THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 17, Issue of April 23, pp. 17905–17913, 2004

Printed in U.S.A.

Pearson Correlation Analysis of Microarray Data Allows for the Identification of Genetic Targets for Early B-cell Factor*[S]

Received for publication, January 20, 2004
Published, JBC Papers in Press, February 11, 2004, DOI 10.1074/jbc.M400589200

Robert Månsson, Panagiotis Tsapogas, Mikael Åkerlund, Anna Lagergren, Ramiro Gisler, and Mikael Sigvardsson‡

From the Department of Heamatopoetic Stemcell Biology, Stemcell Center, Lund University, S221 84 Lund, Sweden

B lymphocyte development is a complex biological process critically dependent on the transcription factor early B cell factor (EBF). To deepen understanding of the roles for EBF in this process, we have used Pearson correlation analysis to evaluate microarray data from a set of mouse B lymphoid cell lines representing different stages of development. Comparing the expression pattern of EBF to that of the other genes in the data set revealed that *VpreB1*, *mb-1*, and *λ5*, all known target genes, presented high correlation values to EBF. High correlations were also seen for the *VpreB3* and *CD19* genes and biochemical as well as functional data supported that they are target genes for EBF even though the expression of *CD19* was critically dependent of Pax-5. We also obtained evidence for extensive collaborative actions of EBF and E47 even though microarray analysis of hematopoetic progenitor cells ectopically expressing these proteins suggested that they activated only a subset of pre-B cell restricted genes.

B cell development proceeds from a multipotent progenitor in the bone marrow into a highly specialized immunoglobulin secreting plasma cell. The process can be divided into several stages based on the recombination status of the immunoglobulin genes and gene expression patterns (1, 2). The differentiation pathway is dependent on the action of transcription factors that act to coordinate the expression of these stage-specific genes (3). In the earliest stages of B cell development there exists an apparent need of the coordinated action of transcription factors Pu.1 (4), EBF[1] (5), and E2A (6) for the formation of the earliest progenitors, whereas the transcription factor BSAP (7, 8) is crucial for lineage commitment (9–11) and progression into the pre-B cell stage (12, 13). Homologous disruptions of the genes encoding these proteins in mice have proven their importance *in vivo* (12–19) but even though several target genes are identified (3, 20, 21), there is still a need to elucidate how they exert their biological functions to establish and promote B cell development. One possibility to obtain information about genetic programs and coordinated gene expression is by the use of microarray technology that allows for the simultaneous measurements of the expression levels of several thousand genes. We have earlier used a set of B cell lines arrested at different stages of their development to establish a crude map over gene expression patterns in B cell differentiation (22). The analysis of control gene expression suggested that this approach allowed for a reasonable approximation of expression patterns (22) giving us a tool to investigate also the coordination of stage-specific gene activation. We have used part of these data to estimate the relative importance of transcription factors in the activation of the mouse *mb-1* (Igα) promoter by Pearson correlation analysis (23). This promoter contains binding sites for EBF, E47, BSAP, and Ets proteins (23–27) and when comparing the expression levels of the *mb-1* message to the levels of the transcription factors, EBF gave a 98% correlation, whereas BSAP and Ets proteins resulted in lower values (23). This indicated to us that the expression of the *mb-1* gene was highly dependent of EBF and that this type of analysis can provide important biological information.

To increase the understanding of the role of EBF in early B cell development we decided to investigate if global Pearson correlation analysis of microarray data can be used to identify genetic targets. The expression levels of EBF displayed a high correlation with a number of known EBF target genes but also of other genes including the *VpreB3* and *CD19* genes. A role for EBF in the regulation of these genes could also be supported by biochemical and functional assays suggesting that correlation analysis can be used to investigate genetic networks and identify targets for transcription factors in complex biological systems.

MATERIALS AND METHODS

Tissue Culture Conditions—All cells were grown at 37 °C and 5% CO_2 in RPMI supplemented with 7.5% fetal calf serum, 10 mM HEPES, 2 mM pyruvate, 50 μM 2-mercaptoethanol, and 50 μg/ml gentamycin (all purchased from Invitrogen AB). The pro-B cell lines were grown in RPMI as above supplemented with 10% of interleukin-3 containing WEHI3-conditioned media. The Ba/F3 cells were kind gifts from Dr. R. Grosschedl (Gene Center, Munich).

Gene Expression Analysis—RNA was prepared using Trizol (Invitrogen) and 7.5 μg of total RNA was annealed to a T7-oligo T primer by denaturation at 70 °C for 10 min followed by 10 min of incubation of the samples on ice. First strand synthesis was performed for 2 h at 42 °C using 20 units of SuperScript Reverse Transcriptase (Invitrogen) in buffers and nucleotide mixtures according to the manufacturer's instructions. This was followed by a second strand synthesis for 2 h at 16 °C, using RNase H, *Escherichia coli* DNA polymerase I, and *E. coli* DNA ligase (all from Invitrogen), according to the manufacturer's instructions. The obtained double stranded cDNA was then blunted by the addition of 20 units of T4 DNA polymerase and incubated for 5 min at 16 °C. The material was then purified by phenol:cloroform:isoamyl alcohol extraction followed by precipitation with NH_4Ac and ethanol. The cDNA was then used in an *in vitro* transcription reaction for 6 h at 37 °C using a T7 *in vitro* translation kit and biotin-labeled ribonucleotides. The obtained cRNA was purified from unincorporated nucleo-

* This work was supported by the Swedish Medical Research Council (VRM), Cancerfonden, Crafoords Foundation, Kocks Foundation, Österlunds Foundation, and Barncancer fonden. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains the Appendix.

‡ To whom correspondence should be addressed: BMC B12, S-221 84 Lund, Sweden. Tel.: 46-0-46-2223829; Fax: 46-0-46-2223600; E-mail: mikael.sigvardsson@stemcell.lu.se.

[1] The abbreviations used are: EBF, early B cell factor; MES, 4-morpholineethanesulfonic acid; EMSA, electrophoretic mobility shift assay; RT, reverse transcriptase.

tides on an RNAeasy column (Qiagen). The eluted cRNA was then fragmented by incubation of the products for 2 h in fragmentation buffer (40 mM Tris acetate, pH 8.1, 100 mM KOAc, 150 mM MgOAc). 20 μg of the final fragmented cRNA was hybridized to the Affymetrix chip U74Av2 (Affymetrix) in 200 μl of hybridization buffer (100 mM MES buffer, pH 6.6, 1 M NaCl, 20 mM EDTA, 0.01% Tween 20) supplemented with herring sperm DNA (100 μg/ml) and acetylated bovine serum albumin (500 μg/ml) in an Affymetrix Gene Chip hybridization oven 320. The chip was then developed by the addition of fluorescein isothiocyanate-streptavidin followed by washing using an Affymetrix Gene Chip Fluidics Station 400. Scanning was performed using a Hewlett-Packard Gene Array scanner.

Data Analysis—Affymetrix gene chip arrays were analyzed using the dChip software (28) and obtained model-based expression values (based on the PM/MM model) were exported. Standard Pearson correlation coefficients were determined for each gene *versus* all other genes using data from 16 different Affymetrix arrays. As a measurement of random/average correlation between genes the average number of correlations for all given values were calculated. (Process was automated and programs can be made available on request.) The slope of the fitted line, K_i, used to compare expression levels of EBF to the rest of the data set, was calculated using linear regression minimizing the sum of the square errors.

Protein Extracts and Electrophoretic Mobility Shift Assay (EMSA)—Nuclear extracts were prepared according to Schreiber *et al.* (29). DNA probes were labeled with [γ-^{32}P]ATP by incubation with T4 polynucleotide kinase (Roche Applied Science), annealed with the complementary strand and purified on a 5% polyacrylamide TBE gel. 5–10 μg of nuclear extract or 0.5–2 μl *in vitro* transcribed/translated protein was incubated with the labeled probe (20,000 cpm, 3 fmol) for 30 min at room temperature in binding buffer (10 mM HEPES, pH 7.9, 70 mM KCl, 1 mM dithiothreitol, 1 mM EDTA, 2,5 mM $MgCl_2$, 5% glycerol) with 0.75 μg of poly(dI-dC) (Amersham Biosciences Inc.). 1 mM $ZnCl_2$ was added to shift assays performed in the presence of EBF. DNA competitors were added 10 min before addition of the DNA probe. The samples were separated on a 6% acrylamide TBE gel, which was dried and subjected to autoradiography. Competitors based on synthetic oligonucleotides were added at molar excesses indicated in the respective figures. Full-length *λ5*, *mb-1*, *CD19*, and *VpreB* promoters were generated by PCR (see below) and were added at molar excesses as indicated in the respective figures.

Oligonucleotides used for electrophoretic mobility shift assays were: mb-1 sense: 5′-AGCCACCTCTCAGGGGAATTGTGG; mb-1 antisense, 5′-CCACAATTCCCCTGAGAGGTGGCT; mutated *mb-1* sense, 5′ AGCCACCTCTCAGCCGTTTTGTGG; mutated *mb-1* antisense, 5′-CCACAAAACGGCTGAGAGGTGGCT; *μE5* sense, 5′-GGCCAGAACACCTGCAGACG; *μE5* antisense, 5′-CGTCTGCAGGTGTTCTGGCC; Oct binding site sense, 5′-CATCTCAAGTGATTTGCATCGCATGAGACG; Oct binding site antisense, 5′-CGTCTCATGCGATGCAAATCACTTGAGATC; CD19-BSAP sense, 5′-GCAGACACCCATGGTTGAGTGCCCTCCAGG; CD19-BSAP antisense, 5′-CCTGGAGGGCACTCAACCATGGGTGTCTGC; VpreB3 EBF site 1 sense, 5′-CTCAGTCCCACTAGGGTAGCTATT; VpreB3 EBF site 1 antisense, 5′-AATAGCTACCCTAGTGGGACTGAG; VpreB3 EBF site 2 sense, 5′-GCCTCAGTCTCCTGGGATTAACAG; VpreB3 EBF site 2 antisense, 5′-CTGTTAATCCCAGGAGACTGAGGC; VpreB3 E47 site 1 sense, 5′-AAGTGCAAGCAGGTGAAGAGAGTG; VpreB3 E47 site 1 antisense, 5′-CACTCTCTTCACCTGCTTGCACTT; VpreB3 E47 site 2 sense, 5′-TCTGAATCACAGGTGTGCCCTCCA; VpreB3 E47 site 2 antisense, 5′-TGGAGGGCACACCTGTGATTCAGA; VpreB1/2 EBF site intron sense, 5′-AAGGAAACTCTTGGGGCCCAGG; VpreB1/2 EBF site intron antisense, 5′-CCTGGGCCCCAAGAGTTTCCTT; VpreB1 E47 site 1 sense, 5′-GACGTTCCAGCAGGTGCTTCCTCCC; VpreB1 E47 site 1 antisense, 5′-GGGAGGAAGCACCTGCTGGAACGTC; VpreB1 E47 site 2 sense, 5′-GTGCTTCCCAGCTGGTCAGGGC; VpreB1 E47 site 2 antisense, 5′-GCCCTGACCAGCTGGGAAGCAC; VpreB1 E47 site 3 sense, 5′-GGGGAGAGCAGGTGCCAAGTGA; VpreB1 E47 site 3 antisense, 5′-TCACTTGGCACCTGCTCTCCCC; VpreB1 E47 site 4 sense, 5′-CCAAGTGACAGGTGTGGAGCAG; VpreB1 E47 site 4 antisense, 5′-CTGCTCCACACCTGTCACTTGG; CD19 EBF site sense, 5′-GAAGCCGGGTGCCCCAGGGAGCCTG; CD19 EBF site antisense, 5′-CAGGCTCCCTGGGGCACCCGGCTTC; CD19 E-box 1 sense, 5′-TTAACTCCACCAGCTGGGAAAC; CD19 E-box 1 antisense, 5′-GTTTCCCAGCTGGTGGAGTTAA; CD19 E-box 2 sense, 5′-AATATTTACAGGTGCCACTATG; CD19 E-box 2 antisense, 5′-CATAGTGGCACCTGTAAATATT. Promoter competitors were obtained by PCR amplifications from the pGL-3 promoter plasmids using the Gl-2 and RV3 primers (Promega).

Plasmids and Constructs—The E47FD plasmid was based on the neomycin encoding eukaryotic expression vector pcDNA3 (Invitrogen) placing the inserted cDNA under the control of a cytomegalovirus promoter. The EBF encoding plasmid mediating puromycin resistance was based on the retroviral vector pBabe (30). The *mb-1* and *λ5* promoter constructs have been reported previously (23, 31), whereas the *VpreB1*, *CD19*, and *VpreB3* promoters were cloned by PCR using genomic mouse DNA as template and cloning of the resulting products in the SmaI site of the luciferase reporter plasmid pGL3 basic (Promega). All constructs were verified by sequencing.

Oligonucleotides used for promoter constructs were: VpreB1 promoter sense, 5′-GATCTGCTTATTTGGGGCTCAG; VpreB1 promoter antisense, 5′-GGCTCCAGACCTGCAGCACCCTG; VpreB3 promoter sense, 5′-GAGTCACAAAACTTATGGAATGTCTC; VpreB3 promoter antisense, 5′-CAGAAAGGCCGACCTAGAGTT; CD19 promoter sense, 5′-GATGCCAGACATTGTGGCATATGC; and CD19 promoter antisense, 5′-GGTAGCCAGGCTCCCTGGGGCAC.

In Vitro Transcription and Translation—Recombinant protein was generated by coupled *in vitro* transcription/translation using a TNT reticulocyte lysate kit (Promega) and pCDNA3 (Invitrogen) template plasmids (32).

Retroviruses and Infection of BaF/3 Cells—The Pax-5 encoding retrovirus virus and the control virus were kind gifts from Professor Sten-Eirik Jacobsen. The viruses were based on murine stem cell virus vectors expressing BSAP and green fluorescent protein, or for the control only green fluorescent protein, as a single bicistronic transcript containing an internal ribosomal entry site. Prior to viral coating 24-well non-tissue culture plates were incubated 2 h in room temperature with retronectin (40 μg/ml, Stemcell Technologies) followed by 30 min in phosphate-buffered saline supplemented with 2% bovine serum albumin. Plates were then incubated with the viral supernatant for 1 h at tissue culture conditions (see above). 50,000 BaF3 cells were seeded into the coated wells in media supplemented with 10% viral supernatant, 6 μg/ml Polybrene (Sigma), and 25 ng/ml murine IL-3 (Stemcell Technologies). Cells were incubated for 24 h before repeating the infection procedure. Day 7 after infection green fluorescent protein-positive cells were sorted on a FACS-Vantage Cell sorter (BD Biosciences, San Jose, CA).

Reverse Transcriptase and Polymerase Chain Reactions—RNA was prepared from cells using Trizol (Invitrogen) and cDNA was generated by annealing 1 μg of total RNA to 0.5 μg of random hexamers in 10 μl of diethyl pyrocarbonate-treated water. Reverse transcriptase reactions were performed with 200 units of SuperScript reverse transcriptase (Invitrogen) in the manufacturer's buffer supplemented with 0.5 mM dNTP, 10 mM dithiothreitol, and 20 units of RNase inhibitor (Roche Diagnostics) in a total volume of 20 μl, at 37 °C for 1 h. One-twentieth of the RT reactions were used in the PCR assays. PCR reactions were performed with 1 unit of Taq polymerase (Invitrogen) in the manufacturer's buffer supplemented with 0.2 mM dNTP, in a total volume of 25 μl. Primers were added to a final concentration of 1 mM.

For all PCRs the program was identical with the exception of the number of cycles (Y) and the annealing temperature (XX).

- 95°C for 2 min
- 95°C for 45 sec
- XX°C for 45 sec
- 72°C for 1 min

} Y cycles

- 72°C for 2 min

STRUCTURE 1

Annealing temperatures and cycles were: 36B4 at 57 °C for 25 cycles; λ5 at 60 °C for 32 cycles; VpreB1 at 60 °C for 32 cycles; VpreB2 at 60 °C for 32 cycles; VpreB3 at 60 °C for 32 cycles; HPRT at 50 °C for 30 cycles; and EBF at 58 °C for 35 cycles.

Oligonucleotides used for RT-PCR were: 36B4, sense, 5′-GAGGAATCAGATGAGGATATGGGA and antisense, 5′-AAGCAGGCTGACTTGGTTGC; λ5, sense, 5′-TGTGAAGTTCTCCTCCTGCTC and antisense, 5′-ACCACCAAAGTACCTGGGTAG; VpreB1, sense, 5′-TGCTCATGCTGCTGCCTAT and antisense, 5′-CTCCGGAGCCCCACGGCA; VpreB2, sense, 5′-CTCATGCTGCTGGCCCAC and antisense, 5′-TTCTAGGCTCTGCAAAGCCC; VpreB3, sense, 5′-TGTGGCAGTCTTTCAGCCAACT

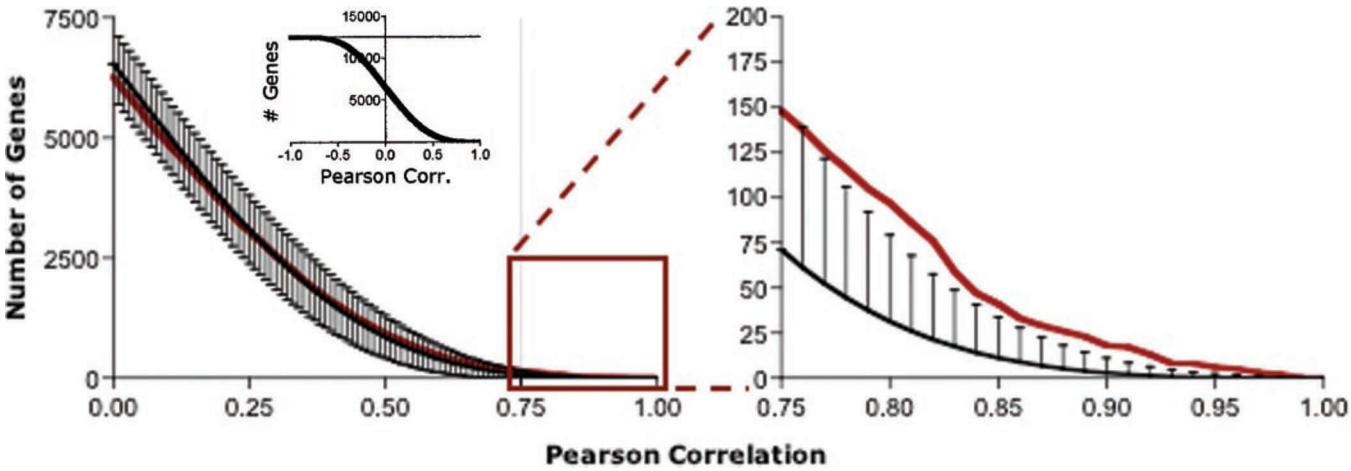

FIG. 1. **EBF displays higher correlation with a set of genes than expected from random events.** The diagram displays a standard curve of Pearson correlation values generated by analysis of all possible correlations in the data set (*filled line*) with standard deviation indicated. The diagram indicates the number of genes (*y* axis) with correlations above a given value (*y* axis). The *small graph* indicates the complete distribution of correlation values in the data set, whereas the first *large panel* displays the distribution of positive correlation values. The *red line* indicates the values obtained by comparing the correlation of EBF with all the genes in the data set. The *second large graph* shows a zoom on the region with the highest correlation values are shown.

and antisense, 5′-AATATCTCCAGCGGTGGCATG; HPRT, sense, 5′-GCTGGTGAAAAGGACCTCT-3′ and antisense, 5′-CACAGGACTAGAACACCTGC-3′; EBF, sense, 5′-GCCCGGGCTCACTTTGAGAAGCAGC-3′ and antisense, 5′-GCTTTTCTTGTCACAACAGCG-3′; CD19, sense, 5′-TACCTTAGTAGGAGGCAGGCC-3′ and antisense, 5′-ATAGTGGGGAGATGTTCCGG-3′.

Q-PCR—RNA was prepared using Trizol (Invitrogen). Total RNA cleanup including DNase treatment was performed with the RNeasy Micro kit (Qiagen). Reverse transcription was performed as described above (see RT and PCR). Q-PCR were performed using 2× TaqMan PCR master mixture (Applied Biosystems) and 20× assay-on-demand TaqMan probes (Applied Biosystems) in a total volume of 20 μl. TaqMan probes used for Q-PCR were: CD19, Mm00515420_m1; Mb-1, Mm00432423_m1; and HPRT, Mm00446968_m1 (all from Applied Biosystems).

Isolation and Purification of Bone Marrow Progenitors and Mature Peripheral B-cells—Bone marrow cells were sorted on a FACS-Vantage Cell Sorter (BD Biosciences, San Jose, CA), equipped with a 488-nm argon ion (Coherent Enterprise II, Santa Clara, CA) and a 633-nm He-Ne (model 127, Spectra-Physics, Mountain View, CA) laser. Antibodies used were B220-APC, CD43-PE, IgM-biotin (streptavidin TRI), CD19-fluorescein isothiocyanate, and CD138 (Syndecan)-PE all from Pharmingen. The purity of all sorted cell populations is reproducible over 95%.

RESULTS

Pearson Correlation Analysis Allows for the Identification of EBF Target Genes—Making the presumption that genes regulated in a similar fashion would have correlating expression patterns, we decided to attempt to define genetic targets for EBF using a data set generated from microarray analysis of 14 mouse B lineage cell lines representing different stages of development (22). These cell lines displayed a wide spectra of EBF expression levels and the relative amount of RNA, as estimated by the microarray analysis, appeared to correlate well to the amount of EBF DNA binding activity observed using nuclear extracts from the same cell lines (data not shown). To investigate the correlations between genes in the data set and EBF, we decided to use standard Pearson correlation analysis that results in a numerical value for how well alterations in expression levels of two genes correlates. This type of analysis will generate a wide variety of correlations between different genes, so to obtain information about the levels of correlations that can be expected for any given gene, either by chance or by true correlation, we calculated the correlation values for the whole data set, *i.e.* ~156 million correlations. We then plotted the average number of correlations $\geq X$ (X being the Pearson correlation values between −1 to +1) found for all the genes in the data set (Fig. 1). This shows that any given gene in average only displays a correlation ≥0.9 to three other genes, *i.e.* about 1 in 4000. Thus, Pearson correlation values above or around 0.9 would only rarely be seen as random events and could therefore indicate a biologically relevant link. To investigate if we would be able to use this approach to define target genes for EBF we performed the same analysis focusing on the correlations obtained between genes within the data set and EBF. The number of correlating genes were plotted into a diagram and compared with the average numbers previously calculated (Fig. 1). The analysis suggested that EBF expression correlated significantly better with a set of genes than what would be expected for an average gene. Fifteen genes were found to correlate with a value >0.9 and another 15 with a value above 0.85, as compared with the predicted value of around 9. Investigating the identity of these highly correlated mRNAs revealed that they included λ*5*, *VpreB1*, and *mb-1* transcripts, all suggested EBF target genes (5, 31). This shows that from the 12,000 genes we were able to mathematically extract 3 of 6 known/suggested EBF target genes among the top 15 (0.12%). The other proposed EBF targets *B29* (33), *Blk* (34), and *BSAP* (35) were found to display a correlation of 0.81, 0.26, and 0.56, respectively, suggesting that the expression of these genes is not closely linked to the expression of EBF in this data set.

To further investigate the highly correlated genes we postulated that to be a genuine genetic target there should be a high dependence of EBF for the expression of the potential target gene. That is, a doubling in EBF expression levels should result in at least a doubling of target gene expression. Thus, by plotting the expression value of genes displaying a high Pearson correlation with EBF toward the levels of EBF we should obtain a graph with a rather steep slope (high K_i). Calculating the K_i value for all the genes in the data set and blotting the correlation values on the x axis and the K_i values on the y axis (Fig. 2, Table I), it became apparent that the majority of the genes that displayed a high K_i were found in the group with a high Pearson correlation value. Among these were λ*5*, *VpreB1*, *mb-1*, *VpreB3*, *IgM*, and *CD19* (Fig. 2). This indicates that all the known EBF target genes with high correlations displayed a high dependence of EBF, whereas some other highly correlated genes did not. One gene that fulfilled the criteria for being a potentially new EBF target was the *VpreB3* gene (36). The

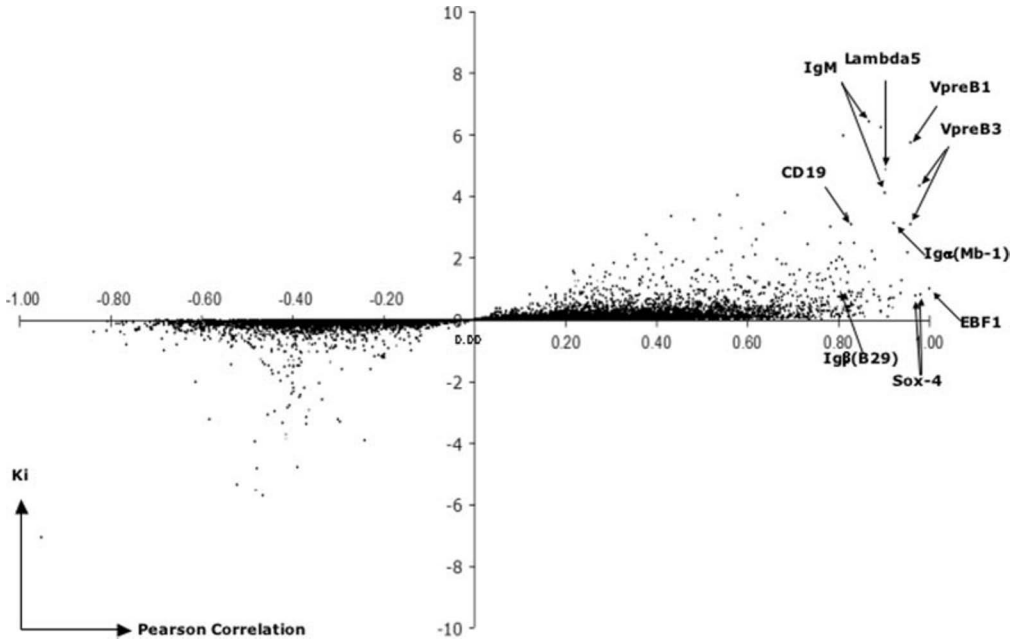

FIG. 2. **Genes with a high dependence of EBF can be identified by estimation of the graphical slope.** The figure displays the Pearson correlation values for EBF against the whole data set on the *x* axis and the obtained slope constant (K_i) on the *y* axis as indicated. The obtained values from all genes in the data set are displayed as *individual dots* and some suggested or potential EBF target genes are indicated.

TABLE I

The table presents the genes showing the highest Pearson correlation with EBF according to the expression patterns as assessed by microarray analysis

For each gene the corresponding K_i value is indicated.

Pearson correlation value	K_i value	Gene annotation
0.98	0.80	SRY-box containing gene 4 (Sox-4)
0.98	4.34	Pre-B lymphocyte gene 3 (VpreB3)
0.97	0.78	SRY-box containing gene 4 (Sox-4)
0.96	5.75	Pre-B lymphocyte gene 1 (VpreB1)
0.95	2.20	Sorting nexin 2
0.94	1.30	Epithelial membrane protein 1
0.92	3.14	Immunoglobulin-associated α (mb-1)
0.92	0.71	GRO1 oncogene
0.92	0.38	Ena-vasodilator stimulated phosphoprotein
0.92	1.08	RIKEN cDNA 1110061M03 gene
0.92	0.72	Glutaryl-coenzyme A dehydrogenase
0.91	0.46	DNA segment, Chr 14, ERATO Doi 813
0.91	0.65	DOPA decarboxylase
0.91	0.10	Vascular endothelial growth factor C
0.91	4.88	Immunoglobulin λ chain 5 (λ5)
0.90	4.13	Immunoglobulin heavy (heavy chain of IgM)
0.89	0.26	Ena-vasodilator stimulated phosphoprotein
0.89	0.25	Metal response element binding transcription factor 2
0.89	6.24	Purine-nucleoside phosphorylase
0.89	0.62	Expressed sequence AU046135
0.89	0.74	Immunoglobulin heavy chain 4 (serum IgG1)
0.88	1.90	Deoxycytidine kinase
0.88	0.81	Interferon-activated gene 202A
0.88	0.97	myc box-dependent interacting protein 1
0.87	2.25	CD19 antigen
0.87	1.22	Lymphocyte antigen 68
0.86	0.61	Interleukin 17 receptor
0.86	1.67	Calponin 3, acidic
0.86	0.20	Bone marrow stromal cell antigen 1

expression of this gene has not been extensively studied so to investigate the expression pattern in primary bone marrow B cells we sorted B220+, CD43+, IgM− pro-B cells, B220+, CD19+, IgM− pre-B cells, B220+, IgM+, CD19+ B cells, and CD19− Syndecan+ plasma cells and performed RT-PCR analysis of the RNA content of these cells (Fig. 3). EBF was expressed at the highest level in the pre-B cell population, whereas some expression was detected also among the B and plasma cell populations. The peak of expression of *λ5* and *VpreB3* was also observed at the pre-B cell stage but while the λ5 message only could be detected in the pre-B cells, *VpreB3* expression followed that of EBF also in primary cells. Thus, our mathematical analysis of microarray data suggests that the expression levels of a limited but well defined set of genes are correlated to the expression of EBF in B cell development.

EBF Shares Several Target Genes with E47—To investigate if the promoters of the potential target genes we defined by the Pearson/K_i analysis had the ability to interact with EBF, we performed EMSA analysis using the *mb-1* promoter EBF site (25) and recombinant *in vitro* translated EBF protein (Fig. 4*A*).

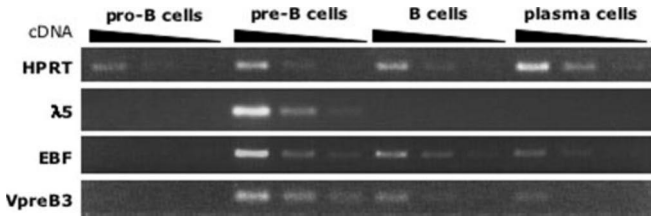

FIG. 3. **RT-PCR analysis of the VpreB3 expression pattern in bone marrow B cells supports a correlation with EBF levels.** The panels show ethidium bromide-stained agarose gels of PCR products obtained from 5-fold serial dilutions of cDNA. Primary cells from different stages of B cell development were isolated and sorted from bone marrow using antibodies against surface markers. RNA was extracted and after cDNA synthesis RT-PCR was performed to detect the levels of *HPRT*, *λ5*, *EBF*, and *VpreB3* transcripts as indicated.

The ability of the target promoters to compete for complex formation between the recombinant protein and the binding site was then assayed by the inclusion of PCR products obtained by amplification using the promoters cloned in a pGL3 vector and primers directed against vector sequences flanking the promoters or the polylinker in the control vector (Fig. 4*A*). The *λ5* and *mb-1* promoters competed as expected for complex formation, whereas the amplified pGL3-polylinker did not. The *VpreB1* promoter has been suggested to be a direct target for EBF (37) and this control element competed for complex formation although with a low efficiency. Also the *CD19* promoter competed for complex formation suggesting that this promoter contains a binding site for EBF. There have been no reports concerning the characterization of the *VpreB3* promoter but the GenBank™ registration of three full-length enriched cDNA sequences (NM009514, AK008794, and BL062250) allowed for the identification of a putative promoter 5′ of the coding gene (36). We cloned this putative promoter into a luciferase reporter plasmid after PCR amplification of genomic DNA and subsequent transfections of reporter constructs under control of this element suggested that it displayed a weak promoter activity in pre-B and B lineage cell lines (data not shown). Using a PCR product from this putative *VpreB3* promoter as competitor resulted in a reduced formation of the labeled complex suggesting that the promoter was able to compete for binding to recombinant EBF protein. Because *VpreB3* and *CD19* represent novel target genes, we investigated their promoter sequences and tested potential EBF binding sites in competition experiments as above (Fig. 4*A*). This allowed us to define EBF binding sites in both the *VpreB3* (VpreB3:EBF) and the *CD19* (CD19:EBF) promoters supporting the idea that they are direct targets for EBF. The limited ability of the *VpreB1* promoter to compete for complex formation made us examine the genomic surroundings of this gene searching for potential additional EBF binding sites. One of these presumed sites located in the *VpreB1* intron was able to interact with EBF (VpreB1:intron, Fig. 4*A*), possibly contributing to the EBF responsiveness of this gene.

The *λ5, VpreB1*, and *mb-1* promoters have been shown to be direct targets for E-proteins (E2A, Heb, and E2–2) (38, 39) as well as EBF (23, 31) and a collaboration between these proteins in gene regulatory events in the early B cells has been supported both *in vitro* (23, 31, 32) and *in vivo* (35). Inspection of the DNA sequence of the *VpreB1*, *CD19*, and *VpreB3* promoters revealed that they all contained E-boxes (CANNTG) (41) suggesting that these genes could be targets for E-proteins such as E47. To investigate if E47 was able to interact with these promoters, we competed for complex formation of recombinant E47 and an E-box from the mouse immunoglobulin heavy chain enhancer (μE5) by the addition of PCR amplified promoter elements (Fig. 4*C*). The addition of the promoter DNA from the *mb-1*, *λ5*, *VpreB1*, *VpreB3*, or *CD19* genes resulted in competition for complex formation, whereas the addition of amplified polylinker did not. This suggests that all the promoters are able to interact with E47 *in vitro*. By competition experiments using oligonucleotides spanning specific E-boxes, we were also able to define functional E-boxes in the *VpreB1*, *CD19*, and *VpreB3* promoters as well as 3′ of the *VpreB1* gene. These data support the idea that several of the genes defined as dependent of EBF from the correlation analysis contain binding sites for E47 and thus might represent targets for coordinated action of these transactivators.

Stable Ectopic Expression of E47 and EBF Induces Expression of Pre-B Cell-restricted Target Genes—To investigate the ability of EBF and E47 to induce B lineage genes in a hematopoetic progenitor cell, we made stably transfected Ba/F3 cells ectopically expressing these transcription factors. Ba/F3 cells represent an interleukin-3-dependent hematopoetic progenitor cell (42) with low expression of B-lineage genes such as *B29* and *mb-1* but without any expression of EBF or surrogate light chain genes (31). These progenitor cells express E47 as judged by Western blot (23), but none of the B lineage restricted the E47 homodimer BCF1 as judged by EMSA analysis (Fig. 5) (23). The lack of BCF1 is believed to be a result of lineage-restricted post-translational modifications of the E47 protein (43), so to achieve the formation of this protein complex in the progenitor cell, we utilized a forced-dimer of E47 (31) to generate stably transfected Ba/F3 cells expressing both BCF1 and EBF. One of these clones has been reported earlier (31), whereas three more (clones 3, 5 and 6, Fig. 5) were newly established by co-transfection of a neomycin resistance encoding E47FD (31) and a puromycin resistance encoding EBF expression plasmid (30). The cells were double selected and clones were screened for protein expression by EMSA using an *mb-1* promoter EBF site or an IgH intron enhancer E-box (μE5) (31) (Fig. 5). Screening of 60 clones resulted in three with significant expression of both E47FD and EBF as judged by EMSA (Fig. 5). RNA was prepared from these cells in parallel with the production of the nuclear extracts and after cDNA synthesis and *in vitro* transcription used to hybridize an Affymetrix U74Av2 chip. As reference we used two selected neo/pyro-resistant lines and one parental Ba/F3 clone. The obtained data were analyzed by treating the different measurements within the control or transfected groups as replicas to reduce the impact of clone-specific features, and the average alterations in expression levels of genes classified as present in at least three of the clones were investigated. This suggested a more than 3-fold increase in expression of 29 genes and a similar decrease in the expression of 15 genes (Table II and Supplementary Materials). Among the regulated genes were several of those we predicted to be genetic targets for EBF and E47 from the correlation analysis. These included *λ5* and *VpreB1*, where two probe sets for each gene detected induction and also *mb-1, Blk*, and IgM were modestly induced in the transfected cells. To verify the induction of some presumed target genes we performed a RT-PCR analysis of the newly generated clones (Fig. 6, data not shown). This indicated that *VpreB1* and *λ5*, as expected, were induced to a low level in EBF-transduced cells, whereas no expression was detected in E47FD expressing clones. The expression was, however, more robust in the cells expressing both EBF and E47FD supporting the idea of a functional synergy between these proteins. The same profile was observed when we investigated the expression of the *VpreB2* gene. We could also detect an increased expression of *VpreB3*, even though we were unable to detect this in the microarray analysis. However, in contrast to *λ5* and *VpreB1*, we detected a basal expression of *VpreB3* in the pa-

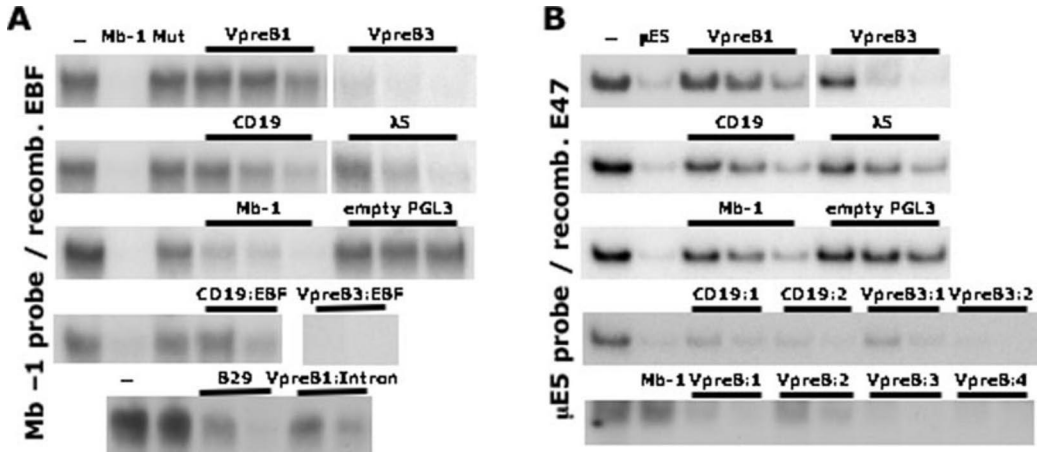

FIG. 4. **Promoters of presumed EBF target genes contain EBF and E47 binding sites.** *Panel A* displays EMSAs where the complex formation between recombinant EBF and the EBF binding sites from the mouse *mb-1* promoter is competed for either by the addition of either the indicated PCR amplified promoter (50-, 150-, or 450-fold excess) elements or duplex oligonucleotides (150- or 500-fold excess). *Panel B* displays EMSAs where the complex formation between the μE5 E-box and recombinant E47 is competed for by the addition of promoters or duplex oligonucleotides as in *panel A*. The gels are representative of at least three competition experiments and the gels have been cut to display only the protein-DNA complexes.

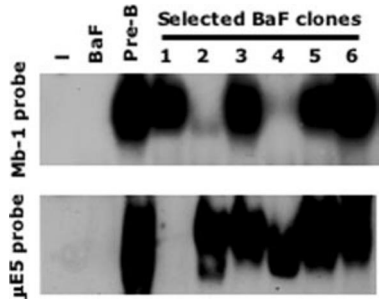

FIG. 5. **EBF- and E47-transfected cells express significant amounts of EBF and BCF1 proteins.** The figure shows the complex formation between 5 μg of nuclear proteins and either the *mb-1* promoter EBF site or the immunoglobulin heavy chain enhancer μE5 E47 binding site. *BaF/3* represents the parental cell line and the 7023 pre-B cell line. The gels have been cut to display only the protein-DNA complexes.

rental BaF3 cells. We were also able to verify a modest up-regulation of *mb-1* expression by real time PCR (Fig. 6*B*). Even though a set of pre-B cell-specific genes were induced, we were unable to detect expression of *Rag-1*, *Pax-5*, *CD19*, or *TdT* by PCR (Fig. 7 and data not shown) in any of the clones. Thus, we suggest that EBF has the ability to induce expression of several defined target genes and that the function of EBF in early B cell development is highly dependent of the presence of BCF1.

EBF/E47 and Pax-5 Cooperate in the Induction of the CD19 Gene—Even though we were able to find both EBF and E47 binding sites in the *CD19* promoter we were unable to detect any expression of *CD19* message in our stably transfected clones. The expression of *CD19* has been shown to be highly dependent on Pax-5 (12, 13) and to investigate if EBF/E47 and Pax-5 coordinately were involved in the regulation of the *CD19* gene, we transduced two of our stably transfected EBF/E47FD expressing BaF/3 clones and one E47FD expressing clone with a Pax-5/green fluorescent protein encoding retrovirus. The expression of the transcription factors after 7 days of cultivation was then assayed by EMSA (Fig. 7*A*). RT-PCR analysis suggested that Pax-5 induced a low expression of the *CD19* gene but cells expressing all three proteins displayed a more robust *CD19* expression (Fig. 7*B*). This was totally dependent of Pax-5 because control virus-infected cells from the same clones did not express any detectable amounts of *CD19* message even after 40 cycles of PCR (Fig. 7, *B* and *C*). To get a more accurate quantification of the induction levels we performed real time PCR analysis of *CD19* expression (Fig. 7*C*). This revealed that the expression levels were 10–12-fold higher in the cell lines expressing EBF and E47 together with Pax-5 indicating that the *CD19* gene may be a target for coordinated activation by Pax-5 and EBF/E47.

TABLE II
This table displays a selection of genes up regulated in BaF/3 cells expressing EBF and E47FD

The induction value is based on the average increase in expression in four independently generated stably transfected clones as compared to the level of tree control samples. For a gene to be included in the analysis we demanded present calls in at least three of the samples unless indicated by an asterisk (*).

Fold induction	Gene annotation
10.0	Immunoglobulin λ chain 5 (λ5)
9.0	Immunoglobulin λ chain 5 (λ5)
5.6	CEA-related cell adhesion molecule 1
4.6	Pre-B lymphocyte gene 1 (VpreB1)*
4.4	Casein γ
4.0	Hemoglobin α, adult chain 1
3.8	FBJ osteosarcoma oncogene B
3.8	Glycoprotein 9 (platelet)
3.7	Neurogranin (protein kinase C substrate, RC3)
3.6	Interleukin 6
3.6	Hemoglobin X, α-like embryonic chain in Hba complex
3.5	Microtubule-associated protein tau
3.3	FBJ osteosarcoma oncogene B
3.3	CEA-related cell adhesion molecule 2
3.2	CEA-related cell adhesion molecule 1
3.2	Pre-B lymphocyte gene 1 (VpreB1)
3.2	Adenylate cyclase 4
3.2	Heat shock 70-kDa protein-like 1
3.2	Laminin, γ 2
3.1	Clusterin
3.0	LIM only 2 (LMO2)
2.8	LIM only 2 (LMO2)
2.5	Inhibitor of DNA binding 2 (Id-2)
2.2	Protein kinase C α
2.2	B lymphoid kinase (BLK)
2.2	Immunoglobulin-associated α (Igα/Mb-1)
2.2	IgM (germ line gene fragment for μ-immunoglobulin)
2.2	IgM heavy chain (immunoglobulin heavy chain 6)

DISCUSSION

EBF is a crucial transcription factor in early B cell development and we here report of the identification of additional genetic targets for this protein. The identification of the *VpreB2*, *VpreB3*, and *CD19* genes as targets for EBF and E47

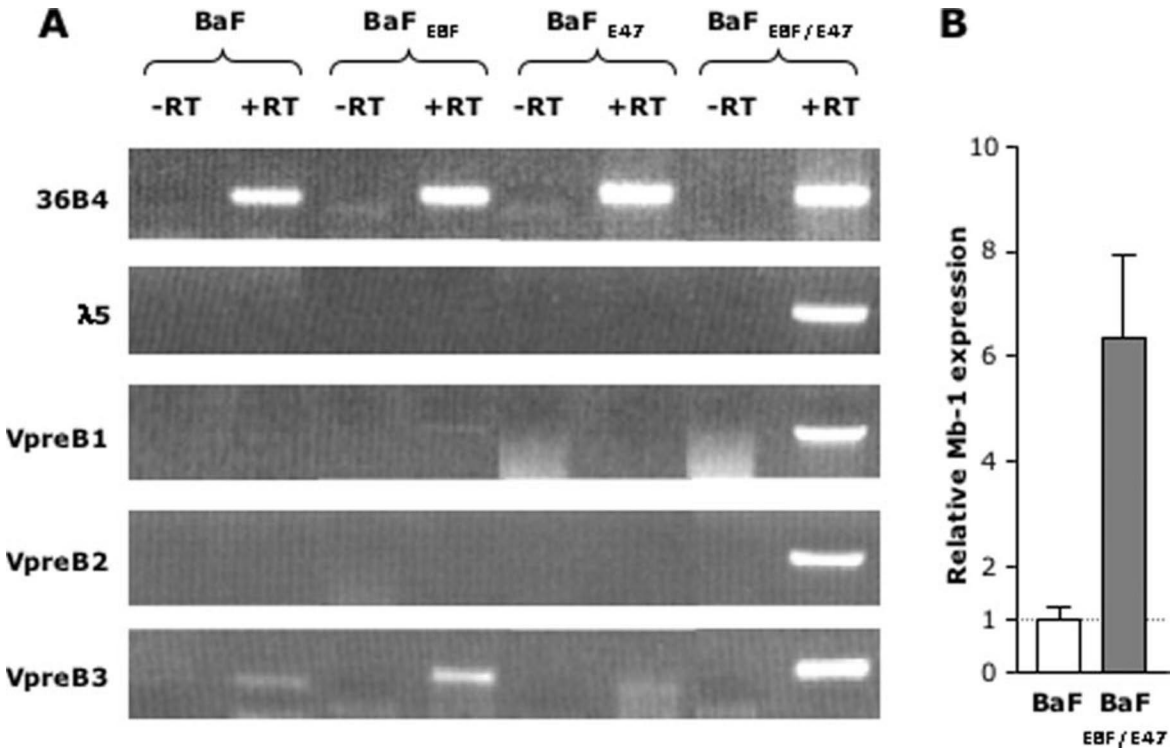

FIG. 6. **Ectopic expression of EBF and E47 induces a subset of pre-B cell restricted genes.** *Panel A* shows ethidium bromide-stained agarose gels with the resulting PCR products after amplification of either the control gene 36B4 or potential EBF target genes using cDNA from control or EBF-, E47-, or EBF/E47-transfected BaF/3 cells as indicated. The *−RT* reaction contains the input material that has not been converted into cDNA by addition of reverse transcriptase. The experiment is representative for several independent RT-PCR experiments and at least two clones of each type were analyzed with comparable results. *Panel B* displays a real time PCR analysis of *mb-1* expression in parental or EBF/E47 expressing BaF/3 cells.

confirms that these two proteins often act in concert. This has now been shown for several genes in cell line systems (23, 31, 44, 45) but also by *in vivo* studies of mice that are transheterozygote for mutations in the EBF and E47 encoding E2A genes (35). These mice display a more pronounced impairment of B cell development than mice carrying either of the heterozygote mutations. The exact molecular mechanisms of this functional synergy has not yet been resolved even though some studies indicate an increased stability of the DNA binding complexes in the presence of both factors *in vitro* (23, 32). *λ5* and *VpreB1* are encoded from genes closely linked in the genome (46) and the coordinate activation of these genes could be because of the activation of common regulatory elements. However, the coordinated activation of the *VpreB2* gene, highly homologous to the *VpreB1* gene but located in another position at the same chromosome (46, 47), suggests that *VpreB1* and *VpreB2* are both regulated in a similar manner by control elements in the conserved regions. This was further supported by the identification of both conserved E47 and EBF binding sites also in the intron and 3′ of the coding gene. We could also detect modulation of expression levels of a set of other genes, some of which were dependent on EBF alone, whereas the majority of these responded to the coordinated expression of the transcription factors (data not shown, Table II). The majority of the genes were not associated with the pre-B cell stage and even though this does not exclude that they indeed are target genes for EBF and E47 we have not investigated them further. One gene that could be of potential interest is the *CEACAM1* gene that was up-regulated in the stably transfected cells (Table II). In contrast to what could be observed for most of the other genes, CEACAM1 expression was enhanced also in the absence of E47FD (data not shown). Even though this gene is broadly expressed, this surface protein has been shown to enhance B cell receptor signaling (48) and thus, may play a role in B-cell development. Even though it is likely that the function of EBF and E47 can be affected by both the transcription factor context and by the epigenetic status of the parental cell line, we believe that these two proteins themselves induce a limited set of genes in the developing B lymphocyte. Another potential role for EBF and E47 could be to aid Pax-5 in the activation of pre-B cell restricted target genes. Several lines of investigation, including studies in Pax-5-deficient pro-B cells, support the idea that *CD19* expression is critically dependent on Pax-5 (12, 13). Our studies in BaF/3 cells support the idea of a crucial role for Pax-5 but also provide some data suggesting that EBF and E47 may support the activity of Pax-5 on its target genes. This would then provide another important mechanism in B cell differentiation where early acting factors provide an environment for increased activity of the commitment factor Pax-5.

We also report that putting numerical values on the relative correlations between a transcription factor and a whole gene set using data generated by microarray analysis, can aid in the identification of genetic targets for the protein. Using this type of mathematical analysis will to a certain degree produce false correlations both because of random events and simply overlapping expression patterns but the numbers of randomly generated positive correlations appear to decrease dramatically around 0.9. It also appears as the K_i value can be used to further reduce the number of potential target genes down to reasonable numbers. To obtain useful information from this approach it demands that the transcription factor on which we focus acts in a dominant manner so to create a high correlation and the usefulness of the method will be highly dependent on the factor under observation. Correlation analysis of E47 expression *versus* the other genes in the data set was, for instance, uninformative because we were unable to observe any strong correlations to the expression of the E2A mRNA. This

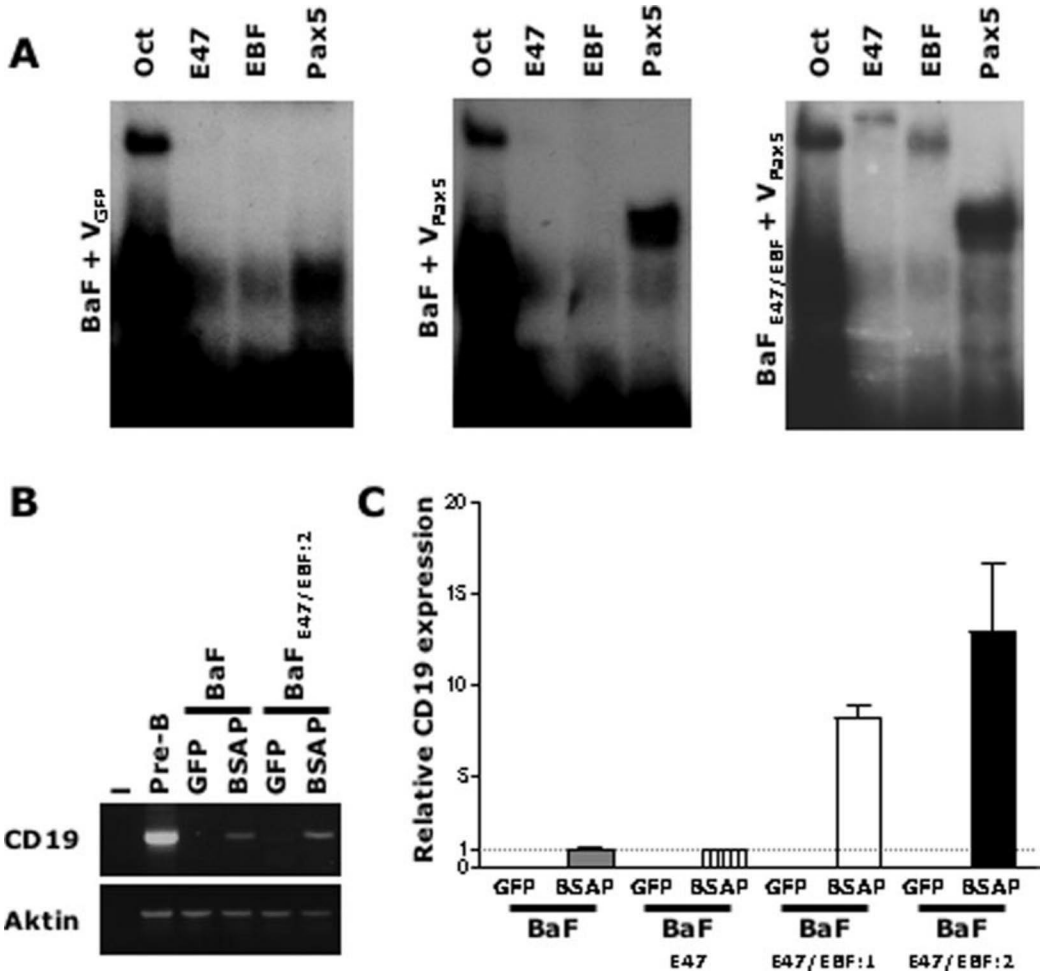

FIG. 7. **The expression of the *CD19* gene is critically dependent on expression of Pax-5.** *Panel A* displays EMSAs where nuclear extracts from transfected/infected cells have been assayed for the expression of Oct, EBF, E47, and Pax-5 proteins. *Panel B* shows ethidium bromide-stained agarose gels where the expression of the CD19 message has been analyzed by 40 cycles of PCR in the presence or absence of Pax-5. The diagram in *panel C* shows one representative real time PCR experiment of transfected/transduced cells as indicated using triplicate reactions.

might be a result of the fact that E47 is a part of a network involving both repressors and redundant activators of transcription (38). The functional relevance of this network in B cell development has been established both in transgenic mice models where expression of the E47 inhibitor Id-1 in pro-B cells resulted in impaired B cell development (49) and in mice lacking combinations of E-proteins (18, 19, 50, 51). The role of a specific factor in such networks will be hard to elucidate by correlation analysis unless several dimensions of the data set are analyzed simultaneously. The K_i value could also provide information about the amplification loops within a defined genetic system as well as of genes upstream of the transcription factor (low K_i). Optimal usage of the large amounts of information that are generated in microarray experiments represents a large challenge and even though the method we describe in this report has limitations and may not be useful in all model systems, it allowed us to isolate three known and two novel genetic targets for EBF in the top 0.12% of the investigated genes. This provides proof of concept and it is likely that this approach can be used also in other model systems to provide clues of genetic networks and transcription factor target genes. Even though we here report of additional EBF target genes, none of those defined to date are likely to explain either the complete lack of mature B cells in EBF-deficient mice (17) or the apparent ability of EBF to reduce T cell development (40) and a continued search for additional genetic targets will be needed to understand the full function of this factor.

Acknowledgments—We thank the Swegen microarray facility for help with chip hybridizations, personal at the Department for Stem Cell Biology for help with cell sorting, and members of the Department for Complex Systems for helpful discussions.

REFERENCES

1. Ghia, P., ten Boekel, E., Rolink, A. G., and Melchers, F. (1998) *Immunol. Today* **19,** 480–485
2. Hoffmann, R., and Melchers, F. (2003) *Curr. Opin. Immunol.* **15,** 239–245
3. Schebesta, M., Heavey, B., and Busslinger, M. (2002) *Curr. Opin. Immunol.* **14,** 216–223
4. Klemsz, M. J., McKercher, S. R., Celada, A., van Beveren, C., and Maki, R. A. (1990) *Cell* **61,** 113–124
5. Hagman, J., Belanger, C., Travis, A., Turck, C. W., and Grosschedl, R. (1993) *Genes Dev.* **7,** 760–773
6. Murre, C., McCaw, P. S., and Baltimore, D. (1989) *Cell* **56,** 777–783
7. Kozmik, Z., Wang, S., Dorfler, P., Adams, B., and Busslinger, M. (1992) *Mol. Cell. Biol.* **12,** 2662–2672
8. Adams, B., Dorfler, P., Aguzzi, A., Kozmik, Z., Urbanek, P., Maurer-Fogy, I., and Busslinger, M. (1992) *Genes Dev.* **6,** 1589–1607
9. Mikkola, I., Heavey, B., Horcher, M., and Busslinger, M. (2002) *Science* **297,** 110–113
10. Nutt, S. L., Heavey, B., Rolink, A. G., and Busslinger, M. (1999) *Nature* **401,** 556–562
11. Souabni, A., Cobaleda, C., Schebesta, M., and Busslinger, M. (2002) *Immunity* **17,** 781–793
12. Nutt, S. L., Urbanek, P., Rolink, A., and Busslinger, M. (1997) *Genes Dev.* **11,** 476–491
13. Urbanek, P., Wang, Z. Q., Fetka, I., Wagner, E. F., and Busslinger, M. (1994) *Cell* **79,** 901–912
14. McKercher, S. R., Torbett, B. E., Anderson, K. L., Henkel, G. W., Vestal, D. J., Baribault, H., Klemsz, M., Feeney, A. J., Wu, G. E., Paige, C. J., and Maki, R. A. (1996) *EMBO J.* **15,** 5647–5658
15. Olson, M. C., Scott, E. W., Hack, A. A., Su, G. H., Tenen, D. G., Singh, H., and Simon, M. C. (1995) *Immunity* **3,** 703–714
16. Scott, E. W., Simon, M. C., Anastasi, J., and Singh, H. (1994) *Science* **265,** 1573–1577

17. Lin, H., and Grosschedl, R. (1995) *Nature* **376,** 263–267
18. Zhuang, Y., Soriano, P., and Weintraub, H. (1994) *Cell* **79,** 875–884
19. Bain, G., Maandag, E. C., Izon, D. J., Amsen, D., Kruisbeek, A. M., Weintraub, B. C., Kroop, I., Schlissel, M. S., Feeney, A. J., van Roon, M., van der Valk, M., te Riel, H. P. J., Berns, A., and Murre, C. (1994) *Cell* **79,** 885–892
20. Quong, M. W., Romanow, W. J., and Murre, C. (2002) *Annu. Rev. Immunol.* **20,** 301–322
21. Liberg, D., Sigvardsson, M., and Akerblad, P. (2002) *Mol. Cell. Biol.* **22,** 8389–8397
22. Tsopogas, P., Breslin, T., Bilke, S., Månsson, R., Lagergren, A., Liberg, D., Peterson, C., and Sigvardsson, M. (2003) *J. Leukocyte Biol.* **74,** 102–110
23. Sigvardsson, M., Clark, D. R., Fitzsimmons, D., Doyle, M., Åkerblad, P., Breslin, T., Bilke, S., Liu, Y., Yeamans, C., Zhang, G., and Hagman, J. (2002) *Mol. Cell. Biol.* **22,** 8539–8551
24. Fitzsimmons, D., Hodsdon, W., Wheat, W., Maira, S. M., Wasylyk, B., and Hagman, J. (1996) *Genes Dev.* **10,** 2198–2211
25. Hagman, J., Travis, A., and Grosschedl, R. (1991) *EMBO J.* **10,** 3409–3417
26. Travis, A., Hagman, J., and Grosschedl, R. (1991) *Mol. Cell. Biol.* **11,** 5756–5766
27. Feldhaus, A. L., Mbangkollo, D., Arvin, K. L., Klug, C. A., and Singh, H. (1992) *Mol. Cell. Biol.* **12,** 1126–1133
28. Li, C., and Wong, W. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 31–36
29. Schreiber, E., Matthias, P., Müller, M., and Schaffner, W. (1989) *Nucleic Acids Res.* **17,** 6419
30. Akerblad, P., Lind, U., Liberg, D., Bamberg, K., and Sigvardsson, M. (2002) *Mol. Cell. Biol.* **22,** 8015–8025
31. Sigvardsson, M., O'Riordan, M., and Grosschedl, R. (1997) *Immunity* **7,** 25–36
32. Sigvardsson, M. (2000) *Mol. Cell. Biol.* **20,** 3640–3654
33. Akerblad, P., Rosberg, M., Leanderson, T., and Sigvardsson, M. (1999) *Mol. Cell. Biol.* **19,** 392–401
34. Akerblad, P., and Sigvardsson, M. (1999) *J. Immunol.* **163,** 5453–5461
35. O'Riordan, M., and Grosschedl, R. (1999) *Immunity* **11,** 21–31
36. Shirasawa, T., Ohnishi, K., Hagiwara, S., Shigemoto, K., Takebe, Y., Rajewsky, K., and Takemori, T. (1993) *EMBO J.* **12,** 1827–1834
37. Persson, C., Martensson, A., and Martensson, I. L. (1998) *Eur. J. Immunol.* **28,** 787–798
38. Massari, M. E., and Murre, C. (2000) *Mol. Cell. Biol.* **20,** 429–440
39. Greenbaum, S., and Zhuang, Y. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 15030–15035
40. Zhang, Z., Cotta, C. V., Stephan, R. P., deGuzman, C. G., and Klug, C. A. (2003) *EMBO J.* **22,** 4759–4769
41. Ephrussi, A., Church, G. M., Tonegawa, S., and Gilbert, W. (1985) *Science* **227,** 134–140
42. Palacios, R., and Steinmetz, M. (1985) *Cell* **41,** 727–734
43. Sloan, S. R., Shen, C. P., McCarrick-Walmsley, R., and Kadesch, T. (1996) *Mol. Cell. Biol.* **16,** 6900–6908
44. Kee, B. L., and Murre, C. (1998) *J. Exp. Med.* **188,** 699–713
45. Romanow, W. J., Langerak, A. W., Goebel, P., Wolvers-Tettero, I. L., van Dongen, J. J., Feeney, A. J., and Murre, C. (2000) *Mol. Cell* **5,** 343–353
46. Kudo, A., and Melchers, F. (1987) *EMBO J.* **6,** 2267–2272
47. Bauer, S. R., D'Hoostelaere, L. A., and Melchers, F. (1988) *Curr. Top. Microbiol. Immunol.* **137,** 130–135
48. Greicius, G., Severinson, E., Beauchemin, N., Obrink, B., and Singer, B. B. (2003) *J. Leukocyte Biol.* **74,** 126–134
49. Sun, X.-H. (1994) *Cell* **79,** 893–900
50. Zhuang, Y., Cheng, P., and Weintraub, H. (1996) *Mol. Cell. Biol.* **16,** 2898–2905
51. Zhuang, Y., Barndt, R. J., Pan, L., Kelley, R., and Dai, M. (1998) *Mol. Cell. Biol.* **18,** 3340–3349