTTTTAT	16	32	48	Y-box binding protein 1
	GACCTGGGTC	19	20	39	Interferon regulatory factor 7
	TAAAAGTTCT	18	15	33	ELF-1
	CGCGGACTAG	20	13	33	Jun D
	CTCCCGCTGG	10	22	32	CBP interacting transactivator 4
	TTTGTITGTT	16	14	30	c-jun
	GGTGTGTTGT	13	13	26	Chromobox homolog 4 (CBX4)
	TAGCAATCAA	13	11	24	Interferon regulatory factor 2
	GACCTGGTAA	9	10	19	T-BET
	TAAAGTTTTT	10	9	19	TCRa EBF interacting protein 1
	GAAATAAAGA	4	14	18	MAX
	GAACTGTGC	8	8	16	Embigin (MCP-1)
	AATGGCCCTT	0	9	9	X-box binding protein 1
	AAGGACGCCA	1	7	8	GATA 3
	TAATAAGCAC	1	6	7	Max binding protein
	TTCTCAAGTG	3	4	7	NFAT-5
	TGATCATACT	1	6	7	TBP
	GTACCTTTAT	5	1	6	ATF-1
	CTGCAGCGGC	3	3	6	RFLAT-1

**C**

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	AAGAGTTTTT	121	226	347	Rgs1
	GTCTGCTGAT	150	133	283	GNBP $\beta$ 2 related sequence 1
	TCTGAATGTG	78	84	162	LAD/RIBP
	TTGGGATACC	83	62	145	Rac2
	TCTCTGTGTG	67	75	142	Lck
	ATCAACACCG	64	56	120	GNBP $\alpha$ stimulating, extra large
	AAGCAGAAGG	63	54	117	S100a10 (calpactin light chain)
	ACCCACACCG	54	46	100	Calponin 2
	TATCCACGCG	49	48	97	S100A11 (calgizzarin)
	CTGAATGCTC	50	44	94	Calcium-binding protein pp52
	GCTGGCAGCC	47	45	92	Ser/Thr phosphatase type 1 $\alpha$
	ACTCGGAGCC	44	45	89	Calmodulin
	CCTGGCAGAG	42	46	88	DAP10
	GGCTGGGGCT	32	54	86	ZAP-70
	GCACAACCTG	33	40	73	Calmodulin 2
	TGCTGAGAAT	40	26	66	Protein phosphatase 1 $\gamma$
	CATCAGCCTT	24	41	65	PTP (70zshp)
	AGTTAGGAAG	34	23	57	Lck-interacting protein (LIME)
	GAAATGATGA	21	33	54	Prefoldin 5 (c-myc interacting)
	TTGATCATCA	25	24	49	GDP-dissociation inhibitor
	ATTCCTTTATA	22	27	49	Cdc42
	TATTTATTCC	29	19	48	SLAP
	GGGGCTCTGG	15	30	45	DAP12
	CGCGTATGAG	18	27	45	PTP non-receptor type 16
	CTCTCTTTCA	25	17	42	Chat H SH2-containing protein 3
	TCTGTGGAA	17	22	39	Deltex

Figure 3. Most Abundant Immunological Effectors, Transcription Factors, and Signaling Molecules

The layout of the data is as described for Figure 2, with the exception that the color code reflects the abundance class of the tag, as shown in the key and as defined in Figure 2B. The most abundant tags for immunological effectors (A), transcription factors (B), and signaling molecules (C) are shown.

**Categories of Surface Receptors**

**A**  
Antigen  
Receptors

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	AAGACCCCA	82	263	345	TCR $\gamma$ C1, C2
	TTCTAGGACG	87	74	161	TCR $\beta$ C1
	AATACACCAG	83	72	155	CD3 $\epsilon$
	ACATGATTA	90	54	144	TCR $\alpha$
	AGACCGGAAG	53	73	126	CD3 $\delta$
	GAATCTGGGC	25	62	87	CD3 $\zeta$
	GAGTAAAAA	36	41	77	TCR $\beta$ C2
	GCCCAGCCCC	1	69	70	TCR $\delta$
	GATGGGAAC	13	23	36	TCR $\gamma$ chain
	ATTTTGTAT	2	16	18	TCR $\gamma$ C4

**B**  
NK-like  
Receptors

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	GGATTAGCC	21	25	46	LAG-3 (class II receptor)
	GCAGTTCTA	5	13	18	CD94
	TAACACTGTG	8	7	16	2B4
	ACAGAGCAGA	9	1	10	Multiple Ly49 members
	GTTTCTGGG	3	3	6	NKR-P1 (Ly55c)
	TCTACAGCTT	2	3	5	NKR-P1 (Ly55a)
	AAAACAAGGA	<5	<5	<5	Ly49E OR Ly49C
	TCTATAATT	<5	<5	<5	NKG2-A
	AGACAGGGGA	<5	<5	<5	NKG2-D
	AAAATAAAG	<5	<5	<5	NKR-P1 (Ly55b)

**C**  
Cytokine  
Receptors

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	CTTCTGATA	23	42	65	TNFR1
	GTCCTTCTCT	26	28	54	Common gamma chain
	AGCCTGTGGC	19	20	39	IL 21R
	AGTTATTTTA	5	26	31	IL 2R beta
	GTGACTCACA	16	14	30	LIF receptor
	TTGGTGAAG	15	14	29	IFN gamma R
	ACCCTGTGAC	8	11	19	TSLP-R
	TGGGTTTTT	3	8	11	c-kit
	AGAAAAAATA	6	4	10	CNTF-R
	GTCAGTCACA	7	3	10	IL 12R beta1
	AACATTTTGG	5	3	8	IL 10R beta
	GCCTCCTGG	2	5	7	IL 3R alpha
	CTGGAAGCCT	3	4	7	TNFR II
	ACCCTAATCT	6	0	6	IL 10R alpha
	CTCAAAAAATA	3	2	5	IL 12R beta 2
	GAGCTTTATG	4	1	5	IL 4R alpha
	CGCTATTTTC	<5	<5	<5	IL 7R alpha
	GAGAGTACG	<5	<5	<5	IFN alpha/beta R2
	GAACAGTAC	<5	<5	<5	IL 15R alpha
	GCTTCTTCC	<5	<5	<5	IL 2R alpha

**D**  
Costimulatory  
Molecules

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	GTGGAGGGAG	54	53	107	BY55
	ACTCCTGGAC	11	12	23	4-1BB
	GTGTTGGATT	14	6	20	PD-1L
	TTTCTGTGGG	13	2	15	CD82
	TATAATATTT	10	3	13	CTLA-4
	TTGACAGCAG	4	4	8	PD-1
	CTAGCAGCTG	7	0	7	Ox 40
	CTAAAAAATA	2	5	7	CD81
	GACTTTTTAT	1	4	5	ICOS ligand
	ATTTAGAAC	<5	<5	<5	CD28
	ACTTTTCTTT	<5	<5	<5	4-1BB ligand
	GTGCTATTA	<5	<5	<5	ICOS

**E**  
Surface  
and CD  
Antigens

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	GATTGAGAAT	326	257	583	MHC class I
	CATCCCAAC	147	307	454	CD7
	TTTTCAAAA	187	235	422	Beta-2 microglobulin
	CCTGATCTTT	159	128	287	Laminin receptor
	CTTGTGCCTC	88	97	185	CD8 alpha
	TCTGCTTCTC	84	75	159	P- $\alpha$ lectin ligand 1
	CGTTTTGATG	70	84	154	Integrin alpha E
	AAAACCTGTT	69	68	137	CD45
	CTGCCCGAG	53	56	109	Qa-2 (Q7/Q9)
	CAGCCAGCGG	61	45	106	Integrin beta 7
	AGAGTTGCTC	47	40	87	Qa-2 (Q6/Q8)
	GAGGACTGCC	45	30	75	Ly6E
	TGCAATAAAG	34	28	62	CD69
	TCTATGTTTT	28	24	52	CD53
	TATAATACCA	29	10	39	T $\gamma$ 1
	TTTTCATATC	14	13	27	CD48

**F**  
Chemokine  
Receptors

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	CAGTTTCGGA	21	12	33	CCR5
	AAATAAAAT	8	8	16	CCR9
	TACTTGTGTT	8	5	13	SDF1 R
	CCCTATCTTG	6	3	9	CXCR3
	TGAATGAGTG	<5	<5	<5	CXCR4
	TTCTGCTTTT	<5	<5	<5	CCR7

**Associated Signaling Pathways**

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	TCTCTGTGTG	67	75	142	Lck
	GGGTGGGGCT	32	64	86	ZAP-70
	AAATCATCTT	23	15	38	HS1
	TGACATAAAA	26	8	34	LAT
	AGGGTGAGGA	7	21	28	Grb10
	TGGGGATAAC	7	12	19	SLP-76
	AAGCCTTGCT	8	5	13	Grb2
	GATTGCATCC	3	8	11	TRIM
	CTCTTGCAA	3	2	5	TRIM alt 3' end

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	CCTGGCAGAG	42	46	88	DAP10
	GGGGCTCTGG	15	30	45	DAP12
	TCGTTTTTTA	16	6	21	Akt
	GTGCTTGTTA	7	6	13	PI3K p85
	AAGAAGAGTA	<5	<5	<5	SAP (2B4 adaptor)
	TCCTCCCTGT	<5	<5	<5	Syk

CODE	TAG	$\alpha\beta$	$\gamma\delta$	NET	DESCRIPTION
	ACCTGTGCCC	15	19	34	Jak3
	GCATCCTGTT	14	13	27	Stat 3
	GCTTCATAGA	19	6	25	Cish
	GATCCTAAGC	12	12	24	Stat 6
	GATTGAATAA	8	9	17	Stat 1
	CCCAGCCTCA	8	8	16	Stat 2
	GGGAAGCGGA	2	14	16	Tyk2
	TCAAAAATAA	6	5	11	Jak2
	CAATTACCTG	7	4	11	Cish3/SOCS3
	GTGACTGGAA	3	7	10	Jak1
	GGTGTGGTC	3	5	8	Stat 5a
	CTGTGCTCTC	4	4	8	Stat 5b
	TTTACATATT	<5	<5	<5	Cish1/SOCS1

Abundance Classes

- Very High
- High
- Moderate
- Low
- Rare

Figure 4. Expression of Surface Markers and Associated Signaling Pathways

Expression profiles for antigen receptors (A), NK-like receptors (B), cytokine receptors (C), costimulatory molecules (D), surface and CD antigens (E), and chemokine receptors (F). For each of (A)–(C), the left panel shows the abundance for a given group of surface receptors and the right panel shows expression levels for related downstream signaling molecules. The assignment of signaling molecules avoids duplication, although we acknowledge redundancy. For example, PI3Kp85 is used by more than one set of receptors.

slightly higher in  $\gamma\delta$  DN,  $\gamma\delta$  CD8 $\alpha\alpha$ , and  $\alpha\beta$  CD8 $\alpha\alpha$  cells than it is in  $\alpha\beta$  CD8 $\alpha\beta$  cells, which show the greatest resemblance to activated splenic CD8 $^+$  cells.

### The Regulatory Potential of IELs

The “very high”, “high”, and “moderate” abundance classes contain tags for particular immunoregulatory molecules. The third most common tag is for the inflammatory chemokine, RANTES ( $\alpha\beta$  582;  $\gamma\delta$  933) (Figures 2 and 3). Although RANTES expression by human intestinal and murine epidermal IELs has been observed (Mazzucchelli et al., 1996; Boismenu et al., 1996), previous reports did not note its “very high” abundance. The “very high” expression of GzmB and RANTES may reflect the fact that both are stored in the same cytolytic granules. Again, RT-PCR analysis showed the RANTES tag frequency to be generally representative of IELs, although as for GzmB, RANTES expression was somewhat less in  $\alpha\beta$  CD8 $\alpha\beta$  cells that, in this case, resembled resting CD8 $^+$  splenocytes (Figure 5B).

IELs also express MIP1 $\alpha$  ( $\alpha\beta$  18;  $\gamma\delta$  27) and MIP1 $\beta$  ( $\alpha\beta$  39;  $\gamma\delta$  38) that, like RANTES, are Th1-associated chemokines that bind to CCR5 (Figure 3). The high constitutive levels of CCR5 ligand expression support the hypothesis that IELs may in humans antagonize CCR5-mediated cell entry of HIV across mucosal epithelia. Physiologically, it suggests that IELs may play an important role in the recruitment into epithelia of CCR5 $^+$  cells including immature dendritic cells (DC), effector/memory T cells, and NK cells. IELs also express lymphotactin ( $\alpha\beta$  15;  $\gamma\delta$  14) and macrophage migration inhibitory factor (MIF;  $\alpha\beta$  16;  $\gamma\delta$  12), chemokines that regulate CD8 $^+$  CTL activity.

By contrast, chemokines expressed in the “low” or “rare” classes include Scya 19 and 21, which bind CCR7 on naïve T cells and activated DC; Scya 20, which promotes DC localization to mucosal Peyer’s patches; Syca 11 (eotaxin), Syca 1, and other Th2 cell-associated chemokines; and CXCL chemokines such as Scyb 1–3 (Gro  $\alpha$ ,  $\beta$ , and  $\gamma$ ) and Scyb6, which attract neutrophils. Furthermore, the fifth most abundant tag is for prothymosin  $\beta$ 4 (pT $\beta$ 4) ( $\alpha\beta$  521;  $\gamma\delta$  603) that, in addition to encoding a ubiquitously expressed actin binding protein that regulates the equilibrium between globular and filamentous actin (Carlier and Pantaloni, 1994), is a potent inhibitor of neutrophil infiltration (Young et al., 1999). An immunoregulatory role of pT $\beta$ 4 may explain the existence of a lymphoid-specific alternatively spliced transcript (Rudin et al., 1990). Once more, lymphoid pT $\beta$ 4 expression is representative of all IEL subsets (Figure 5B), although as for RANTES, expression was less in  $\alpha\beta$  CD8 $\alpha\beta$  cells, which were again more similar to resting CD8 $^+$  splenic T cells.

Epithelial cells regularly produce IL8 and other neutrophil attractants. Nonetheless, neutrophil infiltration of epithelia is rare, and pathognomonic for conditions such as psoriasis. In part, control over inflammation may be attributable to TGF $\beta$ , which is broadly expressed in the mucosal epithelium. However, the constitutive production of pT $\beta$ 4, coupled with the “low”/“rare” expression of neutrophil-recruiting chemokines, suggests an additional mechanism whereby neutrophil infiltration may be suppressed by IELs. Consistent with this, neutrophil

accumulation is a common pathological feature of *TCR $\delta^{-/-}$*  mice (D’Souza et al., 1997; Fu et al., 1994; M. Girardi, A.C.H., and R.E. Tigelaar, submitted). Furthermore, topical application of pT $\beta$ 4 to rats following full-thickness epithelial wounding enhanced healing by 42%–61% over a 4–7 day period. This improvement featured increased keratinocyte migration and collagen deposition (Malinda et al., 1999). Thus, the “very high” tag frequency of pT $\beta$ 4 may in part explain the reported role for IELs in epithelial maintenance. Consistent with a requirement for prior cell activation (Boismenu and Havarin, 1994), no tags for FGFVII were detected (Figure 3).

That highly conserved molecules such as pT $\beta$ 4 have both ubiquitous cytoplasmic functions and specialized immunoregulatory roles has precedent in indoleamine-2,3-dioxygenase, which suppresses T cell activation by depletion of tryptophan (Munn et al., 1998). Another candidate for such “dual activity” is a broadly expressed ubiquitin-like polypeptide, MNSF $\beta$  (monoclonal nonspecific suppressor factor), found at “high” abundance ( $\alpha\beta$  182;  $\gamma\delta$  155) (Figures 2 and 3). MNSF released by activated T cells inhibits lymphocyte proliferation and IL4 production, particularly in pro-inflammatory contexts (Nakamura et al., 1995).

By contrast, most conventional cytokine genes, including *IFN $\gamma$* , *IL2*, *IL4*, and *IL10* are expressed in the “low”/“rare” abundance classes (Figure 3). That this is a general property of TCR $\gamma\delta^+$  and TCR $\alpha\beta^+$  IELs was confirmed by intracellular staining of monensin-treated cells from uninfected mice, examined directly or after 6 hr of activation with PMA and ionomycin. Only negligible levels of intracellular cytokines or IL4 were detected in any IEL subset (E. Ramsburg and A.C.H., unpublished data). Instead, the most highly sampled cytokine genes were *lymphotoxin*  $\beta$  ( $\alpha\beta$  22;  $\gamma\delta$  8) and *TGF $\beta$ -1* ( $\alpha\beta$  5;  $\gamma\delta$  6) (which may contribute to the cytolytic/immunoregulatory capacities of IELs), and *IL17B* ( $\alpha\beta$  5;  $\gamma\delta$  13) (Figure 3), a recently discovered cytokine reported to be expressed in small intestine and that also may regulate neutrophil influx, possibly indirectly (Li et al., 2000; Shi et al., 2000). The identification of the IL17B tag exemplifies the capacity of SAGE to report on the expression of novel genes as they are discovered. Heretofore, the gut-associated source of IL17B had not been clarified.

In sum, SAGE has identified an effector hierarchy in which cytolytic mediators, selective chemokines, and nonspecific immunoregulators are constitutively highly expressed and thereby capable of rapid mobilization, whereas significant expression of most conventional cytokines and growth factors will presumably require additional cell activation and hence form part of a more delayed response. Where examined by flow cytometry and RT-PCR, these conclusions apply to all major IEL subsets, although the  $\alpha\beta$  CD8 $\alpha\beta$  subset consistently behaves slightly differently than the others.

### The Activation State of IELs

The overtly high expression of particular effectors (Figure 3; Figure 5) suggests that  $\alpha\beta^+$  and  $\gamma\delta^+$  IELs are strongly activated, consistent with which expression of CD69 is “moderate” ( $\alpha\beta$  34;  $\gamma\delta$  28) (Figure 4E). Likewise, the “moderate” expression of CCR5 ( $\alpha\beta$  21;  $\gamma\delta$  12) is similar to that on activated Th1 cells, consistent with

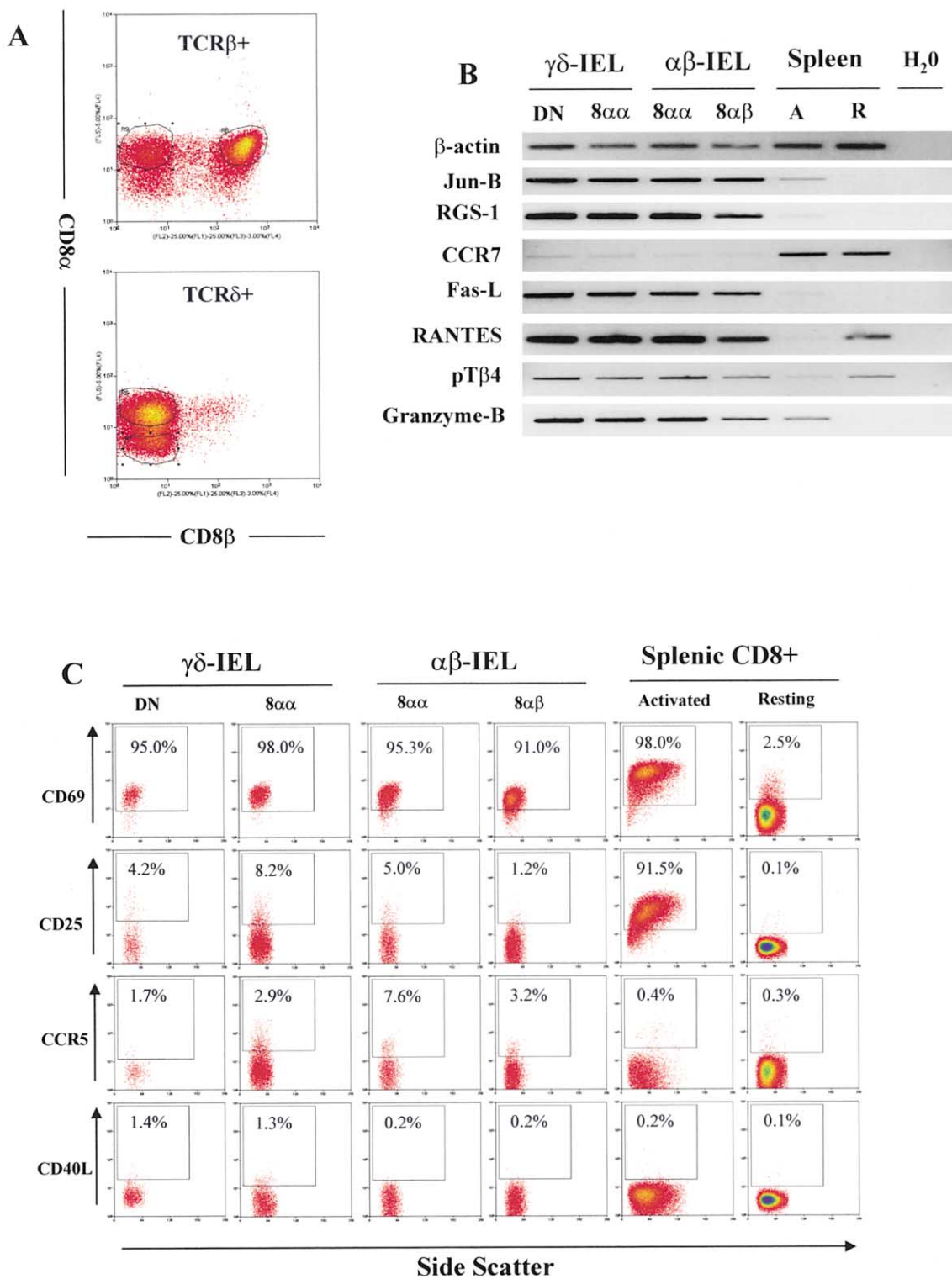


Figure 5. Comparison of IEL Subsets with Activated and Resting Splenic CD8 T Cells

(A) IEL subsets were sorted to generate cDNAs for quantitative RT-PCR. IEL preparations were stained with a five-color combination of mAbs to TCR $\alpha\beta$ , TCR $\gamma\delta$ , CD4, CD8 $\alpha$ , and CD8 $\beta$ . CD4<sup>+</sup> cells were gated out as a first step. TCR $\alpha\beta$ <sup>+</sup> and  $\gamma\delta$ <sup>+</sup> cells were separately gated and each analyzed for the expression of CD8 $\alpha$  and CD8 $\beta$ . Four-way sorting was used to purify the populations circled in the two dot plots and used in (B) and (C).

(B) RT-PCR was performed on cDNAs generated from the four sorted IEL populations. Expression of  $\beta$ -actin was used to normalize all cDNAs. For comparison, splenic TCR $\alpha\beta$ <sup>+</sup>CD8 $\beta$ <sup>+</sup> cells were sorted directly ex vivo, or after a 48 hr activation of bulk splenocytes with plate-bound anti-CD3 $\epsilon$ ; and anti-CD28. Abbreviations for the IEL lanes are as in the text; for the splenic populations: A, activated; R, resting.



which CCR7 (expressed on naïve or recently activated T cells) was “rare” (Figure 4F; Figure 5B).

Conversely, a resting state is implied by the rarity of tags for conventional cytokines and for high-affinity cytokine receptor chains such as IL15R $\alpha$ , IL7R $\alpha$ , and IL2R $\alpha$  (CD25), which ordinarily are strongly responsive to activation (Figure 4C). Likewise, there was “low” expression of IL12R $\beta$ 1 (10 tags) and IL12R $\beta$ 2 (5 tags), which are ordinarily induced by signals from the IFN- $\gamma$ R (Figure 4C).

To examine whether this “activated yet resting” phenotype was generally representative of IELs or whether it reflects a mixture of activated cells and resting cells, each of the four IEL subsets was stained for CD69, CD25, and CCR5 (Figure 5C). Greater than 90% of cells in each subset were CD69<sup>+</sup> CD25<sup>-</sup> and small, as judged by side scatter profiles. Surface CCR5 was present on 3%–8% of TCR $\alpha\beta$ <sup>+</sup> IELs and 1.5%–3% of TCR $\gamma\delta$ <sup>+</sup> IELs, consistent with tag frequencies ( $\alpha\beta$  21;  $\gamma\delta$  12). (Note, however, that surface CCR5 expression levels are probably underestimated, since CCR5 is rapidly downmodulated on exposure to its ligands which are highly expressed by IELs [M. Marsh, personal communication]). In sum, these staining profiles provided a further, general distinction of IELs from both resting and activated splenic CD8<sup>+</sup> T cells (Figure 5C).

#### Inside IELs

SAGE data on signaling molecules and transcription factors provided additional evidence for a signatory IEL phenotype. For example, the most abundant identified transcription factor tag is junB ( $\alpha\beta$  151;  $\gamma\delta$  161) (Figure 3B). This expression is representative of all subsets, and is higher than in resting or activated splenic CD8<sup>+</sup> cells (Figure 5B). The role of junB in T cells has remained uncertain, with recent data showing it regulates cytokine expression in Th2 cells (Li et al., 1999). However, because IELs are biased toward Th1, not Th2 (see Transcription Factors, Cell Differentiation, and a “Th1 Phenotype”), “high” junB expression more likely reflects its reported role in maintaining highly differentiated cells in a resting state. Thus, JunB inactivation in myeloid cells leads to leukaemias with enhanced cytokine-mediated proliferation (Passegué et al., 2001). Moreover, JunB and c-jun were shown to antagonistically regulate KGF and GM-CSF transcription in dermal fibroblasts (Szabowski et al., 2000), with junB suppressing expression. Therefore, the 10-fold excess of JunB over c-jun ( $\alpha\beta$  16;  $\gamma\delta$  14) correlates well with the lack of conventional cytokine and KGF expression in IELs (Figure 3).

Likewise, the most abundantly expressed identified signaling molecule is RGS1 (regulator of G protein signaling) ( $\alpha\beta$  121;  $\gamma\delta$  226) (Figure 3C). Again, “high” expression is representative of all subsets (Figure 5B) and is greater than that in either activated or resting splenic CD8<sup>+</sup> cells (Figure 5B). RGS molecules are encoded by  $\geq$ 20 genes, and suppress GTP-dependent signaling downstream of G protein-coupled receptors such as

chemokine receptors (Reif and Cyster, 2000). Since IELs show “moderate” CCR5 expression, they are potentially responsive to the large amounts of CCR5 ligands, RANTES, MIP1 $\alpha$ , and Mip1 $\beta$  that they make. Interestingly, RGS1 was noted to be active in B cells, whereas activated systemic T cells were characterized by RGS2 and RGS16 that in IELs occur at tags ( $\alpha\beta$  11;  $\gamma\delta$  19) and ( $\alpha\beta$  6;  $\gamma\delta$  12), respectively (Beadling et al., 1999; Oliveira-Dos-Santos et al., 2000). These data, combined with those in Figure 5B, suggest that “high” RGS1 expression is a further feature distinguishing IELs from systemic T cells.

Several other tags indicate a constitutive resting state of IELs. For example, there is “high” expression of B cell translocation gene (BTG) 1 ( $\alpha\beta$  116;  $\gamma\delta$  103), and “moderate” expression of Schlafen 2 ( $\alpha\beta$  27;  $\gamma\delta$  28) and BTG2 ( $\alpha\beta$  20;  $\gamma\delta$  26) (sws), all of which are downregulated on lymphocyte activation and can arrest the cell cycle in G0/G1 (Schwarz et al., 1998). IELs also show “high” frequency for ornithine decarboxylase antizyme ( $\alpha\beta$  91;  $\gamma\delta$  107) (sws), a regulated gene encoding an inhibitor of DNA synthesis. Additionally, it was recently reported that interference with cytokine receptor signaling is achieved by CD45 via dephosphorylation of jak kinases (Irie-Sasaki et al., 2001). CD45 is the fifth most abundantly expressed CD tag ( $\alpha\beta$  69;  $\gamma\delta$  68) (Figure 4E).

Another aspect of IELs revealed by SAGE is that they express only “low”/“rare” levels of well-characterized antiapoptotic genes, such as *Bcl2* ( $\alpha\beta$  7;  $\gamma\delta$  11) (sws). Indeed, junB has been shown to suppress *Bcl2* and *Bcl<sub>L</sub>*. Proapoptotic effectors (caspase 3 [35 tags]; caspase 8 [32 tags]) and regulators (Bax [27 tags]; BAK [11 tags]; Bad [7 tags]; and BID [6 tags]) were equally well represented, possibly explaining the high sensitivity of IELs to apoptosis (Viney et al., 1990).

#### IEL Activation

Because the data depict IELs as cells expressing high levels of certain effectors but with a potential for further activation, the question arises as to how IELs are activated in situ. Therefore, the databases were examined for surface receptor gene profiles. Among the most highly expressed such genes were those for TCR $\alpha\beta$ , TCR $\gamma\delta$ , and CD3. Indeed, the least expressed of these were CD3 $\zeta$  ( $\alpha\beta$  13;  $\gamma\delta$  23) and its recently identified associated molecule, TRIM (16 tags; Figure 4A). Also present are tags for other TCR-associated signaling molecules, such as Lck, LAT, Fyn, Fyb, ZAP70, SLP76, HS1, and NFAT-C (Figure 4; sws), some of which are “high”, such as LAD ( $\alpha\beta$  78;  $\alpha\beta$  84), lck ( $\alpha\beta$  67;  $\gamma\delta$  75), and ZAP70 ( $\alpha\beta$  32;  $\gamma\delta$  54) (Figure 3C).

By contrast, there is less compelling evidence for the primary control of IEL activation via NK receptors (Halary et al., 1997). Although the NK receptor-associated signaling molecules DAP10 ( $\alpha\beta$  42;  $\gamma\delta$  46) and DAP12 ( $\alpha\beta$  15;  $\gamma\delta$  30) are expressed, Ly49, CD94, 2B4, and NKG2A expression are “low”. IELs also lack significant expression of Toll-like receptors (TLRs) 1, 2, 4, 5, 6, 7, and 9

(C) IELs were isolated and stained with the indicated mAbs in conjunction with CD8 $\alpha$ , CD8 $\beta$ , and TCR $\alpha\beta$  or TCR $\gamma\delta$ . The IEL subsets shown in the various panels were gated similarly to those in (A). Side scatter is used on the abscissa to indicate differences in cell size between IELs and splenic cells.

and CD14, which delivers microbial antigens to TLRs (a lack of 3' end data precludes the assessment of TLRs 3 and 8). Also germane to bacterial recognition and the host response is the observation that  $\beta$ -defensin and cryptidin expression are "rare" (sws).

Activation of T cells via the TCR requires metabolic upregulation provided via the PI3-kinase-Grb2-Akt pathway by costimulators such as CD28, CD40L, and ICOS. The conspicuous underrepresentation of these (Figure 4D) is a general phenotype of IELs (Figure 5C). By contrast, IELs express 4-1BB ( $\alpha\beta$  11;  $\gamma\delta$  12) (Figure 4D). Since it was reported that preactivated systemic cytolytic cells can be costimulated by the engagement of 4-1BB (Shuford et al., 1997; DeBenedette et al., 1997), future studies should examine the contribution of this molecule to IEL activation.

A further candidate for costimulation is BY55 ( $\alpha\beta$  54;  $\gamma\delta$  53) (Figure 4D), an NKR-like molecule expressed by human IELs (Anumanthan et al., 1998) and functionally implicated in costimulation of CD28<sup>-</sup> cells (Agrawal et al., 1999). Nonetheless, since BY55 is GPI linked, its signaling capacity presumably requires association with a transmembrane protein. Ligands for costimulators can also downregulate T cells. One such—CTLA4—is expressed at "low" levels ( $\alpha\beta$  10;  $\gamma\delta$  3), as is the PD-PD1 ligand-receptor pairing recently implicated in T cell inhibition (Freeman et al., 2000).

The TCR-mediated activation of IELs is usually thought to be provoked by epithelial cells. Consistent with such interactions, IELs show "high" expression of integrins  $\alpha_E$  ( $\alpha\beta$  70;  $\gamma\delta$  84) and  $\beta_7$  ( $\alpha\beta$  61;  $\gamma\delta$  45) that engage E-cadherin on enterocytes. Activated enterocytes secrete IL15. However, while there is "moderate" constitutive expression in both IEL subsets of the common  $\gamma$  chain  $\gamma_C$  (54 tags) and the IL2R $\beta$  chain (31 tags; indicating that IELs could potentially respond to one or more of IL2, IL4, IL7, IL9, and IL15), the high-affinity IL15R $\alpha$  tag (like other ILR $\alpha$  tags) is only "rare". Most likely, TCR-mediated activation induces IL15R $\alpha$  (as well as other ILR $\alpha$  chain expression), facilitating IL15 responses via JAKs and STATs that are expressed (Figure 4). This pathway may upregulate NKG2d, which would provide further costimulation via DAP10 (Groh et al., 2001; Lanier, 2001). Additionally, tags in both subsets imply IEL responsiveness via TNF and IFN $\gamma$  receptors (Figure 4C). Interestingly, the recently characterized IL21R is the third most abundant cytokine receptor chain (Figure 4). IL21 most resembles IL2 and IL15 (Parrish-Novak et al., 2000), and the possibility that this cytokine may be important for IEL homeostasis and/or activation is under study.

Activated enterocytes also upregulate MHC class II, in which regard it is interesting that IELs show "moderate" expression of LAG-3 ( $\alpha\beta$  21;  $\gamma\delta$  25) (Figure 4B), a CD4-related molecule that binds class II MHC (Avicce et al., 1999). Possibly, LAG-3 mediates interaction of IELs with other MHC class II<sup>+</sup> cells such as local DC, which might then be functionally regulated by IEL products such as RANTES and Flt3 ligand ( $\alpha\beta$  13;  $\gamma\delta$  23).

#### The Consequences of IEL Activation

IELs express genes to facilitate TCR- and cytokine-induced signal transduction. One pathway for which

each component member is clearly represented is the vav-regulated rhoA/Rac2-cdc42-WASP axis. Following activation by either the TCR or by integrins, the pathway targets the cytoskeleton, and has consequent effects on cell-cell adhesion. Thus, its activity may be important for controlling exocytosis of secretory granules containing granzymes and RANTES, and/or for regulating interactions of IELs within epithelia. The channeling of signal transduction through to the cytoskeleton is also suggested by "high" tags for cofilin 1 ( $\alpha\beta$  115;  $\gamma\delta$  95) (sws), a protein that is dephosphorylated after T cell costimulation, and that then destabilizes filamentous actin, possibly in concert with pT $\beta$ 4.

#### Transcription Factors, Cell Differentiation, and a "Th1 Phenotype"

The transcription factor profile is broad, including genes implicated downstream of the TCR (e.g., *c-fos* and *c-jun*), and in TCR gene regulation (e.g., *ELF-1*; Figure 3B; sws). With the exception of junB, most are "moderate"/"low" abundance. Nonetheless, there is a Th1 bias. Thus, T-bet (Szabo et al., 2000) is present in both subsets ( $\alpha\beta$  9;  $\gamma\delta$  10), whereas tags for Th2-associated c-MAF/MAF-B (Kim et al., 1999) are "rare". Although GATA3 (also associated with Th2 differentiation; Kuo and Leiden, 1999) is expressed at "rare" to "low" abundance (with a bias toward TCR $\gamma\delta^+$  IELs), it is known to be expressed in Th1-type cells, albeit at approximately five times lower levels than in Th2 cells (Li et al., 2000). Conversely, the same authors showed Rac2 to be overexpressed in Th1 cells, and this tag is present at "high" levels ( $\alpha\beta$  83;  $\gamma\delta$  62) (Figure 3C). Also in support of a Th1 bias are the expression of RANTES, lymphotactin, LAG3, CCR5, and CXCR3 (Figure 4F), contrasted with the rarity of CCR3, 4, and 8 (sws), T1/ST2 (the signatory Th2 marker in mouse; Lohning et al., 1998), and IL10R $\alpha$  (Figure 4C).

The Notch pathway has been implicated in lineage determination and in cell survival induced by cell-cell contact (Deftos et al., 1998). IELs show "low"/"rare" representation of tags for Notch 1, 3, or 4 and for Notch ligands, Delta-like and jagged 1 or 2 (sws). Interestingly, however, a deltex 1 homolog is expressed in both subsets ( $\alpha\beta$  17;  $\gamma\delta$  22), implying that signaling via a Notch pathway may have recently occurred.

#### Genes Specific to TCR $\gamma\delta^+$ IELs

The majority of identified tags are expressed at similar levels irrespective of TCR type. This is equally true for genes underlying cell metabolism and for genes encoding T cell-associated signaling and effector molecules, expressed across the spectrum of abundance classes. The most highly differentially expressed gene is *TCR $\delta$*  ( $\alpha\beta$  1;  $\gamma\delta$  69) (Figure 4A). This skewing will be at the "high end" of differential gene expression because the *TCR $\delta$*  gene is deleted in most  $\alpha\beta$  T cells. *TCR $\gamma$*  expression is also skewed toward TCR $\gamma\delta^+$  IELs (Figure 4A). The excess expression of *TCR $\alpha$*  in TCR $\alpha\beta^+$  IELs is only ~2-fold (Figure 4A), suggesting that the *TCR $\alpha$*  locus silencer is compromised in TCR $\gamma\delta^+$  IELs. The expression of *TCR $\beta$*  in  $\gamma\delta$  cells (Figure 4A) has been previously considered (Dudley et al., 1994).

Other genes with skewed tag frequencies are listed in Figure 6. Confidence in differential TCR gene expression

TAG	$\alpha\beta$	$\gamma\delta$	NET	fold	DESCRIPTION
GCCCAGCCCC	1	69	70	69.00	TCR $\delta$
TGTGCCTGTC	27	0	27	27.00	Ly6C
CTTGTATCTG	16	1	17	16.00	Unknown
GGAAGTTATG	1	14	15	14.00	Unknown
GTAATTGGAT	14	0	14	14.00	Unknown
AGCCCCCTCC	13	1	14	13.00	Schlafen 1
GTCTCTGTTG	12	1	13	12.00	Cellular apoptosis susceptibility protein
TGCACAGTGC	12	1	13	12.00	Mts1 protein (Ca <sup>++</sup> binding)
GCAGCCCTAC	1	11	12	11.00	Unknown
GTGAAGATTC	1	11	12	11.00	Unknown
GAGAGTGTGA	0	11	11	11.00	Kdap
CGCGCAACA	1	10	11	10.00	Death-associated kinase 3 (Dapk3)
GACAAATGAAA	1	10	11	10.00	Isocitrate dehydrogenase gamma subunit
AGCCAAATAC	1	10	11	10.00	Unknown
TTCAAATTTG	1	10	11	10.00	Unknown
TATCCTGAAT	10	1	11	10.00	Ly6E.1
AGCTCAGAGA	10	1	11	10.00	Unknown
GCCCGTAGGA	10	1	11	10.00	Unknown
GCAGTGGTTC	28	3	31	9.33	Lymphocyte antigen 116
CTTGCTTCCA	27	3	30	9.00	Calpain2
AGTGCCGAGC	2	18	20	9.00	Unknown
CAATGACTTG	1	9	10	9.00	Axin
AAGGTGAAGG	1	9	10	9.00	Unknown
GCCCTTAAAA	1	9	10	9.00	Unknown
TTTTTTTTCT	1	9	10	9.00	Unknown
ACAGAGCAGA	9	1	10	9.00	Multiple Ly49 family members
CTCTATGAAA	9	1	10	9.00	Unknown
TAATGCTAGA	9	1	10	9.00	Unknown
AATAAACGTG	9	1	10	9.00	Tripeptidyl peptidase II
AATGCCCTT	0	9	9	9.00	X-box binding protein 1
AATAAAGTTT	17	2	19	8.50	CD8 $\beta$
ATTTTTGTAT	2	16	18	8.00	TCR $\gamma$ C4
CTTGGGAAC	16	2	18	8.00	Unknown
CCTTGCTGGA	1	8	9	8.00	Unknown
ACCTCCAGGA	1	8	9	8.00	Unknown
CATCTGTGTA	1	8	9	8.00	Unknown
TGTGAGTCC	1	8	9	8.00	Unknown
ACAACAGGGC	8	1	9	8.00	Unknown
CTCTTGCTT	8	1	9	8.00	Unknown
CTTCTCTTC	8	1	9	8.00	Unknown
GGGGCATCTT	8	1	9	8.00	Unknown
TAAGATTGAG	8	1	9	8.00	Unknown
TACCTAAGGT	8	1	9	8.00	Unknown
TAGATTTGGG	8	1	9	8.00	Unknown
TCTGGAAAAT	8	1	9	8.00	Unknown
TTCAAAGAGT	8	1	9	8.00	Unknown
ATGAAGTTAT	0	8	8	8.00	Unknown
AATTACTATT	0	8	8	8.00	Unknown
CCCTTTGGCT	0	8	8	8.00	TEC kinase
GATTTTGCTG	8	0	8	8.00	PU.1 interaction partner (Spip)



  $\gamma\delta$  overexpressed  
  $\alpha\beta$  overexpressed

Figure 6. The 50 Most Differentially Expressed TAGs between  $\alpha\beta$  and  $\gamma\delta$  IELs

The fold differences between TCR $\alpha\beta^+$  IELs and TCR $\gamma\delta^+$  IELs are listed in a single column highlighted in green for tags found only or predominantly in  $\gamma\delta$  cells and in blue for  $\alpha\beta$  cells. Tags were sorted in descending order based on fold difference.

derives partly from their “high” abundance and consequent high sampling size. In the “low” class, apparent differential gene expression may be due to sampling variation. Therefore, candidates for differentially expressed genes were retested by quantitative, radiolabeled, linear range RT-PCR on independently derived IEL cDNA. For example, the protease Kdap is expressed in TCR $\gamma\delta^+$  IELs, but in contrast to the tag assignment ( $\alpha\beta$  0  $\gamma\delta$  11) (Figure 6), it is clearly detected in TCR $\alpha\beta^+$  IELs (Figure 7). The same was true for Tyk2 ( $\alpha\beta$  2;  $\gamma\delta$  14) (Figure 7). Further validation of candidate differentially expressed genes is ongoing.

#### Genes Specific to TCR $\alpha\beta^+$ IELs

The differential expression of several  $\alpha\beta$ -enriched tags was largely validated by RT-PCR (Figure 7). For example, Schlafen 1 (13:1) is expressed at >5-fold excess in TCR $\alpha\beta^+$  IELs, while Ly6C (27:0) is expressed at ~8- to 16-fold excess (Figure 7). Flow cytometry was then used to further validate bias, and to test whether this bias

was generally representative of TCR $\alpha\beta^+$  IELs or whether it identified heterogeneity among them.

“High” expression of Ly6C protein was much more evident in TCR $\alpha\beta^+$  cells (Figure 7B), and was enriched in  $\alpha\beta^+$ CD8 $\beta^+$  cells (Figure 7C). This distinction adds to the other differences in gene expression seen in  $\alpha\beta^+$ CD8 $\beta^+$  cells compared to other IELs (Figure 5B). However, Ly6C expression is not TCR $\alpha\beta$  specific, since splenic  $\gamma\delta$  cells and splenic  $\alpha\beta$  T cells express comparable Ly6C RNA (data not shown). Recently, Ly6C was identified by Ahmed and colleagues to be overrepresented in memory T cells compared to naïve T cells (R. Ahmed, personal communication). Thus, its differential expression in IELs suggests that  $\alpha\beta^+$ CD8 $\alpha^+\beta^-$ ,  $\gamma\delta^+$ CD8 $\alpha^+$ , and  $\gamma\delta^+$ DN cells are not preactivated memory cells, whereas some characteristics of memory cells may be present in the  $\alpha\beta^+$ CD8 $\beta^+$  subset. Not surprisingly, the skewed tag frequency of CD8 $\beta$  itself was validated by flow cytometry, as were those for CD5, CD4, and CD2 (Figure 7D).

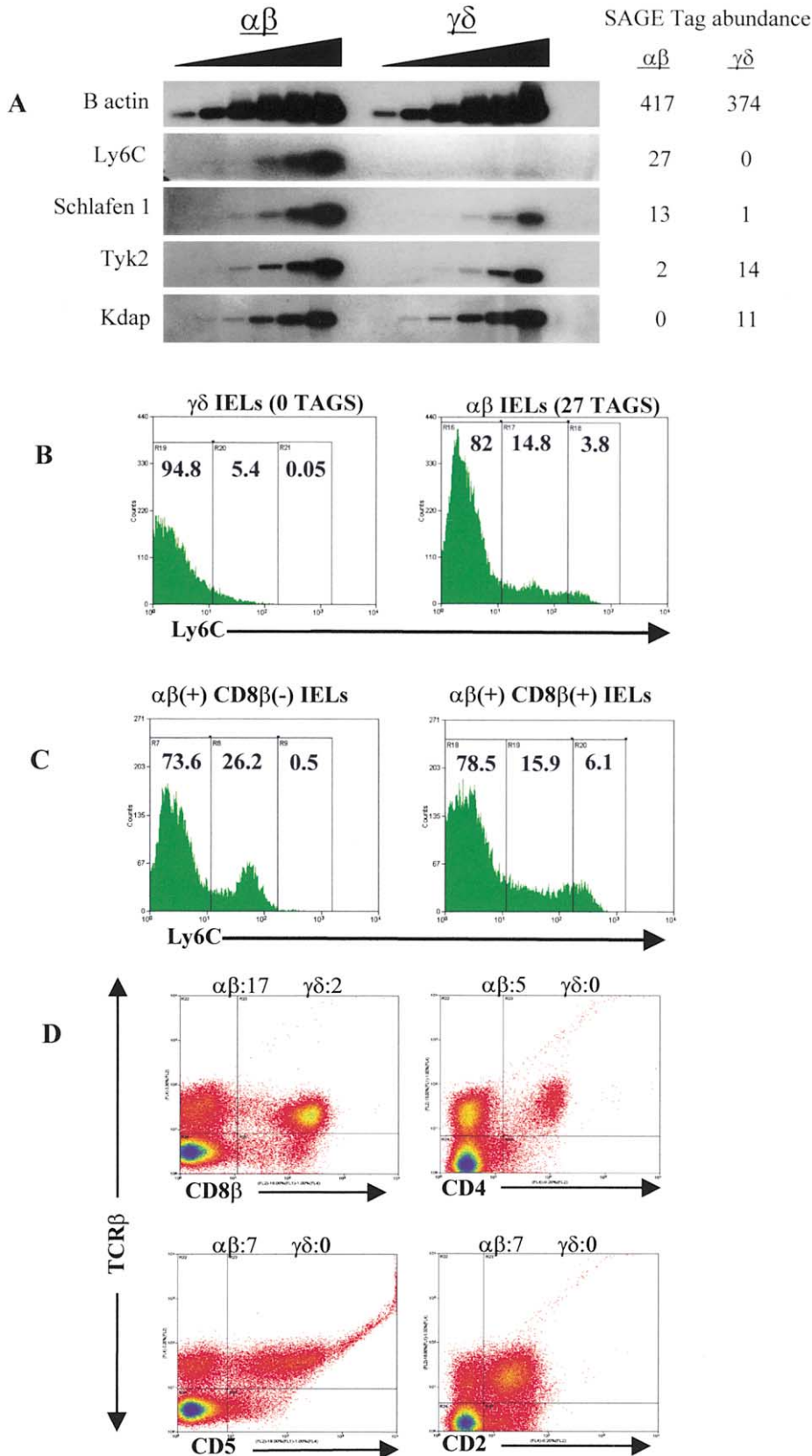


Figure 7. Validation of Differentially Expressed Genes

(A) Cycle time course RT-PCR was performed on independently derived  $\alpha\beta$  and  $\gamma\delta$  IEL cDNA. Individual PCR reactions were aliquotted from a common master mix and were put through 18, 20, 22, 24, 26, and 28 cycles.  $\alpha^{32}\text{P}$ -dCTP was included in the master mix. Products were then run out on an agarose gel, fixed, dried, and exposed to autoradiography film.

There have been two major (although not mutually exclusive) hypotheses regarding the selective pressure for the conservation of  $\gamma\delta$  cells: by one view,  $\gamma\delta$  cells are fundamentally distinct cells that fulfill unique effector functions; by the other view,  $\gamma\delta$  cells are similar to other T cells in their signaling and effector capabilities, but by virtue of their unique TCR, they enable the host to respond to a different universe of antigens. While neither of these hypotheses can as yet be discounted, this SAGE analysis failed to provide obvious support for the first of these hypotheses. Further analysis of candidates for differentially expressed genes, coupled with a SAGE analysis of activated cells, will provide further useful information on this issue.

### T Cells in the Gut

The SAGE databases characterize IEL gene expression and distinguish it from systemic cytolytic T cells. The reasons underlying the differences may include lineage differences in the development of IELs and systemic T cells; inductive cell-cell interactions in the gut; and the preeminence within the mucosal epithelium of TGF $\beta$  that can be expressed by epithelial cells (Dignass and Podolsky, 1993; Kilshaw and Murant, 1991; Suemori et al., 1991) as well as IELs. Indeed, junB, integrin  $\alpha\text{E}\beta 7$ , and most likely Btg1, are TGF $\beta$  induced.

### Qualifications

This analysis has provided an expanded characterization of IELs, revealing a signature phenotype that is distinct from well-characterized systemic T cells. This phenotype is generally applicable to individual subsets of IELs. Nonetheless, among such subsets there appear to be greater differences between TCR $\alpha\beta^+$ CD8 $\beta^+$  IELs and TCR $\alpha\beta^+$ CD8 $\alpha\alpha^+$  IELs than between TCR $\alpha\beta^+$ CD8 $\alpha\alpha^+$  and TCR $\gamma\delta^+$  IELs. This notwithstanding, the analysis is a genomic one. For some genes, correlation with levels of functional protein will be poor, although the functional clustering and flow cytometry data provided here suggest that such cases will be the exception.

Although SAGE has identified over 15,000 transcript tags, their concentration at the 3' ends of mRNA does not readily identify genes whose 3' end sequence is not in available databases. Therefore, many tags currently registered as unknown (sws) will include characterized genes. Conversely, a tag may be identified with a single gene, but could also be associated with a hitherto uncloned gene with an identical 3' oligomer. Alternatively, a tag may not be identified with a gene because the gene bank 3' sequence information is incorrect. This has occurred several times in the course of this analysis, and correct 3' end assignments required alignment with multiple ESTs from several sources. There will also be 3' ends derived from nontranslated "read-through" transcripts.

In "low" and "rare" abundance classes, the reliability of expression differences between cell types is reduced, emphasizing the need for independent validation (Figure 7). Independent validation is also required to confirm that some genes were not induced or stabilized by the extraction process, despite the precautions taken. Additionally, "low" representation might reflect a gene highly expressed in a contaminating cell population. Nevertheless, there was little evidence of this (e.g., tags for biochemical, enterocyte-associated metabolism) in the IEL libraries.

### Experimental Procedures

#### IEL Preparation

Twenty-four 6-week-old C57BL/6J female mice were purchased from Harlan, UK. On each of three occasions, IELs were isolated from intestinal epithelia of eight mice as described (Findly et al., 1993). Unless noted, all steps were performed with ice-cold reagents. Briefly, after removal of Peyer's patches, intestines were flushed with media before being opened longitudinally, cut into ~1 cm lengths, and washed repeatedly. Intestinal pieces were then collected and resuspended in prewarmed calcium- and magnesium-free media containing 10% heat-inactivated FCS and dithioerythritol. Samples were placed into a 37°C incubator with rotation (120 rpm) for 20 min. Supernatants were decanted and fresh media was added, and this was repeated a total of three times. Total cellular eluates were resuspended in RPMI 1640 media and passed over nylon wool columns. Columns were washed with RPMI 1640 and the remaining cells were resuspended in 44% Percoll (Pharmacia) solution in 1× HBSS/RPMI. The cell suspensions were layered onto a 67% Percoll cushion and centrifuged for 20 min at 400 g. IELs were collected from the 44%/67% interphase.

#### SAGE

SAGE was performed as previously described (Velculescu et al., 1995), with modifications. Several aspects of the method were reoptimized to allow the use of SAGE on smaller amounts of starting material. The complete protocol is available upon request, and is part of a manuscript in preparation. Briefly, mRNA was isolated with oligo-dT magnetic beads (Dyna). First strand cDNA synthesis was performed with a biotinylated phosphorothioate oligo-dT. The phosphorothioate inhibited the loss of the 5'-terminal biotinylated nucleotides during second strand cDNA synthesis, discovered to be a cause of significantly reduced throughput of input material at this step. Additionally, the first strand reaction was carried out using optimized enzyme and temperature conditions. The PCR amplification of colonies containing concatemers was done at reduced primer concentrations and increased cycle number which consumed the bulk of the nucleotide and primers, allowing the PCR reactions to be diluted 1:10 and taken straight into cycle sequencing without further purification. Cycle sequencing was done with BigDye Ready Reaction sequencing mix (ABI) and automated sequencing was performed on a 96-lane ABI377 sequencer. Analysis of sequenced SAGE libraries was done using the SAGE300 software (provided by Kenneth Kinzler) and eSAGE (Margulies and Innis, 2000).

#### RT-PCR

Total RNA was isolated using the RNeasy Total RNA Isolation Kit (Qiagen) following the manufacturer's protocol, and cDNA was synthesized using SuperScript II RT (Life Technologies). All primer pairs used in these studies were optimized over a 15°C temperature range

(B) IELs were harvested from C57BL/6J mice using standard methods. Cells were treated with serum and Fc-Block (Pharmingen) and then three-color stained with mAb for TCR $\alpha\beta$ ,  $\gamma\delta$ , and Ly6C (Pharmingen). Histograms were gated on  $\gamma\delta^+$  and  $\alpha\beta^+$  cells, respectively. The numbers in the different regions indicate the percentage of cells staining positively for Ly6C.

(C) IELs in both panels were gated for TCR $\beta$  expression. Histograms show Ly6C expression for CD8 $\beta^+$  and CD8 $\beta^-$   $\alpha\beta$  IELs.

(D) IELs were stained with markers for TCR $\alpha\beta$  and either CD2, CD4, CD5, or CD8 $\beta$ . Dot plots show that in all cases, these markers are predominantly associated with the TCR $\alpha\beta$  IEL subset. Numbers above the dot plots show the corresponding SAGE tag abundance for each marker tested.

and with four Mg<sup>2+</sup> concentrations using the Tetrad Gradient thermal cycler (MJ Research). For the subset-specific RT-PCRs, a series of dilutions of each of the cDNA was tested for  $\beta$ -actin expression, and dilutions giving equivalent amplification were selected. Each reaction was done in a 25  $\mu$ l total volume, in the presence of 0.25  $\mu$ M of each of the forward and reverse primers, 250  $\mu$ M of each of the dNTPs (Abgene), 1.5–3.0 mM MgCl<sub>2</sub> depending on the primer pair used, and 0.75 U Taq (Qiagen). Quantitative cycle time course RT-PCR (CT-PCR) was performed in 10  $\mu$ l reactions, and each reaction was supplemented with 0.333  $\mu$ l of <sup>32</sup>P-dCTP (10 mCi/ml, 3000 Ci/mmol; Amersham). Products were amplified for 35 cycles (subset RT-PCR) or 18–20–22–24–26 or 28 cycles (CT-PCR) as follows: 94°C for 30 s, 60°C–64.5°C for 30 s, and 72°C for 45 s and analyzed on a 2% agarose gel in TAE. Bands were visualized by ethidium bromide staining (subset RT-PCR), or following electrophoresis of the entire CT-PCR reaction, the gels were fixed in 7% trichloroacetic acid, dried on a gel drier (model 583; Bio-Rad), and the bands were visualized by autoradiography on X-ray film (X-Omat AR; Kodak). The primers used were as follows: beta-actin forward: 5'-TCCCTG TATGCTCTGGTCGTACCAC-3'; beta-actin reverse: 5'-CAGGATC TTCATGAGGTAGTCTGTCTAG-3'; Ly6C forward: 5'-CTGCAGTGC TACGAGTGCTATG-3'; Ly6C reverse: 5'-GTCTGCAGGACGACT GAGCTCA-3'; Tyk2 forward: 5'-TTGGGATTCCTGAGTCTATTCG-3'; Tyk2 reverse: 5'-ATGCTGCCTGTCTCCGCTTCCC-3'; Slnf1 forward: 5'-AGACAAGATCAATAGTCTTGAT-3'; Slnf1 reverse: 5'-CTCATG AAGCAGCAGTGAGCTTG-3'; Kdap forward: 5'-TCCCTCTCAACG AATCCACCTTG-3'; Kdap reverse: 5'-GCATCCAGCGGAGGCTG AACTCCG-3'; perforin forward: 5'-ACTGCCAGCGTAATGTGGCCG CAG-3'; perforin reverse: 5'-CAAGTACTTCGACGTGACGCTCAC-3'; JunB forward: 5'-GACTGGGAGCTCATACCCGACGGC-3'; JunB reverse: 5'-TGGCAGCCGTTGTGACATGGGTC-3'; RGS1 forward: 5'-ACCTGAGATCGATGATCCACATCT-3'; RGS1 reverse: 5'-CTG TCGATTCTCGAGTATGGAAGTC-3'; CCR7 forward: 5'-ATCATCCG TACCTTGCTCCAGGCAC-3'; CCR7 reverse: 5'-TGTCACCTGACT GGCCAGAAATGC-3'; FasL forward: 5'-TCTCTGGAGCAGTCAGCG TCAGAG-3'; FasL reverse: 5'-GGTTCCTGTTAAATGGCCACAC-3'; RANTES forward: 5'-CTCCCTGCTTTGGCTACCTCTC-3'; RANTES reverse: 5'-CTAGCTCATCTCCAAATAGTTGATG-3'; lymphoid pT $\beta$ 4 forward: 5'-TGCCCTGTCCAGCGCAGGCACTTG-3'; lymphoid pT $\beta$ 4 reverse: 5'-CTCTGTAGCCAGACCATCAGATG-3'; granzyme B forward: 5'-ATGAAGATCCTCTGCTACTGCTGA-3'; granzyme B reverse: 5'-AGCTCTAGCTCTTGGCCCTACTC-3'.

#### Flow Cytometry

All antibodies used in this study were monoclonal and were purchased from either Pharmingen or Caltag. Each mAb was titrated to determine its optimal working concentration. All FACS analysis and cell sorting was performed on a MoFlo (Cytomation). Antibody staining was generally done on ice, and cells were pretreated with Fc-Block (Pharmingen). Staining was done in PBS supplemented with 2% FCS.

#### Acknowledgments

This work was supported by the Wellcome Trust, in facilities supported by the Dunhill Medical Trust. We thank Scott Roberts and Elizabeth Ramsburg for assistance with IEL preparations, Wayne Turnbull and Robert McCord for flow cytometry, Nigel Grindley for advice on reverse transcription, Bob Tigelaar and particularly Susan John for an immense input of advice and expertise on the databases, Ken Kinzler for provision of the original SAGE protocol and assistance with software, Elliott Marguiles for assistance with eSAGE and database handling, Sue Chin for statistical analyses, Robert Shires for help with EST database handling, and Suzanne Creighton and Jeremy Cridland for assistance with writing and computing. Additionally, we are indebted to the curators of the public databases at the NCBI, EMBL, and RIKEN, without which the work would not have been possible.

Received February 2, 2001; revised July 16, 2001.

#### References

Agrawal, S., Marquet, J., Freeman, G.J., Tawab, A., Bouteiller, P.L., Roth, P., Bolton, W., Ogg, G., Bousmell, L., and Bensussan, A. (1999).

Cutting edge: MHC class I triggering by a novel cell surface ligand costimulates proliferation of activated human T cells. *J. Immunol.* 162, 1223–1226.

Allison, J.P., and Havran, W.L. (1991). The immunobiology of T cells with invariant gamma delta antigen receptors. *Annu. Rev. Immunol.* 9, 679–705.

Anumanthan, A., Bensussan, A., Bousmell, L., Christ, A.D., Blumberg, R.S., Voss, S.D., Patel, A.T., Robertson, M.J., Nadler, L.M., and Freeman, G.J. (1998). Cloning of BY55, a novel Ig superfamily member expressed on NK cells, CTL, and intestinal intraepithelial lymphocytes. *J. Immunol.* 161, 2780–2790.

Avice, M.N., Sarfati, M., Triebel, F., Delespesse, G., and Demeure, C.E. (1999). Lymphocyte activation gene-3, a MHC class II ligand expressed on activated T cells, stimulates TNF-alpha and IL-12 production by monocytes and dendritic cells. *J. Immunol.* 162, 2748–2753.

Beadling, C., Druey, K.M., Richter, G., Kehrl, J.H., and Smith, K.A. (1999). Regulators of G protein signaling exhibit distinct patterns of gene expression and target G protein specificity in human lymphocytes. *J. Immunol.* 162, 2677–2682.

Beresford, P.J., Xia, Z., Greenberg, A.H., and Lieberman, J. (1999). Granzyme A loading induces rapid cytolysis and a novel form of DNA damage independently of caspase activation. *Immunity* 10, 585–594.

Boismenu, R., and Havran, W. (1994). Modulation of epithelial cell growth by intraepithelial gamma delta T cells. *Science* 266, 1253–1255.

Boismenu, R., Feng, L., Xia, Y.Y., Chang, J.C., and Havran, W.L. (1996). Chemokine expression by intraepithelial gamma delta T cells. Implications for the recruitment of inflammatory cells to damaged epithelia. *J. Immunol.* 157, 985–992.

Carlier, M.F., and Pantaloni, D. (1994). Actin assembly in response to extracellular signals: role of capping proteins, thymosin beta 4 and profilin. *Semin. Cell Biol.* 5, 183–191.

Das, G., Gould, D.S., Augustine, M.M., Fragoso, G., Scitto, E., Stroynowski, I., Van Kaer, L., Schust, D.J., Ploegh, H., and Janeway, C.A., Jr. (2000). Qa-2-dependent selection of CD8alpha/alpha T cell receptor alpha/beta(+) cells in murine intestinal intraepithelial lymphocytes. *J. Exp. Med.* 192, 1521–1528.

DeBenedette, M.A., Shahinian, A., Mak, T.W., and Watts, T.H. (1997). Costimulation of CD28 – T lymphocytes by 4-1BB ligand. *J. Immunol.* 158, 551–559.

Deftos, M.L., He, Y.W., Ojala, E.W., and Bevan, M.J. (1998). Correlating notch signaling with thymocyte maturation. *Immunity* 9, 777–786.

Dignass, A.U., and Podolsky, D.K. (1993). Cytokine modulation of intestinal epithelial cell restitution: central role of transforming growth factor beta. *Gastroenterology* 105, 1323–1332.

D'Souza, C.D., Cooper, A.M., Frank, A.A., Mazzaccaro, R.J., Bloom, B.R., and Orme, I.M. (1997). An anti-inflammatory role for gamma delta T lymphocytes in acquired immunity to *Mycobacterium tuberculosis*. *J. Immunol.* 158, 1217–1221.

Dudley, E.C., Petrie, H.T., Shah, L.M., Owen, M.J., and Hayday, A.C. (1994). T cell receptor beta chain gene rearrangement and selection during early thymocyte development in adult mice. *Immunity*, 1, 83–93.

Findly, R.C., Roberts, S.J., and Hayday, A.C. (1993). Dynamic response of murine gut intraepithelial T cells after infection by the coccidian parasite *Eimeria*. *Eur. J. Immunol.* 23, 2557–2564.

Freeman, G.J., Long, A.J., Iwai, Y., Bourque, K., Chernova, T., Nishimura, H., Fitz, L.J., Malenkovich, N., Okazaki, T., Byrne, M.C., et al. (2000). Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. *J. Exp. Med.* 192, 1027–1034.

Fu, Y.X., Roark, C.E., Kelly, K., Drevets, D., Campbell, P., O'Brien, R., and Born, W. (1994). Immune protection and control of inflammatory tissue necrosis by gamma delta T cells. *J. Immunol.* 153, 3101–3115.

Fujihashi, K., McGhee, J., Kweon, M., Cooper, M., Tonegawa, S., Takahashi, I., Hiroi, T., Mestecky, J., and Kiyono, H. (1996). gamma/

- delta T cell-deficient mice have impaired mucosal immunoglobulin A responses. *J. Exp. Med.* **183**, 1929–1935.
- Groh, V., Rhinehart, R., Randolph-Habecker, J., Topp, M.S., Riddell, S.R., and Spies, T. (2001). Costimulation of CD8 alpha beta T cells by NKG2d via engagement by MIC induced on virus-infected cells. *Nat. Immunol.* **2**, 255–260.
- Halary, F., Peyrat, M.A., Champagne, E., Lopez-Botet, M., Moretta, A., Moretta, L., Vie, H., Fournie, J.J., and Bonneville, M. (1997). Control of self-reactive cytotoxic T lymphocytes expressing gamma delta T cell receptors by natural killer inhibitory receptors. *Eur. J. Immunol.* **27**, 2812–2821.
- Havran, W.L., Chien, Y.H., and Allison, J.P. (1991). Recognition of self antigens by skin-derived T cells with invariant gamma delta antigen receptors. *Science* **252**, 1430–1432.
- Hayday, A., and Viney, J.L. (2000). The ins and outs of body surface immunology. *Science* **290**, 97–100.
- Irie-Sasaki, J., Sasaki, S., Matsumoto, W., Opavsky, A., Cheng, M., Welstead, G., Griffiths, E., Krawczyk, C., Richardson, C.D., Aitken, K., et al. (2001). CD45 is a JAK phosphatase and negatively regulates cytokine receptor signalling. *Nature* **409**, 349–354.
- Janeway, C.A., Jones, B., and Hayday, A.C. (1988). Specificity and function of cells bearing gamma delta T cell receptors. *Immunol. Today* **9**, 73–76.
- Ke, Y., Pearce, K., Lake, J.P., Ziegler, K.H., and Kapp, J.A. (1997). Gamma delta T lymphocytes regulate the induction and maintenance of oral tolerance. *J. Immunol.* **158**, 3610–3618.
- Kilshaw, P.J., and Murrant, S.J. (1991). Expression and regulation of beta 7(beta p) integrins on mouse lymphocytes: relevance to the mucosal immune system. *Eur. J. Immunol.* **21**, 2591–2597.
- Kim, J.I., Ho, I.C., Grusby, M.J., and Glimcher, L.H. (1999). The transcription factor c-Maf controls the production of interleukin-4 but not other Th2 cytokines. *Immunity* **10**, 745–751.
- King, D.P., Hyde, D.M., Jackson, K.A., Novosad, D.M., Ellis, T.N., Putney, L., Stovall, M.Y., Van Winkle, L.S., Beaman, B.L., and Ferrick, D.A. (1999). Protective response to pulmonary injury requires gamma delta T lymphocytes. *J. Immunol.* **162**, 5033–5036.
- Komano, H., Fujiura, Y., Kawaguchi, M., Matsumoto, S., Hashimoto, Y., Obana, S., Mombaerts, P., Tonegawa, S., Yamamoto, H., Itoharu, S., et al. (1995). Homeostatic regulation of intestinal epithelia by intraepithelial gamma delta T cells. *Proc. Natl. Acad. Sci. USA* **92**, 6147–6151.
- Kuo, C.T., and Leiden, J.M. (1999). Transcriptional regulation of T lymphocyte development and function. *Annu. Rev. Immunol.* **17**, 149–187.
- Lanier, L.L. (2001). On guard—activating NK cell receptors. *Nat. Immunol.* **2**, 23.
- Lepage, A.C., Buzoni-Gatel, D., Bout, D.T., and Kasper, L.H. (1998). Gut-derived intraepithelial lymphocytes induce long term immunity against *Toxoplasma gondii*. *J. Immunol.* **161**, 4902–4908.
- Li, B., Tournier, C., Davis, R.J., and Flavell, R.A. (1999). Regulation of IL-4 expression by the transcription factor JunB during T helper cell differentiation. *EMBO J.* **18**, 420–432.
- Li, H., Chen, J., Huang, A., Stinson, J., Heldens, S., Foster, J., Dowd, P., Gurney, A.L., and Wood, W.I. (2000). Cloning and characterization of IL-17B and IL-17C, two new members of the IL-17 cytokine family. *Proc. Natl. Acad. Sci. USA* **97**, 773–778.
- Lin, T., Yoshida, H., Matsuzaki, G., Guehler, S.R., Nomoto, K., Barrett, T.A., and Green, D.R. (1999). Autospecific gamma delta thymocytes that escape negative selection find sanctuary in the intestine. *J. Clin. Invest.* **104**, 1297–1305.
- Lohning, M., Stroehmann, A., Coyle, A.J., Grogan, J.L., Lin, S., Gutierrez-Ramos, J.C., Levinson, D., Radbruch, A., and Kamradt, T. (1998). T1/ST2 is preferentially expressed on murine Th2 cells, independent of interleukin 4, interleukin 5, and interleukin 10, and important for Th2 effector function. *Proc. Natl. Acad. Sci. USA* **95**, 6930–6935.
- Malinda, K.M., Sidhu, G.S., Mani, H., Banaudha, K., Maheshwari, R.K., Goldstein, A.L., and Kleinman, H.K. (1999). Thymosin beta4 accelerates wound healing. *J. Invest. Dermatol.* **113**, 364–368.
- Margulies, E.H., and Innis, J.W. (2000). eSAGE: managing and analysing data generated with serial analysis of gene expression (SAGE). *Bioinformatics* **16**, 650–651.
- Mazzucchelli, L., Hauser, C., Zraggen, K., Wagner, H.E., Hess, M.W., Laissue, J.A., and Mueller, C. (1996). Differential in situ expression of the genes encoding the chemokines MCP-1 and RANTES in human inflammatory bowel disease. *J. Pathol.* **178**, 201–206.
- McVay, L.D., Jaswal, S.S., Kennedy, C., Hayday, A., and Carding, S. (1998). The generation of human gamma delta T cell repertoires during fetal development. *J. Immunol.* **160**, 5851–5860.
- Muller, S., Buhler-Jungo, M., and Mueller, C. (2000). Intestinal intraepithelial lymphocytes exert potent protective cytotoxic activity during an acute virus infection. *J. Immunol.* **164**, 1986–1994.
- Munn, D.H., Zhou, M., Attwood, J.T., Bondarev, I., Conway, S.J., Marshall, B., Brown, C., and Mellor, A.L. (1998). Prevention of allogeneic fetal rejection by tryptophan catabolism. *Science* **281**, 1191–1193.
- Nagata, S., and Golstein, P. (1995). The Fas death factor. *Science* **267**, 1449–1456.
- Nakamura, M., Xavier, R.M., Tsunematsu, T., and Tanigawa, Y. (1995). Molecular cloning and characterization of a cDNA encoding monoclonal nonspecific suppressor factor. *Proc. Natl. Acad. Sci. USA* **92**, 3463–3467.
- Oliveira-Dos-Santos, A.J., Matsumoto, G., Snow, B.E., Bai, D., Houston, F.P., Whishaw, I.Q., Mariathasan, S., Sasaki, T., Wakeham, A., Ohashi, P.S., et al. (2000). Regulation of T cell activation, anxiety, and male aggression by RGS2. *Proc. Natl. Acad. Sci. USA* **97**, 12272–12277.
- Parrish-Novak, J., Dillon, S.R., Nelson, A., Hammond, A., Sprecher, C., Gross, J.A., Johnston, J., Madden, K., Xu, W., West, J., et al. (2000). Interleukin 21 and its receptor are involved in NK cell expansion and regulation of lymphocyte function. *Nature* **408**, 57–63.
- Passegué, E., Jochum, W., Schorpp-Kistner, M., Möhle-Steinlein, U., and Wagner, E.F. (2001). Chronic myeloid leukemia with increased granulocyte progenitors in mice lacking junB expression in the myeloid lineage. *Cell* **104**, 21–32.
- Reif, K., and Cyster, J.G. (2000). RGS molecule expression in murine B lymphocytes and ability to down-regulate chemotaxis to lymphoid chemokines. *J. Immunol.* **164**, 4720–4729.
- Roberts, S., Smith, A.L., West, A.B., Wen, L., Findly, R.C., Owen, M.J., and Hayday, A.C. (1996). T-cell receptor alpha-beta (+) and gamma-delta (+) deficient mice display abnormal but distinct phenotypes toward a natural, widespread infection of the intestinal epithelium. *Proc. Natl. Acad. Sci. USA* **93**, 11774–11779.
- Rocha, B., Vassalli, P., and Guy-Grand, D. (1994). Thymic and extrathymic origins of gut intraepithelial lymphocyte populations in mice. *J. Exp. Med.* **180**, 681–686.
- Rudin, C.M., Engler, P., and Storb, U. (1990). Differential splicing of thymosin beta 4 mRNA. *J. Immunol.* **144**, 4857–4862.
- Schwarz, D.A., Katayama, C.D., and Hedrick, S.M. (1998). Schlafen, a new family of growth regulatory genes that affect thymocyte development. *Immunity* **9**, 657–668.
- Shi, Y., Ullrich, S.J., Zhang, J., Connolly, K., Grzegorzewski, K.J., Barber, M.C., Wang, W., Wathen, K., Hodge, V., Fisher, C.L., et al. (2000). A novel cytokine receptor-ligand pair. Identification, molecular characterization, and in vivo immunomodulatory activity. *J. Biol. Chem.* **275**, 19167–19176.
- Shiohara, T., Moriya, N., Hayakawa, J., Itoharu, S., and Ishikawa, H. (1996). Resistance to cutaneous graft-vs.-host disease is not induced in T cell receptor delta gene-mutant mice. *J. Exp. Med.* **183**, 1483–1489.
- Shresta, S., Graubert, T.A., Thomas, D.A., Raptis, S.Z., and Ley, T.J. (1996). Granzyme A initiates an alternative pathway for granulocyte-mediated apoptosis. *Immunity* **10**, 595–605.
- Shuford, W.W., Klussman, K., Tritchler, D.D., Loo, D.T., Chalupny, J., Siadak, A.W., Brown, T.J., Emswiler, J., Raecho, H., Larsen, C.P., et al. (1997). 4-1BB costimulatory signals preferentially induce CD8+ T cell proliferation and lead to the amplification in vivo of cytotoxic T cell responses. *J. Exp. Med.* **186**, 47–55.

- Suemori, S., Ciacci, C., and Podolsky, D.K. (1991). Regulation of transforming growth factor expression in rat intestinal epithelial cell lines. *J. Clin. Invest.* *87*, 2216–2221.
- Szabo, S.J., Kim, S.T., Costa, G.L., Zhang, X., Fathman, C.G., and Glimcher, L.H. (2000). A novel transcription factor, T-bet, directs Th1 lineage commitment. *Cell* *100*, 655–669.
- Szabowski, A., Maas-Szabowski, N., Andrecht, S., Kolbus, A., Schorpp-Kistner, M., Fusenig, N.E., and Angel, P. (2000). c-Jun and JunB antagonistically control cytokine-regulated mesenchymal-epidermal interaction in skin. *Cell* *103*, 745–755.
- Taguchi, T., Aicher, W.K., Fujihashi, K., Yamamoto, M., McGhee, J.R., Bluestone, J.A., and Kiyono, H. (1991). Novel function for intestinal intraepithelial lymphocytes: murine CD3<sup>+</sup>, gamma/delta TCR<sup>+</sup> T cells produce IFN-gamma and IL-5. *J. Immunol.* *147*, 3736–3744.
- Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. (1995). Serial analysis of gene expression. *Science* *270*, 484–487.
- Viney, J.L., Kilshaw, P.J., and MacDonald, T.T. (1990). Cytotoxic alpha/beta<sup>+</sup> and gamma/delta<sup>+</sup> T cells in murine intestinal epithelium. *Eur. J. Immunol.* *20*, 1623–1626.
- Young, J.D., Lawrence, A.J., MacLean, A.G., Leung, B.P., McInnes, I.B., Canas, B., Pappin, D.J., and Stevenson, R.D. (1999). Thymosin beta 4 sulfoxide is an anti-inflammatory agent generated by monocytes in the presence of glucocorticoids. *Nat. Med.* *5*, 1424–1427.