Gene Ontology-driven transcriptional analysis of CD34+ cell-initiated megakaryocytic cultures identifies new transcriptional regulators of megakaryopoiesis

Peter G. Fuhrken,1* Chi Chen,2* Pani A. Apostolidis,1 Min Wang,2 William M. Miller,1,2,3 and Eleftherios T. Papoutsakis1,3,4

1Department of Chemical and Biological Engineering, 2Interdepartmental Biological Sciences Program, and 3Robert H. Lurie Comprehensive Cancer Center, Northwestern University, Evanston, Illinois; and 4Department of Chemical Engineering and Delaware Biotechnology Institute, University of Delaware, Newark, Delaware

Submitted 14 June 2007; accepted in final form 29 January 2008

Fuhrken PG, Chen C, Apostolidis PA, Wang M, Miller WM, Papoutsakis ET. Gene Ontology-driven transcriptional analysis of CD34+ cell-initiated megakaryocytic cultures identifies new transcriptional regulators of megakaryopoiesis. Physiol Genomics 33: 159–169, 2008. First published February 5, 2008; doi:10.1152/physiolgenomics.00127.2007.—Differentiation of hematopoietic stem and progenitor cells is an intricate process controlled in large part at the level of transcription. While some key megakaryocytic transcription factors have been identified, the complete network of megakaryocytic transcriptional control is poorly understood. Using global gene expression microarray analysis, Gene Ontology-based functional annotations, and a novel interlineage comparison with parallel, isogenic granulocytic cultures as a negative control, we closely examined the mRNA level of transcriptional regulators in megakaryocytes derived from human mobilized peripheral blood CD34+ hematopoietic cells. This approach identified 199 differentially expressed transcription factors or transcriptional regulators. We identified and detailed the transcriptional kinetics of most known megakaryocytic transcription factors including GATA1, FLI1, and MAFG. Furthermore, many genes with transcription factor activity or transcription factor binding activity were identified in megakaryocytes that had not previously been associated with that lineage, including BTEB1, NR4A2, FOXO1A, MEF2C, HDAC5, VDR, and several genes associated with the tumor suppressor p53 (HIPK2, FHL2, and TADA3L). Protein expression and nuclear localization were confirmed in megakaryocytic cells for four of the novel candidate megakaryocytic transcription factors: FHL2, MXD1, E2F3, and RFX5. In light of the hypothesis that transcription factors expressed in a particular differentiation program are important contributors to such a program, these data substantially expand our understanding of transcriptional regulation in megakaryocytic differentiation of stem and progenitor cells.

gene expression; megakaryocytopoiesis; thrombopoiesis

THE UNIQUE PROCESS OF DIFFERENTIATION of hematopoietic stem cells into megakaryocytic (Mk) cells and the subsequent generation of platelets is under complex and strict transcriptional regulation. Gain- and loss-of-function studies and examination of various leukemias have revealed pivotal roles and coordinated control of lineage-specific and universally expressed transcription factors (TFs) in this regulation. Basic methods to study hematopoietic TFs include analysis of cis-regulation of cell-specific genes, isolation and gene targeting during lineage differentiation, and complementary in vitro studies. However, to date few TFs have been associated with megakaryopoiesis and thrombopoiesis, and their molecular action and kinetics of expression are not well understood.

Mk cells go through sequential differentiation stages under the control of multiple TFs as reviewed previously (8, 42, 50). First, bipotent erythrocytic-Mk progenitors become committed Mk progenitors under the regulation of FOG-1 and either GATA1 or GATA2 (5, 58). Mice with germ line absence of FOG-1 show complete ablation of Mk progenitors and have low levels of PF4 and GPIIbα, which indicates deficiency in the Mk lineage (17). While FOG-1 can drive Mk gene expression in the absence of GATA1 (17), this appears to be dependent on the redundant function of GATA2 (5). Evidence also supports a key role for SCL/TAL1 in Mk commitment (20). Fli-1, an Ets TF family member, transactivates the GPIX promoter, and Fli-1−/− mouse embryos exhibit impaired megakaryopoiesis (22). Later, committed Mk progenitors turn into early-maturing megakaryocytes, with positive regulation by RUNX1, TEL, GATA1, and NFE2 (25, 27, 52). GATA1-deficient megakaryocytes consist of hyperproliferating abnormal cells and show reduced expression of Mk-related genes (60). In the absence of NFE2, there is reduced proliferation of early megakaryocytes but normal endomiosis (34). As Mk cells undergo endomiosis and cytoplasmic maturation, GATA1 again plays an essential role with cofactor FOG-1. GATA1-deficient megakaryocytes display impaired cytoplasmic maturation (53). GPIIbα (ITGA2B) and the p45 subunit of NFE2 may be targets for the GATA1/FOG-1 complex (17). Fli-1 also cooperates with GATA1 to synergistically promote expression of Mk-related genes in this stage (11). Late megakaryocytes go through terminal differentiation and extend proplatelets. Platelet release is severely hampered in NFE2 knockout mice (31).

In an effort to expand the set of potential TFs involved in hematopoietic stem and progenitor cell (HSC) differentiation toward the Mk lineage, we have used a Gene Ontology-driven genomic approach whereby several comparisons of temporal microarray data from CD34+ cell-initiated Mk and isogenic granulocytic (G) cultures were employed (1). The key hypothesis is that TFs are differentially expressed in a particular differentiation program only if such TFs are active participants in such a program. This approach is unique among recent Mk cell transcriptional analyses (7, 44) in that it utilizes curated functional associations, based on Gene Ontology terminology

* P. G. Fuhrken and C. Chen contributed equally to this work.

Address for reprint requests and other correspondence: E. T. Papoutsakis, Dept. of Chemical Engineering and Delaware Biotechnology Inst., Univ. of Delaware, Newark, DE 19711 (e-mail: epaps@udel.edu).
(1), to specifically identify potential transcriptional regulators in Mk cells. Earlier microarray-based analyses of Mk differentiation identified stage-specific molecular markers of murine Mk differentiation and expanded the known regulon of the key Mk TF NFE2 (7). Another study compared the transcriptome of culture-derived human Mk cells separated by ploidy class and identified a set of ~400 genes, representing a range of functions that distinguished cells of different ploidy classes. In addition, we recently (6) used this same Gene Ontology-based approach to probe the apoptosis-related transcriptome of differentiating Mk cells, successfully identifying both individual genes and pathways involved in Mk apoptosis.

After the identification of several potential sets of transcriptional regulators based on their patterns of expression, we verified the expression patterns by quantitative (Q)-RT-PCR and examined the protein expression and localization of a select set of TFs, which have not been previously associated with Mk differentiation. These and other genes identified in this study are valuable leads for further study of megakaryopoiesis.

MATERIALS AND METHODS

All materials were obtained from Sigma Aldrich except where noted.

Hematopoietic cultures. Mk and G cultures were initiated with fresh (AllCells; Berkeley, CA) or frozen (Hematopoietic Cell Processing Core, Fred Hutchinson Cancer Center, Seattle, WA) granulocyte colony-stimulating factor (G-CSF)-mobilized human peripheral blood CD34+ cells from normal donors. Mk cultures employed X-VIVO 20 serum-free medium (BioWhittaker; Walkersville, MD) supplemented either with a cytokine cocktail including 50 ng/ml thrombopoietin (Tpo; Genentech, South San Francisco, CA), 50 ng/ml Flt-3 ligand (PeproTech, Rocky Hills, NJ), and 5 ng/ml IL-3 (R&D Systems, Minneapolis, MN) or with 100 ng/ml Tpo only. To ensure saturation, Tpo (50 ng/ml) was added to the cultures every 5 days. Mk cultures were seeded at 70,000 cells/ml and maintained for up to 21 days at 37°C in a fully humidified environment containing 5% CO2 and 20% O2. Cells were fed by dilution to maintain the cell density at 1–2 × 10^6 cells/ml to minimize changes in pH and PO2. G cultures were performed in serum-containing human long-term medium (23) supplemented with 50 ng/ml stem cell factor (SCF, R&D Systems), 10 ng/ml IL-6 (PeproTech), 10 ng/ml G-CSF (Amgen, Thousand Oaks, CA), and 10 ng/ml IL-3 (R&D Systems) at 37°C in a fully humidified environment containing 5% CO2 and 5% O2.

Microscopy. Approximately 2 × 10^6 cells from the Mk cultures were washed with PBS-2% BSA and subsequently deposited onto microscope slides by cytocentrifugation (Cytospin 4, Shandon, Pittsburgh, PA). The cells were fixed in a 4% paraformaldehyde (Electron Microscopy Sciences, Ft. Washington, PA) solution in PBS for 10 min, washed in PBS-0.1 M glycine five times for 3 min each, and then permeabilized in a 0.3% Triton X-100 solution in PBS for 5 min. After blocking for 1 h at room temperature with either 10% normal donkey or goat serum (both from Jackson ImmunoResearch, West Grove, PA) in PBS-2% BSA, cells were costained overnight at 4°C with mouse anti-human CD41a monoclonal IgG1k, clone P2 (Beckman Coulter, Fullerton, CA) and goat anti-human FHL2, goat anti-human regulatory factor X5 (RFX5), rabbit anti-human E2F3, or rabbit anti-human MAX dimerization protein 1 (MXD1) polyclonal IgG antibodies (all from Santa Cruz Biotechnology, Santa Cruz, CA). All of the primary antibodies were used at a 1:50 dilution. Cells were washed and then incubated with either 1) fluorescein isothiocyanate (FITC)-conjugated donkey anti-goat IgG (1:100) and Texas red-conjugated rabbit anti-mouse IgG (1:50) secondary antibodies in PBS-2% donkey serum or 2) FITC-conjugated goat anti-rabbit IgG (1:100) and Texas red-conjugated goat anti-mouse IgG (1:50) secondary antibodies in PBS-2% BSA-2% goat serum for 3 h in the dark at room temperature. All secondary antibodies came from Jackson ImmunoResearch. After another washing step, the cells were labeled with Prolong Gold antifade reagent containing DAPI (Molecular Probes, Eugene, OR) and the slides were mounted. All microscopy was performed with a ×63 (numerical aperture = 1.32 in oil) objective on a Leica Model DMIRE2 inverted microscope (Wetzlar, Germany) equipped with a Q Imaging Retiga Exi charge-coupled device camera (Burnaby, BC, Canada) and captured into OpenLab imaging software (Improvison; Lexington, MA).

Flow cytometry. For intracellular detection of GATA1 and NFE2, cells were first stained with FITC-conjugated mouse anti-human CD41a (Beckman Coulter) and then fixed, permeabilized, and stained as described previously (40). Primary antibodies for GATA1 and NFE2, along with isotype-matched controls, were obtained from Santa Cruz Biotechnology, and secondary antibody was obtained from Jackson ImmunoResearch. All flow cytometry data were acquired on a FACScan flow cytometer and analyzed with CellQuest software (BD Biosciences, San Jose, CA).

Purification of CD41a+ cells. Positive selection of CD41a+ cells was performed with MiniMACS MS columns as described previously (6). The purity of the selected cells was assessed by flow cytometry and found to exceed 97% in all cases (data not shown).

Microarray experiments and data analysis. Experiments and analysis were performed as described previously (6). Briefly, cell samples from Mk cytokine cocktail cultures were collected at day 0 for CD34+ cells, days 1, 2, 3, 4, 5, 7, 9, and 12 for Mk cultures, and days 1, 2, 3, 4, 5, 7, 9, and 11 for G cultures, and total RNA was isolated. Sample RNA and Human Universal Reference RNA (Stratagene, La Jolla, CA) were linearly amplified and labeled as described previously (16). Three parallel Mk and G culture experiments were analyzed separately (biological replication), and approximately one-third of the individual microarrays were randomly selected for replication (technical replication). This level of technical replication was previously established (6, 16) based on the highly reproducible microarray outcomes in our laboratory with the Agilent technology (correlation coefficients typically exceeding 0.9; data not shown). Raw microarray data for the Tpo-only cultures were obtained from a previously published study from our laboratory (18) and were processed with the same procedures described below for Mk cocktail cultures and G cultures.

Raw and normalized microarray data were deposited in the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/; Mk cells: GSE3839, G cells: GSE5917). All subsequent data analyses were performed with the MultiExperiment Viewer 3.0 (MeV; Institute for Genomic Research, Rockville, MD) (47). Differentially expressed genes were identified with the statistical analysis of microarrays (SAM; for Mk temporal comparisons) or analysis of variance (ANOVA; for comparison of Mk to G cultures) with a false discovery rate <5% and P < 0.05, respectively. Gene Ontology annotations were retrieved from the Gene Ontology consortium website (www.geneontology.org).

Quantitative RT-PCR. cDNA was obtained from total RNA samples with the High-Capacity cDNA Archive Kit, and Q-RT-PCR was performed with Assays-on-Demand kits (Applied Biosystems, Foster City, CA) as described previously (6, 18). The amount of mRNA for each sample was normalized with the average of two housekeeping genes [glucuronidase-β and ribosomal protein (large) P0] (6, 16, 18). The primer codes were as follows: GUSB (Hs99999908_m1), RPLP0 (Hs99999902_m1), FHL2 (Hs00179935_m1), MXD1 (Hs00231137_m1), E2F3 (Hs00605457_m1), and RFX5 (Hs00230841_m1).
RESULTS

Megakaryocytic and granulocytic differentiation of CD34+ cells. Mobilized peripheral blood CD34+ cells were cultured with a cocktail of cytokines designed to induce both expansion of myeloid progenitor cells and Mk differentiation. As we have previously shown (39, 63), these cells underwent Mk commitment, as assessed by CD41a expression, polyplloidization, and constitutive apoptosis (data not shown). Because this culture system yields a mixed population of cells, samples were further enriched for CD41a+ Mk cells by immunomagnetic selection on days 5–12 (>97% purity) before gene expression analysis. Microarray analysis was performed on uncultured samples (day 0), cultured but unselected samples (days 1–4), and cultured, selected samples (days 5–12). In preliminary experiments (not shown or discussed), we tested for changes in gene expression patterns resulting from the selection process and found no significant differences. Parallel, isogenic G cultures, initiated from the same population of donor cells, were also analyzed as a negative control. These G cultures showed extensive G commitment as assessed by flow cytometry of G markers: on days 5 and 11, the G cell cultures were 40% and 87% CD15+, 26% and 70% CD11b+, and 10% and 65% CD66b+, respectively (data not shown).

Comparative, dynamic analysis scheme for identifying Mk transcription factors and related genes yields four distinct clusters of expression. To study the global gene expression profiles of TFs in Mk differentiation, we generated a list of 939 genes that are represented by 1,103 probes on the Agilent Human1A array and are related to TF activity and TF binding activity (i.e., TF coactivators and repressors) by Gene Ontology classification. To distinguish expression patterns, we used four different comparisons: 1) Mk culture samples from days 1–12 were compared to uncultured CD34+ cells; 2) Mk cultures from days 5–12 were compared to the average from Mk culture samples on days 1–4; 3) Mk samples from days 5–12 were compared to equivalent time point samples from parallel G cultures; 4) G culture samples from days 5–11 were compared to the average from G culture samples on days 1–4. We note that for days 5–12 cells were enriched for CD41a+ cells before RNA extraction, while cells from all other days or the G cultures were not selected before RNA extraction. As we have shown before for the case of Mk-related and apoptosis-related genes (6), these comparisons provide meaningful discrimination in temporal expression patterns: the comparison to day 0 CD34+ cells reveals expression differences between differentiating Mk cells and quiescent progenitors; the comparison to early cells reveals transcriptional differences between late Mk cells and early, actively cycling Mk culture cells; and the comparison to G cells reveals expression differences between Mk and G lineages, which allows the discernment of general hematopoietic differentiation-related genes from those that are Mk specific. We have shown (6) that this comparative analysis accurately captures key, known transcriptional events during Mk differentiation and platelet biogenesis including upregulation of CD41b, CD61, CD62p, and the components of the von Willebrand factor receptor.

Using this approach, we identified a list of 199 TF genes (211 probes) that were differentially expressed in at least one of the comparisons (Fig. 1). Hierarchical clustering revealed distinct expression patterns for these 199 genes and allowed us to divide them into four clusters: cluster A, expression higher in Mk cells compared with unselected early Mk culture cells and G cells and generally higher in Mk cells than in day 0 CD34+ stem and progenitor cells (Fig. 1A); cluster B, late Mk expression—higher expression in late Mk cells compared with unselected early cells and G cells but lower than in CD34+ stem and progenitor cells (Fig. 1B); cluster C, early expression—lower after day 5 in Mk cells compared with unselected early cells but higher compared with CD34+ progenitors (Fig. 1C); cluster D, consistently downregulated in early and late Mk cultures (Fig. 1D).

Exploring transcriptional regulation of known Mk TFs. We first examined the expression patterns of the known Mk-related TFs. The transcriptional patterns of these known Mk TFs have not been systematically reported in primary Mk cultures, and in fact, for many of them, their regulation at the transcriptional level has not been firmly established in such cultures. This analysis thus aimed at exploring the transcriptional regulation of these known Mk TFs. It also aimed at examining the validity of our analytical approach, based on whether we would capture the regulation of those known to be transcriptionally regulated, as well as by select analysis of protein-level expression. As expected, GATA1, TAL1, FLI1, and MAFG were upregulated in these experiments (Fig. 1, A and B), GATA1 and TAL1 showed early upregulation (cluster A), with GATA1 expression preceding that of TAL1 by 1–2 days. This early upregulation agrees with their known role in early megakaryopoiesis and suggests that their expression is a primary transcriptional response to the signaling induced by this cytokine cocktail. Flow cytometric analysis confirmed early and increasing expression of GATA1 protein (Fig. 2), both consistent with the microarray data. In contrast, FLI1 was upregulated only in the selected CD41a+ cells (cluster B, Fig. 1B), suggesting that it is more restricted to the Mk lineage and functions at a later stage than GATA1 and TAL1. Interestingly, MAPG, a component of the NFE2 TF complex in maturing Mk cells, clustered with the early upregulated genes (Fig. 1A). As may be expected, MAFG expression continued to increase after day 5. Other important Mk TFs (FOG1, ETv6, RUNX1, ETS1, and NFE2) were not found to be regulated at the transcriptional level as assayed by our microarray (data not shown). Among these, FOG1, ETv6, and NFE2 were already expressed at very high levels in day 0 CD34+ cells (>90th percentile intensity among all probes), while ETS1 and RUNX1 showed only moderate expression (30–70th percentile; data not shown). Furthermore, in the case of NFE2, we hypothesized that it is not expressed until later in culture after cessation of microarray sampling. Indeed, protein-level expression of NFE2, as assayed by flow cytometry, remained low until day 12, then increased dramatically at later time points, and reached a peak at day 17 (Fig. 2). Only one known Mk TF, MAfK, an alternative NFE2 complex component, was found among the downregulated genes (cluster D; Fig. 1D).

Upregulated TF genes include several TFs not previously associated with Mk cells. Cluster A includes 59 genes that were upregulated specifically in Mk cells compared with CD34+ cells and unselected early cells (Fig. 1A). These genes would be expected to be important in Mk differentiation. Most of these upregulated TFs either have not been associated with megakaryopoiesis or their role in the Mk lineage remains ambiguous. These include FHL2, E2F3 and RFX5, which are
Expression pattern of differentially expressed transcription factor genes during megakaryocytic (Mk) cell differentiation. Expression profiles of transcription factor/transcriptional regulatory genes (per Gene Ontology) that were differentially expressed temporally in Mk cultures and/or between Mk and granulocytic (G) cultures are shown. Genes were divided into 4 clusters (A–D) according to their distinct expression patterns based on hierarchical clustering using the euclidean distance metric. Color denotes degree of differential expression: saturated red, 4-fold upregulation; saturated green, 4-fold downregulation; black, unchanged (see scale). First block shows average expression ratios across the biological replicates ($n=3$) for the designated Mk culture samples compared with day 0 CD34$^+$ cells; 2nd block shows Mk cell expression ratios, averaged across the biological replicates ($n=3$), compared with the average expression from days 1–4; 3rd block shows average Mk cell expression profiles compared with equivalent-day G cells (day 12 for Mk vs. day 11 for G cells) ($n=2$); 4th block shows expression profiles of G cells with respect to average expression of G cells on days 1–4 ($n=2$).
strongly upregulated in Mk cells but not G cells, and MXD1 (also known as MAD1), which is particularly upregulated at later Mk time points and also upregulated, though to a lesser extent, in G cells. We used Q-RT-PCR to verify the expression patterns of four of these TFs, namely FHL2, E2F3, MXD1 (Fig. 3), and RFX5 (data not shown), using cells from separate Mk cultures. We desired at the same time to examine the generality of the Mk TF data, and thus we chose to use Tpo-only Mk stimulation for these cultures. Since Tpo-only cultures produce stronger Mk commitment (higher fraction of CD41a<sup>+</sup>/H11001 cells) (18), we chose to use the cells from these cultures without CD41a<sup>+</sup>/H11001 cell selection. These Q-RT-PCR data show similar (but not quantitatively identical, as would be expected) expression patterns and thus confirm the generality of the microarray data, especially in view of the different Mk stimulation and the lack of CD41a<sup>+</sup>/H11001 cell selection; the latter would lead to a lesser level of apparent differential expression, and thus these Q-RT-PCR data underestimate the level of differential expression in the CD41a<sup>+</sup>/H11001 cells. Q-RT-PCR analysis of RFX5 in unselected cells from these Mk cultures gave lower levels of differential expression (data not shown) likely due to the impact of RFX5 expression in G culture cells (although at considerably lower levels than in Mk culture cells) as well (Fig. 1A). The potential roles of these TFs in Mk cells is further explored below. Another upregulated TF, MEF2C, has been mainly associated with the regulation of muscle differentiation and also plays roles in neuronal survival and T-cell apoptosis (36). MEF2C can be activated by either the p38 or ERK5 pathway (21, 30). Notably, HDAC5, an MEF2C repressor (62), is also highly upregulated (Fig. 1B). HDAC5 belongs to the class IIa histone deacetylase family and inhibits MEF2C transcriptional activity by sumoylation (19). It is not clear from these data, but should be subject to further study, which of these opposing forces would predominate and whether MEF2C activates transcription in Mk cells.

Several genes in *cluster A* are associated with transcriptional activity of the tumor suppressor p53. Homeodomain-interacting protein kinase 2 (HIPK2; Fig. 1A) induces p53 activity (10), and FHL2 (Fig. 1A) can promote p53-induced transcription activation by association with HIPK2 (32). TADA3L (Fig. 1A), which encodes part of the PCAF histone acetylase complex, is required for p53 transcriptional activity and p53-mediated apoptosis (31, 61). Such an increase in p53 transcriptional activity during Mk differentiation has been suggested previously by our group (16) and is presently the subject of further study.

Two other genes in this cluster have been previously linked to Mk cells, but their functional roles remain undetermined. VDR (Fig. 1A) encodes the receptor for 1,25-dihydroxyvitamin D₃, and its signaling induces both differentiation and cell cycle arrest (14, 41). Although VDR expression has been reported in megakaryoblastic cell lines and treatment with 1,25-dihydroxyvitamin D₃ has been shown to induce formation of multilobated nuclei in HiMeg cells (54), the effects of VDR stimulation in cultured primary Mk cells have not, to our knowledge, been reported. CBFA2T3 (Fig. 1A) has been re-

![Fig. 2. Flow cytometric analysis of GATA1 and NFE2 expression during Mk cell cultures. Protein expression level of GATA1 (○) and NFE2 (●) in CD41a<sup>+</sup> cells as represented by geometric mean fluorescence intensity (MFI) of specific antibody-stained samples relative to cells stained with isotype-matched control antibodies. Data shown are from 1 representative experiment (n = 3).](http://physiolgenomics.physiology.org/)

![Fig. 3. Quantitative (Q)-RT-PCR analysis of mRNA expression of the 3 putative Mk transcription factors FHL2 (A), MXD1 (B), and E2F3 (C). mRNA was isolated from cells (without CD41a<sup>+</sup> selection) collected at the indicated days from 2 independent primary Mk cultures started with human CD34<sup>+</sup> cells stimulated with 100 ng/ml thrombopoietin (Tpo). Data are averaged from 2 independent biological experiments. Error bars represent SE. These data show expression patterns qualitatively similar to the microarray expression patterns of Fig. 1A, which were obtained from a separate set of Mk cultures stimulated with a Tpo-containing cytokine cocktail (Tpo-IL-3-Flt-3L), and to CD41a<sup>+</sup>-selected cells, thus confirming the broader validity of the microarray data.](http://physiolgenomics.physiology.org/)
ported to interact with the early Mk-associated TF TAL1 in both erythroid (E) and Mk lineages and acts as a repressor of E-cell differentiation, but its role in Mk differentiation is not clear (49). Thus we identified a set of TFs that, along with some of their cofactors, may be involved in regulating megakaryopoiesis.

Immunofluorescence microscopy confirms protein expression and nuclear localization of four potential new Mk TFs: FHL2, MXD1, E2F3, and RFX5. As a first step to assessing the functional significance of some of the novel findings among these early upregulated TF genes, we employed immunofluorescence microscopy (IFM) to confirm the expression and subcellular localization of FHL2, MXD1, E2F3, and RFX5 at the protein level in Mk cultures initiated either with Tpo alone or with the cocktail of cytokines, as well as in G cultures. At least three independent biological experiments were conducted for each TF in cultures initiated with Tpo alone or with the cocktail of cytokines, and representative data are discussed below. Similar expression patterns were observed in both types of Mk cultures, and this is consistent with the finding that our Q-RT-PCR data (Fig. 3; Tpo-only Mk cultures) agree qualitatively with the microarray data of Fig. 1 (cocktail Mk cultures). The choice of Tpo-only cultures was made based on two findings from our earlier study involving human mobilized peripheral blood CD34+/H11001 cells stimulated with Tpo alone (18).

First, these cultures characteristically yielded a higher purity of CD41a+/H11001 cells (70% by day 9; this is an important attribute for the IFM studies described below), and the cells underwent more extensive maturation, including polyploidization and pro-platelet formation, compared with the cytokine cocktail cultures (18). Second, microarray analysis (18) (using the same microarray platform, experimental design, and reference RNA as in the present study) of the Tpo-only cultures with a different set of donor samples and without CD41a+ cell enrichment showed similar gene expression patterns (18). Therefore, we directly compared the results of the two types of cultures. The goal was to assess whether the overall patterns of TF expression were preserved between the two culture systems. Using SAM with a false discovery rate of 1%, we identified a set of 178 TF genes that were differentially expressed in either the Tpo-only cultures (days 0, 5, 6, 8, and 10) or the cytokine cocktail cultures (days 0, 5, 7, 9, and 12 only) or both. Qualitatively, the data from the two culture systems agreed for most of these TF genes (see Supplemental Fig. S1), and certainly so for the four genes discussed in this section.

The relatively few TF expression differences can be attributed either to the presence of additional cytokines in the cocktail cultures or to the lack of CD41a+ cell enrichment in the Tpo-only cultures.

FHL2 expression was weak and diffuse in CD34+ HSPCs (Fig. 4) and after 2 days of stimulation with Tpo only (data not shown). On further Mk differentiation, FHL2 was expressed in the nucleus with a punctate pattern and weakly in the cytoplasm (days 5 and 7; Fig. 4). The same strong nuclear punctate distribution of FHL2 was also found in cytokine cocktail-
stimulated Mk cultures (data not shown). Later in culture, FHL2 was more diffuse but still significantly expressed (day 9; Fig. 4). Cultured G cells also transiently expressed FHL2 at early time points. However, FHL2 was largely absent in more differentiated G cells (data not shown).

MXD1 was not expressed in day 0 CD34+ HSPCs (Fig. 5). By day 2, the cultured Mk cells displayed weak, primarily nuclear MXD1 expression (data not shown). On further Mk differentiation, nuclear MXD1 expression intensified and MXD1 became detectable in the cytoplasm of Mk cells (days 5 and 7; Fig. 5). Of note, MXD1 was also expressed on the protein level during G differentiation (data not shown), in agreement with its upregulation at the transcriptional level (Fig. 1A).

E2F3 was expressed mostly in the nucleus in CD34+ HSPCs (Fig. 6) but was weaker in the nucleus and the cytoplasm after 2 days of stimulation with Tpo (data not shown). On further Mk differentiation, the expression of E2F3 intensified and was mainly nuclear and low in the cytoplasm (day 5) and somewhat weaker on day 7. Increased nuclear accumulation of E2F3 was observed in late Mk cells (day 9) and on day 12 (data not shown). In agreement with our findings at the transcriptional level, expression of E2F3 was initially high on HSPCs stimulated with IL-3, IL-6, SCF, and G-CSF to undergo G differentiation (days 5, 7, 9, and 12; data not shown).

RFX5 was not expressed in CD34+ HSPCs (day 0; Supplemental Fig. S2), but early Mk differentiation under the effect of either Tpo or the cytokine cocktail resulted in nuclear localization and a punctate distribution at day 2 (not shown) and day 5 (Supplemental Fig. S2). Further Mk differentiation resulted in lower nuclear expression, and the expression was particularly low in late polyploid Mk cells (Supplemental Fig. S2). RFX5 was also expressed with a strong nuclear punctate pattern in early G cultures, and RFX5 expression significantly declined with progression along the G lineage (data not shown). The similarity of protein expression between Mk and G cells suggests that RFX5 may play a role in both Mk and G differentiation.

In summary, the protein expression patterns observed by IFM (Figs. 4–6) agreed overall quite well with the mRNA expression patterns of our microarray analysis. Furthermore, the nuclear localization of all four TFs examined by IFM and the punctate organization of FHL2 and RFX5 during Mk differentiation reinforce the likely functional activity of these TFs during megakaryopoiesis.

Gene expression clusters B through D provide additional clues regarding transcriptional regulation in Mk cells. Cluster B includes 28 genes (28 probes) that were upregulated in Mk cells compared with G cells and early cells but were expressed at a lower level than in CD34+ cells (Fig. 1B). These genes may be important in both Mk and hematopoietic progenitor regulation. Among the established Mk TFs, FLI1 was present in this cluster. BTEB1 exhibited an expression pattern very similar to that of FLI1. BTEB1 has been studied in neuronal development and was found to regulate the level of neurite branching (4). Thus it will be interesting to examine the role of

![MXD1/DAPI, CD41/DAPI, MXD1](image-url)
BTEB1 in proplatelet formation. NR4A2 is a nuclear hormone receptor widely expressed in the central nervous system the silencing of which, in cancer cell lines, induces apoptosis (35)—a program of paramount importance in Mk cells. The forkhead transcription factor family member FOXO1A participates in phosphatidylinositol 3-kinase/Akt signaling and induces cell cycle arrest (12). While FOXO1A has not, to our knowledge, been associated with Mk cells, its family member FOXO3A, which is also upregulated (Fig. 1A), has been shown to be present in mature human Mk cells and to be phosphorylated in response to Tpo (56).

Cluster C includes 14 genes (14 probes) that were expressed at a higher level in Mk cells than in CD34+ cells but lower compared with early culture (Fig. 1C). Some of these genes may be necessary for differentiation of other lineages induced by our cytokine cocktail rather than the Mk lineage. In support of this, we note the inclusion of the erythroid TF Kruppel-like factor 1 (KLF1) in this cluster. It has been shown that KLF1 and FLI1 antagonize each other in E/Mk development (55), and thus KLF1 downregulation may be necessary for Mk differentiation. Furthermore, glycophorin A+ E cells are a major contaminating population in cultures treated with the Mk cytokine cocktail (data not shown) and would be expected to contribute to the detected gene expression among the mixed-population samples from early time points.

Cluster D includes 110 genes (112 probes) that were consistently downregulated in Mk cells (Fig. 1D). These genes would be expected to be dispensable for Mk differentiation. In agreement with this, the myeloid TFs CEBPA, CEBPB, and CEBPE were generally downregulated in Mk cells, although CEBPE was slightly upregulated in early Mk cultures. GFI1, which was also downregulated, is essential for neutrophil differentiation (29) and also restricts HSPC proliferation (24). The downregulation of G-specific TFs from this cluster reflects the transcriptional decision of cells that undergo unilineage differentiation. Members of the class I homeobox gene (HOX) family were downregulated in both Mk and G cells, demonstrating a downregulation of HSPC-associated genes during lineage differentiation. Many zinc finger proteins are also included in this cluster, but their functions remain elusive.

DISCUSSION

In this work, we sought to identify TFs and transcriptional regulators not previously associated with megakaryopoiesis. We approached this problem by analyzing global time course microarray data from ex vivo Mk cultures, comparing the results with parallel, isogenic G cell cultures, and leveraging Gene Ontology associations to focus our attention on those genes associated with the regulation of gene transcription. This approach captured most of what was previously known about Mk transcription factors and extended this knowledge base in two ways: first, by presenting detailed kinetic gene expression information on the previously known Mk TFs under two different Mk-inducing cytokine treatments; and second, by identifying a new set of TFs not previously associated with megakaryocytes. This new TF set was further explored by examining protein-level expression and subcellular localization via IFM. These validation studies strengthened the argument that these genes are active during Mk differentiation. We
have also shown that these TF mRNA expression patterns are independent of the cytokines used to differentiate CD34+ cells into megakaryocytes.

FHL2 is one of the better-studied members of the four-and-a-half-LIM-only protein family, with roles in transcriptional activation, signal transduction, cell cycle control, and apoptosis (28). Its presence is well established mainly in heart but also in brain, ovarian, liver, and lung tissue (28). FHL2 may activate the transcriptional program of the p53 tumor suppressor through association with HIPK2 (32). We previously proposed (16) a role for the transcriptional activation of p53 in Mk differentiation.

MXD1 is a member of the MAD family of genes known to dimerize with MAX, thus preventing it from activating the MYC transcription factors (26, 64). In agreement with this, after transient expression at early time points, MYC was downregulated after day 5 in our Mk cultures (Fig. 1C). Furthermore, MYC levels decline in K562 cells undergoing Mk differentiation, whereas MAX levels remain stable, which is reminiscent of the competition between MYC and MAD for association with MAX (33). Studies in differentiating U937 monocyctic cells (3) and a study of granulopoiesis in mouse knockouts of MXD (13) have led to the hypothesis that MXD-mediated repression of MYC represents a mechanism that arrests proliferation and initiates terminal differentiation. Indeed, overexpression of MYC results in pronounced proliferation of the immature Mk compartment and a concomitant drop in the fraction of polyploid Mk cells in vivo (57).

E2F3, a member of the E2F family of transcription factors (48), is composed of two isoforms, E2F3a, known to promote proliferation via cell cycle progression and also function as a transcriptional activator, and E2F3b, which is associated with inhibition of the cell cycle and transcriptional repression (9). Interestingly, E2F3 has been shown by yeast two-hybrid screens to strongly interact with FHL2 (48). Of note, E2F3b was demonstrated to impinge on the p53 tumor suppressor pathway by repressing ARF, which is known to repress MDM2 that associates with p53, leading to polyubiquitination and degradation of p53 (2). Ectopic expression of E2F3 was also shown to induce apoptosis in Rat1 fibroblasts (59). Loss of E2F4, another member of the E2F family, leads to aberrations in the erythrocytic, granulocytic, monocyctic, and lymphocytic lineages such as increased numbers of immature cells and increased apoptosis (46).

RFX5 is expressed by a panel of unstimulated leukemic cell lines with Mk potential including MEG-01, UT7/GM, CMK11-5, and HEL (37); however, its expression in primary Mk cells has not been previously demonstrated. RFX5 is best known as a transcriptional regulator of major histocompatibility complex class II (MHC II) genes and is often mutated in patients suffering from bare lymphocyte syndrome (38, 45). MHC II gene expression is normally restricted to antigen-presenting cells, and, as expected, MHC-II genes were not upregulated in our Mk or G cultures (data not shown). While our microscopy and gene expression analysis support a potential role for RFX5 in Mk cells, its function and target genes in this context remain to be elucidated.

Of note, our analysis predicted that KLF1, a zinc finger transcription factor with a well-established role in erythropoiesis (43), is downregulated in late Mk differentiation. In agreement with our finding, a very recent independent study has demonstrated downregulation of KLF1 in Mk cells after the bipotent MkE progenitor stage (15).

In summary, we have leveraged global transcriptional analysis and Gene Ontology associations to identify numerous candidate TF and TF-related genes expressed in differentiating Mk cells. Global approaches to such biological questions as “What is the complete repertoire of transcription factors orchestrating megakaryopoiesis?” inevitably identify more candidate genes than can be fully validated in any one study and paper, and here our goal is to develop the methodology for identifying the likely repertoire of Mk TFs. Further follow-up studies to address the specific role of these newly identified Mk TFs on megakaryopoiesis are ongoing in our laboratory, and we hope this work will provide useful data and methodological ideas for further work by others.

ACKNOWLEDGMENTS

We acknowledge the use of instruments in the Keck Biophysics Facility, the Biological Imaging Facility, and the Center for Genetic Medicine at Northwestern University. We gratefully acknowledge Genentech for the donation of recombinant human thrombopoietin.

GRANTS

This work was supported by National Heart, Lung, and Blood Institute Grant HL-48276 and the Robert H. Lurie Comprehensive Cancer Center. P. G. Fuhrken was supported by a National Science Foundation Graduate Research Fellowship. P. A. Apostolidis was supported by the A. S. Onassis Foundation.

REFERENCES


