Characterization of the Murine mafF Gene*

Ko Onodera‡, Jordan A. Shavit‡, Hozumi Motohashi, Fumiki Katsuoka§, Jun-etsu Akasaka§, James Douglas Engel‡,§ and Masayuki Yamamoto§

From the ‡Department of Biochemistry, Molecular Biology, and Cell Biology, Northwestern University, Evanston Illinois 60208-3500 and §Institute of Basic Medical Sciences and Center for TARA, University of Tsukuba, 1-1 Tennoudai, Tsukuba 305, Japan

Small Maf proteins are obligatory heterodimeric partner molecules of mammalian Cap’n’Collar proteins that together control a wide variety of eukaryotic genes. Although both MafK and MafG are expressed in overlapping but distinct tissue distribution patterns during embryonic development, the physiological consequences of loss-of-function mutations in either gene are modest. This suggested that compensation by the third small Maf protein, MafF, might be a major reason for such mild phenotypes and that further analysis of MafF might therefore provide important insights for understanding small Maf regulatory function(s). We therefore cloned, mapped, transcriptionally and developmentally characterized, and finally disrupted the mafF gene. We show that murine mafF is transcriptionally regulated by three different promoters and is most abundantly expressed in the lung. The lacZ gene inserted into the mafF locus revealed prominent expression sites in the gut, lung, liver, outflow tract of the heart, cartilage, bone membrane, and skin but not in hematopoietic cells at any developmental stage. Homozygous mafF null mutant mice were born in a normal Mendelian ratio and displayed no obvious functional deficiencies, indicating that MafF activity may be dispensable even in tissues where the expression of other small Maf proteins is quite low.

How a single cell differentiates into a variety of different tissues and organs to form a coherent multicellular organism is the central question of developmental biology. Transcription factors have been shown both genetically and biochemically to be key mediators of cellular differentiation, because they bind to specific DNA sequences to activate specific target gene transcription and often appear to be direct targets of cellular responses to inductive cues from the environment. The c-Maf transcription factor, originally discovered as a retroviral transduced oncogene (1), was the founding member of a family of six related proteins that have been identified in vertebrate organisms. All of the Maf family members possess a conserved basic region closely flanked by a heptad repeat motif (bZip), which mediates DNA binding and dimer formation, respectively. These proteins commonly recognize a specific palindromic sequence, TGCTGACTCAGCA or TGCTGACGTCAGCA, called a MARE (maf recognition element (2, 3)) through which they potentiate tissue-specific gene transcription.

The Maf family is subdivided according to molecular size. The large Maf proteins, c-Maf, MafB, and NRL, have similar conserved sequence domains and also possess potent trans-activation sequences (2, 4, 5). The small Maf proteins, MafF, MafG, and MafK, do not possess trans-activation activity, and indeed these three genes encode proteins with little more than the basic and heptad repeat motifs, which identify them as members of the Maf protein family (6, 7).

The small Mafs can affect transcription either by forming heterodimers with other bZip factors or by forming homodimers and heterodimers among themselves. When two small Maf proteins homo- or heterodimerize, they bind to MAREs and negatively regulate transcription, either indirectly, by competitively inhibiting the binding of activating molecules, or directly, by repressing transcription from the basal apparatus (8).

Over the last several years, three groups of small Maf interacting molecules, which can form either productive or unproductive complexes, have been described (3, 9). The most physiologically important of these partner molecules are probably the vertebrate CNC1 transcription factors (45, Nrf1, Nrf2, and Nrf3), so named originally for their sequence homology to the Drosophila Cap’n’Collar protein (10). The mafF gene showed that murine mafF is transcriptionally regulated by three different promoters and is most abundantly expressed in the lung. The lacZ gene inserted into the mafF locus revealed prominent expression sites in the gut, lung, liver, outflow tract of the heart, cartilage, bone membrane, and skin but not in hematopoietic cells at any developmental stage. Homozygous mafF null mutant mice were born in a normal Mendelian ratio and displayed no obvious functional deficiencies, indicating that MafF activity may be dispensable even in tissues where the expression of other small Maf proteins is quite low.

The Maf family is subdivided according to molecular size. The large Maf proteins, c-Maf, MafB, and NRL, have similar conserved sequence domains and also possess potent trans-activation sequences (2, 4, 5). The small Maf proteins, MafF, MafG, and MafK, do not possess trans-activation activity, and indeed these three genes encode proteins with little more than the basic and heptad repeat motifs, which identify them as members of the Maf protein family (6, 7).

The small Mafs can affect transcription either by forming heterodimers with other bZip factors or by forming homodimers and heterodimers among themselves. When two small Maf proteins homo- or heterodimerize, they bind to MAREs and negatively regulate transcription, either indirectly, by competitively inhibiting the binding of activating molecules, or directly, by repressing transcription from the basal apparatus (8).

Over the last several years, three groups of small Maf interacting molecules, which can form either productive or unproductive complexes, have been described (3, 9). The most physiologically important of these partner molecules are probably the vertebrate CNC1 transcription factors (45, Nrf1, Nrf2, and Nrf3), so named originally for their sequence homology to the Drosophila Cap’n’Collar protein (10). The mafF gene showed that murine mafF is transcriptionally regulated by three different promoters and is most abundantly expressed in the lung. The lacZ gene inserted into the mafF locus revealed prominent expression sites in the gut, lung, liver, outflow tract of the heart, cartilage, bone membrane, and skin but not in hematopoietic cells at any developmental stage. Homozygous mafF null mutant mice were born in a normal Mendelian ratio and displayed no obvious functional deficiencies, indicating that MafF activity may be dispensable even in tissues where the expression of other small Maf proteins is quite low.

The Maf family is subdivided according to molecular size. The large Maf proteins, c-Maf, MafB, and NRL, have similar conserved sequence domains and also possess potent trans-activation sequences (2, 4, 5). The small Maf proteins, MafF, MafG, and MafK, do not possess trans-activation activity, and indeed these three genes encode proteins with little more than the basic and heptad repeat motifs, which identify them as members of the Maf protein family (6, 7).

The small Mafs can affect transcription either by forming heterodimers with other bZip factors or by forming homodimers and heterodimers among themselves. When two small Maf proteins homo- or heterodimerize, they bind to MAREs and negatively regulate transcription, either indirectly, by competitively inhibiting the binding of activating molecules, or directly, by repressing transcription from the basal apparatus (8).

Over the last several years, three groups of small Maf interacting molecules, which can form either productive or unproductive complexes, have been described (3, 9). The most physiologically important of these partner molecules are probably the vertebrate CNC1 transcription factors (45, Nrf1, Nrf2, and Nrf3), so named originally for their sequence homology to the Drosophila Cap’n’Collar protein (10). The mafF gene showed that murine mafF is transcriptionally regulated by three different promoters and is most abundantly expressed in the lung. The lacZ gene inserted into the mafF locus revealed prominent expression sites in the gut, lung, liver, outflow tract of the heart, cartilage, bone membrane, and skin but not in hematopoietic cells at any developmental stage. Homozygous mafF null mutant mice were born in a normal Mendelian ratio and displayed no obvious functional deficiencies, indicating that MafF activity may be dispensable even in tissues where the expression of other small Maf proteins is quite low.

How a single cell differentiates into a variety of different tissues and organs to form a coherent multicellular organism is the central question of developmental biology. Transcription factors have been shown both genetically and biochemically to be key mediators of cellular differentiation, because they bind to specific DNA sequences to activate specific target gene transcription and often appear to be direct targets of cellular responses to inductive cues from the environment. The c-Maf transcription factor, originally discovered as a retroviral transduced oncogene (1), was the founding member of a family of six related proteins that have been identified in vertebrate organisms. All of the Maf family members possess a conserved basic region closely flanked by a heptad repeat motif (bZip), which mediates DNA binding and dimer formation, respectively. These proteins commonly recognize a specific palindromic sequence, TGCTGACTCAGCA or TGCTGACGTCAGCA, called a MARE (maf recognition element (2, 3)) through which they potentiate tissue-specific gene transcription.

The Maf family is subdivided according to molecular size. The large Maf proteins, c-Maf, MafB, and NRL, have similar conserved sequence domains and also possess potent trans-activation sequences (2, 4, 5). The small Maf proteins, MafF, MafG, and MafK, do not possess trans-activation activity, and indeed these three genes encode proteins with little more than the basic and heptad repeat motifs, which identify them as members of the Maf protein family (6, 7).

The small Mafs can affect transcription either by forming heterodimers with other bZip factors or by forming homodimers and heterodimers among themselves. When two small Maf proteins homo- or heterodimerize, they bind to MAREs and negatively regulate transcription, either indirectly, by competitively inhibiting the binding of activating molecules, or directly, by repressing transcription from the basal apparatus (8).

Over the last several years, three groups of small Maf interacting molecules, which can form either productive or unproductive complexes, have been described (3, 9). The most physiologically important of these partner molecules are probably the vertebrate CNC1 transcription factors (45, Nrf1, Nrf2, and Nrf3), so named originally for their sequence homology to the Drosophila Cap’n’Collar protein (10). The mafF gene showed that murine mafF is transcriptionally regulated by three different promoters and is most abundantly expressed in the lung. The lacZ gene inserted into the mafF locus revealed prominent expression sites in the gut, lung, liver, outflow tract of the heart, cartilage, bone membrane, and skin but not in hematopoietic cells at any developmental stage. Homozygous mafF null mutant mice were born in a normal Mendelian ratio and displayed no obvious functional deficiencies, indicating that MafF activity may be dispensable even in tissues where the expression of other small Maf proteins is quite low.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM and EBI Data Bank with accession number(s) AB009694 and AB009693.

* This work was supported by fellowships from the Japan Society for the Promotion of Science (to K. O.) and National Institutes of Health (NIH) Medical Scientist Training Program Grant GM08152 to Northwestern University (to J. A. S.). Research support was provided by the Ministry of Education, Science, and Culture, CREST, and JSPS as well as core facility support from Robert H. Lurie Comprehensive Cancer Center Grant CA60553 and NIH Grant CA80088 (to J. D. E.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The abbreviations used are: CNC, Cap’n’Collar protein; ES, embryonic stem cells; RACE, 5′-rapid amplification of cDNA ends; kbp, kilobase pair(s); nls, nuclear localization signal; RT-PCR, reverse transcription-polymerase chain reaction; dpc, days post-coitus; RFLP, restriction fragment length polymorphism.

21162 This paper is available on line at http://www.jbc.org
Loss of function mutants for several of the CNC proteins have been generated by targeted mutagenesis in embryonic stem (ES) cells followed by the generation of germ line mutant mice. p45 mutant mice displayed impaired platelet formation but little disturbance of erythropoiesis (15). nrf1 germ line mutation conferred a noncell autonomous mesodermal embryonic lethal defect (20, 21), whereas nrf2 null mutants were viable (22) but had impaired reaction to xenobiotics (23). Because all three of these CNC proteins are capable of activating transcription from MARE elements in tissue culture co-transfection assays, the lack of prominent erythroid phenotypes in any of the mutant mice indicated that these activities could be compensating for one another in erythroid cells. This conclusion in turn indicated that unraveling the transcriptional responses mediated by individual MARE-binding proteins was going to be complicated.

To clarify in vivo functions of this extensively interacting regulatory network, we previously described null mutations in both the murine mafG and mafK loci. These studies showed that the mafK loss of function resulted in no discernable phenotype in homozygous mutant animals (24, 25), whereas mafG mutant mice displayed both hematopoietic and behavioral deficiencies (25). Analysis of the mafK- and mafG-directed expression patterns (following insertion of lacZ into both chromosomal loci) showed, to our surprise, that both factors were expressed in highly tissue-restricted spatial and temporal patterns during embryonic development. This conclusion clearly contradicted our earlier perception that the small Maf proteins were ubiquitously expressed (13, 26). However, the lack of severe phenotypes in these mutant mice implied that the pattern of expression of the small Maf proteins might be extensively overlapping, suggesting that one small Maf might be able to compensate for the loss of another during development. If this hypothesis was correct, it would be very difficult to ascribe specific functional roles to individual small Maf family members because the effects of mutation of one protein would be masked by the expression of another.

To complete the initial analysis of this family to determine the contribution(s) of the small Maf proteins to this growing regulatory network, we cloned and characterized the gene encoding the third small Maf protein, mafF. MafF, like MafK and MafG, is expressed in a highly tissue-restricted pattern during embryonic development and is also found in some tissues that overlap expression of the other small Mafs in pattern, but it is also unique to a subset of tissues where the other proteins do not significantly accumulate. Homozygous mafF mutant mice were viable and apparently normal and healthy. The data show that although MafF, like MafK and MafG, is expressed in a distinct pattern during development, its loss of function in vivo is probably fully compensated (3).

EXPERIMENTAL PROCEDURES

Cloning the Murine mafF Gene—The mafF gene was cloned from a 129/SvJ genomic DNA library in XFix vector (Stratagene). The chicken mafF cDNA clone (6) was used as a probe under low stringency screening conditions. Filter hybridization was carried out in 50% formamide, 5 × SSC (0.15 M NaCl and 0.015 M sodium citrate), 20 mm Tris-Cl, pH 7.6, 1 × Denhardt’s solution, 10% dextran sulfate, and 0.1% SDS at 42 °C overnight. Filters were washed in 2 × SSC, 0.1% SDS at 37 °C for 1 h. Positive clones were purified by three additional plaque purifications. The remaining positive clones were subcloned into the Bluescript II (Stratagene). DNA sequences were determined by automated analysis (ABI, model 373A). Sequences of the murine mafF and mafG clones are available from Genbank® (accession numbers AB009684 and AB009693, respectively).

Rapid Amplification of cDNA Ends Assay—Rapid amplification of cDNA ends (5′-RACE) assays were performed utilizing the Marathon cDNA amplification kit following the manufacturer’s instructions (CLONTECH). The sequences of the gene-specific primers (mMafF-F2 and mMafF-F1 nested primers) are shown in Table I. RACE products were subcloned into pGEM-T (Promega), and their sequences were determined as described above.

Chromosomal Mapping of mafF and mafG—DNA was obtained from the (C57Bl/6J × 129SvJ)F1 mice (25) by the method of Wang et al. (26). The chromosomal loci of mafF, mafG, and mafK were mapped in interspecific backcrosses generated by The Jackson Laboratory (strain Bar Harbor Mus musculus, Jackson-son BKS). For the mafF locus, a 1.0-kbp XbaI fragment 5′ to the second exon was used to detect a TaqI restriction fragment length polymorphism (RFLP). The more intensely hybridizing Mus spretus and C57Bl/6J bands were 6.5 kbp and 1.9 kbp, respectively, with 9.8 and 0.8 kbp faintly hybridizing fragments. Taq1 digestion of DNA from each of the two parental species was performed, followed by electrophoresis and blotting to a Nylon membrane. Hybridization and washing were performed using standard conditions. For the mafG locus, a 5.3-kbp BamHI fragment that spanned the second and third exons was used to detect the Taq1 RFLP. The M. spretus fragment was 4.0 kbp, whereas C57Bl/6J DNA showed two bands of 1.9 and 1.7 kbp. In addition, there was a nonpolymorphic 1.3-kbp hybridizing band.

The typing data for the segregation of these two polymorphisms were submitted to The Jackson Laboratory Backcross DNA panel mapping resource and compared with approximately 3000 loci that had already been mapped. Linkage was observed for mafF to markers on chromosome 15 and for mafG to the distal end of chromosome 11.

Construction of the mafF Targeting Vector—To generate the mafF ES cell targeting vector, a 7.8-kbp Spel-BamHI genomic fragment and a 1.9-kbp Spel-XbaI fragment were used. This replacement strategy removes a 1.1-kbp BamHI-SphI fragment that includes the entire mafF coding sequence except for the first three amino acids.

In the first step of the cloning strategy, nuclear localization signal (nls)-lacZ was introduced 3′ to the 7.8-kbp Spel-BamHI fragment, producing plasmid p7.8SB-nls-lacZ. To confirm that no mutations were introduced during cloning, the insert was sequenced to ensure that the nls-lacZ gene was translated in-frame from the mafF initiation codon. Simultaneously, a loxP-flanked McI1promoter-neomycin resistance gene-poly(A)-McI1promoter-thymidine kinase-poly(A) (called MC1 neoMC1tk (25)) cassette and a McI1promoter-diphtheria toxin (MC1-DT (27)) cassette were cloned 5′ and 3′ to the 1.9-kbp Spel-XbaI genomic fragment, producing loxPM1neoMC1tk-1.9X-DT. To minimize the possibility of affecting nls-lacZ expression, the loxPM1neoMC1tk cassette was introduced in reverse orientation with respect to mafF. In the second cloning step, the insert of p7.8SB-nls-lacZ was introduced into ploxPM1neoMC1tk-1.9X-DT, producing the final mafF targeting vector.

Generation of mafF Mutant Mice—The targeting vector was linearized and introduced into RI ES cells (28) by electroporation. Recombined ES cells were selected for G418 resistance (215 μg of active component/ml; Life Technologies, Inc.). After 8 days of selection, G418-resistant colonies were picked and subjected to PCR screening. After confirming homologous recombination by Southern blot analysis, positive ES cell clones were aggregated with CD-1 eight-cell stage embryos. After overnight culture, the blastocysts were transferred into the uterus of pseudopregnant foster females. Fourteen highly chimeric 129/SvEv+/−;CD1-1 mice (>80%) male mice were born starting with four different ES cell clones.

To confirm germ line transmission, the chimeric mice were intercrossed to CD1 females. Three of the ten chimeric animals transmitted the mutation through the germ line, and each was derived from a different ES cell clone. F1 and F2 heterozygous mafF mutants were intercrossed to determine any effects on viability, and the resulting pups were genotyped by PCR. The results were the same for all three mafF mutant lines.

Neonycin Gene Excision by Cre Recombinase—We removed the MC1neoMC1tk cassette using Cre-loxP-mediated recombination, because the McI1 promoter 5′ to the neomycin or thymidine kinase genes could, in principal, affect nls-lacZ expression from the endogenous mafF promoter. For this purpose, we injected a Cre expression plasmid into MC1neoMC1tk-specific embryonic stem cells. The resulting embryonic stem cell clones were subjected to PCR analysis, and progeny were analyzed with PCR and Southern blotting (data not shown). One mouse (line 801) was found to have precisely excised the targeted cassette from the MC1neoMC1tk allele. Initially, the loxPM1neoMC1tk allele was used to determine the segregation of these polymorphisms. The more intensely hybridizing 3′ RFLP. The 3′ RFLP was used to confirm the presence of the MC1neoMC1tk allele.

RT-PCR and Southern Blot Analysis—For the detection of mRNA expression, we used real time PCR analysis using a probe labeled with a reporter fluorescent dye (TaQMan probe). It was demonstrated that

http://lena.jax.org/.
Murine mafF Characterization

### Table I

| Small Maf oligonucleotide primers | 5'-RACE | RT-PCR |
|-----------------------------------|---------|--------|
| mMafF-F1:                         | TTC AGG GCT TCG GAT GAA GGA TCC A | CAT CAG GGC TTA TAC CGA |
| mMafF-F2:                         | TGT GCT CGC TCA ACT CGC GCT TGA TCT T | (FAM)-TGA AGG TCA AGC GCC AGT TGA GC-(TAMRA) |
| mMafF-P:                          | GTC GAT CTC TCA AGC AAA G | CAT CAG GGC TTA TAC CGA |
| mMafG-F:                          | CAT CAG GGC TTA TAC CGA | GAC AGC GCC GGC GGT ATG |
| mMafG-R:                          | AGC TCA TCA TCT CTA AGA ACA GG | AGC TCA TCA TCT CTA AGA ACA GG |
| mMafK-F:                          | AGA GGT CAT CTT CTG CGA AA | CAG CTT GTC CAC TTC AGC CT |
| mMafK-R:                          | GCC TAT AAA GGA AAC AAG GCA | CAG CTT GTC CAC TTC AGC CT |
| mmfu1:                            | GAC AGC GCC GGC GGT ATG | (FAM)-TGA AGG TCA AGC AGG GCG-(TAMRA) |
| mmfd1:                            | (FAM)-TGA AGG TCA AGC AGG GCG-(TAMRA) | (FAM)-TGA AGG TCA AGC AGG GCG-(TAMRA) |

this analysis provides a quantitative measurement for mRNA levels (29).

cDNAs were synthesized from the RNA samples using murine leukemia virus reverse transcriptase (Superscript II; Life Technologies, Inc.) and random hexamer oligonucleotide primers. TaqMan probes, primers, and cDNAs were added to the master mixture that contained all reagents for PCR (Perkin-Elmer). Standard curves for MafF, MafG, MafK, and ribosomal RNA were generated by serial dilution of cDNA derived from intestine RNA. Sequences of the probes and primers are described in Table I. The primers for detection of MafF cDNA were mMafF-F and mMafF-R, and the probe was mMafF-F, mMafF-R (primers), and mMafF-P (probe) were used for MafF. mMafK-F, mMafK-R (primers), and mMafK-P (probe) were used for MafK. The ribosomal RNA primers and probe (TaqMan Ribosomal RNA control reagents) were purchased from Perkin-Elmer.

The relative increase in reporter fluorescent dye emission because of this analysis provides a quantitative measurement for mRNA levels (29).

cDNAs were synthesized from the RNA samples using murine leukemia virus reverse transcriptase (Superscript II; Life Technologies, Inc.) and random hexamer oligonucleotide primers. TaqMan probes, primers, and cDNAs were added to the master mixture that contained all reagents for PCR (Perkin-Elmer). Standard curves for MafF, MafG, MafK, and ribosomal RNA were generated by serial dilution of cDNA derived from intestine RNA. Sequences of the probes and primers are described in Table I. The primers for detection of MafF cDNA were mMafF-F and mMafF-R, and the probe was mMafF-F, mMafF-R (primers), and mMafF-P (probe) were used for MafF. mMafK-F, mMafK-R (primers), and mMafK-P (probe) were used for MafK. The ribosomal RNA primers and probe (TaqMan Ribosomal RNA control reagents) were purchased from Perkin-Elmer.

The relative increase in reporter fluorescent dye emission because of the progression of the PCR reaction was monitored in real time by the use of TaqMan probes, primers, and cDNAs were added to the master mixture that contained all reagents for PCR (Perkin-Elmer). Standard curves for MafF, MafG, MafK, and ribosomal RNA were generated by serial dilution of cDNA derived from intestine RNA. Sequences of the probes and primers are described in Table I. The primers for detection of MafF cDNA were mMafF-F and mMafF-R, and the probe was mMafF-F, mMafF-R (primers), and mMafF-P (probe) were used for MafF. mMafK-F, mMafK-R (primers), and mMafK-P (probe) were used for MafK. The ribosomal RNA primers and probe (TaqMan Ribosomal RNA control reagents) were purchased from Perkin-Elmer.

The relative increase in reporter fluorescent dye emission because of the progression of the PCR reaction was monitored in real time by the use of TaqMan probes, primers, and cDNAs were added to the master mixture that contained all reagents for PCR (Perkin-Elmer). Standard curves for MafF, MafG, MafK, and ribosomal RNA were generated by serial dilution of cDNA derived from intestine RNA. Sequences of the probes and primers are described in Table I. The primers for detection of MafF cDNA were mMafF-F and mMafF-R, and the probe was mMafF-F, mMafF-R (primers), and mMafF-P (probe) were used for MafF. mMafK-F, mMafK-R (primers), and mMafK-P (probe) were used for MafK. The ribosomal RNA primers and probe (TaqMan Ribosomal RNA control reagents) were purchased from Perkin-Elmer.

The relative increase in reporter fluorescent dye emission because of the progression of the PCR reaction was monitored in real time by the use of TaqMan probes, primers, and cDNAs were added to the master mixture that contained all reagents for PCR (Perkin-Elmer). Standard curves for MafF, MafG, MafK, and ribosomal RNA were generated by serial dilution of cDNA derived from intestine RNA. Sequences of the probes and primers are described in Table I. The primers for detection of MafF cDNA were mMafF-F and mMafF-R, and the probe was mMafF-F, mMafF-R (primers), and mMafF-P (probe) were used for MafF. mMafK-F, mMafK-R (primers), and mMafK-P (probe) were used for MafK. The ribosomal RNA primers and probe (TaqMan Ribosomal RNA control reagents) were purchased from Perkin-Elmer.

The relative increase in reporter fluorescent dye emission because of the progression of the PCR reaction was monitored in real time by the use of TaqMan probes, primers, and cDNAs were added to the master mixture that contained all reagents for PCR (Perkin-Elmer). Standard curves for MafF, MafG, MafK, and ribosomal RNA were generated by serial dilution of cDNA derived from intestine RNA. Sequences of the probes and primers are described in Table I. The primers for detection of MafF cDNA were mMafF-F and mMafF-R, and the probe was mMafF-F, mMafF-R (primers), and mMafF-P (probe) were used for MafF. mMafK-F, mMafK-R (primers), and mMafK-P (probe) were used for MafK. The ribosomal RNA primers and probe (TaqMan Ribosomal RNA control reagents) were purchased from Perkin-Elmer.

### RESULTS

The mafF Gene Is Transcribed from Three Different Promoters—The murine mafF locus was cloned from a 129/SvJ genomic DNA library using a chicken MafF cDNA (6) and screening under low stringency conditions. Three recombinants were identified, and all three were found to overlap the mafF locus. After detailed structural analysis (below), the murine mafF locus was found to be composed of three exons—exon Ia, exon Ib, and exon II. The 5'-RACE analysis was performed (“Experimental Procedures”). Exon Ib was found using adult lung mRNA as the initial template, and two additional first exons called Ia and Ic (as well as Ib) were found when 16.5 dpc whole embryo mRNA was used as the initial reverse transcriptase substrate (Fig. 2). The proximal promoter region of mafF exon Ib shows extensive homology with the chicken mafF first exon.3 We conclude that the murine mafF locus is organized with two coding exons, in similar fashion to the mafG and mafK genes, but that three different promoters are alternatively spliced to the common exon II to generate three MafF mRNAs that encode the same protein product.

Chromosomal Mapping of the mafF and mafG Genes—Interspecific backcross mapping was employed to map the chromosomal position of the mafF locus. A TaqI RFLP was used for screening an interspecific backcross panel generated by The Jackson Laboratory (“Experimental Procedures”). After Southern blotting of the DNA from 94 backcross progeny, the results were submitted to a data base search containing approximately 3000 previously typed markers from the same backcross, thereby placing the mafF locus on chromosome 15 (Fig. 3). At the same time, we also determined that the mafG locus was on mouse chromosome 11 (Fig. 3). Together with a prior report describing the position of the mafK locus on chromosome 5 (30), these data complete the chromosomal positioning of all three small maf genes. Neither of the positions corresponds to a murine locus or human disease phenotype that might be because of the loss of mafF or mafG function (below).

Targeting of the lacZ Gene to Replace mafF Coding Sequences—To generate mafF null mutant mice, a targeting vector was designed as depicted in Fig. 4A. Except for the first three amino acids, the entire coding region of mafF was deleted by substitution with an in-frame lacZ fusion bearing a nuclear localization sequence (“Experimental Procedures”). At the same time, we introduced tandemly repeated loxP sites into the targeting vector, surrounding the neomycin resistance/thymidine kinase selection cassette, anticipating later excision by Cre recombinase, because it has been shown that the introduction of foreign promoters into genomic loci can disrupt normal gene regulation (e.g., Ref. 31).

The targeting vector (“Experimental Procedures”) was electroporated into R1 ES cells, and clones were positively selected by culturing in the presence of G418. Of 240 ES cell transformants, 11 were found to be homologous recombinants by PCR and Southern blotting (Fig. 4B). Several of these ES clones were then aggregated with CD1 8-cell stage embryos to generate chimeras (32), and then F1 mice were generated by intercrossing the chimeras to CD1 females. From 10 initial chimeric...
males derived from the aggregation of four different ES cell clones, three transmitted the mutant allele through the germ line. F1 and subsequent generation mice heterozygous for the mafF mutant allele were used for further intercross analysis. All three lines (in addition to a fourth, in which the positive selection cassette was removed by Cre microinjection into eggs of F1 backcrosses (33)) that were established using the three different ES cell clones gave identical genetic and expression results (below).

**mafF Is Expressed in a Highly Restricted Pattern during Embryonic Development**—To clarify possibly unique versus potentially compensated MafF function(s), we analyzed lacZ expression in heterozygous or homozygous gene-targeted mafF mutant embryos as a reflection of natural MafF mRNA temporal- and tissue-specific accumulation. We then compared this pattern to those generated by the mafK- and mafG/lacZ mutant embryos (25).

LacZ expression was detected in the egg cylinder of mafF homozygous mutant embryos as early as 6.5 dpc, in the boundary between the extraembryonic and embryonic regions (Fig. 5A). This pattern is clearly distinct from that conferred by the mafG gene, which is expressed strongest in the embryonic region, and mafK, which is far more abundant in extraembryonic tissue. At 8.5 dpc, mafF/lacZ was weakly expressed in the future gut, allantois, and yolk sac endoderm but more strongly in the ectoplacental cone (Fig. 5B). Once again this pattern is in contrast to mafG expression at the same stage, which is strongly expressed throughout the embryo proper, whereas mafK expression is strong in the yolk sac endoderm and ectoplacental cone (25). At 9.5 dpc, strong mafF expression was observed in the primordial gut (Fig. 5C) and presumptive fetal liver (not shown), in the floorplate of the myelencephalon as well as in neural crest cells (not shown), and in spongiotrophoblasts and giant cells of the placenta (Fig. 5C). Although mafF-directed β-galactosidase synthesis was detected in embryonic hematopoietic organs, at 9.5 dpc, mafF/lacZ-expressing cells in the yolk sac were clearly endodermal rather than hematopoietic (Fig. 5D). These data suggested that MafF might not play a significant role in embryonic hematopoietic gene regulation, in contrast to MaF and Maf0, which are both strongly expressed in hematopoietic cells (25). At 12.5 dpc, strong lacZ expression was also detected in the outflow tract of the heart (Fig. 5F) in a specific subset of dorsal root ganglia (Fig. 5G) and cranial nerve ganglia (not shown) as well as in the lung primordium and in the epithelium of the expiratory tract (below). Thus all three small maf gene expression patterns are quite distinct, but often overlapping, during embryogenesis.

In newborn animals, strong mafF expression was observed in keratinocytes (Fig. 6A) and in the cartilage and bone membrane (Fig. 6B). Intense expression persists in bronchial epithelia (Fig. 6C). As in earlier periods during development, both the bone marrow (Fig. 6D) and spleen (not shown) display scant
β-galactosidase staining, underscoring the possibility that mafF may not play a significant role in hematopoiesis.

Finally, we examined expression of all three small Maf mRNAs using a quantitative fluorescence ("Taqman") PCR assay. After preparing total RNA from various tissues in which we had previously detected expression of one or another of these mRNAs (or of the β-galactosidase reporter gene), we found that the primers used for PCR amplification of the three cDNAs interfered (to some extent) both with one another and with an internal control included in the reactions. We therefore quantified each mRNA relative only to its own peak expression level, and these results are presented in Fig. 7. Not unexpectedly, each of the small Maf mRNAs displayed a unique expression pattern: although MafK and MafF were most abundantly expressed in the lung, MafG appeared to be most abundant in the heart. Although MafG abundance differs by no more than 10-fold in any of the tissues examined, MafF abundance in different tissues varied by more than 65-fold.

The important conclusion that we can draw from this very sensitive assay is that one can detect each small Maf mRNA in each organ system or individual tissue that was analyzed. Whether or not this expression is representative of the multiple different cell types that constitute each organ system or whether in some cases low level expression might simply represent the level of contamination (e.g., of residual blood in muscle or heart) is not known at the present time. However, these data, when taken together with the lacZ expression data, suggest that although mafF might play important transcriptional regulatory roles during embryonic development, its loss of function could be masked in many, if not all, developing organs by compensation from the mafK and/or mafG genes.

mafF Null Mutant Mice Are Viable and Fertile—Homozygous mutant mice were generated by intercrossing CD1 F1 and subsequent generation mafF heterozygous mutant animals. The genotype of the pups was confirmed by PCR and Southern blotting (Fig. 4C). As anticipated, embryonic expression anal-
Murine mafF Characterization

**DISCUSSION**

In this study, we report the cloning, transcriptional, and expression analysis of the third gene encoding small Maf proteins, *mafF*, to clarify the activities of this family during embryonic development. As are the two other small Maf proteins, MafF is highly restricted in its temporal and tissue specificity, thus finally and definitively dispelling the early notion that the small Maf proteins were simply ubiquitously expressed partners of the more highly restricted CNC family trans-activating molecules. In fact, the converse appears to be more general, because both Nrf1 and Nrf2 are far more widely expressed than the small Maf proteins (21, 22).

When we initially cloned the mouse *mafF* gene, structural analysis revealed that it could be transcribed from three different promoters lying close to one another but separated by about 8 kbp from the first coding exon, exon II, used by all three promoters. The *mafF* gene structure is thus similar to the *mafK* gene, structural organization of *mafF* is highly restricted in its temporal and tissue specificity, whereas MaF and MaK are both expressed in only the small Maf proteins (21, 22).

When we initially cloned the mouse *mafF* gene, structural analysis revealed that it could be transcribed from three different promoters lying close to one another but separated by about 8 kbp from the first coding exon, exon II, used by all three promoters. The *mafF* gene structure is thus similar to the *mafK* gene, structural organization of *mafF* is highly restricted in its temporal and tissue specificity, whereas MaF and MaK are both expressed in only the small Maf proteins (21, 22).

When we initially cloned the mouse *mafF* gene, structural analysis revealed that it could be transcribed from three different promoters lying close to one another but separated by about 8 kbp from the first coding exon, exon II, used by all three promoters. The *mafF* gene structure is thus similar to the *mafK* gene, structural organization of *mafF* is highly restricted in its temporal and tissue specificity, whereas MaF and MaK are both expressed in only the small Maf proteins (21, 22).
and/or mating), MafF and MafK must confer a selective advantage to an animal bearing the gene over one in which it has been lost.

Previous studies have indicated how the small Maf proteins may be involved in both erythroid- and megakaryocyte-specific gene regulation. In consideration of the embryonic expression profile of the mafF/lacZ profile as well as the hematological analyses of mice missing the factor (Table III), it seems likely that MafF does not play a significant role in hematopoiesis. Given this conclusion then, our previous results from analysis of the mafK and mafG homozygous loss of function mutations may be interpreted as follows. In erythroid lineage cells, both MafK and MafG provide equivalent, fully redundant, activity for binding to, and activating transcription from, erythroid MARE elements. In this lineage, each homozygous mutant mouse (in mafK or mafG) displays no phenotype because of full compensation by the other small Maf protein. In contrast, MafG confers the predominant small Maf activity in megakaryocytes, and thus, MafK cannot fully compensate for MafG loss in thrombopoiesis (25). Furthermore, this explanation predicts that compound loss of MafK and MafG activity would result in both an exacerbated megakaryocytic phenotype as well as erythroid deficiencies, whereas loss of both MafF and MafG or MafF and MafK function may result in no new phenotypes. We are testing these predictions at the present time.

The results of completion of the small Maf family loss of function mutant analysis that we report here also suggests

**TABLE II**

| Genotypic distribution of animals recovered from mafF<sup>+/−</sup> intercrosses |
|-----------------------------------------------|
| ES clone/genotype | +/+ | +/− | −/− | Total |
| 2D12 | 34 (24.1%) | 69 (48.9%) | 38 (27.0%) | 141 (100%) |
| 1D11 | 22 (24.7%) | 46 (51.7%) | 21 (23.6%) | 89 (100%) |
| 1C12 | 14 (24.6%) | 31 (54.4%) | 12 (21.0%) | 57 (100%) |
| S01 (−neo) | 25 (20.5%) | 71 (58.2%) | 26 (21.3%) | 122 (100%) |

**TABLE III**

| Hematological analysis of mafF mutant mice |
|------------------------------------------|
| Each value represents the mean ± S.D. of measurement of blood samples from three different animals. WBC, white blood cells; RBC, red blood cells. 500 μl of blood was drawn from 7–16-week-old mice and placed in a tube containing 7 μl of 0.5 M EDTA. The blood samples were analyzed by Analytics, Inc. (Gaithersburg, MD). |

| Parameter/ genotype | +/+ | +/− | −/− | Total |
|---------------------|-----|-----|-----|-------|
| WBC (×10<sup>3</sup>/μl) | 7.8 ± 2.01 | 8.4 ± 1.48 | 7.8 ± 1.80 |
| RBC (×10<sup>6</sup>/μl) | 8.48 ± 0.53 | 8.04 ± 0.76 | 8.85 ± 0.22 |
| Hemoglobin (g/dl) | 15.8 ± 0.43 | 14.4 ± 0.80 | 15.2 ± 0.30 |
| Hematocrit (%) | 43.5 ± 1.54 | 39.6 ± 2.28 | 42.3 ± 0.30 |
| Platelets (×10<sup>3</sup>/μl) | 1105 ± 137.2 | 994 ± 93.3 | 1121 ± 50.1 |

**FIG. 6.** Expression of mafF in homozygous mutant pups. A, MafF is abundantly expressed in the skin of newborn animals, as is MafK (25). MafF is uniquely expressed in bone membrane (B) and is also strongly expressed in the bronchial epithelia (C). Labeling of cells in the bone marrow of mafF mutant pups (D) is infrequent.

**FIG. 7.** Comparison of tissue-specific mRNA accumulation for small Maf family members using quantitative RT-PCR. cDNAs were synthesized from total RNA samples isolated from adult animals using murine leukemia virus reverse transcriptase. Standard curves for MafF, MafG, MafK, and ribosomal RNA were generated by serial dilution of intestine cDNA. The small Maf gene expression levels were normalized for ribosomal RNA expression levels measured in the same cDNA preparation. Values for the tissues that displayed the highest level of mRNA accumulation were arbitrarily set at 100. Values represent the mean ± S.D. of the ratios from three independent experiments.
several general conclusions about the complexity of gene regulation that might be elicited through the MARE element by this growing array of basic/zipper proteins that can interact with one another (3, 9). First, MafG alone may be able to fully sustain embryonic and postnatal development, including hematopoiesis. The expression profile shown in Fig. 5 supports the notion that MafG is probably the protein fulfilling small Maf function in most tissues, and this is also most consistent with both earlier (25, 35) and the present (Fig. 7) studies.

Second, it is somewhat puzzling that even though MafF and MafK appear to be strongly expressed in at least a few tissues in which there is no significant accumulation of MafG (Fig. 5), there is nonetheless no phenotype associated with either individual loss of function mutation. One hypothesis that would resolve this conundrum is by proposing that a fourth small Maf protein exists, and this is also the hypothetical protein that compensates for the mafF or mafK mutations in those tissues where the other two are uniquely expressed. However, numerous previous studies, including chicken and human cDNA library screens as well as screens performed here have failed to recover clones encoding any other than three known small Maf cDNAs.

If it is true that all of the small Maf proteins have been discovered, we must logically draw one of three possible conclusions. The first is that expression of all the small Mafs in unique expression sites is biologically meaningless; this intellectually dissatisfying conclusion is almost certainly wrong but cannot be refuted by reference to the currently available data alone. Second, it is possible that low level expression of other small Maf proteins, below the level of detection in the lacZ staining assay, is sufficient to rescue the loss of function of a more abundant family member. Finally, it is possible that other proteins in this interactive network can partially complement the missing small Maf activity (3). With regard to this final possibility, we might imagine that c-Fos or c-Jun, which normallly comprise transcription factor AP-1, are able to compensate for a loss of function mutation in mafF or mafK, because it has been shown that c-Fos or c-Jun can form heterodimers with c-Maf, Nrf1, and Nrf2 (2, 7, 18, 19). However, heterodimers formed between AP-1 components and the large Mafs or CNC family members do not appear to exert productive regulatory effects when tested in co-transfection assays.

Finally, when members of these various leucine zipper molecules were initially cloned, it was implicitly assumed that the some or all CNC family members might be able to exert functions at MARE sites as homodimers, not requiring the cooperation of small Maf partner proteins. However, all the evidence accumulated to date suggests that homodimers of p45, Nrf1, Nrf2, Nrf3, Bach1, and Bach2 do not efficiently associate with MAREs when compared with CNC/small Maf heterodimers (11, 14, 16).

The complexity of all possible combinations of heterodimer formation between the small Mafs, large Mafs, CNC family, and Fos/Jun family members makes it exceedingly difficult to predict what the in vivo consequences of a mutation in any of these genes might be. This complication is exacerbated when mutations are examined in an outbred genetic background, such as the CD1 mice we used in the studies described here. To address the possible biological consequences that might arise as a result of these mutations in a more controlled environment, it would almost certainly be advantageous to examine these loss of function alleles in a constant genetic background.

To this end, we are breeding each of these small maf mutant alleles into congenic backgrounds of several different inbred strains of mice. These studies may or may not result in the presentation of new phenotypic abnormalities in the small maf mutant animals, but if each of these proteins do fulfill specific roles during development, these functions are much more likely to be revealed in a congenic genetic background.

Acknowledgments—We are grateful to our laboratory colleagues for stimulating discussions. We thank Jie Fan and Weimin Song for technical assistance with some of these experiments and Frank Grosveld (Erasmus University) for the gift of the nls-lacZ plasmid.

REFERENCES

1. Nishizawa, M., Kataoka, K., Goto, N., Fujiwara, K., and Kawai, S. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 7711–7715

2. Kataoka, K., Fujiwara, K. T., Noda, M., and Nishizawa, M. (1994) Mol. Cell. Biol. 14, 7571–7579

3. Motohashi, H., Shavit, J. A., Igarashi, K., Yamamoto, M., and Engel, J. D. (1997) Nucleic Acids Res. 25, 2853–2859

4. Kataoka, K., Nishizawa, M., and Kawai, S. (1993) J. Virol. 67, 2133–2141

5. Swaroop, A., Xu, J., Pawar, H., Jackson, A., Scolini, C., and Agarwal, N. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 266–270

6. Fujiwara, K., Kataoka, T. K., and Nishizawa, M. (1993) Oncogene 8, 2971–2980

7. Kataoka, K., Igarashi, K., Itoh, K., Fujiwara, K. T., Noda, M., Yamamoto, M., and Nishizawa, M. (1995) Mol. Cell. Biol. 15, 2180–2180

8. Nagai, T., Igarashi, K., Akasaka, J., Furuyama, K., Fujita, H., Hayashi, N., Yamamoto, M., and Sassa, S. (1998) J. Biol. Chem. 273, 5583–5586

9. Blank, V., and Andrews, N. C. (1997) Trends Biochem. Sci. 22, 437–440

10. Mohler, J., Vani, K., Leung, S., and Epstein, A. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 10–14

11. Shivdasani, R. A., Rosenblat, M. F., Zucker-Franklin, D., Jackson, C. W., Hunt, P., Saris, C. J., and Orkin, S. H. (1995) Cell 81, 695–701

12. Oyake, T., Itoh, K., Motohashi, H., Hayashi, N., Hoshino, H., Nishizawa, M., Yamamoto, M., and Igarashi, K. (1996) Mol. Cell. Biol. 16, 6083–6095

13. Muto, A., Hoshino, H., Madisen, L., Yanai, N., Obinata, M., Karayama, H., Hayashi, N., Nakauchi, H., Yamamoto, M., Groudine, M., and Igarashi, K. (1998) EMBO J. 17, 5734–5743

14. Kataoka, K., Noda, M., and Nishizawa, M. (1994) Mol. Cell. Biol. 14, 700–712

15. Kataoka, K., Noda, M., and Nishizawa, M. (1996) Oncogene 13, 53–62

16. Farmer, S. C., Sun, C. W., Winnier, G. E., Hogan, B. L., and Townes, T. M. (1997) Genes Dev. 11, 786–798

17. Chan, J. Y., Kwong, M., Lu, R., Chang, J., Wang, B., Yen, T. S., and Kan, Y. W. (1989) EMBO J. 17, 1779–1787

18. Igarashi, K., Lu, R., Chan, J. C., and Kan, Y. W. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 13943–13948

19. Itoh, K., Chiba, T., Takahashi, S., Ishii, T., Igarashi, K., Kato, Y., Oyake, T., Hayashi, N., Satoh, K., Hayata, Y., Yamamoto, M., and Nabeshima, Y. (1997) Biochem. Biophys. Res. Commun. 236, 313–322

20. Kotkov, K. J., and Orkin, S. H. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 3514–3518

21. Shavri, J. A., Motohashi, H., Onodera, K., Akasaka, J., Yamamoto, M., and Engel, J. D. (1998) Genes Dev. 12, 2164–2174

22. Igarashi, K., Itoh, K., Hayashi, N., Nishizawa, M., and Yamamoto, M. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 7445–7449

23. Yagi, T., Nada, S., Watanabe, T., Yamamoto, H., Kohmura, N., Ikawa, Y., and Aizawa, S. (1995) Anal. Biochem. 214, 77–86

24. Nagy, A., Rossant, J., Nagy, R., Abramow, N. W., and Roder, J. C. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 8424–8428

25. Gilson, U. E. M., Heid, C. A., and Williams, P. M. (1996) Genome Res. 6, 995–1001

26. Peters, R. L., and Eicher, E. M. (1994) Genomics 22, 490–491

27. Kim, C. G., Epner, E. M., Forrester, W. C., and Groudine, M. (1992) Genes Dev. 6, 928–938

28. Wood, S. A., Allen, N. D., Rossant, J., Auerbach, R., and Nagy, A. (1993) Nature 365, 87–89

29. Aizawa, S. (1993) Anal. Biochem. 216, 223–238

30. Nakafuku, M., Nishizawa, M., Engel, J. D. (1998) Nature 398, 344–348

31. Motohashi, H., Igarashi, K., Onodera, K., Takahashi, S., Ohtani, H., Nakafuku, M., Nishizawa, M., Engel, J. D., and Yamamoto, M. (1996) Genes Cells 1, 223–238

32. Blank, V., Kim, M. J., and Andrews, N. C. (1997) Blood 89, 3925–3935