Expression analysis of primary mouse megakaryocyte differentiation and its application in identifying stage-specific molecular markers and a novel transcriptional target of NF-E2

Zhao Chen,[1,2] Michael Hu,[1] and Ramesh A. Shivdasani[1-3]

[1]Dana-Farber Cancer Institute, Boston, MA; [2]Department of Medicine, Harvard Medical School, Boston, MA; [3]Department of Medicine, Brigham & Women's Hospital, Boston, MA

Megakaryocyte (MK) differentiation is well described in morphologic terms but its molecular counterparts and the basis for platelet release are incompletely understood. We profiled mRNA expression in populations of primary mouse MKs representing successive differentiation stages. Genes associated with DNA replication are highly expressed in young MKs, in parallel with endomitosis. Intermediate stages are characterized by disproportionate expression of genes associated with the cytoskeleton, cell migration, and G-protein signaling, whereas terminally mature MKs accumulate hemostatic factors, including many membrane proteins. We used these expression profiles to extract a reliable panel of molecular markers for MKs of early, intermediate, or advanced differentiation and establish the value of this marker panel using mouse models of defective thrombopoiesis resulting from absence of GATA1, NF-E2, or tubulin β1. Computational analysis of the promoters of late-expressed MK genes identified new candidate targets for NF-E2, a critical transcriptional regulator of platelet release. One such gene encodes the kinase adaptor protein LIMS1/PINCH1, which is highly expressed in MKs and platelets and significantly reduced in NF-E2–deficient cells. Transactivation studies and chromatin immunoprecipitation implicate *Lims1* as a direct target of NF-E2 regulation. Attribution of stage-specific genes, in combination with various applications, thus constitutes a powerful way to study MK differentiation and platelet biogenesis. (Blood. 2007;109:1451-1459)

Introduction

Megakaryocytes (MKs) undergo sequential morphologic and functional transitions that culminate in platelet assembly within proplatelets. Surprisingly little is known about the molecular chronology of MK maturation, and the field has relied on cytomorphologic criteria that often lack agreement or precision. Furthermore, the molecular underpinnings of many human thrombocytopenias remain unknown.

Early MK progenitors feature a rim of basophilic cytoplasm. Repeated cycles of endomitosis increase DNA content as the nucleus becomes convoluted and the cytoplasm less basophilic.[1,2] Thereafter, MKs develop a prominent cytoplasm that contains first innumerable α granules and, subsequently, dense bodies and a network of invaginated internal membranes.[3] Evidence suggests that this demarcation membrane system provides the reserve for proplatelet membranes and that nascent platelets assemble de novo during elaboration of proplatelets.[4-7] Thus, the maturing MK's duty is to assemble the signaling and structural components required to extend proplatelets.

Animal and human models of thrombocytopenia provide useful insights into mechanisms of platelet biogenesis. MKs require the transcription factor (TF) GATA1 to traverse early steps in maturation; in its absence, mouse MKs develop with a rudimentary, organelle-poor cytoplasm,[8,9] and *GATA1* mutations in man cause inherited thrombocytopenias.[10-12] Absence of the TF NF-E2, also associated with severe thrombocytopenia in mice,[13] spares early MK differentiation but imposes a profound block in proplatelet formation.[14] Between these 2 extremes, some mouse coat-color mutants manifest storage-pool disorders without affecting platelet counts, although the *RabGGT* mutation in *gunmetal* mice appears to affect both aspects.[15,16] *β1 tubulin*, an MK-specific gene and apparent target of NF-E2 regulation,[17] is required to assemble platelets in proper numbers and shape[18] but cannot alone restore proplatelet formation in NF-E2–deficient MKs.[17] Some genes responsible for human disorders of platelet synthesis are defined, including components of the glycoprotein (GP) Ibα/Ibβ/IX complex, the myosin heavy chain *MYH9,* or TF *Fli1,* in the Bernard-Soulier,[19] May-Hegglin,[20,21] and Paris-Trousseau syndromes,[22,23] respectively. Although these findings shed light on thrombopoietic mechanisms, prominent gaps remain in the present understanding.

The experience with NF-E2 illustrates one such gap. Genes that might operate downstream of NF-E2 include enzymes (thromboxane synthase),[24] cytoskeletal proteins (β1 tubulin),[17] signaling factors (αIIb-β3 and Rab27b),[16,25] and effectors of apoptosis (caspase 12).[26] In most cases, it is unclear whether NF-E2 regulates the transcript directly, as some candidates could be markers of MK maturation whose absence in NF-E2–null cells merely reflects differentiation arrest. One barrier is the absence of a reliable catalog of gene-expression changes that occur as MKs progress in maturity. Such a resource could provide the molecular framework to place individual genes or pathways in functional context.

Expression profiling of human platelets[27] and $CD34^+$ cell–derived MKs[28-31] has enumerated abundant transcripts, but published studies have not compared successive stages in MK differentiation or established associations with cellular function. Here we

Submitted August 1, 2006; accepted September 25, 2006. Prepublished online as *Blood* First Edition Paper, October 17, 2006; DOI 10.1182/blood-2006-08-038901.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

relate gene-expression profiles during mouse MK differentiation to cell morphology, molecular markers of representative differentiation stages, and expression in well-established mouse models of thrombocytopenia. Our results reveal broad, stage-specific arrests in differentiation of MKs that lack GATA1 or NF-E2 and imply that most mRNAs reduced in the mutant cells represent stage-specific markers rather than direct transcriptional targets. We also studied promoters of differentiation-related genes that may be regulated directly by NF-E2 and identified an adaptor protein for the integrin-linked kinase, LIMS1/PINCH1, as a novel candidate component of NF-E2–dependent thrombopoietic pathways.

Materials and methods

Cell culture

Mouse fetal livers, collected on embryonic day 14 (E14), were processed into single-cell suspension by successively passing through 20- and 22-gauge needles, followed by filtering with a 70-μm cell strainer. Cell culture and first-pass enrichment for MKs were performed as described previously,[14] except that a triple gradient of bovine serum albumin (BSA; 4%, 3%, and 1.5%) was applied for improved separation. COS cells were maintained in Dulbecco modified Eagle medium (DMEM; Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (FBS).

Flow cytometry

MK cultures collected at different days were stained with FITC-labeled CD41 monoclonal antibody (BD Biosciences Pharmingen, San Diego, CA) for 40 minutes on ice, washed in cold phosphate-buffered saline (PBS), and resuspended in DMEM containing 10% FBS. Stage-specific MKs were isolated in a MoFlo cell sorter (Dako, Carpinteria, CA) into sterile tubes containing 1 mL FBS. For DNA content analysis, cells were incubated for 30 minutes in staining solution (50 μg/mL propidium iodide; 4 mM sodium citrate; 0.2 mg/mL RNaseA; 0.1% Triton X-100, pH 7.8) and analyzed by FACscan flow cytometry (Becton Dickinson, Franklin Lakes, NJ).

Microarray analysis

Total cellular RNA was isolated with Trizol reagent (Invitrogen) and its quality assessed by spectrophotometry and microfluidic size fractionation (Agilent Technologies, Palo Alto, CA). RNA was reverse transcribed and amplified in the presence of biotin-conjugated nucleotides to produce biotinylated cRNA. Fragmented cRNAs were hybridized to Affymetrix (Santa Clara, CA) MOE430A microarray chips and washed; bound cRNA was stained with fluorescent streptavidin and detected by laser. MK subpopulations were tested in duplicate; only results for MKs isolated on the sixth culture day (MK-6) derive from 3 independent isolates. Hybridization data were analyzed by MAS 5.0 software (Affymetrix) and also normalized by the dCHIP invariant set algorithm based on a perfect-match/mismatch (PM/MM) model,[32] although expression values are based on the PM model only.[33] Self-organizing maps (SOMs),[34] quality threshold (QT) clusters,[35] and one-way analysis of variance (ANOVA) were generated with multiexperiment viewer (MEV) software.[36] QT clustering was performed with Pearson distance metric at a diameter of 0.1 and SOM with Pearson distance metric at 1.0 radius; α-level was set at 0.01. Sample comparison, hierarchical clustering, and gene-ontology analyses were carried out using dCHIP software. Gene-set enrichment analysis (GSEA) was performed using GSEA software.[37]

Real-time quantitative (q) PCR

Following reverse transcription (RT) of isolated RNA, we compared polymerase chain reaction (PCR)–amplified products in real time by the amount of integration of SYBR green dye using either Biorad iCycler (Hercules, CA) or ABI 7300 (Norwalk, CT) instruments. Primers for PCR were designed with Primer3 software (http://frodo.wi.mit.edu/primer3) to amplify approximately 200–base pair products for each transcript of interest. Efficiency of each primer pair was confirmed by serial dilutions of templates, and sequences are provided in Table S3 (available on the *Blood* website; see the Supplemental Materials link at the top of the online article). A dissociation curve was generated for each reaction to confirm amplification of single products, which were further assessed for product size and purity by agarose gel electrophoresis. Results were normalized to the level of *GAPDH* mRNA between samples.

Luciferase reporter assay

Lims1 promoter variants were cloned into the vector PGL3Basic (Promega, Madison, WI). COS cells were transfected at 90% confluence using Lipofectamine 2000 (Invitrogen) and cultured in 24-well plates. Besides the reporter construct and p45 NF-E2 expression plasmid cloned in a pEF vector,[38] CMV promoter–driven *Renilla* luciferase plasmid (Promega) was included in every well as the internal control for transfection efficiency. Firefly and *Renilla* luciferase activities were measured 48 hours after transfection using the Stop&Glo Dual Luciferase assay kit (Promega). All expressed values are normalized for transfection efficiency.

ChIP

Mature MKs were isolated on day 5 of E13.5 mouse fetal liver culture over a discontinuous 4% + 3% BSA gradient. Chromatin immunoprecipitation (ChIP) analysis was performed essentially as described in detail in a previous report.[16] Primers for PCR analysis (sequences listed in Table S3) were designed to amplify approximately 200-bp products that flank putative NF-E2 binding sites or 4-kb 3′ of these sites and in irrelevant genes (*Apo1a, histoneH4*) as controls.

Immunofluorescence analysis

MK cytospin preparations were fixed in 4% paraformaldehyde (PFA) for 15 minutes; purified mouse platelets were fixed in 4% PFA for 5 minutes before centrifugation at 18*g* onto poly-L-lysine–coated coverslips. Cells were permeabilized with 0.5% Triton X-100 in PBS for 5 minutes and blocked with 1% BSA and 0.1% goat serum in PBS for 1 hour. Antibodies (Abs) used in staining included FITC-conjugated anti-CD41 (BD Biosciences) and rabbit anti-Lims1 (gift from Dr Chuanyue Wu, University of Pennsylvania, Philadelphia) followed by Cy3-labeled goat anti–rabbit IgG (Jackson Immunoresearch, West Grove, PA). Slides were mounted with Fluoromount-G (Southern Biotech, Birmingham, AL) and examined on an Olympus IX70 inverted fluorescence microscope (Olympus, Melville, NY) using a 60×/1.40 NA oil-immersion objective lens with a 0.1-mm aperture for platelets and a 40×/1.35 NA plain objective lens for MKs. Images were acquired with a CM350 CCD camera (Applied Precision, Issaquah, WA) using DeltaVision software (Applied Precision) and processed with Photoshop 7.0 software (Adobe Systems, San Diego, CA).

Results

Isolation of MK populations according to the degree of differentiation

MK progenitors expand in mouse bone marrow or fetal liver cell preparations cultured with thrombopoietin and mature over 5 to 6 days in vitro. After 3 days of culture, a significant fraction of cells shows features of committed MK progenitors, and numerous terminally mature, proplatelet-forming MKs appear by day 5. Because isolation of cell populations that correspond to sequential stages in MK differentiation is hindered by the lack of synchrony in primary MK cultures, we applied flow cytometry to harvest subpopulations that are substantially enriched for MKs with defined properties. High surface expression of the lineage marker CD41 identified MKs, whereas forward-scatter (FSC) properties distinguished cells on the basis of size.

We collected thrombopoietic culture suspensions from days 3, 4, and 6 and sorted populations by flow cytometry (Figure 1A). On

culture days 3 and 4, a sizeable population of cells corresponds to young MKs, as judged by distinct, albeit submaximal, CD41 expression. By morphology, these cells (designated MK-P) were the smallest among the several populations we collected and they displayed features of MK progenitors, including large nuclei and small basophilic cytoplasm (Figure 1B). To verify this assignment, we cultured sorted MK-P cells in TPO for 3 additional days and observed that the vast majority progressed to mature MKs, recognized by their size, high surface CD41 levels, and acetylcholinesterase enzyme activity (data not shown). There was little difference between MK-P isolated on either days 3 or 4 of fetal liver cell culture. Cells isolated by flow cytometry on culture day 6 were highly enriched for terminally mature MKs, as revealed by the nuclear morphology and enormous cytoplasm. Of the 4 populations we isolated, this one (designated MK-6) showed the shortest time (1 day) to, and the highest frequency (~30%) of, proplatelet formation. Two intermediate populations isolated on the third or fourth culture days, MK-3 and MK-4, carried dense nuclei and considerably more cytoplasm than MK-P (Figure 1B), with limited differences in morphology between one another. DNA content shifted toward higher ploidy across the 4 isolated cell populations (Figure 1C), implying that they represent points along a continuum of MK maturation. We thus isolated populations that are significantly enriched for cells with the ploidy, morphology, and functions associated with advancing MK differentiation.

Molecular interrogation of isolated MK populations

To reduce the effect of any variation between primary cell preparations, replicate RNA samples were isolated from separate sorted cell preparations so that each sample represents an independent cell culture. As the sorted populations derive from unsynchronized cultures, they overlap in some features and diploid contaminants were detected in each population; MK-6 included cells that may have completed platelet release. Nevertheless, our separation of successive stages in MK maturation was sound, and to identify corresponding alterations in gene expression, we used their RNAs to probe Affymetrix MOE430A mouse oligonucleotide arrays (22 690 probe sets; 14 417 genes).

Gene filtering identified 1160 genes (1400 probe sets) with significant differences between sample sets (standard deviation [SD] > 0.41). All replicates clustered together, except that one MK-4 set was related more closely to MK-3 than to its MK-4 duplicate. Reducing the gene-set size by increasing SD requirements did not alter cluster distributions, and to improve distinction between MK populations, we removed MK-4 from comparison studies. The arrays conservatively estimated approximately 1.5-fold differences (highest confidence limit; the upper boundary of the confidence limits were substantially higher) in expression of von Willebrand factor (VWF) and CD41. As the latter was readily detected by flow cytometry and constituted one basis for cell separation, we adopted 1.5-fold change as an empirical and stringent threshold for differences. Three hundred eighty-six transcripts were reliably increased and 284 genes showed reduced expression beyond the stage designated MK-P, with median false-discovery rates of 0 by permutation analysis. Among the 386 up-regulated transcripts, fewer than half are expressed at the highest levels in MK-6 or at roughly equal levels in MK-3 and MK-6; the rest peak in expression in MK-3. Many up-regulated genes are known to be highly expressed in MKs or important in platelet function (eg, the first 11 genes listed in Figure 2A), indicating that the sorted cell populations were enriched as predicted.

Unsupervised clustering with self-organizing maps (SOMs) partitions genes into a designated number of clusters based on correlation of expression. To accommodate the 3 experimental groups, we set SOM dimensions at 3×3 (9 clusters). The expression patterns assigned the 1160 filtered genes from these 9 clusters to 3 groups that show peak levels in MK-P, MK-3, or MK-6 (Figure S1). Thus, an unbiased analysis partitioned the filtered gene list into clusters that correspond to MK maturation stages and validates the experimental approach.

Functions gathered from MK expression profiles

To reveal molecular and cellular functions, we subjected each group derived from SOMs to hierarchical clustering, followed by gene-ontology analysis; this provided a statistical basis to identify functional pathways enriched within individual groups. Group 1 (507 genes) contains genes with a trend of declining expression over the course of MK differentiation (lowest expression in MK-6), and gene-ontology analysis uncovered striking enrichment of genes involved in DNA replication, nuclear activity, and metabolism (Table S1). TFs that mark hematopoietic progenitor cells (eg, Figure 2B), including c-kit, c-Myc, and c-Myb, also decline with MK maturation. Peak expression of this group corresponds to

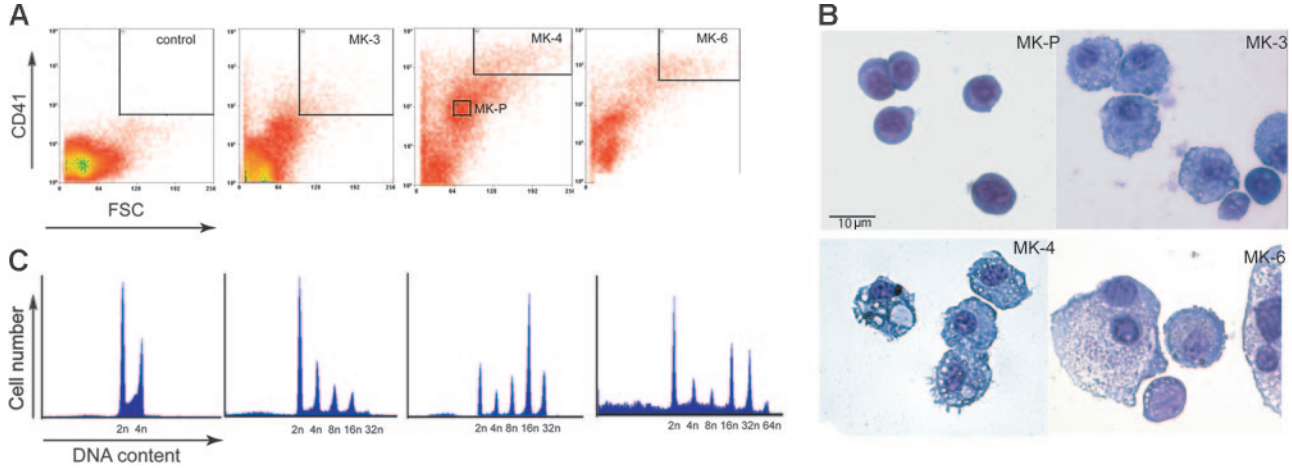

Figure 1. Isolation of mouse MK populations at successive stages of maturation. (A) Cells harvested on different days of fetal liver culture were labeled with CD41 antibody and sorted by flow cytometry, using gates suited to the forward-scatter (FSC; size) and CD41 expression level (fluorescence) of MKs. Parameters chosen for isolating MK-3, MK-P, MK-4, and MK6 are boxed within the corresponding plots. (B) May-Grünwald-Giemsa–stained cells representing the 4 MK populations sorted by flow cytometry. (C) Profiles of the DNA content of the sorted MK populations, labeled with propidium iodide and analyzed by flow cytometry.

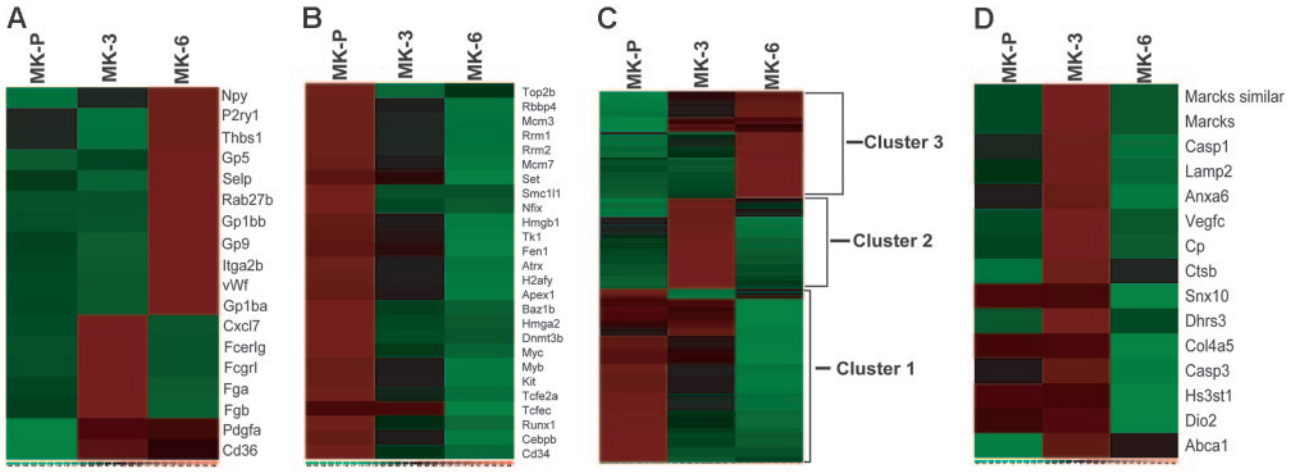

Figure 2. Gene-expression analysis of MK differentiation. Relative mRNA expression of individual genes (rows) is shown for 3 MK populations (columns) on a scale from low (green) to high (red) levels. (A) Relative expression of known MK/platelet transcripts in the 3 populations, indicating that abundant platelet genes are highly represented in mature MKs compared with MK-P and that some markers are activated earlier than others. (B) Partial representation of transcripts that are highly expressed in MK-P and subsequently extinguished. (C) Hierarchical clustering of the 594 probe sets identified through analysis of variance (ANOVA) in MK populations, revealing clusters that correspond well with successive stages. (D) Of the 32 genes implicated in retinoic acid signaling, 15 are highly expressed in MK-3, which represents a significant overlap (highest GSEA rank in our analysis) and encompasses factors that function in diverse processes.

residual cell proliferation and to endomitosis, a prominent feature in early MK ontogeny.

Group 2 (382 genes) carries genes that peak in MK-3 and is most notably enriched for 5 cellular functions (Table S2): (1) cytoskeletal and cytoskeleton protein binding; (2) G-protein–coupled signal transduction; (3) response to external stimuli; (4) chemotaxis; and (5) defense response. MK maturation and platelet release require assembly and reorganization of an extensive cytoskeleton,[5] probably in response to local environmental cues, including chemokines.[39,40] The elevation of mRNA levels in classes 1 to 4 agrees with prevailing models, highlights MK-3 as a cell population in transition toward platelet assembly, and identifies many candidate effectors of thrombopoiesis. Expression of genes related to host defense and in other categories may reflect platelet attributes or a possible role in platelet biogenesis.

Group 3 (271 genes) is the smallest and includes genes that appear at their highest levels in the most morphologically mature MK-6 population or roughly equally in MK-3 and MK-6 but significantly increased over MK-P. Well-characterized MK genes such as *CD41, VWF, GPIb*, and *GPV* fall in this group, which is highly enriched for hemostatic factors and includes genes responsible for binding to receptors and calcium ions (Table S3).

To test the significance of these results, we performed analysis of variance (ANOVA) for the 3 different MK populations. Five hundred seven of the 1160 genes were found to differ the most significantly ($P < .05$) in expression among populations, and hierarchical clustering yielded 3 distinctive clusters that peak in expression in MK-P, MK-3, or MK-6 (Figure 2C). Although fewer than half of the genes met the more stringent statistical criteria, functional pathways enriched in each cluster exactly match those identified independently by SOMs. To enhance stringency even further, we subjected the 507 genes identified by ANOVA to quality threshold (QT) clustering, which identifies sets of genes with highly correlated expression patterns across MK maturation. This analysis returned 9 significant subclusters (Figure S2), and the functional categories within these tight groupings, judged by gene-ontology criteria, confirmed the same signature pathways for maturation stages (Figure S3) as determined in the unsupervised clustering by SOMs.

To improve the search for pathways represented in our data, we applied gene-set enrichment analysis (GSEA), a comprehensive

Table 1. Gene sets enriched in MK-3 compared with all other isolated MK populations

Gene set	FDR	Rank
Signaling, cell motility, cytoskeleton		
RAR up-regulation	0.27	1
SIG PIP3 signaling in B lymphocytes	0.28	2
Rac1 pathway	0.20	3
GLUCO	0.18	5
CR cytoskeleton	0.17	6
Insulin signaling	0.14	10
Ptdins pathway	0.13	15
mCalpain pathway	0.13	16
Cell motility	0.13	19
Egf pathway*	0.14	29
Ras pathway	0.15	32
Igf1 pathway	0.15	33
Rho pathway	0.15	38
Transcriptional regulation		
CBFB MYH11 downing AML	0.16	4
HOXA9 down-regulation	0.16	7
Down-regulation by HOXA9	0.12	14
Up-regulation of G1ER	0.13	18
MLL chimeric fusion downing AML	0.13	20
GATA1 WEISS	0.13	21
Calcineurin pathway	0.14	31
Pparα pathway	0.16	45
NFAT pathway	0.15	46
General cell properties		
Bcl2 family and regulatory network	0.15	9
Bad pathway	0.13	11
Intrinsic pathway	0.13	17
Fatty acid metabolism	0.12	22
GO reactive O_2 species	0.12	23
ST dictyostelium cAMP chemotaxis	0.12	24
NK-cell pathway	0.14	30
HL60 ATRA BEN	0.15	34
Parkin pathway	0.14	35

Listed are 31 of the top 46 functional domains identified by GSEA as being enriched in MK-3.

FDR indicates false discovery rate; and AML, acute myeloid leukemia.

*$P = .05$; $P < .001$ for all other gene sets.

approach that links expression with an extended library of functional pathways.[37] Although this strategy did not expand the functions we had already associated with MK-P, it highlighted specific cell cycle–related genes that may be related to endomitosis, a hallmark of early MKs (Figure S4). Moreover, in comparing MK-3 with the other sorted MK populations, this analysis uncovered additional functions associated with MKs at midmaturity, particularly in pathways of retinoic acid, phosphatidylinositol P_3, and small-GTPase signaling; HoxA9-regulated transcription; and Bcl2/Bad-mediated cell survival (Table 1). Functions known to be important in MK maturation, such as cytoskeletal rearrangement and GATA1-regulated transcription, were readily evident in this analysis (Table 1) and enhance confidence in the validity of novel findings. In particular, MK-3 transcripts correlated best with a diverse array of genes previously associated with retinoic acid signaling (Figure 2D) and include the retinoic acid receptor *Rara* and the retinoid metabolic enzyme *Dhrs3*. These observations attach new importance to retinoid signaling in MK maturation.

These data collectively demonstrate the scope of our analysis, constitute a detailed atlas of MK gene expression, highlight dominant cellular processes recruited during thrombopoiesis, and permit empiric derivation of candidate markers and regulators of MK maturation. To identify specific genes that might be especially important in the dominant cellular functions, we used ANOVA to extract genes whose expression dynamics across the sorted cell populations correlate highly ($P < .05$) with those of CD41, a bona fide marker of mature MKs and platelets. Although CD41 mRNA levels for mature MKs showed a less than 2-fold increase over MK-P in the microarray analysis, this value was computed conservatively, represents the lower confidence limit, and likely underestimates true expression differences, as flow cytometry indicated substantially greater differences in surface protein levels. Nevertheless, the change in CD41 transcript levels across MK stages should serve as a suitable benchmark for molecules associated with differentiation. We thus identified 386 transcripts whose apparent change in expression between MK-P, MK-3, and MK-6 parallels that of CD41. Gene-ontology criteria place 22 of these CD41-correlated genes in the categories of cell communication and G-protein–coupled or small GTPase-mediated signal transduction (Table 2).

Identification and validation of new molecular markers of stages in MK differentiation

Because each MK population was unavoidably contaminated with cells from other stages, it is challenging to identify transcripts that might uniquely mark early, midstage, or advanced MKs. Within enriched cell populations, however, one can expect to identify genes with overrepresented or depleted expression. After excluding genes with low absolute expression (< 80 units), we compared the replicates in each group with every other sample to isolate transcripts with a greater than 1.5-fold difference in expression level. Among the 470 genes that change during MK maturation, the expression of 44 could potentially be restricted to discrete populations (8 in MK-P, 16 in MK-3, and 20 in MK-6); we then discarded 10 transcripts where redundant probe sets showed discordant hybridization and 12 genes that were poorly annotated. Using fresh MK populations, sorted according to the same parameters as before (Figure 1A), we used real-time quantitative (q) RT-PCR to confirm expression dynamics for the remaining candidate markers of MK maturation (Figure 3). In all, 17 genes fulfilled strict criteria for stage selectivity: 3 genes in MK progenitors, 10 in MKs of intermediate maturity, and 4 in the most advanced cells ($P < .05$ in *t* test comparisons of samples from each stage versus all other samples). All 17 markers appear among the top-ranked genes for corresponding stages by GSEA (data not shown). This identification of stage-enriched transcripts meets one of our goals but is not intended to be comprehensive, and our primary data, which are publicly available,[49] can serve as a resource to identify additional stage-selective genes.

Genes expressed highest in MK-3 included α-fetoprotein, albumin 1, *H19*, and transthyretin, which are especially abundant in

Table 2. Genes implicated in MK differentiation by virtue of a temporal expression pattern that correlates best with that of CD41

Gene symbol	Gene name	Accession no.
G-protein–coupled signaling		
Rgs10	Regulator of G-protein signaling 10	NM_026418
Rgs18	Regulator of G-protein signaling 18	NM_022881
Gpr56	G-protein–coupled receptor 56	NM_018882
Ptger3	Prostaglandin E receptor 3 (subtype EP3)	NM_011196
Ptgir	Prostaglandin I receptor (IP)	AV373771
Gnaz	Guanine nucleotide-binding protein, alpha z subunit	AI326356
Gna13	Guanine nucleotide-binding protein, alpha 13	NM_010303
Small GTPase-mediated signal transduction		
Rab10	RAB10, member RAS oncogene family	BF465974
Rab11a	RAB11a, member RAS oncogene family	BI083615
Rab6	RAB6, member RAS oncogene family	BB699846
Rab27b	RAB27b, member RAS oncogene family	NM_030554
Kras2	Kirsten rat sarcoma oncogene 2, expressed	BC010202
Cdc423p5	CDC42 effector protein (Rho GTPase binding) 5	NM_021454
Cell communication		
Itga2b	Integrin alpha 2b	NM_010575
Lims1	LIM and senescent cell antigen-like domains 1	BM213930
Col18a1	Procollagen, type XVIII, alpha 1	NM_009929
Gp1ba	Glycoprotein 1b, alpha polypeptide	NM_010326
Gp1bb	Glycoprotein 1b, beta polypeptide	NM_010327
Plek	Pleckstrin	NM_019549
Itga4	Integrin alpha 4	NM_010576
Stim1	Stromal interaction molecule 1	NM_009287
Esa1	Endothelial-cell–selective adhesion molecule	AF361882

The results represent analysis of variance (ANOVA) across all tested MK differentiation stages (ie, MK-P, MK-3, and MK-6). This table highlights the 22 transcripts (out of a total of 386 CD41-correlated genes) implicated in G-protein–coupled signaling, small GTPase-mediated signal transduction, or cell communication.

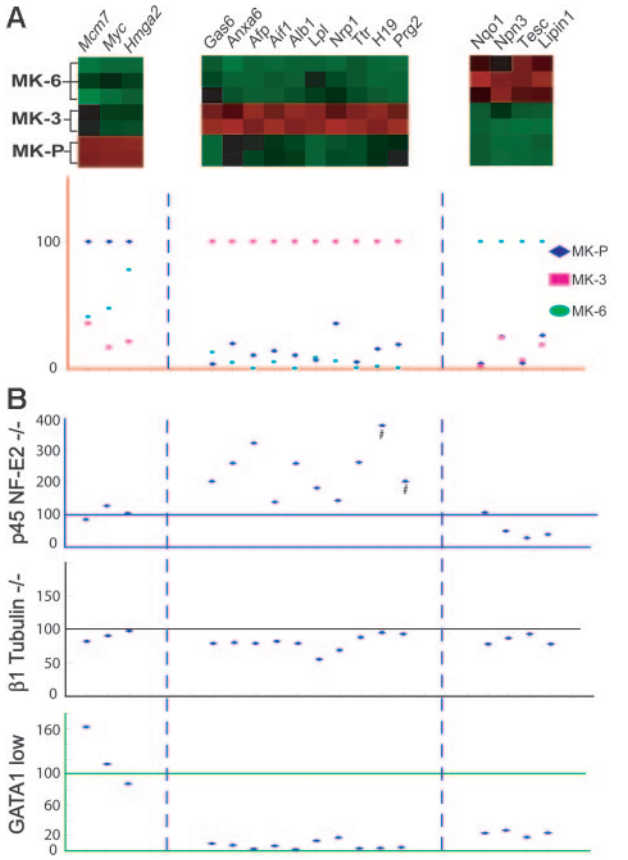

Figure 3. Molecular markers of MK differentiation stages. (A) Distinctive profiles of putative stage–selective markers revealed in microarray analysis (red = high, green = low levels) and confirmed by qRT-PCR in independent flow-sorted wild-type MK populations. To display the results, peak expression of each marker is represented at 100% in the stage with which it is best associated (although absolute levels of each may differ significantly), and transcript levels in the other 2 populations are expressed in relative terms. (B) Graphic representation of real-time qRT-PCR analysis of candidate molecular markers in unsorted (BSA gradient-purified) MKs derived from mouse models of defective MK differentiation. Individual genes are represented in the same columns throughout the figure and the levels of each are expressed in relation to those observed in wild-type (littermate) MK cultures, which is set at 100%. p45 NF-E2$^{-/-}$ MKs carry higher levels of genes that normally mark MK-3, with minimal change in MK-P markers and reduced levels of MK-6 genes. GATA1^{lo} MKs show significantly reduced expression of genes that normally express well in MK-3 and MK-6, whereas changes in expression of this marker gene panel in β1 tubulin–null MKs are negligible. Error bars are omitted for ease of viewing but are included in Figure S5. #Plotted at one third of actual values to fit the scale.

adult liver and hence raise the concern that their appearance represents contamination of our MK populations by fetal hepatocytes. We therefore used qRT-PCR to assess expression of these genes in freshly isolated fetal liver cells with that in MK fractions. α-fetoprotein, albumin 1, *H19,* and transthyretin mRNAs were considerably enriched in MKs purified over 2 successive BSA triple-gradients as well as in mature MKs cultured from adult mouse bone marrow (data not shown). They are thus unlikely to reflect cell contamination and instead show unexpected expression in maturing MKs of genes regarded to be liver specific.

One value of stage-selective molecular markers is to help define steps at which specific mutations arrest cell differentiation. To test this role, first we isolated MKs from p45 NF-E2$^{-/-}$ and littermate (p45 NF-E2$^{+/-}$) fetal livers. Whereas qRT-PCR analysis of an unsynchronized, mixed MK population from control animals showed high expression of all genes, NF-E2–null MKs showed elevated levels of nearly every marker assigned to MK-3 and at least 2-fold reduction of transcripts linked to MK-6 (Figure 3); these results provide distinct and novel molecular correlates for the arrested state of NF-E2–null MKs. Next we considered β1 tubulin, a factor that is absent from NF-E2–deficient cells and partially responsible for their failure to extend proplatelets[18]; absence of a structural protein like β1 tubulin may be predicted to have fewer effects on MK maturation than that of the TF NF-E2. Indeed, β1 tubulin$^{-/-}$ MKs showed little to no change in transcript levels of the candidate stage markers (Figure 3). Finally, we tested MKs lacking GATA1, a TF that operates upstream of NF-E2 and whose absence arrests MK maturation early with prominent progenitor expansion.[8,9] GATA1-null MKs show significantly reduced expression of markers of intermediate or advanced maturity (Figure 3), exactly according to prediction. These studies validate the panel of stage-selective genes identified by expression profiling and can be applied to characterize the growing number of murine or human conditions associated with arrested MK differentiation.

Integration of expression data with biologic functions

The main value of gene-expression studies in cell differentiation is to identify coordinately regulated transcripts, a value that is best realized when gene groups are considered further in the light of appropriate biologic models. For instance, although NF-E2 executes vital functions near the terminal phase of MK maturation, few of its many presumed transcriptional targets are known; these targets ought to be considerably enriched among genes expressed late in the MK life cycle. Thus, besides enabling future characterization of the components and mechanisms of platelet assembly, our gene-expression data can be applied to advance the current priority of identifying putative NF-E2 target genes. To this end, we integrated our microarray-based definition of late-expressed MK genes with prior knowledge of the NF-E2–specific *cis*-element (G/A)NTGA(C/G)TCA.[41,42] Functional copies of this sequence are found in a handful of erythro-MK promoters,[43] including that for *Rab27b,* which is expressed abundantly and almost exclusively in MKs but lost in p45 NF-E2$^{-/-}$ cells.[16]

Identification of candidate targets of NF-E2 regulation

We first extracted 2500 bp of sequence upstream of the transcription start sites for 13 500 mouse and 19 000 annotated human genes[44] (Figure 4A). For each putative promoter region, we used the PatScan pattern-matching algorithm[45] to locate all NF-E2 binding sites, (G/A)NTGA(C/G)TCA. Five thousand one hundred ninety-seven human and 2409 mouse promoters harbor 1 or more NF-E2 motifs. We considered only the 429 genes in which both mouse and human homologs[46] carry NF-E2 binding sites and determined how many were represented among the 386 genes that are expressed significantly more in MK-3 or MK-6 compared with MK-P (249 of these genes are reliably annotated). Eighteen genes emerged as candidates for NF-E2 regulation, including *Rab27b,* which we had previously identified by genetic and chromatin immunoprecipitation (ChIP) studies, as a likely direct transcriptional target.[16] To focus attention further, we considered only the 6 genes that carry at least 2 nonoverlapping NF-E2 binding motifs in the mouse gene promoter. qRT-PCR evaluation of these candidates revealed modest (2- to 3-fold) increases in the levels of 3 transcripts in p45 NF-E2$^{-/-}$ MKs. Because NF-E2 activates genes, we concentrated on the single transcript (*Lims1*) that is reduced in NF-E2–null MKs to roughly the same degree (≥ 10-fold) as the well-characterized *Rab27b* and β1 tubulin mRNAs (Figure 4B).

Lims1/Pinch1 is implicated in cell migration and spreading,[47] functions thought to be important in platelet release. Immunofluorescence and immunoblot studies confirmed that LIMS1 is highly expressed in MKs and platelets (Figure 4C). In agreement with

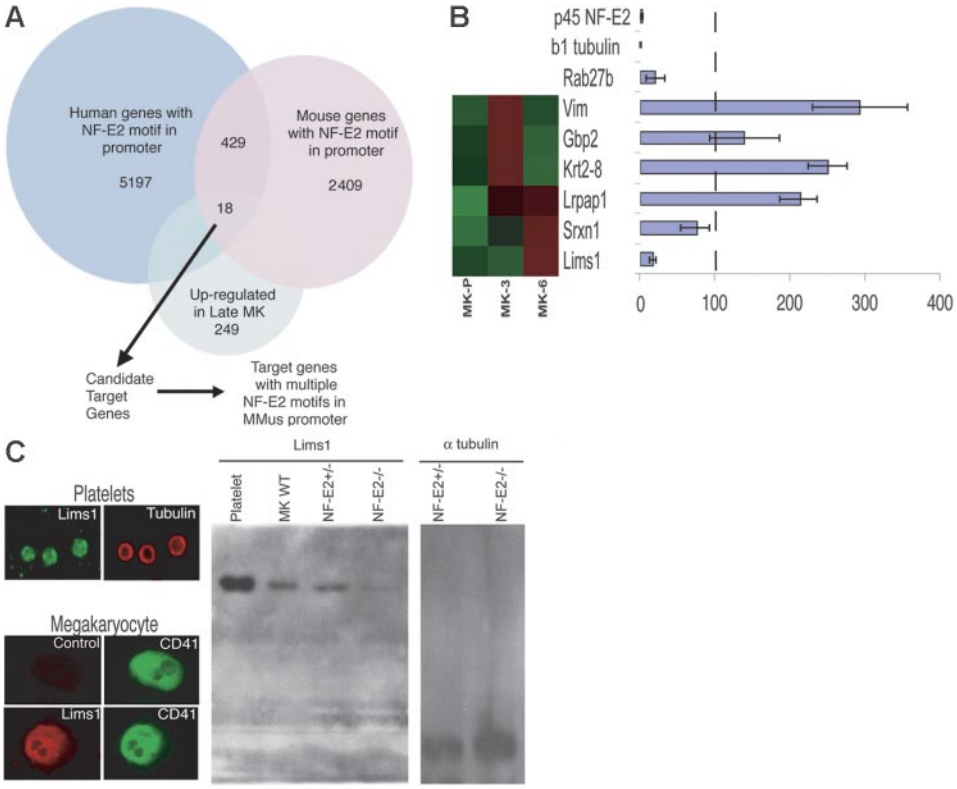

Figure 4. Identification of candidate NF-E2 transcriptional targets. (A) Outline of the computational approach to identify targets of NF-E2 regulation among transcripts that are significantly enriched in mature MKs. Eighteen genes carried NF-E2 binding sites in their promoters and 6 of these promoters contain 2 or more NF-E2 motifs. (B) Candidate NF-E2 target genes. Microarray results of MK expression are shown on the left (high expression, red; low levels, green) and results of qRT-PCR in p45 NF-E2$^{-/-}$ MKs on the right. For each gene, expression in p45 NF-E2$^{+/-}$ MKs is set to 100% (vertical dashed line) and relative mRNA levels in NF-E2–null MKs are represented in the bar graph. *Lims1* mRNA is reduced to levels comparable to those of known NF-E2–dependent genes, *Rab27b* and β1 tubulin. (C) LIMS1 protein is highly expressed in MKs and blood platelets and levels are reduced in the absence of NF-E2. A representative mature (CD41^{hi}) MK is shown stained for LIMS1. Immunoblot analysis also reveals LIMS1 expression in platelets and MKs, with significantly reduced amounts in p45 NF-E2$^{-/-}$ MKs compared with wild-type (WT) or heterozygous cells. Lanes to the far right represent the loading control (α-tubulin) for heterozygous and nullizygous mutant MKs, respectively. Error bars in (B) represent standard deviation. Images in (C) were captured as described in "Materials and Methods."

reduced mRNA in p45 NF-E2$^{-/-}$ MKs, LIMS1 protein levels are also significantly lower (Figure 4C). Thus, a combined computational and experimental approach identified a new candidate target of NF-E2 gene regulation.

NF-E2 regulates the *Lims1/Pinch1* gene

The 5′ flanking region of *Lims1* contains 2 potential NF-E2 binding sites, at −495 bp and −1939 bp with respect to the transcription start site (Figure 5A). To assess the relevance of these sites in gene regulation, we performed luciferase reporter assays in COS cells. As the basal promoter activity is substantially higher than that of the empty vector, we assume that factors responsible for the bulk of isolated promoter activity are present in COS cells (Figure 5B). The proximal NF-E2 site in the *Lims1* promoter did not reveal additive activity when p45 NF-E2 was forcibly expressed, whereas the distal site conferred over 2-fold increase with the addition of p45 NF-E2 (Figure 5B). Although this effect is small, it was detected consistently in 3 independent experiments, each with at least 3 replicates, and disappeared when a small region flanking the NF-E2 binding site was deleted from the reporter construct.

To establish *Lims1* as a direct transcriptional target of NF-E2, we used p45 NF-E2–specific antisera to perform ChIP on primary cultured MKs. After isolation of DNA fragments coprecipitated with NF-E2–specific antisera, primers encompassing the distal (−1939 bp) NF-E2 site on the *Lims1* promoter consistently amplified a product, whereas the proximal NF-E2 site and negative controls showed no association with NF-E2 (Figure 5C). These results were measured by qPCR, and the PCR products were also resolved by agarose gel electrophoresis (Figure 5D). Negative controls included preimmune serum and, for PCR, primers that amplify the promoter regions of 2 irrelevant genes, *Apo1a* and HistoneH4a, or a randomly chosen fragment 4.4 kb downstream of the *Lims1* transcription start site. Taken together, these results implicate the distal NF-E2 consensus element in the *Lims1* locus in positive transcriptional regulation by NF-E2. Thus, like other late, NF-E2–dependent MK products such as β1 tubulin and Rab27b, this adaptor protein for integrin-linked kinases is an excellent candidate to mediate cellular mechanisms of platelet assembly and release.

Discussion

The molecular underpinnings of progressive MK differentiation and platelet assembly are insufficiently characterized. We have studied, and here report, gene-expression dynamics during mouse MK maturation and highlight features that shed light on dominant underlying mechanisms. At each step our analysis uncovered the few gene products or pathways previously implicated in MK differentiation or platelet functions, results we take as indication that the bulk of the data are correspondingly robust. We can hence state with increasing confidence that cellular pathways of phosphatidylinositol, retinoic acid, and the Rho, Rac, and Rab small-GTPases likely play vital roles in MK differentiation. Platelet biogenesis would also appear to represent a culmination of cytoskeletal activities that are better defined through the data reported in this paper. Previously we demonstrated a requirement for Rab27b in platelet formation,[16] and others reported a role for the regulator of G-protein signaling RGS16 in MK differentiation.[48] Within the large protein families that accommodate these factors, our analysis identifies additional candidates (Table 2) that may be especially worthy of investigating for specific cellular roles. Additionally, although MK maturation and platelet release require cell movement and significant morphogenesis, there is as yet limited correlation with specific molecular pathways; through the identification of selected adhesion and signaling molecules and genes that regulate shape and motility of other cells, our analysis can begin to address this deficiency.

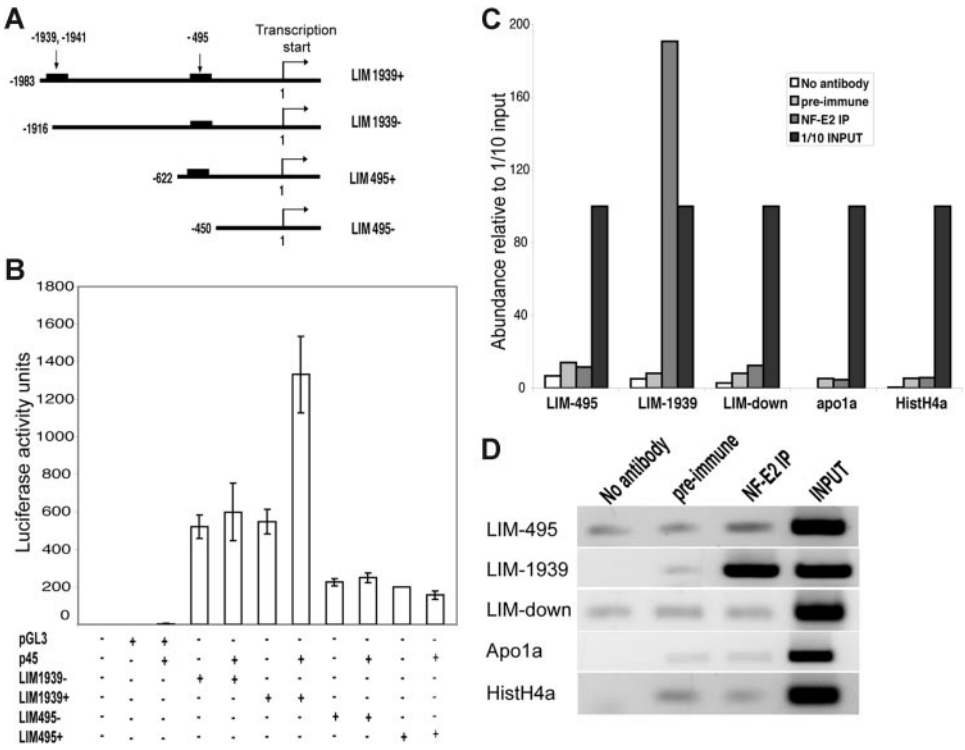

Figure 5. *Lims1* is a direct transcriptional target of NF-E2. (A) Schema of reporter constructs with presence or deletion of putative NF-E2 binding sites, which are indicated in terms of their position relative to the transcription start site (+1). (B) Results of luciferase reporter assays for the *Lims1* promoter region and its deletions variants. COS cells were transfected with the reporter (pGL3 or LIM series) construct in the presence or absence of p45 NF-E2 expression plasmid. (C) Quantitation of MK chromatin IP (ChIP) results for the 2 putative NF-E2 binding sites in the *Lims1* promoter and controls. Immunoprecipitated DNA was amplified by qPCR and the abundance of individual fragments is normalized to that of input (pre-IP) DNA. LIM-495 and LIM-1939 amplify corresponding fragments in the *Lims1* promoter and LIM-down interrogates sequences approximately 4 kb downstream; other controls test for IP of promoter fragments in the *Apo1a* and HistoneH4a genes. (D) Agarose gel electrophoresis of PCR fragments after ChIP. The signals for input DNA with all primer sets and for LIM-1939 after NF-E2 IP were saturated well before the 33 cycles of PCR applied in this experiment; all other signals seem to represent background. Error bars in (B) represent standard deviation.

Our complete data are publicly available[49] and can serve as a resource to investigate platelet ontogeny. However, many transcripts expressed in advanced MKs must represent platelet cargo and not functions required for platelet assembly per se; conversely, some mRNAs found in blood platelets may be stable remnants of MK-specific functions. Thus, there are some inherent barriers to the goal of identifying genes required for MK maturation and platelet release. Nevertheless, our premise, materials, and approach carry significant advantages over prior efforts in MK expression profiling.[29,30] These studies started with human $CD34^+$ cells, which yield MKs with low DNA content and low proplatelet output, hence limiting the ability to characterize platelet biogenesis, a late event in MK maturation. Thus, although our results correlate positively with the published data, they also extend the scope of potential insights into MK ontogeny. One reason for the improvement in this regard is that MK-P contains a more advanced progenitor pool than human $CD34^+$ cells. We also note that, counter to expectation, many more genes peak in expression in the MK-3 population than in MK-6, even though the latter population contains cells with the most developed morphology. Although platelets contain a large protein store, their RNA content is relatively low, and the platelet transcriptome is dominated by mitochondrial RNAs.[27] Taken together, these observations suggest that terminal MK differentiation is accompanied by significant, and apparently selective, mRNA degradation that produces discordant steady-state levels between some RNAs and their protein products.

A small group of TFs, including GATA1, Runx1/CBFβ, and NF-E2, is known to regulate MK maturation,[50] and the corresponding framework made a 3-fold contribution to our analysis. First, among the many TF pathways assessed by GSEA, our analysis highlighted those regulated by GATA1 and CBFβ (Table 1). These results enhance confidence in the pathway analysis and suggest that these TFs help MKs cross from differentiation stages most represented in the MK-P population to those represented in MK-3. Second, we could use MKs derived from GATA1- and NF-E2–deficient MKs, which arrest at different steps,[8,13] to validate a panel of stage-selective MK molecular markers (Figure 3). Together, the results provide support and molecular correlates for the idea that certain lineage-restricted TFs and their corresponding transcriptional programs are key elements in step-wise transitions in MK maturation. Where critical steps previously were defined largely in morphologic terms, our results will enable more precise definition of stages in MK differentiation and, ultimately, improved molecular understanding of platelet biogenesis. Third, we used NF-E2 as the example to identify potential transcriptional targets among genes that are highly expressed in mature MKs, when the need for NF-E2 becomes apparent.[13,14] By combining computational promoter analysis with experimental studies, we identified a novel target, *Lims1*, the gene for an adaptor protein that bridges integrin-linked kinases with cytoskeletal proteins, helps form focal adhesions, and regulates cell shape and motility.[47] The multi-faceted defects in platelet biogenesis in NF-E2–null MKs may result in part from defects in migration or in interactions with the microenvironment, and LIMS1 is a good candidate to mediate these functions. Although early embryonic lethality in LIMS1-null mice, before blood formation,[51,52] currently precludes investigation of its putative thrombopoietic functions in vivo, the approach outlined in this report can help characterize other pertinent pathways.

Note added in proof: Expression data described in this report are deposited in the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/projects/geo; verified January 5, 2007) under the accession series number GSE6593.

Acknowledgments

This work was supported by the National Institutes of Health (R01HL63143). We thank Maris Handley for expert assistance with flow cytometry; Ed Fox, Eleanor Howe, Carol Chang, and John Quackenbush for advice on microarray analysis; Zhe Li for RNA preparation; Chuanyue Wu for Lims1 antibody; and members of our laboratory for helpful discussions and critical review of the manuscript.

Authorship

Contribution: Z.C. and R.A.S. designed and performed the research and wrote the paper; M.H. contributed new analytical tools.

Conflict-of-interest: The authors declare no competing financial interests.

Correspondence: Ramesh A. Shivdasani, Dana-Farber Cancer Institute, One Jimmy Fund Way, Boston, MA 02115; e-mail: ramesh_shivdasani@dfci.harvard.edu.

References

1. Ebbe S, Stohlman FJ, Overcash J, Donovan J, Howard D. Megakaryocyte size in thrombocytopenic and normal rats. Blood. 1968;32:383-392.
2. Odell TT, Jackson CW. Polyploidy and maturation of rat megakaryocytes. Blood. 1968;32:102-111.
3. Zucker-Franklin D. Megakaryocytes and platelets. In: Zucker-Franklin D, Greaves MF, Grossi CE, Marmont AM, eds. Atlas of Blood Cells: Function and Pathology. Vol 1. 2nd ed. Philadelphia, PA: Lea & Febiger; 1989:623-693.
4. Radley JM, Haller CJ. The demarcation membrane system of the megakaryocyte: a misnomer? Blood. 1982;60:213-219.
5. Italiano JE, Lecine P, Shivdasani RA, Hartwig JH. Blood platelets are assembled principally at the ends of proplatelet processes produced by differentiated megakaryocytes. J Cell Biol. 1999;147:1299-1312.
6. Richardson JL, Shivdasani RA, Boers C, Hartwig JH, Italiano JE Jr. Mechanisms of organelle transport and capture along proplatelets during platelet production. Blood. 2005;106:4066-4075.
7. Schulze H, Korpal M, Hurov J, et al. Characterization of the megakaryocyte demarcation membrane system and its role in thrombopoiesis. Blood. 2006;107:3868-3875.
8. Shivdasani RA, Fujiwara Y, McDevitt MA, Orkin SH. A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. EMBO J. 1997;16:3965-3973.
9. Vyas P, Ault K, Jackson CW, Orkin SH, Shivdasani RA. Consequences of GATA-1 deficiency in megakaryocytes and platelets. Blood. 1999;93:2867-2875.
10. Nichols KE, Crispino JD, Poncz M, et al. Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. Nat Genet. 2000;24:266-270.
11. Freson K, Devriendt K, Matthijs G, et al. Platelet characteristics in patients with X-linked macrothrombocytopenia because of a novel GATA1 mutation. Blood. 2001;98:85-92.
12. Mehaffey MG, Newton AL, Gandhi MJ, Crossley M, Drachman JG. X-linked thrombocytopenia caused by a novel mutation of GATA-1. Blood. 2001;98:2681-2688.
13. Shivdasani RA, Rosenblatt MF, Zucker-Franklin D, et al. Transcription factor NF-E2 is required for platelet formation independent of the actions of thrombopoietin/MGDF in megakaryocyte development. Cell. 1995;81:695-704.
14. Lecine P, Villeval J-L, Vyas P, Swencki B, Xu Y, Shivdasani RA. Mice lacking transcription factor NF-E2 validate the proplatelet model of thrombocytopoiesis and show a platelet production defect that is intrinsic to megakaryocytes. Blood. 1998;92:1608-1616.
15. Detter JC, Zhang Q, Mules EH, et al. Rab geranylgeranyl transferase alpha mutation in the gunmetal mouse reduces Rab prenylation and platelet synthesis. Proc Natl Acad Sci U S A. 2000;97:4144-4149.
16. Tiwari S, Italiano JE, Barral DC, et al. A role for Rab27b in NF-E2-dependent pathways of platelet formation. Blood. 2003;102:3970-3979.
17. Lecine P, Italiano JE Jr, Kim SW, Villeval JL, Shivdasani RA. Hematopoietic-specific beta 1 tubulin participates in a pathway of platelet biogenesis dependent on the transcription factor NF-E2. Blood. 2000;96:1366-1373.
18. Schwer HD, Lecine P, Tiwari S, Italiano JE, Hartwig JH, Shivdasani RA. A lineage-restricted and divergent beta-tubulin isoform is essential for the biogenesis, structure and function of blood platelets. Curr Biol. 2001;11:579-586.
19. Clemetson KJ, McGregor JL, James E, Dechavanne M, Luscher EF. Characterization of the platelet membrane glycoprotein abnormalities in Bernard-Soulier syndrome and comparison with normal by surface-labeling techniques and high-resolution two-dimensional gel electrophoresis. J Clin Invest. 1982;70:304-311.
20. Kelley MJ, Jawien W, Ortel TL, Korczak JF. Mutation of MYH9, encoding non-muscle myosin heavy chain A, in May-Hegglin anomaly. Nat Genet. 2000;26:106-108.
21. Consortium TM-HFS. Mutations in MYH9 result in the May-Hegglin anomaly, and Fechtner and Sebastian syndromes. Nat Genet. 2000;26:103-105.
22. Hart A, Melet F, Grossfeld P, et al. Fli-1 is required for murine vascular and megakaryocytic development and is hemizygously deleted in patients with thrombocytopenia. Immunity. 2000;13:167-177.
23. Raslova H, Komura E, Le Couedic J, et al. FLI1 monoallelic expression combined with its hemizygous loss underlies the Paris-Trousseau/Jacobsen thrombopenia. J Clin Invest. 2004;114:77-84.
24. Deveaux S, Cohen-Kaminsky S, Shivdasani RA, et al. p45 NF-E2 regulates the expression of thromboxane synthase in megakaryocytes. EMBO J. 1997;16:5654-5661.
25. Shiraga M, Ritchie A, Aidoudi S, et al. Primary megakaryocytes reveal a role for transcription factor NF-E2 in integrin alpha IIb beta 3 signaling. J Cell Biol. 1999;147:1419-1430.
26. Kerrigan SW, Gaur M, Murphy RP, Shattil SJ, Leavitt AD. Caspase-12: a developmental link between G-protein-coupled receptors and integrin alphaIIbbeta3 activation. Blood. 2004;104:1327-1334.
27. Gnatenko DV, Dunn JJ, McCorkle SR, Weissmann D, Perrotta PL, Bahou WF. Transcript profiling of human platelets using microarray and serial analysis of gene expression. Blood. 2003;101:2285-2293.
28. Kim JA, Jung YJ, Seoh JY, Woo SY, Seo JS, Kim HL. Gene expression profile of megakaryocytes from human cord blood CD34(+) cells ex vivo expanded by thrombopoietin. Stem Cells. 2002;20:402-416.
29. Shim MH, Hoover A, Blake N, Drachman JG, Reems JA. Gene expression profile of primary human CD34+CD38lo cells differentiating along the megakaryocyte lineage. Exp Hematol. 2004;32:638-648.
30. Tenedini E, Fagioli ME, Vianelli N, et al. Gene expression profiling of normal and malignant CD34-derived megakaryocytic cells. Blood. 2004;104:3126-3135.
31. Balduini A, d'Apolito M, Arcelli D, et al. Cord blood in vitro expanded CD41 cells: identification of novel components of megakaryocytopoiesis. J Thromb Haemost. 2006;4:848-860.
32. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc Natl Acad Sci U S A. 2001;98:31-36.
33. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. Genome Biol. 2001;2:research0032.0031-0032.0011.
34. Tamayo P, Slonim D, Mesirov J, et al. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. Proc Natl Acad Sci U S A. 1999;96:2907-2912.
35. Heyer LJ, Kruglyak S, Yooseph S. Exploring expression data: identification and analysis of coexpressed genes. Genome Res. 1999;9:1106-1115.
36. Saeed AI, Sharov V, White J, et al. TM4: a free, open-source system for microarray data management and analysis. Biotechniques. 2003;34:374-378.
37. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545-15550.
38. Mizushima S, Nagata S. pEF-BOS, a powerful mammalian expression vector. Nucleic Acids Res. 1990;18:5322-5327.
39. Avecilla ST, Hattori K, Heissig B, et al. Chemokine-mediated interaction of hematopoietic progenitors with the bone marrow vascular niche is required for thrombopoiesis. Nat Med. 2004;10:64-71.
40. Emadi S, Clay D, Desterke C, et al. IL-8 and its CXCR1 and CXCR2 receptors participate in the control of megakaryocytic proliferation, differentiation, and ploidy in myeloid metaplasia with myelofibrosis. Blood. 2005;105:464-473.
41. Andrews NC, Erdjument-Bromage H, Davidson MB, Tempst P, Orkin SH. Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. Nature. 1993;362:722-728.
42. Lecine P, Blank V, Shivdasani R. Characterization of the hematopoietic transcription factor NF-E2 in primary murine megakaryocytes. J Biol Chem. 1998;273:7572-7578.
43. Mignotte V, Eleouet F, Raich N, Romeo P-H. Cis- and trans-acting elements involved in the regulation of the erythroid promoter of the human porphobilinogen deaminase gene. Proc Natl Acad Sci U S A. 1989;86:6548-6551.
44. National Center for Biotechnology Information. RefSeq: The Reference Sequence. http://www.ncbi.nlm.nih.gov/RefSeq. Accessed January 30, 2005.
45. Argonne National Laboratory. PatScan. http://www-unix.mcs.anl.gov/compbio/PatScan. Accessed February 5, 2005.
46. National Center for Biotechnology Information. HomoloGene. www.ncbi.nlm.nih.gov/HomoloGene. Accessed February 5, 2005.
47. Wu C. PINCH, N(i)ck and the ILK: network wiring at cell-matrix adhesions. Trends Cell Biol. 2005;15:460-466.
48. Berthebaud M, Riviere C, Jarrier P, et al. RGS16 is a negative regulator of SDF-1-CXCR4 signaling in megakaryocytes. Blood. 2005;106:2962-2968.
49. http://research2.dfci.harvard.edu/dfci/shivdasani/. Accessed December 13, 2005.
50. Schulze H, Shivdasani RA. Molecular mechanisms of megakaryocyte differentiation. Semin Thromb Hemost. 2004;30:389-398.
51. Liang X, Zhou Q, Li X, et al. PINCH1 plays an essential role in early murine embryonic development but is dispensable in ventricular cardiomyocytes. Mol Cell Biol. 2005;25:3056-3062.
52. Li S, Bordoy R, Stanchi F, et al. PINCH1 regulates cell-matrix and cell-cell adhesions, cell polarity and cell survival during the peri-implantation stage. J Cell Sci. 2005;118:2913-2921.

2007 109: 1451-1459
doi:10.1182/blood-2006-08-038901 originally published online October 17, 2006

Expression analysis of primary mouse megakaryocyte differentiation and its application in identifying stage-specific molecular markers and a novel transcriptional target of NF-E2

Zhao Chen, Michael Hu and Ramesh A. Shivdasani

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.