DISRUPTION of RAS-MAPK SIGNALLING in HUMAN NEUROCUTANEOUS DISORDERS

(Thesis format: Integrated Article)

by

Laura Marie McDonell

Graduate Program in Biochemistry with specialization in Human and Molecular Genetics

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor in Philosophy

The Faculty of Medicine Graduate and Postdoctoral Studies

The University of Ottawa, Ottawa, Ontario, Canada

Abstract

Ras-MAPK signalling regulates key cellular processes such as proliferation, differentiation and survival. Unsurprisingly, mutations in RAS genes are now recognized as potent oncogenic drivers. However, disruption of this pathway during development is associated with a family of disorders termed the *Rasopathies*. Shared clinical features include cutaneous, neurological and cardiac anomalies. At the outset of this study, the genetic etiology of three neurocutaneous disorders, microcephaly-capillary malformation syndrome (MIC-CAP), encephalocraniocutaneous lipomatosis (ECCL) and PHACE (Posterior fossa malformations, facial Hemangiomas, cerebral Arterial anomalies, Cardiovascular defects and Eye abnormalities) syndrome had not yet been established. This thesis identifies mutations in STAM-binding protein (*STAMBP*) in a cohort of individuals with MIC-CAP syndrome using whole-exome sequencing (WES). This gene encodes a deubiquitinating isopeptidase that regulates cell surface receptor-mediated endocytosis and sorting. Cell lines of individuals with MIC-CAP show reduced STAMBP expression, associated with accumulation of ubiquitinated protein aggregates, increased apoptosis and constitutive activation of the Ras-MAPK and PI3K-AKT pathways. WES also enabled the identification of post-zygotic mutations within the tyrosine kinase domain of fibroblast growth factor receptor 1 (*FGFR1)* in individuals with ECCL. Fibroblasts from affected individuals showed increased phosphorylation of the FGFRs consistent with receptor activation as well as insensitive signal transduction through the Ras-MAPK pathway. Neurocutaneous syndromes can feature striking vascular lesions such as the cerebral vasculopathy and large segmented facials hemangiomas seen in PHACE syndrome. The asymmetric and patchy vascular malformations coupled with a sporadic incidence and absence of familial recurrence suggested that PHACE might be caused by post-zygotic mutations. Interrogation of a discordant sib-pair using copy number analysis and WES did not identify

causative mutations indicating the need for a comprehensive and targeted –omic approach to elucidate the molecular mechanism of this syndrome. Taken together, these findings expand the spectrum of the *Rasopathies* while providing novel pathomechanistic insights into the regulation of cellular proliferation and survival during development.

Acknowledgements

This thesis would not have been possible without the participation of the many patients and families whose stories contribute to the work presented herein. I would like to thank all the clinicians, scientists, laboratory technicians, genetic counsellors and other team members who have, and continue to, support the CareforRare initiative.

I am forever grateful to Dr. Kym Boycott whose leadership, infallible work ethic and unwavering dedication to her field has been both inspiration and motivation. You have provided mentorship, academic challenges and innumerable learning opportunities and have contributed to my growth as a clinician-scientist trainee. I will proudly carry forward the lessons you have taught me. Thank you.

I would like to express my deepest gratitude to Dr. Dennis Bulman, my lab father, who first encouraged me to apply to the MD/PhD program. You provided me with perspective, guidance and steadfast encouragement. Your incredible knowledge of molecular genetics was a wonderful asset throughout my graduate studies.

I am grateful to all our collaborators whose patients, expertise and labour colour these pages. In particular, I would like to thank Dr. Mark O'Driscoll and his lab, whose cell biology work provided invaluable depth to our stories and who was willing to take me in.

I would like to thank every single lab member, past and present, whose names and stories could fill many pages. Thank you for supporting me all of these years. Each of you holds a special place in my heart and I will not forget all of the laughs, challenges, success and popcorn we have shared.

And finally, thank you to my family, friends and loved one for showering me with their unconditional love and endless support. It would not have been possible without you.

Table of Content

Abbreviations

4E-BP1	Eukaryotic translation initiation factor 4E-binding protein 1
AAF	Alternate allele fraction
AB	Acid box
AMSH-LP	AMSH-like protein
AT	Ataxia telangiectasia
ATM	Ataxia-telangiectasia mutated
BAD	B-cell lymphoma-associated agonist of cell death
BAQ	Base alignment quality
BCL2	B-cell lymphoma
BWA	Burrows-Wheeler aligner
CA	Celiac artery
CCDS	Consensus coding sequence
CFC	Cardiofaciocutaneous
CHMP	Charged multivesicular body proteins
CIHR	Canadian Institutes of Health Research
CNS	Central nervous system
CNV	Copy number variations
COSMIC	Catalogue of Somatic Mutations in Cancer
DAPI	4′,6-diamidino-2-phenylindole
dbSNP	NCBI database

DC	Diaphragmatic crura
DNA	Deoxyribonucleic acid
DRTK	Developmental receptor tyrosine kinasopathies
DUB	Deubiquitinating isopeptidase
DUR	Distal ubiquitin recognition site
ECCL	Encephalocraniocutaneous lipomatosis
eIF4E	Eukaryotic translation initiation factor 4E
ERK	Extracellular signal-regulated kinases
ERK1/2	ERK1 and ERK2
ES	Exome sequencing
ESCRT	Endosomal sorting complexes required for transport
EVS	Exome variant server
ExAC	Exome Aggregation Consortium
FGF	Fibroblast growth factor
FGFR1	Fibroblast growth factor receptor 1
FK2	Conjugated ubiquitin
FOXO	Forkhead box O
FRS2	Fibroblast growth factor receptor substrate 2
GAB	GRB2-associated binding protein
GAP	GTPase activating protein

GATK	Genome Analysis Toolkit
GEF	Guanine nucleotide-exchange factor
gnomAD	Genome Aggregation Database
GRB2	Growth-factor-receptor bound protein
Hg	Hedgehog
HGMD	Human Gene Mutation Database
HS	Heparan-sulfate
HSB	Heparan-sulfate binding site
Ig	Immunoglobulin
IGH	Immunoglobulin heavy
IH	Infantile hemangiomas
ILVs	Intralumenal vesicles
Indels	Insertions and deletions
JAK/STAT	Janus kinase/signal transducer and activator of transcription
JAMM	JAB1/MPN/MOV34
LADD	Lacrimo-auriculo-dento-digital syndrome
LC3-II isoform	Autophagosome-associated phosphatidylethanolamine-conjugated microtubule-associated light chain 3
LCL	Lymphoblastoid cell line
LOH	Loss of heterozygosity
LRA	Left renal artery

M-CAP	Megalencephaly-capillary malformation syndrome
MAPK	Mitogen-activated protein kinase
MEK	MAPK/ERK Kinase
MEK1/2	MEK1 and MEK2
MEN2B	Multiple endocrine neoplasia type IIB
MIC-CAP	Microcephaly-capillary malformation syndrome
MIT	Microtubule-interacting and transport
MRI	Magnetic resonance imaging
mTORC1	Mechanistic target of rapamycin complex 1
NCBI	National Center for Biotechnology Information
NF-κB	nuclear factor kappa-light-chain-enhancer of activated B cells
NF1	Neurofibromin 1
NF1/2	NF1 and NF2
NHLBI	National Heart, Lung, and Blood Institute
NLS	Nuclear localization signal
OCCS	Oculocerebrocutaneous syndrome
OES	Oculoectodermal syndrome
OMIM	Online Mendelian Inheritance in Man
PHACE	Posterior fossa malformations, facial Hemangiomas, cerebral Arterial anomalies, Cardiovascular defects and Eye abnormalities
PI3K	Phosphatidylinositol 3-kinase

PIK3CA	PI3K catalytic subunit alpha
PIP_2	Phosphatidylinositol 4,5-biphosphate
PIP_3	Phosphatidylinositol 3,4,5-triphosphate
PKB/AKT	Protein kinase B
PKD1	Phosphoinositide-dependant kinase 1
PLCε	Phospholipase Cε
PNS	Peripheral nervous system
PSI	Percent spliced in
PTB	Phosphotyrosine binding
PTEN	Phosphatase and tensin homolog
PTP	Protein tyrosine phosphatase
PTPN11	Tyrosine-protein phosphatase non-receptor type 11
qRT-PCR	Real Time PCR
RA	Renal arteries
RASA1	p120-RasGAP
RHEB	Ras homolog enriched in the brain
RNA	Ribonucleic acid
RRA	Right renal artery
RSK	Ribosomal S6 kinases
RTK	Receptor tyrosine kinase
S6K1	S6 kinase 1

SBM	SH3 binding motif
SHC	SH-2 containing protein
SHH	Sonic hedgehog
siRNA	Short interfering RNA
SMA	Superior mesenteric artery
smMIP	Single molecule molecular inversion probes
SNV	Single nucleotide variant
SOS	Son of sevenless
STAMBP	STAM-binding protein
TD	Thanatophoric dysplasia
TGF-ß	Transforming growth factor-ß
TK	Tyrosine kinase
TK1	Tyrosine kinase inhibitor
TKD	Tyrosin kinase domain
TM	Transmembrane
TSC1	Tuberous sclerosis complex 1
TSC2	Tuberous sclerosis complex 2
WCE	Whole cell extract
WES	Whole-exome sequencing
Wnt	Wingless related
WT	Wild-type

List of Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 5

Chapter 6

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 6

List of Appendices

Chapter 1

1. General Introduction

Human development is an incredibly dynamic process that spans early gestation and natal life.[1] It entails the timely orchestration of complex molecular and cellular events that support the emergence of new structures and functions. Proteins that participate in complex signalling cascades, and, together with extracellular stimuli, guide cellular differentiation, morphogenesis and organogenesis during the embryonic, fetal and post-natal periods.[1,2] As such, these pathways play an important role during development and throughout life. Their disruption can cause multiple developmental syndromes and contribute to the pathogenesis of many aggressive cancers.[3,4] Mutations in the genes encoding the proteins of the Ras-MAPK pathway are associated with a family of developmental disorders termed the *RASopathies* and share phenotypic features that can include neurological, cardiac, facial and cutaneous abnormalities.[3,4] The work presented herein resolves the genetic etiology of two striking neurocutaneous disorders, microcephaly-capillary malformation syndrome (MIC-CAP) and encephalocraniocutaneous lipomatosis (ECCL) and characterizes altered Ras-MAPK signalling in cell lines of individuals with MIC-CAP and ECCL. In addition, this work devises an approach to solving PHACE (*P*osterior fossa malformations, facial *H*emangiomas, cerebral *A*rterial anomalies, *C*ardiovascular defects and *E*ye abnormalities) syndrome and explores the challenges posed by mosaic *RASopathie*s.

1.1 Ras-MAPK signalling

Cell to cell interactions, mediated by intracellular transduction pathways, are vital for ensuring synchronized developmental program progression.[2] Briefly, ligand binding to transmembrane

receptors recruits intracellular proteins that transmit signals via successive reversible protein changes ultimately altering gene expression.[1,2] Strikingly, only a few central pathways are involved during development and include Hedgehog (Hh), wingless related (Wnt), transforming growth factor-ß (TGF-ß), Notch, Janus kinase/signal transducer and activator of transcription (JAK/STAT), nuclear hormone and receptor tyrosine kinase (RTK) pathways.[1]

Signaling through the RTKs is not linear; rather, each receptor, adaptor, scaffold protein or downstream effector can interact with multiple different signaling molecules which enables cross-talk between pathways and highlights the dynamic complexity of RTK signaling.[5] A notable target of the RTKs is the well-characterized Ras-mitogen-activated protein kinase (MAPK) pathway, which regulates crucial cellular processes such as proliferation, differentiation, survival and cell cycle progression.[6-8]

1.1.1 Ras family

The Ras GTPase superfamily includes over 150 low-molecular-weight GTP-binding/hydrolyzing proteins and is commonly divided into five major families, the Arf, Ran, Rho/Rac, Rab and Ras families.[9] The three highly conserved canonical Ras family isoforms HRAS, KRAS, and NRAS share a highly homologous G domain that binds GDP and GTP.[9] This domain also includes two switches that regulate binding to RAS regulators and effectors.[9] Post-transcriptional modifications of the hypervariable C-terminal domain specify membrane localization.[3]

The Ras GTPases act like molecular switches that couple extracellular signals to intracellular signalling networks by cycling between an inactive GDP-bound and an active GTP-bound state (Figure 1).[6-8,10,11] Broadly, Ras signalling is initiated by ligand binding to cell surface receptors

which activates guanine nucleotide-exchange factors (GEFs) that subsequently displace GDP from GDP-bound Ras to allow passive GTP binding and Ras activation.[12,13] Conversely, Ras proteins are negatively regulated by GTPase activating proteins (GAPs) that stimulate Ras GTPase activity. [1,3,13]

1.1.2 Signal transduction in the Ras pathway

Growth factor binding to cell surface receptors such as the RTKs, initiates an intricate phosphorylation cascade that activates both the receptor and receptor-associated complexes (Figure 1).[14,15] These complexes contain adaptor proteins such as SH-2 containing protein (SHC), growth-factor-receptor bound protein (GRB2) and GRB2-associated binding protein (GAB).[15] These proteins recruit GEFs such as tyrosine-protein phosphatase non-receptor type 11 (PTPN11) and son of sevenless (SOS) that increase intracellular levels of Ras-GTP.[15] The conversion of Ras-GTP to Ras-GDP is facilitated by GAPs such as neurofibromin 1 (NF1).[3] Activated Ras-GTP acts like a central signalling node binding with differential affinity to a plethora of downstream effectors.[3]

1.1.3 Downstream effectors of Ras-MAPK signalling

Activated Ras-GTP can interact with over 20 downstream effectors that regulate essential cellular responses such as proliferation, differentiation and survival.[3] Notably, Ras activates Raf, phosphatidylinositol 3-kinase (PI3K), the enzyme phospholipase Cε (PLCε) and several Rac and Ral exchange factors.[12,16] The well-characterized Ras-MAPK cascade is composed of three core tiers, MAP3K, MAP2K and MAPK.[14] Signalling is transmitted through the cascade when one or more kinases in a given tier phosphorylates and activates components in the succeeding tier until regulatory target proteins, such as transcription factors, are phosphorylated and initiate required cellular processes.[12]

For the purpose of this work, only the basis of the ERK cascade will be explored, although other distinct MAPK signalling cascades are recognized (JNK, p38, BMK) and reviewed elsewhere.[17-20] The three MAP3K tier Raf kinases (BRAF, RAF1 and ARAF), once activated, phosphorylate and activate MAPKK tier proteins MEK1/2 and they, in turn, phosphorylate and activate the MAPK tier kinases ERK1/2 (Figure 1).[3,14] The localization of both active and inactive ERK kinases strongly relies on their interaction with regulatory proteins and, as such, most activated ERKs translocate to the nucleus while a subset migrates to distinct cellular compartments and accomplishes location-specific functions.[20,21] Notable substrates phosphorylated by the ERKs include Elk, c-Fos, p53 and c-Myc as well as other kinases such as 90 kDa ribosomal S6 kinases (RSKs).[20] Once activated, RSKs phosphorylate cytosolic and nuclear substrates directly contributing to the regulation of transcription, translation, survival and cell-cycle progression.[22]

ERKs also phosphorylate tuberous sclerosis complex 2 (TSC2), which disrupts the GAP activity of the TSC1/TSC2 complex and increases the amount of GTP-bound Ras homolog enriched in the brain (RHEB).[23,24] Activated RHEB directly interacts with the mechanistic target of rapamycin complex 1 (mTORC1) to stimulate its kinase activity resulting in the phosphorylation of a large number of functionally diverse substrates such as the eukaryotic translation initiation factor 4E (eIF4E)-binding protein 1 (4E-BP1) and S6 kinase 1 (S6K1), both translational regulators.[24] In addition to promoting protein synthesis, mTORC1 can control lipid synthesis, positively regulate cellular metabolism and negatively regulate autophagy and lysosome biogenesis.[24] The TSC1/TSC2 complex integrates many upstream signals that converge on mTORC1.[24] For example, TSC2 is phosphorylated not only by ERK, but also by protein kinase B (PKB, also known as AKT) and by stress response mediators.[23,24]

AKT is an effector of class I PI3K signalling activated by phosphoinositide-dependant kinase 1 (PDK1) at the plasma membrane following phosphatidylinositol 3,4,5-triphosphate (PIP_3) recruitment and binding (Figure 1).[25] PIP_3 is generated from phosphatidylinositol 4,5-biphosphate (PIP_2) in a reaction catalyzed by activated PI3K and negatively regulated by the phosphatase and tensin homolog (PTEN).[26] Class I PI3Ks are heterodimeric and consist of a catalytic (p110α, p110β or p110δ) subunit tightly bound to a regulatory subunit (p85α, p85β, p55α, p55γ or p50α).[27] Signal transduction through class I PI3K can be initiated by activated RTKs, G protein-coupled receptors and activated RAS.[28,29]

The effectors of both Ras-MAPK and PI3K-AKT signalling often converge and act on the same substrates.[23] For instance, activated ERK, RSK, AKT and S6K act as switches to regulate the expression of cell cycle regulators in addition to apoptotic, pro-survival and growth genes.[30] These substrates include forkhead box O (FOXO) and c-Myc transcription factors, B-cell lymphoma (BCL2)-associated agonist of cell death (BAD) as well as GSK3.[23,30] Together, Ras-MAPK and PI3K-AKT govern fundamental cellular mechanisms such as metabolism, proliferation, differentiation and survival.[23,30] With such diverse roles, it is not surprising that signalling alterations have profound impact on cell-fate decisions and as such, must be tightly regulated.[3,28,31]

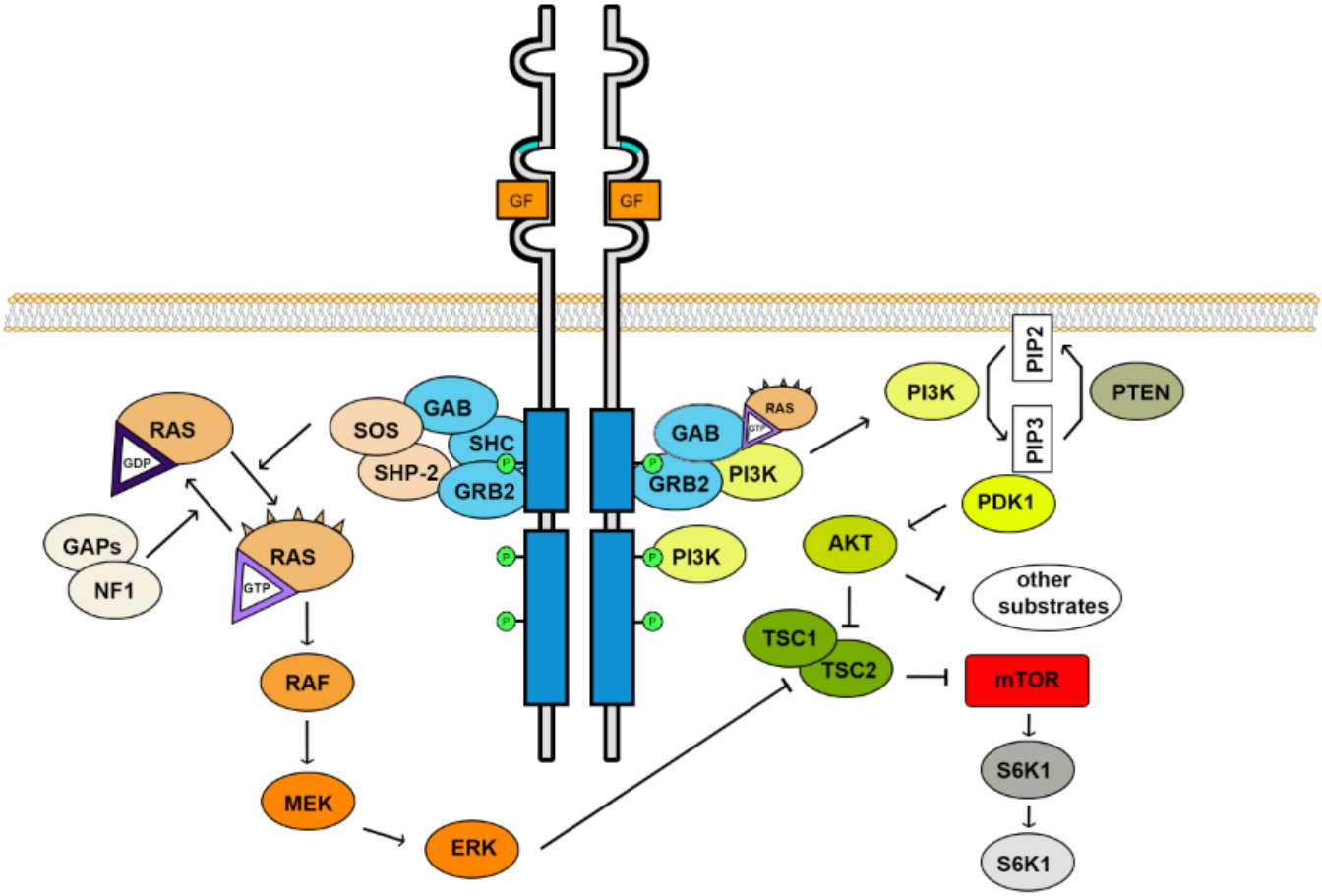


Figure 1. The interconnected Ras-MAPK and PI3K-AKT signalling pathways. Extracellular growth factor (GF) binding to cell-surface receptors results in receptor autophosphorylation (P) and activation. Receptor associated complexes such as SHC, GRB2 and GAB are then activated and recruit SHP2 and SOS. Guanine nucleotide-exchange factors such as SOS, displace GDP from GDP-bound Ras to allow GTP binding and Ras activation. This is negatively regulated by GTPase activating proteins (GAPs) such as NF1. Active Ras can trigger a phosphorylation cascade through the MAPK/ERK network (RAF, MEK, ERK). Ras can also activate the PI3K-AKT pathway. In addition, class I PI3Ks can be activated directly at cell surface receptors, in association with adaptor molecules such as GRB2 and GAB or by G-coupled proteins. PIP_3 generated by PI3K activates PDK1 and AKT. The effectors of these two signalling pathways converge and can act on the same substrates such as mTOR.

1.1.4 Ras-MAPK signalling in oncogenesis

Deregulated cellular growth, differentiation, survival and migration are hallmarks of oncogenesis.[32] Given the involvement of both Ras-MAPK and PI3K-AKT signalling in regulating these cellular processes, it is not surprising that mutations in components of these pathways are associated with cancer.[3,31] In fact, missense gain-of-function mutations in Ras isoforms are found in 27% of human cancers.[33] Strikingly, though KRAS, HRAS and NRAS are widely expressed in adult tissues and tumours, KRAS is the most mutated Ras family member (85%).[3,33] *KRAS* mutations predominate in pancreatic, colorectal, small intestine, biliary tract, endometrial and lung cancers whereas *NRAS* and *HRAS* mutations are prevalent in melanoma and urinary tract cancer, respectively.[3]

In addition to the oncogenic transformation of Ras genes, activation of Ras signalling can occur following loss of GAP proteins, hyperactivation of RTKs and by mutations of any downstream effector.[3] Similarly, alterations in the PI3K signalling network confer cell-specific oncogenic properties and are believed to contribute to cancer-promoting aspects of the tumour environment such as angiogenesis and inflammatory cell recruitement.[34] Mutations in this pathway are common and in fact, the PI3K catalytic subunit alpha (*PIK3CA*) encoding p110α and *PTEN* are among the most mutated human oncogene and tumour suppressor gene, respectively.[34] A testament to the extensive crosstalk between signalling networks, tumours expressing normal PI3K and PTEN may still benefit from hyperactive PI3K signalling through activating mutations in RTKs, AKT, and the Ras network.[35] Interestingly, many genes and mutations driving oncogenic transformation in human cancer are also associated with human developmental disorders.[4]

1.1.5 Ras-MAPK signalling in development and post-natal life

Germline and somatic mutations disrupting Ras-MAPK signalling have been identified as the cause of a number of congenital malformation syndromes termed the *RASopathies*.[4] These mutations encode novel RTKs, protein tyrosine phosphatases (PTPs), GAPs, GEFs, GTPases and MAPKs.[3] With the exception of TSC1/2, NF1, p120-RasGAP (RASA1) and other GAPs, these disorders are mostly caused by gain-of-function alleles.[3] Nonetheless, inactivating mutations in GAPs can still lead to hyperactive downstream signalling as loss of GAP activity hinders GTPase inhibition.[13] Characteristic facial, cutaneous, cardiac and neurological abnormalities define the *RASopathies* such that the term "neuro-cardio-facial-cutaneous" syndromes has also been proposed.[36]

Mutations in the PI3K-AKT axis typically result in increased signalling and lead to disorders exhibiting segmental overgrowth.[36] Though certain disorders such as Cowden are autosomal dominant, many result from post-zygotic somatic mosaicism.[37,38] As such, this family of disorders typically presents with clinical features that are striking both in appearance and distribution.[38] Of note, this signalling cascade is also targeted by mutations leading to impaired signalling as exemplified by SHORT syndrome, caused by inactivating mutations in *PIK3R1*.[39] In contrast to the overgrowth commonly seen in disorders caused by increased signalling through PI3K-AKT, individuals with SHORT syndrome typically have short stature, hyperextensibility, lipodystrophy and delayed dentition.[39] A selection of developmental syndromes associated with alterations of Ras-MAPK and PI3K-AKT signalling is summarized in Table 1.

Interestingly, individuals with mutations in specific components of the Ras-MAPK and PI3K-AKT network are more likely to develop cancer.[4] For instance, individuals with Noonan have a higher predisposition for juvenile myelomonocytic leukemia whereas patients with Cowden have an

increased risk of breast, thyroid and endometrial cancer.[3] Likewise, individuals with neurofibromatosis have an increased incidence of neurofibromas, astrocytomas and pheochromocytomas.[3] Fascinatingly, cardiofaciocutaneous (CFC) syndrome, commonly caused by activating mutations in *BRAF*, is not associated with cancer even though half of all melanomas harbour activating *BRAF* mutations.[3]

Phenotypic heterogeneity is widely observed in disorders of Ras-MAPK and PI3K signalling.[3] This occurs when the same gene causes multiple different disorders. For example, heterozygous mutations in *BRAF* cause LEOPARD, Noonan and CFC syndrome whereas mutations in *AKT1* cause Proteus and Cowden syndrome.[37,40] Remarkably, genetic alteration of different genes can also cause the same disorder. For instance, mutations in *PTEN*, *PIK3CA* and *AKT1* cause Cowden syndrome, whereas mutations in *PTPN11*, *SOS1/2*, *KRAS*, *NRAS*, *RAF1*, and *BRAF* cause Noonan syndrome.[3,37,40,41]

Overall, the range and extent of clinical disease resulting from mutations in the Ras-MAPK and PI3K pathways highlights the complex equilibrium of many factors including the position and nature of the mutation, the function of the given protein, and the mutation's impact on the intensity and duration of signalling, all in the context of the individual's genetic background. Given the extensive and complex regulatory circuits for Ras and PI3K signalling, it is not surprising that many molecules within a given pathway lead to similar clinical presentations.

Table 1. Selection of developmental disorders associated with Ras-MAPK and PI3K-AKT signalling *

Disorder	Causative gene	Inheritance	Mechanism
Encephalocraniocutaneous lipomatosis	*FGFR1, KRAS*	somatic	GOF
Microcephaly-capillary malformation syndrome	*STAMBP*	AR	LOF
Leopard	*PTPN11, BRAF, RAF1*	AD	GOF
Noonan	*PTPN11, SOS1, SOS2, KRAS, NRAS RAF1, BRAF*	AD	GOF
Sturge-Webber	*GNAQ*	somatic	GOF
Neurofibromatosis	*NF1, NF2*	AD	LOF
Capillary malformation arteriovenous malformation syndrome	*RASA1*	AD	LOF
Costello	*HRAS*	AD, sporadic	GOF
Occuloectodermal syndrome	*KRAS*	somatic	GOF
Cardiofaciocutaneous syndrome	*KRAS, BRAF, MEK1, MEK2*	AD, sporadic	GOF
Tuberous sclerosis	*TSC1, TSC2*	AD	LOF
Cowden	*PIK3CA, PTEN, AKT1*	AD, sporadic	GOF, LOF
Proteus	*AKT1*	somatic	GOF
Congenital lipomatous overgrowth, vascular malformations, and epidermal nevi	*PIK3CA*	somatic	GOF
Megalencephaly-capillary malformation-polymicrogyria syndrome	*PIK3CA*	somatic	GOF
Megalencephaly-polymicrogyria-polydactyly-hydrocephalus syndrome	*PIK3R2, AKT3*	AD	GOF

* Data extracted from OMIM (https://www.omim.org, November 2017)

1.2 Neurocutaneous disorders

Neurocutaneous disorders are a heterogeneous collection of congenital syndromes with primary involvement of the skin and of the central nervous system (CNS).[42] They represent a phenotypically striking and diverse group of disorders characterized by developmental abnormalities of ectodermal and at times, mesodermal origin.[43] Most are classified as single-gene disorders with autosomal dominant, recessive or X-linked inheritance patterns.[42] However, in keeping with the patchy distribution of anomalies often seen in these patients, an increasing number of post-zygotic/mosaic conditions have now been recognized.[44] In individuals with neurocutaneous disorders, the cutaneous manifestations are often present at birth and act as a diagnostic tool that can prompt an early workup and diagnosis. [43]

1.2.1 Cutaneous manifestations

The constellation of skin features seen in neurocutaneous syndromes can sometimes be explained by the embryonic origin of the affected cell type.[45] During early development, neural crest cells emerge from the embryonic ectoderm at the dorsal neural plate and migrate within the embryo where they differentiate into multiple different cell types including neuroendocrine cells, melanocytes, cells of the peripheral nervous system (PNS) and vascular smooth muscle cells of the cardiac outflow track.[46] In addition, cephalic neural crest cells can give rise to the ectomesenchyme, which in vertebrates, contributes to building the craniofacial skeleton and it's associated tendons, dermis, adipose tissue as well as glandular and muscular connective tissue.[47] Mesenchymal tissue in the remaining embryo derives from the mesoderm.[47] With this understanding, skin lesions in the neurocutaneous syndromes can be classified as either pigmentary, vascular, adnexal, hyperkeratotic, connective tissue, photosensitive or inflammatory.[45]

Examples of hyperpigmented lesions of melanocytic origin include café-au-lait spots, lentigines and melanocytic nevi seen in NF1/2, Leopard syndrome and CFC syndrome, respectively.[45] Peripheral nerve sheath tumours such as neurofibromas, commonly associated with NF1, can present as flesh coloured, pink or hyperpigmented papulonodules.[45] Vascular lesions such as telangiectasia, capillary malformations, capillary venous malformations and hemangiomas can be appreciated in neurocutaneous syndromes.[45] These can be subtle such as the telangiectasia of the ears, cheeks and upper trunk seen in ataxia telangiectasia (AT) or they can be striking such as the large segmented facial hemangiomas associated with PHACEs.[45,48] Ichthyosis, characterized by dry hyperkeratotic, scaly skin is associated with multiple disorders including X-linked recessive ichthyosis and Sjögren-Larson syndrome.[45] Finally, the presentation of connective tissue anomalies in neurocutaneous disorders can be quite varied. For instance, we must consider the large lipomas of ECCL, the rhabdomyomatous mesenchymal hamartomas associated with oculocerebrocutaneous (OCCS) syndrome or the lax acral skin seen in cutis laxa and Costello syndrome.[45,49,50]

The examples listed above represent a small selection of skin anomalies seen in neurocutaneous syndromes and attempts to highlight the phenotypic variability of these conditions. It is important to note that many of these disorders present with multiple histologically distinct cutaneous lesions, as will be highlighted in this thesis, testament to the pluripotent potential of the embryonic progenitor cells.

1.2.2 Neurological involvement

The skin and the CNS share the same ectodermal beginnings and as such, cutaneous anomalies often act as a diagnostic window to the CNS.[51] In keeping with the cutaneous phenotypic variability, there is a large spectrum of CNS anomalies associated with neurocutaneous disorders.[51]

For many syndromes, anomalies of the CNS can represent internal manifestations of externally visible lesions.[45,52] Such is the case for familial cerebral cavernous malformation syndrome where both skin and cerebral capillary malformations can be found.[45,52] Unlike their typically benign cutaneous counterparts, vascular lesions of the CNS may lead to neurological dysfunction as demonstrated by the progressive moyamoya-like vasculopathy of PHACE.[53] Tumours of the CNS are not uncommon in neurocutaneous syndromes and are enriched in disorders involving genetic alterations in oncogenic pathways.[54] For instance, patients with NF2 are predisposed to a number of tumours including meningiomas, vestibular schwannomas and spinal cord ependymomas.[45] Whereas patients with von Hippel Lindau syndrome may develop hemangioblastomas of the brain, spinal cord and retina in addition to multiple other non-CNS malignancies.[55]

Malformations of cerebral cortical development can be appreciated in some neurocutaneous conditions.[56] These can represent malformations secondary to abnormal neuronal and glial proliferation or apoptosis.[56] Examples would include the progressive cortical atrophy of MIC-CAP and the progressive brain growth of megalencephaly-capillary malformation (M-CAP) syndrome.[56] Malformations of abnormal neuronal migration or postmigrational development are also identified.[56] In fact, both nodular heterotopia and polymicrogyra are commonly identified in individuals with OCCS.[57]

CNS involvement can manifest with neurologic complications such as epilepsy, neurocognitive dysfunction and behavioural challenges.[43,51] For instance, 80-90% of individuals with tuberous sclerosis will have seizures and 70% of them will have their first episode within their first year of life.[58] In these patients, early-onset epilepsy and refractory epilepsy is correlated with poor

cognitive and behavioural outcomes.[59] These complications are not limited to tuberous sclerosis as they are seen in many other neurocutaneous syndromes including but not limited to NF1, Noonan syndrome, Leopard syndrome, CFC, OCCS, MIC-CAP and Sturge-Weber syndrome.[45]

It is important to recognize that developmentally as well as anatomically, the retina and axons of the optic nerve are extensions of the CNS.[60] Unsurprisingly then, ocular manifestations are quite common and may provide diagnostic clues as well as potential sources of morbidity.[60] As such, there must be appropriate investigation and management of ocular complications in these individuals. In addition to CNS manifestations, there may be involvement of the PNS and several neurocutaneous syndromes with peripheral neuropathy either exclusively or in conjunction with a CNS phenotype are recognized.[61] There is a large spectrum of CNS and PNS manifestations identified in individuals with neurocutaneous disorders and it is the specific constellation of neurological features that may inform the diagnosis in a given individual and guide appropriate genetic testing. [61]

1.2.3 Contributing pathways

As described above, a large subset of neurocutaneous disorders are caused by alterations to genes in the interconnected Ras-MAPK and PI3K-AKT signalling cascades.[3,37,38] Within these networks, a selection of genes encodes proteins that regulate distinct and essential cellular functions.[3,37,38] Alterations of ataxia-telangiectasia mutated (ATM), a protein kinase encoded by a member of the PI3K family, *ATM,* are associated with AT and result in an impaired response to double stranded DNA breaks and defective cell-cycle checkpoints.[62,63] It is unsurprising given the association between genomic instability and cancer that individuals with mutations in DNA damage response genes would have a predisposition to cancer and in fact, 10% of individuals with AT develop cancer

mostly lymphomas and leukemias.[62,64,65] Nijmegen breakage syndrome, associated with mutations in *Nibrin* also encodes another double stranded DNA repair protein.[66]

Neurocutaneous conditions can also be caused by disruption to other signalling networks such as the prototypical inflammatory nuclear factor kappa-light-chain-enhancer of activated B cells pathway (NF-κB) or the sonic hedgehog (SHH) signaling pathway.[67-69] Incontinentia pigmenti, an X-linked dominant syndrome that typically presents with antenatal male lethality is associated with a complex rearrangement in *IKBKG* in 85% of affected individuals that results in absent IKBKG protein and absent NF-κB activation.[68,70] Mutations in this gene as well as other genes are also associated with hypohidrotic/anhidrotic ectodermal dysplasias.[68] Moreover, dominant mutations in *PTCH1* or *SUFU* encoding proteins of the SHH cascade cause nevoid basal cell carcinoma syndrome and affected individuals carry an elevated life-time risk of developing medulloblastomas.[67,69]

Neurocutaneous manifestations can arise secondary to metabolic disease as seen in lysosomal storage disorders, peroxisomal disorders, carboxylase deficiency, cystathionine ß-synthase deficiency and copper deficiency. These disorders are not thought to be purely developmental and as such, are classified as secondary neurocutaneous syndromes.[61]

1.2.4 Mosaic neurocutaneous disorders

In mammals, the unicellular zygote represents the first and potentially last time a homogeneous genetic signature may be found in a developing organism.[71] Any alteration to this original sequence, be it genomic or epigenetic, will produce diverting cell lines and lead to mosaic states.[71] Unsurprisingly, every human is, to a certain extent, a mosaic.[72] Whether or not these *de novo* events

lead to an obvious clinical phenotype will depend on the type of genomic or epigenetic alteration and its overall impact on a given cellular lineage at a specific time point.[73]

Mosaic disorders caused by post-zygotic events in the affected individual can be categorized into two distinct clinical categories.[73] First, they can be mosaic presentations of known Mendelian disorders (Table 1; somatic mutations).[73] For instance, post-zygotic de novo mutations in *NF1* are identified in individuals with manifestations of NF1 limited to specific segments of their bodies.[74] Secondly, mosaic disorders can be caused by mutations only seen in mosaic states and this is likely explained by the fact that these mutations, if constitutional, are embryonically lethal.[73] This class was initially based on the existence of sporadic disorders without familial recurrence where patients presented with distinctly patterned skin lesions.[73] This was further supported by the recognition of discordant monozygotic twins.[73]

Any cell, tissue or organ system has the potential to exist in a mosaic state. The skin, being a visible organ, is where mosaicism can most easily be noticed. To date, five archetypical patterns of cutaneous mosaicism have been identified including narrow lines of Blaschko, broad lines of Blaschko lines, patchy without midline separation as well as the "checkerboard" and "phylloid" patterns. Of note, other non-classified patterns are also recognized. Indeed, identification of such a pattern in an individual may represent the first clue of an underlying mosaic etiology and should prompt careful consideration of the approach used for detecting the causative mosaic alteration.[73]

Awareness of the common molecular classes of mosaicism including large scale chromosomal abnormalities, copy number variants, small insertions/deletions and point mutations may direct the employed detection technology ranging from cytogenetics to microarrays and finally, to high-

throughput sequencing. Irrespective of the favoured detection method, tissue type considerations will remain, as mosaicism may be tissue-restricted.[73]

1.3 High-throughput sequencing

The advent of high-throughput sequencing technologies revolutionized the identification of genes underlying rare genetic disorders.[75] Prior to the arrival of these sequencing methodologies, solving rare disorders relied on large multi-generational families with many affected individuals and employed genetic linkage, positional cloning and sequential sequencing strategies.[75] Though these studies lead to landmark discoveries such as the identification of the cystic fibrosis transmembrane conductance regulator gene as the cause of cystic fibrosis, they were labour intensive and time consuming.[75,76] It became evident that solving novel genetic disorders affecting a limited number of individuals would require technologies able to analyze DNA on a larger scale, at a greater resolution and a faster rate.[75]

1.3.1 Clinical application of whole-exome sequencing

Whole-exome sequencing (WES) is a high-throughput sequencing approach enabling the capture and sequencing of the protein-coding regions of the genome.[77,78] It provides reproducible high depth of sequence coverage for most protein coding genes accounting for approximately 1% of the total human genome.[77,78] To date, WES has shown itself to be a cost-effective technology generating manageable data sets and quick turnaround times.[75] Indeed, WES has enabled the elucidation of an unprecedented and at least initially, exponential, number of rare genetic disorders.[75]

This technology is not without limitations including its inability to interrogate intergenic variants important for transcriptional regulation and splicing as well as its limited ability to comprehensively represent structural variants. Furthermore, by its capture-based design, selection of interrogated regions is guided by the current understanding of the genome, and as such, genomic sequences not currently recognized as gene-encoding can be missed. Despite these limitations, WES remains a great tool for identifying novel genes. By delivering low-hanging fruit in a time and cost-effective manner, it can be used for triage: WES first and then whole genome sequence if the former is unrevealing. [79]

1.4 Thesis Overview

The overarching objective of this study was to identify and characterize novel RAS-MAPK genes involved in congenital neurocutaneous disorders using high throughput sequencing.

1.4.1 Rationale and Hypothesis

Disruption of Ras-MAPK signalling network during development is associated with a family of neurocutaneous disorders termed the *Rasopathies.* Clinically, they are recognized by their spectrum of skin and CNS manifestations in addition to other commonly associated ocular, and cardiac anomalies. Though great strides have been made in identifying the genetic etiology of neurocutaneous syndromes, not all have been solved. Genetic insight into neurocutaneous syndromes provides an opportunity to delineate the phenotypic and genotypic spectrum of these conditions while gaining an understanding for the complex signalling networks involved during development and post-natal life. I therefore hypothesized that novel *Rasopathies* remain to be identified and that discovery of the respective causative genes of the Ras-MAPK network would facilitate study of the underlying disease biology.

1.4.2 Chapter 2: Microcephaly-capillary malformation syndrome

This study describes MIC-CAP, a novel neurocutaneous disorder, characterized by severe microcephaly with progressive cortical atrophy, intractable epilepsy, profound developmental delay and multiple small capillary malformations on the skin. The DNA of five unrelated patients was sequenced using WES and recessive mutations in the STAM-binding protein (*STAMBP*) gene, encoding a deubiquitinating isopeptidase (DUB), were identified. Patient cell lines showed reduced or absent expression of STAMBP and elevated levels of ubiquitin-conjugated protein aggregates. In keeping with the clinical findings, the cell lines also showed elevated apoptosis and insensitive

activation of the Ras-MAPK and PI3K-AKT-mTOR pathways. This was the first description of a congenital human disorder caused by a defective DUB and implicated ubiquitin-conjugate aggregation and elevated apoptosis as factors potentially influencing the progressive neuronal loss in MIC-CAP patients.

1.4.3 Chapter 3: Encephalocraniocutaneous lipomatosis

ECCL is sporadic condition characterized by ocular, cutaneous, and central nervous system anomalies. Patients typically present with a hairless fatty nevus on the scalp, benign ocular tumours, and CNS lipomas. Epilepsy, spasticity, and neurocognitive deficits may also be present. ECCL has been hypothesized to be caused by post-zygotic somatic mutations given the patchy malformations, absence of familial recurrence and normal sex ratio.

1.4.3.1 Encephalocraniocutaneous lipomatosis and *FGFR1*

We set out to use WES sequencing of DNA from multiple affected tissues from five unrelated individuals with ECCL. We identified two mosaic mutations within the tyrosine kinase domain of *FGFR1*, in two affected individuals each. Targeted resequencing of *FGFR1* in multiple tissues from an independent cohort of individuals with ECCL identified an additional *FGFR1* mutation. Interestingly, sequencing the pilocytic astrocytoma sample of one affected individual revealed a second missense substitution in *FGFR1*. Functional studies of patient-derived fibroblast cell lines showed increased levels of phosphorylated FGFRs and its direct substrate, FRS2, in addition to activation of Ras-MAPK signalling. Overall, these findings identified the molecular etiology of ECCL while shedding light on the overlap between mosaic neurocutaneous disorders and oncogenesis.

1.4.3.2 Encephalocraniocutaneous lipomatosis and *KRAS*

This study reports an individual with ECCL and diaphragmatic crura hypertrophy causing compression of the right renal, celiac and superior mesenteric arteries. High throughput sequencing using a clinical gene panel identified a somatic missense mutation in *KRAS* in affected tissue thus expanding the clinical phenotype of mosaic *Rasopathies*.

1.4.4 Chapter 4: PHACE syndrome

PHACE syndrome is a striking neurocutaneous disorder characterized by neurologic, arterial, cardiac, ocular, and sternal anomalies associated with segmental infantile hemangiomas (IH). The phenotypic overlap between PHACE and disorders of the Ras-MAPK network suggest that it may indeed belong to the *Rasopathy* family. Copy number variation analysis and WES discordant monozygotic twins was unrevealing highlighting the complexities faced when undertaking the investigation of hypothesized mosaic syndromes.

1.5 References

1 Pires-daSilva, A. & Sommer, R. J. The evolution of signalling pathways in animal development. *Nat. Rev. Genet.* **4**, 39-49, (2003).
2 Gerhart, J. 1998 Warkany lecture: signaling pathways in development. *Teratology* **60**, 226-239, (1999).
3 Schubbert, S., Shannon, K. & Bollag, G. Hyperactive Ras in developmental disorders and cancer. *Nat. Rev. Cancer* **7**, 295-308, (2007).
4 Bellacosa, A. Developmental disease and cancer: biological and clinical overlaps. *Am. J. Med. Genet. A* **161A**, 2788-2796, (2013).
5 Lemmon, M. A. & Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117-1134, (2010).
6 Vetter, I. R. & Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299-1304, (2001).
7 Donovan, S., Shannon, K. M. & Bollag, G. GTPase activating proteins: critical regulators of intracellular signaling. *Biochim. Biophys. Acta* **1602**, 23-45, (2002).
8 Boguski, M. S. & McCormick, F. Proteins regulating Ras and its relatives. *Nature* **366**, 643-654, (1993).
9 Castellano, E. & Santos, E. Functional specificity of ras isoforms: so similar but so different. *Genes Cancer* **2**, 216-231, (2011).
10 Bourne, H. R., Sanders, D. A. & McCormick, F. The GTPase superfamily: a conserved switch for diverse cell functions. *Nature* **348**, 125-132, (1990).
11 Bourne, H. R., Sanders, D. A. & McCormick, F. The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349**, 117-127, (1991).
12 Mitin, N., Rossman, K. L. & Der, C. J. Signaling interplay in Ras superfamily function. *Curr. Biol.* **15**, R563-574, (2005).
13 Bos, J. L., Rehmann, H. & Wittinghofer, A. GEFs and GAPs: critical elements in the control of small G proteins. *Cell* **129**, 865-877, (2007).
14 Shaul, Y. D. & Seger, R. The MEK/ERK cascade: from signaling specificity to diverse functions. *Biochim. Biophys. Acta* **1773**, 1213-1226, (2007).
15 Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **103**, 211-225, (2000).
16 Patel, M. & Cote, J. F. Ras GTPases' interaction with effector domains: Breaking the families' barrier. *Commun. Integr. Biol.* **6**, e24298, (2013).
17 Rouse, J. *et al.* A novel kinase cascade triggered by stress and heat shock that stimulates MAPKAP kinase-2 and phosphorylation of the small heat shock proteins. *Cell* **78**, 1027-1037, (1994).
18 Lee, J. D., Ulevitch, R. J. & Han, J. Primary structure of BMK1: a new mammalian map kinase. *Biochem. Biophys. Res. Commun.* **213**, 715-724, (1995).
19 Kyriakis, J. M. *et al.* The stress-activated protein kinase subfamily of c-Jun kinases. *Nature* **369**, 156-160, (1994).
20 Plotnikov, A., Zehorai, E., Procaccia, S. & Seger, R. The MAPK cascades: signaling components, nuclear roles and mechanisms of nuclear translocation. *Biochim. Biophys. Acta* **1813**, 1619-1633, (2011).
21 Wortzel, I. & Seger, R. The ERK Cascade: Distinct Functions within Various Subcellular Organelles. *Genes Cancer* **2**, 195-209, (2011).

22 Anjum, R. & Blenis, J. The RSK family of kinases: emerging roles in cellular signalling. *Nat. Rev. Mol. Cell Biol.* **9**, 747-758, (2008).
23 Mendoza, M. C., Er, E. E. & Blenis, J. The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. *Trends Biochem. Sci.* **36**, 320-328, (2011).
24 Laplante, M. & Sabatini, D. M. mTOR signaling in growth control and disease. *Cell* **149**, 274-293, (2012).
25 Wick, M. J., Dong, L. Q., Riojas, R. A., Ramos, F. J. & Liu, F. Mechanism of phosphorylation of protein kinase B/AKT by a constitutively active 3-phosphoinositide-dependent protein kinase-1. *J. Biol. Chem.* **275**, 40400-40406, (2000).
26 Maehama, T. & Dixon, J. E. The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. *J. Biol. Chem.* **273**, 13375-13378, (1998).
27 Vanhaesebroeck, B., Ali, K., Bilancio, A., Geering, B. & Foukas, L. C. Signalling by PI3K isoforms: insights from gene-targeted mice. *Trends Biochem. Sci.* **30**, 194-204, (2005).
28 Martini, M., De Santis, M. C., Braccini, L., Gulluni, F. & Hirsch, E. PI3K/AKT signaling pathway and cancer: an updated review. *Ann. Med.* **46**, 372-383, (2014).
29 Vadas, O., Burke, J. E., Zhang, X., Berndt, A. & Williams, R. L. Structural basis for activation and inhibition of class I phosphoinositide 3-kinases. *Science signaling* **4**, re2, (2011).
30 Castellano, E. & Downward, J. RAS Interaction with PI3K: More Than Just Another Effector Pathway. *Genes Cancer* **2**, 261-274, (2011).
31 Cully, M., You, H., Levine, A. J. & Mak, T. W. Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis. *Nat. Rev. Cancer* **6**, 184-192, (2006).
32 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, (2011).
33 Hobbs, G. A., Der, C. J. & Rossman, K. L. RAS isoforms and mutations in cancer at a glance. *J. Cell Sci.* **129**, 1287-1292, (2016).
34 Fruman, D. A. & Rommel, C. PI3K and cancer: lessons, challenges and opportunities. *Nat. Rev. Drug Discov.* **13**, 140-156, (2014).
35 Engelman, J. A. Targeting PI3K signalling in cancer: opportunities, challenges and limitations. *Nat. Rev. Cancer* **9**, 550-562, (2009).
36 Bentires-Alj, M., Kontaridis, M. I. & Neel, B. G. Stops along the RAS pathway in human genetic disease. *Nat. Med.* **12**, 283-285, (2006).
37 Vahidnezhad, H., Youssefian, L. & Uitto, J. Molecular Genetics of the PI3K-AKT-mTOR Pathway in Genodermatoses: Diagnostic Implications and Treatment Opportunities. *J. Invest. Dermatol.* **136**, 15-23, (2016).
38 Keppler-Noreuil, K. M., Parker, V. E., Darling, T. N. & Martinez-Agosto, J. A. Somatic overgrowth disorders of the PI3K/AKT/mTOR pathway & therapeutic strategies. *Am. J. Med. Genet. C Semin. Med. Genet.* **172**, 402-421, (2016).
39 Dyment, D. A. *et al.* Mutations in PIK3R1 cause SHORT syndrome. *Am. J. Hum. Genet.* **93**, 158-166, (2013).
40 Sarkozy, A. *et al.* Germline BRAF mutations in Noonan, LEOPARD, and cardiofaciocutaneous syndromes: molecular diversity and associated phenotypic spectrum. *Hum. Mutat.* **30**, 695-702, (2009).
41 Pandit, B. *et al.* Gain-of-function RAF1 mutations cause Noonan and LEOPARD syndromes with hypertrophic cardiomyopathy. *Nat. Genet.* **39**, 1007-1012, (2007).
42 Rosser, T. Neurocutaneous Disorders. *Continuum (Minneap Minn)* **24**, 96-129, (2018).

43 Klar, N., Cohen, B. & Lin, D. D. M. Neurocutaneous syndromes. *Handb. Clin. Neurol.* **135**, 565-589, (2016).
44 Ruggieri, M. & Pratico, A. D. Mosaic Neurocutaneous Disorders and Their Causes. *Semin. Pediatr. Neurol.* **22**, 207-233, (2015).
45 Chernoff, K. A. & Schaffer, J. V. Cutaneous and ocular manifestations of neurocutaneous syndromes. *Clin. Dermatol.* **34**, 183-204, (2016).
46 Liu, J. A. & Cheung, M. Neural crest stem cells and their potential therapeutic applications. *Dev. Biol.* **419**, 199-216, (2016).
47 Dupin, E. & Sommer, L. Neural crest progenitors and stem cells: from early development to adulthood. *Dev. Biol.* **366**, 83-95, (2012).
48 Frieden, I. J., Reese, V. & Cohen, D. PHACE syndrome. The association of posterior fossa brain malformations, hemangiomas, arterial anomalies, coarctation of the aorta and cardiac defects, and eye abnormalities. *Arch. Dermatol.* **132**, 307-311, (1996).
49 Sanchez, R. L. & Raimer, S. S. Clinical and histologic features of striated muscle hamartoma: possible relationship to Delleman's syndrome. *J. Cutan. Pathol.* **21**, 40-46, (1994).
50 Moog, U. Encephalocraniocutaneous lipomatosis. *J. Med. Genet.* **46**, 721-729, (2009).
51 Kurlemann, G. Neurocutaneous syndromes. *Handb. Clin. Neurol.* **108**, 513-533, (2012).
52 Herron, J., Darrah, R. & Quaghebeur, G. Intra-cranial manifestations of the neurocutaneous syndromes. *Clin. Radiol.* **55**, 82-98, (2000).
53 Tortora, D. *et al.* Moyamoya Vasculopathy in PHACE Syndrome: Six New Cases and Review of the Literature. *World Neurosurg.* **108**, 291-302, (2017).
54 Ullrich, N. J. Neurocutaneous Syndromes and Brain Tumors. *J. Child Neurol.* **31**, 1399-1411, (2016).
55 Maher, E. R., Neumann, H. P. & Richard, S. von Hippel-Lindau disease: a clinical and scientific review. *Eur. J. Hum. Genet.* **19**, 617-623, (2011).
56 Barkovich, A. J., Guerrini, R., Kuzniecky, R. I., Jackson, G. D. & Dobyns, W. B. A developmental and genetic classification for malformations of cortical development: update 2012. *Brain* **135**, 1348-1369, (2012).
57 Moog, U., Jones, M. C., Bird, L. M. & Dobyns, W. B. Oculocerebrocutaneous syndrome: the brain malformation defines a core phenotype. *J. Med. Genet.* **42**, 913-921, (2005).
58 Thiele, E. A. Managing epilepsy in tuberous sclerosis complex. *J. Child Neurol.* **19**, 680-686, (2004).
59 Joinson, C. *et al.* Learning disability and epilepsy in an epidemiological sample of individuals with tuberous sclerosis complex. *Psychol. Med.* **33**, 335-344, (2003).
60 London, A., Benhar, I. & Schwartz, M. The retina as a window to the brain-from eye research to CNS disorders. *Nat. Rev. Neurol.* **9**, 44-53, (2013).
61 Roach, E. S. & Miller, V. S. *Neurocutaneous disorders*. (Cambridge University Press, 2004).
62 McKinnon, P. J. ATM and ataxia telangiectasia. *EMBO reports* **5**, 772-776, (2004).
63 Savitsky, K. *et al.* A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science* **268**, 1749-1753, (1995).
64 Hoeijmakers, J. H. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366-374, (2001).
65 van Gent, D. C., Hoeijmakers, J. H. & Kanaar, R. Chromosomal stability and the DNA double-stranded break connection. *Nat. Rev. Genet.* **2**, 196-206, (2001).
66 Varon, R. *et al.* Nibrin, a novel DNA double-strand break repair protein, is mutated in Nijmegen breakage syndrome. *Cell* **93**, 467-476, (1998).

67 Marsh, A., Wicking, C., Wainwright, B. & Chenevix-Trench, G. DHPLC analysis of patients with Nevoid Basal Cell Carcinoma Syndrome reveals novel PTCH missense mutations in the sterol-sensing domain. *Hum. Mutat.* **26**, 283, (2005).
68 Smahi, A. *et al.* The NF-kappaB signalling pathway in human diseases: from incontinentia pigmenti to ectodermal dysplasias and immune-deficiency syndromes. *Hum. Mol. Genet.* **11**, 2371-2375, (2002).
69 Smith, M. J. *et al.* Germline mutations in SUFU cause Gorlin syndrome-associated childhood medulloblastoma and redefine the risk associated with PTCH1 mutations. *J. Clin. Oncol.* **32**, 4155-4161, (2014).
70 Smahi, A. *et al.* Genomic rearrangement in NEMO impairs NF-kappaB activation and is a cause of incontinentia pigmenti. The International Incontinentia Pigmenti (IP) Consortium. *Nature* **405**, 466-472, (2000).
71 Taylor, T. H. *et al.* The origin, mechanisms, incidence and clinical consequences of chromosomal mosaicism in humans. *Hum. Reprod. Update* **20**, 571-581, (2014).
72 Happle, R. *Mosaicism in human skin : understanding nevi, nevoid skin disorders, and cutaneous neoplasia*. (Springer, 2013).
73 Biesecker, L. G. & Spinner, N. B. A genomic view of mosaicism and human disease. *Nat. Rev. Genet.* **14**, 307-320, (2013).
74 Lara-Corrales, I. *et al.* Mosaic Neurofibromatosis Type 1 in Children: A Single-Institution Experience. *J. Cutan. Med. Surg.* **21**, 379-382, (2017).
75 Boycott, K. M., Vanstone, M. R., Bulman, D. E. & MacKenzie, A. E. Rare-disease genetics in the era of next-generation sequencing: discovery to translation. *Nat. Rev. Genet.* **14**, 681-691, (2013).
76 Rommens, J. M. *et al.* Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science* **245**, 1059-1065, (1989).
77 Choi, M. *et al.* Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19096-19101, (2009).
78 Ng, S. B., Nickerson, D. A., Bamshad, M. J. & Shendure, J. Massively parallel sequencing and rare disease. *Hum. Mol. Genet.* **19**, R119-124, (2010).
79 Biesecker, L. G., Shianna, K. V. & Mullikin, J. C. Exome sequencing: the expert view. *Genome Biol.* **12**, 128, (2011).

Chapter 2: Microcephaly-capillary malformation syndrome

2.1 Preface

The following chapter consists of the manuscript titled "Mutations in *STAMBP*, encoding a deubiquitinating enzyme, cause microcephaly–capillary malformation syndrome" published in Nature Genetics by Laura M. McDonell, Ghayda M. Mirzaa, Diana Alcantara, Jeremy Schwartzentruber, Melissa T. Carter, Leo J. Lee, Carol L. Clericuzio, John M. Graham Jr, Deborah J. Morris-Rosendahl, Tilman Polster, Gyula Acsadi, Sharron Townshend, Simon Williams, Anne Halbert, Bertrand Isidor, Albert David, Christopher D. Smyser, Alex R. Paciorkowski, Marcia Willing, John Woulfe, Soma Das, Chandree L. Beaulieu, Janet Marcadier, FORGE Canada Consortium, Michael T. Geraghty, Brendan J. Frey, Jacek Majewski, Dennis E. Bulman, William B. Dobyns, Mark O′Driscoll and Kym M. Boycott.

2.2 Statement of permission for use of copyrighted material

Statement of permission for the reproduction of copyrighted material from Springer Nature can be found in Appendix B.

2.3 Contributions

The specific contributions of each author are as follows:

Laura M. McDonell

Analyzed the sequencing data and confirmed causative MIC-CAP mutations. Using Sanger sequencing, validated WES-identified mutations, performed co-segregation analysis and screened additional affected individuals. Performed tissue culture for cell lines of patients 5.1 and 7.1 Identified region of loss-of-heterozygosity in patient 5.1 as well as deep intronic

mutation causing inclusion of pseudo-exon which was quantified using qRT-PCR. Used qRT-PCR to quantify alternative splicing in patient 5.1. Performed microsatellite analysis to reveal maternal isodisomy in Family 4. Coordinated splicing analysis. Performed protein modeling. Co-wrote manuscript, generated figures, contributed to tables and responded to reviewer comments.

Ghayda M. Mirzaa

Contributed clinical data, co-write clinical portions of paper.

Diana Alcantara

Performed protein biochemistry and cell biology to characterize STABMP expression and Ras-MAPK, PI3K-AKT signalling. Investigated ubiquitin aggregation, apoptosis and autophagic flux. Generated figures.

Jeremy Schwartzentruber

Performed bioinformatics analysis of the Illumina sequencing data and generated list of candidate genes. Wrote bioinformatics analysis methods section of the manuscript. Generated bioinformatics table for the manuscript.

Melissa T. Carter

Contributed clinical data.

Leo J. Lee

Performed computational analysis of intronic MIC-CAP mutations.

Carol L. Clericuzio, John M. Graham Jr, Deborah J. Morris-Rosendahl, Tilman Polster, Gyula Acsadi, Sharron Townshend, Simon Williams, Anne Halbert, Bertrand Isidor, Albert David, Christopher D. Smyser, Alex Paciorkowski, Marcia Willing, John Woulfe, Soma Das

Contributed clinical data.

Chandree L. Beaulieu

Project Manager of the FORGE Canada Consortium.

Janet Marcadier

Clinical Coordinator of the FORGE Canada Consortium.

FORGE Canada Consortium

Provided critical infrastructure and resources for this project.

Michael T. Geraghty

Provided clinical data.

Brendan J. Frey

Oversaw computational analysis of intronic MIC-CAP mutations.

Jacek Majewski

Oversaw bioinformatics analysis of the Illumina sequencing data.

Dennis E. Bulman

Directed the study.

William B. Dobyns

Directed the study. Contributed clinical data. Edited the manuscript. Oversaw response to reviewer comments.

Mark O'Driscoll

Directed the cell biology and protein biochemistry studies. Contributed to writing and editing the manuscript. Oversaw response to reviewer comments.

Kym M. Boycott

Directed the study. Contributed clinical data. Contributed to writing and editing the manuscript. Oversaw response to reviewer comments.

2.4 Main Manuscript

Mutations in *STAMBP*, encoding a deubiquitinating enzyme, cause microcephaly–capillary malformation syndrome

Laura M McDonell[1], Ghayda M Mirzaa[2], Diana Alcantara[3], Jeremy Schwartzentruber[4], Melissa T Carter[5], Leo J Lee[6], Carol L Clericuzio[7], John M Graham Jr[8], Deborah J Morris-Rosendahl[9], Tilman Polster[10], Gyula Acsadi[11], Sharron Townshend[12], Simon Williams[13,14], Anne Halbert[15], Bertrand Isidor[16], Albert David[16], Christopher D Smyser[17], Alex R Paciorkowski[18], Marcia Willing[19], John Woulfe[20], Soma Das[2], Chandree L Beaulieu[1], Janet Marcadier[1], FORGE Canada Consortium[21], Michael T Geraghty[1], Brendan J Frey[6], Jacek Majewski[22], Dennis E Bulman[1], William B Dobyns[23–26], Mark O'Driscoll[3,26] & Kym M Boycott[1,26]

Microcephaly–capillary malformation (MIC-CAP) syndrome is characterized by severe microcephaly with progressive cortical atrophy, intractable epilepsy, profound developmental delay and multiple small capillary malformations on the skin. We used whole-exome sequencing of five patients with MIC-CAP syndrome and identified recessive mutations in *STAMBP*, a gene encoding the deubiquitinating (DUB) isopeptidase STAMBP (STAM-binding protein, also known as AMSH, associated molecule with the SH3 domain of STAM) that has a key role in cell surface receptor–mediated endocytosis and sorting. Patient cell lines showed reduced STAMBP expression associated with accumulation of ubiquitin-conjugated protein aggregates, elevated apoptosis and insensitive activation of the RAS-MAPK and PI3K-AKT-mTOR pathways. The latter cellular phenotype is notable considering the established connection between these pathways and their association with vascular and capillary malformations. Furthermore, our findings of a congenital human disorder caused by a defective DUB protein that functions in endocytosis implicates ubiquitin-conjugate aggregation and elevated apoptosis as factors potentially influencing the progressive neuronal loss underlying MIC-CAP syndrome.

MIC-CAP syndrome was recently described in six children, including one brother-sister pair, who all presented with small scattered capillary malformations, severe congenital microcephaly, early onset intractable epilepsy, profound global developmental delay, spastic quadriparesis, hypoplastic distal phalanges and poor growth[1–4]. Capillary malformations, sometimes referred to as port-wine stains, are nonregressing cutaneous vascular abnormalities[5] that are seen in a growing number of congenital syndromes linked to dysregulated RAS-MAPK (RAS-mitogen activated protein kinase) function; these are collectively termed 'RASopathies'. For example, mutations in *RASA1*, encoding p120-RasGAP, a negative regulator of the RAS pathway, have been found in patients with capillary malformation–arteriovenous malformation syndrome[6], and mutations in *KRIT1*, encoding a RAS-related protein 1A interactant, cause hyperkeratotic cutaneous capillary–venous malformations associated with cerebral capillary malformations[7]. Sequencing of *RASA1* in two patients with MIC-CAP syndrome did not show any mutations, and sequencing of *KRIT1* was not pursued[7]. Until now, the genetic mechanism responsible for this devastating disorder has been unknown.

We studied ten affected individuals from nine families with MIC-CAP syndrome (**Fig. 1** and **Table 1**). Brain magnetic resonance imaging scans of the affected individuals showed enlarged extra-axial spaces and other changes suggesting prenatal-onset cerebral atrophy with relative sparing of the cerebellum (**Fig. 1a–c**). The gyral pattern was universally simplified and was associated with variable degrees of diffuse hypomyelination and hippocampal hypoplasia. We found

[1]Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, Ontario, Canada. [2]Department of Human Genetics, University of Chicago, Chicago, Illinois, USA. [3]Genome Damage and Stability Centre, University of Sussex, Brighton, UK. [4]McGill University and Genome Quebec Innovation Centre, Montréal, Quebec, Canada. [5]Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, Ontario, Canada. [6]Department of Electrical and Computer Engineering, Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario, Canada. [7]Department of Genetics, University of New Mexico Health Sciences Center, Albuquerque, New Mexico, USA. [8]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [9]Institute of Human Genetics, University Clinic Freiburg, Freiburg, Germany. [10]Bethel Epilepsy Center, Krankenhaus Mara, Bielefeld, Germany. [11]Department of Neurology, Connecticut Children's Medical Center, Hartford, Connecticut, USA. [12]Genetics Service of Western Australia, King Edward Memorial Hospital, Perth, Western Australia, Australia. [13]Department of Neurology, Princess Margaret Hospital, Perth, Western Australia, Australia. [14]Department of Pediatric Rehabilitation, Princess Margaret Hospital, Perth, Western Australia, Australia. [15]Department of Pediatric Dermatology, Princess Margaret Hospital for Children, Subiaco, Western Australia, Australia. [16]Centre Hospitalier Universitaire Nantes, Service de Génétique Médicale, Nantes, France. [17]Department of Neurology, Washington University, St. Louis, Missouri, USA. [18]Department of Neurology, University of Washington and Seattle Children's Research Institute, Seattle, Washington, USA. [19]Department of Pediatrics, Washington University, St. Louis, Missouri, USA. [20]Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada. [21]Membership of the Steering Committee for the Consortium is provided in the Acknowledgments section. [22]Department of Human Genetics, McGill University, Montréal, Quebec, Canada. [23]Department of Pediatrics, University of Washington, Seattle, Washington, USA. [24]Department of Neurology, University of Washington, Seattle, Washington, USA. [25]Center for Integrative Brain Research, Seattle Children's Hospital, Seattle, Washington, USA. [26]These authors jointly directed this work. Correspondence should be addressed to K.M.B. (kboycott@cheo.on.ca) or M.O. (m.o-driscoll@sussex.ac.uk).

Received 12 January; accepted 7 March; published online 31 March 2013; doi:10.1038/ng.2602

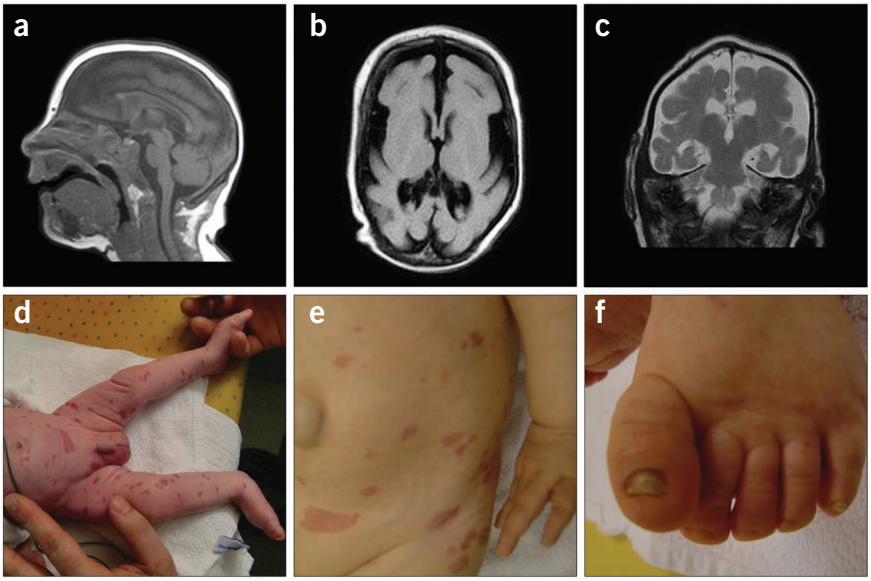


Figure 1 Neuroimaging and clinical features of MIC-CAP syndrome in patient 9.1. (**a**–**c**) T1-weighted sagittal (**a**), and axial (**b**) and T2-weighted coronal (**c**) images of the brain of patient 9.1 at 3 months of age. Note the low-sloping forehead, simplified gyral pattern, increased extra-axial space, diffuse hypomyelination and hippocampal hypoplasia. (**d**–**f**) Photos of patient 9.1 at 3 weeks (**d**) and 18 months (**e**) of age showing generalized capillary malformations of variable sizes and hypoplastic toenails (**f**).

that all individuals with MIC-CAP syndrome had intractable epilepsy, severe developmental delay and profound intellectual disability. Other distinguishing features of MIC-CAP syndrome include infantile spasms, hypoplasia of the distal phalanges characterized by variable degrees of nail and toe hypoplasia and capillary malformations (**Fig. 1d**–**f**). The capillary malformations were striking in appearance and visible at birth in all patients. They were also generalized in distribution and tended to vary from small (2–3 mm) to large (15–20 mm) lesions. Interestingly, limited evidence suggests that the vascular anomalies are not restricted to skin capillary malformations; one patient (designated P3.1 in this study) had a cerebellar angioma[1], and another patient (9.1) had possible vascular malformations of the liver as determined by ultrasound (data not shown).

To establish the genetic cause of MIC-CAP syndrome, we performed exome sequencing on DNA samples from five individuals (**Table 1**) diagnosed with MIC-CAP syndrome. The two affected children in family 1, from nonconsanguineous parents, suggested a recessive mode of inheritance for this disorder. Therefore, we focused on identifying genes in which a maximal number of patients had two rare protein-altering variants that were absent from dbSNP131, the 1000 Genomes Project and 159 in-house control exomes. In four of the five patients studied by exome sequencing (P1.1, P1.2, P2.1 and P3.1), including the two siblings from family 1, we identified two variants in *STAMBP* in each individual (**Fig. 2a**, **Supplementary Figs. 1a**, **2** and **Supplementary Table 1**). Analysis of an additional three affected individuals (P6.1, P8.1 and P9.1) by Sanger sequencing identified two coding *STAMBP* variants in each patient (**Supplementary Table 2**). Co-segregation analysis confirmed an autosomal-recessive mode of inheritance in all families (**Supplementary Fig. 3**). Protein blot analysis of whole-cell extracts from patient-derived lymphoblastoid cell lines (LCLs) did not detect STAMBP expression in patient 1.2 (p.[Glu42Gly];[Arg178*]) (**Fig. 2b**). Patient 3.1 (p.[Phe100Tyr];[Arg424*]) showed a reduction of STAMBP expression compared to wild-type (WT) controls (**Fig. 2b**).

We identified one coding mutation in *STAMBP* in patient 7.1. Analysis by protein blotting did not detect STAMBP expression in this individual (**Fig. 2b**), and further sequencing of the gene revealed an intronic mutation (c.203+5G>A) believed to lead to an increase in skipping of the first coding exon (**Table 1** and **Supplementary Figs. 1b** and **4a**–**d**).

In patient 5.1, we did not identify any coding mutations using exome sequencing. The depth of coverage across the exons of *STAMBP* did not suggest a deletion. However, analysis of SNP data from an Illumina Human Omni2.5 array, which contains 25 probes within *STAMBP*, suggested a 40-Mb region of copy-neutral homozygosity spanning *STAMBP*. Protein blotting revealed a severe reduction in STAMBP expression in this patient (**Fig. 2b**), suggesting that P5.1 has MIC-CAP syndrome secondary to noncoding mutations in *STAMBP*. Sequencing of patient-derived complementary DNA showed the presence of a 108-bp pseudoexon containing a premature stop codon (**Supplementary Fig. 4e**,**f**). Deep intronic sequencing identified a homozygous mutation (c.1005+358A>G). Application of a computational model of splicing regulation[8] predicted that this mutation would activate a new donor site, as well as a cryptic AG acceptor site 114 bp upstream ($P = 8.7 \times 10^{-7}$, sign test). We believed that this mutation caused the leaky splicing of the full-length transcript and showed that patient cells have a threefold reduction of full-length transcript expression (**Supplementary Figs. 1c** and **4g**).

Sanger sequencing in patient 4.1 identified a homozygous stop mutation encoding p.Arg424*. Co-segregation analysis was not consistent with the suspected autosomal-recessive mode of inheritance in this nonconsanguineous family, as only the mother was heterozygous for the mutation causing p.Arg424*. We analyzed ten microsatellite markers spanning chromosome 2 and found all markers to be homozygous; a diagnostic array performed using DNA extracted from whole blood showed no evidence of copy number variation across chromosome 2. Therefore, we suspect that the mechanism of MIC-CAP syndrome in this patient is secondary to maternal isodisomy (**Supplementary Fig. 5**). In summary, we identified two mutations in *STAMBP* in a total of ten patients: six missense variants, two nonsense mutations, two translational frameshift mutations predicted to cause a premature truncation of the STAMBP protein and three intronic mutations leading to alternative splicing of the *STAMBP* transcript (**Fig. 2a**).

STAMBP is a JAMM-family DUB containing a microtubule-interacting and transport (MIT) domain and a STAM-binding domain, both of which interact with the endosomal sorting and trafficking machinery (**Fig. 2a** and **Supplementary Fig. 6a**)[9–11]. STAMBP is recruited to the endosomal sorting complexes required for transport (ESCRTs), a group of distinct macromolecule assemblies that mediate the sorting and trafficking of ubiquitinated proteins from endosomes to lysosomes. STAMBP functions in regulating endosomal sorting of ESCRT machinery and ubiquitinated receptor cargo[9,12–17]. Endosomal sorting is a highly dynamic process that is fundamental to regulating protein homeostasis through the active regulation of receptor-mediated signal transduction and enabling processes such

as autophagy[18,19]. Impaired ESCRT function is associated with the intracellular accumulation of ubiquitinated proteins. Brain lesions containing ubiquitinated protein aggregates have been noted in *Stambp*$^{-/-}$ mice[20], suggesting this to be a probable mechanism influencing microcephaly and its progression in MIC-CAP syndrome. Consistent with this, we observed elevated amounts of conjugated-ubiquitin aggregates after short interfering RNA (siRNA)-mediated silencing of *STAMBP* in the human medullablastoma line T98G using indirect immunofluoresence with an antibody that specifically detects conjugated ubiquitin (FK2) and not free ubiquitin (**Fig. 3a** and **Supplementary Fig. 6b**). Notably, we also observed elevated amounts of conjugated-ubiquitin aggregates in several LCLs

Table 1 Clinical characteristics and molecular findings in patients with MIC-CAP[a]

	MIC-CAP									
Patient	P1.1	P1.2	P2.1	P3.1	P4.1	P5.1	P6.1	P7.1	P8.1	P9.1
Exome sequencing	+	+	+	+	–	+	–	–	–	–
Validated mutations[b] (protein alteration (cDNA))	p.Glu42Gly (c.125A>G)	p.Glu42Gly (c.125A>G)	p.Arg38Cys (c.112C>T)	p.Phe100Tyr (c.299T>A)	p.Arg424* (c.1270C>T)	c.1005+ 358A>G	p.Lys378Asnfs*2 (c.1134_1138 delACTAA)	p.Arg38Cys (c.112C>T)	p.Arg38Cys (c.112C>T)	p.Tyr63Cys (c.188A>G)
	p.Arg178* (c.532C>T)	p.Arg178* (c.532C>T)	c.279+5G>T	p.Arg424* (c.1270C>T)	p.Arg424* (c.1270C>T)	c.1005+ 358A>G	p.Thr313Ile (c.938C>T)	c.203+5G>A	p.Ile138Serfs*12 (c.411del C)	p.Arg14Pro (c.41G>C)
Gender	F	M	M	M	M	F	M	F	F	M
Age at assessment	2 y	9 mo	12 mo	2 y	22 d	5 y 4 mo	2 mo	28 mo	8 mo	15 mo
Ethnicity	African American	African American	European descent	European descent	European descent	European descent	European descent	European descent	Polynesian	European descent
Gestational age (weeks)	39	39	36+5	36	37	Term	36	37+2	37+6	35
Clinical features										
Progressive congenital MIC	+	+	+	+	+	+	+	+	+	+
Small for gestational age	+	+	+	+	+	–	–	+	–	+
Birth OFC (s.d.)	–6	–4	–2	–2	–4	–1.8	–2	–8	–5	–2
Birth weight (s.d.)	–1.5	–1.5	–1.5	–2	–2	+1.8	–1.5	–4	–1.5	–1.5
Birth length (s.d.)	–4	–2	0.5	ND	–1.5	0 (mean)	–2	–4	–1 to –2	–2
Later age	2.5 y	9 mo	12 mo	17 mo	22 d	2 y 10 mo	2 mo	28 mo	8 mo	18 mo
Later OFC (s.d., age)	–8	–6	–6	–4	–3	–2.5	–6	–8	–4	–4
Later weight (s.d.)	+1.5 to +2	–4	+1 to +2	–1 to –2	–3	+2	–2	–2.5	+1	–3
Later length (s.d.)	ND	ND	–1 to –2	–2	–3	0 (mean)	–3	–4	ND	–3 to –4
Generalized capillary malformations	+	+	+	+	+	+	+	+	+	+
Early onset intractable seizures	+	+	+	+	+	+	+	+	+	+
Infantile spasms	+	–	+	–	–	–	–	ND	+	+
Hypoplastic distal phalanges	+	+	–	+	+	+	+	+	+	+
Global DD	+	+	+	+	+	+	+	+	+	+
Spastic quadriparesis	+	+	+	+	+	–	–	+	+	+
Myoclonus	+	–	+	+	–	–	+	+	Intermittent dyskinetic and choreiform movements	+
Optic atrophy	+	+	+	+	+	ND	ND	+	–	–
Neuroimaging features										
Simplified gyral pattern	+	+	+	+	+	ND	+	+	+	+
Increased extra-axial space	+	+	+	+	+	ND	+	+	+	+
Hippocampal hypoplasia	+	+	+	ND	+	ND	–	+	+	ND
Hypomyelination	–	–	+	+	+	ND	ND	+	–	+
Published	+ (ref. 3)	+ (ref. 3)	+ (ref. 3)	+ (ref. 1)	+ (ref. 1)	+ (ref. 2)	–	–	–	–

[a]This table summarizes the clinical findings in the study participants. [b]The numbering of the mutations and alterations is relative to NM_006463.4 (gene) and NP_006454.1 (protein), respectively. OFC, occipitofrontal circumference; s.d., values are shown by their s.d. value from the mean; ND, not determined; DD, developmental delay.

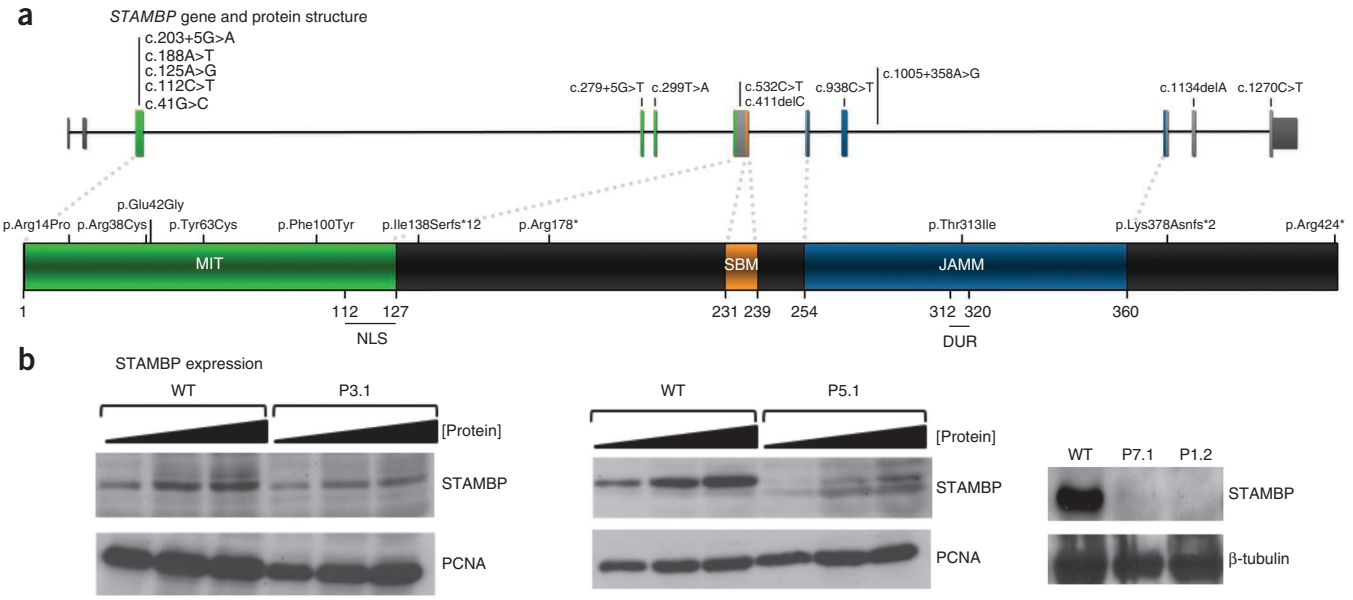


Figure 2 Mutations in *STAMBP* cause MIC-CAP syndrome. (**a**) The *STAMBP* gene (top; chromosome 2, hg19 74,056,114–74,094,295, RefSeq NM_006463.4) and STAMBP protein (bottom; NP_006454.1) indicating MIC-CAP mutations and the resulting alterations. STAMBP contains a MIT domain[9,10], a SH3 binding motif (SBM) (PX[V/I][D/N]RXXP)[26], a JAMM (JAB1/MPN/MOV34) motif[12], a nuclear localization signal (NLS)[11] and the distal ubiquitin recognition site (DUR)[27]. For c.279+5G>T in P2.1 (tissue from this patient was not available), a computational splicing model predicted the inclusion of an extra codon in exon 4 ($P = 1.9 \times 10^{-9}$, sign test). We validated this model using the known mutation in P7.1 ($P = 1.9 \times 10^{-9}$, sign test) (**Supplementary Fig. 1a**). Five out of six missense alterations are located in the MIT domain, which is required for the interaction of STAMBP with CHMP3, an ESCRT-III subunit[28]. The sixth alteration, p.Thr313Ile, located in the distal ubiquitin binding site within the JAMM domain, eliminates a hydrogen bond between the ubiquitin carbon backbone and STAMBP, probably decreasing ubiquitin binding to STAMBP (**Supplementary Fig. 2**). Two alterations were recurrent in multiple unrelated MIC-CAP families; p.Arg424*, detected in patients 3.1 and 4.1, and p.Arg38Cys, detected in individuals P2.1, P7.1 and P8.1, suggestive of mutational hotspots in *STAMBP*. Within the ~5,000 exomes in the NHLBI Exome variant server, only p.Arg38Cys was represented in 2 of 10,756 alleles, suggesting a carrier frequency of approximately 1:5,000 in a population of African and European ancestry, consistent with the prevalence of this very rare disorder. (**b**) Protein blot analysis of whole-cell extracts of LCLs from P3.1, P5.1, P7.1 and P1.2 showing equivocal (P3.1), reduced (P5.1) or absent STAMBP expression (P7.1 and P1.2).

from patients with STAMBP alterations compared to WT control LCLs after serum starvation (**Fig. 3b**). This phenotype was reversed after stable lentiviral transduction of patient LCLs with *STAMBP* (**Supplementary Fig. 6c**,**d**). Furthermore, this phenotype was also associated with apoptosis induction, denoted by elevated amounts of cleaved caspase 3 (**Fig. 3c**) and annexin V staining (**Fig. 3d**) in the LCLs of patients with STAMBP alterations compared to WT LCLs. STAMBP functions with the ESCRT machinery to facilitate autophagy. Autophagic flux can be monitored by detection of the expression of the autophagosome-associated phosphatidylethanolamine-conjugated microtubule-associated light chain 3 (LC3-II isoform) in the presence of the autophagy inhibitor bafilomycin A. Consistent with increased autophagic flux (that is, increased amounts of autophagosomes), we found elevated amounts of LC3-II in LCLs from patients with STAMBP alterations compared to WT control LCLs (**Fig. 3e**).

ESCRT-mediated endocytosis of activated cell-surface receptors (for example, activated receptor tyrosine kinases or G protein–coupled receptors) controls receptor distribution and coordinates signal transduction amplitude and duration[17]. Endocytosed ubiquitinated receptors are either recycled to the cell surface or targeted for degradation in the lysosome, leading to the proteolysis and termination of receptor signaling[21]. STAMBP interacts with key components of receptor signaling pathways, such as Grb2 (**Fig. 4a**)[10,22]. Considering the known role of STAMBP in regulating receptor-mediated endocytosis, sorting and trafficking, we investigated aspects of the interconnected RAS-MAPK and PI3K-AKT-mTOR signal transduction pathways in our MIC-CAP LCLs, as mutations in components of these networks are associated with congenital capillary malformation disorders[6,7,23]. We found elevated amounts of GTP-bound RAS (active RAS) in extracts from LCLs of patients with STAMBP alterations compared with WT LCLs, which is suggestive of elevated signaling through this pathway (**Fig. 4b**). Similarly, we found elevated amounts of phosphorylated active phosphoinositol 3-kinase (PI3K) in cell extracts from LCLs from patients with STAMBP alterations relative to WT cells, even after serum starvation (**Fig. 4c**). Collectively, these data suggest elevated and insensitive active signal transduction in these interconnected pathways associated with defective STAMBP in patient LCLs.

To further characterize signaling abnormalities, we examined the response of patient LCLs to serum starvation for both of these pathways using a selection of substrates. Serum starvation induced a substantial reduction in C-RAF phosphorylation at Ser338 in WT LCLs, consistent with inhibition of C-RAF activity under these conditions (**Fig. 5a**). LCLs from patients with STAMBP alterations maintained C-RAF phosphorylation at Ser338 in the absence of serum, indicating persistent activation and insensitivity of this pathway. Further evidence suggesting insensitive signal transduction in the RAS-MAPK pathway in *STAMBP*-mutated LCLs is given by the relative insensitivity of these cells to the MEK1 and MEK2 (MEK1/2) inhibitor U0126. Active C-RAF phosphorylates and activates MEK1/2 kinase, which then phosphorylates and activates ERK1 and ERK2 (ERK1/2) (**Fig. 4a**). We repeatedly found elevated amounts of phosphorylated ERK1/2 in exponentially growing *STAMBP*-mutated LCLs compared to WT LCLs after a short treatment (1 h) with U0126 (**Fig. 5b** and **Supplementary Fig. 6e**). The excess of phosphorylated ERK1/2 in

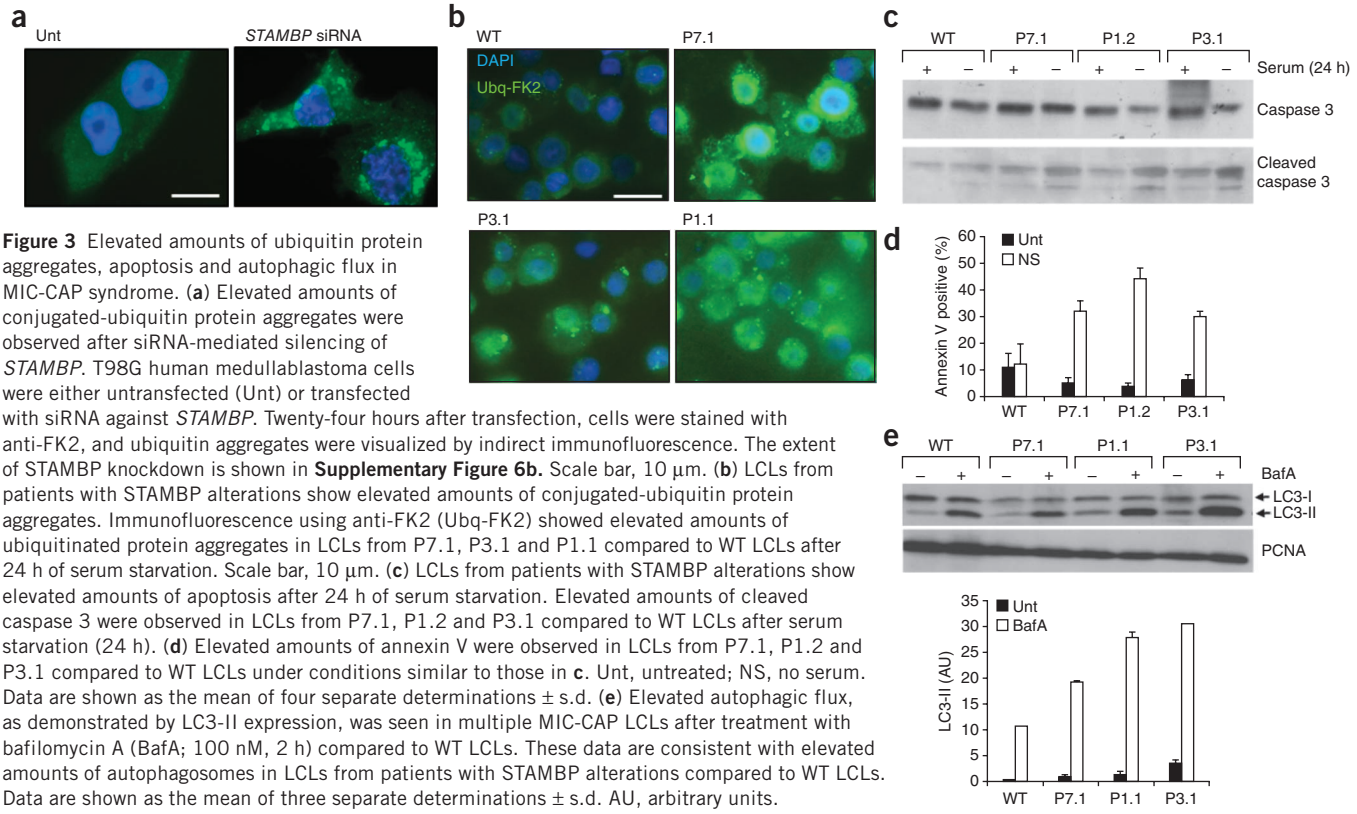


Figure 3 Elevated amounts of ubiquitin protein aggregates, apoptosis and autophagic flux in MIC-CAP syndrome. (**a**) Elevated amounts of conjugated-ubiquitin protein aggregates were observed after siRNA-mediated silencing of *STAMBP*. T98G human medullablastoma cells were either untransfected (Unt) or transfected with siRNA against *STAMBP*. Twenty-four hours after transfection, cells were stained with anti-FK2, and ubiquitin aggregates were visualized by indirect immunofluorescence. The extent of STAMBP knockdown is shown in **Supplementary Figure 6b.** Scale bar, 10 µm. (**b**) LCLs from patients with STAMBP alterations show elevated amounts of conjugated-ubiquitin protein aggregates. Immunofluorescence using anti-FK2 (Ubq-FK2) showed elevated amounts of ubiquitinated protein aggregates in LCLs from P7.1, P3.1 and P1.1 compared to WT LCLs after 24 h of serum starvation. Scale bar, 10 µm. (**c**) LCLs from patients with STAMBP alterations show elevated amounts of apoptosis after 24 h of serum starvation. Elevated amounts of cleaved caspase 3 were observed in LCLs from P7.1, P1.2 and P3.1 compared to WT LCLs after serum starvation (24 h). (**d**) Elevated amounts of annexin V were observed in LCLs from P7.1, P1.2 and P3.1 compared to WT LCLs under conditions similar to those in **c**. Unt, untreated; NS, no serum. Data are shown as the mean of four separate determinations ± s.d. (**e**) Elevated autophagic flux, as demonstrated by LC3-II expression, was seen in multiple MIC-CAP LCLs after treatment with bafilomycin A (BafA; 100 nM, 2 h) compared to WT LCLs. These data are consistent with elevated amounts of autophagosomes in LCLs from patients with STAMBP alterations compared to WT LCLs. Data are shown as the mean of three separate determinations ± s.d. AU, arbitrary units.

STAMBP-mutated LCLs under these robust inhibition conditions is further supportive of a hyperactive and insensitive RAS-MAPK pathway in these cells. Analysis of several endpoints in the PI3K-AKT-mTOR pathway under identical conditions indicated a similar insensitive activation of this pathway. Serum starvation of WT LCLs reduced the phosphorylation of AKT at Thr308, of the AKT-dependent Thr1462 of TSC2 and of Ser240 and Ser244 of S6 protein. This is consistent with pathway inactivation under these conditions in WT LCLs (**Fig. 5c**,**d**). In contrast, LCLs from patients with STAMBP alterations maintained phosphorylation of all three proteins under these conditions (**Fig. 5c**,**d**). Furthermore, stable lentiviral transduction of patient LCLs with *STAMBP* resulted in the reconstitution of a normal signaling response to serum starvation (**Supplementary Fig. 6f**).

The RAS-MAPK and PI3K-AKT-mTOR pathways regulate crucial cellular processes, including cell growth, cell-cycle progression and differentiation. Disorders characterized by hyperactivity of the RAS-MAPK network, including Noonan and Costello syndromes, present with growth delay[24]. Considering the marked postnatal growth retardation and capillary abnormalities seen in MIC-CAP syndrome, hyperactive RAS-MAPK signaling may be a major biological

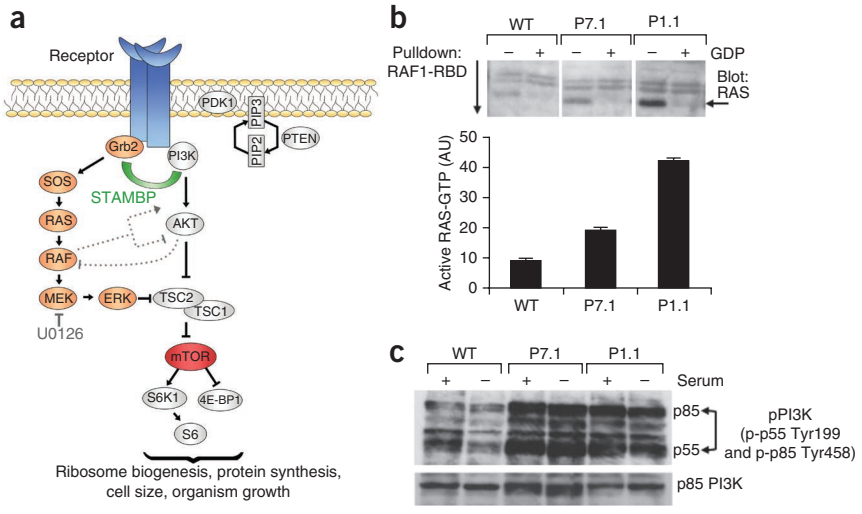


Figure 4 Elevated RAS-GTP (active RAS) and activated PI3 kinase in MIC-CAP syndrome. (**a**) Schematic overview of the core components of the RAS-MAPK and PI3K-AKT-mTOR networks highlighting the interconnectivity. As well as interacting with the ESCRT machinery and STAM, STAMBP has been shown to interact with other important components of these signal transduction pathways, including the Grb2 adaptor and the class II PI3 kinase catalytic subunit. (**b**) GTP-bound active RAS was precipitated from whole-cell extracts using recombinant RAF1-RBD (RAS binding domain) GST (glutathione S-transferase) beads followed by protein blotting for RAS. GDP was shown to effectively outcompete any interaction. Elevated amounts of RAS-GTP were pulled down from LCLs from P7.1 and P1.1 compared to WT LCLs. On the bottom is an ImageJ–based quantification of active RAS-GTP from three separate experiments. Data are shown as the mean ± s.d. AU, arbitrary units. (**c**) Serum starvation (24 h) reduced PI3 kinase activation in WT LCLs as monitored by phosphorylation of the PI3K subunits p55 at Tyr199 (p-p55) and p85 at Tyr458 (p-p85). Amounts of phosphorylated PI3K (pPI3K) were found to be elevated in extracts of LCLs from P7.1 and P1.1 either endogenously or after serum starvation, which is suggestive of hyperactive and insensitive PI3K activity.

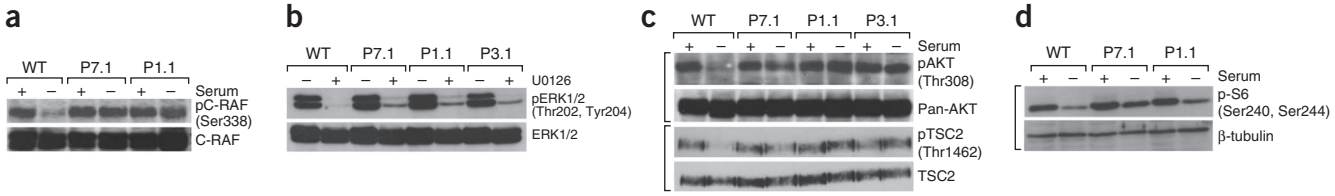


Figure 5 Elevated and insensitive RAS-MAPK and PI3K-AKT-mTOR signaling in MIC-CAP syndrome. (**a**) Serum starvation (24 h) inhibits C-RAF activation in WT LCLs, in contrast to LCLs from P7.1 and P1.1. pC-RAF, phosphorylated C-RAF. (**b**) LCLs were either treated (+) or not treated (–) with 10 μM U0126, a specific MEK1/2 inhibitor, for 1 h (**Fig. 4a**). Cells were harvested, and whole-cell extracts were probed for phosphorylation of ERK1/2 (pERK1/2), which is mediated by MEK. Insensitivity to this treatment (as measured by relative amounts of pERK1/2 remaining after treatment with the MEK inhibitor) would reflect the magnitude and intensity of signal transduction from RAF to MEK to ERK (**Fig. 4a**). Residual pERK1/2 (Thr202 and Tyr204) signal (MEK-dependent phosphorylation) was seen in MIC-CAP LCLs in contrast to WT LCLs. This phenotype is underscored after titration of U0126 in various MIC-CAP LCLs compared to WT LCLs (**Supplementary Fig. 6e**). Collectively, these data indicate a greater strength of MEK1/2 activity in MIC-CAP LCLs compared to WT cells. (**c**) Serum starvation (24 h) reduces phosphorylation (activation) of AKT at Thr308 (pAKT) and of TSC2 at Thr1462 (pTSC2) and AKT-dependent inhibitory phosphorylation of TSC2 in WT LCLs in contrast to LCLs from P7.1, P1.1 and P3.1. The TSC1 and TSC2 complex is the principal negative regulator of the mTOR kinase complex (**Fig. 4a**). These data are consistent with active signal transduction from PI3K-AKT-mTOR in MIC-CAP cells under these conditions. (**d**) S6 protein is phosphorylated by S6 kinase in an mTOR-dependent fashion (**Fig. 4a**). Consistent with active signal transduction in this pathway under serum starvation conditions, LCLs from P7.1 and P1.1 maintained S6 phosphorylation (p-S6) at Ser240 and Ser244 in contrast to WT LCLs.

consequence induced by impaired STAMBP function in humans, suggesting that *STAMBP*-mutated MIC-CAP syndrome may have an overlapping pathomechanism with the RASopathies. Furthermore, as the PI3K-AKT-mTOR pathway also has a role in angiogenesis and vascularization, and considering the interconnectivity between these networks (**Fig. 4a**), it is possible that the combined insensitive activation of both these networks may contribute to the MIC-CAP phenotype.

In summary, we identify mutations in *STAMBP* in MIC-CAP syndrome, a recently described severe developmental disorder. Analysis of LCLs from patients with MIC-CAP syndrome demonstrated elevated ubiquitin-conjugated protein aggregation and apoptosis activation. These data are consistent with elevated ubiquitin-conjugated protein aggregate–induced progressive apoptosis as a potential underlying mechanism for the microcephaly in this disorder. This is consistent with brain imaging and human pathological analysis of MIC-CAP syndrome[1] and of the knockout mouse model of *Stambp*[25]. Furthermore, we document elevated autophagosome content and active and insensitive RAS-MAPK and PI3K-AKT-mTOR pathways as previously unidentified consequences of defective STAMBP, potentially contributing to the vasculature and growth characteristics of MIC-CAP syndrome. This work presents the first example, to our knowledge, of a human disorder caused by a congenitally defective DUB isopeptidase functioning in the endocytosis pathway, providing important new insights into the pathophysiology of human microcephaly and capillary malformation.

URLs. National Heart, Lung, and Blood Institute (NHLBI) Exome variant server, http://evs.gs.washington.edu/EVS/; FASTX-Toolkit, http://hannonlab.cshl.edu/fastx_toolkit/; Picard tools, http://picard.sourceforge.net/; SAMtools, http://samtools.sourceforge.net/.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

The authors would like to thank the study patients and their families, without whose participation this work would not be possible. This work was funded by the Government of Canada through Genome Canada, the Canadian Institutes of Health Research (CIHR) and the Ontario Genomics Institute (OGI-049) (to K.M.B.). Additional funding was provided by Genome Quebec and Genome British Columbia (to K.M.B.), the US National Institutes of Health under National Institute of Neurological Disorders and Stroke (NINDS) grant NS058721 (to W.B.D.), as well as National Institute of Child Health and Human Development (NICHD) grant HD36657 and National Institute of General Medicine Sciences (NIGMS) grant 5-T32-GM08243 (to J.M.G.) and the Leukaemia Lymphoma Research (UK), Medical Research Council (UK) and Cancer Research UK (CR-UK) (to M.O.). The authors acknowledge the contribution of the high-throughput sequencing platform of the McGill University and Génome Québec Innovation Centre, Montréal, Canada, as well as M. Moellers, (Pediatric Radiology, Evangelisches Krankenhaus Bielefeld). This work was selected for study by the FORGE Canada Steering Committee, consisting of K. Boycott (University of Ottawa), J. Friedman (University of British Columbia), J. Michaud (University of Montreal), F. Bernier (University of Calgary), M. Brudno (University of Toronto), B. Fernandez (Memorial University), B. Knoppers (McGill University), M. Samuels (University of Montreal) and S. Scherer (University of Toronto). L.M.M. is supported by a Frederick Banting Graduate Scholarship from CIHR. M.O. is a CR-UK Senior Cancer Research Fellow. K.M.B. is supported by a Clinical Investigatorship Award from the CIHR Institute of Genetics.

AUTHOR CONTRIBUTIONS

K.M.B., M.O., W.B.D. and D.E.B. directed the study. M.T.C., L.J.L., C.L.C., J.M.G., D.J.M.-R., T.P., G.A., S.T., S.W., A.H., B.I., A.D., C.D.S., A.R.P., M.W., J.W., S.D., M.T.G., G.M.M., W.B.D. and K.M.B. provided clinical data. L.M.M. performed Sanger sequencing, genotyping studies and variant analysis supervised by K.M.B. and D.E.B. D.A. performed the protein biochemistry and cell biology studies, which were directed by M.O. J.S. and J. Majewski performed exome variant calling analysis. The manuscript was written by L.M.M., G.M.M., M.O. and K.M.B. FORGE Canada Consortium provided the clinical and bioinformatic infrastructure under the direction of K.M.B. assisted by C.L.B. and J. Marcadier. All authors reviewed the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Carter, M.T. *et al.* A new syndrome with multiple capillary malformations, intractable seizures, and brain and limb anomalies. *Am. J. Med. Genet. A.* **155A**, 301–306 (2011).
2. Isidor, B., Barbarot, S., Bénéteau, C., Le Caignec, C. & David, A. Multiple capillary skin malformations, epilepsy, microcephaly, mental retardation, hypoplasia of the distal phalanges: report of a new case and further delineation of a new syndrome. *Am. J. Med. Genet. A.* **155A**, 1458–1460 (2011).
3. Mirzaa, G.M. *et al.* The microcephaly-capillary malformation syndrome. *Am. J. Med. Genet. A.* **155A**, 2080–2087 (2011).
4. Carter, M.T. & Boycot, K.M. Microcephaly-capillary malformation syndrome: a story of rapid emergence of a new recognizable entity. *Am. J. Med. Genet. A.* **155A**, 2078–2079 (2011).

5. Jacobs, A.H. & Walton, R.G. The incidence of birthmarks in the neonate. *Pediatrics* **58**, 218–222 (1976).
6. Eerola, I. *et al.* Capillary malformation-arteriovenous malformation, a new clinical and genetic disorder caused by *RASA1* mutations. *Am. J. Hum. Genet.* **73**, 1240–1249 (2003).
7. Eerola, I. *et al.* KRIT1 is mutated in hyperkeratotic cutaneous capillary-venous malformation associated with cerebral capillary malformation. *Hum. Mol. Genet.* **9**, 1351–1355 (2000).
8. Barash, Y. *et al.* Deciphering the splicing code. *Nature* **465**, 53–59 (2010).
9. Sierra, M.I., Wright, M.H. & Nash, P.D. AMSH interacts with ESCRT-0 to regulate the stability and trafficking of CXCR4. *J. Biol. Chem.* **285**, 13990–14004 (2010).
10. Tsang, H.T.H. *et al.* A systematic analysis of human CHMP protein interactions: additional MIT domain-containing proteins bind to multiple components of the human ESCRT III complex. *Genomics* **88**, 333–346 (2006).
11. Tanaka, N. *et al.* Possible involvement of a novel STAM-associated molecule 'AMSH' in intracellular signal transduction mediated by cytokines. *J. Biol. Chem.* **274**, 19129–19135 (1999).
12. McCullough, J. *et al.* Activation of the endosome-associated ubiquitin isopeptidase AMSH by STAM, a component of the multivesicular body-sorting machinery. *Curr. Biol.* **16**, 160–165 (2006).
13. Agromayor, M. & Martin-Serrano, J. Interaction of AMSH with ESCRT-III and deubiquitination of endosomal cargo. *J. Biol. Chem.* **281**, 23083–23091 (2006).
14. Mizuno, E., Kobayashi, K., Yamamoto, A., Kitamura, N. & Komada, M. A deubiquitinating enzyme UBPY regulates the level of protein ubiquitination on endosomes. *Traffic* **7**, 1017–1031 (2006).
15. Kim, M.S., Kim, J.A., Song, H.K. & Jeon, H. STAM-AMSH interaction facilitates the deubiquitination activity in the C-terminal AMSH. *Biochem. Biophys. Res. Commun.* **351**, 612–618 (2006).
16. Kyuuma, M. *et al.* AMSH, an ESCRT-III associated enzyme, deubiquitinates cargo on MVB/late endosomes. *Cell Struct. Funct.* **31**, 159–172 (2007).
17. Raiborg, C. & Stenmark, H. The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins. *Nature* **458**, 445–452 (2009).
18. Komada, M. Controlling receptor downregulation by ubiquitination and deubiquitination. *Curr. Drug Discov. Technol.* **5**, 78–84 (2008).
19. Wright, M.H., Berlin, I. & Nash, P.D. Regulation of endocytic sorting by ESCRT-DUB–mediated deubiquitination. *Cell Biochem. Biophys.* **60**, 39–46 (2011).
20. Suzuki, S. *et al.* AMSH is required to degrade ubiquitinated proteins in the central nervous system. *Biochem. Biophys. Res. Commun.* **408**, 582–588 (2011).
21. Williams, R.L. & Urbé, S. The emerging shape of the ESCRT machinery. *Nat. Rev. Mol. Cell Biol.* **8**, 355–368 (2007).
22. Sowa, M.E., Bennett, E.J., Gygi, S.P. & Harper, J.W. Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**, 389–403 (2009).
23. Boon, L.M., Mulliken, J.B. & Vikkula, M. RASA1: variable phenotype with capillary and arteriovenous malformations. *Curr. Opin. Genet. Dev.* **15**, 265–269 (2005).
24. Tidyman, W.E. & Rauen, K.A. The RASopathies: developmental syndromes of Ras/MAPK pathway dysregulation. *Curr. Opin. Genet. Dev.* **19**, 230–236 (2009).
25. Ishii, N. *et al.* Loss of neurons in the hippocampus and cerebral cortex of AMSH-deficient mice. *Mol. Cell Biol.* **21**, 8626–8637 (2001).
26. Kato, M., Miyazawa, K. & Kitamura, N. A deubiquitinating enzyme UBPY interacts with the Src homology 3 domain of Hrs-binding protein via a novel binding motif PX(V/I)(D/N)RXXKP. *J. Biol. Chem.* **275**, 37481–37487 (2000).
27. Davies, C.W., Paul, L.N., Kim, M.I. & Das, C. Structural and thermodynamic comparison of the catalytic domain of AMSH and AMSH-LP: nearly identical fold but different stability. *J. Mol. Biol.* **413**, 416–429 (2011).
28. Ma, Y.M. *et al.* Targeting of AMSH to endosomes is required for epidermal growth factor receptor degradation. *J. Biol. Chem.* **282**, 9805–9812 (2007).

ONLINE METHODS

Study participants. All families provided written informed consent, and this study was approved by the ethics review boards at the Children's Hospital of Eastern Ontario, the University of Chicago and Seattle Children's Hospital. We studied a cohort of ten affected individuals from nine families with MIC-CAP syndrome. Genomic DNA was extracted from the whole blood of affected subjects and their family using standard techniques.

Sequencing technology and variant calling pipeline. Using target capture with the Agilent SureSelect 50 Mb All Exon kit (Agilent Technologies, Santa Clara, CA) and sequencing of 100-bp paired-end reads on Illumina HiSeq, we generated over 15 Gb of sequence for each sample such that approximately 90% of the coding bases of the exome defined by the consensus coding sequence (CCDS) project were covered by at least 20 reads. Reads were first quality trimmed from the 3′ end using the Fastx toolkit and were then aligned to hg19 with BWA[29]. Duplicate reads were marked using Picard and excluded from downstream analyses. For each sample, single nucleotide variants (SNVs) and short insertions and deletions (indels) were called using SAMtools pileup and varFilter[30] with the base alignment quality (BAQ) adjustment disabled and were then quality filtered to require that at least 20% of reads supported the variant call. Coverage of the exome was determined using the Genome Analysis Toolkit (GATK). Variants were annotated using both Annovar[31] and custom scripts to identify whether they affected protein-coding sequence and whether they had previously been included in dbSNP131 or in the 1000 Genomes pilot release (Nov. 2010).

Genetic analysis. To elucidate the molecular mechanism of MIC-CAP syndrome in family 4, we PCR amplified ten polymorphic microsatellite markers spanning the length of chromosome 2, four on the short arm and six on the long arm. The amplification products were resolved using the IR^2 DNA Analyzer and interpreted using SAGA software (LI-COR). Analysis of the intronic mutations was performed using the computational model of splicing regulation first described by Barash *et al.*[8]. Real-time PCR was performed using the Mastercycler Realplex (Eppendorf) in the presence of SYBR Green PCR Mastermix reagents (Life Technologies, Applied Biosystems). Standard protocol was followed for the optimization of the real-time PCR primers; however, reactions were scaled to 25 µl per reaction. The PCR conditions were standard, and all reagents, excluding the template and primers, were provided in the SYBR Green PCR Mastermix kit. Standard curves were generated using β-2 microglobulin (NM_004048.2) as a control.

Functional analysis. All antibodies used in this section can be found in **Supplementary Table 3**. STAMBP expression in patient-derived LCLs (P1.1, P7.1 and P3.1) was assessed by protein blotting using anti-STAMBP (H-4) with an epitope directed to amino acids 131–270 of STAMBP (Santa Cruz Biotechnology, Santa Cruz, CA). The caspase 3 antibody was from Cell Signaling Technology (Beverly, MA). For annexin V apoptosis assessment, we used the Single Channel Annexin V Apoptosis Kit (Alexa Fluor 488–conjugated anti-annexin V with SyTOX Green) from Life Technologies LTD (Paisley, UK) according to the manufacturer's instructions. The anti-LC3 was from Cell Signaling (D50G8 XP(R)), and bafilomycin A was from Sigma-Aldrich (Poole, UK). Amounts of active RAS-GTP were determined using the RAS activation assay kit (17-218) from Millipore according to the manufacturer's instructions.

For siRNA-mediated silencing of *STAMBP*, we used ONTARGET plus SMARTpool human STAMBP (L-012202-00-0005) from Dharmacon-Thermo Fisher Scientific (UK) and performed transfection using Metafectene-Pro from Canbio (Cambridge, UK) according to the manufacturer's instructions. Cells were analyzed 24 h after transfection. The SMARTpool is a mixture of four oligonucleotides with distinct target sequences (**Supplementary Table 2**).

For indirect immunofluorescence, LCLs were pelleted, swollen in 75 mM KCl (10 min), immobilized onto polylysine-coated slides by cytospinning (CytoSpin, Shandon), permeabilized (0.1% Triton X-100 in 5% BSA and PBS for 2 min) and blocked in 5% BSA and PBS (10 min) before sequential incubation with primary and secondary antibodies. Slides were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) and preserved in antifade mounting medium (Vectashield). Slides were analyzed using the Zeiss AxioPlan platform, and images were captured using SimplePCI software at constant exposure times. Anticonjugated ubiquitin mouse monoclonal clone FK2 was from Enzo Lifesciences UK LTD (Exeter, UK).

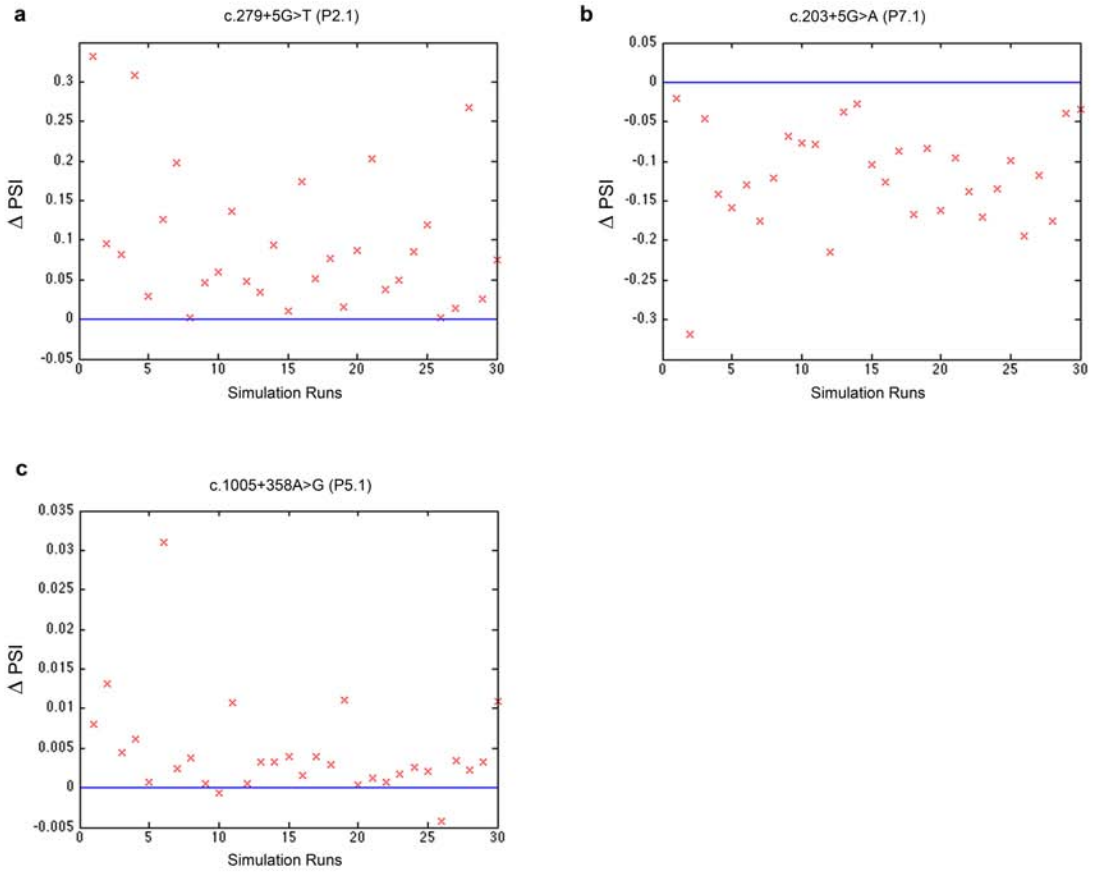
To interrogate RAS-MAPK pathway function, patient-derived LCLs were grown exponentially in complete medium in the presence or absence of fetal bovine serum for 24 h. Antibodies, including phospho-specific antibodies to pC-Raf (Ser33) and pMAPK and pERK1/2 (Thr202 and Tyr204, respectively), along with their corresponding native antibodies, were from Cell Signaling Technology (Beverly, MA). The MEK1/2 inhibitor U0126 was used at 10 µM for 1 h. Whole-cell extracts were prepared by sonication in urea buffer (9 M urea, 50 mM Tris-HCl, pH 7.5, and 10 mM β-mercaptoethanol).

For lentiviral transduction of LCLs, high-titer Precision LentiORF viral particles derived from the pLOC system were obtained from Thermo Scientific (Open Biosystems) and used according to the manufacturer's instructions. Stable STAMBP-expressing clones were obtained after blasticidin S selection of transduced populations.

29. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
30. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
31. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).

2.5 Supplementary

SUPPLEMENTARY FIGURES AND LEGENDS

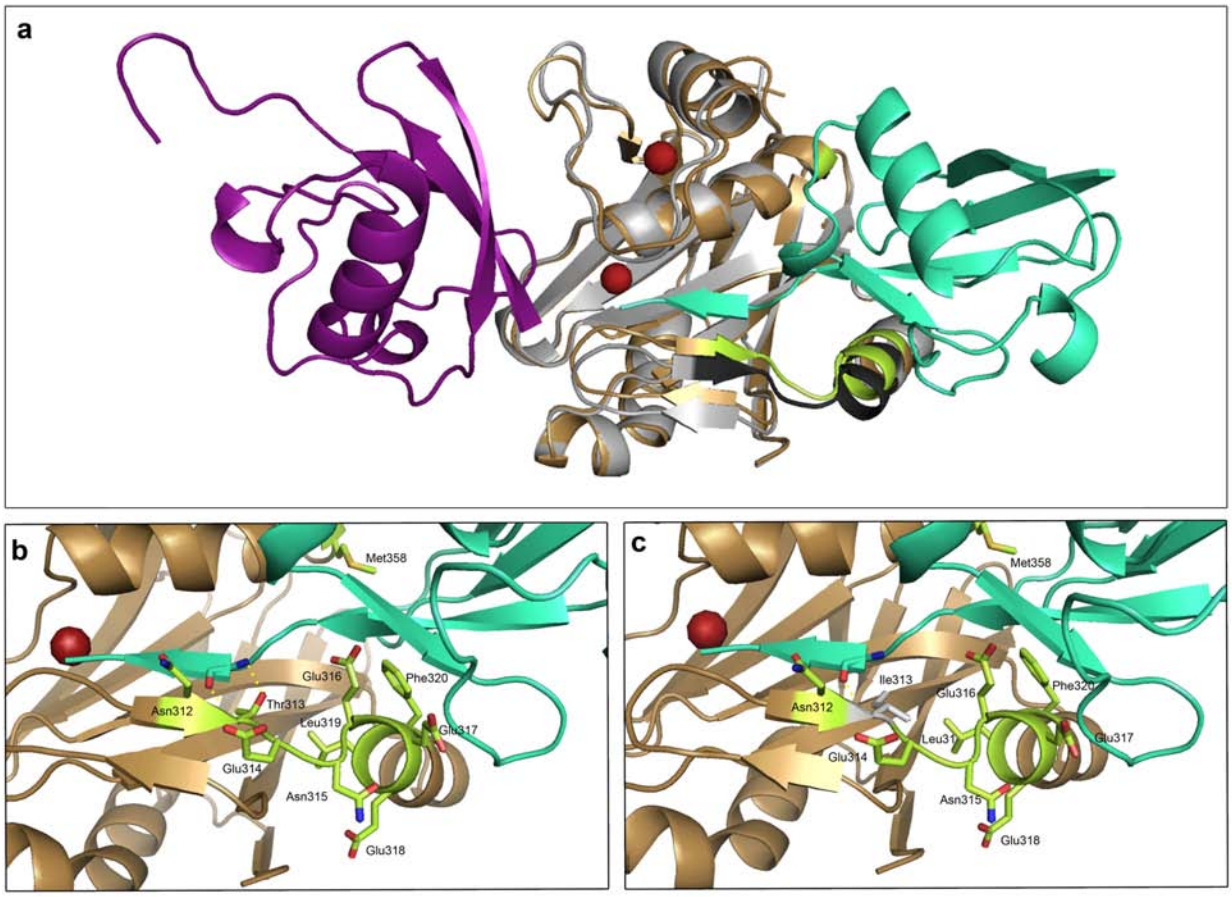


Supplementary Figure 1. Analysis of the intronic MIC-CAP mutations using a computational model of splicing regulation[31]. In order to analyze the effects of intronic mutations on splicing, we use the recently assembled human splicing code that is able to accurately predict the effects of simple mutations (manuscript in preparation), and is built upon the previously published mouse splicing code <http://www.ncbi.nlm.nih.gov/pubmed/20445623>. The splicing code models the Percent Spliced In (PSI) level of each exon as a random distribution between 0 and 1, and the effect of a mutation is quantified by ΔPSI, the difference between its predicted PSI value and that of the WT. We performed 30 random simulations for each of the three intronic

Nature Genetics: doi:10.1038/ng.2602

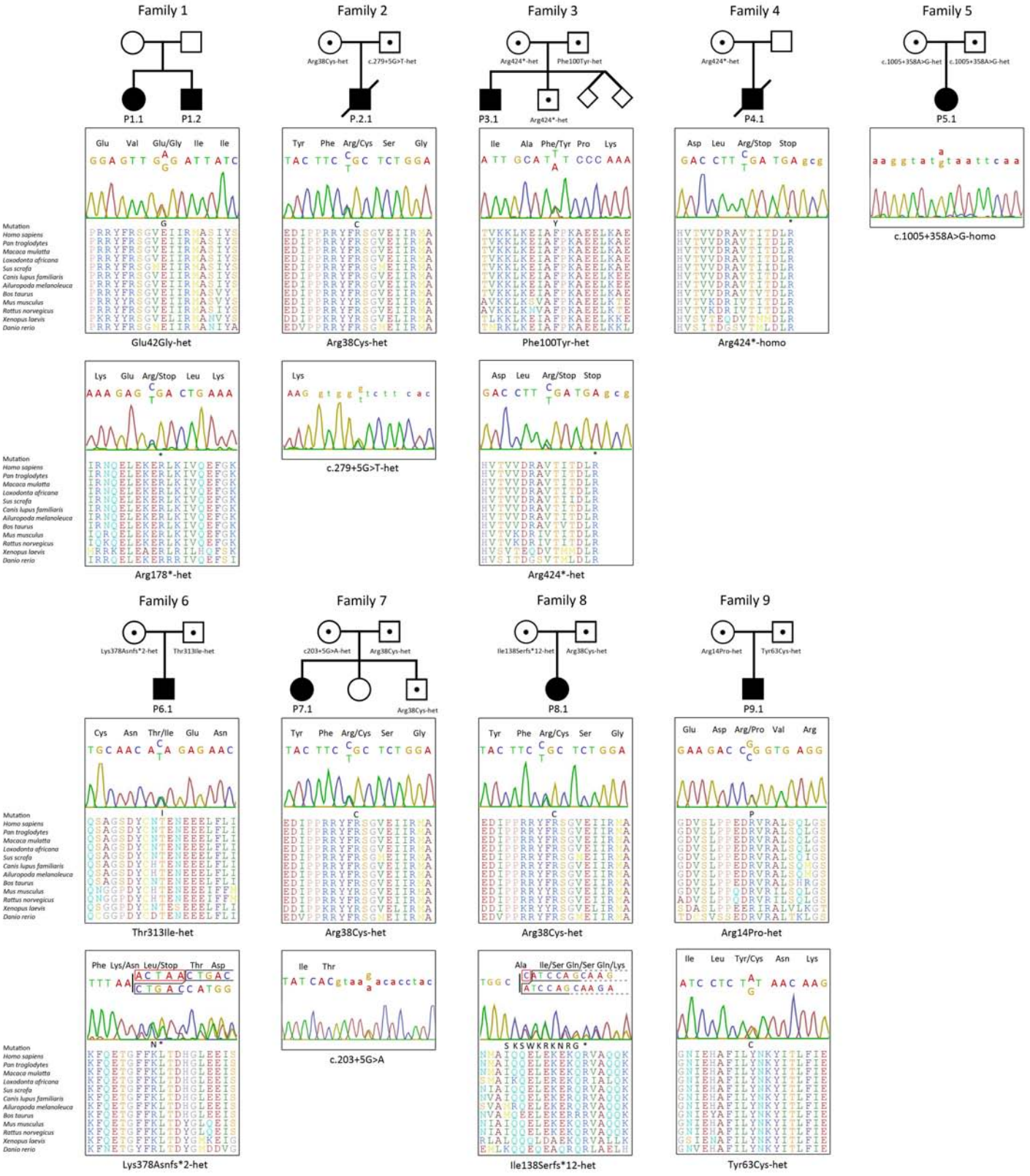
mutations. (**a**) For mutation c.279+5G>T (P2.1) the code predicts that the mutation causes the inclusion of an extra codon in exon 4 with significant confidence (p = 1.86e-9, the sign test), supporting its pathological effect. (**b**), (**c**) For mutations c.203+5G>A (P7.1) and c.1005+358A>G (P5.1) the predicted ΔPSI's are consistent with the experimental evidence (see **Supplementary Fig. 4**) with highly significant p-values (8.68e-7 and 1.86e-9 respectively, the sign test).

Nature Genetics: doi:10.1038/ng.2602



Supplementary Figure 2. Thr313Ile protein modeling. (**a**) Structural comparison of the catalytic JAMM domain. Backbone superimposition of STAMBP/AMSH (gold) (Protein Data Bank (PDB) code: 3RZU) and AMSH-*l*ike *p*rotein (grey) (AMSH-LP) (PDB code: 2ZNV) bound to a K63-linked ubiquitin dimer. The distal ubiquitin recognition site (DUR) highlighted in STAMBP (lime) and AMSH-LP (black) interacts with the distal ubiquitin monomer (teal). Also represented are proximal ubiquitin monomer (purple) and two bound Zn^{2+} atoms (red spheres). Superimposition of the two proteins was achieved using DeepView / Swiss-PdbViewer and visualized using Pymol. (**b**) Expanded view of the conserved residues implicated in distal ubiquitin binding with wild-type Thr313. Hydrogen bonds are shown as yellow dashes. (**c**) Expanded view of the DUR site with Patient 6.1 mutation (Thr313Ile) showing ablation of the H bond between residue 313 and the backbone of the ubiquitin monomer.

Nature Genetics: doi:10.1038/ng.2602

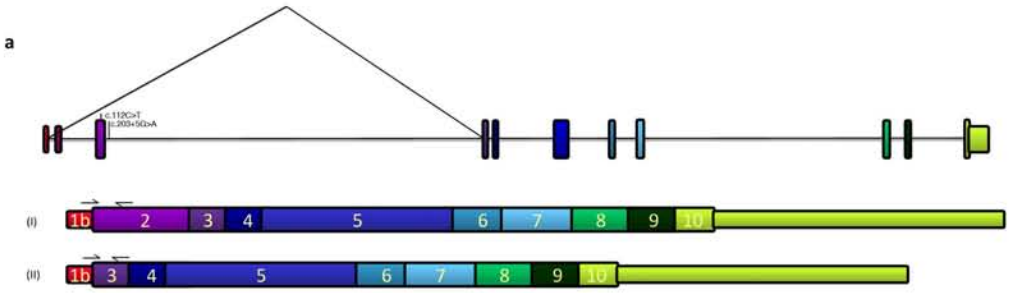


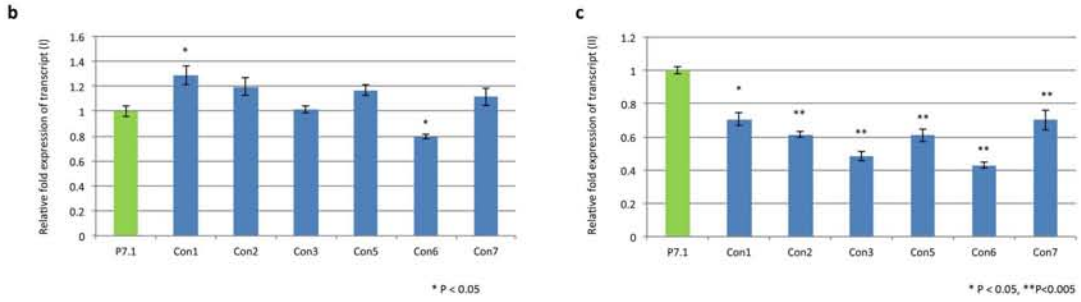
Supplementary Figure 3. Family pedigrees of the nine MIC-CAP families. For the nine families affected individuals are indicated by black square/circles and carriers by dotted

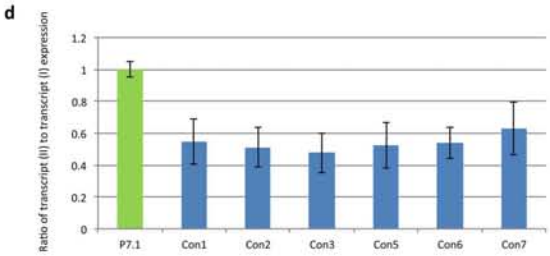
Nature Genetics: doi:10.1038/ng.2602

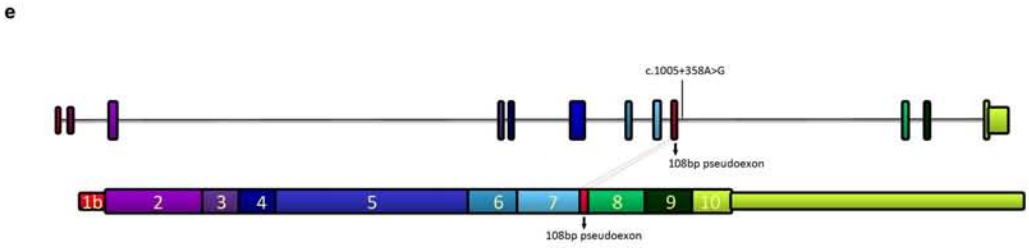
squares/circles. Parental DNA was sequenced and showed appropriate segregation for all families, however, parental DNA for Family 1 was not available for this study. To confirm that MIC-CAP was a recessive trait in Family 1, we cloned, isolated and sequenced STAMP cDNA from P1.2 and showed that c.125A>G and c.532C>T are *in trans* (**data not shown).** Two variants per individual were validated by Sanger sequencing; exceptions are P4.1 and P5.1 whose homozygous mutations are a result of maternal isodisomy and identity-by-descent, respectively. Sequence alignment of STAMBP to 11 orthologues is shown; all variants are found in highly conserved residues.

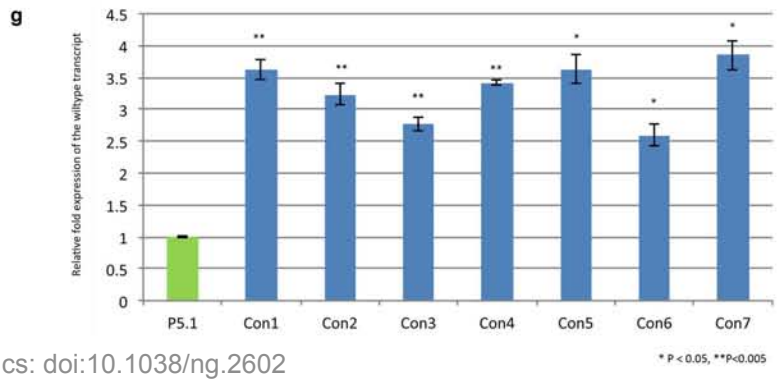
Nature Genetics: doi:10.1038/ng.2602







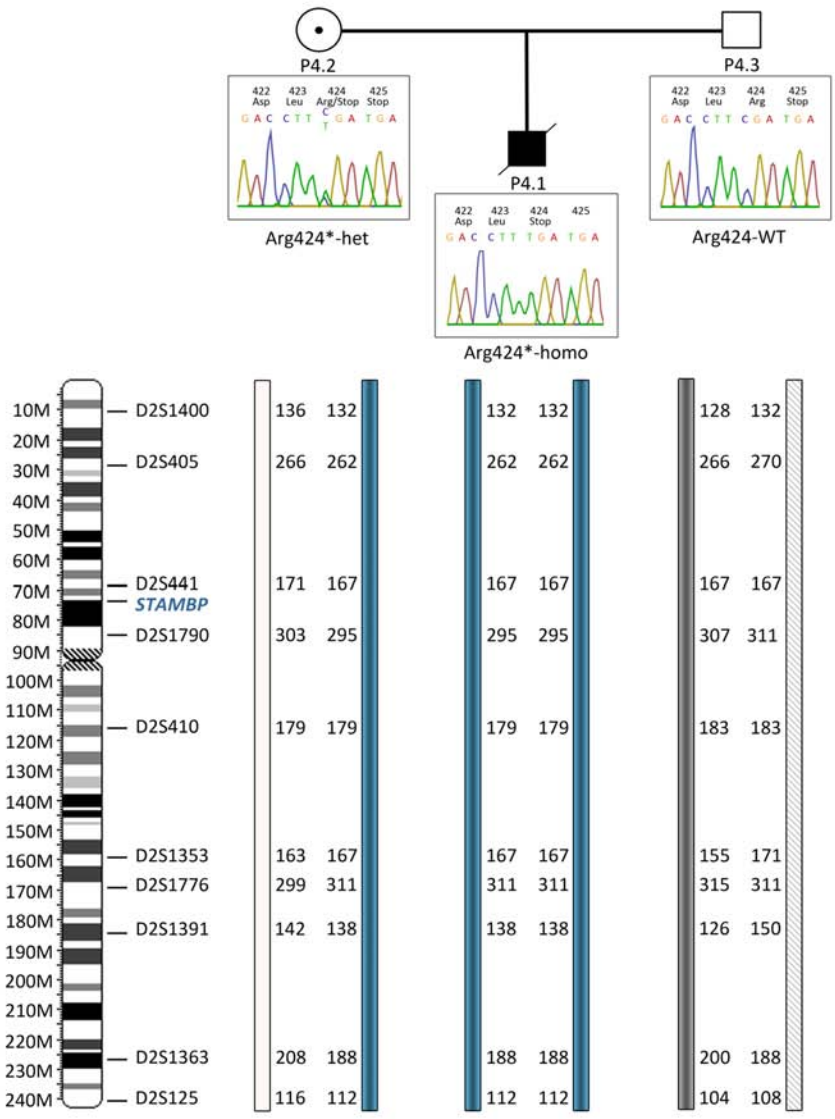




Nature Genetics: doi:10.1038/ng.2602

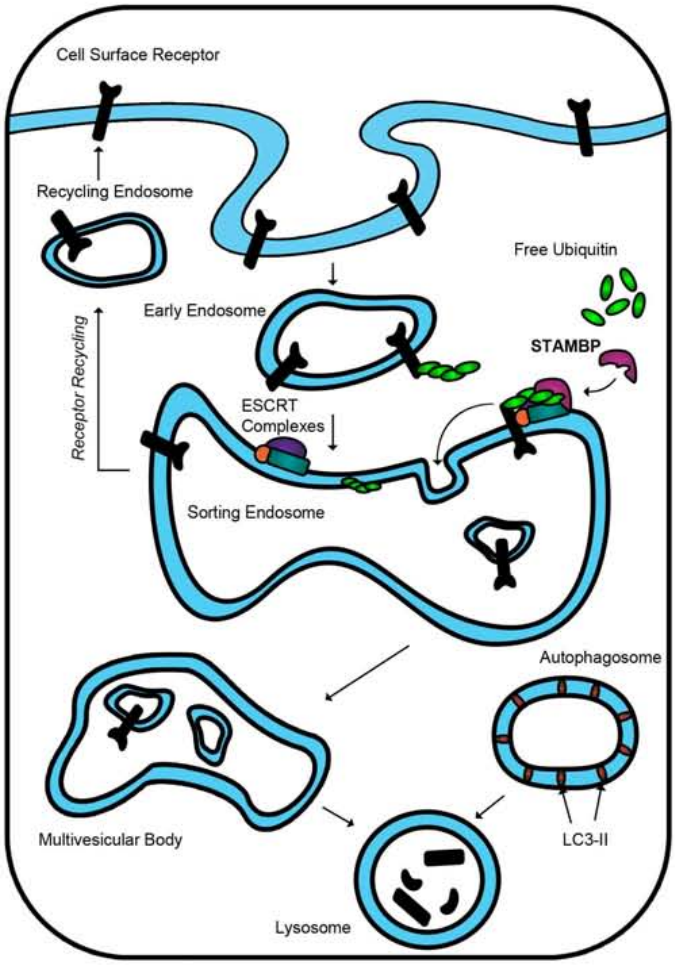
Supplementary Figure 4. Compound heterozygous and homozygous mutations cause MIC-CAP in Patient 7.1 and Patient 5.1, respectively. (**a**) *STAMBP* gene (upper) showing the two mutations identified in P7.1 (c.112C>T and c.203+5G>A) and STAMBP transcripts (lower) showing the full-length transcript (I) and its alternatively spliced isoform (II). To quantify the levels of each transcript, primers were designed to span the exon-exon boundaries between exons 1b-2 and 1b-3 and Real Time PCR (qRT-PCR) was performed using patient derived LCLs. (**b**) and (**c**) qRT-PCR analysis of *STAMBP* transcripts (I) and (II) expression, respectively. The y-axis represents the relative expression ratio between the patient (Patient 7.1) and 6 controls (mean ± SEM for triplicates). (**d**) Ratio of transcript (II) to transcript (I) expression showing a marked increase in skipping of the first coding exon in the patient cells (Patient 7.1) compared to controls. **(e)** A deep homozygous intronic mutation (c.1005+358A>G) leads to the splicing of a 108bp pseudoexon that contains a portion of intron 7 in Patient 5.1. (**f**) Sequencing of the *STAMBP* transcript in Patient 5.1 identifies the 108bp pseudoexon found between exons 7 and 8. This event leads to a premature stop codon 10 amino acids into the sequence. (**g**) Primers were designed to span the exon-exon boundary between exons 7 and 8 and qRT-PCR analysis performed to determine the relative expression of the full-length transcript (mean ± SEM for triplicates). Patient 5.1 cells showed a 3-fold reduction of full-length transcript expression .

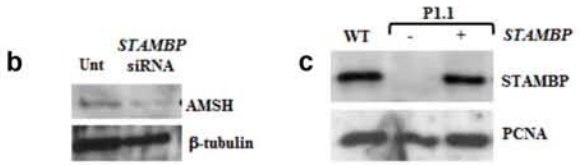
Nature Genetics: doi:10.1038/ng.2602

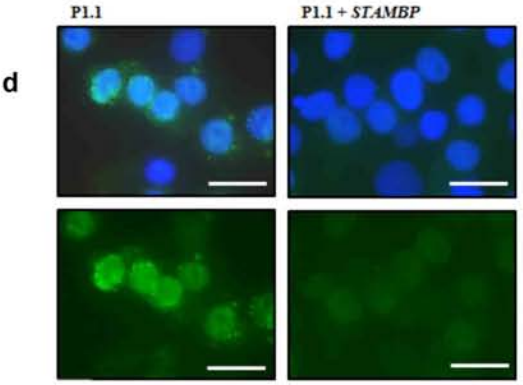


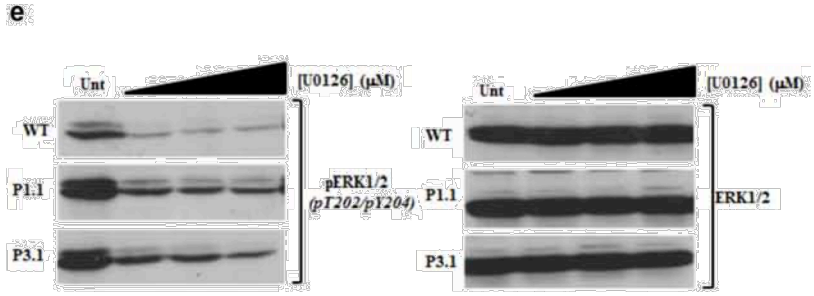
Supplementary Figure 5. Maternal isodisomy unmasks a *STAMBP* mutation in Family 4. Patient (black square) has 10 homozygous microsatellite markers that are maternally derived and thus do not demonstrate paternal contribution. A diagnostic clinical microarray excluded monosomy as the mechanism.

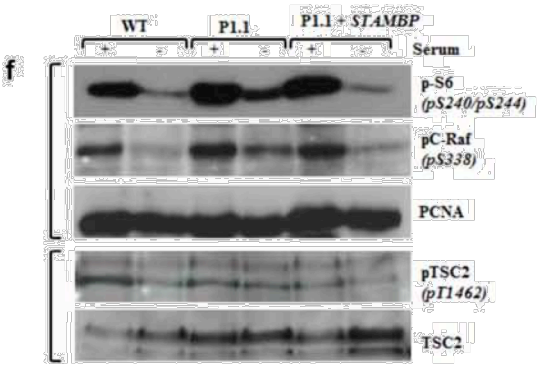
Nature Genetics: doi:10.1038/ng.2602


a
Cell Surface Receptor
Recycling Endosome
Free Ubiquitin
Early Endosome
STAMBP
Receptor Recycling
ESCRT
Complexes
Sorting Endosome
Autophagosome
Multivesicular Body
LC3-II
Lysosome


b
STAMBP
Unt siRNA
AMSH
β-tubulin
c
P1.1
WT - +
STAMBP
STAMBP
PCNA


d
P1.1
P1.1 + STAMBP


e
Unt
[U0126] (μM)
WT
P1.1
P3.1
pERK1/2
(pT202/pT204)
Unt
[U0126] (μM)
WT
P1.1
P3.1
ERK1/2


f
WT
P1.1
P1.1 + STAMBP
Serum
p-S6
(pS240/pS244)
pC-Raf
(pS338)
PCNA
pTSC2
(pT1462)
TSC2

Supplementary Figure 6. (a) Regulation of cell surface receptor distribution and signal transduction occurs via endocytotic processing. Activated cell surface receptors are internalized by invagination of the plasma membrane and incorporated into sorting endosomes. The receptors can then be recycled back to the cell surface or targeted for lysosomal degradation. Ubiquitination of the receptors and ESCRT-machinery represent fundamental post-translational modifications directing the incorporation of the receptors into intralumenal vesicles (ILVs) to form multivesicular bodies that fuse with lysosomes and terminate receptor signaling[21]. STAMBP, is an important DUB with a specificity for Lys63-linked ubiquitin chains, influencing receptor processing by interacting with STAM that together with Hrs, constitute the ESCRT-0 complex[21,26]. STAMBP also interacts with ESCRT-III charged multivesicular body proteins (CHMPs) such as CHMP3, thereby likely functioning at early and late stages of the endocytic pathway[28]. **(b)** STAMBP/AMSH expression in T98G cells 24hrs post-transfection with siRNA SmartPool oilgo's against *STAMBP* (*STAMBP* siRNA). Unt: untransfected. **(c)** P1.1 MIC-CAP patient LCLs were transduced with either pLOC only (-) or pLOC-*STAMBP* (+) lentiviral particles, selected in Blasticidin S (according to the manufacturer's instructions), sub-cloned and probed for STAMBP/AMSH expression, compared to WT. Full complementation was achieved as indicated by STAMBP expression. **(d)** Lentiviral complementation of P1.1 LCLs (P1.1+*STAMBP*) significantly reduced conjugated ubiquitin aggregation compared to P1.1 LCLs upon serum starvation (24hrs) consistent with functional complementation. **(e)** Exponentially growing LCLs were either untreated (Unt) or treated with increasing concentrations (2,4,8 μM) of the MEK1/2 inhibitor U0126 for 1hrs prior to extraction. Whole cell extracts were then blotted for phospho-ERK1/2 using anti-pERK1/2(pT202/pY204) (left hand panels) and re-probed using native anti-ERK1/2 (right hand panels). In contrast to MIC-CAP patient LCLs P1.1 and P3.1,

Nature Genetics: doi:10.1038/ng.2602

even the lowest concentration of U0126 significantly reduced levels of MEK1-dependent phospho-ERK1/2 in WT LCLs. This is consistent with active and insensitive MEK1 activity in the MIC-CAP LCLs. **(f)** Functional lentiviral complementation of P1.1 LCLs (P1.1+*STAMBP*) reconstituted the normal wild-type (WT) response to serum starvation for S6, C-RAF and TSC2 phosphorylation, compared to P1.1 LCLs.

Nature Genetics: doi:10.1038/ng.2602

SUPPLEMENTARY TABLES

Variants identified in 5 MIC-CAP patients (P1.1, P1.2, P2.1, P3.1 and P5.1; Table 1) were compiled and filtered to identify rare indels and nonsynonymous/splicing changes whereby 1722 genes with rare variants were identified. Next, we filtered our list of genes with rare variants looking for genes with rare homozygous or multiple heterozygous (het) variants shared between multiple MIC-CAP patients (P1.1, P1.2, P2.1, P3.1 and P5.1).

Three samples (P1.1, P1.2 and P3.1.) had heterozygous variants in *STAMBP* and *DNHD1* (**Supplementary Table 1**). The latter, which is the 37th most frequent gene seen mutated at our sequencing centre, was discounted from further analysis. Additional analysis of the intronic splice site junctions revealed the second mutation in Patient 2.1.

Supplementary Table 1. Gene filtering for the identification of disease causing *STAMBP* (recessive model)

GENES with **rare homozygous or multiple het** variants present in X individual samples:	**Number of genes**	**Gene names**
5 samples	0	
4 samples	0	
3 samples	2	*STAMBP, DNHD1*
2 samples	39	
1 sample	103	

Nature Genetics: doi:10.1038/ng.2602

Supplementary Table 2. Oligonucleotides used for the amplification, sequencing and siRNA mediated silencing of *STAMBP*

Name	Coding/Non-coding	Forward	Reverse
Stambp_e1 [a]	non-coding	GACTCAGAAAGCCGCAAGCATATC	GTTCGCTCCTCATTGGTCGGATTT
Stambp_e2 [a]	coding	TGAGGGTCTCAGCCTTCTTG	TCTTTGAACGGCAAGGACAG
Stambp_e34 [a]	coding	CCAGCTGGGTGTATTAGCTCTC	AACCCTGAGCCTAACCACTG
Stambp_e5 [a]	coding	CTGTGGCCAGATAGAAAGGG	TAGGTTTCAGCAAAGGCCAC
Stambp_e6 [a]	coding	GGCCCAAGATTGGAGAAGAC	AAGTGACATGAAATTATGCAGTCC
Stambp_e7 [a]	coding	ACAGAGTACCCCAGGCAATG	CAATAGTCAAGTATGCTAAGG
Stambp_e8 [a]	coding	TTTCCTATACATGCATACTTGCG	TCATGATATGGGGCTTAAAAGG
Stambp_e9 [a]	coding	GAAGAGACCTGTGAGGCAGC	TCTATTCCTGTCCCACACTGC
Stambp_e10 [a]	coding	AGGCAGAAGAATCGCTTG	TTTCCTTCACCTTTCCA
Stambp_e10.1 [a]	coding	AGAAATTTGGAAGCCATTTAGA	ACACCATTATTTCTAGGCCACC
Stambp_e10.2 [a]	non-coding	GCCTTTGTTTGGAGTACC	AAACAAAGAGACTTGGTAAA
Stambp_e10.3 [a]	non-coding	ATTTCCTGCTGCAGCCCAGAG	TAGACCATCCTATGCAGAGTGCCT
Stambp_e10.4 [a]	non-coding	TCAGCCAGGTCAACAACATATCA	ACACTGGCCCAAGGGAATGTTTCT
Stambp_e10.5 [a]	non-coding	CCAGGTGCAGAGAGAATCAGAGGA	TAGGAAGTTAAGTGATTTGCTGGTG
Stambp_e10.6 [a]	non-coding	GCCCAGAGTAGAACTTCTGAGTCAA	TGAATGATGCCGGACACACCTGTA
Stambp_e10.7 [a]	non-coding	TCCAGCCCACTCTTGTTTGAACCT	TGGTGTCACTATCTCCTTGTCTCC
Stambp_e10.8 [a]	non-coding	GCCAACAGCGAATTTAATGGGTGG	ACAACCTCCACTTCCCAGGTTCAA
J-012202-05 STAMBP [b]	n/a [c]	GAUGAGCGUUUGAGUCCAA	
J-012202-06 STAMBP [b]	n/a	UAAACUAACUGACCAUGGA	
J-012202-07 STAMBP [b]	n/a	UAUAUCACGCUCUUUAUUG	
J-012202-08 STAMBP [b]	n/a	UCACACAACUGUAAGGCCA	

a. Oligonucleotide used for the amplification and sequencing of *STAMBP*

b. SMARTpool oligonucleotide used for the siRNA mediated silencing of *STAMBP*

c. Abbreviation : n/a: not applicable

Supplementary Table 3. Source, catalogue number and dilution of antibodies used for protein blotting

Antibody Name	Provider (catalogue number)	Dilution used for protein blot
STAMBP Antibody (H-4)	Santa Cruz Biotechnlogy sc-271641	1:1000
Caspase 3	Cell Signaling 9662	1:500
LC3A	Cell Signaling 4599	1:500
pC-RAF (S338)	Cell Signaling 9421	1:1000
C-RAF	Cell Signaling 9422	1:1000
pPI3K p85 (Tyr458)/p55(Tyr199)	Cell Signaling 4228	1:500
PI3K p85	Cell Signaling 4257	1:1000
pERK1/2 (T202/Y204)	Cell Signaling 4377	1:1000
ERK1/2	Cell Signaling 9102	1:1000
pAKT (Thr308)	Cell Signaling 2965	1:1000
AKT	Cell Signaling 4691	1:1000
pTSC2 (Thr1462)	Cell Signaling 3611	1:1000
Tuberin/TSC2	Cell Signaling 3990	1:1000
pS6 (S240/S244)	Cell Signaling 2215	1:1000
β-tubulin	Santa Cruz Biotechnology sc-9104	1:2000

Nature Genetics: doi:10.1038/ng.2602

URL

DeepView / Swiss-PdbViewer, http://spdbv.vital-it.ch/ ; Pymol, http://www.pymol.org/

Nature Genetics: doi:10.1038/ng.2602

Chapter 3: Encephalocraniocutaneous lipomatosis and *FGFR1*

3.1 Preface

The following chapter consists of the manuscript titled "Mosaic Activating Mutations in FGFR1 Cause Encephalocraniocutaneous Lipomatosis" published in the American Journal of Human Genetics by James T. Bennett, Tiong Yang Tan, Diana Alcantara, Martine Tétrault, Andrew E. Timms, Dana Jensen, Sarah Collins, Malgorzata J.M. Nowaczyk, Marjorie J. Lindhurst, Katherine M. Christensen, Stephen R. Braddock, Heather Brandling-Bennett, Raoul C.M. Hennekam, Brian Chung, Anna Lehman, John Su, SuYuen Ng, David J. Amor, University of Washington Center for Mendelian Genomics, Care4Rare Canada Consortium, Jacek Majewski, Les G. Biesecker, Kym M. Boycott, William B. Dobyns, Mark O′Driscoll, Ute Moog and Laura M. McDonell.

3.2 Statement of permission for use of copyrighted material

Statement of permission for the reproduction of copyrighted material from the American Journal of Human Genetics can be found in Appendix C

3.3 Contributions

The specific contributions of each author are as follows:

L. M. McDonell

Analyzed the sequencing data and confirmed causative ECCL mutations. Using Sanger sequencing, validated WES-identified mutations, screened additional tissue. Performed tissue culture and some initial protein biochemistry. Co-wrote manuscript, generated figures and contributed to tables.

James T. Bennett

Provided clinical data. Oversaw cloning and smMIPs studies. Co-wrote manuscript, generated tables and responded to reviewer comments.

Tiong Yang Tan

Contributed clinical data.

Diana Alcantara

Performed protein biochemistry to characterize FGFR1 activity and Ras-MAPK, PI3K-AKT signalling.

Martine Tétrault

Performed bioinformatics analysis of the Illumina sequencing data and generated list of candidate genes.

Andrew E. Timms, Dana Jensen, Sarah Collins, Malgorzata J.M. Nowaczyk, Marjorie J. Lindhurst, Katherine M. Christensen, Stephen R. Braddock, Heather Brandling-Bennett, Raoul C.M. Hennekam, Brian Chung, Anna Lehman, John Su, SuYuen Ng, David J. Amor

Contributed clinical data.

University of Washington Center for Mendelian Genomics

Provided critical infrastructure and resources for this project.

Care4Rare Canada Consortium

Provided critical infrastructure and resources for this project.

Jacek Majewski

Oversaw bioinformatics analysis of the Illumina sequencing data for a subset of patients.

Les G. Biesecker

Directed the study. Contributed clinical data. Oversaw bioinformatics analysis of the Illumina sequencing data for a subset of patients.

Kym M. Boycott

Directed the study. Contributed to editing the manuscript.

William B. Dobyns

Directed the study. Provided clinical data. Contributed to editing the manuscript.

Mark O′Driscoll

Directed the cell biology and protein biochemistry studies. Contributed to writing and editing the manuscript. Oversaw response to reviewer comments.

Ute Moog

Directed the study. Contributed clinical data.

3.4 Main Manuscript

Mosaic Activating Mutations in *FGFR1* Cause Encephalocraniocutaneous Lipomatosis

James T. Bennett,[1,2,18] Tiong Yang Tan,[3,18] Diana Alcantara,[4] Martine Tétrault,[5] Andrew E. Timms,[6] Dana Jensen,[2] Sarah Collins,[2] Malgorzata J.M. Nowaczyk,[7] Marjorie J. Lindhurst,[8] Katherine M. Christensen,[9] Stephen R. Braddock,[9] Heather Brandling-Bennett,[10] Raoul C.M. Hennekam,[11] Brian Chung,[12] Anna Lehman,[13] John Su,[14] SuYuen Ng,[14] David J. Amor,[3] University of Washington Center for Mendelian Genomics, Care4Rare Canada Consortium, Jacek Majewski,[5] Les G. Biesecker,[8] Kym M. Boycott,[15,19] William B. Dobyns,[1,2,16,19] Mark O'Driscoll,[4,19,*] Ute Moog,[17,19,*] and Laura M. McDonell[15,19]

Encephalocraniocutaneous lipomatosis (ECCL) is a sporadic condition characterized by ocular, cutaneous, and central nervous system anomalies. Key clinical features include a well-demarcated hairless fatty nevus on the scalp, benign ocular tumors, and central nervous system lipomas. Seizures, spasticity, and intellectual disability can be present, although affected individuals without seizures and with normal intellect have also been reported. Given the patchy and asymmetric nature of the malformations, ECCL has been hypothesized to be due to a post-zygotic, mosaic mutation. Despite phenotypic overlap with several other disorders associated with mutations in the RAS-MAPK and PI3K-AKT pathways, the molecular etiology of ECCL remains unknown. Using exome sequencing of DNA from multiple affected tissues from five unrelated individuals with ECCL, we identified two mosaic mutations, c.1638C>A (p.Asn546Lys) and c.1966A>G (p.Lys656Glu) within the tyrosine kinase domain of *FGFR1*, in two affected individuals each. These two residues are the most commonly mutated residues in *FGFR1* in human cancers and are associated primarily with CNS tumors. Targeted resequencing of *FGFR1* in multiple tissues from an independent cohort of individuals with ECCL identified one additional individual with a c.1638C>A (p.Asn546Lys) mutation in *FGFR1*. Functional studies of ECCL fibroblast cell lines show increased levels of phosphorylated FGFRs and phosphorylated FRS2, a direct substrate of FGFR1, as well as constitutive activation of RAS-MAPK signaling. In addition to identifying the molecular etiology of ECCL, our results support the emerging overlap between mosaic developmental disorders and tumorigenesis.

Congenital malformations featuring asymmetry, focal anomalies, or segmental overgrowth have long been hypothesized to be due to post-zygotic (mosaic) mutations.[1] Gene discovery for these disorders has been challenging due to the absence of familial recurrence, difficulty obtaining affected tissues, and the challenge of detecting low-frequency genetic variation. Encephalocraniocutaneous lipomatosis (ECCL; [MIM 613001]) is a sporadic neurocutaneous disorder characterized by patchy, asymmetric malformations and absence of familial recurrence.[2] Given this presentation, as well as an equal sex ratio and the occurrence of discordant monozygotic twins, ECCL has been hypothesized to be due to mosaic mutations.[3–5] ECCL is characterized by cutaneous, ocular, and central nervous system (CNS) abnormalities, and in the absence of known genetic cause, diagnosis has been based on the presence of characteristic clinical features.[2,6] The most characteristic skin anomaly in ECCL is nevus psiloliparus, a well-demarcated, alopecic fatty tissue nevus on the scalp seen in 80% of affected individuals.[2] Other dermatologic features include frontotemporal or zygomatic subcutaneous fatty lipomas, non-scarring alopecia, focal dermal hypoplasia or aplasia of the scalp, periocular skin tags, and pigmentary abnormalities following the lines of Blaschko. Choristomas of the eye (epibulbar dermoids or lipodermoids) are also frequent (80% of individuals with ECCL), and can be unilateral or bilateral.[2] Characteristic CNS features in ECCL include intracranial and intraspinal lipomas (61% of affected individuals), and less often cerebral asymmetry, arachnoid cysts, enlarged ventricles, and leptomeningeal angiomatosis.[7] A predisposition to low-grade gliomas has also been observed.[8–12] Seizures and

[1]Department of Pediatrics (Genetics), University of Washington, Seattle, WA 98195, USA; [2]Center for Integrative Brain Research, Seattle Children's Research Institute, Seattle, WA 98101, USA; [3]Victorian Clinical Genetics Services, Murdoch Children's Research Institute, Department of Paediatrics, University of Melbourne, Melbourne, VIC 3052, Australia; [4]Genome Damage and Stability Centre, University of Sussex, Brighton BN19RQ, UK; [5]Department of Human Genetics, McGill University, Montreal, QC H3A0G4 Canada; [6]Center for Developmental Biology and Regenerative Medicine, Seattle Children's Research Institute, Seattle, WA 98101, USA; [7]Department of Pathology and Molecular Medicine, McMaster University, Hamilton, ON L8S 4J9, Canada; [8]National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA; [9]Department of Pediatrics, Cardinal Glennon Children's Medical Center, St. Louis, MO 63104, USA; [10]Departments of Pediatrics and Medicine (Dermatology), University of Washington, Seattle, WA 98195, USA; [11]Department of Pediatrics, Academic Medical Centre, University of Amsterdam, 1105AZ Amsterdam, Netherlands; [12]Department of Paediatrics and Adolescent Medicine, Queen Mary Hospital, University of Hong Kong, 21 Sassoon Road, Hong Kong, China; [13]Department of Medical Genetics, University of British Columbia, Vancouver, BC V6H3N1, Canada; [14]Monash University, Eastern Health, Department of Dermatology, Box Hill, VIC 3128, Australia; [15]Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, ON K1H5B2, Canada; [16]Department of Neurology, University of Washington, Seattle, WA 98195, USA; [17]Institute of Human Genetics, Heidelberg University, 69120 Heidelberg, Germany
[18]These authors contributed equally to this work
[19]These authors contributed equally to this work
*Correspondence: m.o-driscoll@sussex.ac.uk (M.O.), ute.moog@med.uni-heidelberg.de (U.M.)
http://dx.doi.org/10.1016/j.ajhg.2016.02.006.

Table 1. Clinical Features of 5 Individuals with ECCL in Whom an *FGFR1* Mutation Was Detected

	LR12-068	LR13-278	IN_0039	NIH_183	LR14-261
FGFR1 mutation	c.1966A>G (p.Lys656Glu)	c.1638C>A (p.Asn546Lys)	c.1638C>A (p.Asn546Lys)	c.1966A>G (p.Lys656Glu)	c.1638C>A (p.Asn546Lys)
Mutation discovery method	ES	ES	ES	ES	smMIP
Age at last assessment	7 y	15 y	17 m	2 y 8m	5 y
Gender	M	M	M	M	F
Neurocognitive function	normal	delayed, in special skills class	normal	normal	normal
Seizures	no	yes	no	no	no
Intracranial lipomas	yes	yes	yes	no	yes
Spinal lipomas	no	not assessed	yes (T2/3 and L5/S1)	no	no
CNS Other	Pilomyxoid/ pilocytic astrocytoma WHO I	Tectal tumor, left temporal cortical dysplasia	no	Pilocytic/ pilomyxoid astrocytoma WHO II	no
Nevus psiloliparus	yes	yes	yes	yes	yes
Alopecia	yes	yes	yes	yes (right parietal)	yes
Subcutaneous lipoma	yes (fronto-temporal)	yes	yes (fronto-temporal)	yes (parietal)	yes
Focal scalp aplasia	yes	yes	yes	yes	no
Skin tags	yes (eyelid)	yes	yes	yes (right eyelid, anterior to right ear)	yes
Choristoma	yes (bilateral)	yes	yes (right)	no	yes
Coloboma	no	yes (left upper eyelid)	yes (left upper eyelid, iris and bilateral retinal)	no (but segmental iris heterochromia present)	no
Prior Publication	no	yes[59]	no	yes[8]	yes[60]

Abbreviations are as follows: ES (exome sequencing) and mMIP (single molecule molecular inversion probes).

intellectual disability are common but normal intellect is seen in a third of affected individuals.[2] Skeletal manifestations include bone cysts and jaw tumors, such as odontomas, osteomas, and ossifying fibromas.[13] ECCL had been proposed to be a localized form of Proteus syndrome (MIM 176920), although diagnostic criteria suggest that the two conditions are clinically distinct.[2]

To identify the molecular etiology of ECCL, we performed exome sequencing (ES) on DNA samples from five unrelated ECCL probands (IN_0039, LR12-068, LR13-278, LR13-175, NIH_183). Written informed consent to participate in this study was obtained for each participant. This study was approved by ethics review boards at the Children's Hospital of Eastern Ontario, Seattle Children's Hospital, and the National Human Genome Research Institute. Clinical features of these affected individuals are described in Table 1 and highlighted in Figures 1A–1D. To maximize the likelihood of detecting low frequency, tissue-restricted mosaic variants, we sequenced DNA at high coverage (64-172X) from probands' affected and unaffected tissue where possible. ES was also performed on blood-derived DNA from parents of probands LR12-068, LR13-278, LR13-175, NIH_183, and from the unaffected monozygotic twin sibling of IN_0039. ES platforms and data analyses are detailed in Tables S1 and S2. Genomic alterations identified by ES were screened against variants in the NHLBI Exome Sequencing Project Exome Variant Server (EVS), the Exome Aggregation Consortium (ExAC), the NCBI database (dbSNP), and in-house variant databases. Variants inherited from a parent, or present in the unaffected twin in the case of IN_0039, were also filtered out.

Two rare missense variants, c.1638C>A (p.Asn546Lys) and c.1966A>G (p.Lys656Glu), located within the intracellular tyrosine kinase domain of *FGFR1* (NM_023110.2), were identified in four of the five probands (Figures 1E and 1F). In IN_0039, the affected proband of a monozygotic twin pair discordant for ECCL, the p.Asn546Lys substitution was identified in fibroblasts cultured from biopsies of both unaffected skin (23% alternate allele fraction, AAF) and a scalp lesion (33% AAF), but was absent (0/76 reads at this position) from the unaffected twin's blood. In individual LR13-278, the p.Asn546Lys substitution was identified in fibroblasts cultured from biopsies of unaffected skin (35% AAF), scalp nevus (42% AAF), and eyelid dermoid (54% AAF). In proband NIH_183, the p.Lys656Glu substitution was identified in fibroblasts cultured from a scalp lesion (45% AAF) but was not detected in blood. In proband LR12-068, the p.Lys656Glu substitution was identified in fibroblasts

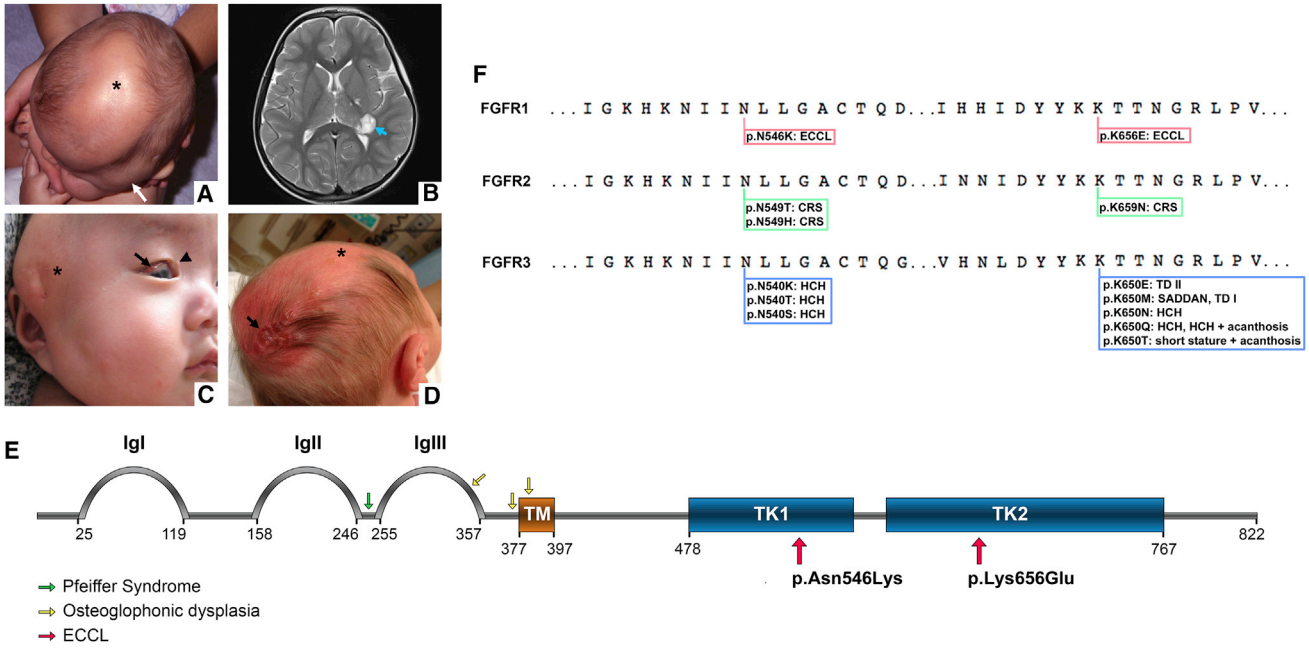


Figure 1. Exome Sequencing Identifies *FGFR1* Mutations in Four Individuals with ECCL

(A) Photograph of LR13-278, showing nevus psiloliparus (asterisk) and subcutaneous lipoma (arrow).
(B) Horizontal T2 MRI of LR12-068, showing pilocytic astrocytoma (light blue arrow) adjacent to posterior left lateral ventricle.
(C) Photograph of IN_0039, showing large subcutaneous lipoma (asterisk), epibulbar dermoid (arrow), and eyelid skin tag (arrowhead).
(D) Photograph of NIH_183 showing several regions of focal skin hypoplasia over vertex (arrow) and nevus psiloliparus anteriorly (asterisk).
(E) Protein structure of FGFR1. The three extracellular Ig-like domains, the transmembrane (TM) domain, and the two-part tyrosine kinase (TK1 and TK2) domain are shown. Locations of mutations for two other syndromes due to activating FGFR1 substitutions are shown: Pfeiffer syndrome in green (p.Pro252Arg) and osteoglophonic dysplasia in yellow (p.Asn330Ile, p.Tyr374Cys, and p.Cys381Arg). The two ECCL associated substitutions (p.Asn546Lys and p.Lys656Glu) are located in the cytoplasmic kinase domain.
(F) Amino acid sequences of FGFR1, 2, and 3 (P11362.3, P22607.1, P21802.1) were aligned using MUSCLE Alignment with the Geneious software.[55] In addition to the two ECCL substitutions in FGFR1, disorders associated with substitutions in paralogous amino acids in FGFR2[41,56] and FGFR3[26,34,57,58] are also shown. Abbreviations: CRS (craniosynostosis), HCH (hypochondroplasia), TD (thanatophoric dysplasia), and SADDAN (Severe Achondroplasia with Developmental Delay and Acanthosis Nigricans)

cultured from a scalp nevus (47% AAF), and from a pilocytic astrocytoma (32% AAF). In each case the *FGFR1* variant detected by exome sequencing was confirmed by Sanger sequencing. Neither of these two variants was present in EVS, ExAC, or dbSNP. No rare non-synonymous variants were identified in *FGFR1* in LR13-175. Coverage information for all eleven exome samples is included in Table S1. On the basis of finding four unrelated individuals with the same rare phenotype who shared one of two missense mutations in the same gene, we considered these variants in *FGFR1* to be pathogenic and causative of ECCL.

ES identified an additional *FGFR1* variant, c.1681G>A (p.Val561Met), in LR12-068, in 45/87 reads (45% AAF). This variant was present in the pilocytic astrocytoma but not in cultured skin fibroblasts (0/183 reads). Interestingly, this variant has been reported to confer resistance to lucitanib, a tyrosine kinase inhibitor (TKI) currently in phase II trials for FGFR-dependent tumors.[14,15] However, to mediate TKI resistance, the c.1681G>A (p.Val561Met) variant must be in *cis* with the primary *FGFR1* activating mutation.[14] We hypothesized that p.Val561Met was a second hit that arose during tumorigenesis in *cis* with this individual's primary *FGFR1* mutation, c.1966A>G (p.Lys656Glu). To test this, we subcloned DNA from the tumor sample. Briefly, a 1,408 basepair fragment containing both c.1681G>A and c.1966A>G was amplified from tumor DNA (primers listed in Table S3), subcloned into a plasmid (pCR2.1-TOPO, Life Technologies) using TOPO-TA cloning, and used to transform competent cells. Colonies containing the fragment were identified by PCR, expanded in liquid culture, and genotyped by Sanger sequencing. Of 20 clones isolated, 16 possessed neither variant, two possessed only the p.Lys656Glu variant, and two possessed both variants. These results suggest that the c.1681G>A (p.Val561Met) variant is in *cis* with the c.1966A>G (p.Lys656Glu) mutation, and possibly arose during tumorigenesis.

To facilitate the identification of mutations in *FGFR1* in additional individuals suspected of having ECCL, we developed an approach using single molecule Molecular Inversion Probes (smMIPs) because low-frequency mosaic mutations could be missed using conventional Sanger sequencing. smMIPs are an inexpensive and highly sensitive next generation sequencing method that have been reported to detect alleles present as low as 0.1%,[16] lower than the typical Sanger cutoff of 20%. smMIPs allows

independent molecular capture events to be distinguished, so that smMIP coverage is reported as independent reads, each of which represents an individual capture event.[16] Briefly, smMIPs were designed to capture all coding regions of *FGFR1* plus at least ten bases of flanking sequence. A pool of 47 smMIPs (sequences in Table S4) was hybridized with 120 ng of DNA from each sample in the cohort. Each smMIP contained a 5 nucleotide degenerate "molecular tag" used to distinguish independent molecular capture events. Sample-specific eight-base barcodes were introduced in subsequent PCR amplification steps, and pooled libraries were sequenced using a 101 cycle paired end protocol on an Illumina MiSeq. Reads were aligned to the human assembly hg19 using BWA, and GATK was used to refine local alignments and call variants (SNVs and indels). Reads with the same molecular barcode were collapsed to form independent reads, and we required the presence of a variant in three or more independent reads. We used smMIPs to screen multiple tissues from two probands (LR13-278 and IN_0039, see Table S5) with mutations in *FGFR1* detected by ES to determine the tissue distribution of the mutations. In LR13-278, the c.1638C>A (p.Asn546Lys) mutation was detected in DNA derived from fibroblasts (affected and unaffected skin), but was absent in blood- or saliva-derived DNA at a depth of 153 and 27 independent reads, respectively. This same mutation was detected in DNA derived from fibroblasts (affected skin) from individual IN_0039, but was absent in saliva, buccal swab, and blood-derived DNA, at depths of 114, 40, and 51 independent reads, respectively. At a depth of 27 independent reads, the smMIPs assay should be able to detect variants at a frequency as low as 11% (3/27). At a depth of 153 independent reads, the detection limit is as low as 2% (3/153). Because we were unable to detect *FGFR1* mutations in blood, saliva, or buccal swab derived DNA in two individuals with known mutations present at high levels (31%–55% AAF, see Table S5) in biopsied tissues, we suspect that the tissue distribution of *FGFR1* mutations in individuals with ECCL is skewed. Although it is possible that the *FGFR1* mutations are present in blood, saliva, or buccal swab at levels below our detection limit, these results suggest that the negative predictive value of *FGFR1* sequencing of these non-biopsied samples might be low for ECCL and that sequencing of skin-biopsy derived DNA will provide a higher diagnostic yield.

Using the same smMIP assay, we screened an independent cohort of four individuals with ECCL (LR14-261, LR04-090, LR09-120, and IN_0025, see Table S5) for whom tissue biopsy-derived DNA was available. We identified one additional individual (LR14-261) with the c.1636C>A (p.Asn546Lys) mutation in *FGFR1*, present at an allele fraction of 55% (110 of 199 independent reads) in DNA isolated from cultured fibroblasts from a scalp nevus, but was not detected in saliva (0/36 independent reads, see Table S5). Clinical details about this individual are listed in Table 1. No other *FGFR1* mutations were identified within any samples from these four individuals in which tissue biopsy-derived DNA was available. An additional group of three individuals (LR04-093, LR09-252, and LR14-210) with ECCL were screened using the smMIP assay, but for these three individuals only blood or saliva derived DNA was available (see Table S5). No additional *FGFR1* mutations were detected in this group, but since we did not have tissue biopsy-derived DNA available in this group, *FGFR1* mutations cannot be excluded. Clinical phenotypes of the individuals in which an *FGFR1* mutation was not detected were not different from those of individuals with an *FGFR1* mutation (data not shown). The number of independent reads at each of the two *FGFR1* mutation sites, for each tissue tested, is shown in Table S5.

Receptor tyrosine kinases (RTKs) regulate a wide range of complex biological functions including cell growth, differentiation, tissue patterning, and organogenesis.[17,18] Fibroblast growth factor receptors (FGFRs) represent an RTK subfamily comprising four homologous receptors encoded by four FGFR genes. The encoded proteins share a basic structure consisting of three extracellular ligand-binding immunoglobulin domains (IgI, IgII, IgIII) linked to a cytoplasmic protein kinase core (TK1 and TK2) via a single-pass transmembrane domain (TM) (Figures 1E and 2A).[19] The two recurrent FGFR1 substitutions are located within the cytoplasmic kinase core (Figures 1E and 1F). FGFRs function by binding their respective ligands and heparan accessory molecules to induce dimerization and conformational changes.[17,20] Following ligand binding, *trans*-phosphorylation of the cytoplasmic domains between dimer pairs releases *cis*-autoinhibition and enables catalytic kinase activity.[20–22] Phosphorylation of additional tyrosine sites in the kinase domain creates high affinity binding sites for proteins containing phosphotyrosine binding (PTB) domains and Src-homology 2 domains.[21] Catalytically active receptors initiate intracellular signaling through several pathways, including the RAS-MAPK network (Figure 2A), resulting in phosphorylation of downstream targets such as ERK1, ERK2, and C-RAF (HUGO gene names are *MAPK3*, *MAPK1*, and *RAF1*, respectively).

To determine the effect of ECCL mutations on FGFR activity, we conducted Western blot analysis of whole cell extracts from several fibroblast lines derived from LR13-278, who harbors the p.Asn546Lys substitution. Using antibodies that detect phosphorylation of FGFR1-4 on Tyr653 and Tyr654 (pFGFR-Y653/Y654), we observed spontaneously elevated levels of phosphorylated FGFRs in exponentially growing fibroblasts derived from the skin, eyelid, and scalp of LR13-278, compared to wild-type (WT) cells (Figure 2B). We next examined signal transduction in these cells following prolonged serum deprivation, compared to exponentially growing cells. WT fibroblasts showed the expected reduction in phosphorylation of FGFR (Figure 2C) and ERK1/2 phosphorylation (pERK1/2-T202/Y204) upon serum starvation (Figure 2C). In contrast, fibroblasts from LR13-278 exhibited elevated phosphorylation of FGFR and ERK1/2

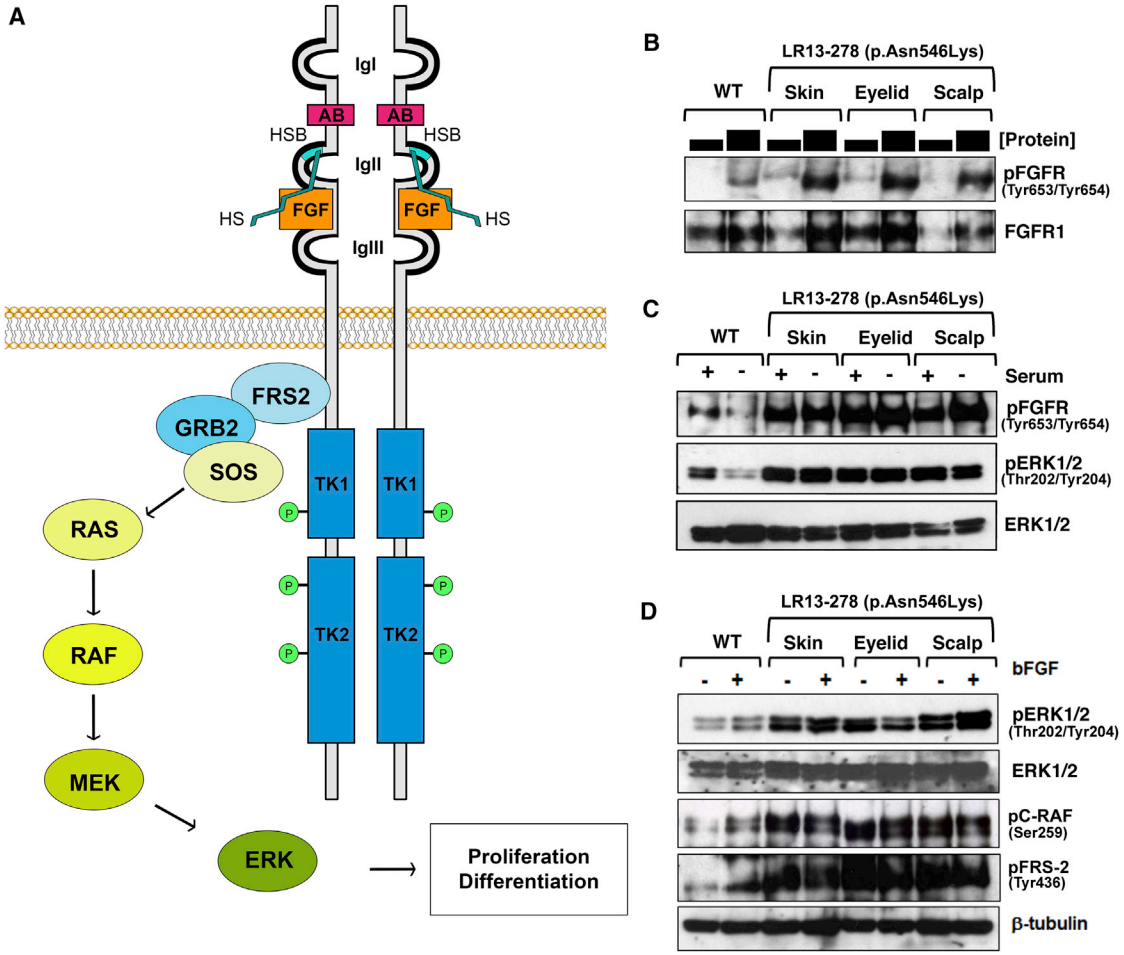


Figure 2. Hyperphosphorylation of FGFR and RAS-MAPK Activation in an Individual with ECCL Due to p.Asn546Lys Substitution
(A) Ligand and heparan-sulfate binding induces FGFR dimerization and conformational changes followed by *trans*-phosphorylation and activation of the cytoplasmic kinase domain. Phosphorylation of additional tyrosine sites in the kinase domain creates high affinity binding sites for downstream effector proteins such as FRS2, which recruits GRB2 and initiates RAS-MAPK signaling. The three extracellular Ig-like domains, the acid box (AB), heparan-sulfate (HS), heparan-sulfate binding site (HSB), fibroblast growth factor (FGF), and the two-part tyrosine kinase (TK1 and TK2) domain are shown.
(B) Differing amounts of whole-cell extract (WCE) from exponentially growing wild-type (WT) and ECCL fibroblasts derived from various tissues from LR13-278 were probed for FGFR-phosphorylation using pan-FGFR phosphorylation antibodies (pFGFR-Tyr653/Tyr654).
(C) WCE was prepared from exponentially growing cells (+ serum) and from cells that were serum starved for 72 hr (− serum) from wild-type (WT) fibroblasts and from various tissues from LR13-278. These were blotted using antibodies to detect pan-FGFR transphosphorylation and ERK1/2 phosphorylation.
(D) All fibroblasts were serum starved for 72 hr and then either untreated (−) or treated (+) with bFGF (10 nM for 15 min). WCE from wild-type (WT) and LR13-278 fibroblasts from various tissues were blotted to detect ERK1/2 and C-RAF phosphorylation and also for FRS2 phosphorylation. All antibodies were obtained from Cell Signaling Technology: anti-pFGFR-Tyr653/Tyr654 (Cat #3471S), anti-FGFR-1 (Cat #9740), anti-pFRS2-Tyr463 (Cat #3861S), anti-pERK1/2-Thr202/Tyr204 (Cat #9101S), anti-ERK1/2 (Cat #4695S), and anti-pC-RAF-Ser259 (Cat #9421).

compared to WT in the presence or absence of serum (Figure 2C). Similar results were observed in fibroblasts derived from the thigh and scalp of IN_0039 (data not shown). Finally, we examined FGFR-dependent signal transduction in LR13-278 fibroblasts in response to acute treatment with recombinant basic fibroblast growth factor (bFGF) following prolonged serum deprivation. WT fibroblasts treated with bFGF showed elevated levels of phosphorylated ERK1/2 and C-RAF, another RAS-pathway effector (Figure 2D). In contrast, fibroblasts from LR13-278 showed elevated levels of phosphorylated ERK1/2 and C-RAF even in the absence of bFGF stimulation (Figure 2D), suggesting ligand-independent activation of FGFR signaling. Because phosphorylated ERK1/2 and C-RAF can reflect increased activity of a variety of RTKs, we also examined FRS2, whose activating phosphorylation is mainly FGFR-dependent.[23] Similar to ERK1/2 and C-RAF, phosphorylated FRS2 is increased by bFGF stimulation in WT fibroblasts, but in LR13-278 elevated levels of phosphorylated FRS2 are present even in the absence of bFGF stimulation (Figure 2D). Collectively, these results demonstrate elevated autophosphorylation of FGFRs, the

FGFR-dependent substrate FRS2, and the RAS-pathway components C-RAF and ERK1/2, in multiple proband-derived fibroblasts with the p.Asn546Lys substitution (Figures 2B–2D and data not shown). A proband-derived fibroblast line harboring the p.Lys656Glu substitution was unavailable for this study.

We have shown that mosaic, activating substitutions at two residues (p.Asn546Lys and p.Lys656Glu) in the cytoplasmic tyrosine kinase domain of *FGFR1* cause ECCL. The involvement of FGFRs in human disease is well documented.[24,25] Germline gain-of-function mutations in *FGFRs* cause craniosynostosis (*FGFR1-3*)[26–31] and skeletal dysplasia (*FGFR1* and *3*),[32–34] while loss-of-function mutations cause hypogonadotrophic hypogonadism (*FGFR1*, [MIM 615465]) and Hartsfield syndrome (*FGFR1*, [MIM 615465]).[35,36] Lacrimoauriculodentodigital syndrome [MIM 149730] is caused by mutations in *FGFR2*, *FGFR3*, and *FGF10*,[37] and somatic activating mutations in *FGFR3* are present in some epidermal nevi.[38] Both activating mutations and whole gene amplification of *FGFR1* contribute to the pathogenesis of cancer.[24,39] Although activating mutations in the tyrosine kinase domain of *FGFR2* and *FGFR3* have been reported,[25] this is the first report, to our knowledge, of activating mutations in this domain in *FGFR1* associated with a developmental disorder.

Strikingly, the mutations identified in this study in *FGFR1* are paralogous to mutations in *FGFR2* and *FGFR3* that cause craniosynostosis and skeletal dysplasia (Figure 1F).[17,25] The p.Lys650Glu substitution in FGFR3 causes thanatophoric dysplasia II (MIM 187601), and is paralogous to the ECCL-associated p.Lys656Glu substitution in FGFR1.[34] The p.Asn540Lys substitution in FGFR3, paralogous to p.Asn546Lys in FGFR1, is the most common cause of hypochondroplasia (MIM 1460000).[40] Similarly, paralogous substitutions of Asn549 and Lys659 in FGFR2 have been reported in individuals with syndromic craniosynostosis.[41] The identification, in individuals with ECCL, of amino acid substitutions in FGFR1 that are identical to substitutions in other FGF receptors provides additional support for the pathogenicity of these variants, and highlights the distinct roles FGFR1, 2, and 3 signaling during human development.

The findings presented here highlight an emerging link between recurrent somatic activating mutations in tumors and mosaic developmental disorders that frequently have an increased risk of cancer.[42] ECCL represents the first known example of a developmental disorder in the FGFR family with an increased risk for cancer, specifically low-grade gliomas.[8–12] RTKs are one of the most commonly mutated gene families in cancer and their contribution to tumorigenesis is widely recognized.[43] Not surprisingly, both the c.1638C>A (p.Asn546Lys) and c.1966A>G (p.Lys656Glu) mutations in *FGFR1* are known oncogenic mutations,[44–47] and are the two most commonly mutated residues among *FGFR1* mutation-containing tumors in the COSMIC (Catalogue of Somatic Mutations in Cancer) database.[48] Interestingly, most of the tumors associated with substitutions in these two residues are central nervous system gliomas, including pilocytic astrocytomas,[48,49] the same type of tumor seen at increased frequency in individuals with ECCL. In the pilocytic astrocytoma sample from LR12-068, ES identified a second missense substitution, p.Val561Met, also in the tyrosine kinase domain and in *cis* with the p.Lys656Glu substitution. Previous studies have shown that p.Val561Met confers a 38-fold increase in phosphorylation of the FGFR1 receptor, as well as resistance to lucitanib, an FGFR inhibitor currently in phase II clinical trials for FGFR-dependent tumors.[14,15] Whether the p.Val561Met substitution actively contributes to tumorigenesis remains to be elucidated. In individuals with ECCL who develop low-grade gliomas, knowledge of causative *FGFR1* mutations could lead to informed treatment choices with targeted RTK inhibitors and improved clinical management.

The RAS-MAPK pathway regulates crucial cellular processes including DNA synthesis, cell growth, and differentiation. Mutations in components of this pathway cause a variety of developmental syndromes.[50] Oculoectodermal syndrome (OES; [MIM 600268]) is characterized by congenital abnormalities of the scalp (cutis aplasia and focal alopecia) and eyes (eyelid skin tags and epibulbar dermoids), features that are also seen in ECCL.[51] OES has been proposed to be a milder form of ECCL, which is distinguished from OES by the presence of CNS lipomas.[51] Notably, somatic mutations in *KRAS* have recently been associated with OES.[52] Considering the striking phenotypic overlap between ECCL and OES, hyperactive RAS-MAPK signaling might represent a common mechanism underlying these two disorders. The absence of CNS lipomas in OES could be due to the relatively small number of individuals with OES who have had brain imaging, or could reflect the tissue distribution of these somatic mutations. Specific differences in pathway activation due to mutations in *KRAS* versus *FGFR1* might also play a role. Sequencing of *KRAS* in individuals with ECCL, and *FGFR1* in individuals with OES, will be helpful in addressing this question.

In summary, we identified two recurrent mutations in *FGFR1* in individuals with ECCL, a rare neurocutaneous disorder. We developed a smMIP assay to facilitate screening of individuals with suspected ECCL and showed that DNA derived from fibroblasts provides the highest yield for identification of mutations in *FGFR1*. We identified a total of five *FGFR1* individuals with *FGFR1* mutations within our cohort of nine individuals for whom biopsy-derived fibroblast DNA was available. We did not detect any mutations among three individuals for whom only blood- or saliva-derived DNA was available, but this does not rule out the possibility of an *FGFR1* mutation in other tissues. Potential explanations for the individuals in the cohort for whom an *FGFR1* mutation was not detected include (1) mutations present at a level below the limit of detection of our smMIP assay, (2) underlying locus

heterogeneity, and (3) absence of available biopsy-derived DNA for testing. With the exception of the brain tumor from individual LR12-068, all of the samples that possessed an *FGFR1* mutation were from cultured fibroblasts, so that the mutation levels detected might reflect selection for activating *FGFR1* mutations in cell culture. This might explain why the level of mutation in DNA derived from individual LR12-068's brain tumor (32%) is lower than that of his scalp nevus (47%). Sequencing of DNA from uncultured tissue samples from individuals with ECCL will help address this issue. The phenotypes of the individuals without detectable *FGFR1* mutations do not differ significantly from the individuals with *FGFR1* mutations (data not shown). Given the phenotypic similarities between OES and ECCL, screening these individuals for *KRAS* mutations is a logical next step. Our functional analysis of fibroblast cell lines harboring the p.Asn546Lys substitution showed hyperphosphorylation of FGFRs and downstream dependent substrates, consistent with elevated activation of the receptor. Interestingly, elevated FGFR1 signaling is implicated in both proliferation of human mesenchymal stem cells and human preadipocytes and might explain the striking nevus psiloliparus seen in individuals with ECCL.[53,54] We do not currently understand how activating mutations in a single gene can cause ECCL, craniosynostosis, and skeletal dysplasias. It seems likely that the developmental timing and tissue specific location of the post-zygotic *FGFR1* mutation might play an important role. Clearly different activating mutations in *FGFR1* can lead to distinct phenotypes, and further studies are needed to understand the pleiotropic effects of gain-of-function mutations in *FGFR1*. Finally this work adds another gene to the growing number of disorders due to mosaic mutations impacting the RAS-MAPK pathway and further supports the emerging overlap between mosaic developmental disorders and tumorigenesis.

Supplemental Data

Supplemental Data include five tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2016.02.006.

Acknowledgments

We are grateful to the individuals and families who participated in this study. Research reported in this publication was supported by the Care4Rare Canada Consortium (Enhanced Care for Rare Genetic Diseases in Canada) funded by Genome Canada, the Canadian Institutes of Health Research, the Ontario Genomics Institute, Ontario Research Fund, Genome Quebec, and Children's Hospital of Eastern Ontario Foundation; and by the National Institute of Neurological Disorders and Stroke (NINDS) under award numbers 1R01NS092772 (W.B.D.) and the National Human Genome Research Institute of the National Institutes of Health. Some sequencing in this study was provided by the University of Washington Center for Mendelian Genomics (UW-CMG) and was funded by the National Human Genome Research Institute and the National Heart, Lung and Blood Institute grant U54HG006493 to Drs. Debbie Nickerson, Jay Shendure, and Michael Bamshad. M.J.L. and L.G.B. are supported by the Intramural Research Program of the National Human Genome Research Institute. We also acknowledge the contributions of the high-throughput sequencing platform of the McGill University and Genome Quebec Innovation Centre, Montreal, Canada. M.T. received a post-doctoral fellowship from the Reseau de Medecine Genique Appliquee. L.M.M. is supported by a scholarship from the (CIHR) and Consortium National de Formation en Santé. T.Y.T. was supported by an Australian National Health and Medical Research Council Postdoctoral Overseas Training Scholarship (#607431). D.A. and M.O'D. are supported by Cancer Research UK. None of the authors have any conflicts of interest to declare.

Received: November 4, 2015
Accepted: February 9, 2016
Published: March 3, 2016

Web Resources

The URLs for data presented here are as follows:

COSMIC, http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/
ExAC Browser, http://exac.broadinstitute.org/
Mutalyzer, https://mutalyzer.nl/index
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
OMIM, http://www.omim.org/
SeattleSeq Annotation 137, http://snp.gs.washington.edu/SeattleSeqAnnotation137/
UCSC Genome Browser, http://genome.ucsc.edu
FGFR1: NM_023110.2

References

1. Happle, R. (1986). Cutaneous manifestation of lethal genes. Hum. Genet. *72*, 280.
2. Moog, U. (2009). Encephalocraniocutaneous lipomatosis. J. Med. Genet. *46*, 721–729.
3. Hall, J.G. (1988). Review and hypotheses: somatic mosaicism: observations related to clinical genetics. Am. J. Hum. Genet. *43*, 355–363.
4. Hamm, H. (1999). Cutaneous mosaicism of lethal mutations. Am. J. Med. Genet. *85*, 342–345.
5. Happle, R. (1993). Mosaicism in human skin. Understanding the patterns and mechanisms. Arch. Dermatol. *129*, 1460–1470.
6. Hunter, A.G. (2006). Oculocerebrocutaneous and encephalocraniocutaneous lipomatosis syndromes: blind men and an elephant or separate syndromes? Am. J. Med. Genet. A. *140*, 709–726.
7. Moog, U., Jones, M.C., Viskochil, D.H., Verloes, A., Van Allen, M.I., and Dobyns, W.B. (2007). Brain anomalies in encephalocraniocutaneous lipomatosis. Am. J. Med. Genet. A. *143A*, 2963–2972.
8. Bieser, S., Reis, M., Guzman, M., Gauvain, K., Elbabaa, S., Braddock, S.R., and Abdel-Baki, M.S. (2015). Grade II pilocytic astrocytoma in a 3-month-old patient with encephalocraniocutaneous lipomatosis (ECCL): case report and literature review of low grade gliomas in ECCL. Am. J. Med. Genet. A. *167A*, 878–881.

9. Brassesco, M.S., Valera, E.T., Becker, A.P., Castro-Gamero, A.M., de Aboim Machado, A., Santos, A.C., Scrideli, C.A., Oliveira, R.S., Machado, H.R., and Tone, L.G. (2010). Low-grade astrocytoma in a child with encephalocraniocutaneous lipomatosis. J. Neurooncol. *96*, 437–441.
10. Kocak, O., Yarar, C., and Carman, K.B. (2015). Encephalocraniocutaneous lipomatosis, a rare neurocutaneous disorder: report of additional three cases. Childs Nerv. Syst., 1–4. Published online August 1, 2015.
11. Phi, J.H., Park, S.H., Chae, J.H., Wang, K.C., Cho, B.K., and Kim, S.K. (2010). Papillary glioneuronal tumor present in a patient with encephalocraniocutaneous lipomatosis: case report. Neurosurgery *67*, E1165–E1169.
12. Valera, E.T., Brassesco, M.S., Scrideli, C.A., de Castro Barros, M.V., Santos, A.C., Oliveira, R.S., Machado, H.R., and Tone, L.G. (2012). Are patients with encephalocraniocutaneous lipomatosis at increased risk of developing low-grade gliomas? Child's nervous system: ChNS: official journal of the International Society for Pediatric Neurosurgery *28*, 19–22.
13. Moog, U., Roelens, F., Mortier, G.R., Sijstermans, H., Kelly, M., Cox, G.F., Robson, C.D., and Kimonis, V.E. (2007). Encephalocraniocutaneous lipomatosis accompanied by the formation of bone cysts: Harboring clues to pathogenesis? Am. J. Med. Genet. A. *143A*, 2973–2980.
14. Sohl, C.D., Ryan, M.R., Luo, B., Frey, K.M., and Anderson, K.S. (2015). Illuminating the molecular mechanisms of tyrosine kinase inhibitor resistance for the FGFR1 gatekeeper mutation: the Achilles' heel of targeted therapy. ACS Chem. Biol. *10*, 1319–1329.
15. Soria, J.C., DeBraud, F., Bahleda, R., Adamo, B., Andre, F., Dienstmann, R., Delmonte, A., Cereda, R., Isaacson, J., Litten, J., et al. (2014). Phase I/IIa study evaluating the safety, efficacy, pharmacokinetics, and pharmacodynamics of lucitanib in advanced solid tumors. Ann. Oncol. *25*, 2244–2251.
16. Hiatt, J.B., Pritchard, C.C., Salipante, S.J., O'Roak, B.J., and Shendure, J. (2013). Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation. Genome Res. *23*, 843–854.
17. Beenken, A., and Mohammadi, M. (2009). The FGF family: biology, pathophysiology and therapy. Nat. Rev. Drug Discov. *8*, 235–253.
18. Dorey, K., and Amaya, E. (2010). FGF signalling: diverse roles during early vertebrate embryogenesis. Development *137*, 3731–3742.
19. Mohammadi, M., Olsen, S.K., and Ibrahimi, O.A. (2005). Structural basis for fibroblast growth factor receptor activation. Cytokine Growth Factor Rev. *16*, 107–137.
20. Heldin, C.H. (1995). Dimerization of cell surface receptors in signal transduction. Cell *80*, 213–223.
21. Lemmon, M.A., and Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. Cell *141*, 1117–1134.
22. Schlessinger, J. (2000). Cell signaling by receptor tyrosine kinases. Cell *103*, 211–225.
23. Gotoh, N. (2008). Regulation of growth factor signaling by FRS2 family docking/scaffold adaptor proteins. Cancer Sci. *99*, 1319–1325.
24. McDonell, L.M., Kernohan, K.D., Boycott, K.M., and Sawyer, S.L. (2015). Receptor tyrosine kinase mutations in developmental syndromes and cancer: two sides of the same coin. Hum. Mol. Genet. *24* (R1), R60–R66.
25. Wilkie, A.O. (2005). Bad bones, absent smell, selfish testes: the pleiotropic consequences of human FGF receptor mutations. Cytokine Growth Factor Rev. *16*, 187–203.
26. Bellus, G.A., Gaudenz, K., Zackai, E.H., Clarke, L.A., Szabo, J., Francomano, C.A., and Muenke, M. (1996). Identical mutations in three different fibroblast growth factor receptor genes in autosomal dominant craniosynostosis syndromes. Nat. Genet. *14*, 174–176.
27. Jabs, E.W., Li, X., Scott, A.F., Meyers, G., Chen, W., Eccles, M., Mao, J.I., Charnas, L.R., Jackson, C.E., and Jaye, M. (1994). Jackson-Weiss and Crouzon syndromes are allelic with mutations in fibroblast growth factor receptor 2. Nat. Genet. *8*, 275–279.
28. Muenke, M., Gripp, K.W., McDonald-McGinn, D.M., Gaudenz, K., Whitaker, L.A., Bartlett, S.P., Markowitz, R.I., Robin, N.H., Nwokoro, N., Mulvihill, J.J., et al. (1997). A unique point mutation in the fibroblast growth factor receptor 3 gene (FGFR3) defines a new craniosynostosis syndrome. Am. J. Hum. Genet. *60*, 555–564.
29. Muenke, M., Schell, U., Hehr, A., Robin, N.H., Losken, H.W., Schinzel, A., Pulleyn, L.J., Rutland, P., Reardon, W., Malcolm, S., et al. (1994). A common mutation in the fibroblast growth factor receptor 1 gene in Pfeiffer syndrome. Nat. Genet. *8*, 269–274.
30. Reardon, W., Winter, R.M., Rutland, P., Pulleyn, L.J., Jones, B.M., and Malcolm, S. (1994). Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. Nat. Genet. *8*, 98–103.
31. Wilkie, A.O., Slaney, S.F., Oldridge, M., Poole, M.D., Ashworth, G.J., Hockley, A.D., Hayward, R.D., David, D.J., Pulleyn, L.J., Rutland, P., et al. (1995). Apert syndrome results from localized mutations of FGFR2 and is allelic with Crouzon syndrome. Nat. Genet. *9*, 165–172.
32. Rousseau, F., Bonaventure, J., Legeai-Mallet, L., Pelet, A., Rozet, J.M., Maroteaux, P., Le Merrer, M., and Munnich, A. (1994). Mutations in the gene encoding fibroblast growth factor receptor-3 in achondroplasia. Nature *371*, 252–254.
33. Shiang, R., Thompson, L.M., Zhu, Y.Z., Church, D.M., Fielder, T.J., Bocian, M., Winokur, S.T., and Wasmuth, J.J. (1994). Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. Cell *78*, 335–342.
34. Tavormina, P.L., Shiang, R., Thompson, L.M., Zhu, Y.Z., Wilkin, D.J., Lachman, R.S., Wilcox, W.R., Rimoin, D.L., Cohn, D.H., and Wasmuth, J.J. (1995). Thanatophoric dysplasia (types I and II) caused by distinct mutations in fibroblast growth factor receptor 3. Nat. Genet. *9*, 321–328.
35. Dodé, C., Levilliers, J., Dupont, J.M., De Paepe, A., Le Dû, N., Soussi-Yanicostas, N., Coimbra, R.S., Delmaghani, S., Compain-Nouaille, S., Baverel, F., et al. (2003). Loss-of-function mutations in FGFR1 cause autosomal dominant Kallmann syndrome. Nat. Genet. *33*, 463–465.
36. Simonis, N., Migeotte, I., Lambert, N., Perazzolo, C., de Silva, D.C., Dimitrov, B., Heinrichs, C., Janssens, S., Kerr, B., Mortier, G., et al. (2013). FGFR1 mutations cause Hartsfield syndrome, the unique association of holoprosencephaly and ectrodactyly. J. Med. Genet. *50*, 585–592.
37. Rohmann, E., Brunner, H.G., Kayserili, H., Uyguner, O., Nürnberg, G., Lew, E.D., Dobbie, A., Eswarakumar, V.P., Uzumcu, A., Ulubil-Emeroglu, M., et al. (2006). Mutations in different components of FGF signaling in LADD syndrome. Nat. Genet. *38*, 414–417.

38. Hafner, C., Hartmann, A., and Vogt, T. (2007). FGFR3 mutations in epidermal nevi and seborrheic keratoses: lessons from urothelium and skin. J. Invest. Dermatol. *127*, 1572–1573.
39. Weiss, J., Sos, M.L., Seidel, D., Peifer, M., Zander, T., Heuckmann, J.M., Ullrich, R.T., Menon, R., Maier, S., Soltermann, A., et al. (2010). Frequent and focal FGFR1 amplification associates with therapeutically tractable FGFR1 dependency in squamous cell lung cancer. Sci. Transl. Med. *2*, 62ra93.
40. Prinos, P., Costa, T., Sommer, A., Kilpatrick, M.W., and Tsipouras, P. (1995). A common FGFR3 gene mutation in hypochondroplasia. Hum. Mol. Genet. *4*, 2097–2101.
41. Kan, S.H., Elanko, N., Johnson, D., Cornejo-Roldan, L., Cook, J., Reich, E.W., Tomkins, S., Verloes, A., Twigg, S.R., Rannan-Eliya, S., et al. (2002). Genomic screening of fibroblast growth-factor receptor 2 reveals a wide spectrum of mutations in patients with syndromic craniosynostosis. Am. J. Hum. Genet. *70*, 472–486.
42. Bellacosa, A. (2013). Developmental disease and cancer: biological and clinical overlaps. Am. J. Med. Genet. A. *161A*, 2788–2796.
43. Blume-Jensen, P., and Hunter, T. (2001). Oncogenic kinase signalling. Nature *411*, 355–365.
44. Cancer Genome Atlas Research Network (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature *455*, 1061–1068.
45. Hart, K.C., Robertson, S.C., Kanemitsu, M.Y., Meyer, A.N., Tynan, J.A., and Donoghue, D.J. (2000). Transformation and Stat activation by derivatives of FGFR1, FGFR3, and FGFR4. Oncogene *19*, 3309–3320.
46. Lew, E.D., Furdui, C.M., Anderson, K.S., and Schlessinger, J. (2009). The precise sequence of FGF receptor autophosphorylation is kinetically driven and is disrupted by oncogenic mutations. Sci. Signal. *2*, ra6.
47. Rand, V., Huang, J., Stockwell, T., Ferriera, S., Buzko, O., Levy, S., Busam, D., Li, K., Edwards, J.B., Eberhart, C., et al. (2005). Sequence survey of receptor tyrosine kinases reveals mutations in glioblastomas. Proc. Natl. Acad. Sci. USA *102*, 14344–14349.
48. Forbes, S.A., Tang, G., Bindal, N., Bamford, S., Dawson, E., Cole, C., Kok, C.Y., Jia, M., Ewing, R., Menzies, A., et al. (2010). COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. Nucleic Acids Res. *38*, D652–D657.
49. Jones, D.T., Hutter, B., Jäger, N., Korshunov, A., Kool, M., Warnatz, H.J., Zichner, T., Lambert, S.R., Ryzhova, M., Quang, D.A., et al.; International Cancer Genome Consortium PedBrain Tumor Project (2013). Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat. Genet. *45*, 927–932.
50. Schubbert, S., Shannon, K., and Bollag, G. (2007). Hyperactive Ras in developmental disorders and cancer. Nat. Rev. Cancer *7*, 295–308.
51. Ardinger, H.H., Horii, K.A., and Begleiter, M.L. (2007). Expanding the phenotype of oculoectodermal syndrome: possible relationship to encephalocraniocutaneous lipomatosis. Am. J. Med. Genet. A. *143A*, 2959–2962.
52. Peacock, J.D., Dykema, K.J., Toriello, H.V., Mooney, M.R., Scholten, D.J., 2nd, Winn, M.E., Borgman, A., Duesbery, N.S., Hiemenga, J.A., Liu, C., et al. (2015). Oculoectodermal syndrome is a mosaic RASopathy associated with KRAS alterations. Am. J. Med. Genet. A. *167*, 1429–1435.
53. Dombrowski, C., Helledie, T., Ling, L., Grünert, M., Canning, C.A., Jones, C.M., Hui, J.H., Nurcombe, V., van Wijnen, A.J., and Cool, S.M. (2013). FGFR1 signaling stimulates proliferation of human mesenchymal stem cells by inhibiting the cyclin-dependent kinase inhibitors p21(Waf1) and p27(Kip1). Stem Cells *31*, 2724–2736.
54. Widberg, C.H., Newell, F.S., Bachmann, A.W., Ramnoruth, S.N., Spelta, M.C., Whitehead, J.P., Hutley, L.J., and Prins, J.B. (2009). Fibroblast growth factor receptor 1 is a key regulator of early adipogenic events in human preadipocytes. Am. J. Physiol. Endocrinol. Metab. *296*, E121–E131.
55. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics *28*, 1647–1649.
56. Wilkie, A.O., Bochukova, E.G., Hansen, R.M., Taylor, I.B., Rannan-Eliya, S.V., Byren, J.C., Wall, S.A., Ramos, L., Venâncio, M., Hurst, J.A., et al. (2006). Clinical dividends from the molecular genetic diagnosis of craniosynostosis. Am. J. Med. Genet. A. *140*, 2631–2639.
57. Berk, D.R., Spector, E.B., and Bayliss, S.J. (2007). Familial acanthosis nigricans due to K650T FGFR3 mutation. Arch. Dermatol. *143*, 1153–1156.
58. Wilcox, W.R., Tavormina, P.L., Krakow, D., Kitoh, H., Lachman, R.S., Wasmuth, J.J., Thompson, L.M., and Rimoin, D.L. (1998). Molecular, radiologic, and histopathologic correlations in thanatophoric dysplasia. Am. J. Med. Genet. *78*, 274–281.
59. Nowaczyk, M.J., Mernagh, J.R., Bourgeois, J.M., Thompson, P.J., and Jurriaans, E. (2000). Antenatal and postnatal findings in encephalocraniocutaneous lipomatosis. Am. J. Med. Genet. *91*, 261–266.
60. Kupsik, M., and Brandling-Bennett, H. (2013). An infant with an alopecic plaque on the scalp and ocular choristomas: case presentation. Diagnosis: Encephalocraniocutaneous lipomatosis (ECCL). Pediatr. Dermatol. *30*, 491–492.

3.5 Supplementary

SUPPLEMENTAL DATA

TABLE S1. Capture methods and coverage summary of exome data

	LR12-068		LR13-278			LR13-175		IN_0039		NIH_183	
	Tumor*	Scalp nevus	Unaffected skin	Scalp nevus	Eyelid dermoid	Scalp nevus	Lipoma	Scalp nevus	Unaffected skin	Scalp nevus	Blood*
Capture Method	SeqCap EZ Exome Library v2.0 kit							SureSelect All Exon V5		SeqCap EX Exome+UTR	
Mean Coverage	122X	160X	161X	160X	150X	169X	172X	106X	105X	65X	64X
% Covered > 20X	95.9	97.7	97.3	97.1	96.6	97.8	97.9	94.7	94.7	83.1	84.4
c.1638C>A (p.Asn546Lys)	0% (47)	0% (98)	35% (74)	42% (99)	54% (92)	0% (93)	0% (105)	33% (76)	23% (61)	0% (24)	0% (39)
c.1966A>G (p.Lys656Glu)	32% (127)	47% (182)	0% (219)	0% (172)	0% (181)	0% (202)	0% (205)	0% (70)	0% (90)	45% (29)	0% (40)

All DNA isolated from cultured fibroblasts cultured from biopsied tissue except those with asterisk (*), in which DNA was directly isolated from tissue without culture.

TABLE S2: Exome sequencing and variant filtering pipelines

Sample	**LR12-068, LR13-278, LR13-175**	**IN_0039**	**NIH_183**
Sequencing platform	HiSeq 2000 (Illumina)	HiSeq 2000 (Illumina)	HiSeq 2500 (Illumina)
Sequence alignment	Burrows-Wheeler Aligner	Burrows-Wheeler Aligner	Novoalign
Variant calling & annotation	Unified Genotyper, SeattleSeq Annotation Server	GATK, SAMtools, BCFtools, custom script	Shimmer, Mutect, Somatic Sniper
Filtering	missense,nonsense, & splice variants with < 1% in EVS, ExAC, & dbSNP	missense,nonsense, & splice variants with < 1% in EVS, ExAC, & dbSNP. Not seen in > 5 in-house exomes	missense, nonsense and splice variants with < 2% ClinSeq™ frequency*

*ClinSeq™ frequency is defined as the number of individuals with alternative allele frequency ≥1%, divided by the number of individuals with at least ten reads at that position. This is a population frequency based filter that is not limited to constitutional variants (as is the case with EVS, EXAC, and dbSNP), and is based on the NIH in house ClinSeq dataset (www.genome.gov/25521305)

TABLE S3: Primers used for subcloning 1408 basepair fragment containing c.1681G>A (p.Val561Met) and c.1966A>G (p.Lys656Glu)

Name	**Sequence**
FGFR1_ex14_F	CTTTGAGGTGAAGCCAAACC
FGFR1_ex15_R	ACCCCACTCCTTGCTTCTC

TABLE S4: Sequences for *FGFR1* smMIPs

Name	Sequence
FGFR1_01	GAGCTCTGGCTCTGGCACGGGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGGTGTCGGGAAAGCTGGGGG
FGFR1_02	CACGCCCTCCCCAGACTCCACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTTCCCGACACCCGGAGCTCTACGT
FGFR1_03	GGGCCCCTCCTCCCTGCTCAGGGAGGTGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNATGAGAGAAGACGGAA
FGFR1_04	CCCACTGCGTGCACGCACCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTACATGATGATGCGGGACTGCTGGC
FGFR1_05	CGTCTCCTGGAGATGGATACTCTCTAGTCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCCTGGTTGGAGGTCAA
FGFR1_06	GCAAATGGGCGGAGAGCCACAGGGTGTTACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGTGCAGTTACTGGG
FGFR1_07	TGAGCCAGGCCTTGGGGCACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGGAGATCTTCACTCTGGGCGG
FGFR1_08	CGCATGGACAAGCCCAGTAACTGCACCAACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAACACCCTGTGGCTCT
FGFR1_09	TGGCCCCAGGCAGGGCCATGACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCTCTATCCCACACCTCCCTGGCA
FGFR1_10	CCTCTGTCACCAGGACATTCCTGGCTGCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCACTCCTTGCTTCTCA
FGFR1_11	CGCACGGGACATTCACCACATCGACTACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCAGAGCCTTCCAGCTC
FGFR1_12	GGGTTGTGGCTGGGGTTGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAAGACTAGGGGGGCTCTGTTCCCAC
FGFR1_13	AGCAGCTCTCCTCCAAGGACCTGGTGTCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCTTCTCCTGTGCCTG
FGFR1_14	CCACCCCCAGCAGCACACCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGGGCCTCCGGGCCTGCAGGTACTCC
FGFR1_15	GCTAGGGAAGGGGGTTAAGAGAGGCTGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGGAAGCATAAGAATA
FGFR1_16	CGCAGGATGGTGGGTGCCGGCCAGACTGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCAAGTAAATGAGTCTCA
FGFR1_17	CCCATTCCAAGCAAACAGCAGGCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGGAGAGGGCTGCTTTGGG
FGFR1_18	CGTGTGACCAAAGTGGCTGTGAAGATGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAACTTCACCAGCCCCAA
FGFR1_19	CGAACCAGAAGAACCCCAGAGTTCATGGACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAATGCCTTCAAAAAGT
FGFR1_20	GCAAGGAGGGGGGACGGGGTGACTCCTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCATACTCAGAGACCC
FGFR1_21	TGCACACTCAGCACCACCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNATCTCCTGCATGGTGGGGTCGGTCATC
FGFR1_22	GGTACCAAGAAGAGTGACTTCCACAGCCACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCACTGACTCAGCCCTG
FGFR1_23_SNPa	AGGCC**G**GCAGTGATGACCTCGCCCCTGTACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCATGGTTCTTCTCCCT
FGFR1_23_SNPb	AGGCC**A**GCAGTGATGACCTCGCCCCTGTACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCATGGTTCTTCTCCCT
FGFR1_24_SNPa	CGTGCC**C**GTGGCGAGGGCAGGACATCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCTAGGGAAGCTCTTCTC
FGFR1_24_SNPb	CGTGCC**T**GTGGCGAGGGCAGGACATCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCTAGGGAAGCTCTTCTC
FGFR1_25	GGGGAGACACAGAGGCAGGAGAGCTGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTGCTTGGCGGGTAAC
FGFR1_26	CGGAAGCAAAATGGACAAGCACAGGACCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGAGTGATGGGAGAGTC
FGFR1_27	CGTCACTGGGGCTTTGGGGTCAGCTACACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNACACACACTCCATCTCA
FGFR1_28	AGTAACAGAGGTCACAAAGTGGAGGTGAGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGGAAGGAGACCACT
FGFR1_29	CGGGGGCTCAAGTTCCTGTGGGCAGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGCACATCCAGTGGCTAAAG
FGFR1_30	CGGTGAGGGGACCGCTCTGTGGACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAATCTTGCTCCCATTCACCTC
FGFR1_31	GGGTGGGCTCACCTGCGCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTACACATGAACTCCACGTTGCTACC
FGFR1_32	AAGAGCCAGGCTTGGAGAACACAGCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGACTCTGTGGTGCCCT
FGFR1_33	GGCAACTACACCTGCATTGTGGAGACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAGGTGCCACGGGGTGC
FGFR1_34	AGGGGAGGCCGAGTTAGGAAGTCCTGATTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCAAATGCCCTTCCAGT
FGFR1_35	GGGACCCCAAACCCCACACCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAGCTGCTCCTCTCCACCCTGCCT
FGFR1_36	GGGAAGAAGAAGGGGCACTGAGGTTCCTCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCAGACCCAAAGGGCAG
FGFR1_37	GGACACCCTCCCCATGGGGATCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTTCTCCTCTGAAGAGGAGTCA
FGFR1_38_SNPa	GGGCAC**G**GAGTCCTGCACCTCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGCAGAGAGGGCTGGAGGGGG
FGFR1_38_SNPb	GGGCAC**A**GAGTCCTGCACCTCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGCAGAGAGGGCTGGAGGGGG
FGFR1_39	TGCTCTGCACATCGTCCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGTGTCACTGCCCGAGGGGCTGCTG
FGFR1_40	TCAACTGGCTGCGGGACGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCTCTCGCCCCTTGGCTTCCCTTC
FGFR1_41	CGTAATAAAAAAACCTCTGCAGAGGGCCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTTTGGGGCTCTCTCC
FGFR1_42	GGGCAGCCTGGACTCCTGGGCCTTGGGACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCACCAACCTCTAACT
FGFR1_43	GGATGTGGAGCTGGAAGTGCCTCCTCTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCCCCTACCTAGACCCT
FGFR1_44	CCCTCTGATGAGTGGGAAACTGAGATGTGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCTATCTCACCAGACAC

Sequences of all 47 smMIPs used in this study are listed. The string of five N's represents the degenerate molecular tag. Three smMIPs overlapped a common SNP, so smMIPs complimentary to both alleles (in red) were used, and labeled "a" and "b"

TABLE S5: Coverage depth at the two *FGFR1* mutation sites for each sample sequenced by smMIPs

Cohort	Individual	Tissue	c.1638C>A (p.Asn546Lys)	c.1966A>G (p.Lys656Glu)
Exome sequencing	**LR13-278**	Unaffected Skin	50/160 (31%)	0/22
		Scalp Nevus	52/108 (48%)	0/12
		Eyelid Dermoid	42/76 (55%)	low coverage
		Blood	0/153	0/29
		Saliva	0/27	0/28
	IN_0039	Scalp Nevus	30/83 (36%)	low coverage
		Saliva	0/114	low coverage
		Buccal	0/40	low coverage
		Blood	0/51	low coverage
Tissue biopsy available	**LR14-261**	Scalp Nevus	110/199 (55%)	0/67
		Saliva	0/36	low coverage
	LR04-090	Unaffected skin	0/119	0/22
		Saliva	0/228	0/36
		Blood	0/119	0/23
	LR09-120	Scalp Nevus	0/35	0/19
		Saliva	0/124	0/24
	IN_0025	Lipoma	0/117	0/22
		Blood	0/211	0/26
		Saliva	0/94	0/25
Blood/Saliva Only	**LR04-093**	Blood	0/152	0/39
	LR09-252	Saliva	0/105	0/28
	LR14-210	Blood	0/227	0/49

Low coverage was defined as less than 10 independent reads

Chapter 4: Encephalocraniocutaneous lipomatosis and *KRAS*

4.1 Preface

The following chapter consists of the manuscript titled "Mosaic KRAS mutation in a patient with encephalocraniocutaneous lipomatosis and renovascular hypertension" submitted to the American Journal of Medical Genetics, Part A by Laura M. McDonell, Gordon Ka-Chun Leung, Hussein Daoud, Janice Ip, Stella Chim, Ho Ming Luk, Lawrence La, Care4Rare Canada Consortium, Kym M. Boycott and Brian Hon-Yin Chung.

4.2 Statement of permission for use of copyrighted material

This manuscript has been submitted to the American Journal of Medical Genetics, Part A. As such, an embargo will be applied at time of final submission.

4.3 Contributions

The specific contributions of each author are as follows:

Laura M. McDonell

Analyzed the sequencing data and confirmed causative identified KRAS mutation. Validated with Sanger, and screened additional tissue and parents. Performed tissue culture. Wrote the manuscript and generated the figures.

Gordon Ka-Chun Leung

Contributed clinical data

Hussein Daoud

Performed high-throughput sequencing and generated bioinformatics data.

Janice Ip, Stella Chim, Ho Ming Luk, Lawrence La

Contributed clinical data.

Care4Rare Canada Consortium

Provided critical infrastructure and resources for this project.

Kym M. Boycott

Oversaw study. Edited the manuscript.

Brian Hon-Yin Chung

Oversaw study, contributed clinical data. Edited the manuscript.

4.4 Main Manuscript

Mosaic *KRAS* mutation in a patient with encephalocraniocutaneous lipomatosis and renovascular hypertension

AUTHORS

Laura M.McDonell[1], Gordon Ka-Chun Leung[2], Hussein Daoud[3], Janice Ip[4], Stella Chim[2], Ho Ming Luk[5], Lawrence Lan[6], Care4Rare Canada Consortium, Kym M. Boycott[1,7] and Brian Hon-Yin Chung[2,7]

AFFILIATIONS

1. Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, Ontario, Canada

2. Department of Paediatrics and Adolescent Medicine, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China

3. Department of Genetics, Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada

4. Department of Radiology, Queen Mary Hospital, Hong Kong, China

5. Clinical Genetic Service, Department of Health, Hong Kong, China

6. Division of Paediatric Surgery, Department of Surgery, The University of Hong Kong, Hong Kong, China

7. Co-corresponding authors

CO-CORRESPONDING AUTHORS:

Dr. Brian Hon-Yin Chung,

Department of Paediatrics and Adolescent Medicine,

102 Pokfulam Road

Hong Kong, China

Email: bhychung@hku.hk

Dr. Kym M. Boycott,

Department of Genetics

Children's Hospital of Eastern Ontario

401 Smyth Road

Ottawa, Ontario, K1H 8L1, Canada

Email: kboycott@cheo.on.ca

KEY WORDS

Neurocutaneous Syndromes/diagnosis; Lipomatosis/genetics; Genes, Ras; Proto-oncogene; Mutation, Missense; Hypertension, Renovascular; Renal Artery Obstruction; Diaphragm/diagnostic imaging

To the Editor:

Encephalocraniocutaneous lipomatosis (ECCL) (OMIM 613001) is a sporadic disorder characterized by congenital asymmetric ocular, cutaneous, and central nervous system malformations (Hunter, 2006; Moog, 2009). The most common dermatological feature is the nevus psiloliparus, a striking alopeic fatty nevus of the scalp seen in up to 80% of individuals with ECCL (Moog, 2009). Other skin anomalies include orbital and periorbital skin tags, subcutaneous fatty masses of the frontotemporal or zygomatic regions, linear or patchy non-scarring alopecia, focal dermal hypoplasia or aplasia of the scalp, and rarely, cutaneous hyperpigmentation (Moog, 2009). Ocular findings may be unilateral or bilateral and include choristomas, microphthalmia, calcification of the globe, and anomalies of the anterior chamber (Moog, 2009). Skeletal findings in ECCL include macrocephaly, lytic bone lesions, and jaw tumours, histologically reported as osteomas, odontomas, and ossifying fibromas (Moog, 2009; Moog, Roelens, et al., 2007). Intracranial and spinal lipomas are the most frequent CNS manifestation of ECCL, however, arachnoid cysts, cerebral hemiatrophy, ventriculomegaly, and leptomeningeal angiomatosis are documented (Moog, Jones, et al., 2007). The extent of these intracranial lesions does not predict the degree of intellectual disability or severity of epileptic activity, though both are common (Moog, 2009). Interestingly, both low- and high-grade gliomas have been described in several affected individuals highlighting a predisposition toward carcinogenesis in ECCL (Valera et al., 2012).

There is significant clinical overlap between ECCL and oculoectodermal syndrome (OES) (OMIM 600268). Shared features include focal dermal aplasia of the scalp, arachnoid cysts of the brain, ocular choristomas such as epibulbar dermoids, periorbital skin tags, linear hyperpigmentation, macrocephaly, and benign bone tumours (Ardinger, Horii, & Begleiter, 2007).

OES has been proposed as a mild presentation of ECCL though lack of intracranial lipomas in OES separates the two diagnoses (Ardinger et al., 2007; Moog, 2009). Given these similarities, it was hypothesized that ECCL and OES are caused by mutations in the same gene or at the very least, genes in the same pathway.

The patchy asymmetric lesions typical of ECCL, normal sex ratio, and absence of familial recurrence suggested that this sporadic disorder is caused by post zygotic mutations (Hall, 1988; Happle & Steijlen, 1993). This hypothesis was confirmed by the identification of mosaic mutations within the tyrosine kinase domain of *FGFR1* in a cohort of individuals with ECCL (Bennett et al., 2016). Fibroblasts from these affected individuals showed increased activation of the FGFRs and insensitive signal transduction through the RAS-MAPK pathway (Bennett et al., 2016). Interestingly, mosaic mutations in *KRAS* have been identified in patients with OES (Peacock et al., 2015) and more recently ECCL (Boppudi et al., 2016), highlighting the molecular overlap between these two syndromes and broadening the phenotypic spectrum of RAS-related disorders (*Rasopathies)*. Moreover, these studies confirmed the importance of tissue-type considerations when approaching suspected mosaic conditions, as causative mutations are likely tissue-restricted. Indeed, Bennett et al. (2016) successfully identified mosaic mutations in *FGFR1* in cultured fibroblasts (derived from scalp nevus, eyelid dermoid, affected scalp and unaffected skin) and pilocytic astrocytoma. This tissue selection mirrored that of Bopuddi et al. (2016) who confidently identified mosaic *KRAS* mutations in an epibulbar dermoid, scalp lesion biopsies and scalp-biopsy derived fibroblasts. Peacock et al. (2015) interrogated non-ossifying fibromas, muscle, periosteum and hyperpigmented skin to identify mosaic alterations in *KRAS* and strikingly, in one affected individual, were able to identify the causative mutation in skin as well as blood and bone marrow.

Here, we report a Chinese male with noted anomalies at birth, who presented for genetic investigation at age 14 years with nevus psiloliparus, temporal fatty masses, an isolated

subcutaneous lipoma of the left scalp and focal alopecia (Fig 1A). These features coupled with intracranial lipomas and prominent ventricles on MRI led to a preliminary diagnosis of ECCL. Cerebral hemiatrophy, intracranial cysts, and spinal lipomas were not identified on imaging. However, lytic bones lesions of the scalp, femur, and tibia were eventually recognized. Bilateral ocular findings consisted of retinal hypopigmentation, corneal opacities, and choristomas in addition to left upper eyelid coloboma and epibulbar dermoid. He also had marked cutaneous hyperpigmentation following the lines of Blashko on the trunk and all four limbs (Fig 1B).

At age 8 years, the proband was investigated for obesity and chronic hypertension exceeding the 95th percentile which revealed normal cortisol, catecholamine, and thyroid profiles. Fasting glucose and HbA1c levels were also normal. Assessment for renovascular hypertension included Doppler ultrasonography of his kidneys, which showed elevated flow at the proximal right renal artery (RRA) suspicious for arterial stenosis. A follow-up CT angiogram, performed at 13 years, confirmed significant stenosis of the RRA (~62% stenosis) in addition to moderate stenosis of both the superior mesenteric (SMA; ~45% stenosis) celiac (CA; ~50% stenosis) arteries (Fig 1C-E). Mild stenosis of the proximal left renal artery (LRA; ~30%) was noted (Fig 1F). Interestingly, the diaphragmatic crura (DC), the musculotendinous extensions that tether the diaphragm to the vertebral column, showed marked muscular hypertrophy resulting in compression of the RRA, CA, and SMA. No atherosclerotic plaques were appreciated.

The family provided written informed consent and this study was approved by the ethics review boards at the Children's Hospital of Eastern Ontario and the University of Hong Kong/Hong Kong West Cluster (Hospital Authority of Hong Kong). To maximize the likelihood of detecting tissue-restricted somatic mutations, high throughput sequencing was performed on genomic DNA, which was extracted from the proband's scalp lipoma according to standard methods. The DNA sample was enriched for 4813 clinically significant genes using the TruSight One Sequencing Panel

kit (Illumina Inc.) and sequenced on the MiSeq platform (Illumina Inc.) according to the manufacturer's recommendations for paired-end 150 base pair reads. The MiSeq Reporter Software (Illumina Inc.) was used to trim the adaptors and generate FASTQ files. Alignment, variant calling, and annotation were performed with the NextGene software (SoftGenetics LLC). Variants with an allele frequency greater than 1% in either the 1000 Genomes Project or the NHLBI Exome Variant Server were filtered.

Genes associated with ECCL were interrogated to identify causal pathogenic variants. One rare missense variant, c.436G>A (p.Ala146Thr) was identified in *KRAS* (NM_004985.3) (34% alternate allele frequency) (Fig 2A). No other variants were identified in *KRAS* or *FGFR1* and visual inspection of the sequencing data using the Integrative Genomics Viewer (Broad Institute) confirmed coverage of all exons in both genes. The c.436G>A alteration in *KRAS* was confirmed by conventional Sanger sequencing in DNA extracted from the lipoma but was not detected in the proband's blood, saliva, buccal swab, or fibroblasts cultured from a scalp biopsy, in keeping with a post-zygotic somatic alteration (Fig 2B). Neither parent carried the variant (Fig 2B). This mutation was not present in the Genome Aggregation Database.

KRAS, a Ras GTPase, cycles between an inactive guanosine diphosphate (GDP)-bound and an active guanosine triphosphate (GTP)-bound state to regulate key cellular processes such as proliferation, differentiation, and survival (Vetter & Wittinghofer, 2001). Given the importance of these processes in development, it is unsurprising that disruption of RAS signalling results in developmental defects. Indeed, gain-of-function germline and mosaic mutations in *KRAS* are recognized in multiple developmental syndromes and congenital skin disorders (Bennett et al., 2016; Hafner & Groesser, 2013; Niemela et al., 2011; Peacock et al., 2015). Somatic gain-of-function *KRAS* mutations are also heavily implicated in cancer and their tissue- and cell-type

specific contributions to oncogenic signalling during tumorigenesis are just starting to be understood (Schneider, Schmidt-Supprian, Rad, & Saur, 2017).

The p.Ala146Thr substitution in KRAS identified in this study was first described in human colorectal cancer but has since been detected in several other forms of neoplasia (hematopoetic, pancreatic, biliary, , others) (Forbes et al., 2010; Orita et al., 1991). The p.Ala146Thr substitution has been shown to increase RAS-GTP expression, activate RAS signalling and be transformative though the mechanism for this mutation's activity is not completely understood (Janakiraman et al., 2010; Smith et al., 2010). Recently, the mosaic c.436G>A variant (p.Ala146Thr) was identified in lesional skin biopsies of an individual with OES as well as in lesional skin and cultured fibroblasts of an individual with ECCL (Boppudi et al., 2016). Neither of these patients had reported renovascular hypertension , to date, stenosis of the renal arteries has not been documented in ECCL or OES.

The two most common causes of renal-artery stenosis are atherosclerosis (90% of cases) and fibromuscular dysplasia of the arterial wall (<10% of cases) (Safian & Textor, 2001). Entrapment of the renal arteries (RA) by the DC remains a rare cause of renal-artery stenosis and of renovascular hypertension (Deglise et al., 2007). The proband described in this report showed stenosis of the CA, SMA, and RRA secondary to compression by hypertrophic DC. Such multivessel involvement has rarely been documented (Deglise et al., 2007). Despite potential compromise of vascular flow through the SMA and CA, which could result in decreased visceral organ perfusion, the proband remains free of gastrointestinal symptoms and to date, his hypertension has been managed conservatively with multiple antihypertensive medications (metoprolol tartrate 150mg twice daily, Exforge (5mg/160mg) 1 tablet twice daily and enalapril maleate 7.5mg twice daily). Given the high failure rate of angioplasty, definitive management for patients with vessel entrapment secondary to DC hypertrophy is surgical and includes DC

debulking, reanastomosis and bypass (Gaebel, Hinterseher, Saeger, & Bergert, 2009; Thony et al., 2005).

Renovascular hypertension has been described in the *Rasopathies* (Eom et al., 2015; Oderich et al., 2007). Notably, vascular malformations are well recognized in neurofibromatosis syndrome type 1 (NF1); stenosis of the RA, SMA, and CA have all been reported (Oderich et al., 2007). Unlike the extrinsic entrapment described here, the etiology of RA stenosis in NF1 is a result of vessel wall dysplasia (Friedman et al., 2002). Hallmark features of ECCL are characterized by anomalies of ectodermal and mesodermal origin (Moog, 2009). DC arise from the myoblastic infiltration of the dorsal mesentery (Merrell & Kardon, 2013). Conceivably, DC hypertrophy may represent another possible manifestation of mesodermal dysgenesis in ECCL and without investigation for renovascular hypertension, or incidental finding on imaging, asymptomatic hypertrophic DC would go undiagnosed. For this reason, hypertension in a patient with ECCL should prompt further investigation into a renovascular cause.

One of the hallmarks distinguishing ECCL from OES is intracranial and intraspinal lipomatosis. Nonetheless, the clinical diagnosis of ECCL may still be given to an individual without CNS lipomas if they present with two or more CNS abnormalities as delineated by Moog et al. (2009). Two individuals with the recurrent mosaic c.436G>A (p.Ala146Thr) alteration in *KRAS* have been described (Boppudi et al., 2016). These two individuals, diagnosed with ECCL and OES, respectively, did not show evidence of CNS lipomas on imaging, suggesting that mosaic mutations in *FGFR1* and not *KRAS* result in CNS lipomatosis (Bennett et al., 2016; Boppudi et al., 2016; Peacock et al., 2015). Our findings demonstrate that an individual affected with a mosaic mutation in *KRAS* may indeed present with intracranial lipomas. Importantly, intracranial lipomas represent a congenital malformation rather than a true neoplasm and result from the persistence and maldifferentiation of the primitive meninges or meninx primitiva during development (Truwit &

Barkovich, 1990). As such, increased signalling through RAS-MAPK, whether initiated by gain-of-function mutations in *FGFR1* or *KRAS*, may lead to CNS lipomatosis if it occurs in the meninx primitiva at a specific developmental window. Thus, the clinical manifestations of the overlapping OES/ECCL mutations likely represent developmental cell fate decisions in the context of increased RAS-MAPK signalling in specific tissues at given time points.

In conclusion, deep sequencing of affected tissue in a patient with ECCL and extrinsic compression of the RRA, SMA, and CA has identified the recurrent c.436G>A (p.Ala146Thr) mosaic mutation in *KRAS*. This is the first report of hypertrophic DC in ECCL and may expand the phenotypical spectrum of mosaic Rasopathies. It is evident that the timing, type, and location of *KRAS* mutations contribute to alternate, but in some instances, converging phenotypic and oncogenic outcomes. Given the enrichment of the p.Ala146Thr substitution in colorectal cancer, longitudinal reporting of oncogenic outcomes in ECCL patients remains important.

ACKNOWLEDGMENTS

Research in this article was supported by The Hong Kong Society for the Relief of Disabled Children, a Canadian Institutes of Health Research Doctoral Research Award, and the Baxter and Alma Ricard Foundation. This work was performed under the Care4Rare Canada Consortium funded by Genome Canada, the Canadian Institutes of Health Research (CIHR), the Ontario Genomics Institute, Ontario Research Fund, Génome Québec, and Children's Hospital of Eastern Ontario Foundation. We would like to thank the proband and his family for their ongoing participation in this study.

CONFLICTS OF INTEREST

The authors declare no competing financial or professional interests and take responsibility for the writing of this paper.

REFERENCES

Ardinger, H. H., Horii, K. A., & Begleiter, M. L. (2007). Expanding the phenotype of oculoectodermal syndrome: possible relationship to encephalocraniocutaneous lipomatosis. *American Journal of Medical Genetics Part A, 143A*(24), 2959-2962.

Bennett, J. T., Tan, T. Y., Alcantara, D., Tetrault, M., Timms, A. E., Jensen, D., . . . McDonell, L. M. (2016). Mosaic Activating Mutations in FGFR1 Cause Encephalocraniocutaneous Lipomatosis. *American Journal of Human Genetics, 98*(3), 579-587.

Boppudi, S., Bogershausen, N., Hove, H. B., Percin, E. F., Aslan, D., Dvorsky, R., . . . Zenker, M. (2016). Specific mosaic KRAS mutations affecting codon 146 cause oculoectodermal syndrome and encephalocraniocutaneous lipomatosis. *Clinical Genetics, 90*(4), 334-342.

Deglise, S., Corpataux, J. M., Haller, C., Binaghi, S., Meuwly, J. Y., & Qanadli, S. D. (2007). Bilateral renal artery entrapment by diaphragmatic crura: a rare cause of renovascular hypertension with a specific management. *Journal of Computer Assisted Tomography, 31*(3), 481-484.

Eom, M.-Y., Kim, W.-J., Kim, K.-H., Kim, Y.-N., Choi, W., Jung, Y.-L., & Cho, H.-A. (2015). A Case of Noonan Syndrome Presenting with Malignant Hypertension in an Adult. *Korean J Med, 89*(4), 433-438.

Forbes, S. A., Tang, G., Bindal, N., Bamford, S., Dawson, E., Cole, C., . . . Futreal, P. A. (2010). COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. *Nucleic Acids Research, 38*(Database issue), D652-657.

Friedman, J. M., Arbiser, J., Epstein, J. A., Gutmann, D. H., Huot, S. J., Lin, A. E., . . . Korf, B. R. (2002). Cardiovascular disease in neurofibromatosis 1: report of the NF1 Cardiovascular Task Force. *Genetics in Medicine, 4*(3), 105-111.

Gaebel, G., Hinterseher, I., Saeger, H. D., & Bergert, H. (2009). Compression of the left renal artery and celiac trunk by diaphragmatic crura. *Journal of Vascular Surgery, 50*(4), 910-914.

Hafner, C., & Groesser, L. (2013). Mosaic RASopathies. *Cell Cycle, 12*(1), 43-50.

Hall, J. G. (1988). Review and hypotheses: somatic mosaicism: observations related to clinical genetics. *American Journal of Human Genetics, 43*(4), 355-363.

Happle, R., & Steijlen, P. M. (1993). Encephalocraniocutaneous lipomatosis. A non-hereditary mosaic phenotype. *Hautarzt, 44*(1), 19-22.

Hunter, A. G. (2006). Oculocerebrocutaneous and encephalocraniocutaneous lipomatosis syndromes: blind men and an elephant or separate syndromes? *American Journal of Medical Genetics Part A, 140*(7), 709-726.

Janakiraman, M., Vakiani, E., Zeng, Z., Pratilas, C. A., Taylor, B. S., Chitale, D., . . . Solit, D. B. (2010). Genomic and biological characterization of exon 4 KRAS mutations in human cancer. *Cancer Research, 70*(14), 5901-5911.

Merrell, A. J., & Kardon, G. (2013). Development of the diaphragm -- a skeletal muscle essential for mammalian respiration. *FEBS Journal, 280*(17), 4026-4035.

Moog, U. (2009). Encephalocraniocutaneous lipomatosis. *Journal of Medical Genetics, 46*(11), 721-729.

Moog, U., Jones, M. C., Viskochil, D. H., Verloes, A., Van Allen, M. I., & Dobyns, W. B. (2007). Brain anomalies in encephalocraniocutaneous lipomatosis. *American Journal of Medical Genetics Part A, 143A*(24), 2963-2972.

Moog, U., Roelens, F., Mortier, G. R., Sijstermans, H., Kelly, M., Cox, G. F., . . . Kimonis, V. E. (2007). Encephalocraniocutaneous lipomatosis accompanied by the formation of bone cysts:

Harboring clues to pathogenesis? *American Journal of Medical Genetics Part A, 143A*(24), 2973-2980.
Niemela, J. E., Lu, L., Fleisher, T. A., Davis, J., Caminha, I., Natter, M., . . . Oliveira, J. B. (2011). Somatic KRAS mutations associated with a human nonmalignant syndrome of autoimmunity and abnormal leukocyte homeostasis. *Blood, 117*(10), 2883-2886.
Oderich, G. S., Sullivan, T. M., Bower, T. C., Gloviczki, P., Miller, D. V., Babovic-Vuksanovic, D., . . . Stanson, A. (2007). Vascular abnormalities in patients with neurofibromatosis syndrome type I: clinical spectrum, management, and results. *Journal of Vascular Surgery, 46*(3), 475-484.
Orita, S., Higashi, T., Kawasaki, Y., Harada, A., Igarashi, H., Monden, T., . . . Miyoshi, J. (1991). A novel point mutation at codon 146 of the K-ras gene in a human colorectal cancer identified by the polymerase chain reaction. *Virus Genes, 5*(1), 75-79.
Peacock, J. D., Dykema, K. J., Toriello, H. V., Mooney, M. R., Scholten, D. J., 2nd, Winn, M. E., . . . Steensma, M. (2015). Oculoectodermal syndrome is a mosaic RASopathy associated with KRAS alterations. *American Journal of Medical Genetics Part A, 167*(7), 1429-1435.
Safian, R. D., & Textor, S. C. (2001). Renal-artery stenosis. *New England Journal of Medicine, 344*(6), 431-442.
Schneider, G., Schmidt-Supprian, M., Rad, R., & Saur, D. (2017). Tissue-specific tumorigenesis: context matters. *Nature Reviews: Cancer, 17*(4), 239-253.
Smith, G., Bounds, R., Wolf, H., Steele, R. J., Carey, F. A., & Wolf, C. R. (2010). Activating K-Ras mutations outwith 'hotspot' codons in sporadic colorectal tumours - implications for personalised cancer medicine. *British Journal of Cancer, 102*(4), 693-703.
Thony, F., Baguet, J. P., Rodiere, M., Sessa, C., Janbon, B., & Ferretti, G. (2005). Renal artery entrapment by the diaphragmatic crus. *European Radiology, 15*(9), 1841-1849.
Truwit, C. L., & Barkovich, A. J. (1990). Pathogenesis of intracranial lipoma: an MR study in 42 patients. *AJR: American Journal of Roentgenology, 155*(4), 855-864; discussion 865.
Valera, E. T., Brassesco, M. S., Scrideli, C. A., de Castro Barros, M. V., Santos, A. C., Oliveira, R. S., . . . Tone, L. G. (2012). Are patients with encephalocraniocutaneous lipomatosis at increased risk of developing low-grade gliomas? *Child's Nervous System, 28*(1), 19-22.
Vetter, I. R., & Wittinghofer, A. (2001). The guanine nucleotide-binding switch in three dimensions. *Science, 294*(5545), 1299-1304.

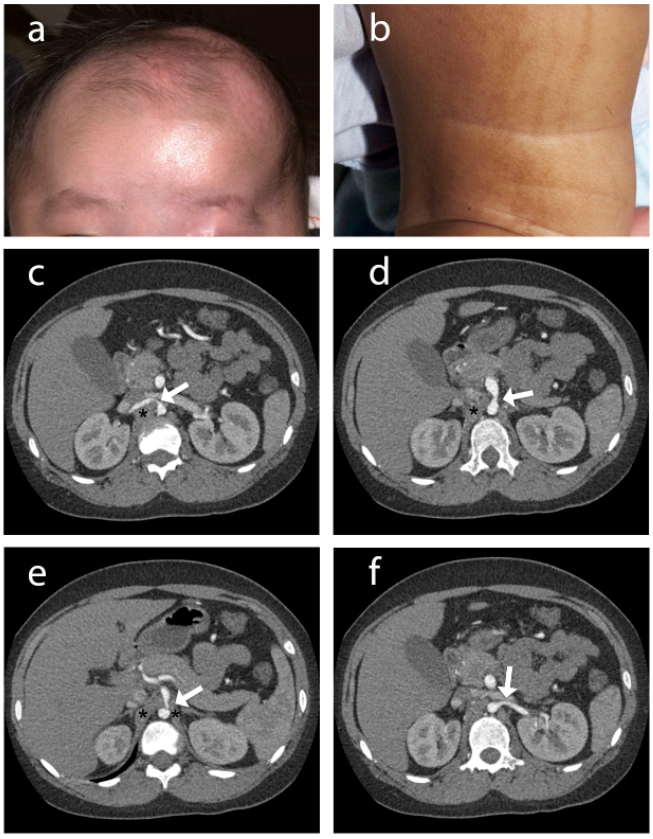


FIGURE 1 Clinical features and axial CT imaging of the proband at 4 month and 13 years respectively. (a) Photograph showing asymmetric temporal fatty masses, alopecia and a parietal fatty accumulation suspicious for nevus psiloliparus. Disrupted left eyebrow are also visible. (b) Linear hyperpigmentation of the posterior right thigh. (c) Hypertrophied diaphragmatic crus (*) indenting into the proximal right renal artery (white arrow). (d) Hypertrophied diaphragmatic crus (*) impinging the proximal superior mesenteric artery (white arrow). (e) Bilateral hypertrophied diaphragmatic crura (*) compressing the proximal celiac artery (white arrow). (f) Mild stenosis of the proximal left renal artery (white arrow) in absence of extrinsic compression.

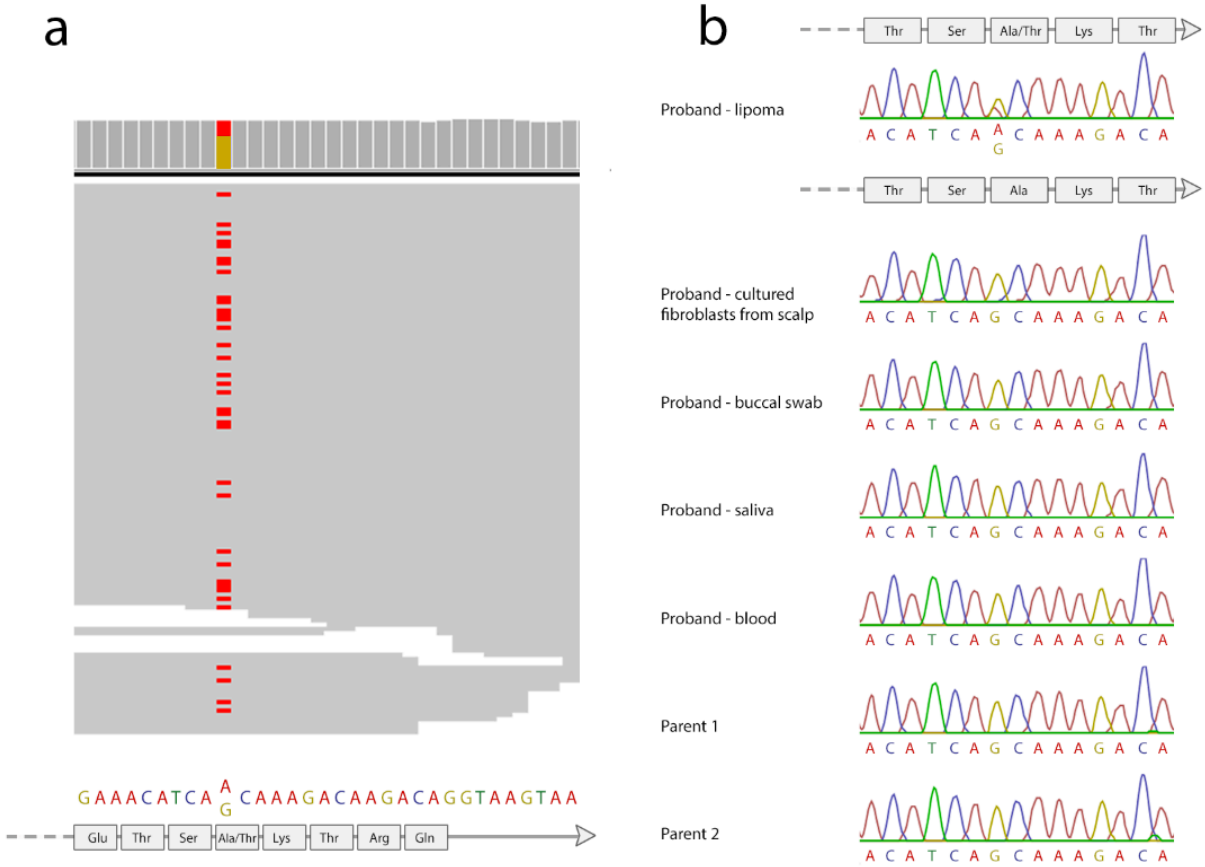


FIGURE 2 Mosaic *KRAS* missense variant identified in the proband's lipoma. (a) MiSeq generated sequencing alignment showing the c.436G>A (p.Ala146Thr) alteration at 166X coverage and 34% alternate allele frequency. (b) The c.436G>A alteration is detected in lipoma by Sanger sequencing but is absent from the proband's cultured fibroblasts (scalp biopsy), buccal swab, saliva, and blood, and parental blood samples.

Chapter 5: PHACE syndrome

5.1 Preface

The following chapter consists of the manuscript titled "Understanding the molecular etiology of PHACE syndrome: Approaches and considerations" that will be submitted to Molecular Genetics and Genomic Medicine by Laura M. McDonell, Erik Bareke, Sergey A. Naumenko, Jean McGowan-Jordan, Sharan L. Goobie, Jacek Majewski, Care4Rare Canada Consortium, Kym M. Boycott

5.2 Statement of permission for use of copyrighted material

This manuscript will be been submitted to Molecular Genetics and Genomic Medicine. As such, an embargo will be applied at time of final submission.

5.3 Contributions

The specific contributions of each author are as follows:

Laura M. McDonell

Analyzed the sequencing data, interpreted microarray analysis and performed tissue culture for available cell lines. Wrote the manuscript, generated tables.

Erik Bareke

Performed bioinformatics analysis of the Illumina sequencing.

Sergey A. Naumenko

Performed bioinformatics analysis of the Illumina sequencing.

Jean McGowan-Jordan

Oversaw genotyping and performed microarray analysis.

Sharan L. Goobie

Contributed clinical data.

Jacek Majewski

Oversaw bioinformatics analysis of the Illumina sequencing data.

Care4Rare Canada Consortium

Provided critical infrastructure and resources for this project.

Kym M. Boycott

Oversaw the study and edited the manuscript.

5.4 Main Manuscript

Understanding the molecular etiology of PHACE syndrome: approaches and considerations

AUTHORS

Laura M. McDonell[1], Erik Bareke[2,3], Sergey A. Naumenko[4], Jean McGowan-Jordan[5], Sharan L. Goobie[6], Jacek Majewski[2,3], Care4Rare Canada Consortium, Kym M. Boycott[1]

AFFILIATIONS

1. Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, Canada

2. Department of Human Genetics, McGill University, Montreal, Canada

3. Department of Pediatrics, Montreal Children's Hospital, McGill University Health Centre, Montreal, Canada

4. Genetics and Genome Biology Program, The Hospital for Sick Children, Toronto, Canada

5.Department of Genetics, Children's Hospital of Eastern Ontario, Ottawa, Canada

6. Maritime Medical Genetics, Department of Pediatrics, Izaak Walton Killam Health Centre, Halifax, Canada

CORRESPONDING AUTHOR:

Dr. Kym M. Boycott,

Children's Hospital of Eastern Ontario Research Institute,

University of Ottawa,

Ottawa, Ontario, Canada

Email: kboycott@cheo.on.ca

ABSTRACT

Purpose

PHACE syndrome is a rare sporadic neurocutaneous disorder characterized by the association of infantile hemangioma with distinct vascular, neurologic, and ocular anomalies. Despite a large number of reported cases, the genetic etiology of this syndrome remains unknown. Given the patchy and asymmetric distribution of lesions in PHACE syndrome, it is hypothesized that this disorder is caused by post-zygotic mosaic mutations. Here, we investigate a monozygotic twin pair discordant for PHACE syndrome using whole exome sequencing (WES) and genome-wide genotyping arrays.

Methods

To investigate possible causative variants in coding regions and intron/exon boundaries, we performed WES of tissue from the proband and her unaffected twin. We subsequently genotyped the proband and her unaffected parents using Affymetrix's CytoScan HD platform and focused our analysis on identification of copy-neutral loss of heterozygosity (LOH) as well as copy number variations (CNVs) that were absent from parents and controls.

Results

In total, eight CNVs were detected in the proband, however, these were all found in her parents and healthy controls and as such, are likely benign. Additionally, three large regions of copy-neutral LOH were detected in the proband exclusively but no rare variants were identified by WES within these regions. Though WES enabled the identification of variants unique to the proband and not found in the unaffected twin, most were false positives and no convincing candidate variants were identified.

Conclusion

This study highlights the challenges faced when undertaking the investigation of hypothesized mosaic syndromes. Ultimately, the availability of appropriate tissues, cohort type and size as well as access to appropriate sequencing technologies will contribute to successful outcomes for the remaining unsolved but presumed genetic neurocutaneous disorders.

KEY WORDS

neurocutaneous syndromes; mosaicism; PHACE association; genotyping techniques; whole exome sequencing

1 | INTRODUCTION

PHACE syndrome (OMIM 606519) is a sporadic neurocutaneous disorder characterized by neurologic, arterial, cardiac and ocular anomalies associated with segmental infantile hemangiomas (IH). PHACE is an acronym that refers to the association of *P*osterior fossa malformations, facial *H*emangiomas, cerebral *A*rterial anomalies, *C*ardiovascular defects and *E*ye abnormalities. In the presence of ventral defects such as *S*ternal clefting or *S*upraumbilical raphe, the term PHACE(S) is used.[1] Structural brain anomalies such as posterior fossa defects including Dandy-Walker malformations and focal cerebellar hypoplasia are frequently seen.[2] Cerebrovascular anomalies involving the arterial system are common and can result in progressive arteriopathy and ischemic stroke.[3,4] Though there is limited histopathology describing the underlying vascular defects of PHACE syndrome, radiological studies have revealed a large spectrum of vascular anomalies including persistent embryonic arteries, aberrant arterial origin/course as well as arterial hypoplasia, dysplasia, agenesis and stenosis.[2,3,5-7] Interestingly, post-mortem analysis of one individual with PHACE syndrome showed disease limited to medium and large arteries with sparing of arterioles, capillaries, veinules and veins.[8] The most common observation was degeneration of the medial architecture of the affected vessels with or without associated fibro-intimal proliferation, and overall, these findings were suspicious for a developmental arterial dysplasia.[8]

Infantile hemangiomas are the most common benign tumours of infancy and are associated with 4-5% of live births.[9] Of children with large segmental facial hemangiomas, a subset, estimated to be as high as 31%, have extracutaneous findings and meet the diagnostic criteria for PHACE syndrome.[10] Moreover, there is a striking female predominance amongst individuals with PHACE syndrome.[7,10] In keeping with this, the female-to-male ratio is estimated to be as high as 9:1 compared to the 2.8:1 ratio of infantile hemangioma.[7,10,11] Intriguingly, a study of X chromosome

inactivation comprising 31 female individuals with PHACE syndromes and their mothers failed to show statistically significant skewed X inactivation.[12]

To date, over 300 individuals with PHACE syndrome have been described and despite this elevated number, no genetic cause has been identified.[2] The segmental hemangiomas, arteriopathy and absence of familial recurrence suggest that this disorder is caused by post-zygotic mosaic mutations. This is further supported by the occurrence of discordant monozygotic twins, presented herein. In this study, we pursue pathogenic copy number variations (CNV) using genome-wide genotyping arrays and employ whole-exome sequencing (WES) to pursue *de novo* coding variants in a monozygotic twin pair discordant for PHACE syndrome. Furthermore, we discuss signalling pathways that may be contributing to the clinical phenotype and devise approaches to solving hypothesized mosaic neurocutaneous conditions.

2 | MATERIALS AND METHODS

2.1 Ethical considerations

The family provided written and informed consent for genetic studies and DNA extraction from blood and saliva samples. This study was approved by the ethics review boards at the Children's Hospital of Eastern Ontario and the London Health Sciences Centre.

2.2 Study participants

The proband belongs to a female monozygotic twin pair discordant for PHACE syndrome. She meets the diagnostic criteria for this syndrome as outlined by Metry *et al*. (2009). DNA was isolated from available tissue according to standard methods and used for genotyping and high-throughput sequencing studies.

2.3 Exome sequencing

We performed WES of blood- and saliva-derived DNA from the proband and her unaffected monozygotic twin. Genomic DNA was fragmented and exomes captured with the SureSelect All Exon V5 Panel (Agilent, Santa Clara, CA, U.S.A.). The captured libraries were subsequently sequenced on the Illumina HiSeq2000. Raw paired-end reads were aligned to the human reference genome (UCSC hg19) using BWA.[13] Single-nucleotide variants and small insertions/deletions (indels) were detected using Samtools/BCFtools[14], GATK[15] and custom scripts. Variants were annotated using Ensemble Variant effect predictor (http:// www.ensembl.org/info/docs/tools/vep/index.html), Gemini18, and custom annotations (https://github.com/naumenko-sa/cre) including OMIM (http://www.omim.org) and Orphanet (http://www.oprha.net/consor/cgi-bin/index.php) databases. We prioritized variants using gnomAD[16] and frequencies in the internal Care4Rare database. For the purpose of identifying mosaic mutations, variants at any depth of coverage were considered.

2.4 Microarray analysis

Microarray analysis followed WES and unfortunately, by this time, DNA from the unaffected twin had been exhausted obliging us to use parental DNA. Furthermore, cultured fibroblast lines derived from biopsies of the proband's skin were established after WES and given the hypothesized mosaicism, we favoured using this tissue for genotyping. As such, DNA was extracted from blood samples provided by each parent as well as from the proband's fibroblasts. Genotyping was carried out on the Affymetrix CytoScan HD platform (Santa Clara, CA, U.S.A.) according to the manufacturer's instruction and the resulting data was analyzed using the ChAS software package. The size cut-off for detection of duplications and deletions was set at 500kb and 200kb,

respectively. For regions of known clinical significance, the threshold for both duplications and deletions was lowered to 50kb. Analysis focused on identifying CNVs identified in the fibroblast DNA from the affected proband but absent from either parent.

3 | RESULTS

3.1 Exome variants

We performed multiple phases of analysis of the WES data generated from blood and saliva provided by the proband and her unaffected twin. Each phase focused on identifying coding or exon/intron variants in the blood and/or saliva samples provided by the proband. Any variant, detected at a depth of coverage of two or more reads was considered. Once a candidate variant was identified, we verified the alignments to ensure its validity, as indels located in repeat-heavy regions are often false positives.

First, we sought variants present in both the proband's blood and saliva but absent from the unaffected twin, in-house controls and gnomAD. Though variants in five genes were identified, all were false positives seen in the unaffected twin on alignment.

For our second analysis, we identified variants present in the proband's saliva but absent from blood, the unaffected twin, in-house controls and gnomAD. We detected, variants in 10 genes (*SETD18, KRT77, PTGER2, GOLGA6L2, MTX2, ZFYVE28, KCNH2, LY6K, TPD52 and GOLGA2*).

We then looked for variants present in the proband's blood but absent from saliva, the unaffected twin, in-house controls and gnomAD. We identified variants in 4 genes (*KCNH1, RBM45, MUC21* and *HLA-DRB)*.

Finally, we filtered the data to reveal homozygous variants found in a heterozygous state in the unaffected twin. We identified one intronic two-basepair deletion in *ESCO1* located in a repeat-rich region. Six in-house controls also harboured this variant, making it an unlikely candidate variant.

3.2 CNVs

Overall, eight CNVs were identified in the DNA from the fibroblasts derived from the proband (Table 1). Of these, one was detected in both parents, four were detected in one parent, one was located in the variable immunoglobulin heavy (IGH) locus, and finally, two were in regions with poor probe coverage and likely inherited. All eight CNVs were found in the Database of Genomic Variants (http://dgv.tcag.ca/dgv/app/home) and the Ontario Population Genomics Platform (http://www.tcag.ca/facilities/cyto_population_control_DNA.html). Altogether, this data suggests, that the CNVs identified in the proband are not involved in the pathophysiology of PHACE syndrome.

3.3 Loss of heterozygosity

The proband harboured three large regions of copy-neutral loss of heterozygosity (LOH) on chromosome X (Table 2). Each region spanned approximately 5Mb, contained multiple genes and was not detected in her mother (Table 2). Interestingly, no rare variants were detected by WES in these regions.

4 | DISCUSSION

Neurocutaneous disorders are a heterogeneous collection of congenital conditions recognized for their central nervous system (CNS) and skin involvement.[17] A subset of these conditions present with striking vascular anomalies of the skin.[18] For instance, patients with megalencephaly-capillary

malformation syndrome (OMIM 602501), caused by *de novo* mosaic mutations in *PIK3CA*, have capillary malformations of the face, limbs and trunk.[19,20] Moreover, Sturge-Weber syndrome (OMIM 185300), caused by somatic mutations in *GNAQ*, is characterized by extensive angiomatous vascular malformations of the face, eyes and CNS.[21] Both germline and somatic mutations are identified in the neurocutaneous syndromes and many affect genes of the interconnected Ras-MAPK and PI3K-AKT signalling networks.[22,23] These pathways govern fundamental cellular processes such as metabolism, proliferation, differentiation and survival, as such, their disruption during development can have profound effects on cell-fate decisions and contribute to developmental defects.[23-27] Disruption of these signalling networks is associated with a family of neurocutaneous syndromes, termed the *Rasopathies*, recognized for their spectrum of cutaneous, ocular, cardiac and CNS anomalies.[23]

Interestingly, the striking segmental hemangiomas, arterial malformations and CNS defects of PHACE syndrome suggest that it too, may belong to the *Rasopathies*. This knowledge suggests that variants in genes of the Ras-MAPK and PI3K-AKT pathways identified by high-throughput sequencing technologies should be carefully examined. Variants in other relevant pathways, such as VEGF and HIF-1α, should also be considered. No such variants were identified in our data.

We are not the first group to pursue the molecular etiology of PHACE syndrome. As described above, X chromosome inactivation was investigated to test whether X-linked mutations inherited from an unaffected mother with favourable skewing could, in daughters with random patterns of X-inactivation, unmask the PHACE phenotype.[12] Analysis revealed skewed X chromosome-inactivation in a subset of the mothers but these results were not statistically significant.[12] In another study, interrogation of CNVs in 98 individuals with PHACE syndrome failed to detect rare CNVs shared by more than one individual.[28] Importantly, both studies were undertaken with DNA isolated from blood and saliva. This is not surprising, given the accessibility

of these tissues, however, in disorders resulting from post-zygotic mosaicism, the causative mutation is likely tissue-restricted.

In fact, for many mosaic conditions, the causative mutation is observed at lower, if not absent, levels of mosaicism in blood.[19,29] As such, identification of pathogenic mosaic mutation in blood and saliva has relied on targeted ultra-deep sequencing (>10 000 fold).[19] To maximize the likelihood of identifying mosaic mutations, WES of affected uncultured tissue is favoured as culturing artifacts may be introduced in cell lines of affected tissue. Regardless, sequencing of both cultured and uncultured affected tissue has successfully identified pathogenic low-frequency mosaic mutations in neurocutaneous disorders.[19,29,30] Access to high-yield tissue can sometimes be challenging and in the case of PHACE syndrome, biopsies of dysplastic arteries and hemangiomas would provide the ideal material for deep sequencing. Unfortunately, sequencing in this study relied on blood- and saliva-derived DNA for WES. As such, it is likely that the causative mutation, even if present at very low frequency in blood or saliva, was not detected given the depth of coverage provided by WES compared to targeted deep sequencing.

Discordant monozygotic twins provide the ideal group in which to assess mosaicism as the unaffected twin can acts like a built-in internal control allowing the rapid identification of *de novo* variants while reducing the number of alignment and variant calling artifacts. As discussed above, the ability to deeply sequence an appropriate tissue in the discordant twins, as well as appropriate tissues from additional affected individuals would improve the likelihood of successfully identifying the disease gene.

Elucidating the genetic etiology of PHACE has proven to be a challenging endeavour complicated by restricted access to appropriate tissue, a small cohort size and overall limitations of the depth of standard WES. Ultimately, structural rearrangements or regulatory mutations may be responsible for PHACE syndrome and in that case, whole genome sequencing with de novo

alignment would be the most appropriate technology for their detection. We anticipate that identifying the molecular cause of PHACE syndrome will provide novel insights into the disease spectrum of the *Rasopathies* and the proteins regulating normal human development.

ACKNOWLEDGMENTS

Research in this article was supported by a Canadian Institutes of Health Research Doctoral Research Award, and the Baxter and Alma Ricard Foundation. This work was performed under the Care4Rare Canada Consortium funded by Genome Canada, the Canadian Institutes of Health Research (CIHR), the Ontario Genomics Institute, Ontario Research Fund, Génome Québec, and Children's Hospital of Eastern Ontario Foundation. We would like to thank the proband and their family for their ongoing participation in this study as well as the staff members in the clinical recruiting centers.

CONFLICTS OF INTEREST

None declared.

REFERENCES

1 Frieden, I. J., Reese, V. & Cohen, D. PHACE syndrome. The association of posterior fossa brain malformations, hemangiomas, arterial anomalies, coarctation of the aorta and cardiac defects, and eye abnormalities. *Arch. Dermatol.* **132**, 307-311, (1996).

2 Metry, D. *et al.* Consensus Statement on Diagnostic Criteria for PHACE Syndrome. *Pediatrics* **124**, 1447-1456, (2009).

3 Burrows, P. E. *et al.* Cerebral vasculopathy and neurologic sequelae in infants with cervicofacial hemangioma: report of eight patients. *Radiology* **207**, 601-607, (1998).

4 Drolet, B. A. *et al.* Early stroke and cerebral vasculopathy in children with facial hemangiomas and PHACE association. *Pediatrics* **117**, 959-964, (2006).

5 Hess, C. P. *et al.* Cervical and intracranial arterial anomalies in 70 patients with PHACE syndrome. *AJNR Am. J. Neuroradiol.* **31**, 1980-1986, (2010).

6 Heyer, G. L. *et al.* The cerebral vasculopathy of PHACES syndrome. *Stroke* **39**, 308-316, (2008).

7 Metry, D. W. *et al.* PHACE syndrome: current knowledge, future directions. *Pediatr. Dermatol.* **26**, 381-398, (2009).

8 Chad, L. *et al.* Postmortem vascular pathology in PHACES syndrome: a case report. *Pediatr. Dev. Pathol.* **15**, 507-510, (2012).

9 Kilcline, C. & Frieden, I. J. Infantile hemangiomas: how common are they? A systematic review of the medical literature. *Pediatr. Dermatol.* **25**, 168-173, (2008).

10 Haggstrom, A. N. *et al.* Risk for PHACE syndrome in infants with large facial hemangiomas. *Pediatrics* **126**, e418-426, (2010).

11 Metry, D. W. *et al.* A prospective study of PHACE syndrome in infantile hemangiomas: demographic features, clinical findings, and complications. *Am. J. Med. Genet. A* **140**, 975-986, (2006).

12 Sullivan, C. T. *et al.* X Chromosome-Inactivation Patterns in 31 Individuals with PHACE Syndrome. *Molecular syndromology* **4**, 114-118, (2013).

13 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760, (2009).

14 Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987-2993, (2011).

15 Van der Auwera, G. A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* **43**, 11 10 11-33, (2013).

16 Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285-291, (2016).

17 Ruggieri, M. & Pratico, A. D. Mosaic Neurocutaneous Disorders and Their Causes. *Semin. Pediatr. Neurol.* **22**, 207-233, (2015).

18 Puttgen, K. B. & Lin, D. D. Neurocutaneous vascular syndromes. *Childs Nerv. Syst.* **26**, 1407-1415, (2010).

19 Riviere, J. B. *et al.* De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. *Nat. Genet.* **44**, 934-940, (2012).

20 Mirzaa, G. M. & Dobyns, W. B. The "megalencephaly-capillary malformation" (MCAP) syndrome: the nomenclature of a highly recognizable multiple congenital anomaly syndrome. *Am. J. Med. Genet. A* **161A**, 2115-2116, (2013).
21 Shirley, M. D. *et al.* Sturge-Weber syndrome and port-wine stains caused by somatic mutation in GNAQ. *N. Engl. J. Med.* **368**, 1971-1979, (2013).
22 Keppler-Noreuil, K. M., Parker, V. E., Darling, T. N. & Martinez-Agosto, J. A. Somatic overgrowth disorders of the PI3K/AKT/mTOR pathway & therapeutic strategies. *Am. J. Med. Genet. C Semin. Med. Genet.* **172**, 402-421, (2016).
23 Schubbert, S., Shannon, K. & Bollag, G. Hyperactive Ras in developmental disorders and cancer. *Nat. Rev. Cancer* **7**, 295-308, (2007).
24 Castellano, E. & Downward, J. RAS Interaction with PI3K: More Than Just Another Effector Pathway. *Genes Cancer* **2**, 261-274, (2011).
25 Mendoza, M. C., Er, E. E. & Blenis, J. The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. *Trends Biochem. Sci.* **36**, 320-328, (2011).
26 Martini, M., De Santis, M. C., Braccini, L., Gulluni, F. & Hirsch, E. PI3K/AKT signaling pathway and cancer: an updated review. *Ann. Med.* **46**, 372-383, (2014).
27 Cully, M., You, H., Levine, A. J. & Mak, T. W. Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis. *Nat. Rev. Cancer* **6**, 184-192, (2006).
28 Siegel, D. H. *et al.* Copy number variation analysis in 98 individuals with PHACE syndrome. *J. Invest. Dermatol.* **133**, 677-684, (2013).
29 Lindhurst, M. J. *et al.* A mosaic activating mutation in AKT1 associated with the Proteus syndrome. *N. Engl. J. Med.* **365**, 611-619, (2011).
30 Bennett, J. T. *et al.* Mosaic Activating Mutations in FGFR1 Cause Encephalocraniocutaneous Lipomatosis. *Am. J. Hum. Genet.* **98**, 579-587, (2016).

Table 1. Copy number variations detected in the proband

Copy Number	Gain, loss or LOH	Chr	Nucleotide start	Nucleotide stop	Size (kb)	Seen in >2 controls (DGV)	Seen in >2 controls (OPGP)	Seen in in-house controls	Seen in parent	Gene
1	Loss	8	39247097	39352609	106	Yes	Yes	1	Mother, father	ADAM5, ADAM3A
1	Loss	10	46287821	47149411	862	Yes	Yes	Yes	Likely in father	FAM21C, AGAP4, PTPN20A,PTPN20B, FRMPD2, SYT15, GPRIN2, ANXA8, NPY4R, LINC00842
3	Gain	12	19467158	19577135	110	Yes	Yes	No	Mother	PLEKHA5
3	Gain	12	52685758	52785919	100	Yes	Yes	No	Father	KRT81, KRT86, KRT83, KRT85, KRT84
3	Gain	14	106270138	106716050	446	Yes	Yes	>2	IGH	IGHE, IGHG1, IGHD, FLJ00382, KIAA0125, ADAM6
1	Loss	15	43888261	43976407	88	Yes	Yes	Yes	Mother	CKMT1B, STRC, CATSPER2, PPIP5K1
3	Gain	17	34424638	34477480	53	Yes	Yes	No	Mother	CCL4
1	Loss	17	36283806	36410559	127	Yes	Yes	No	Likely in mother	TBC1D3

Abbreviations are as follows: IGH (immunoglulin heavy chain locus), Chr (chromosome), DGV (Database of Genomic Variants) and OPGP (Ontario Population Genomics Platform)

Table 2. Loss of heterozygosity detected in the proband

Copy Number	Gain, loss or LOH	Chr	Nucleotide start	Nucleotide stop	Size (kb)	Gene
2	LOH	X	49444309	54782812	5339	PAGE1, PAGE4, CLCN5, AKAP4, CCNB3, DGKK, SHROOM4, BMP15, NUDT10, NUDT11, MAGED1, MAGED4, MAGED4B, XAGE2, XAGE1A, XAGE1B, XAGE1C, XAGE1D, SSX2, SSX2B, SPANXN5, XAGE5, XAGE3, FAM156A, GPR173, TSPYL2, KDM5C, IQSEC2, SMC1A, RIBC1, HUWE, HSD17B10, PHF8, FAM120C, WNK3, TSR2, GNL3L, FGD1, ITIH6
2	LOH	X	61932503	67174031	5242	SPIN4, ARHGEF9, AMER1, ASB12, MTMR8, ZC4H2, LAS1L, MSN, VSIG4, HEPH, EDA2R, AR
2	LOH	X	125659991	130753416	5093	ACTRT1, CXORF64, DCAF12L1, SMARCA1, OCRL, APLN, XPNPEP2, SASH3, ZDHHC9, UTP14A, BCORL1, ELF4, AIFM1, RAB33A, ZNF280C, SCL25A14, GPR119, RBMX2, ENOX2, ARHGAP36, IGSF1, OR13H1

Abbreviations are as follows: Chr (chromosome),

Chapter 6: Discussion

6.1 Preface

The following chapter consists of excerpts from the manuscript titled "Receptor tyrosine kinase mutations in developmental syndromes and cancer: two sides of the same coin " submitted as an invited review to Human Molecular Genetics by Laura M. McDonell, Kristin D. Kernohan, Kym M. Boycott and Sarah L. Sawyer.

6.2 Statement of permission for use of copyrighted material

Statement of permission for the reproduction of copyrighted material from Oxford University Press can be found in Appendix A.

6.3 Contributions

The specific contributions of each author are as follows:

L. M. McDonell

Performed the literature review and data mining. Generated the table and figure. Wrote the manuscript.

Kristin D. Kernohan

Edited the manuscript.

Kym M. Boycott

Oversaw and edited the manuscript.

Sarah L. Sawyer

Oversaw and edited the manuscript.

6.4 Discussion

The introduction of high-throughput sequencing approaches revolutionized genetic research and our understanding of molecular genetics. It has contributed to a veritable explosion of information that has flooded the fields of rare disorders and cancer research. From this sea of information has emerged new knowledge about the proteins and pathways contributing to normal human development and the processes extending to post-natal life. With each novel discovery, a new thread of understanding has been weaved through our genetic tapestry, and, in time, we have come to appreciate the junction between a subset of rare developmental disorders and oncogenesis. The recognition of distinct oncological phenotypes in patients with genetic disorders was the first indication that a molecular overlap between these two fields may exist and indeed, many individuals with neurocutaneous disorders are predisposed to cancer.

In this work, the genetic etiology of two neurocutaneous disorders, MIC-CAP syndrome and ECCL is resolved. MIC-CAP is caused by loss-of-function mutations in *STAMBP* encoding a DUB with key roles in cell-surface receptor endocytosis and sorting. On the other hand, activating mutations in *FGFR1*, a receptor tyrosine kinase, and *KRAS*, a known oncogene, causes ECCL. Interestingly, patients with ECCL have a predisposition to low-grade CNS tumours, whereas, no cases of malignancy have been reported in MIC-CAP.

Given these findings, herein, the molecular relationship between cancer and rare developmental disorders will be further explored. The overlapping genomic alterations found in developmental syndromes and oncogenesis will be further investigated with a specific focus on genetic disorders caused by mutations in receptor tyrosine kinases (RTKs) since their activity often represent the first

step in modulating important signalling pathways such as Ras-MAPK and PI3K-AKT, as demonstrated in ECCL. This family of cell surface receptors is also of particular relevance given STAMBP's potential role in their endocytosis and ensuing signalling termination. Insights gained from this exploration will contribute to the elucidation of other unsolved neurocutaneous developmental disorders, such as PHACE syndrome.

6.4.1 Receptor Tyrosine Kinases

Receptor tyrosine kinases (RTKs) are a subclass of tyrosine kinases, which are involved in mediating *inter*cellular communication and orchestrating a wide range of complex biological functions.[1] Genetic studies have demonstrated a role for RTK signaling in developmental and acquired human disease.[2,3] Germline mutations in genes encoding RTKs cause several developmental syndromes, while somatic alterations contribute to the pathogenesis of many aggressive cancers. [2,3]

There are 58 RTKs identified to date, which can be subdivided into 20 subfamilies, all of which share a basic structure consisting of an *extra*cellular ligand-binding domain linked to an *intra*cellular protein kinase core via a single-pass transmembrane domain (Figure 1).[1,4] RTK activation is a complex biological process and has been reviewed elsewhere. [1] Briefly, canonical RTKs function by binding their specific ligand to induce dimerization and conformational changes.[1] As such, ligand activation leads to *trans*-autophosphorylation of tyrosine residues in the dimer/oligomer and activation of RTK catalytic activity.[5] RTK phosphorylation occurs in two phases, *intra*molecular followed by *inter*molecular.[1,6] In the first phase, *trans*-autophosphorylation between the dimer pair destabilizes *cis*-inhibition, permitting RTK catalytic activity.[1,6] Autophosphorylation of the tyrosine kinase domain (TKD) continues creating

phosphotyrosine-based binding sites. In the second phase, phosphotyrosine recognition motif containing cytoplasmic signalling proteins are recruited.[1,6] Once bound, these proteins are activated by phosphorylation to initiate *intra*cellular signalling pathways such as Ras-MAPK and PI3K-AKT that modulate cell proliferation, growth, survival, apoptosis, differentiation, morphogenesis, cell-cycle progression, migration, and autophagy.[1,6] In addition to dimerization and i*ntra*molecular phosphorylation, RTK signalling is also regulated by positive and negative feedback mechanisms, tissue-specific splicing, RTK post-translational modifications, endocytosis, and ligand availability.[1,7-9] In particular, RTK signalling is terminated by rapid internalization and sorting of the RTK-ligand complex to lysosomes.[8,9] Ubiquitination mediates the interaction of the RTKs and the sorting machinery at the cell surface and the endosomes; this occurs through RTK ubiquitination/deubiquitination by ubiquitin ligases and DUBs such as STAMBP.[8-10] Together, these mechanisms prevent unwanted protein kinase activation, and enable temporal and tissue specific kinase activity.

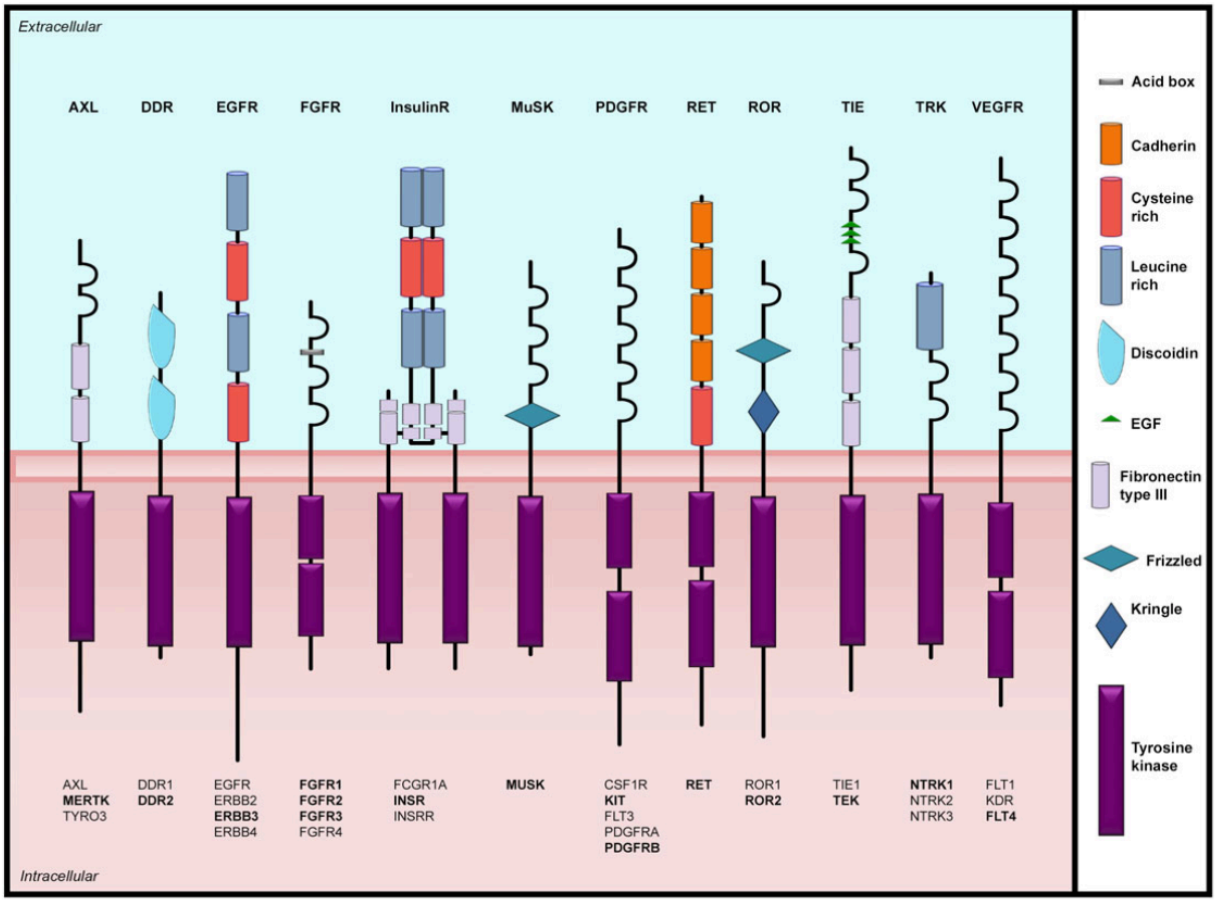


Figure 1. Receptor tyrosine kinase families involved in human developmental disorders. Schematic representation of the RTKs with all family members listed below each receptor. Receptors involved in developmental disease are indicated in bold. Structural domains are marked according to the key. This figure is reproduced from McDonell *et al*. (2015).

6.4.2 Kinasopathies

Germline mutations disrupting RTK signalling pathways have been identified as the cause of a number of congenital malformation syndromes; we refer to this collection of disorders as the 'developmental receptor tyrosine kinasopathies' (DRTKs). To date, at least 35 DRTKs have been described in the Online Mendelian Inheritance in Man (OMIM) database, caused by mutations in 15 RTK genes (Table 1).[11] Interestingly, skeletal abnormalities are over-represented in DRTKs; mutations in *DDR2*, *FGFR1, FGFR2, FGFR3* and *ROR2* are associated with over 20 clinically distinct skeletal dysplasias. Specific mutations in these genes during embryogenesis cause defects in osteoblast/osteoclast/chondrocyte cell proliferation, growth, differentiation and apoptosis resulting in abnormal bone morphogenesis.[12-15] Other affected systems include the nervous and endocrine systems. For example, a mutation in *ERBB3* has been associated with lethal congenital contractural syndrome type 2 [16], and a number of insulin resistance conditions are caused by dysregulated INSR signaling, including familial hyperinsulinemic hypoglycaemia [17] and Donohue [17] and Rabson-Mendenhall syndromes [18,19].

To evaluate the range of causative mutations in DRTKs, I conducted a review of coding pathogenic variants reported for each of the associated disorders (Table 1); of the more than 500 disease-associated mutations reported in the Human Gene Mutation Database (HGMD)[20] and ClinVar[21], over two thirds are missense mutations. Notably, both loss-of-function and gain-of-function mutations are observed, the latter being more common. For some of the dominant disorders, the mutations cluster exclusively in a given functional domain. For example, all 38 reported mutations causing autosomal dominant Hereditary lymphedema type 1A are localized within the two *intra*cellular TKDs of *FLT4* and have been shown to reduce receptor activation, suggesting a domain dependent mechanism for this disorder.[20-22] Extreme clustering is observed in

Achondroplasia, the most frequent form of skeletal dysplasia with short stature; the Gly380Arg mutation in *FGFR3* is observed in approximately 97% of patients.[23] Clearly, for a subset of the DRTKs, the type and location of the mutation has a very specific impact on the phenotypic outcome.

Table 1. Examples of developmental receptor tyrosine kinasopathies

Gene family	Gene symbol	OMIM number	Developmental disease (OMIM number)	Inheritance	Suspected Mechanism	Disease mutation and COSMIC mutation overlap. (Known drivers are bolded)
AXL	*MERTK*	604705	Retinitis pigmentosa 38 (613862)	AR	LOF	
DDR	*DDR2*	191311	Short limb-hand spondylometaepiphyseal dysplasia (271665)	AR	LOF	E113K, R752C
EGFR	*ERBB3*	190151	Lethal congenital contractural syndrome 2 (607598)	AR	LOF	
FGFR	*FGFR1*	136350	Hartsfield syndrome (615465)	AD/AR	?	
			Hypogonadotrophic hypogonadism 2 (147950)	AD	LOF	R250W, A343V, G703S, V795I
			Trigoncephaly (190440)	AD	?	
			Pfeiffer syndrome (101600)	AD	GOF	
			Osteoglophonic dysplasia (166250)	AD	GOF	
			Encephalocraniocutaneous lipomatosis (613001)	somatic	GOF	**K656E, N546K**
	FGFR2	176943	Antley-Bixler syndrome (207410)	AD	?	**W290C**
			Apert syndrome (101200)	AD	GOF	**S252W**, **P253R**
			Beare-Stevenson cutis gyrata syndrome (123790)	AD	GOF	S372C, Y375C
			Bent bone dysplasia syndrome (614592)	AD	?	
			Crouzon syndrome (123500)	AD	GOF	**S267P**, W290R, D549H, R678G
			Jackson-Weiss syndrome (123150)	AD	GOF	
			LADD syndrome (149730)	AD	LOF	A648T
			Pfeiffer syndrome (101600)	AD	GOF	**W290C**
	FGFR3	134934	Achondroplasia (100800)	AD	GOF	**G380R**
			Severe achondroplasia with developmental delay and acanthosis nigricans	AD	GOF	**K650M**
			Crouzon syndrome with acanthosis nigricans (612247)	AD?	GOF	**A391E**
			Hypochondroplasia (146000)	AD	GOF	**N540S, K650N, K650T, K650Q**
			LADD syndrome (149730)	AD	LOF	
			Muenke craniosynostosis (602849)	AD	GOF	
			Thanatophoric dysplasia I (187600)	AD	GOF	**R248C, S249C**, **G370C, S371C, Y373C**, **K65**
			Thanatophoric dysplasia II (187600)	AD	GOF	**K650E**
INSR	*INSR*	147670	Donohue syndrome (246200)	AR	LOF	R924*
			Rabson-Mendenhall syndrome (262190)	AR	LOF	
MUSK	*MUSK*	601296	Myasthenic syndrome 9 (616325)	AR	LOF	V790M
			Fetal akinesia deformation sequence (208150)	AR	LOF	
PDGFR	*KIT*	164920	Piebaldism (172800)	AD	LOF	W557*, F584L, G664R, R796G
	PDGFRß	173410	Infantile myofibromatosis (228550)	AD	?	
RET	*RET*	164761	Hirschsprung disease (142623)	AD	LOF	R77C, V145G, V202M, R231H, T278N, R330Q, R330W, R360W, A373V, E480K, R844W, G894S, R912Q, E921K, M980T
			Multiple endocrine neoplasia 2B	AD	GOF	**M918T**
ROR	*ROR2*	602337	Brachydactyly type B1 (113000)	AD	GOF	
			Robinow syndrome (268310)	AR	LOF	
TIE	*TEK*	600221	Multiple cutaneous and mucosal venous malformations (600195)	AD/somatic	GOF	R849W
TRK	*NTRK1*	191315	Insensitivity to pain, congenital, with anhidrosis (256800)	AR	LOF	
VEGFR	*FLT4*	136352	Hereditary lymphadema type IA (153100)	AD	LOF	G1024E, R1041Q, R1041W, R1114L, P113

*as reviewed in ref [24,25]

6.4.3 Phenotypic heterogeneity in developmental receptor tyrosine kinasopathies

Receptor tyrosine kinases possess intricate mechanisms to direct quantitatively and qualitatively distinct cell-type specific responses in precise developmental windows; these can include involvement of an accessory molecule, as well as differences in receptor and ligand expression levels and splice isoforms.[1,26-28] This perhaps explains how mutations in some RTK genes cause multiple developmental syndromes (Table 1). For example, gain-of-function mutations in the TKD of *FGFR1* cause ECCL[29] whereas gain-of-function mutations in the extracellular immunoglobulin domain of *FGFR1* are associated with Pfeiffer syndrome [30] and osteoglophonic dysplasia [31]. Interestingly, loss-of-function mutations in both the extracellular domain and TKD of *FGFR1* cause Hypogonadotrophic hypogonadism [32]. Similarly, gain-of-function mutations in *FGFR3* cause Achondroplasia, Severe achondroplasia with acanthosis nigricans, Crouzon syndrome with acanthosis nigricans, Hypochondroplasia, Muenke craniosynostosis, and Thanatophoric dysplasia type (TD) I and II (Table 1). Interestingly, specific substitutions of *FGFR3* at Lys650, have been reported to cause Hypochondroplasia with Lys650Asn/Gln mutations, TDII with Lys650Glu, and TD1 or Severe achondroplasia with acanthosis nigricans with Lys650Met, demonstrating that the nature and severity of the disease can be influenced by the specific change in a single amino acid.[33] Overall, the range of clinical disease resulting from mutations in each RTK is likely a balance of many factors such as the location and type of the mutation, the function of the given kinase isoform, and the mutation's impact on receptor integrity and kinase activity in the context of the individual's genetic background. These observations emphasize the complexity of phenotype-genotype associations in the RTKs and further investigation will be required to more fully understand these intricacies.

6.4.4 Genetic heterogeneity in developmental receptor tyrosine kinasopathies

Just like different mutations in the same RTKs can cause very different diseases, mutations in different RTKs can cause the same disease. This genetic heterogeneity is widely observed in the FGFR family. For example, heterozygous mutations in either *FGFR2* or *FGFR3* cause Crouzon syndrome, *FGFR1* and *FGFR2* cause Pfeiffer syndrome, and *FGFR2* and *FGFR3* cause lacrimo-auriculo-dento-digital (LADD) syndrome (Table 1). Additionally, mutations in signaling components which dysregulate RTK pathways can contribute to genetic heterogeneity. For example, a subset of LADD syndrome is caused by heterozygous mutations in *FGF10*, a FGFR ligand that interacts with *FGFR2*.[34,35] LADD-associated mutations in any of *FGFR2*, *FGFR3*, or *FGF10* result in reduced downstream signalling and this developmental disorder.[34,36] Moreover, mutations in signalling pathways downstream of the RTKs can also result in the same clinical presentation, as demonstrated by the pathogenic KRAS and FGFR1 mutations causative for ECCL. Overall, given the extensive and complex regulatory circuits for RTK signalling, there are often many molecules within a given pathway that can result in a similar clinical presentation.

6.4.5 Somatic mutations in receptor tyrosine kinasopathies and developmental syndromes

Save for the mosaic *FGFR1* mutations identified in individuals with ECCL, germline mutations underlie the majority of the DRTKs recognized to date (Table 1). There is an inherent bias for identification of germline mutations as they are readily detected in DNA extracted from blood, the DNA source used in most gene discovery studies. Somatic mutations are more challenging to identify and require high degree of clinical suspicion, access to appropriate patient tissue samples, and analysis by deep sequencing. Therefore, it is possible that a number of developmental disorders caused by somatic mutations in RTKs have yet to be identified. These disorders may present as a

milder or atypical form of a known disease, or as a novel condition. Mutations in *FGFR3* highlight this interesting paradigm; the Arg248Cys substitution typically results in TD1, though an individual with somatic mosaicism for that same substitution was reported with atypical features of Achondroplasia at 2 years of age.[37] Remarkably, this same *FGFR3* Arg248Cys substitution was identified in epidermal nevi (OMIM 162900) and absent from adjacent normal skin in a small number of individuals.[38] Moreover, germline mutations in *FGFR1* are associated with multiple disorders of skeletal malformation, however, mosaic mutations in this same gene lead to a novel phenotype as seen with the fatty alopeic nevus and CNS lipomas of ECCL. The contrast between the skeletal and epidermal phenotypes with these mutations is striking and likely reflects the type, timing and the point in embryonic lineage where the mutation arose.[29,31,39] It is quite probable that as deep sequencing becomes more widely available, novel developmental syndromes will be identified that are secondary to somatic mutations in RTKs.

6.4.6 Cancer predisposition and the developmental receptor tyrosine kinasopathies

Several RTKs have been recognized to contribute to both cancer and developmental syndromes (Table 1). This raises the question of whether patients with DRTKs, especially those with mutations also implicated in tumorigenesis, would be predisposed to certain types of cancer. Multiple endocrine neoplasia type IIB (MEN2B), caused by recurrent germline mutations in *RET,* is characterized by early aggressive medullary thyroid cancer, pheochromocytoma, mucosal neuromas, and a Marfanoid body habitus with dysmorphic facies.[40] Before the genetic etiology of ECCL was resolved, MEN2B was, to my knowledge, the only example of a DRTK with an inherited predisposition towards tumorigenesis. Patients with ECCL may present with intracranial and intraspinal lipomas, in addition to a spectrum of bony tumours including osteomas, odontomas,

and ossifying fibromas.[41] Moreover, low-grade gliomas have been reported in multiple individuals with ECCL.[42-46] Excluding individuals with MEN2B and ECCL, there is no clear evidence to support that patients with other DRTKs are at increased risk for cancer, though some cases have been reported. For example, two patients with Apert syndrome with germline Pro253Arg mutations in *FGFR2* were reported with cancer, one with early-onset low-grade papillary carcinoma of the bladder [47], and the other an ovarian dysgerminoma [48]. The Pro253Arg mutation has been reported in endometrial carcinomas and has been demonstrated to be oncogenic.[49] Notably, many of the DRTKs are characterized by a shortened lifespan, which would preclude cancer formation (eg. TD1 is associated with mortality in the neonatal period). So while RTKs have a clear and emerging role in cancer pathology, it seems that with the exception of MEN2B and ECCL, cancer is not prevalent in known DRTKs; the explanation for this apparent discordance is unknown but may reflect the timing and specific cellular environment of the mutation.

6.4.7 RTKs and cancer

In contrast to the ordered proliferation and differentiation of development, cancer represents an accumulation of genetic and epigenetic changes resulting in a disregard for the constraints of differentiation, proliferation, programmed cell death, and localization. By the time cancers reach an advanced state, genomic instability often results in hundreds of mutations, which can be categorized as either 'driver' mutations, those conferring a selective growth advantage to cells and are instrumental in cancer initiation or progression, or 'passenger' mutations, which are functionally neutral.[50,51] A number of the driver mutations identified occur in genes involved in key developmental pathways, such as gastrulation, angiogenesis, and patterning, and contribute to specific malignant phenotypes.[51,52]

Protein kinases, including RTKs, are one of the most frequently mutated gene families implicated in cancer, which has prompted numerous studies on their role in cancer pathogenesis (reviewed in [2,3,53]). There are four main mechanisms of RTK dysregulation in human cancers: genomic rearrangements, autocrine activation, overexpression, and gain- or loss-of-function mutations.[1,3] Unchecked RTK signaling can disrupt the balance between cell growth, cell-cycle progression, and apoptosis and when coupled with factors such as timing, location, duration, and strength of dysregulated RTK signaling, may sensitize cells to oncogenic transformation or trigger RTK-induced oncogenesis.[1,51]

6.4.8 Driver mutations in receptor tyrosine kinase genes

High throughput DNA sequencing of tumor tissues has begun to shed light on the complex genomic landscape of human cancers. Initiatives such as the Catalogue Of Somatic Mutations In Cancer (COSMIC) database archive genetic sequence with the ultimate goal of elucidating the molecular determinants of cancer.[54] Discovery and functional characterization of driver mutations is changing the understanding of cancer formation and progression, and providing opportunities for targeted treatments.[51,53] Distinguishing passenger from driver mutations is a central challenge in cancer genome analysis. Statistical approaches may be able to identify candidate cancer genes, but are not always able to predict the tumorigenic potential of individual mutations.[53] For example, activating mutations in the RTK *FLT* have been recognized to cause a common class of acute myeloid leukemia (AML).[55] Subsequent high-throughput *FLT3* sequencing in a cohort of AML patients followed by functional characterization for each identified variant confirmed constitutively active kinase activity in a subset of mutations (driver mutations) but revealed that many were likely passenger mutations.[56] Computationally, these passenger mutations could not be distinguished from

driver mutations, highlighting the need for functional validation studies and novel strategies to identify driver mutations.[56]

6.4.9 Kinases, kinasopathies and cancer: two sides of the same coin

Mutations in RTKs contribute to the pathogenesis of both cancer and developmental syndromes.[2,53] Notably, it has been recognized that some mutations which cause DRTKs are also drivers in somatic cancer (Table 1).[24,25] For example, the Gly380Arg mutation in *FGFR3* that causes Achondroplasia has been demonstrated to be a driver mutation in bladder cancer.[24,57] Similarly, somatic mutations in *FGFR2* were present in 12% of endometrial carcinomas [49], two of which (Ser252Trp, Pro253Arg) have been identified as driver mutations and are identical to germline mutations reported in Apert and Crouzon syndromes.[24,49,58] Moreover, the activating mosaic mutations (Asn546Lys and Lys656Glu) identified in *FGFR1* in individuals with ECCL are enriched in glioblastoma and pilocytic astrocytoma.[59,60] Interestingly, using WES, we identified a second *FGFR1* mutation (p.Val561Met) found exclusively in the pilocytic astrocytoma of an individual with ECCL who already harbored a known pathogenic *FGFR1* mosaic mutation (Lys656Glu).[29] The Val561Met substitution is thought to confer a 38-fold increase in phosphorylation of the FGFR1 receptor, as well as resistance to lucitanib, an FGFR inhibitor.[61,62]

I set out to investigate the overlapping subset of mutations that cause DRTKs and have also been identified in tumor tissue in COSMIC. Interestingly, mutations shared between developmental syndromes and cancer were either seen in a very small or very large number of tumor samples. For example, of the *FGFR3* gain-of-function mutations reported to cause TDI, six are in COSMIC and are found in 3105 samples; five of these mutations (Arg248Cys, Ser249Cys, Gly370Cys, Ser371Cys and Tyr373Cys) showed a high distribution in urinary tract neoplasia, while one

mutation (Lys650Met) was found more frequently in skin cancer.[54] FGFR variants that cause DRTKs have been shown to be oncogenic in several tumour types.[24] For instance, Lys650Glu (TDII), a recognized driver mutation in *FGFR3,* results in constitutively elevated kinase activity has been identified in spermatocytic seminoma [63], bladder carcinoma [64], multiple myeloma [65] and seborrheic keratosis [66]. Conversely, some variants showed a strong association with tissue-specific pathology. Unsurprisingly, ECCL mutations (Asn546Lys, Lys656Glu) were enriched in CNS tumors such as oligodendroglioma, primitive neuroectodermal medulloblastoma, ependymoma and low- to high-grade astrocytoma.[54] Interestingly, despite the activating potential of the p.Val561Met mutation, it was not found in COSMIC.

6.5 Conclusion and future directions

This data mining exercise allows us to reaffirm the genetic link between rare developmental disorders and cancer. More specifically, it confirms that overlapping genetic mutations contribute to developmental conditions and oncogenesis and that mutation timing, location, duration, and strength of dysregulated signaling modulate these outcomes. This may, in part, explain why individuals with DRTKs caused by known driver mutations, do not develop cancer. We can leverage the knowledge of these shared molecular profiles in two ways. First, mutations that cause rare developmental disorders, are inherently pathogenic and contribute, in some way, to altering cellular function. As such, finding a disorder-causing mutation in a tumor dataset may support a potential oncogenic contribution of that specific mutation. This may lead to the identification of new driver mutations in cancer and accelerate the development and use of targeted therapeutics. Conversely, knowledge of oncogenic mutations may help elucidate the genetic etiology of unsolved rare developmental disorders. Though high-throughput sequencing has greatly increased our

understanding of rare genetic disorders it has done so by flooding us with an overwhelming amount of data requiring painstaking filtering and analysis; this is only further exacerbated when hunting for mosaic mutations. This exercise has shown us that tumor mutation catalogues such as COSMIC, may be used as tools to prioritize variants and perhaps, in some cases, identify causative mutations. Had this approach been used in the ECCL cohort, it may have accelerated the identification of the *FGFR1* mutations. It was, however, by cross-referencing COSMIC that the pathogenic *KRAS* mutation was identified in an individual with ECCL that did not have alterations in *FGFR1*. It is clear that this approach will only benefit a subset of rare disorders, and indeed, no MIC-CAP mutations are found in COSMIC. In this instance, the role of the causative gene, *STAMBP* is to facilitate sorting of cell surface receptors and ensure appropriate signal termination. As discovered in this work, though disruption of STAMBP's activity leads to increased signaling through Ras-MAPK and PI3K-AKT this may be overshadowed, at least clinically, by the apoptotic phenotype resulting from cytoplasmic ubiquitinated protein aggregation. This approach for developmental disorders characterized by predisposition to cancer, overgrowth or clinical features suggestive of mosaicism will likely yield the most discoveries. PHACE syndrome, despite our best efforts, remains unsolved; the lessons provided herein, hopefully contributing to the eventual discovery of the pathogenic PHACE syndrome mutations.

6.6 References

1 Lemmon, M. A. & Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117-1134, (2010).

2 Lahiry, P., Torkamani, A., Schork, N. J. & Hegele, R. A. Kinase mutations in human disease: interpreting genotype-phenotype relationships. *Nat. Rev. Genet.* **11**, 60-74, (2010).

3 Blume-Jensen, P. & Hunter, T. Oncogenic kinase signalling. *Nature* **411**, 355-365, (2001).

4 Robinson, D. R., Wu, Y. M. & Lin, S. F. The protein tyrosine kinase family of the human genome. *Oncogene* **19**, 5548-5557, (2000).

5 Heldin, C. H. Dimerization of cell surface receptors in signal transduction. *Cell* **80**, 213-223, (1995).

6 Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **103**, 211-225, (2000).

7 Casaletto, J. B. & McClatchey, A. I. Spatial regulation of receptor tyrosine kinases in development and cancer. *Nat. Rev. Cancer* **12**, 387-400, (2012).

8 Haglund, K. *et al.* Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. *Nat. Cell Biol.* **5**, 461-466, (2003).

9 Adrain, C. & Freeman, M. Regulation of receptor tyrosine kinase ligand processing. *Cold Spring Harb. Perspect. Biol.* **6**, (2014).

10 McDonell, L. M., Kernohan, K. D., Boycott, K. M. & Sawyer, S. L. Receptor tyrosine kinase mutations in developmental syndromes and cancer: two sides of the same coin. *Hum. Mol. Genet.* **24**, R60-66, (2015).

11 Amberger, J. S., Bocchini, C. A., Schiettecatte, F., Scott, A. F. & Hamosh, A. OMIM.org: Online Mendelian Inheritance in Man (OMIM(R)), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* **43**, D789-798, (2015).

12 Labrador, J. P. *et al.* The collagen receptor DDR2 regulates proliferation and its elimination leads to dwarfism. *EMBO reports* **2**, 446-452, (2001).

13 Ornitz, D. M. & Marie, P. J. FGF signaling pathways in endochondral and intramembranous bone development and human genetic disease. *Genes Dev.* **16**, 1446-1465, (2002).

14 Billiard, J. *et al.* The orphan receptor tyrosine kinase Ror2 modulates canonical Wnt signaling in osteoblastic cells. *Mol. Endocrinol.* **19**, 90-101, (2005).

15 Maeda, K. *et al.* Wnt5a-Ror2 signaling between osteoblast-lineage cells and osteoclast precursors enhances osteoclastogenesis. *Nat. Med.* **18**, 405-412, (2012).

16 Narkis, G. *et al.* Lethal congenital contractural syndrome type 2 (LCCS2) is caused by a mutation in ERBB3 (Her3), a modulator of the phosphatidylinositol-3-kinase/Akt pathway. *Am. J. Hum. Genet.* **81**, 589-595, (2007).

17 Hojlund, K. *et al.* A novel syndrome of autosomal-dominant hyperinsulinemic hypoglycemia linked to a mutation in the human insulin receptor gene. *Diabetes* **53**, 1592-1598, (2004).

18 Psiachou, H. *et al.* Leprechaunism and homozygous nonsense mutation in the insulin receptor gene. *Lancet* **342**, 924, (1993).

19 Thiel, C. T. *et al.* Two novel mutations in the insulin binding subunit of the insulin receptor gene without insulin binding impairment in a patient with Rabson-Mendenhall syndrome. *Mol. Genet. Metab.* **94**, 356-362, (2008).

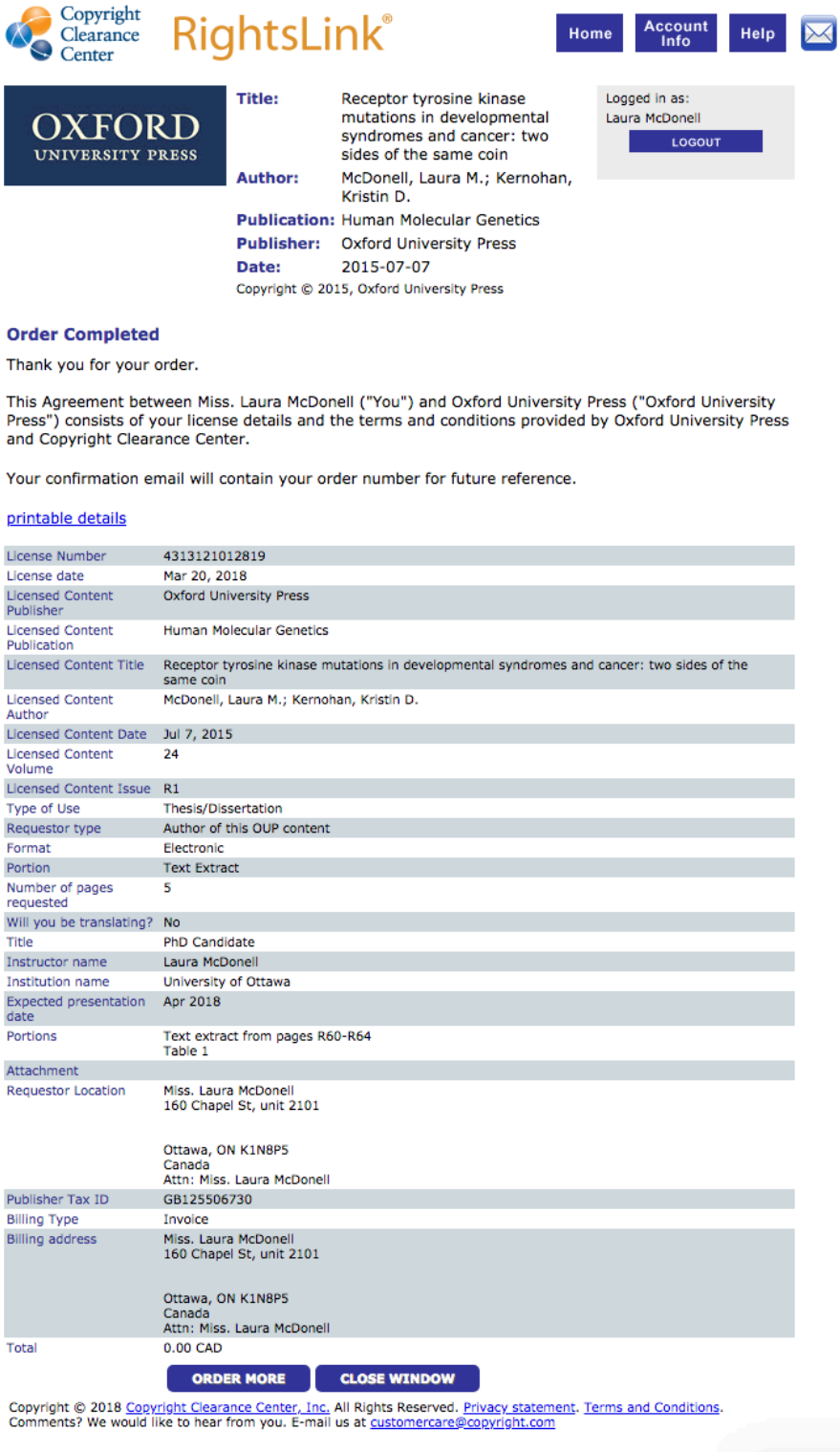
20 Stenson, P. D. *et al.* The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum. Genet.* **133**, 1-9, (2014).
21 Landrum, M. J. *et al.* ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* **42**, D980-985, (2014).
22 Karkkainen, M. J. *et al.* Missense mutations interfere with VEGFR-3 signalling in primary lymphoedema. *Nat. Genet.* **25**, 153-159, (2000).
23 Bellus, G. A. *et al.* Achondroplasia Is Defined by Recurrent G380R Mutations of FGFR3. *Am. J. Hum. Genet.* **56**, 367-373, (1995).
24 Greulich, H. & Pollock, P. M. Targeting mutant fibroblast growth factor receptors in cancer. *Trends Mol. Med.* **17**, 283-292, (2011).
25 Jhiang, S. M. The RET proto-oncogene in human cancers. *Oncogene* **19**, 5590-5597, (2000).
26 Beenken, A. & Mohammadi, M. The FGF family: Biology, pathophysiology and therapy. *Nat. Rev. Drug Discov.* **8**, 235-253, (2009).
27 Li, A. W. & Murphy, P. R. Expression of alternatively spliced FGF-2 antisense RNA transcripts in the central nervous system: regulation of FGF-2 mRNA translation. *Mol. Cell. Endocrinol.* **162**, 69-78, (2000).
28 <Regulation of Endocytic Sorting by ESCRT–DUB-Mediated.pdf>.
29 Bennett, J. T. *et al.* Mosaic Activating Mutations in FGFR1 Cause Encephalocraniocutaneous Lipomatosis. *Am. J. Hum. Genet.* **98**, 579-587, (2016).
30 Schell, U. *et al.* Mutations in FGFR1 and FGFR2 cause familial and sporadic Pfeiffer syndrome. *Hum. Mol. Genet.* **4**, 323-328, (1995).
31 White, K. E. *et al.* Mutations that cause osteoglophonic dysplasia define novel roles for FGFR1 in bone elongation. *Am. J. Hum. Genet.* **76**, 361-367, (2005).
32 Pitteloud, N. *et al.* Mutations in fibroblast growth factor receptor 1 cause both Kallmann syndrome and normosmic idiopathic hypogonadotropic hypogonadism. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 6281-6286, (2006).
33 Bellus, G. A. *et al.* Distinct missense mutations of the FGFR3 lys650 codon modulate receptor kinase activation and the severity of the skeletal dysplasia phenotype. *Am. J. Hum. Genet.* **67**, 1411-1421, (2000).
34 Shams, I. *et al.* Lacrimo-auriculo-dento-digital syndrome is caused by reduced activity of the fibroblast growth factor 10 (FGF10)-FGF receptor 2 signaling pathway. *Mol. Cell. Biol.* **27**, 6903-6912, (2007).
35 Rohmann, E. *et al.* Mutations in different components of FGF signaling in LADD syndrome. *Nat. Genet.* **38**, 414-417, (2006).
36 Lew, E. D., Bae, J. H., Rohmann, E., Wollnik, B. & Schlessinger, J. Structural basis for reduced FGFR2 activity in LADD syndrome: Implications for FGFR autoinhibition and activation. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 19802-19807, (2007).
37 Takagi, M., Kaneko-Schmitt, S., Suzumori, N., Nishimura, G. & Hasegawa, T. Atypical achondroplasia due to somatic mosaicism for the common thanatophoric dysplasia mutation R248C. *Am. J. Med. Genet. A* **158A**, 247-250, (2012).
38 Hafner, C. *et al.* Mosaicism of activating FGFR3 mutations in human skin causes epidermal nevi. *J. Clin. Invest.* **116**, 2201-2207, (2006).
39 Muenke, M. *et al.* A common mutation in the fibroblast growth factor receptor 1 gene in Pfeiffer syndrome. *Nat. Genet.* **8**, 269-274, (1994).
40 Morrison, P. J. & Nevin, N. C. Multiple endocrine neoplasia type 2B (mucosal neuroma syndrome, Wagenmann-Froboese syndrome). *J. Med. Genet.* **33**, 779-782, (1996).

41 Moog, U. Encephalocraniocutaneous lipomatosis. *J. Med. Genet.* **46**, 721-729, (2009).
42 Brassesco, M. S. *et al.* Low-grade astrocytoma in a child with encephalocraniocutaneous lipomatosis. *J. Neurooncol.* **96**, 437-441, (2010).
43 Bieser, S. *et al.* Grade II pilocytic astrocytoma in a 3-month-old patient with encephalocraniocutaneous lipomatosis (ECCL): case report and literature review of low grade gliomas in ECCL. *Am. J. Med. Genet. A* **167A**, 878-881, (2015).
44 Kocak, O., Yarar, C. & Carman, K. B. Encephalocraniocutaneous lipomatosis, a rare neurocutaneous disorder: report of additional three cases. *Childs Nerv. Syst.* **32**, 559-562, (2016).
45 Valera, E. T. *et al.* Are patients with encephalocraniocutaneous lipomatosis at increased risk of developing low-grade gliomas? *Childs Nerv. Syst.* **28**, 19-22, (2012).
46 Phi, J. H. *et al.* Papillary glioneuronal tumor present in a patient with encephalocraniocutaneous lipomatosis: case report. *Neurosurgery* **67**, E1165-1169, (2010).
47 Andreou, A. *et al.* Early-onset low-grade papillary carcinoma of the bladder associated with Apert syndrome and a germline FGFR2 mutation (Pro253Arg). *Am. J. Med. Genet. A* **140**, 2245-2247, (2006).
48 Rouzier, C. *et al.* Ovarian dysgerminoma and Apert syndrome. *Pediatr. Blood Cancer* **50**, 696-698, (2008).
49 Dutt, A. *et al.* Drug-sensitive FGFR2 mutations in endometrial carcinoma. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 8713-8717, (2008).
50 Negrini, S., Gorgoulis, V. G. & Halazonetis, T. D. Genomic instability--an evolving hallmark of cancer. *Nat. Rev. Mol. Cell Biol.* **11**, 220-228, (2010).
51 Pon, J. R. & Marra, M. A. Driver and passenger mutations in cancer. *Annu. Rev. Pathol.* **10**, 25-50, (2015).
52 Bellacosa, A. Developmental disease and cancer: biological and clinical overlaps. *Am. J. Med. Genet. A* **161A**, 2788-2796, (2013).
53 Torkamani, A., Verkhivker, G. & Schork, N. J. Cancer driver mutations in protein kinase genes. *Cancer Lett.* **281**, 117-127, (2009).
54 Forbes, S. A. *et al.* COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. *Nucleic Acids Res.* **38**, D652-657, (2010).
55 Reindl, C. *et al.* Point mutations in the juxtamembrane domain of FLT3 define a new class of activating mutations in AML. *Blood* **107**, 3700-3707, (2006).
56 Frohling, S. *et al.* Identification of driver and passenger mutations of FLT3 by high-throughput DNA sequence analysis and functional assessment of candidate alleles. *Cancer Cell* **12**, 501-513, (2007).
57 van Rhijn, B. W. *et al.* Novel fibroblast growth factor receptor 3 (FGFR3) mutations in bladder cancer previously identified in non-lethal skeletal disorders. *Eur. J. Hum. Genet.* **10**, 819-824, (2002).
58 Pollock, P. M. *et al.* Frequent activating FGFR2 mutations in endometrial carcinomas parallel germline mutations associated with craniosynostosis and skeletal dysplasia syndromes. *Oncogene* **26**, 7158-7162, (2007).
59 Jones, D. T. *et al.* Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nat. Genet.* **45**, 927-932, (2013).
60 Rand, V. *et al.* Sequence survey of receptor tyrosine kinases reveals mutations in glioblastomas. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 14344-14349, (2005).

61 Sohl, C. D., Ryan, M. R., Luo, B., Frey, K. M. & Anderson, K. S. Illuminating the molecular mechanisms of tyrosine kinase inhibitor resistance for the FGFR1 gatekeeper mutation: the Achilles' heel of targeted therapy. *ACS Chem. Biol.* **10**, 1319-1329, (2015).
62 Soria, J. C. *et al.* Phase I/IIa study evaluating the safety, efficacy, pharmacokinetics, and pharmacodynamics of lucitanib in advanced solid tumors. *Ann. Oncol.* **25**, 2244-2251, (2014).
63 Goriely, A. *et al.* Activating mutations in FGFR3 and HRAS reveal a shared genetic origin for congenital disorders and testicular tumors. *Nat. Genet.* **41**, 1247-1252, (2009).
64 Cappellen, D. *et al.* Frequent activating mutations of FGFR3 in human bladder and cervix carcinomas. *Nat. Genet.* **23**, 18-20, (1999).
65 Chesi, M. *et al.* Activated fibroblast growth factor receptor 3 is an oncogene that contributes to tumor progression in multiple myeloma. *Blood* **97**, 729-736, (2001).
66 Logie, A. *et al.* Activating mutations of the tyrosine kinase receptor FGFR3 are associated with benign skin tumors in mice and humans. *Hum. Mol. Genet.* **14**, 1153-1160, (2005).

Appendices

Appendix A. Statement of permission for the reproduction of copyrighted material from Oxford University Press



Copyright Clearance Center RightsLink®

Home | Account Info | Help

OXFORD UNIVERSITY PRESS

Title: Receptor tyrosine kinase mutations in developmental syndromes and cancer: two sides of the same coin

Author: McDonell, Laura M.; Kernohan, Kristin D.

Publication: Human Molecular Genetics

Publisher: Oxford University Press

Date: 2015-07-07

Copyright © 2015, Oxford University Press

Logged in as:
Laura McDonell
LOGOUT

Order Completed

Thank you for your order.

This Agreement between Miss. Laura McDonell ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

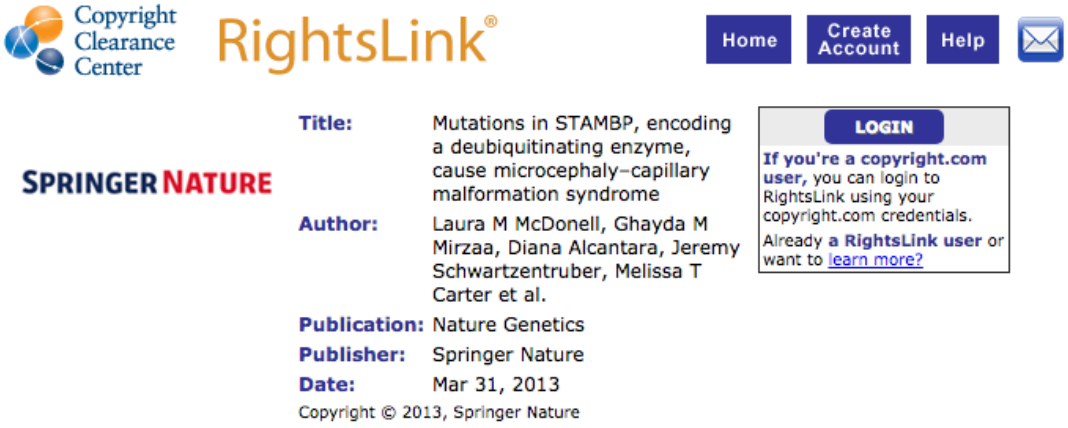
printable details

License Number	4313121012819
License date	Mar 20, 2018
Licensed Content Publisher	Oxford University Press
Licensed Content Publication	Human Molecular Genetics
Licensed Content Title	Receptor tyrosine kinase mutations in developmental syndromes and cancer: two sides of the same coin
Licensed Content Author	McDonell, Laura M.; Kernohan, Kristin D.
Licensed Content Date	Jul 7, 2015
Licensed Content Volume	24
Licensed Content Issue	R1
Type of Use	Thesis/Dissertation
Requestor type	Author of this OUP content
Format	Electronic
Portion	Text Extract
Number of pages requested	5
Will you be translating?	No
Title	PhD Candidate
Instructor name	Laura McDonell
Institution name	University of Ottawa
Expected presentation date	Apr 2018
Portions	Text extract from pages R60-R64 Table 1
Attachment	
Requestor Location	Miss. Laura McDonell 160 Chapel St, unit 2101 Ottawa, ON K1N8P5 Canada Attn: Miss. Laura McDonell
Publisher Tax ID	GB125506730
Billing Type	Invoice
Billing address	Miss. Laura McDonell 160 Chapel St, unit 2101 Ottawa, ON K1N8P5 Canada Attn: Miss. Laura McDonell
Total	0.00 CAD

ORDER MORE | CLOSE WINDOW

Copyright © 2018 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement. Terms and Conditions.
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Appendix B. Statement of permission for the reproduction of copyrighted material from Springer Nature



Copyright Clearance Center RightsLink®

Home | Create Account | Help

SPRINGER NATURE

Title: Mutations in STAMBP, encoding a deubiquitinating enzyme, cause microcephaly–capillary malformation syndrome

Author: Laura M McDonell, Ghayda M Mirzaa, Diana Alcantara, Jeremy Schwartzentruber, Melissa T Carter et al.

Publication: Nature Genetics

Publisher: Springer Nature

Date: Mar 31, 2013

Copyright © 2013, Springer Nature

LOGIN

If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials.

Already **a RightsLink user** or want to learn more?

Author Request

If you are the author of this content (or his/her designated agent) please read the following. If you are not the author of this content, please click the Back button and select no to the question "Are you the Author of this Springer Nature content?".

Ownership of copyright in original research articles remains with the Author, and provided that, when reproducing the contribution or extracts from it or from the Supplementary Information, the Author acknowledges first and reference publication in the Journal, the Author retains the following non-exclusive rights:

To reproduce the contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).

The author and any academic institution, where they work, at the time may reproduce the contribution for the purpose of course teaching.

To reuse figures or tables created by the Author and contained in the Contribution in oral presentations and other works created by them.

To post a copy of the contribution as accepted for publication after peer review (in locked Word processing file, of a PDF version thereof) on the Author's own web site, or the Author's institutional repository, or the Author's funding body's archive, six months after publication of the printed or online edition of the Journal, provided that they also link to the contribution on the publisher's website.

Authors wishing to use the published version of their article for promotional use or on a web site must request in the normal way.

If you require further assistance please read Springer Nature's online author reuse guidelines.

For full paper portion: Authors of original research papers published by Springer Nature are encouraged to submit the author's version of the accepted, peer-reviewed manuscript to their relevant funding body's archive, for release six months after publication. In addition, authors are encouraged to archive their version of the manuscript in their institution's repositories (as well as their personal Web sites), also six months after original publication.

v1.0

BACK | CLOSE WINDOW

Appendix C. Statement of permission for the reproduction of copyrighted material from the American Journal of Human Genetics

Copyright Clearance Center RightsLink®

Home | Create Account | Help

Title: Mosaic Activating Mutations in FGFR1 Cause Encephalocraniocutaneous Lipomatosis

Author: James T. Bennett,Tiong Yang Tan,Diana Alcantara,Martine Tétrault,Andrew E. Timms,Dana Jensen,Sarah Collins,Malgorzata J.M. Nowaczyk,Marjorie J. Lindhurst,Katherine M. Christensen,Stephen R. Braddock,Heather Brandling-Bennett,Raoul C.M. Hennekam et al.

Publication: The American Journal of Human Genetics

Publisher: Elsevier

Date: 3 March 2016

Copyright © 2016 The American Society of Human Genetics. Published by Elsevier Inc. All rights reserved.

LOGIN

If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials.

Already **a RightsLink user** or want to learn more?

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights

BACK | CLOSE WINDOW

LAURA M. MCDONELL

mcdonelllaura@gmail.com
lmcdo050@uottawa.ca

EDUCATION

2012 - 2018	**MD/PhD in Biochemistry and Molecular Genetics**, University of Ottawa Thesis title: Disruption of RAS-MAPK signalling in human neurocutaneous disorders Supervisors: Dr. Kym Boycott and Dr. Dennis Bulman
2010 - 2012	MSc in Biochemistry, University of Ottawa (transferred to PhD in 2012)
2006 - 2010	BSc in Biopharmaceutical Sciences with Honours, University of Ottawa

NOTABLE AWARDS

2013, 2016, 2017	Baxter and Alma Ricard Foundation Scholarship
2012 - 2017	Faculty of Medicine Priority Scholarship
2014 - 2016	CIHR Doctoral Award
2015, 2016	Audrey J. Boyce Research Fellowship
2015	CNFS Doctoral Award
2013, 2014	CIHR summer Studentship Award
2011	CIHR Frederick Banting Scholarship
2011	National Excellence Scholarship

COMMUNITY + FACULTY INVOLVEMENT

2016	Volunteer, Celtic Festival
2016	Guest speaker, Operation Medical School
2009 - 2014	Volunteer, Bruyère Continuing Care
2014	Organizer, uOttawa medical student wellness conference
2012 - 2014	Research supervisor, undergradute honours students
2013	Educator, Brain Health Day,
2013	Educator, uOttawa Mini Medical School,
2013	Volunteer, Résidence St-Louis' weekend respite program
2011 - 2012	Volunteer, Genetics Department, CHEO
2010 - 2012	Educator, Let's Talk Science
2007 - 2012	Volunteer, Canadian Blood Services

LANGUAGES

French	Spoken and written
English	Spoken and written
German	Spoken and written with questionable grammar

PUBLICATIONS

James T. Bennett, et al. Mosaic Activating Mutations in FGFR1 Cause Encephalocraniocutaneous Lipomatosis. **Am J Hum Genet** 98, 579-587 (2016) - *Published as Senior Author*

McDonell, L.M., et al. Receptor tyrosine kinase mutations in developmental syndromes and cancer: two sides of the same coin. **Hum Mol Genet** 24, R60-66 (2015)

McDonell, L.M. et al. A novel mutation in WDR62 gene is associated with autosomal recessive primary microcephaly in French Canadian adult siblings. **BMC Neurology** 14(22) (2014)

Carter, M.T., Ghayda, M., McDonell, L.M., Boycott, K.M., Microcephaly-Capillary Malformation Syndrome. **GeneReviews** (2014)

McDonell, L.M. et al Mutations in STAMBP, encoding a deubiquitinating enzyme, cause Microcephaly-Capillary Malformation syndrome. **Nat Genet** 45(5):556-562 (2013)

McDonell, L. & Drouin, G. The abundance of processed pseudogenes derived from glycolytic genes is correlated with their expression level. **Genome** 55, 147-151 (2012).

Armour, C. M. et al. 17p13.3 microduplications are associated with split-hand/foot malformation and long-bone deficiency (SHFLD). **Eur. J. Hum. Genet.** 19, 1144-1151 (2011)

ACTIVE MANUSCRIPTS

Accepted To Cell:
A mild X mutation is associated with adult-onset ataxia, whereas haploinsufficiency causes developmental delay and seizures

Submitted
Atypical Hepatic Mesenchymal Hamartoma: Histologic Appearance, Immunophenotype, and Molecular Findings

A mutation in X causes growth restriction, craniofacial dysmorphism, alopecia, and hypoplastic kidneys in New Mexican Hispanic families

Mosaic X Mutation in a Patient with Encephalocraniocutaneous Lipomatosis and Renovascular Hypertension

Ready for submission

Understanding the molecular etiology of PHACE syndrome: approaches and considerations