This document is the accepted version of a published work that appeared in final form in Am. J. Hum. Genet., after technical editing by the publisher. To access the final edited and published work, see https://doi.org/10.1016/j.ajhg.2015.03.001

Mutations impairing GSK3-mediated MAF phosphorylation cause cataract, deafness, intellectual disability, seizures, and a Down syndrome-like facies

Marcello Niceta,[1,2,22] Emilia Stellacci,[1,22] Karen W. Gripp,[3,23] Giuseppe Zampino,[4,23] Maria Kousi,[5] Massimiliano Anselmi,[6] Alice Traversa,[1,7] Andrea Ciolfi,[1] Deborah Stabley,[8] Alessandro Bruselles,[1] Viviana Caputo,[7] Serena Cecchetti,[9] Sabrina Prudente,[10] Maria T. Fiorenza,[11] Carla Boitani,[12] Nicole Philip,[13] Dmitriy Niyazov,[14] Chiara Leoni,[4] Takaya Nakane,[15] Kim Keppler-Noreuil,[16] Stephen R. Braddock,[17] Gabriele Gillessen-Kaesbach,[18] Antonio Palleschi,[6] Philippe M. Campeau,[19] Brendan H.L. Lee,[20] Celio Pouponnot,[21] Lorenzo Stella,[6] Gianfranco Bocchinfuso,[6] Nicholas Katsanis,[5] Katia Sol-Church,[8] and Marco Tartaglia[1*]

[1]Dipartimento di Ematologia, Oncologia e Medicina Molecolare, Istituto Superiore di Sanità, Rome, 00161 Italy. [2]Polo di Ricerca - Malattie rare, Ospedale Pediatrico Bambino Gesù IRCSS, Rome, 00146 Italy. [3]Division of Medical Genetics, A.I. duPont Hospital for Children, Wilmington, DE 19803, USA. [4]Istituto di Pediatria, Università Cattolica del Sacro Cuore, Rome, 00168 Italy. [5]Center for Human Disease Modeling, Department of Cell Biology, Duke University, Durham, NC 27710, USA. [6]Dipartimento di Scienze e Tecnologie Chimiche, Università “Tor Vergata”, Rome, 00133 Italy. [7]Dipartimento di Medicina Sperimentale, Università “La Sapienza”, 00161 Rome, Italy. [8]Center for Pediatric Research, A.I. duPont Hospital for Children, Wilmington, DE 19803, USA. [9]Dipartimento

di Biologia Cellulare e Neuroscienze, Istituto Superiore di Sanità, Rome, 00161 Italy. [10]Mendel Laboratory, IRCCS Casa Sollievo della Sofferenza, Rome, 00198 Italy. [11]Dipartimento di Psicologia, Sezione di Neuroscienze, Università "La Sapienza", Rome, 00161 Italy. [12]Sezione di Istologia e Embriologia Medica, Dipartimento di Scienze Anatomiche, Istologiche, Medico-legali e dell'Apparato Locomotore, Università "La Sapienza", Rome, 00161 Italy. [13]Département de Génétique Médicale, Hôpital d'Enfants de la Timone, Marseille, 13385 France. [14]Division of Medical Genetics, Ochsner Health System, New Orleans, LA 70121, USA. [15]Department of Pediatrics, Center for Genetic Medicine, University of Yamanashi, Chuo, Yamanashi, 409-3898 Japan. [16]Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute/NIH, Bethesda, MD 20892, USA. [17]Department of Pediatrics, Saint Louis University School of Medicine, St. Louis, MO 63104, USA. [18]Institut für Humangenetik, Universität zu Lübeck, Lübeck, 23538 Germany. [19]Department of Pediatrics, Sainte-Justine Hospital, University of Montreal, Montreal, H3T 1C5 Canada. [20]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA. [21]Institut Curie Centre de Recherche, CNRS UMR 3347, INSERM U1021, Paris Sud University, Orsay, 91405 France. [22]These authors contributed equally to this project. [23]These authors contributed equally to this project.

*Correspondence:

Marco Tartaglia, PhD

Dipartimento di Ematologia, Oncologia e Medicina Molecolare

Istituto Superiore di Sanità

Viale Regina Elena, 299

00161 Rome, Italy

phone: +39 06 49902569

fax: + 39 06 49902850

email: marco.tartaglia@iss.it

Abstract

Transcription factors operate in developmental processes to mediate inductive events and cell competence, and perturbation of their function or regulation can dramatically affect morphogenesis, organogenesis, and growth. We report that a narrow spectrum of amino acid changes within the transactivation domain of the v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog (MAF), a leucine zipper-containing transcription factor of the AP1 superfamily, profoundly affect development. Seven different *de novo* missense mutations involving conserved residues of the four GSK3 phosphorylation motifs were identified in eight unrelated individuals. The distinctive clinical phenotype, for which we propose the eponym Aymé-Gripp syndrome, is not limited to lens and eye defects as previously reported for MAF/Maf loss-of-function, but includes sensorineural deafness, intellectual disability, seizures, brachycephaly, distinctive flat facial appearance, skeletal anomalies, mammary gland hypoplasia, and reduced growth. Disease-causing mutations were demonstrated to impair proper MAF phosphorylation, ubiquitination and proteosomal degradation, perturbed gene expression in primary skin fibroblasts, and induced neurodevelopmental defects in an *in vivo* model. Our findings nosologically and clinically delineate a previously poorly understood recognizable multisystem disorder, provide evidence for MAF governing a wider range of developmental programs than previously appreciated, and describe a novel instance of protein dosage effect severely perturbing development.

Dual sensory impairment due to cataracts and sensorineural hearing loss is a well-recognized consequence of infectious teratogenic exposure (*i.e.*, fetal rubella syndrome), but only rarely observed as a developmental defect in genetic disease phenotypes. In 1996, Gripp and co-workers described two unrelated subjects with congenital cataracts and sensorineural deafness associated with intellectual disability, short stature, brachycephaly, and a distinctive flat facial appearance, and considered this trait to represent a previously unrecognized syndrome (MIM 601088).[1] Aymé and Philip noted similarities to their individual,[2] and others previously reported, all sharing features with that originally described by Fine and Lubinsky,[3] and suggested the term Fine-Lubinsky syndrome (FLS [MIM 601353]) for the phenotype.[4] Since then, a few additional cases exhibiting features fitting or partially overlapping this(these) condition(s) have been reported,[5-9] and whether these phenotypes represent variable manifestations of a single nosologic entity remained unresolved. Autosomal recessive inheritance was suggested, based on affected siblings.[7] Here, whole exome sequencing (WES) on a single affected individual and Sanger sequencing on a selected cohort of subjects with phenotype suggestive of FLS were used to identify a narrow spectrum of missense mutations in v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog (*MAF* [MIM 177075]) as the molecular cause underlying this previously poorly understood multisystem disorder, and delineate its clinical phenotype. We also provide evidence that this class of *MAF* mutations impair GSK3-mediated MAF phosphorylation, ubiquitination and proteosomal degradation, and are dominantly acting in inducing neurodevelopmental defects in a zebrafish model.

Thirteen subjects were included in this study. All subjects were clinically assessed by experienced clinical geneticists. Among them, nine subjects had previously been reported.[1,2,5-8] Clinical features are described in detail in Table S1. Clinical data and biological material collection and storage were attained from the participating families after written informed consent was secured, following procedures in accordance with the ethical standards of the responsible committees on human

experimentation (institutional and national). Genomic DNA was isolated from peripheral blood leukocytes, skin fibroblasts, hair bulb cells, and/or buccal mucosal epithelial cells, using standard protocols. We performed WES on genomic DNA extracted from circulating leukocytes of a single affected subject (case 8) (Figure 1A, CaGi_UCSC). Exome capture was performed using NimbleGen SeqCap EZ Exome V. 3.0 (Roche), and sequencing by a HiSeq2000 instrument (Illumina). WES data analysis was performed using an in-house implemented pipeline.[10] For sequencing statistics, see Table S2. Data annotation predicted 11,168 high-quality variants having functional impact (*i.e.*, non-synonymous and splice site changes). Among them, 259 private, rare (minor allele frequency < 0.001) or clinically associated changes were retained for further analyses. After excluding the presence of variants compatible with autosomal recessive transmission (Table S3), we reasoned that the clinical symptomatology might be caused by a *de novo* event. Candidates were stratified through a mixed filtering/prioritization strategy taking into account the predicted impact of each variant and the functional relevance of individual genes on the developmental processes altered in the disorder. Only changes (private, clinically associated or having unknown frequency or minor allele frequency < 0.001) predicted to be deleterious by the Combined Annotation Dependent Depletion (CADD)[11] (score > 15.0) or Database for Nonsynonymous SNPs' Functional Predictions (dbNSFP) Support Vector Machine (SVM)[12] (radial score > 0.0) algorithm were retained, and prioritized on the basis of the functional relevance of genes using GeneDistiller. Genes were ranked based on combinations of terms from the OMIM clinical synopsis for MIM 601088 and 601353 (*i.e.*, cataract, deafness, mental retardation, facial dysmorphism, short stature, and seizure) as keywords, using similarity of expression patterns and protein-protein interactions as major weights. We obtained the highest score for *MAF*, a gene whose mutations had previously been reported to cause autosomal dominant congenital cataracts and lens abnormalities (MIM 610202).[13,14] Sanger sequencing confirmed heterozygosity for the c.161C>T (p.Ser54Leu) change in our proband, and sequencing of parental DNAs revealed only the reference

allele, evidence for its *de novo* origin (Figure S1). STR genotyping (AmpFlSTR Identifiler Plus [Life Technologies]) confirmed paternity. The variant was documented in the proband's skin fibroblasts as well as hair bulb and buccal epithelial cell specimens, strongly arguing against the possibility of a somatic event (Figure S1). All the other candidate variants turned out to be inherited from one of the unaffected parents (Table S4).

To confirm the causal involvement of *MAF*, the entire coding sequence of the gene (NM_005360.4 and NM_001031804.2) was scanned for mutations in DNA samples from twelve additional subjects (Table S1), including a sib pair, with features overlapping the conditions delineated by Gripp et al.,[1] and Aymé and Philip,[2] by direct sequencing. Primer pairs designed to amplify the *MAF* coding exons and their intron boundaries (NC_000016.10, 79593848..79600725) are listed in Table S5. We identified heterozygous missense mutations in seven unrelated individuals (Table 1 and Figure S1). Two different changes affected Thr58 and Pro59, while others involved adjacent residues, including the previously identified c.161C>T substitution. In all family trios for which parental DNA samples were available, genotyping documented the *de novo* origin of each mutation, and STR analysis confirmed paternity (Table 1). Sanger sequencing of DNA from available oral mucosal epithelial cells (case 11-1), or epithelial cells and fibroblasts (case 4-1) supported the germline origin of mutations. All changes were predicted to impact protein function by dbNSFP and/or CADD (Table 1), and affected residues conserved among orthologs and paralogs (Figure S2).

MAF is a basic leucine zipper (bZIP)-containing transcription factor of the AP1 superfamily.[15,16] It is important for lens and eye development,[17,18] and controls multiple physiological processes, including mechanosensory function, and chondrocyte and T-cell differentiation.[19-21] Immunohistochemical analyses in mouse embryos documented wide *Maf* expression (Figure S3). Consistent with previous reports, we observed Maf staining in the lens, dorsal spinal cord, dorsal root ganglia, skin, kidney, hypertrophic chondrocytes of vertebrae, rib and limb cartilage, and the cartilage

primordium of the basioccipital bone.[22,23] In line with the sensorineural hearing loss occurring in all mutation-positive subjects, we detected a specific and strong signal in cochlear cells of E14.5 embryos.

Similar to other "large" MAF subfamily members (*i.e.*, MAFA [MIM 610303], MAFB [MIM 608968] and NRL [MIM 162080]), MAF's structure is characterized by a *C*-terminal extended homology region and bZIP domain mediating DNA binding, and a *N*-terminal transactivation domain required for transcriptional activity and regulatory function (Figure 1B). The latter contains four glycogen synthase kinase 3 (GSK3) phosphorylation motifs, highly conserved among large MAF proteins (Figure 1C). In MAFA, the sequential phosphorylation of these serine/threonine residues promotes ubiquitination and rapid degradation, but also increases transactivation potential.[24,25] Remarkably, all identified *MAF* mutations clustered within this motif. Three affected residues, Ser54, Thr58 and Ser62, are known GSK3 phosphorylation target sites.[26] The remainder did not involve phosphorylatable residues, but were predicted to affect GSK3-mediated phosphorylation by altering proline residues adjacent to either a phosphorylation site (Thr58) or the *C*-terminal priming site (Ser70), whose phosphorylation is absolutely required for GSK3 function. To explore the impact of the Pro59His, Pro59Leu and Pro69Arg changes, we performed molecular dynamics (MD) simulations on complexes formed by full-length GSK3 and 10 residue-long peptides of MAF corresponding to the segment that interacts directly with the GSK3 binding cleft, encompassing both the GSK3 target and pSer/pThr primed residues (Table S6). The starting coordinates for the ATP-bound GSK3 were taken from the crystallographic structure of GSK3B complexed with AMP-PNP (PDB entry 1pyx).[27] Each decapeptide was set in an extended conformation along the catalytic cleft of GSK3 as specified in Table S5. The MD simulations were carried out according to the protocol previously described.[10] The 53a6 force field was used with the exception of the partial charges of pSer/pThr,[28,29] and the parameters for ATP, obtained from quantum mechanical calculation of the molecular system reported in Figure S6. For Pro59His and Pro59Leu decamers, the conformation of the trimer comprised between the substrate

and primed residues was rearranged considerably (Figure 2A and Figure S4), with the correct orientation of the substrate residue in the GSK3 active site being destabilized (Figure 2A). Pro69Arg, introducing a cationic residue in the proximity of the positively charged GSK3 priming pocket formed by residues Arg96, Arg180, and Lys205, caused a general rearrangement of the adjacent pSer70, pulling it away from the binding pocket (Figure 2B). Overall, our simulations indicated consistently that all disease-causing *MAF* mutations inhibit GSK3-mediated phosphorylation through impaired association and/or catalysis, by perturbing the interaction with the priming site (Pro69Arg) or the active site (substitutions affecting Pro59).

To explore the mutations' functional impact directly, we evaluated MAF phosphorylation status. The disease-causing p.Ser54Leu, p.Thr58Ala, p.Thr58Ile, p.Pro59Leu, p.Pro59His and p.Pro69Arg (Aymé-Gripp syndrome), and p.Arg288Pro changes, the latter considered as representative of lesions associated with isolated cataract,[13] were introduced into the *MAF* cDNA cloned in pCS2+ vector using the QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies). Consistent with previous reports,[26] Western blot analysis of transiently transfected COS1 cell lysates documented two MAF states: a slower-migrating, fully phosphorylated form, and a faster-migrating, unphosphorylated form. In cells expressing wild-type *MAF*, the phosphorylated protein predominated while unphosphorylated MAF was barely detectable (Figure 3A, upper panel). Similarly, the mutant carrying the Arg288Pro substitution in the DNA binding domain, previously associated with isolated lens and eye defects, was efficiently phosphorylated. This was in sharp contrast to all MAF mutants identified in the present study, which accumulated in cells as unphosphorylated proteins. GSK3-mediated phosphorylation represents a regulatory mechanism promoting MAFA ubiquitination and degradation.[24,25] Based on the high conservation of the GSK3 recognition motif and MAF being a GSK3 substrate, we hypothesized that the amino acid changes in our affected subjects might mediate inefficient protein clearance. On Western blot analyses, we noted increased protein levels (Figure 3A, upper panel) and decreased

ubiquitination (Figure 3A, middle panel) for the disease-causing MAF mutants when compared to the wild-type protein (Figure 3A and 3B). CHX treatment (Figure 3B) showed that the increased protein levels of MAF mutants was due to their enhanced stability, which were not efficiently degraded *via* proteasome, as indicated by their independence on treatment with MG132, which specifically inhibit proteosomal function (Figure 3B). Of note, a partial phosphorylation was apparent for the Pro69Arg MAF mutant, which was consistently associated with increased degradation *via* proteasome, even though less efficient compared to wild-type MAF. This finding suggests a milder perturbing role of the proline-to-arginine substitution on GSK3-mediated phosphorylation at Ser58 compared to the other disease-causing amino acid changes, possibly due to the specific effect of the introduced arginine residue on MAF interaction with the GSK priming site. Confocal microscopy of transfected COS1 cells confirmed the nuclear localization of all tested mutants, and their higher abundance within cells (Figure 3C). Moreover, treatment with CSK buffer prior fixation indicated that the syndrome-associated mutants retained efficient interaction with chromatin suggesting that they bind to DNA, in contrast to the DNA binding-impaired cataract-associated Arg288Pro mutant (Figure 3C and Table S6). Transactivation assays using luciferase as reporter under control of the IL4 promoter documented that COS1 cells transiently expressing the cataract-causing mutant allele had barely detectable reporter induction (Figure 3D). In contrast, cells expressing the Ser54Leu, Thr58Ala, Thr58Ile, Pro59Leu, Pro59His or Pro69Arg MAF coding alleles showed efficient induction of luciferase levels, though not reaching the levels of the wild-type protein, suggesting that, despite their stabilization, these mutants are less active, at least under these specific conditions.

Next, we conducted gene expression profiling analyses on primary skin fibroblasts from two unrelated subjects (4-1 and CaGi_UCSC) to explore more globally the impact of dysregulated MAF function on gene expression (Table S7 for details). Approximately 6% of genes in mutation-positive subject-derived cells were expressed differentially compared to control fibroblasts (ATCC code PCS-

201-012). Gene Ontology enrichment analysis of these differentially expressed genes revealed an overrepresentation of genes coding for proteins associated with developmental programs (tissue morphogenesis, branching morphogenesis, urogenital system and bone development) and cellular processes (cytoskeletal rearrangement, cell migration and adhesion, and response to extracellular stimuli) (Table S7). Publicly available microarray data (ArrayExpress accession code E-GEOD-51231) allowed selection of genes whose expression is positively controlled by MAF. Remarkably, these putative targets were observed to be enriched significantly in up- or down-regulated transcripts in *MAF* mutation-positive fibroblasts (Figure S5), suggesting a complex, promoter-specific dominant dysregulatory function of these mutations.

Since cognitive deficits, with or without brain defects, are invariably present among subjects with *MAF* mutations affecting the GSK3 recognition motif, but absent in individuals with congenital cataracts caused by *MAF* mutations involving the DNA binding domain,[13,30-33] we posited that expression of the mutant alleles here identified should induce defects in neurogenesis, whereas alleles driving dominant cataracts through a haploinsufficiency model should not. To test this hypothesis *in vivo*, we analyzed the impact of these mutation classes on the integrity of the central nervous system (CNS) using a zebrafish model. The optic tectum is a major component of the vertebrate midbrain, comprising a structure equivalent to the mammalian superior colliculus, with external layers collecting sensory information and internal layers having a motor-function. We previously showed that measurement of the area of this structure represents a robust surrogate for brain volume, and that reduction of the absolute area of the tectum correlates with neurodevelopmental defects in humans.[34,35] We injected clutches of 50-100 embryos with human mRNA encoding wild-type *MAF*, each of four alleles carrying mutations discovered in the present study, or the cataract-associated allele. Experimental work was carried out under protocols approved by the Institutional Animal Care and Use Committee of the Duke University, as previously described.[34-36] Following quantitative measurement

of the surface area of the tecta (blind to injection cocktail; replicated), we observed that expression of either wild-type and cataract-associated Arg288Pro *MAF* mRNAs did not induce appreciable brain volume differences. In contrast, injection of human *MAF* mRNA encoding each of the Ser54Leu, Thr58Ala, Pro59Leu or Pro69Arg changes caused a statistically significant reduction of the size of the optic tecta ($p < 0.0001$; Figure 4). These data were concordant with our *in vitro* model and provided further evidence for the differential impact of these two mutation classes on CNS development.

MAF mutation-positive individuals shared congenital or early onset cataracts, sensorineural hearing loss, developmental delay/intellectual disability, seizures, brachycephaly, midfacial hypoplasia, and reduced growth (final height -2.25 to -4 SD) as major characteristics (Figure 1A and Table S1). Hearing loss was diagnosed in early childhood, bilateral and had at least a moderate to severe sensorineural component in all, in some cases requiring hearing aids. Limited information on inner ear imaging is available. Facial features were distinctive and recognizable, and included a short nasal tip and long philtrum, a small mouth, and small/low-set posteriorly angulated ears. Late closing anterior fontanel, radio-ulnar synostosis or limited elbow movement occurred in several subjects, as did pericardial effusion and idiopathic chondrolysis of the hip. The three adult females had mammary gland hypoplasia, and two lacked axillary hair. Two individuals showed hypopigmented retinal lesions. Mesangiocapillary glomerulopathy occurred in the oldest affected individual. While intellectual function was not assessed systematically, impairment was universally present. Adaptive function in adult individuals ranged from severely autistic, nonverbal and requiring medication for behavior problems to working in a supervised community environment and living with guardians. Seizures and abnormal EEG findings consistent with focal as well diffuse abnormal activity were present in all individuals, most often diagnosed in early childhood. Brain imaging studies most often showed prominence of the axial and extra-axial fluid spaces, other anomalies like a Chiari 1 malformation or J-shaped sella occurred in single cases. The absence of mutations among subjects with

only partly overlapping features (*e.g.*, absence of brachycephaly and cataracts in sibs 10 and 11, and facies not suggestive in case 13) implies that mutations affecting residues within the GSK3 recognition motif of MAF cause a clinically homogeneous phenotype, and that cataracts, sensorineural hearing loss, intellectual disability, seizures, brachycephaly, Down syndrome-like facial appearance, and reduced growth are likely to represent cardinal features of the disorder.

Transcription factors operate in developmental processes to mediate inductive events and cell competence, and perturbation of their function or regulation can dramatically affect morphogenesis, organogenesis, and growth.[34] Here we establish that missense mutations impairing GSK3-mediated MAF phosphorylation severely perturb multiple developmental processes. Inactivating missense mutations affecting the MAF DNA binding domain had been reported to cause congenital cataracts, microcornea and iris coloboma.[13,30-33] Contrasting those mutations, we describe a class of missense changes not impairing DNA binding, but instead precluding post-translational modifications required for proteasomal degradation, resulting in mutants being more stable than wild-type MAF. In addition to defective degradation, this *MAF* mutation class showed perturbed *in vitro* transactivation activity, and endogenous expression in primary fibroblasts was associated with both upregulation and downmodulation of genes identified as positively controlled MAF targets. These findings support the idea that, besides promoting degradation, GSK3-mediated MAF phosphorylation impacts protein activity through other mechanisms, as previously demonstrated for MAFA,[24] and suggest a complex pathogenetic mechanism involving protein stability and functional dysregulation. Of note, dominant missense mutations affecting residues in the same regulatory motif of MAFB and NRL, including those homologous to Ser54, Pro59, Thr62 and Pro69, have been reported to cause multicentric carpotarsal osteolysis (MCTO [MIM 166300]) and autosomal dominant retinitis pigmentosa (RP27 [MIM 613750]) (Figure S2), respectively,[35,36] further highlighting the critical role of this domain.

Early onset cataracts and hearing loss have rarely been reported in genetic conditions, with the collagenopathies, Alport (MIM 104200), Stickler (MIM 108300) and Marshall (MIM 154780) syndromes being the most common. This work identified *MAF* mutations as principal cause for a disorder combining cataracts and hearing loss in a multisystem developmental syndrome independently recognized by Gripp et al.,[1] and Aymé and Philip.[2] Though the clinical presentation of subjects within this study resembles that described in the context of FLS, we note that this overlap is only partial. Moreover, FLS has been applied to clinically variable phenotypes of likely heterogeneous etiology (here exemplified by cases 10, 11, and 12). We therefore propose the eponym Aymé-Gripp syndrome for this disorder caused by mutations affecting residues of the GSK3 phosphorylation motifs in order to distinguish the phenotype described here from that reported by Fine and Lubinsky.[3]

By regulating the spatio-temporal expression of tissue-specific genes, MAF proteins act as key regulators of terminal differentiation in many tissues and organs, including bone, brain, kidney, lens, pancreas, and retina.[37,38] While apparently no gross anomalies are associated with Maf haploinsufficiency in mice, *Maf*$^{+/-}$ pups die soon after birth and exhibit defective lens formation and eye development,[17,18] chondrocyte terminal differentiation,[21] as well as differentiation and function of mechanoreceptors and neurons with mechanosensory function.[20,39] In contrast, a semi-dominant, missense mutation (Arg291Gln) affecting the DNA-binding domain of Maf and causing a reduced transactivation activity of the transcription factor has been associated with congenital cataract in heterozygote mice,[14] recapitulating the hypomorphic cataracts-associated *MAF* mutations in humans.[13,30-33] Finally, a functionally distinct missense change (Asp90Val) affecting the *N*-terminal transactivation domain and promoting enhanced transactivation function in Maf has been shown to cause a dominant isolated cataract phenotype.[40] Contrary to these loss-of-function, gain-of-function and haploinsufficiency models, we here showed a distinct, dominantly acting effect of *MAF* mutations underlying a complex developmental disorder affecting multiple organs and tissues. As such, the

pleiotropic effect of impaired MAF phosphorylation in Aymé-Gripp syndrome expands the perturbing consequences of dysregulated MAF function for multiple developmental programs, establishing its role in morphogenesis, CNS development, hearing and growth, and delineates a novel instance of protein dosage effect in human disease.

Supplemental Data

Supplemental data include eight tables, six figures, and supplemental references.

Acknowledgements

We are grateful to the individuals and their families who contributed to this study. We thank Serenella Venanzi and Francesca Maiorca (ISS, Rome, Italy) for skilful technical assistance. We acknowledge submission of several DNA samples to the Center for Mendelian Genomics (CMG) in Seattle, WA; disease gene identification occurred prior to and without the use of data from the CMG. We thank BGI (Hong Kong) for the high-quality sequencing raw data. M.T., G.B. and L.S. acknowledge CINECA for computational resources (WES data and structural analyses). NK is a Distinguished George W. Brumley Professor. The authors declare no financial conflict of interest. This work was supported in part by grants from Telethon (GGP13107 to M.T.), Istituto Superiore di Sanità (Ricerca Corrente 2013 to M.T.), Nemours Foundation (to K.S.C.) and NIH (P20GM103464 and P20GM103446 to K.S.C.; P50MH094268 to NK). This work is dedicated to the memory of Luciano Cianetti (ISS, Rome, Italy).

Web Resources

Accession numbers and the URL for data are as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/; Gene, http://www.ncbi.nlm.nih.gov/gene/; GeneDistiller, http://www.genedistiller.org/; Combined Annotation

Dependent Depletion (CADD), http://cadd.gs.washington.edu/; Database for Nonsynonymous SNPs' Functional Predictions (dbNSFP), https://sites.google.com/site/jpopgen/dbNSFP; RCSB Protein Data Bank, http://www.rcsb.org/pdb/home/home.do; ArrayExpress, https://www.ebi.ac.uk/arrayexpress/.

Accession Numbers

The ArrayExpress accession number for the microarray data generated in this paper is E-MTAB-3250.

References

1. Gripp, K.W., Nicholson, L., and Scott, C.I. (1996) Apparently new syndrome of congenital cataracts, sensorineural deafness, Down syndrome-like facial appearance, short stature, and mental retardation. Am. J. Med. Genet. *61*, 382-386.

2 Aymé, S., and Philip, N. (1996) Fine-Lubinsky syndrome: a fourth patient with brachycephaly, deafness, cataract, microstomia and mental retardation. Clin. Dysmorphol. *5*, 55-60.

3 Fine, B.A., and Lubinsky, M. (1983) Craniofacial and CNS anomalies with body asymmetry, severe retardation, and other malformations. J. Clin. Dysmorphol. *1*, 6-9.

4 Aymé, S., and Philip**,** N. (1997) Apparently new syndrome of congenital cataracts, sensorineural deafness, Down syndrome-like facial appearance, short stature, and mental retardation. Am. J. Med. Genet. *70*, 333-335.

5 Keppler-Noreuil, K., Welch, J., and Baker-Lange, K. (2007) Syndrome of congenital cataracts, sensorineural deafness, Down syndrome-like facial appearance, short stature, and mental retardation: two additional cases. Am**.** J**.** Med**.** Genet**.** A *143A*, 2581-2587.

6 Nakane, T., Mizobe, N., Hayashibe, H., and Nakazawa, S. (2002) A variant of Fine-Lubinsky syndrome: a Japanese boy with profound deafness, cataracts, mental retardation, and brachycephaly without craniosynostosis. Clin. Dysmorphol. *11*, 195-198.

7 Holder, A.M., Graham, B.H., Lee, B., and Scott, D.A. (2007) Fine-Lubinsky syndrome: sibling pair suggests possible autosomal recessive inheritance. Am. J. Med. Genet. A *143A*, 2576-2580.

8 Schoner, K., Bald, R., Fritz, B., and Rehder, H. (2008) Fetal manifestation of the Fine-Lubinsky syndrome. Brachycephaly, deafness, cataract, microstomia and mental retardation syndrome complicated by Pierre-Robin anomaly and polyhydramnios. Fetal Diagn. Ther. *23*, 228-232.

9 Corona-Rivera, J.R., Lopez-Marure, E., Garcia-Cruz, D., Romo-Huerta, C.O., Rea-Rosas, A., Orozco-Alatorre, L.G., and Ramirez-Valdivia, J.M. (2009) Further clinical delineation of Fine-Lubinsky syndrome. Am. J. Med. Genet. A *149A*, 1070-1075.

10 Cordeddu, V., Redeker, B., Stellacci, E., Jongejan, A., Fragale, A., Bradley, T.E., Anselmi, M., Ciolfi, A., Cecchetti, S., Muto, V., et al. (2014) Mutations in ZBTB20 cause Primrose syndrome. Nat. Genet. *46*, 815-817.

11 Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014) A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. *46*, 310-315.

12 Liu, X., Jian, X., and Boerwinkle, E. (2013) dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. Hum. Mutat. *34*, e2393-2402.

13 Jamieson, R.V., Perveen, R., Kerr, B., Carette, M., Yardley, J., Heon, E., Wirth, M.G., van Heyningen, V., Donnai, D., Munier, F., et al. (2002) Domain disruption and mutation of the bZIP transcription factor, MAF, associated with cataract, ocular anterior segment dysgenesis and coloboma. Hum. Mol. Genet. *11*, 33-42.

14 Lyon, M.F., Jamieson, R.V., Perveen, R., Glenister, P.H., Griffiths, R., Boyd ,Y., Glimcher, L.H., Favor, J., Munier, F.L., and Black, G.C. (2003) A dominant mutation within the DNA-binding domain of the bZIP transcription factor Maf causes murine cataract and results in selective alteration in DNA binding. Hum. Mol. Genet. *12*, 585-94.

15 Blank, V., and Andrews, N.C. (1997) The Maf transcription factors: regulators of differentiation. Trends Biochem. Sci. *22*, 437-441.

16 Eychene, A., Rocques, N., and Pouponnot, C. (2008) A new *MAF*ia in cancer. Nat. Rev. Cancer *8*, 683-693.

17 Kim, J.I., Li, T., Ho, I.-C., Grusby, M.J., and Glimcher, L.H. (1999) Requirement for the c-Maf transcription factor in crystallin gene regulation and lens development. Proc. Natl. Acad. Sci. USA *96*, 3781-3785.

18 Ring, B.Z., Cordes, S.P., Overbeek, P.A., and Barsh, G.S. (2000) Regulation of mouse lens fiber cell development and differentiation by the *Maf* gene. Development *127*, 307-317.

19 Ho, I.C., Lo, D., and Glimcher, L.H. (1998) c-maf promotes T helper cell type 2 (Th2) and attenuates Th1 differentiation by both interleukin 4-dependent and -independent mechanisms. J. Exp. Med. *188*, 1859-1866.

20 Wende, H., Lechner, S.G., Cheret, C., Bourane, S., Kolanczyk, M.E., Pattyn, A., Reuter, K., Munier, F.L., Carroll, P. , Lewin, G.R., et al. (2012) The transcription factor c-Maf controls touch receptor development and function. Science *335*, 1373-1376.

21 MacLean, H.E., Kim, J.I., Glimcher, M.J., Wang, J., Kronenberg, H.M., and Glimcher, L.H. (2003) Absence of transcription factor c-maf causes abnormal terminal differentiation of hypertrophic chondrocytes during endochondral bone development. Dev. Biol. *262*, 51-63.

22 Sakai, M., Imaki,J., Yoshida, K., Ogata, A., Matsushima-Hibiya, Y., Kuboki, Y., Nishizawa, M., and Nishi, S. (1997) Rat maf related genes: specific expression in chondrocytes, lens and spinal cord. Oncogene *14*, 745-750.

23 Ogata, A., Shimizu, T., Abe, R., Shimizu, H., and Sakai, M. (2004) Expression of c-maf and mafB genes in the skin during rat embryonic development. Acta Histochem. *106*, 65-67.

24 Rocques, N., Abou Zeid, N., Sii-Felice, K., Lecoin, L., Felder-Schmittbuhl, M.P., Eychène, A., and Pouponnot, C. (2007) GSK-3-mediated phosphorylation enhances Maf-transforming activity. Molecular Cell *28*, 584-597.

25 Han, S.-i., Aramata, S., Yasuda, K., and Kataoka, K. (2007) MafA stability in pancreatic beta cells is regulated by glucose and is dependent on its constitutive phosphorylation at multiple sites by glycogen synthase kinase 3. Mol. Cell. Biol. *27*, 6593-6605.

26 Herath, N.I., Rocques, N., Garancher, A., Eychène, A., and Pouponnot, C. (2014) GSK3-mediated MAF phosphorylation in multiple myeloma as a potential therapeutic target. Blood Cancer J *4*, e175.

27 Bertrand, J.A., Thieffine, S., Vulpetti, A., Cristiani, C., Valsasina, B., Knapp, S., Kalisz, H.M., and Flocco, M. (2003) Structural characterization of the GSK-3b active site using selective and non-selective ATP-mimetic inhibitors. J. Mol. Biol. *333*, 393-407.

28 Oostenbrink, C., Villa, A., Mark, A.E., and van Gunsteren, W.F. (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J. Comput. Chem. *25*, 1656-1676.

29 Hansson, T., Nordlund, P., and Aqvist, J. (1997) Energetics of nucleophile activation in a protein tyrosine phosphatase. J. Mol. Biol. *265*, 118-127.

30 Vanita, V., Singh, D., Robinson, P.N., Sperling, K., and Singh, J.R. (2006) A novel mutation in the DNA-binding domain of MAF at 16q23.1 associated with autosomal dominant "cerulean cataract" in an Indian family. Am. J. Med. Genet. A *140A*, 558-566.

31 Hansen, L., Eiberg, H., and Rosenberg, T. (2007) Novel MAF mutation in a family with congenital cataract-microcornea syndrome. Mol. Vis. *13*, 2019-2022.

32 Hansen, L., Mikkelsen, A., Nürnberg, P., Nürnberg, G., Anjum, I., Eiberg, H., and Rosenberg, T. (2009) Comprehensive mutational screening in a cohort of Danish families with hereditary congenital cataract. Invest. Ophthalmol. Vis. Sci. *50*, 3291-3303.

33 Narumi, Y., Nishina, S., Tokimitsu, M., Aoki, Y., Kosaki, R., Wakui, K., Azuma, N., Murata, T., Takada, F., Fukushima, Y., et al. (2014) Identification of a novel missense mutation of MAF in a Japanese family with congenital cataract by whole exome sequencing: a clinical report and review of literature. Am. J. Med. Genet. A *164A*, 1272-1276.

34 Schulte, E.C., Kousi, M., Tan, P.L., Tilch, E., Knauf, F., Lichtner, P., Trenkwalder, C., Högl, B., Frauscher, B., Berger, K., et al. (2014) Targeted resequencing and systematic in vivo functional testing identifies rare variants in MEIS1 as significant contributors to restless legs syndrome. Am. J. Hum. Genet. *95*, 85-95.

35 Borck, G., Hög, F., Dentici, M.L., Tan, P.L., Sowada, N., Medeira, A., Gueneau, L., Thiele, H., Kousi, M., Lepri, F., et al. (2015) BRF1 mutations alter RNA polymerase III-dependent transcription and cause neurodevelopmental anomalies. Genome Res. *25*, 155-166.

36 Margolin, D.H., Kousi, M., Chan, Y.M., Lim, E.T., Schmahmann, J.D., Hadjivassiliou, M., Hall, J.E., Adam, I., Dwyer, A., Plummer, L., et al. (2013). Ataxia, dementia, and hypogonadotropism caused by disordered ubiquitination. N. Engl. J. Med. *368*, 1992–2003.

34 Gilbert, S.F. (2008) Principles of differentiation and morphogenesis. In Inborn errors of development. The molecular basis of clinical disorders of morphogenesis. Second edition, C.J.

Epstein, R.P. Erickson, A. Wynshaw-Boris, eds. (Oxford, New York: Oxford University Press), pp. 9-24.

35 Zankl, A., Duncan, E.L., Leo, P.J., Clark, G.R., Glazov, E.A., Addor, M.C., Herlin, T., Kim, C.A., Leheup, B.P., McGill, J., et al. (2012) Multicentric carpotarsal osteolysis is caused by mutations clustering in the amino-terminal transcriptional activation domain of MAFB. Am. J. Hum. Genet. *90*, 494-501.

36 Bessant, D.A.R., Payne, A.M., Mitton, K.P., Wang, Q.L., Swain, P.K., Plant, C., Bird, A.C., Zack, D.J., Swaroop, A., and Bhattacharya, S.S. (1999) A mutation in NRL is associated with autosomal dominant retinitis pigmentosa. Nat. Genet. *21*, 355-356.

37 Kataoka, K. (2007) Multiple mechanisms and functions of maf transcription factors in the regulation of tissue-specific genes. J. Biochem. *141*, 775–781.

38 Yang, Y., and Cvekl, A. (2007) Large Maf transcription factors: cousins of AP-1 proteins and important regulators of cellular differentiation. Einstein J. Biol. Med. *23*, 2-11.

39 Hu, J., Huang, T., Li, T., Guo, Z., and Cheng, L. (2012) c-Maf is required for the development of dorsal horn laminae III/IV neurons and mechanoreceptive DRG axon projections. J. Neurosci. *32*, 5362–5373.40 Perveen, R., Favor, J., Jamieson, R.V., Ray, D.W., and Black, G.C. (2007) A heterozygous c-Maf transactivation domain mutation causes congenital cataract and enhances target gene activation. Hum. Mol. Genet. *16*, 1030-1038.

Figure Titles and Legends

Figure 1. *De novo* heterozygous missense mutations within the transactivation domain of MAF cause Aymé-Gripp syndrome. (A) Clinical features of affected subjects. Note the distinctive flat face, brachycephaly, ptosis, short nasal tip, long philtrum, small mouth, low-set and posteriorly angulated ears, nail dystrophy, and mammary gland hypoplasia. Permission to publish photographs was

provided for all subjects shown. (B) Scheme of the MAF domain structure, and location of *MAF* mutations causing human disease. MAF contains an *N*-terminal transactivation domain (yellow) with regulatory function, and a *C*-terminal DNA binding domain, the latter containing an "extended homology" (green), "basic motif" (light blue), and leucine-zipper (pink) regions. The region containing the four in tandem arranged phosphorylation sites recognized by GSK3 (orange) is located within the transactivation domain. Residues mutated in subjects with Aymé-Gripp syndrome (red) and previously reported isolated cataracts/eye defects (black) are shown. (C) Cartoon illustrating the GSK3 recognition motifs, and location of residues affected in Aymé-Gripp syndrome. The GSK3 catalytic domain is depicted with its active site (red) and the site binding to the priming phosphorylated residue (green). To phosphorylate its substrates, GSK3 requires a priming phosphorylation on the substrate four amino acids downstream the residue to be phosphorylated. The serine/threonine residues sequentially targeted by GSK3 are shown (red). Upon phosphorylation, they act as priming residues (green) to allow the subsequent phosphorylation of the upstream Ser/Thr. The kinase phosphorylating Ser70 has not been characterized yet. The residues affected by Aymé-Gripp syndrome-causing mutations (Ser54, Thr58, Pro59, Ser62 and Pro69) are indicated in bold and empty squares.

Figure 2. Molecular dynamics (MD) simulations of the GSK3/MAF decapeptide complexes. (A) Structural effects of the Pro59Leu and Pro59His changes. In both mutants, the conformation of the trimer comprised between the target and primed residues is considerably rearranged. Representative conformations are reported for wild-type MAF (left) and the Pro59Leu mutant (middle). In both mutants, larger and more variable distances are observed between the hydroxyl of Thr58, which is a GSK3 target residue, and the γ-phosphate of ATP (right, top plot) or the carboxyl group of the catalytic residue Asp181 (right, bottom plot). The distribution obtained in the simulations of the wild-type MAF sequence (black) and those referred to the peptides containing the Pro59His (red) and Pro59Leu (blue)

substitutions are shown. (B) Effect of the Pro69Arg change. In the simulations, a stable interaction between pSer70 of the wild-type peptide and the priming site was observed (left), while a displacement of that residue from the site was documented for the peptide carrying the Pro69Arg change (middle). Such structural rearrangements are quantified by the distance occurring between the P atom of pSer70 and the ω-carbon atom in the side chain of the GSK3 priming site residue, Arg180 (wild-type peptide, black; Pro69Arg peptide, green) (right). In the left and middle panels, the surface of GSK is colored in brown, except for the catalytic residue Asp181 (red), and the priming site residues, Arg96, Arg180, and Lys205 (blue). ATP is shown in pink and the MAF backbone in yellow. The side chains of priming, target and mutated MAF residues are shown in sticks representation.

Figure 3. Impact of disease causing mutations on MAF function. (A) Protein levels, ubiquitination, and phosphorylation levels of wild-type and disease-causing mutant MAF proteins in transiently transfected COS1 cells (upper panel). COS1 cells were maintained in high glucose DMEM, plus 10% FBS and supplements, and were transiently transfected to express wild-type *MAF* or each of the disease-causing alleles (FuGENE 6 [Promega]). To assess ubiquitination, immunoprecipitated MAF was probed with an anti-ubiquitin antibody (#8017 [Santa Cruz Biotechnology]) (middle panel). Whole cell extracts were blotted with anti-MAF polyclonal (#7866 [Santa Cruz Biotechnology]), and anti-β-actin monoclonal (#A5441 [Sigma-Aldrich]) antibodies. Western blots are from a representative experiment of three performed. (B) Protein stability and proteasome-dependent degradation were assessed in COS1 cells transfected with the indicated constructs. Twenty-four hours *post* transfection, cells were treated with 20μg/ml cycloheximide (CHX) or 20μM MG132 for the indicated times. MAF protein levels were detected by immunoblotting with anti-MAF antibody. Western blots of a representative experiment of three performed are shown. (C) Confocal laser scanning microscopy analysis performed in COS1 cells transiently expressing wild-type *MAF* or one of three disease-causing

alleles, without (upper panels) or with (lower panels) treatment with CSK buffer prior fixation. Cells were stained with anti-MAF polyclonal antibody and Alexa Fluor-488 goat anti-rabbit secondary antibody (green). Nuclei are DAPI stained (blue). Images are representative of 450 analyzed cells (Table S5). Experiments were performed as previously reported.[10] (D) Transactivation assays were performed in COS1 cells transiently cotransfected with the *IL4* promoter cloned into pGL3 vector reporter construct (kindly provided by Michael Lohoff, University of Marburg, Marburg, Germany) alone (black bar) or together with wild-type MAF (white bar) or each of the disease-causing MAF mutants (blue and red bars) (1:1 ratio), and 1:10 of *Renilla* luciferase control vector DNA (pRL-Act *Renilla*). After transfection (24 h), firefly and *Renilla* luciferase activities were measured by the Dual Luciferase Reporter Assay System (Promega). Normalized luciferase activity (mean ± S.D.) of six experiments performed is reported as fold increase relative to cells not expressing exogenous *MAF*. *P*-values were calculated using two-tail Student's t-test. *, **, *** indicate $P < 0.05$, $P < 0.01$ and $P < 0.001$, respectively.

Figure 4. *In vivo* impact of Aymé-Gripp syndrome- and isolated cataract-causing *MAF* mutations on the integrity of the central nervous system using a zebrafish model. (A) Dorsal views of uninjected zebrafish embryos (left), and embryos injected with the Aymé-Gripp syndrome-causing mutant (p.Ser54Leu) (middle) and wild-type (right) *MAF* capped mRNA (100 pg) at 3 days post fertilization (dpf). Embryos were whole-mount stained using a primary antibody against acetylated tubulin (1:1000, T7451 [Sigma-Aldrich]) that marks neuronal axons, and an Alexa Fluor goat anti-mouse IgG secondary antibody (1:1000, A21207 [Invitrogen]). The circle highlights the area of the optic tectum that was measured. (B) Overexpression of wild-type *MAF* or the congenital cataracts-causing Arg288Pro allele do not induce a significant reduction in the size of the optic tectum. By contrast, overexpression of each of the Aymé-Gripp syndrome-causing alleles results in a statistically

significantly reduction of the size of the optic tectum ($p < 0.0001$). Bars indicate standard errors, and AU denotes arbitrary units. Statistical analysis was performed using two-tail Student's *t*-test. For the measurements performed we scored 86 control embryos, 61 embryos injected with wild-type *MAF* mRNA, and 58-70 embryos with each of the Aymé-Gripp syndrome-causing alleles mRNA. All experiments were performed blind to injection cocktail in duplicate.

Table 1. *MAF* mutations identified in individuals with Aymé-Gripp syndrome.

Subject	Ref.	Nucleotide change[a]	Amino acid change[a]	Protein domain	Inheritance	Functional impact (radial SVM score/ CADD score)
1 (19474)	2	c.161C>T	p.Ser54Leu	TD	*de novo*	1.10/16.27
2 (11-1)	1	c.172A>G	p.Thr58Ala	TD	*de novo*, germline	0.83/16.07
3 (4-1)	1	c.206C>G	p.Pro69Arg	TD	*de novo*, germline	0.98/15.49
4 (ICN_ICW)	5	c.173C>T	p.Thr58Ile	TD	not available[b]	0.93/15.38
5 (14-1)	4	c.176C>A	p.Pro59His	TD	*de novo*	1.10/17.69
6 (10-1)	*p.s.*	c.176C>T	p.Pro59Leu	TD	*de novo*	0.98/9.08
7 (962112)	*p.s.*	c.185C>G	p.Thr62Arg	TD	*de novo*	1.10/17.88
8 (CaGi_UCSC)	*p.s.*	c.161C>T	p.Ser54Leu	TD	*de novo*, germline	1.10/16.27

p.s., present study; TD, transactivation domain.

[a]Nucleotide and amino acid positions refer to transcript variant 1 and protein isoform a (longer isoform) (NM_005360.4, NP_005351.2).

[b]Parental DNAs were not available for molecular analyses.

Figure 1

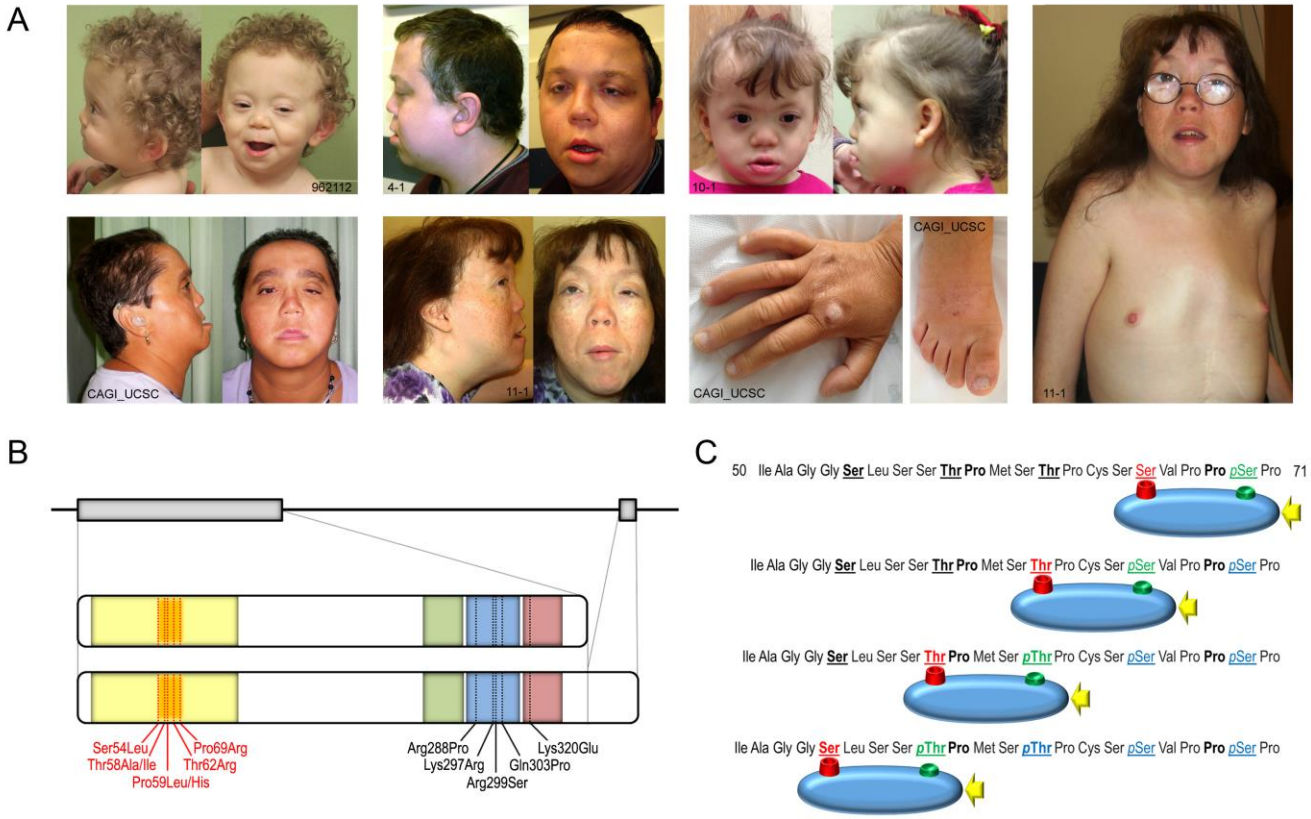


Figure 2

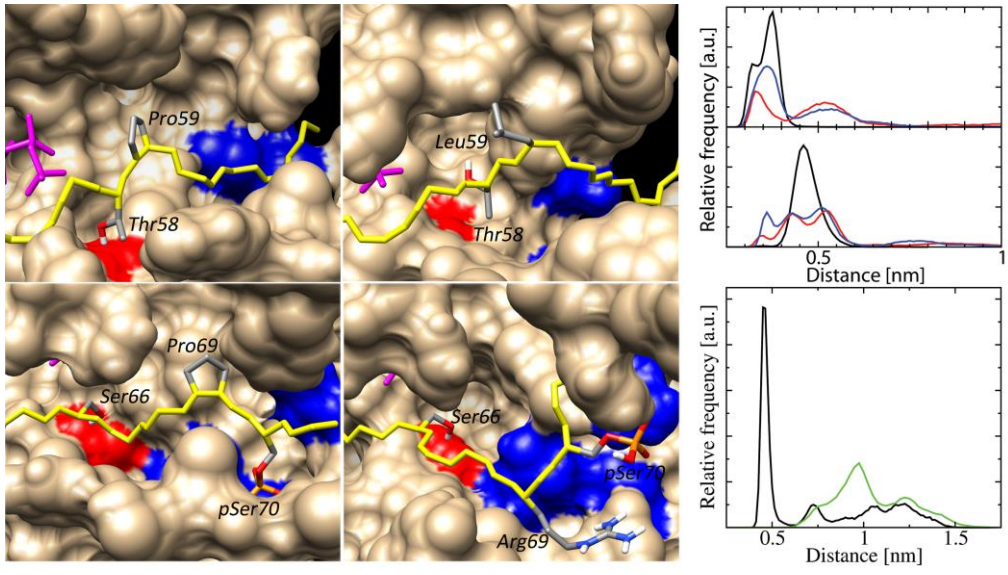


Figure 3

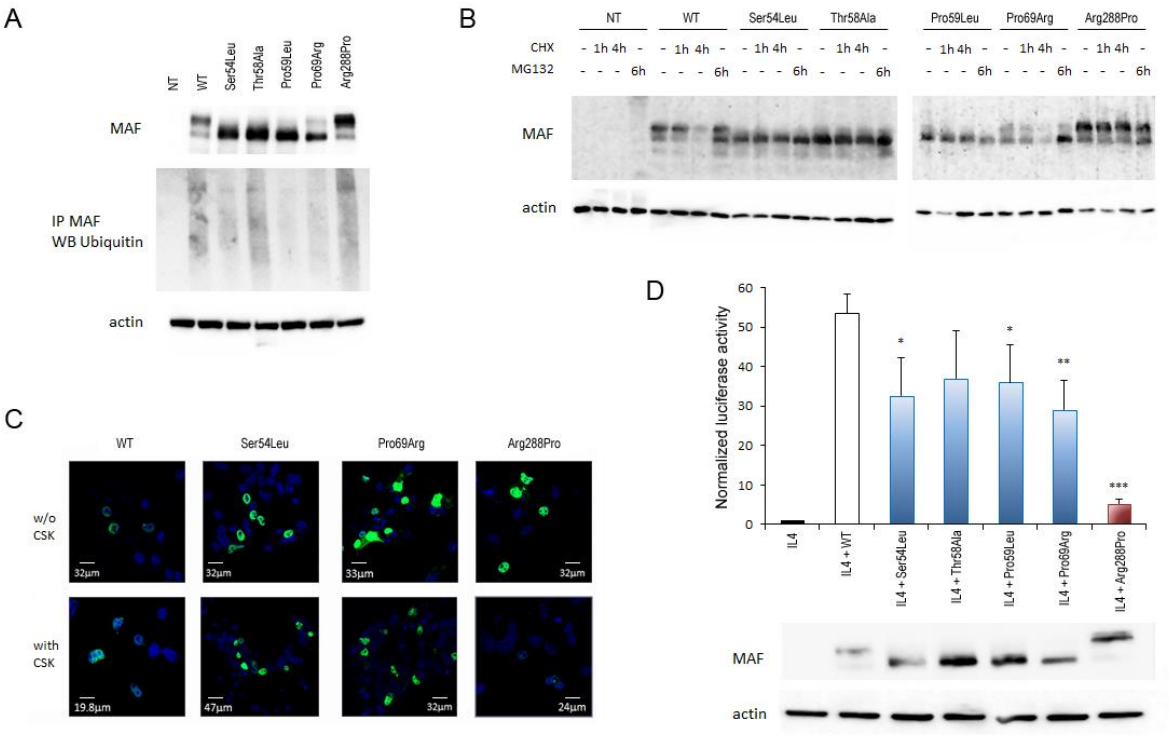
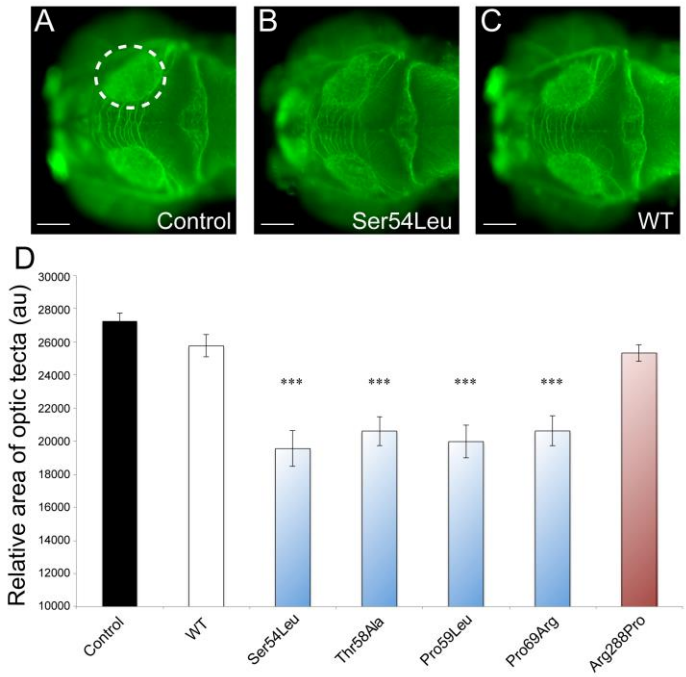


Figure 4



SUPPLEMENTAL DATA

Figure S1. Germline *MAF* mutations causing Aymé-Gripp syndrome.

Figure S2. Alignment of the *N*-terminal region of MAF to its orthologs and paralogs of the "large" MAF subfamily.

Figure S3. Maf protein levels in mouse embryo tissues.

Figure S4. Ramachandran plots of the residues at position 59 as obtained from the simulation of GSK3 and the wild-type, Pro59His, and Pro59Leu MAF peptides.

Figure S5. Putative genes positively controlled by MAF are differentially expressed in skin fibroblasts from *MAF* mutation-positive subjects with Aymé-Gripp syndrome compared to control fibroblasts.

Figure S6. Three-dimensional (above) and two-dimensional (below) chemical structure of the cluster resembling the GSK3 active site, considered in the quantum mechanical calculations to determine the parameters lacking in the force field used in the molecular dynamics simulations of the GSK3/MAF complexes.

Table S1. Clinical features of the subjects included in the study.

Table S2. Whole exome sequencing data output.

Table S3. List of genes having two or more private, rare or clinically associated heterozygous nonsynonymous variants or changes affecting splice sites, together with their predicted radial SVM and CADD scores.

Table S4. Validation and segregation analysis of variants predicted to have a significant impact on protein structure or function, and affecting genes considered as excellent candidates by function.

Table S5. Primer pairs and annealing temperatures used to amplify the *MAF* coding sequence, and sizes of PCR products.

Table S6. MAF peptides included in the molecular dynamics simulations of the GSK3/MAF complexes.

Table S7. Treatment with CSK buffer does not affect nuclear localization of Aymé-Gripp syndrome-causing MAF mutants in transiently transfected COS1 cells.

Table S8. Gene Ontology enrichment analysis of the genes differentially expressed in skin fibroblasts from two *MAF* mutation-positive subjects compared to control fibroblasts.

Supplemental references

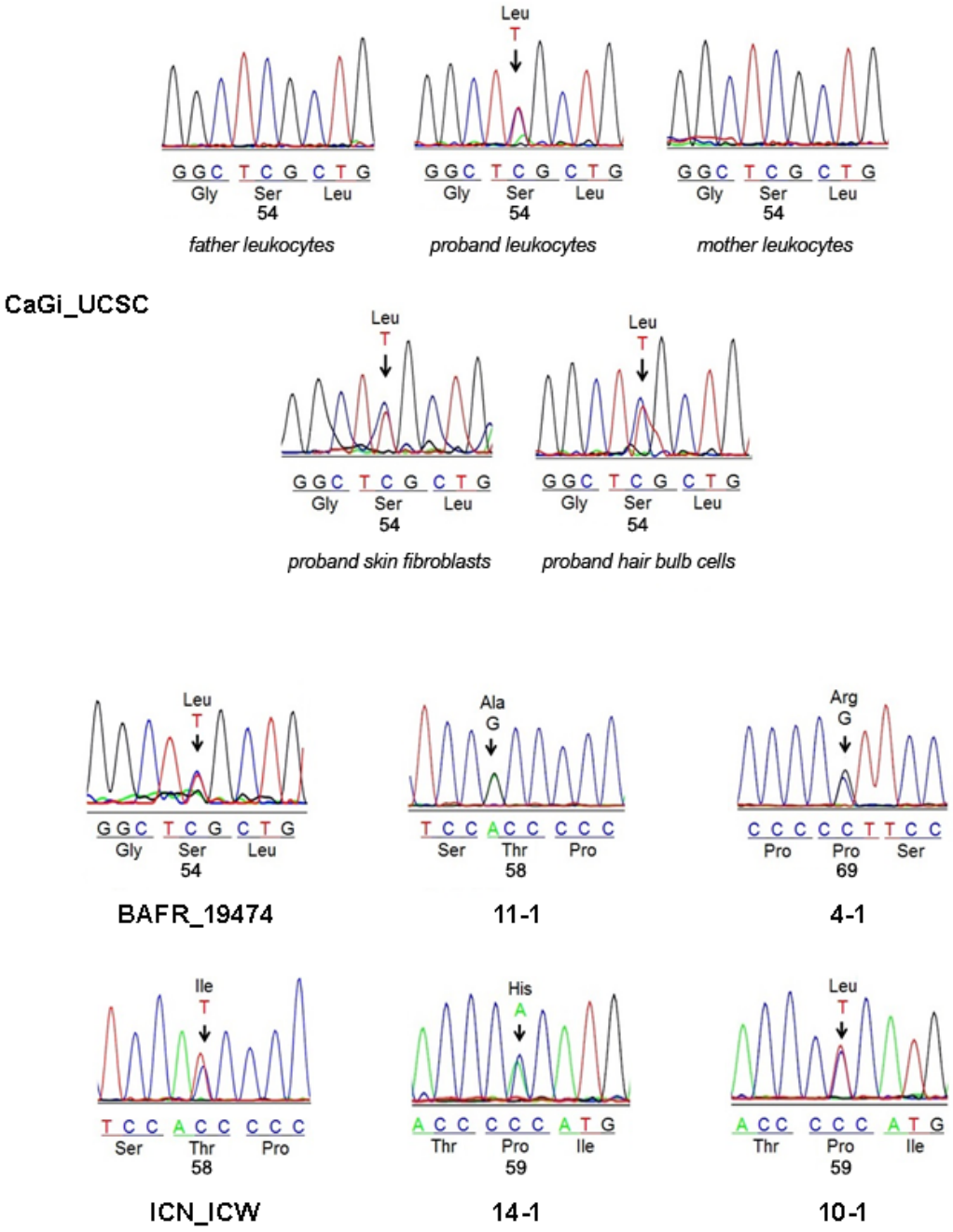
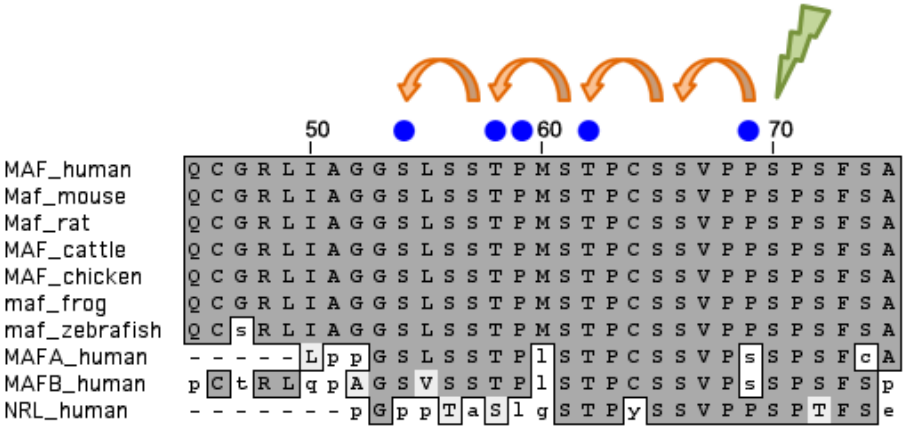


Figure S1. Germline *MAF* mutations causing Aymé-Gripp syndrome.

Sequence chromatograms showing the *de novo* origin of the *MAF* missense changes in patient 8 (sporadic case CaGi_UCSC), and documenting the heterozygous state of mutation in peripheral leukocytes, hair bulb cells and skin fibroblasts. A panel of representative disease-causing mutations is shown.

(A)



(B)

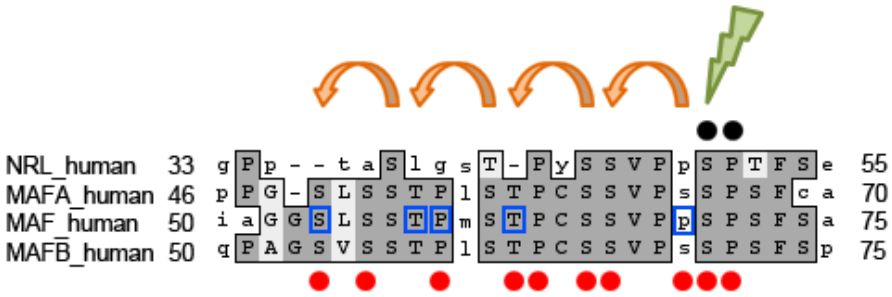


Figure S2. Alignment of the *N*-terminal region of MAF to its orthologs and paralogs of the "large" MAF subfamily.

(A) Mutated residues are indicated by blue circles, while those sequentially targeted by GSK3 are showed by arrows. The priming Serine residue is indicated by a green arrow. (B) The cartoon shows the affected residues in the domain encompassing the GSK3 recognition motifs in MAFB (red circles) and NRL (black circles) causing multicentric carpotarsal osteolysis and autosomal dominant retinitis pigmentosa, respectively. MAF residues mutated in Aymé-Gripp syndrome (blue squares) are reported for comparison.

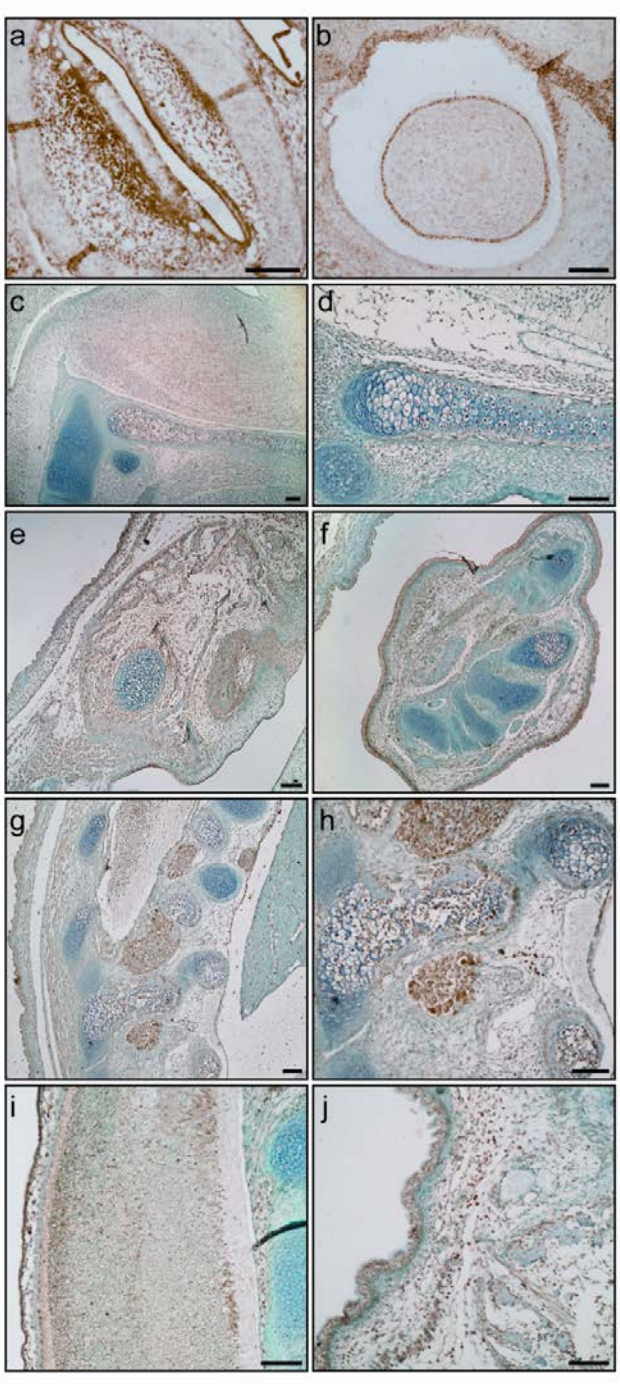


Figure S3. Maf protein levels in mouse embryo tissues.

(a) cochlea; (b) lens; (c, d) hind brain and cartilage primordium of basioccipital bone; (e) Meckel cartilage; (f) limb; (g, h) vertebrae and rib cartilage; (i) spinal cord; (j) skin. Embryos (E12.5-E16.5) were fixed in 4% paraformaldehyde, paraffin embedded, and sectioned (8 μm thickness). Antigen unmasking was performed by incubating sagittal/coronal sections in 10 mM sodium citrate pH 6, 2x 5 min, in a microwave (750 watts). Rabbit anti-MAF (1:30 final dilution) was used as primary antibody, followed by staining with rabbit Vectastain Elite ABC/DAB peroxidase substrate kit (Vector Laboratories). Nuclear counterstaining was performed with 0.5% methylgreen.

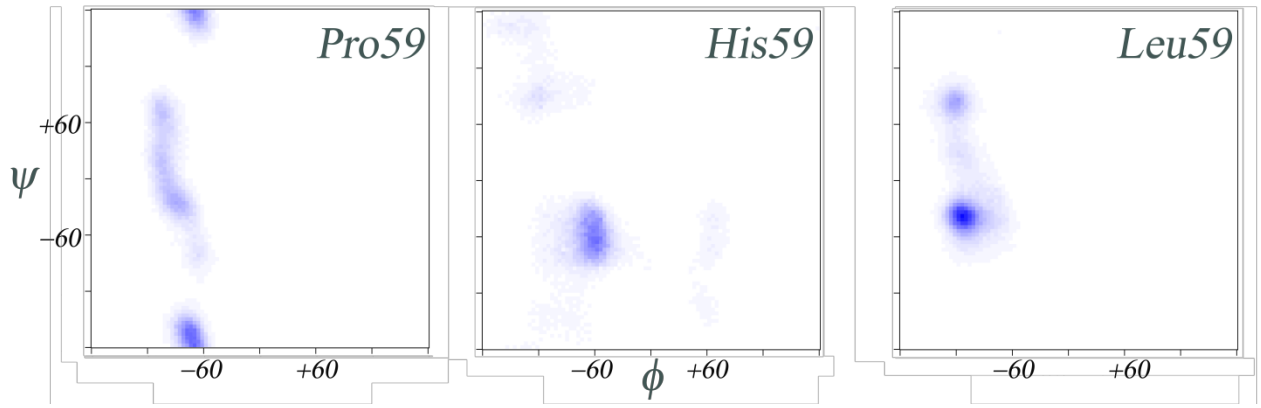


Figure S4. Ramachandran plots of the residues at position 59 as obtained from the molecular dynamics simulations of GSK3 and the wild-type, Pro59His, and Pro59Leu MAF peptides.

In the WT (P59), residue 59 populated predominantly the region corresponding to extended conformations (ψ close to ±180), while in the simulations of Pro59His (H59) and Pro59Leu (L59) the region corresponding to helical conformations was prevalent (ϕ ~ -60° and ψ ~ -50°). This difference is probably due to the peculiar conformational properties of Pro residues. These data indicate that in both P59 mutants the conformations of the trimer comprised between the target and primed residues are significantly different with respect to the WT.

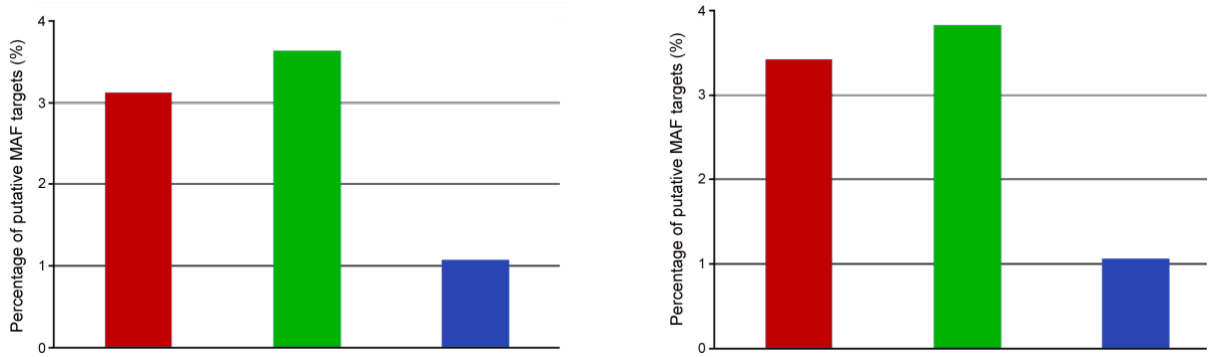


Figure S5. Putative genes positively controlled by MAF are differentially expressed in skin fibroblasts from *MAF* mutation-positive subjects with Aymé-Gripp syndrome compared to control fibroblasts.

Putative MAF targets were identified taking advantage of publicly available microarray data (ArrayExpress, accession code E-GEOD-51231), focusing on genes whose expression is controlled by MAF overexpression in MCF-7 cell lines. To define significantly up-regulated genes we calculated the rank-product (as defined in ref. 1) from the line expressing the short and long MAF isoforms, on RMA-normalized data.[2] The 500 Affymetrix probe-sets with smallest rank-product were annotated with gene symbol. These genes were considered as putative MAF targets. This gene-set was then tested for enrichment among up-, down-, and non-regulated genes in 4-1 and CaGi_UCSC cell lines *versus* control fibroblasts by means of Pearson's χ^2 test, implemented in R (see Table S7). Post-hoc tests were carried out among all pairs of populations adjusting the resulting *p*-values for inflation due to multiple comparisons (using B&H's False Discovery Rate correction).

Histograms represent the percentages of putative MAF targets among genes up-regulated (red), down-modulated (green) or non-differentially expressed (blue) in *MAF* mutation-positive subjects (4-1, left; CaGi_UCSC, right) compared to control fibroblasts (ATCC code PCS-201-012). In both comparisons, the observed differences between distributions were statistically significant (Pearson's χ^2 test, $p < 0.001$).

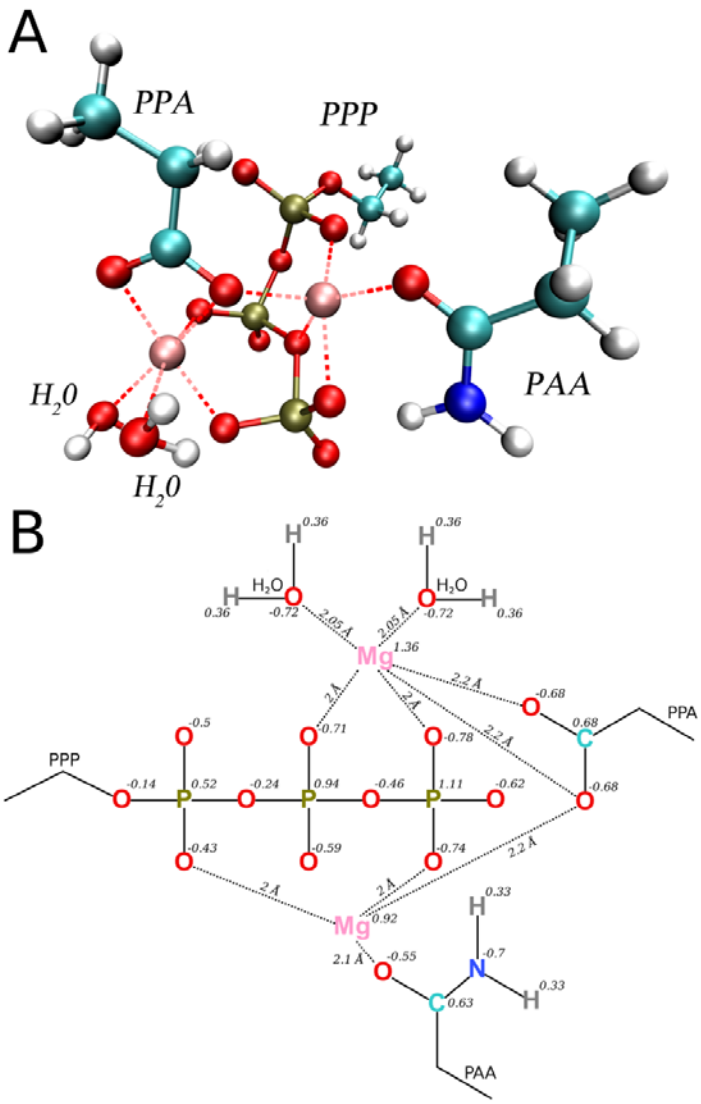


Figure S6. **Three-dimensional (above) and two-dimensional (below) chemical structure of the cluster resembling the GSK3 active site, considered in the quantum mechanical (QM) calculations to determine the parameters lacking in the force field used in the molecular dynamics simulations of the GSK3/MAF complexes.**

Ethyl-triphosphate (PPP), propionic acid (PPA), and propionic acid amide (PAA) have been used in replacement of ATP, and the Asp200 and Asn186 side chains, respectively. In the cluster, the Mg^{2+} ions and their coordinating water molecules (H_2O) have been explicitly considered. After a geometry optimization and a single-point calculation, the point charges were obtained by means of the restrained electrostatic potential (RESP) procedure, according to the CHELPG scheme.[3,4] All QM calculations were performed at B3LYP/cc-pVDZ level using the Gaussian 09 software package.[5-8] The Lennard-Jones parameters for Mg^{2+} ions were taken from the CHARMM27 force field.[9] Atoms are highlighted according the following color code: Mg, pink; P, gold; C, cyan; N, blue; O, red; H, white. Coordination bonds are depicted as dashed lines. Coordination bond lengths and the atomic point charges are reported (below).

Table S1. Clinical features of the subjects included in the study.

Subject	**1** **19474**	**2** **11-1**	**3** **4-1**	**4** **ICN_ICW**	**5** **14-1**	**6** **10-1**	**7** **962112**	**8** **CAGI_UCSC**	**9** **19-1**	**10** **110101**[a]	**11** **110102**[a]	**12** **FP103**	**13** **RML-01**
Reference	10	11	11	12	13	p.s.	p.s.	p.s.	13	14	14	15	p.s.
NT change	c.161C>T	c.172A>G	c.206C>G	c.173C>T	c.176C>A	c.176C>T	c.185C>G	c.161C>T	-	-	-	-	-
AA change	p.Ser54Leu	p.Thr58Ala	p.Pro69Arg	p.Thr58Ile	p.Pro59His	p.Pro59Leu	p.Thr62Arg	p.Ser54Leu	-	-	-	-	-
Sex	F	F	M	M	F	F	M	F	M	M	F	F	F
Age (years, last evaluation)	39	27	33	22 months	9	5	21 months	43	5	11	5 months	aborted fetus	12 months
Growth													
height/lenght (cm)	147	138	154	85	127	96	81	148	104	129	59.8	n.a.	69.5
weight (kg)	34	45	n.a.	9.2	23.5	14.5	10.4	61	16.4	27.4	5.8	n.a.	8.4
OFC (cm)	n.a.	52	55	46	51	51.3	38.4	52.4	54.5	47	40	n.a.	47
Craniofacial features													
brachycephaly	+	+	+	+	+	+	+	+	-	+	-	+	+
craniosynostosis	+[b]	-	-	-	-	-	-	-	-	+[c]	-	-	n.a.
late closing fontanel	n.a.	+ (3 y)	n.a.	+	+ (18 m)	-	-	n.a.	+	n.a.	n.a.	n.a.	+
high forehead	+	-	-	+	+	+	+	+	+	-	-	-	+
prominent frontal bone	+	-	-	+	-	+	+	+	+	+	+	-	+
craniofacial asymmetry	+	-	-	-	-	+	+	+	-	-	-	+	-
midfacial hypoplasia	+	+	+	+	+	+	+	+	+	+	+	+	+[d]
Ears													
sensorineural hearing loss	+	+	+	+	+	+	+	+	+	+	+	n.a.	+
small ears	+	-	-	-	+	-	n.a.	-	-	-	-	-	-
low-set	+	+	-	+	+	+	-	+	+	-	+	+	+
posteriorly angulated ears	+	+	-	-	-	-	-	+	+	-	-	+	+
stenotic external auditory canal	n.a.	-	-	-	-	+	-	-	-	+	+	n.a.	-
Eyes													
cataract	+	+	+	+	+	+	+	+	+	-	-	+	+
sight loss/blindness	n.a.	n.a.	+	-	+	CD	+	+	-	-	-	-	n.a.
megalocornea	-	-	-	-	-	-	-	-	-	-	-	+	-
broad eyebrows	-	+	+	-	+	-	-	+	-	-	-	n.a.	-
downslanting palpebral fissures	-	-	+	-	-	upslanting	upslanting	+	-	-	-	+	-
hypertelorism	+	+	-	+	-	-	-	+	-	-	-	-	+
epicanthal folds	-	-	-	-	+	+	-	+	+	+	-	n.a.	-
palpebral ptosis	-	+	+	-	+	-	-	+	-	-	-	n.a.	-
Nose													
flat nasal bridge	+	-	-	+	+	+	+	+	+	-	+	+	+
broad nasal root	+	+	-	+	-	+	+	+	+	+	+	+	+
short nose	+	+	+	+	+	+	+	+	+	+	+	+	+
Oral region													
long philtrum	+	+	+	-	+	-	+	+	-	+	+	+	+
small mouth/microstomia	+	+	+	-	-	+	+	+	-	-	+	+	+
thin upper lip	+	+	+	-	+	-	-	+	-	-	+	+	+
malar hypoplasia	n.a.	+	+	+	+	+	-	+	+	-	-	+	-
prognathism	n.a.	+	-	-	+	-	-	+	-	-	-	-	-
oligodontia	n.a.	-	-	-	ST	-	-	ALC	-	+	n.a.	n.a.	
alveolar cleft	-	-	-	-	-	-	-	-	+	-	-	+	-
CNS involvement													
intellectual disability	+ (moderate)	+ (DQ 50)	+ (DQ <50)	+ (severe)	+ (DQ 42)	+ (mild)	+ (mild)	+ (severe)	+ (mild)	+	+	n.a.	+
large ventricles/hydrocephaly	+	n.a.	+	-	MV	+	-	+	+	+	n.a.	n.a.	-
cerebral atrophy	-	n.a.	n.a.	-	+	-	-	+	n.a.	+	n.a.	n.a.	n.a.
Chiari 1 malformation	-	+	n.a.	-	-	-	-	+	-	-	n.a.	n.a.	-
febrile seizures	-	+	n.a.	-	-	-	+	+	-	-	-	n.a.	-
epilepsy/abnormal EEG	+	grand mal	grand mal	-	+	+ (3 y)	+ (16 m)	+	-	-	+	n.a.	-
others			autistic	seizures	irritable/SIB		DSC	EST	PCV				CH

Table S1. (continued)

Subject	1 BAFR_19474	2 11-1	3 4-1	4 ICN_ICW	5 14-1	6 10-1	7 962112	8 CAGI_UCSC	9 19-1	10 110101[a]	11 110102[a]	12 FP103	13 RML-01
Reference	10	11	11	12	13	p.s.	p.s.	p.s.	13	14	14	15	p.s.
NT change	c.161C>T	c.172A>G	c.206C>G	c.173C>T	c.176C>A	c.176C>T	c.185C>G	c.161C>T	-	-	-	-	-
AA change	p.Ser54Leu	p.Thr58Ala	p.Pro69Arg	p.Thr58Ile	p.Pro59His	p.Pro59Leu	p.Thr62Arg	p.Ser54Leu	-	-	-	-	-
Sex	F	F	M	M	F	F	M	F	M	M	F	F	F
Age (years, last evaluation)	39	27	33	22 months	9	5	21 months	43	5	11	5 months	aborted fetus	12 months
Ectodermal features													
mammary gland hypoplasia	+	+						+					
thick skin	n.a.	n.a.	n.a.	-	-	-	-	+	n.a.	n.a.	n.a.	n.a.	-
sparse scalp hair	+	-	-	-	-	-	+	+	n.a.	+	+	n.a.	
dystrophic nails	+	-	-	-	-	-	+	+	n.a.	+	+	n.a.	-
Limbs													
brachydactyly	+	-	+	+ (5th finger)	-	-	-	-	-	+	+	+	-
camptodactyly	-	-	-	-	+, great toes	-	-	+ (5th finger)	n.a.	+	+	n.a.	-
clinodactyly	+ (5th finger)	-	+ (5th finger)	-	-	-	-	+ (5th finger)	+	+	-	+	+ (5th finger)
tapered fingers	+	-	-	-	APC	-	+	-	-	+	+	-	+
joint limitations	-	multiple	+ (hips)	-	-	-	CDRH	multiple	-	+ (digits)	+ (digits)	-	-
radio-ulnar synostosis	n.a.	+	n.a.	-	-	-	-	-	-	-	-	-	-
hip chondrolysis	n.a.	+ (HR, 14y)	+ (HR, 16 y)	-	-	-	-	-	n.a	-	-	n.a.	-
prominent heels	n.a.	n.a.	n.a.	-	-	-	-	-	n.a.	+	-	+	n.a.
Genitourinary													
cryptorchidism			CTT	-					-	-			
albuminuria/proteinuria	-	n.a.	n.a.	-	n.a.	-	-	+	n.a.	n.a.	n.a.	n.a.	n.a.
kidney anomalies	-	n.a.	n.a.	-	-	-	-	+ (MG)	-	-	-	-	-
Other clinical features													
cardiac defects	-	-	-	-	-	ASD	-	-	PDA, VM	-	-	-	VSD
gastrointestinal defects	-	hiatal hernia	constipation	-	constipation	-	-	-	-	-	-	-	+
skeletal defects	-	PC	n.a.	-	-	-	-	-	-	-	-	-	-
pericarditis	n.a.	+ (2 y)	n.a.	-	PE (9-11 m)	+ (3 y)	PE (2-3 m)	-	-	-	-	-	-
daphragmatic hernia	n.a.	+ (19-20 m)	n.a.	-	-	n.a.	-	-	-	-	-	-	-
others		HM	HT, LPH			HM		PFP	Del11q				
Previous genetic testing													
Karyotype analysis (blood)	46, XX	46, XX	46, XY		46, XX	46, XX		46, XY	46, XY	46, XY		46, XY	46, XY
Karyotype analysis (fibroblasts)		46, XX				46, XX							
FISH analyses		subtelomeres			subtelomeres				Del11q	telomeres		22q11.2,17p13.3	
Array CGH					+		+	+	Del11q	+			+
SNP analysis													
others					*GJB2* sequencing	oligoarray	hearing loss gene panel			GFR2,FGFR3, TWIST			

[a] Siblings.

[b] Coronal (suspected).

[c] Coronal (surgically repaired).

[d] The overall facial appearance was not suggestive for the conditions reported by Aymé and Philip,[10] and Gripp et al.[11]

CD, cone dystrophy; ST, supernumerary teeth; ALC, agenesis of lateral canine; MV, mild ventriculomegaly; DSC, diffuse mild dilatation of semicircular canals; EST, enlargement of sella turcica; PCV, prominent cavum vergae; CH, cerebellar hypoplasia; APC, abnormal palmar creases; CDRH, bilateral congenital dislocation of the radial heads; HR, hip replacement; CTT, congenital testicular torsion; MG, mesangiocapillary glomerulopathy; ASD, atrial septal defect; PDA, patent ductus arteriosus; VM, ventriculomegaly; VSD, ventricular septal defect; PC, pectus carinatum; PE, pericardial effusion; HM, hypopigmented maculae of the retina; HT, hypothyroidism, LPH, wide neck with low posterior hairline; PFP, prominent fetal pads; Del, deletion.

Table S2. Whole exome sequencing data output.

Total number of reads	120,582,670
Mean read length (bp)	90
Reads mapped to target region[a]	75,598,162
Duplication rate (%)	12.0
Target regions coverage (%)[a]	99.2
Target regions coverage >10x (%)[a]	96.7
Average depth on target	84x
Variants with predicted effect on CDS[b]	11,168
- Annotated variants (dbSNP138) w/o clinical association and minor allele frequency > 0.001	10,144
- Annotated variants (in-house database) with frequency > 0.5%	765
- Novel variants	149
- Variants with clinically associated rs[c]	3
- Variants with low[d] or unknown frequency	107

[a] Nimblegen SeqCap EZ Library v 3.0
[b] Non synonymous SNPs and indels within the coding sequence and splice sites (±5 bases)
[c] APOE (rs769452), SLC5A2 (rs61742739), IKBKG (rs148695964).
[d] Minor allele frequency < 0.001 in dbSNP138.

Table S3. List of genes having two or more private, rare or clinically associated heterozygous nonsynonymous variants or changes affecting splice sites, together with their predicted radial SVM and CADD scores. No homozygous private, rare or clinically associated missense, nonsense, frameshift or splice site variant was identified. Similarly, WES data analyses ruled out occurrence of homozygosity/compound heterozygosity for cataracts- or hearing loss-associated private, rare or clinically associated nonsynonymous/splice site variants. ITGB4 was the only gene having two nonsynonymous/splice site variants satisfying the filtering criteria (dbNSFP scores ≥ 0 or CADD score ≥ 15). Validation and segregation analysis of these variants documented their transmission *in cis* from the paternally inherited allele. The GeneDistiller score, as an estimate of the functional relevance of individual genes is also reported.

gene symbol	position	ref. allele	var. allele	predicted change	novel/ annotated	Radial[a] SVM score	CADD[a] score	Gene Distiller OS	Validation and segregation (Sanger)
UTRN	chr6:144747451	C	T	S145L	rs115198550	0.16	19.20	11933	
UTRN	chr6:144780371	C	T	S863L	novel	-1.01	6.59	11933	
ITGB4	chr17:73725378	C	T	P200L	rs148770294	1.04	16.83	5888	inherited (father)
ITGB4	chr17:73736469	G	A	R826H	rs200694443	0.13	17.4	5888	inherited (father)
TTN	chr2:179438501	G	A	L15055F	rs372309164	-0.33	11.69	3729	
TTN	chr2:179578780	C	T	E7625K	rs375373507	-0.52	14.22	3729	
TTN	chr2:179654748	C	T	G586D	rs150231219	-0.98	11.97	3729	
PWP2	chr21:45544547	G	C	R635P	rs144635092	-0.98	10.24	148	
PWP2	chr21:45550499	A	C	N869T	rs148288077	-1.00	0.18	148	
MUC2	chr11:1081421	C	T	A480V	rs181019515	-0.64	11.24	92	
MUC2	chr11:1098756	G	A	A2372T	rs200767028	0.65	11.43	92	
TSNAXIP1	chr16:67848293	C	T	R46*	rs146214814	n.a	14.83	66	
TSNAXIP1	chr16:67859170	C	T	S201L	rs202171637	-0.91	19.33	66	
CCDC168	chr13:103382321	C	T	S6909N	novel	-0.94	12.39	0	
CCDC168	chr13:103393487	T	A	N3187I	novel	n.a.	3.39	0	

[a] Scores <0 (dbNSFP) or <15 (CADD) predict a negligible impact of the sequence change on protein structure and function.

Table S4. Validation and segregation analysis of variants predicted to have a significant impact on protein structure or function, and affecting genes considered as excellent candidates by function.

Impact of mutations was predicted using CADD and dbNSFP tools, while the functional relevance of each gene was evaluated by GeneDistiller.

gene symbol	position	ref. allele	var. allele	predicted change	novel/ annotated	Radial[a] SVM score	CADD[a] score	Gene Distiller OS	Validation and segregation (Sanger)[b]
MAF	chr16:79633639	G	A	S54L	novel	1.10	16.27	20707	*de novo*
GRM5	chr11:88300327	C	T	V842M	rs144685614	0.63	20.30	14388	inherited (father)
UTRN	chr6:144747451	C	T	S145L	rs115198550	0.16	19.20	11933	inherited (mother)
CRYZ	chr1:75185057	C	T	c.265-1G>A	rs139734636	n.a.	12.00	10148	inherited (mother)
USH1G	chr17:72915993	C	G	G210A	novel	0.97	24.30	9097	inherited (father)
GRM6	chr5:178413188	AG	A	P689fs	novel	n.a.	38.00	6390	inherited (father)
ITGB4	chr17:73725378	C	T	P200L	rs148770294	1.04	16.83	5888	inherited (father)
ITGB4	chr17:73736469	G	A	R826H	rs200694443	0.13	17.4	5888	inherited (father)
ZNF559	chr19:9453372	GA	G	K374fs	novel	n.a.	18.25	4174	inherited (father)
FTO	chr16:53878088	A	G	D258G	rs369122168	-0.94	15.14	3952	inherited (father)
TCL1A	chr14:96178699	C	T	R52H	rs146705048	-0.98	15.88	3390	inherited (father)
KCNA10	chr1:111060845	G	T	P189T	novel	0.02	10.29	3340	inherited (mother)
CD36	chr7:80301310	T	G	L284*	rs56381858	n.a.	37.00	3148	inherited (mother)
SLC9A1	chr1:27428886	CGGT	C	T603del	novel	n.a.	16.37	2994	inherited (father)
ADRM1	chr20:60883142	G	A	E269K	novel	-0.91	18.04	2974	inherited (father)
HCN4	chr15:73614986	C	T	V1150I	novel	0.36	5.11	2916	inherited (father)
PCSK9	chr1:55517957	G	A	G177D	rs368899514	-0.23	20.80	2820	inherited (mother)
TNR	chr1:175334216	C	G	E839D	rs143313777	-0.73	16.93	2639	inherited (father)
CUX1	chr7:101740766	G	A	E105K	novel	-0.35	25.60	2567	inherited (mother)
EPB41L2	chr6:131191149	C	T	G651R	rs139708296	-1.03	18.32	2466	inherited (mother)
AMD1	chr6:111214240	A	G	K159R	rs201112162	-0.97	15.33	2440	inherited (mother)
LAMB2	chr3:49166234	G	A	R584C	novel	-0.50	36.00	2367	inherited (mother)
PRDM16	chr1:3328010	G	A	A417T	novel	-0.58	25.40	2229	inherited (mother)
PAPLN	chr14:73732178	G	A	E1049K	novel	-0.31	28.50	2223	inherited (father)
TNKS1BP1	chr11:57080213	GCCT	G	EA649A	novel	n.a.	0.54	2190	inherited (mother)
LMNB2	chr19:2435112	G	A	R248W	novel	0.15	23.20	2167	inherited (mother)
PLEKHG4	chr16:67319264	A	G	E675G	rs142391556	-0.08	19.49	2017	inherited (father)
SYNE1	chr6:152557441	A	G	c.19987-3T>C	novel	n.a.	9.99	2009	inherited (mother)
ITPR2	chr12:26592167	C	A	R2179L	novel	0.91	33.00	232	
ANKRD17	chr4:74012578	T	C	H478R	novel	-1.06	17.85	187	
NUP205	chr7:135301979	C	T	T1225I	rs151283740	-1.07	15.28	177	
HMGB1	chr13:31035509	ATCT	A	ED210D	novel	n.a.	11.19	170	
MYO1E	chr15:59510221	G	A	R326W	rs150397493	0.98	18.52	162	
CEP72	chr5:620335	G	A	R121H	novel	-0.08	18.74	158	
CYTH2	chr19:48982421	C	T	A385V	novel	-1.10	18.96	154	
IFT140	chr16:1642285	C	T	G176S	rs201849591	-0.16	22.30	151	

Table S4. (continued)

UBE4A	chr11:118250258	C	T	L557F	rs150156682	-0.71	29.30	150	
MRPS28	chr8:80942312	G	A	R58W	rs147337080	-0.80	16.74	147	
XCR1	chr3:46062869	T	C	N191D	novel	-0.03	16.26	140	
PHKG1	chr7:56148922	C	T	R321Q	rs147014225	-1.12	17.95	138	
APBA3	chr19:3751261	CAAT	C	I527del	novel	n.a.	17.74	134	
PARP4	chr13:25049736	TAAA	T	c.1790-3delTTT	novel	n.a.	8.12	131	
SLC11A2	chr12:51390751	A	G	V223A	novel	-0.56	17.61	124	
RSRC1	chr3:158262021	T	C	I263T	rs371320282	-0.92	16.95	123	
DCAKD	chr17:43111687	C	T	G62S	novel	0.50	34.00	118	
PIEZO1	chr16:88793447	C	T	R1152K	novel	-0.24	19.49	115	
RPL34	chr4:109543777	A	G	c.269+3A>G	novel	n.a.	13.16	114	
DHTKD1	chr10:12149921	G	A	W687*	novel	n.a.	38.00	113	
AMBRA1	chr11:46419422	C	T	G1040S	novel	0.43	34.00	111	
ZNF112	chr19:44832439	ACT	A	RV623fs	novel	n.a.	36.00	110	
KBTBD4	chr11:47594766	C	T	A425T	novel	-0.29	28.50	108	
FAM45A	chr10:120883007	G	A	R207Q	novel	-0.08	35.00	106	
CSRNP2	chr12:51467822	A	C	N65K	rs140071191	-1.02	17.47	103	
CHML	chr1:241797994	C	T	A359T	rs199932635	-0.05	16.26	100	
NFASC	chr1:204951037	G	A	V783M	rs147192388	-0.23	17.23	98	
LAMB3	chr1:209807809	G	A	R183C	rs116124880	-0.68	17.51	93	
MUC2	chr11:1098756	G	A	A2372T	rs200767028	0.66	11.43	92	
LLGL1	chr17:18145250	G	A	R940Q	rs367698200	-1.02	16.93	91	
CLSTN3	chr12:7288885	C	A	P276T	novel	-0.99	17.02	89	
IFIT5	chr10:91177303	A	G	Y116C	rs200078352	1.16	13.31	88	
RASGRF2	chr5:80408475	A	C	K629Q	novel	-0.96	27.80	88	
TLL2	chr10:98192642	G	C	R148G	novel	-1.16	19.59	86	
LAMA3	chr18:21494509	C	T	R2433W	rs147463397	0.42	16.74	86	
PADI3	chr1:17599895	T	C	S370P	novel	-0.63	24.80	84	
SDK2	chr17:71397222	C	T	A970T	novel	0.48	25.50	83	
F10	chr13:113793752	T	C	L113S	novel	1.11	0.03	80	
SORBS2	chr4:186570657	G	A	P178L	novel	-0.91	15.56	78	
PIGV	chr1:27124222	C	T	L457F	rs143676075	0.15	19.27	76	
ELP5	chr17:7156268	G	A	V63M	novel	-0.60	20.20	76	
LRP1B	chr2:141116518	G	A	c.11132-3C>T	novel	n.a.	8.86	71	
FAM124B	chr2:225266040	T	G	Y149S	rs150435753	-0.23	25.50	69	
MYH1	chr17:10419954	A	C	S2R	novel	0.23	15.18	67	
TSNAXIP1	chr16:67859170	C	T	S201L	rs202171637	-0.91	14.83	66	
TSNAXIP1	chr16:67848293	C	T	R46*	rs146214814	n.a.	14.83	62	
NACAD	chr7:45120336	C	T	R1526H	rs371048713	-1.07	24.60	61	
PCNX	chr14:71445150	C	T	A699V	novel	0.03	8.57	60	
KIF7	chr15:90173640	G	A	R1066C	rs77474187	0.22	18.89	48	
GPRC5D	chr12:13102481	18 nt	A	PLQGNA274fs	novel	n.a.	19.74	61	
TCP11L1	chr11:33065395	G	A	E26K	rs187732111	-1.09	16.65	46	

Table S4. (continued)

PLEKHG2	chr19:39914225	G	C	R844P	rs201914405	-0.33	22.20	45	
DNAH3	chr16:20975250	T	C	Q3319R	novel	0.17	22.20	44	
FUT1	chr19:49253775	G	A	A255V	rs200299477	0.56	13.82	40	
ELK3	chr12:96617357	A	G	I5V	rs370644588	-1.06	18.76	39	
UBAC1	chr9:138839673	G	A	R138W	novel	-1.08	15.87	39	
SLC15A3	chr11:60704782	G	A	T552M	rs199844407	-0.22	35.00	34	
SBDS	chr7:66460278	C	A	V43L	rs147652512	0.29	19.63	33	
HSPA6	chr1:161496310	GC	G	G621fs	novel	n.a.	31.00	31	
SLC5A8	chr12:101603467	C	T	A54T	rs149707334	0.22	18.81	29	
SEZ6	chr17:27308629	C	T	A162T	rs199557265	-1.09	19.53	28	
DNAH6	chr2:84936627	C	T	T3070I	novel	0.17	19.90	24	
SLC26A8	chr6:35960411	T	C	Y223C	rs148004355	0.97	14.19	21	
LIN54	chr4:83867599	G	C	T106S	rs34929292	-0.99	17.30	20	
ADAMTSL1	chr9:18681858	A	G	I464V	novel	-1.06	15.33	16	
TBRG1	chr11:124502029	GTCT	G	S378del	novel	n.a.	13.50	14	
GLB1L	chr2:220104777	TG	T	F195fs	novel	n.a.	13.77	9	
MIEN1	chr17:37885803	T	C	N107S	novel	-1.05	18.88	9	
ANLN	chr7:36435951	G	A	R32K	rs141973124	-0.43	33.00	8	
POLN	chr4:2172869	C	T	A472T	novel	-1.06	22.30	6	
C2orf71	chr2:29296517	A	G	I204T	novel	-1.05	15.23	6	
TTC25	chr17:40095250	C	T	Q295*	novel	n.a.	6.15	5	
CEP95	chr17:62504753	A	G	I21M	novel	0.99	17.50	5	
PYROXD2	chr10:100144778	C	T	R534H	rs138490501	0.06	32	4	
TMPPE	chr3:33134405	C	T	G291D	rs202056683	0.51	22.70	2	
GFY	chr19:49931331	C	G	R438G	rs140070857	n.a.	11.62	1	
SLC35E4	chr22:31032693	C	A	H86N	novel	-0.17	24.50	1	
TMEM244	chr6:130152529	T	G	M108L	rs140701252	-1.05	18.54	1	
CCDC63	chr12:111342502	T	TG	W406fs	novel	n.a.	23.80	0	
CCDC168	chr13:103393487	T	A	N3187I	novel	n.a.	3.39	0	
C19orf47	chr19:40839772	C	T	R133H	novel	-0,2563	17.73	0	
CYB5R2	chr11:7686684	G	A	P251L	rs143187886	1.08	27.30	-11	
IQCG	chr3:197652966	C	A	G219*	rs140573176	n.a.	23.70	-14	

[a] Scores <0 (dbNSFP) or <15 (CADD) predict a negligible impact of the sequence change on protein structure and function.
[b] Variant validation and segregation analysis was performed for variants encoded by genes with GeneDistiller overall score >2000.

Table S5. Primer pairs and annealing temperatures used to amplify the *MAF* coding sequence, and sizes of PCR products.[a]

Exon	Primer Sequence (5′→3′)		Ann. Temp. (°C)	Product Length (bp)
	Forward	Reverse		
1a	CAAGCTAGAAGCGCCCCAGC	TGGTGGCTGTTGCTGATGAG	64[b]	492
1b	AGCAGAAGGCGCACCTGGAAG	CTTCTGCTTCAGCCGGATCAC	64[b]	692
1c	GCACTTCGACGACCGCTTCT	GAGCATGGCTCTAGAACTAG	64[b]	448
2	ATCCTGAGTAAGTGCCATTC	AGCCTTCTTCTCTAACACAG	60	368

[a] *MAF* encodes two isoforms resulting from alternative transcript processing. The long isoform (NM_005360, NP_005351) counts two coding exons, while the short isoform (NM_001031804 NP_001026974) is encoded by exon 1 only (the transcript retains the downstream intronic sequence, with the termination codon, TGA, located at +2 from the splicing site).
[b] Hot-start protocol using 10% DMSO.

Table S6. MAF peptides included in the molecular dynamics (MD) simulations of the GSK3/MAF complexes.[a]

MAF decamers	Amino acid sequence[b]
Wild-type (residues 54-63)	Ser-Leu-Ser-Ser-Thr-**Pro**-Met-Ser-*pThr*-Pro
Pro59Leu mutant (residues 54-63)	Ser-Leu-Ser-Ser-Thr-**Leu**-Met-Ser-*pThr*-Pro
Pro59His mutant (residues 54-63)	Ser-Leu-Ser-Ser-Thr-**His**-Met-Ser-*pThr*-Pro
Wild-type (residues 62-71)	Thr-Pro-Cys-Ser-Ser-Val-Pro-**Pro**-*pSer*-Pro
Pro69Arg mutant (residues 62-71)	Thr-Pro-Cys-Ser-Ser-Val-Pro-**Arg**-*pSer*-Pro

[a] The selected sequences correspond to MAF stretches, in which residues Pro59 or Pro69 are located between the target and primed residues (Thr58 and pThr62, and Ser66 and pSer70, respectively). p.Pro69Arg could affect also the phosphorylation of Ser70, but this possibility was not investigated as the involved kinase has not been identified yet. In the starting conformations, the peptide was set with the phosphate group pointing toward the T-loop of the protein (residues 207-218) and the target Ser/Thr residue placed in between the ATP cofactor and the side chain of the catalytic Asp181. The complex was solvated in a dodecahedral box containing ~14500 water molecules and counterions.[16] To equilibrate the peptide in the GSK biding cleft, after minimization, the system was processed by means of 100 simulated annealing (SA) cycles of 50 ps each, with temperatures comprised between 300 K and 500 K, by restraining the coordinates of the protein and the peptide-protein distances in the active site. After SA, the structure with the lowest sum of protein-peptide binding energy and peptide conformational energy was equilibrated by two cycles of energy-minimization and thermalization (for a total MD time of 16 ns), while the position restraints were progressively released. Two productive MD runs of 40 ns were carried out at 300 K for each system, using the configurations obtained by repeating twice the last thermalization step, with different initial velocities.

[b] Residues that are substrates of GSK3-mediated phosphorylation are underlined. *pSer* and *pThr* indicate the primed phosphorylated residues. Affected codons are in bold. The peptides in the simulations were capped with an acetyl group at the *N*-terminus, whereas the *C*-terminus was considered uncharged and the phosphate group in the primed residue was considered singly protonated.

Table S7. Treatment with CSK buffer does not affect nuclear localization of Aymé-Gripp syndrome-causing MAF mutants in transiently transfected COS1 cells.

Cells were transfected with the relevant expression construct, and either fixed or treated with CSK buffer prior fixation to remove nuclear soluble protein pools.

MAF construct	w/o CSK buffer		with CSK buffer			
	cell counts	MAF-positive cells[a]	cell counts	MAF-positive cells[a]		*P*[b]
wild-type	450	129 (28.6%)	450	135	(30.0%)	0.66
p.Ser54Leu	450	104 (23.1%)	450	95	(21.2%)	0.47
p.Pro69Arg	450	114 (25.3%)	450	99	(20.7%)	0.24
p.Arg288Pro	450	128 (28.4%)	450	17	(3.8%)	<0.0001

[a] Nuclear localization of MAF proteins was observed in all MAF-positive cells.
[b] *P* values indicate the significance of differences in the number of MAF-positive nuclei between CSK buffer treated and untreated cells (χ^2 statistics based on 2x2 contingency table).

Table S8. Gene Ontology enrichment analysis of the genes differentially expressed in skin fibroblasts from two *MAF* mutation-positive subjects compared to control fibroblasts.

High-quality total RNA (RIN > 8; Bioanalyzer, Agilent Technologies) was obtained from cultured skin fibroblasts of two *MAF* mutation-positive subjects (4-1 and CaGi_UCSC) and control (ATCC code PCS-201-012) (each in duplicates). Expression levels were measured by using the Human Gene 1.0 ST arrays (Affymetrix). Hybridization, washing and staining (GeneChip Fluidics Station 450, Affymetrix), and scanning (GeneChip Scanner 3000) were performed following manufacturer's protocols. Microarray data analysis was performed by means of Bioconductor package v.2.11.[17] CEL files were imported and normalized using RMA algorithm,[2] and differentially expressed genes (DEG) were identified taking advantage of Limma package, imposing B&H's False Discovery Rate < 0.01.[18] Pearson's *r* correlation between the two MAF mutation-positive fibroblast lines was 0.988, in both biological replicates. We analyzed DEG *versus* control fibroblasts in both MAF mutation-positive subjects by means of DAVID functional annotation tool (version 6.7) to cluster them into different functional groups.[19] Enrichment analyses was performed using the Gene Ontology (GO) database (Biological Processes; "GOTERM_BP_FAT") on the 1,245 (4-1) and 1,191 (CaGi_UCSC) genes having statistically significant differential expression between sample groups On all DEG we identified the significantly enriched functional groups selecting the top 20 ranked ones, with an EASE score greater than or equal to 1.3 (corresponding to a *p*-value < 0.05), and setting the clustering stringency at medium level. Annotation clusters are listed in order of Enrichment Score (representative category per cluster is reported).

GO categories	Enrichment score	No. genes	*P*-value	Fold enrichment	FDR
Differentially expressed genes (4-1)					
Annotation cluster: Angiogenesis	*10.7*				
Blood vessel development (GO:0001568)		51	1.0×10^{-12}	3.00	1.9×10^{-9}
Annotation cluster: Control of cell migration	*10.0*				
Regulation of cell motion (GO:0051270)		45	1.9×10^{-12}	3.31	3.6×10^{-9}
Annotation cluster: Lung development	*7.7*				
Tube development (GO:0035295)		51	6.7×10^{-14}	3.29	1.2×10^{-10}
Annotation cluster: cell migration	*5.7*				
Cell motion (GO:0006928)		67	6.8×10^{-8}	2.00	1.2×10^{-4}
Annotation cluster: Response to stimuli	*4.8*				
Response to hormone stimulus (GO:0009725)		50	1.0×10^{-5}	1.93	0.019
Annotation cluster: Tissue morphogenesis	*3.7*				
Tissue morphogenesis (GO:0048729)		32	2.9×10^{-6}	2.52	0.005
Annotation cluster: Regulation of phosphate metabolism	*3.7*				
Regulation of phosphorylation (GO:0009991)		60	6.2×10^{-6}	1.83	0.013
Annotation cluster: Urogenital system development	*3.1*				
Kidney development (GO:0001822)		22	2.4×10^{-6}	3.25	0.004
Annotation cluster: Embryonic development	*2.9*				
Embryonic morphogenesis (GO:0048598)		40	2.9×10^{-4}	1.80	0.523
Annotation cluster: Cell adhesion	*2.9*				
Cell-substrate adhesion (GO:0031589)		18	4.5×10^{-4}	2.61	0.812
Annotation cluster: Skeletal muscle development	*2.9*				
Muscle organ development (GO:0007517)		36	1.6×10^{-6}	2.42	0.003
Annotation cluster: Cell morphogenesis	*2.9*				
Regulation of cell development (GO: 0060284)		35	2.3×10^{-6}	2.42	0.004

Table S8. (continued)

GO categories	Enrichment score	No. genes	*P*-value	Fold enrichment	FDR
Annotation cluster: Cell death	*2.8*				
Cell death (GO: 0008219)		75	5.6×10^{-4}	1.48	1.017
Annotation cluster: Cardiac cell differentiation	*2.8*				
Cardiac muscle tissue development (GO: 0048738)		15	2.9×10^{-5}	3.74	0.054
Annotation cluster: Phosphate metabolic process	*2.7*				
Protein amino acid phosphorylation (GO: 0006468)		71	4.6×10^{-4}	1.51	0.845
Annotation cluster: Bone development	*2.4*				
Skeletal system development (GO:0001501)		44	2.8×10^{-5}	1.96	0.052
Annotation cluster: Response to wounding	*2.3*				
Response to wounding (GO:0009611)		65	1.4×10^{-5}	1.74	0.025
Annotation cluster: Limb morphogenesis	*2.3*				
Limb development (GO:0060173)		18	8.1×10^{-4}	2.48	1.461
Annotation cluster: Branching morphogenesis	*2.3*				
Branching morphogenesis of a tube (GO:0048754)		15	1.4×10^{-4}	3.28	0.251
Annotation cluster: Bone mineralization	*2.1*				
Regulation of ossification (GO:0030278)		16	3.0×10^{-4}	2.91	0.554
Differentially expressed genes (CaGi_UCSC)					
Annotation cluster: cell migration	*7.0*				
Cell motion (GO:0006928)		65	6.2×10^{-8}	2.03	1.1×10^{-4}
Annotation cluster: Response to stimuli	*6.4*				
Response to hormone stimulus (GO:0009725)		51	1.3×10^{-6}	2.07	2.3×10^{-3}
Annotation cluster: Angiogenesis	*5.6*				
Blood vessel development (GO:0001568)		40	3.9×10^{-7}	2.43	7.1×10^{-4}
Annotation cluster: Control of cell migration	*5.3*				
Regulation of cell motion (GO:0051270)		33	1.7×10^{-6}	2.54	3.1×10^{-3}
Annotation cluster: Response to nutrients	*4.6*				
Response to extracellular stimulus (GO:0009991)		35	4.2×10^{-6}	2.37	7.7×10^{-3}
Annotation cluster: Coagulation	*4.2*				
Wound healing (GO:0042060)		34	4.6×10^{-7}	2.63	8.4×10^{-4}
Annotation cluster: Response to steroid hormones	*3.9*				
Response to steroid hormone (GO:0048545)		29	8.4×10^{-5}	2.26	0.153
Annotation cluster: Lung development	*3.3*				
Tube development (GO:0035295)		40	2.0×10^{-8}	2.70	3.7×10^{-5}
Annotation cluster: Response to oxidative stress	*3.2*				
Response to reactive oxygen species (GO:0000302)		16	1.2×10^{-4}	3.17	0.212
Annotation cluster: Response to wounding	*3.0*				
Response to wounding (GO:0009611)		66	1.5×10^{-6}	1.85	0.002
Annotation cluster: Cell adhesion	*3.0*				
Regulation of cell adhesion (GO: 0030155)		25	1.3×10^{-5}	2.71	0.023
Annotation cluster: Regulation of signal transduction	*2.9*				
Positive regulation of cell communication (GO:0010647)		43	4.3×10^{-5}	1.94	0.078
Annotation cluster: Neuron development	*2.9*				
Neuron differentiation (GO:0030182)		52	8.1×10^{-5}	1.76	0.148
Annotation cluster: sensory organ development	*2.7*				
Sensory organ development (GO:0007423)		35	1.0×10^{-5}	2.3	0.019

Table S8. (continued)

GO categories	Enrichment score	No. genes	*P*-value	Fold enrichment	FDR
Annotation cluster: Cell adhesion	*2.6*				
Cell-substrate adhesion (GO:0031589)		17	7.8×10^{-4}	2.58	1.424
Annotation cluster: Embryonic development	*2.6*				
Embryonic morphogenesis (GO:0048598)		37	7.7×10^{-4}	1.79	1.402
Annotation cluster: Immune response	*2.6*				
Immune effector process (GO:0002252)		27	7.8×10^{-7}	2.99	0.001
Annotation cluster: Response to insulin	*2.4*				
Response to peptide hormone stimulus (GO:0043434)		25	9.2×10^{-5}	2.41	0.169
Annotation cluster: Skeletal muscle development	*2.3*				
Muscle organ development (GO:0007517)		32	3.1×10^{-5}	2.25	0.056
Annotation cluster: Response to vitamin	*2.3*				
Response to vitamin (GO:0033273)		14	3.8×10^{-4}	3.15	0.696

Supplemental references

1 Breitling, R., Armengaud, P., Amtmann, A., and Herzyk, P. (2004) Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett. *573*, 83-92.

2 Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., and Speed, T.P. (2003) Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res *31*, e15.

3 Bayly, C.I., Cieplak, P., Cornell, W., and Kollman, P.A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J. Phys. Chem. *97*, 10269-10280.

4 Breneman, C.M., and Wiberg, K.B. (1990) Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. J. Comp. Chem. *11*, 361-373.

5 Becke, A.D. (1993) Density-functional thermochemistry. III. The role of exact exchange. J. Chem. Phys. *98*, 5648-5652.

6 Lee, C., Yang, W., and Parr, R.G. (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys. Rev. B Condens. Matter *37*, 785-789.

7 Dunning, T.H. Jr. (1989) Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J. Chem. Phys. *90*, 1007-1023.

8 Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Mennucci, B., and Petersson, G.A. (2009) Gaussian 09, Revision D.01. Gaussian, Inc, Wallingford CT.

9 MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B *102*, 3586–3616.

10 Aymé, S. and Philip, N. (1996) Fine-Lubinsky syndrome: a fourth patient with brachycephaly, deafness, cataract, microstomia and mental retardation. Clin. Dysmorphol. *5*, 55-60.

11 Gripp, K.W., Nicholson, L., and Scott, C.I. (1996) Apparently new syndrome of congenital cataracts, sensorineural deafness, Down syndrome-like facial appearance, short stature, and mental retardation. Am. J. Med. Genet. *61*, 382-386.

12 Nakane, T., Mizobe, N., Hayashibe, H., and Nakazawa, S. (2002) A variant of Fine-Lubinsky syndrome: a Japanese boy with profound deafness, cataracts, mental retardation, and brachycephaly without craniosynostosis. Clin. Dysmorphol. *11*, 195-198.

13 Keppler-Noreuil, K., Welch, J., and Baker-Lange, K. (2007) Syndrome of congenital cataracts, sensorineural deafness, Down syndrome-like facial appearance, short stature, and mental retardation: two additional cases. Am. J. Med. Genet. A *143A*, 2581-2587.

14 Holder, A.M., Graham, B.H., Lee, B., and Scott, D.A. (2007) Fine-Lubinsky syndrome: sibling pair suggests possible autosomal recessive inheritance. Am. J. Med. Genet. A *143A*, 2576-2580.

15 Schoner, K., Bald, R., Fritz, B., and Rehder, H. (2008) Fetal manifestation of the Fine-Lubinsky syndrome. Brachycephaly, deafness, cataract, microstomia and mental retardation syndrome complicated by Pierre-Robin anomaly and polyhydramnios. Fetal Diagn. Ther. *23*, 228-232.

16 Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., and Hermans, J. (1981) Interaction models for water in relation to protein hydration. In Intermolecular Forces. The Jerusalem Symposia on Quantum Chemistry and Biochemistry. Vol. 14, B. Pullman, ed. (Dordrecht: Reidel Publishing Company), pp. 331-342.

17 Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004) Bioconductor: Open software development for computational biology and bioinformatics. Genome Biol. 5, R80.

18 Smyth, G.K. (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat. Applicat. Genet. Molec. Biol. 3, article 3.

19 Huang, W., Sherman, B.T., and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57.