Lysosomal signaling is required for embryonic stem cell differentiation and human development

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Florian Villegas

Aus Frankreich

Basel, 2018

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultat

auf Antrag von

Prof. Dr. Susan Gasser

Dr. Joerg Betschinger

Dr. David Teis

Basel, den 14.11.2017

Prof. Dr. Martin Spiess (Dekan)

Acknowledgements

The work presented in this thesis would have not been possible without the support and contribution of many people.

First and foremost, I would like to thank my PhD advisor, Doctor Joerg Betschinger, for the opportunity to work in his lab and for being a great mentor. I really enjoyed our endless scientific and non-scientific philosophical discussions. I am grateful for every advice you gave me over the past years.

I want to thank all the lab members, for the great scientific atmosphere and the moments we shared together. I especially would like to thank Daniel Olivieri for mentoring and scientific discussions. Verena Waehle, you have been a great colleague and became a friend. Matyas Flemr, thank you for your collaboration in this work and your precious help. I have learned a lot from you.

I would like to thank the members of my thesis committee, Professor Susan Gasser, Professor Prisca Liberali and Professor David Teis for taking the time to discuss and offer valuable input on my project.

I would like to thank many people who played vital roles along the way. I believe I am not able to identify all of them but essentially all the persons I have met here at the Friedrich Miescher Institute. You all make this institute a very pleasant and stimulating place to do good science.

Je tiens à remercier ma famille pour son éternel soutien, essentiellement pendant les moments les plus difficiles rencontrés au cours de mon Doctorat.
Jessica, certainement la plus belle rencontre que j'ai faite au cours de ces deux dernières années. Merci d'être la chaque jour à mes cotés, sans ton soutien, beaucoup de choses ne se seraient pas aussi bien passées.

Table of content

List of abbreviations

AMPK	AMP-activated protein kinase
BHD	Birt-Hogg-Dubé syndrome
bHLH	basic Helix Loop Helix
BMP	bone morphogenic protein
bp	basepairs
CLEAR	Coordinated Lysosomal Expression and Regulation
CoA	Coenzyme A
ConcA	Concanamycin A
CRISPR	Clustered Regulatory Interspaced Short Palindromic Repeats
DNA	deoxyribonucleic acid
E	embryonic day
EAA	essential amino- acid
ESC	Embryonic stem cell
FLCN	Folliculin
FNIP	Folliculin interacting protein
GAP	GTPase-activating protein
GEF	Guanine nucleotide exchange factor
GDP	Guanosine diphosphate
GTP	Guanosine triphosphate
GSK	glycogen synthase kinase
ICM	inner cell mass
IP	immunoprecipitation
KO	Knock-out
LIF	Leukemia Inhibitory Factor
LOF	Loss of function
LTS	Lysosomal targeting sequence
mTOR	Mammalian target of rapamycin
MS	Mass-spectrometry
NAD	nicotinamide adenine dinucleotide
NEAA	non-essential amino-acid
OxPhos	oxidative phosphorylation
PSC	Pluripotent stem cell
Rag	Ras-related GTP-binding protein
Rheb	Ras homologue enriched in brain
RNA	ribonucleic acid
ROS	reactive oxygen species
TCA	tricarboxylic acid
TSC	Tuberous sclerosis protein
WT	Wild type

1 Summary

Self-renewal and differentiation are the two key properties that characterize stem cells. To initiate differentiation, pluripotent embryonic stem cells (ESC) must be forced out of self-renewal, the transcriptional networks conferring stem cell identity need to be dissolved and lineages must be chosen. Compared to acquisition and maintenance of ESCs pluripotency, the mechanisms driving exit from this cell state are ill defined. Previous work has identified the tumor suppressor Folliculin (Flcn) as a novel gene required for exit from the ESC state during differentiation (Betschinger et al., 2013a). Flcn loss of function in mice is early embryonic lethal consistent with being required for exit from pluripotency *in vivo*. It has been shown in ESCs that Flcn and two associated proteins, Fnip1 and Fnip2, regulate the subcellular localization of the basic helix-loop-helix transcription factor Tfe3. In wildtype ESCs, Tfe3 is ubiquitously detected in the nucleus and cytoplasm. Upon ESC differentiation, Tfe3 is excluded from the nucleus whereas Flcn loss of function ESCs harbor constitutive nuclear Tfe3 localization and are, consequentially, impaired in differentiation. Consistently, ectopic nuclear Tfe3 similarly inhibits ESC differentiation. Recently, Flcn has been identified in a metabolic pathway leading to the activation of the mTORC1 complex in an amino-acid sensing dependent manner after nutrient starvation (Tsun et al., 2013). However, mTORC1 activity is not required for ESC differentiation.

To gain insights into the regulation of this process, we conducted a genome-wide CRISPR/Cas9 screen in ESCs that aimed to identify genes that, similar to Flcn, are required for the maintenance of stem cell identity upon long term exposure to differentiation stimuli. Although molecularly related to amino-acid sensing dependent mTORC1 regulation, our findings suggest an alternate mode of action that may be founded in the non-starving nature of ESC differentiation. In addition, we show in ESC that the lysosome regulates the assembly of a Flcn-Tfe3 axis, whose impairment induces an embryonic developmental arrest.

Additionally, in a collaborative work with human geneticists, we identified five Tfe3 point mutations in a human syndrome called hypomelanosis of Ito. Symptoms associated with this developmental disorder include severe intellectual disability,

coarse facial features, frontonasal dysplasia, obesity, epilepsy and growth retardation, suggesting a pleiotropic developmental disorder. To understand the developmental nature of the disease, we tested these mutations in ESCs and we found them causing defects in exit from self-renewal and nuclear Tfe3 accumulation.

Taken together, the work presented in this thesis has allowed the identification of the lysosome as a critical organelle required for the control of the exit from pluripotency and indicates that lysosomal recruitment of Tfe3 is required for its inactivation, which is a crucial step to enable normal developmental progression.

2 Introduction

2.1 Murine embryonic stem cells

2.1.1 Origin of ESCs

In mammals, embryos develop from the fusion of a sperm cell and the oocyte, two very specialized cell types. After fertilization, the newly created zygote grows up to a very dense cellular sphere and eventually compacts to reach the early blastocyst stage at embryonic day (E) 3.5. At this stage, the blastocyst is a layer of epithelial cells, the trophectoderm, encapsulating the inner cell mass (ICM). Murine pluripotent embryonic stem cells (ESCs) are derived from the mid-blastocyst-stage embryo between E 3.75 and E 4.5. At this stage, the ICM segregates into epiblast and hypoblast and ESCs derivation *in vitro* pauses developmental progression from the epiblast. ESCs can differentiate *in vitro* into all cell types of the embryo as well as *in vivo* when transplanted back into the blastocyst. This unique feature of ESCs recapitulates the pluripotency of the pre-implantation epiblast. Upon developmental progression *in vivo* when the embryo implants into the uterine wall (E 5.5), or upon differentiation *in vitro*, pluripotency potential decreases rapidly. In fact, ESCs self-renew indefinitely *in vitro* while pluripotency exists only transiently *in vivo*. ESCs display unique properties that make them attractive for regenerative medicine and in-depth characterization will help to facilitate their manipulation *in vitro*.

2.1.2 Maintenance of the ground state of pluripotency

Originally ESCs were directly isolated from the late blastocyst and were maintained in culture *in vitro* on a feeder layer of mitotically inactivated fibroblasts in the presence of serum, allowing ESCs to proliferate without losing their pluripotency, which is called self-renewal ((Evans and Kaufman, 1981); (Martin, 1981)). Early studies have shown that feeder cells secrete Leukemia Inhibitory Factor (LIF), a cytokine promoting the maintenance of pluripotency (Smith et al., 1988). Therefore, the feeders have been replaced by addition of LIF directly in the culture medium. Furthermore, bone morphogenic protein 4 (BMP4) was identified as the key

inhibitor of differentiation present in the serum. A fine balance between stimuli promoting self-renewal and stimuli inhibiting differentiation regulates the maintenance of the stem cell pool and needs to be tightly controlled when cells are cultured *in vitro*. ESCs grown in Serum/LIF condition give rise to a heterogeneous cell population of cells in formative pluripotent state that corresponds to the peri-implantation stage *in vivo*. Naïve and primed pluripotent stem cells display key differences in their derivation of germline-competent PSCs, epigenomic states, gene expression of naïve pluripotency markers and lineage-specific genes, signaling requirements to maintain self-renewal, and central carbon metabolism (Davidson et al., 2015). Over the past decade, genetic screens have helped to identify gene regulatory networks governing ESC identity and enabled the development of modulators targeting specific signaling pathways involved in pluripotency regulation. In particular, modulation of both Wnt and the Fgf signaling pathways resulted in a more homogenous ESC culture as compared to culture in the presence of serum and LIF. Mechanistically, Chiron (CHIR99021), a glycogen synthase kinase 3 (GSK3) inhibitor and thus Wnt signaling activator, promotes the pluripotency gene network therefore enhancing the self-renewal ability of ESCs. In contrast, PD03 (PD0325901) impairs the initiation of differentiation by inhibiting Fgf signaling (Figure1). Altogether, this formed the basis for the formulation of culture conditions required for the maintenance of a homogenous population of ESCs in the ground state of pluripotency *in vitro,* so called 2i/LIF (Ying et al., 2008).

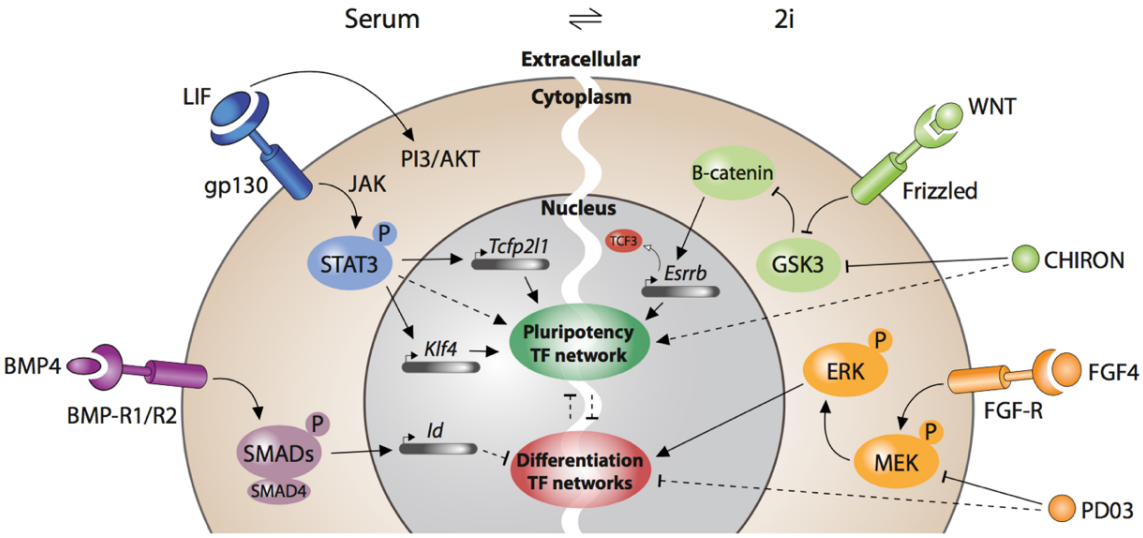


Figure 1: Signaling pathways promoting pluripotency and inhibiting differentiation

Schematic representation of pathways activated in the presence of serum and LIF in the culture medium (serum/LIF condition - left part) or in the presence of 2 inhibitors, Chiron and PDO3 (2i conditions – right part). Both conditions result in the maintenance of a pluripotent state *in vitro*. Present in the serum, BMP4 activates the SMAD signaling cascade, upstream of the Id proteins that inhibit differentiation. Activation of the JAK/STAT signaling pathway by LIF in ESCs leads to activation of Tcfp2l1 and Klf4, key transcription factors of the pluripotency network. Inhibition of GSK3 by Chiron mimics canonical Wnt signaling leading to stabilization of beta-catenin which consequently binds TCF3 and inhibits the repression of Esrrb gene maintaining sustained activation of pluripotency. MAPK pathway is activated by FGF signaling, leading to phosphorylation of MEK and activation of ERK which in turn promotes ESC differentiation that is blocked in the presence of PDO3. Figure taken from (Hackett and Surani, 2014).

An established core network of transcription factors secures the ground state of pluripotency that is rapidly transcriptionally down-regulated after initiation of differentiation and therefore allowed to monitor the loss of stem cell identity during ESC differentiation.

2.1.3 Exit from the ground state of pluripotency

Self-renewal confers to ESCs the unique ability to proliferate without differentiation. To initiate differentiation, ESCs must be forced-out of the self-renewing state, the transcriptional networks conferring stem cell identity need to be dissolved and lineages have to be governed. Compared to acquisition and maintenance of ESC pluripotency, the mechanisms driving exit from this cell state

are ill defined. Upon exposure to appropriate stimuli, *in vitro* cultured ESCs initiate differentiation programs which recapitulate the developmental progression that occurs *in vivo* (Buecker et al., 2014). Traditionally, exit from pluripotency was considered to be a unidirectional process in which a stem or a progenitor cell differentiate to give rise to a specialized cell type with a defined function.

However, this process can be reverted and fibroblast reprogramming towards an induced pluripotent stem cell state was rendered possible chemically or genetically by overexpression of pluripotency genes called the Yamanaka's factors (Zhu et al., 2010) (Takahashi and Yamanaka, 2006). Nonetheless, how exit from the pluripotent state is regulated remains poorly understood. In order to get a better understanding of this process, different approaches have attempted to identify genes whose absence would impair ESCs differentiation. Numerous unbiased genome-wide loss-of-function genetic screens have identified several genes important for the exit of pluripotency and did contribute to a better understanding of stem cell identity ((Betschinger et al., 2013); (Yang et al., 2012); (Westerman et al., 2011); (Leeb et al., 2014)). These include known components of the Fgf and Wnt signaling pathways.

2.2 Metabolic control of cell fate transition

2.2.1 Glycolysis as a predominant pathway for energy production in ESCs

During embryonic development, after morula compaction, naïve ESCs within the ICM of the pre-implantation epiblast rely on oxidative phosphorylation (OxPhos) as their primary source of energy. After implantation, ESCs from the post-implantation epiblast display a strong increase in the metabolic flux and a high glycolytic rate even when cultured outside of the hypoxic blastocyst environment. The naïve-to-primed transition is accompanied by an intrinsic reduction of OxPhos and an increased rate of glycolysis due to down-regulation of the cytochrome c oxidase (COX) genes and high expression of glucose transporters, respectively ((Sperber et al., 2015), (Zhang et al., 2016)). Whether the metabolic switch observed *in vitro* reflects the *in vivo* situation remains controversial. Most of the stem cells rely on

anaerobic glycolysis in their original compartment called the stem cell niche (e.g. hematopoietic stem cells) but contrary to adult stem cells that are quiescent, pluripotent ESCs are highly proliferative (Ito and Suda, 2014). From an energetic aspect, glycolysis offers numerous advantages including minimizing the production of reactive oxygen species (ROS) which cause oxidative damages. Also, the high glycolytic rate of ESCs allows for fast energy production, generation of anabolic intermediates as well as nucleotide production via the pentose phosphate pathway. During ESC differentiation, the mitochondria undergo maturation and develop cristae structures within their inner membrane which is accompanied by an increase in their membrane potential but not reflected in an increase in mitochondrial respiration ((Van Blerkom, 2009); (Zhou et al., 2012)). These changes accompany the metabolic shift from oxidative glycolysis toward mitochondrial oxidative phosphorylation. The hypoxia inducible factor 1 alpha (Hif1a) transcription factor described to be activated in response to hypoxia has been described to induce a switch from bivalent to exclusive glycolytic metabolism during ESC-to-EpiSC transition (Zhou et al., 2012).

This metabolic shift has often been compared to the Warburg effect described in cancer cells that orchestrates a metabolic rewiring towards high glycolytic rates instead of oxidative phosphorylation for energy production (Vander Heiden et al., 2009). This hypothesis has been discarded and it is considered that high glycolytic rates are a common feature of rapidly dividing cells (Moussaieff et al., 2015). However, the metabolic shift toward anaerobic glycolysis after the blastocyst implantation into the uterine wall remains unproven ((Barbehenn et al., 1978); (Brinster and Troike, 1979)).

Like any other cell types, ESCs sense their environment and regulate their metabolism accordingly. During reprogramming of somatic cells toward pluripotent stem cells, a metabolic switch occurs at early stages of this process. It is therefore conceivable that a metabolic checkpoint is activated before ESCs undergo differentiation since this cellular process requires a large amount of energy. In line with this hypothesis, dysregulation of metabolic pathways or perturbations of metabolite concentrations *per se* could impair ESCs differentiation, recapitulating a delay of developmental progression *in vivo*.

2.2.2 Metabolic control of pluripotency

ESCs are very proliferative and retain their self-renewal property upon cell division, maintaining the ESC pool. The predominant dependency of ESCs on glycolysis is a unique characteristic of their pluripotent state which stands in contrast to somatic cells that rely essentially on mitochondrial oxidative phosphorylation (OxPhos) for energy production.

These observations have raised a lot of attention and recent studies have directly linked metabolism to the regulation of the stem cell state (Sperber et al., 2015). For example, successful reprogramming of somatic cells towards induced pluripotent stem cells (iPSCs) is always accompanied by a metabolic shift characterized by an upregulation of glycolysis which occurs even before the acquisition of pluripotency (Folmes et al., 2011).

This strongly supports the importance of metabolic control of stem cell identity. OxPhos takes place within the cristae of mature mitochondria through the electron transport chain. One of the functions of OxPhos is to regenerate NAD+ in order to keep active the TCA cycle. This is crucial to generate sufficient levels of alpha ketoglutarate which is required for epigenetic modifier functions like demethylases (Carey et al., 2014). An intrinsic metabolic control of pluripotency also exists in cultured ESCs. LIF-induced STAT3 activation as well as Esrrb, can promote the transcription of mitochondrial genes which in turn regulate transcription of genes required for mitochondrial Oxphos reaction (Carbognin et al., 2016); in line with this finding, the transcription factor Hif1a was shown to control the activity of the gene Esrrb (Zhou et al., 2012).

Glutamine has also been linked to stem cell regulation. In addition to contributing to the alpha ketoglutarate pool, glutamine controls cellular redox reactions. A recent study has linked glutamine metabolism to the regulation of pluripotency in human ESCs by maintaining high levels of the transcription factor Oct4 in its reduced form, allowing its binding to DNA (Marsboom et al., 2016). In this study, glutamine removal from the medium decreased the pool of the endogenous antioxidant gluthatione (GSH) leading to Oct4 degradation, demonstrating its direct involvement in controlling pluripotency.

In addition, ESCs are highly dependent on threonine metabolism that ensures the regeneration of the S-adenosyl methionine (SAM) pool, which is required as a cofactor for certain post-translational modifications including the repressive histone mark H3K27me3. This illustrates a tight link between metabolic regulation of gene expression and the control of stem cell identity.

2.2.3 mTOR signaling pathway is activated during ESC differentiation

The mammalian target of rapamycin complex (mTORC) is a serine/threonine protein kinase, a master regulator of cellular growth. mTOR belongs to the family of phosphoinositide-3-kinase-related kinases. mTORC kinase promotes growth in response to a large diversity of signaling molecules such as growth factors and its activity is regulated at different levels by cellular stresses including variation in energy levels. Two mTORC complexes do exist, namely mTORC1 and mTORC2, that share the same catalytic sub-unit called mTOR; however, mTORC1 remains the more characterized. mTORC1 is regulated by two distinct mechanisms; on one hand, growth factors signal through the Tsc1/2-Rheb axis (reviewed in (Laplante and Sabatini, 2009)) whereas amino- acids are sensed and signal through the Ragulator protein complex localized at the lysosomal surface (Zoncu et al., 2011) (see section 1.3.2). In this mechanism, the lysosomal recruitment of mTORC1 followed by its activation has been shown to be very specific and accurately controlled by distinct amino- acid residues.

Stimulation of mTORC1 pathway results in activation of cell growth as well as numerous of anabolic processes like protein and lipid synthesis. Its inhibition has the opposite action, resulting in down-regulation of catabolic pathways as well as induction of autophagic process (reviewed in (Laplante and Sabatini, 2009)). Deregulations of mTOR pathway can be found in several pathological conditions. For example, mTOR hyper-activation is a common feature of numerous human cancers and neurological diseases (e.g. epilepsy, autism) whereas decreased mTOR activity is linked to diseases such as muscular atrophy. However, the reasons why changes in mTOR activity have severe consequences remain to be elucidated.

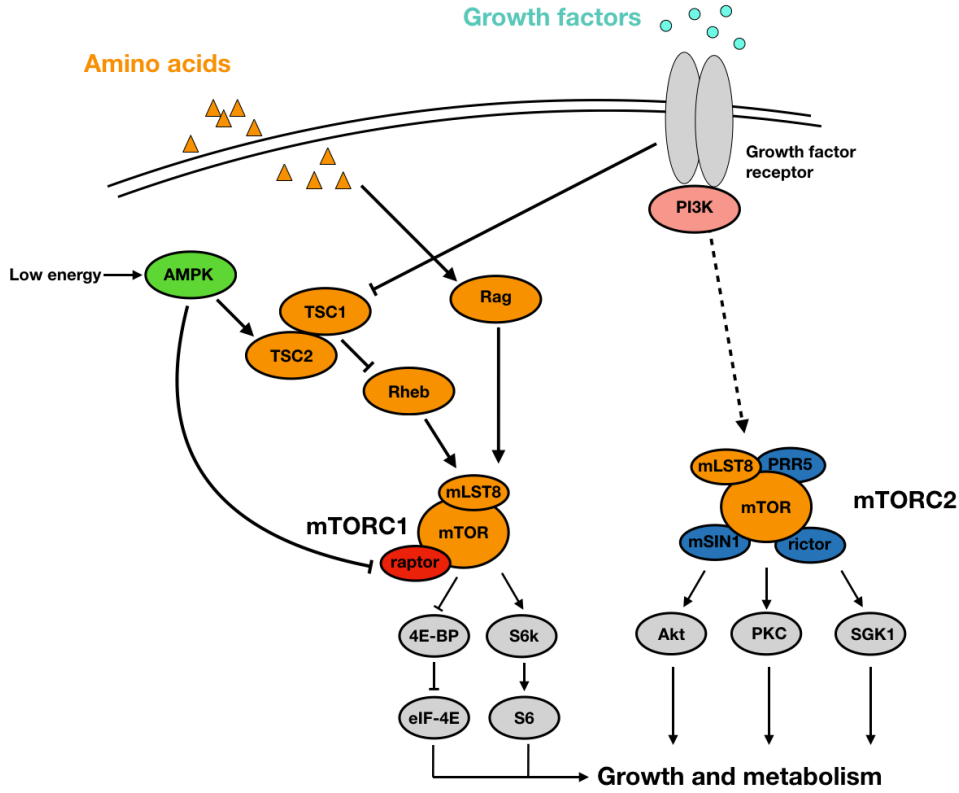


Figure2: The mammalian target of rapamycin (mTOR) signaling pathway

mTOR complex 1 (mTORC1) and mTOR complex 2 (mTORC2) are both controlled by growth factors. Growth factors inhibits the tuberous sclerosis complex (TSC1/2) leading to mTORC1 activation. TSC acts as a GTPase-activating protein (GAP) for the small GTPase RAS homologue enriched in the brain (RHEB), when activated leads to mTORC1 activation. AMP-activated protein kinase (AMPK) is activated in response to a low energy status and phosphorylates the regulatory-associated protein of mTOR (RAPTOR). mTORC1 is also regulated by amino- acids and energy status (detailed on Fig. 3). mTORC2 activation by growth factors is controlled through the classical phosphatidylinositide 3-kinase (PI3K)-AKT pathway. Collectively multiple stimuli modulate mTORC1 and mTORC2 activity in order to maintaining the balance between anabolic and catabolic reactions.

The regulation of mTOR pathway is extremely complex, far from being fully understood. A well characterized upstream mode of regulation of mTORC1 is mediated via the tuberous sclerosis complex (Tsc) composed of two tumor suppressor genes TSC1 and TSC2 (TSC1/2). The Tsc complex is a GTPase activating protein (GAP) acting towards the small GTPase Ras homologue enriched in brain (Rheb). Upon environmental or cellular stresses, Tsc2 is recruited to the lysosome and consequently inhibits Rheb's function, leading to a negative regulation of mTORC1 activity. Tsc2 lysosomal recruitment is a general mechanism in response

to stress and results in mTORC1 down-regulation, defining Tsc2 as a physiological suppressor of mTORC1 (Demetriades et al., 1AD).

In ESCs, mTORC1 activity is relatively low and gets rapidly upregulated upon differentiation (Sampath et al., 2008). In a large-scale loss of function genetic screen aiming to identify genes essential for exit of pluripotency, TSC1/2 were identified as genes required for the maintenance of self-renewal ability of murine ESCs in the presence of differentiation's stimuli. It was therefore surprising to identify an mTORC1 activator as an inhibitor of ESCs differentiation. However, the differentiation defect downstream of TSC2 LOF was reverted by rapamycin treatment, a potent mTORC1 inhibitor (Betschinger et al., 2013).

In line with mTORC1 playing an important role during embryonic development, a recent study has shown that chemical inhibition of mTOR induces diapause, a developmental arrest of the blastocyst corresponding to embryonic dormancy (Bulut-Karslioglu et al., 2016).

2.3 The lysosome is required for normal developmental progression

2.3.1 The lysosome

Lysosomes are organelles localized at the crossroad between anabolic and catabolic reactions. They originate from early endosome and go through a series of maturation processes. The lysosome has a double function in generating fundamental units for anabolic reactions as well as in removing old organelles from the cell. The acidic lysosomal lumen contains several enzymes (e.g. hydrolases) that allow the degradation of most of the substrates (Mindell, 2012). At its surface, the proton pump V-ATPase safeguards the maintenance of the low pH of the lysosomal lumen and impairment of lysosomal acidification leads to lysosomal diseases like lysosomal storage disorders (LSDs). LSDs cover a large scope of syndromes like the well-known "Pompe disease", a neuromuscular disorder causing muscle weakness or also the Wolman's disease with lipid accumulation leading to obesity (Du et al., 1998). The lysosome can adapt to changes occurring in its environment and lysosomal functions can be regulated at the transcriptional level as well at the post-

translational level ((Sardiello et al., 2009); (Settembre et al., 2012)). In an effort to understand how lysosomal turnover is regulated, a system biology approach has identified the so called coordinated lysosomal enhancement and regulation (CLEAR) transcriptional network, a nuclear response element that regulates lysosomal function (Sardiello et al., 2009). This approach has revealed that most of the identified lysosomal genes contain an E-box, the target site for transcription factors having in their structure a basic helix-loop-helix (bHLH). The transcription factor Tfe3 is one of those and has been shown to maintain a pluripotent state in ESCs (Betschinger et al., 2013). Lysosome are key players in autophagy, a conserved pathway by which old organelles and other macro molecules are recycled. Autophagy was shown to have an important role in development and loss of function of beclin1, an important regulator of autophagy leads to early embryonic lethality around the gastrulation stage (Yue et al., 2003). Additionally, a role in the regulation of pluripotency associated proteins has also been described in human ESC (Cho et al., 2014). Therefore, the lysosome is an essential organelle acting in concert with extracellular and intracellular cues to control cellular energy and differentiation.

2.3.2 The amino-acid sensing machinery

The recent finding of the amino-acid sensing machinery revealed a conserved mechanism from yeast to Human (Péli-Gulli et al., 2015); (Sancak et al., 2008) . It acts in sensing nutrient availability required to execute normal cellular processes. Although the complex regulation of this pathway is not yet fully elucidated, the amino-acid sensing machinery acts *in fine* in regulating the activity of mTORC1 kinase through its recruitment at the lysosome surface through the ragulator complex in response to amino-acids stimulation.

The ragulator is a pentameric protein complex constituted of five Late Endosomal Adaptor, MAPK and mTOR activator (Lamtor) proteins localized at the lysosomal surface. It recruits and activates the small Rag GTPases by controlling their nucleotide loading state (Sancak et al., 2010). Lamtor1 acts a scaffold for the entire complex and secures its positioning to the lysosome. Lamtor1 lysosomal localization is secured via post-translational modifications (e.g. lipidation) within

its N-terminus (Nada et al., 2014). In line with its well characterized function in amino- acid-sensing, Lamtor1 and Lamtor2 deficient embryos die at around embryonic day 7-8, demonstrating an important role of these proteins for normal development ((Nada et al., 2009); (Teis et al., 2006)). Furthermore, Lamtor3 was identified in a genetic screen aiming to identify genes essential for ESCs differentiation suggesting an important function so far unknown in ESC biology (Westerman et al., 2011).

Lamtor4 and Lamtor5 were identified only later when the ragulator was shown to act as a Guanine Exchange Factor (GEF) for the small Rag GTPases RagA/B (Bar-Peled et al., 2012). The main function of the ragulator takes place in the transduction of intra lysosomal amino-acid concentrations through the heterodimeric complex of small Rag GTPases consisting of the redundantly acting Rag GTPases RagA or B and RagC or D. In the absence of amino-acids or when concentrations decrease, the Rag GTPases heterodimer is in an inactive conformation as RagA/B-GDP and RagC/D-GTP (Shen et al., 2017). Increased lysosomal amino-acid concentrations are sensed by the ragulator and transduced onto the inactive RagA/B GTPase through the GEF activity of the ragulator, via the lysosomal V-ATPase (Bar-Peled et al., 2012). In contrast, the RagC/D GTPases are activated by the cytoplasmic tumor suppressor Folliculin (Flcn) (Tsun et al., 2013), but how Flcn is activated in the amino-acid sensing machinery remains poorly understood. In fine, the amino-acid mechanism leads to the lysosomal recruitment and activation of mTORC1 (Sancak et al., 2010).

Since its discovery, the knowledge around the ragulator continuously increases. It is now known that the Rag GTPases are regulated by the Gap activity towards Rag (Gator) proteins and Gator1 itself is recruited to the lysosome through the protein complex Kicstor composed of four proteins: **K**PTN, **I**TFG2, **C**12orf66 and **S**ZT2 (Wolfson et al., 2017).

The amino- acid sensing mechanism is regulated at different levels and early studies already suggested a link between this pathway and the regulation of the pluripotent state of ESCs (Westerman et al., 2011).

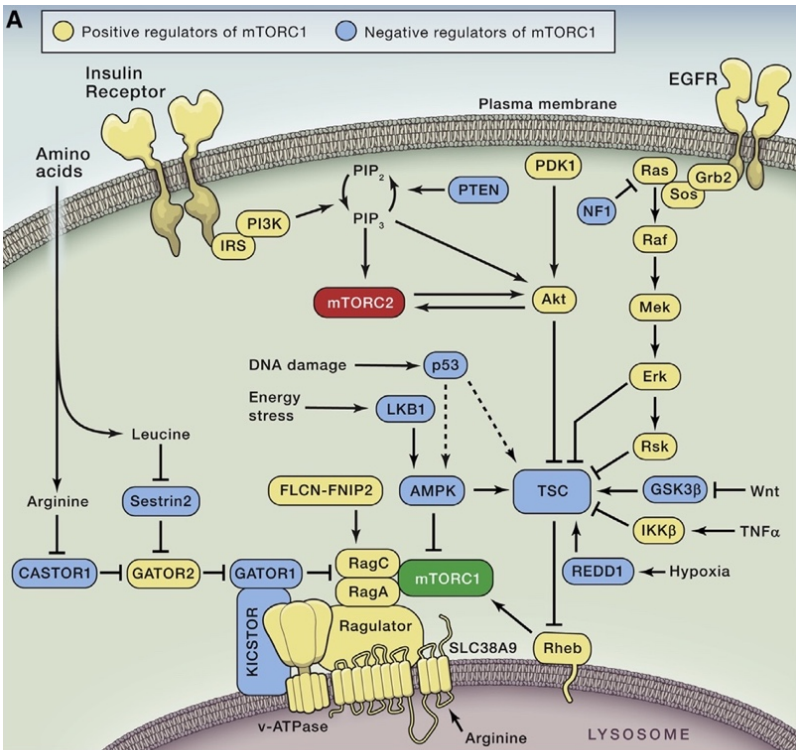


Figure3: mTORC1 complex is positively and negatively regulated

The amino-acid sensing machinery (left part) senses amino-acid within the lysosomal compartment. In the presence of amino-acids, the lysosomal lumen signals to the ragulator through the v-ATPase. Once activated, the ragulator activates RagA/B via its guanine exchange factor (GEF) function. In fine, it leads to the lysosomal recruitment of mTORC1 followed by its activation via Rheb. This pathway has different levels of regulation ensured by positive (yellow) and negative (blue) regulators. Figure taken from (Saxton and Sabatini, 2017)

2.3.3 Folliculin is required for developmental progression and metabolic functions

Developmental progression is often associated with metabolic pathways rewiring. Indeed, tremendous research work was conducted in order to identify intracellular molecular sensors that convey the nutritional availability information from the extracellular toward the intracellular environment. On the above described amino-acid sensing mechanism is mentioned Folliculin (Flcn). Flcn is a tumor suppressor gene, evolutionarily conserved and germ line heterozygous mutations cause the autosomal disorder called the Birt-Hogg-Dubé (BHD) syndrome, associated with renal cell carcinoma (Linehan et al., 2010). Additionally, inactivation of Flcn induces

transcriptional activity of the transcription factor Tfe3 by increasing its nuclear localization (Hong et al., 2010).

In murine, Flcn LOF embryos are not viable and die around E6.5 (Hasumi et al., 2009). Nevertheless, conditional knock out mice carrying a deletion in Flcn locus in muscle were generated to study the potential role of Flcn in metabolism and revealed an increased mitochondrial biogenesis mediated via the transcription factor PGC-1alpha (Hasumi et al., 2012). Flcn LOF associated tumors were reported to display increased mTOR activity (Baba et al., 2008). Interestingly, two independent *in vitro* studies using two different cell lines have shown that Flcn LOF results in down-regulation of mTORC1 activity in one case whereas mTORC1 was up-regulated in the other cell line ((Baba et al., 2006); (Hudon et al., 2010)). These paradoxical observations suggest an intricate mechanism of regulation dictated by Flcn. On one side, Flcn activates mTORC1 in an amino-acid dependent manner whereas Flcn LOF can increase mTORC1 activity. To unravel the mechanism in which Flcn is acting, Flcn LOF ESCs have been generated in a study that has identified a role for Flcn in conferring apoptotic resistance due to loss of Bim expression (Cash et al., 2011). Molecularly, Flcn forms a complex with two cytoplasmic Flcn-interacting partner partners Fnip1/2. It has been suggested that FLCN and FNIP1 are involved in a metabolic pathway involving energy sensing where Fnip1 has been shown to be targeted by AMPK through an unknown mechanism (Baba et al., 2006).

However, despite its implication in several diseases and numerous pathways, the role of Flcn in development remains unclear. Flcn has been identified in a genetic screen using small interference RNA (siRNA) aiming to identify genes required for exit from pluripotency in ESCs. As shown in other cell type, Flcn and its two interacting partners Fnip1/2, regulate the subcellular localization of the bHLH transcription factor Tfe3 and Flcn loss of function ESCs are consequentially impaired for differentiation (Betschinger et al., 2013).

The recent finding of Flcn acting as a GAP towards the small GTPase RagC *in vitro* has allowed to replace Flcn in a metabolic pathway taking place at the lysosome ((Petit et al., 2013); (Tsun et al., 2013)), but whether Flcn acts similarly in ESC differentiation is unclear. Another study has shown that, in normal culture

conditions, a portion of Flcn resides on the lysosome and that Flcn mediates an alternative mTORC1 pathway to regulates browning of adipose tissue involving Tfe3 (Wada et al., 2016). Taken together, this suggests that Flcn's main function might take place around this organelle.

How Flcn acts during early development upstream of Tfe3 and why this pathway appears to be so crucial for ESC differentiation remains an outstanding question.

2.3.4 The MiT family of transcription factor

The MiTF transcription factors family is composed of four evolutionarily conserved homologous proteins encoded by four distinct genes called MITF, TFEB, TFE3 and TFEC, all sharing structural homologies. Their structure is formed by a double helix leucine zipper motif, a transactivating domain and a DNA binding motif that recognizes E-box in the genome. In order to activate transcription, MiT/TFE transcription factors must dimerize with another member of the MiT/TFE family that included homodimer formation. Although Tfe3 and Tfeb have been the most extensively described, the two other family members remain functional. TfeC has a role in monocyte (Rehli et al., 1999) whereas Tfeb plays a role in osteoclast differentiation (Yoneshima et al., 2015). MiTF expression occurs essentially during embryonic development and has a regulatory function for melanocytes differentiation (Levy et al., 2006). Mutations in the MITF locus are associated with a rare genetic disorder called the Waardenburg syndrome characterized by permanent hearing loss and heterochromia of the irises (Baral et al., 2012); (Watanabe et al., 1998).

Although Tfeb LOF in embryo is lethal at an early developmental, Tfe3 KO mice are viable and have been reported to be healthy without any observable phenotype (Steingrimsson et al., 2002). This already suggests distinct functions for Tfeb and Tfe3 during development. Moreover, a recent study involving the tumor suppressor gene Flcn have shown that Tfe3 KO mice display bigger adipose tissues and will eventually become obese at a later stage (Wada et al., 2016).

At the cellular level, Tfe3 and Tfeb are essentially cytoplasmic under normal conditions while they localize exclusively in the nucleus in response to stress (e.g.

nutrient starvation). Supporting this finding, Tfeb overexpression increases lysosomal content and catabolic activity. Tfeb is a master transcriptional regulator of lysosomal biogenesis and autophagy; in the nucleus, Tfeb was shown to bind the Coordinated Lysosomal Expression and Regulation (CLEAR) element response sites involved in lysosomal biogenesis (Sardiello et al., 2009). Importantly, all the MiTF transcription factors can bind CLEAR element response.

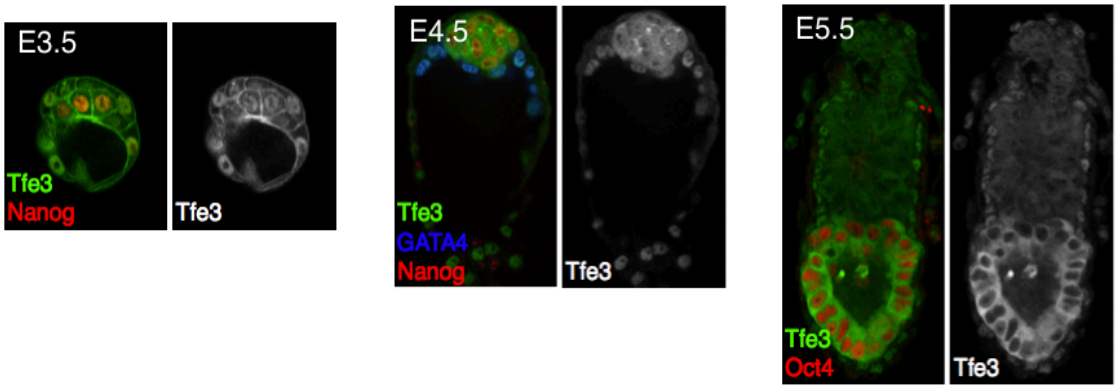


Figure 4: Tfe3 subcellular localization at early developmental stages *in vivo*

At embryonic day (E) 3.5, Tfe3 is essentially found in the nucleus. One day later, within ESCs of the ICM it becomes present in both nuclear and cytoplasm whereas in the post-implantation epiblast, it is found only in the cytoplasm.

Although the MiTF transcription factors display strong similarities, they nevertheless display distinct roles in different biological processes. Together, Tfe3 and Tfeb are at the crossroad with the mTORC1 signaling pathway. Upon lysosomal recruitment, mTORC1 phosphorylates the cytoplasmic protein 14-3-3 (Martina and Puertollano, 2013). When phosphorylated, 14-3-3 can bind Tfeb and Tfe3 in the cytoplasm. Tfe3 and Tfeb have also been described as mTORC1 targets and phosphorylation by mTORC1 leads to their cytoplasmic inactivation (Martina et al., 2014). Interestingly, hyperactivation of mTORC1 in response to Tsc2 LOF in ESCs leads to a differentiation block accompanied by nuclear localization of Tfe3 (Betschinger et al., 2013) . Another mode of regulation of Tfeb is mediated calcium release from the lysosome that activates calcineurin, a calcium and calmodulin dependent serine/threonine protein phosphatase, which binds and dephosphorylates Tfeb, thus promoting its nuclear translocation (Medina

et al., 2015). In addition, Tfeb also contains a nuclear localization signal in its structure that is covered by phosphorylation sites important for its nuclear translocation. In wildtype ESCs, Tfe3 is ubiquitously detected in both nuclear and cytoplasmic compartment, in vitro as well as in in vivo conditions (figure4). Upon differentiation, Tfe3 is excluded from the nucleus and consistently, ectopic nuclear expression of Tfe3 inhibits ESC differentiation in vitro (Betschinger et al., 2013). Tfe3 is ubiquitously expressed and is also the target of numerous chromosomal translocations that have been implicated in diseases (e.g. renal cell carcinoma) (reviewed in (Kauffman et al., 2014)). Tfe3 gene fusion mechanism remains poorly understood but all Tfe3 fusion proteins lead to its over-expression and, similar to Flcn LOF, cause the BHD syndrome.

3 Scope of this thesis

Embryonic stem cells (ESCs) have the unique potential to give rise to any of the three germ layers when exposed to differentiation stimuli. Maintenance of the ground state of pluripotency *in vitro* is well understood but the control of its loss is less so. Previously, a large-scale loss of function genetic screen has identified a set of genes required for the exit from pluripotency, amongst others identifying the tumor suppressor gene Folliculin (Flcn) (Betschinger et al., 2013). Flcn acts upstream of the bHLH transcription factor Tfe3 by controlling its subcellular distribution, thereby constituting a novel pathway in the control of ESC differentiation. However, how Flcn mechanistically regulates this pathway remains to be determined.

In the first part of the work presented here, we aimed to understand the molecular framework in which Flcn and Tfe3 act together. We used a selective genetic loss of function screen to identify genes with a Flcn-like phenotype. To further understand how Tfe3 subcellular localization is regulated, we investigated Tfe3 specific binding partners as well as potential post-translational modifications. In somatic cells, Flcn acts in a metabolic pathway sensing nutrient availability which *in fine* regulates the activity of mTORC1 kinase. We therefore asked if nutrients act instructively upstream of Flcn in ESCs. Although alteration of mTORC1 activity does neither perturb ESCs self-renewal ability nor differentiation, additional analysis revealed that hyperactive mTORC1 dominantly interferes with the Flcn-Tfe3 pathway.

To gain further insight into how this differentiation pathway contributes to human development and disease, we identified uncharacterized Tfe3 patient mutations associated with a developmental disorder.

Statement on my contributions

I wrote the following parts entirely on my own: 1 Summary, 2 Introduction, 3 Scope of this thesis, 5 Addendum, 6 General discussion and conclusions.

The manuscript presented in the chapter 4 was written by Joerg Betschinger, my thesis advisor, and Daphné Lehalle. I performed the experiments presented in figures 1e, 3a, 3b, 3c, 3d, 3g, 3h and Extended figures: 1f, 3d, 4f, 4g, 4h, 4i, 4j, 4k, 4m.

4 Results

4.1 Lysosomal signaling is required for embryonic stem cell differentiation and human development

Submitted manuscript

LYSOSOMAL SIGNALLING IS REQUIRED FOR EMBRYONIC STEM CELL DIFFERENTIATION AND HUMAN DEVELOPMENT

Florian Villegas[1,2,*], Daphné Lehalle[3,4,*], Daniela Mayer[1,2,*], Melanie Rittirsch[1], Michael B. Stadler[1,5], Marietta Zinner[1,2], Daniel Olivieri[1], Pierre Vabres[3,6], Laurence Duplomb-Jego[4], Judith M. de Bont[7], Yannis Duffourd[4], Floor Duijkers[8], Magali Avila[4], David Geneviève[9], Nada Houcinat[3,4], Thibaud Jouan[4], Paul Kuentz[4], Klaske D. Lichtenbelt[10], Christel Thauvin-Robinet[3,4], Judith St-Onge[4,11], Julien Thevenon[3,4], Koen L.I. van Gassen[10], Mieke van Haelst[8], Silvana van Koningsbruggen[8], Daniel Hess[1], Sebastien Smallwood[1], Jean-Baptiste Rivière[4,11,12], Laurence Faivre[3,4], Joerg Betschinger[1]

[1]Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland

[2]Faculty of Sciences, University of Basel, 4003 Basel, Switzerland

[3]Fédération Hospitalo-Universitaire Médecine Translationnelle et Anomalies du Développement (TRANSLAD), Centre Hospitalier Universitaire Dijon et Université de Bourgogne, 21079 Dijon, France

[4]Equipe GAD, INSERM LNC UMR 1231, Faculté de Médecine, Université de Bourgogne Franche-Comté, Dijon, France

[5]Swiss Institute of Bioinformatics, 4058 Basel, Switzerland

[6]Département de Dermatologie, CHU Dijon

[7]Department of Pediatric Neurology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, The Netherlands

[8]Department of Genetics, AMC, Amsterdam, The Netherlands

[9]Department of Clinical Genetics, Montpellier, France

[10]Department of Genetics, University Medical Center Utrecht (UMCU), Utrecht, The Netherlands

[11]Child Health and Human Development Program, Research Institute of the McGill University Health Centre, Montreal, QC H4A 3J1, Canada

[12]Department of Human Genetics, Faculty of Medicine, McGill University, Montreal, QC H3A 1B1, Canada

[*]equal contribution

Addresses for correspondence:

Dr Daphné Lehalle, Centre de Référence Anomalies du Développement et Syndromes Malformatifs, CHU Dijon, 14 rue Gaffarel, 21079 Dijon, France. Tel: 33 (3) 80 29 53 13. Fax: 33 (3) 80 29 32 66. daphne.lehalle@chu-dijon.fr

Pr Laurence Faivre, Centre de Référence Anomalies du Développement et Syndromes Malformatifs, CHU Dijon, 14 rue Gaffarel, 21079 Dijon, France. Tel: 33 (3) 80 29 53 13. Fax: 33 (3) 80 29 32 66. laurence.faivre@chu-dijon.fr

Dr Joerg Betschinger, Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland. Tel: 41 (61) 697 8556. Fax: 41 (61) 697 3976. joerg.betschinger@fmi.ch

Keywords: Embryonic stem cell, pluripotency, development, Ragulator, lysosome, Folliculin, Rag GTPases, Tfe3, differentiation, self-renewal, intellectual disability, hypopigmentation, mosaicism, genotype-first, developmental disorder

Abstract

Self-renewal and differentiation of pluripotent murine embryonic stem cells (ESCs) is balanced by extrinsic signalling pathways [1]. It is less clear if cellular metabolism instructs developmental progression. In an unbiased genome-wide CRISPR/Cas9 screen, we identify lysosomal signalling components as critical drivers of ESC differentiation. Functional analysis revealed that lysosome activity, the Ragulator protein complex and the tumor-suppressor protein Folliculin regulate the Rag GTPases C and D, which in turn bind and sequester the bHLH transcription factor Tfe3 in the cytoplasm. In contrast, ectopic nuclear Tfe3 induces transcriptional programs related to embryonic diapause and antagonizes transcriptional remodelling during differentiation. Importantly, we also identify point mutations in a Tfe3 domain required for cytoplasmic inactivation as causal for a novel human developmental disorder. This reveals a conserved and biomedical relevant role of lysosomal signalling in licensing embryonic cell fate transitions.

Results

Chemical inhibition of glycogen synthetase kinase 3 and fibroblast growth factor (FGF) signalling (2 inhibitors, 2i) confers an undifferentiated and self-renewing ESC state. Release from 2i induces exit from self-renewal and, eventually, terminal differentiation [1]. In order to identify molecular mechanisms essential for this process, we screened for genes whose deletion would block differentiation. Towards this goal we infected Cas9-expressing RGd2 ESCs with a genome-wide lentiviral guide RNA (gRNA) library [2] and continuously passaged these cells in the absence of 2i. To enrich for differentiation resistant mutant cells, we concomitantly selected for expression of the Rex1 (Zfp42) reporter in RGd2 cells, which marks self-renewing ESCs [3] (Fig. 1a). This screen was performed independently in two female ESC clones and identified 19 high-confidence target genes (Fig. 1b, Extended Data Fig. 1a, Supplementary Table 1). We validated the differentiation-resistance phenotype for 15 of those in a male RGd2 ESC line that was transiently transfected with Cas9 and individual gRNAs (Extended Data Fig. 1a). For phenotypic stratification, we also monitored the subcellular localization of the transcription factor Tfe3. Its cytoplasmic redistribution is associated with developmental progression of mouse blastocysts and murine and human pluripotent stem cells, while ectopic nuclear Tfe3 impairs ESC differentiation [3-7]. In contrast to cytoplasmic accumulation in differentiated wildtype cells (Extended Data Fig. 1h) and seven of the

differentiation-resistant mutants, Tfe3 was nuclear in the absence of Folliculin (Flcn), Lamtor1, 2, 4 and 5, and Tuberosclerosis complex (Tsc) 1 and 2 (Fig. 1c, Extended Fig. 1b). Surprisingly, two gRNAs that specifically target exon 4 of Tfe3 generated Tfe3 alleles with ectopic nuclear localization (Fig. 1c). Regulation of exit from pluripotency via subcellular Tfe3 localization by Flcn, Tsc1 and 2 has been described before, yet the regulatory framework they act in ESCs remained undefined [4]. Lamtor3 was suggested to direct the outcome of FGF signalling in ESCs [8], but based on phenotypic similarity we hypothesized that Lamtor proteins and Flcn may function together in exit from self-renewal. We first confirmed the role of Lamtor1, 2, and 3 in an independent assay. Knockdown in Oct4GIP ESCs by siRNA transfection induced nuclear enrichment of Tfe3 and impaired differentiation, similarly to absence of Flcn and the redundantly acting Folliculin-interacting partners Fnip1/2 (Extended Data Fig. 1c-e). To test for Tfe3-dependency, we co-depleted Tfe/Mitf family members expressed in ESCs (data not shown). Removal of Tfe3, but not Tfeb or Mitf, reverted differentiation defects caused by knockdown of Flcn and Lamtor1 and 2, but not Tcf7l1 (Fig. 1d), demonstrating a specific requirement for Tfe3 downstream of these genes. For further analysis, we generated isogenic Flcn and Lamtor1 (LT1) knockout (KO) ESC clones (Extended Data Fig. 1f). These maintained nuclear Tfe3 localization (Extended Fig. 1h) and failed to exit self-renewal under two different differentiation regimes; removal of 2i and the strongly inductive epiblast like cell (EpiLC) differentiation [9] (Fig. 1e). Both phenotypes were rescued by Flcn and Lamtor1 transgene expression in the respective mutants (Fig. 3b-d). Long-term culture of Flcn and Lamtor1 KO ESCs in the absence of 2i gave rise to cell populations retaining self-renewal, which is similar to induction of ectopic nuclear Tfe3 in 4-Hydroxy-Tamoxifen (Tam)-treated Tfe3-ERT2 expressing cells (Extended Data Fig. 1i) [4]. In summary, loss of Lamtor proteins, similarly to loss of Flcn, induces an ESC differentiation block that is dependent on nuclear Tfe3.

To gain insight into Flcn and Lamtor1 function, we compared the transcriptomes of two independent Flcn and Lamtor1 KO ESC lines and their respective controls (Supplementary Table 1). Importantly, transcriptional deregulation in both KO ESCs was strongly correlated (Pearson correlation coefficient R=0.70) (Extended Data Fig. 2a), suggesting that both genes act in the same pathway. To further explore the relation with Tfe3 activity, we profiled Tfe3-ERT2 expressing ESCs after 3 hours (h) and 34h of Tam induction. The short treatment resulted almost exclusively in transcriptional

upregulation, indicating that Tfe3 acts as an activator (Supplementary Table 1). Comparison with prolonged nuclear Tfe3 induction and Flcn and Lamtor1 KO ESCs identified groups of early, late and indirect target genes, and revealed adaption to continuous Tfe3 activation in established KO cell lines (Fig. 2a, Extended Data Fig. 2b-d). To test if transcriptional changes recapitulate a defined stage of embryogenesis, we compared them with published datasets of blastocyst development [10]. In particular the 34h Tfe3 activation in Tfe3-ERT2 ESCs showed a correlation to induction of diapause, an embryonic state of suspended development (Extended Data Fig. 2c). While this correlation is not strong (R=0.22), it is comparable to reported diapause mimicking ESC regimes using chemical inhibitors [11,12]. Direct comparison to these treatments revealed higher similarities to Myc-inhibition [11] than inhibition of mTOR [12] (Extended Data Fig. 2c,e), suggesting regulation of a specific subset of shared target genes.

Transcriptional similarity between diapaused embryos and ESCs with ectopic nuclear Tfe3 may reflect a common causal contribution to the inhibition of developmental progression *in vivo* and *in vitro*. To directly explore the role of Tfe3 in differentiation, we sequenced Flcn KO and nuclear Tfe3-induced Tfe3-ERT2 cells 34h after 2i withdrawal, and compared transcriptional changes to blastocyst development [10] and EpiLC differentiation [13] (Extended Data Fig. 2f, Supplementary Table 1). Principal component analysis separated ESCs and the E4.5 epiblast from 34h differentiated cells, EpiLCs and the E5.5 epiblast along principal component 2 (PC2) (Fig. 2b). In contrast, 34h differentiated Flcn KO and Tfe3-induced cells and their long term differentiated progeny, including Lamtor1 KO cells, clustered close to wildtype ESCs, suggesting transcriptome-wide impairment of peri-implantation programs. For further characterization, we performed k-means clustering of transcripts changing 34h after 2i removal in wildtype cells or being specifically deregulated during differentiation of Flcn KO and Tam-treated Tfe3-ERT2 cells (Fig. 2c, Extended Data Fig. 2b). Dynamics of genes in clusters 9 and 10 were unaffected by perturbations, while cluster 11 contained transiently induced Tfe3-targets. In contrast, clusters 12 and 13 made up more than half of the transcripts that were regulated during wildtype differentiation, but dynamics were significantly dampened upon genetic perturbations (data not shown). Gene ontology (GO) analysis revealed enrichment of developmental and lysosomal function in cluster 11 and 12, respectively. Intriguingly, transcriptional deregulation of these clusters was already detectable in mutant ESCs (Fig. 2c). However, attenuation during differentiation

is likely not a consequence of preceding ESC perturbations since Tam was added to Tfe3-ERT2 cells during differentiation only. Taken together, these findings suggest that ectopic nuclear Tfe3 mitigates large parts of the transcriptional routines underlying exit from ESC self-renewal, and that this repression is independent of the cell state. Identification of Tfe3-regulated differentiation-associated lysosomal target genes furthermore indicates coordination of developmental and metabolic programs by the Flcn-Lamtor-Tfe3 axis.

Lamtor proteins form a lysosomal membrane-associated protein complex called Ragulator [14], which recruits a Rag GTPase heterodimer consisting of the redundantly acting RagA or RagB and RagC or RagD [15]. Combinatorial depletion of RagA/B or RagC/D in ESCs caused nuclear Tfe3 accumulation and Tfe3-dependent differentiation impairment (Fig. 1d, Extended Data Fig. 1c-e), which indicates a crucial role of Rag GTPases in exit from self-renewal. Flcn acts as a GTPase activating protein (GAP) on RagC and RagD *in vitro* [16-18], and overexpression of RagC and RagD mutants preferentially binding GDP (RagC/D-GDP) [16] reverted nuclear Tfe3 localization and differentiation impairment in Flcn KO cells (Extended Data Fig. 3a,b). In contrast, GTPase-inactive mutants of RagC and RagD (RagC/D-GTP) [16] or any nucleotide-loading RagB-mutant failed to do so, suggesting that Flcn controls the nucleotide state of RagC/D but not RagA/B to drive ESC differentiation.

In ESCs, RagC, Lamtor1 and a subfraction of endogeneously tagged Flcn colocalize with the lysosomal protein LAMP1 (Fig. 3a). Lysosomal RagC localization requires Lamtor1, 2 and 3, but not Flcn (Extended Data Fig. 3c). To test the requirement for lysosomal localization we expressed RagB and RagC fused to the N-terminal lysosomal targeting sequence of Lamtor1 (LTS) [19] (Extended Data Fig. 3d). LTS-RagC but not LTS-RagB reverted nuclear Tfe3 localization and differentiation resistance in Lamtor1 KO ESCs (Fig. 3b,c). This indicates that lysosomal RagC/D but not RagA/B is sufficient in the absence of a functional Ragulator complex. Consistent with the notion that Flcn acts upstream of the RagC/D nucleotide state and independently of lysosomal recruitment, LTS-RagC failed to rescue Flcn KO phenotypes, while LTS-RagC-GDP did so (Fig. 3b,d). Conversely, LTS-RagC-GTP dominantly induced nuclear Tfe3 localization and Tfe3-dependent differentiation impairment in wildtype cells (Extended Data Fig. 3e-g). Taken together these results suggest that the principal role of Ragulator in exit from ESC self-renewal is the lysosomal recruitment of RagC/D, which subsequently serve as a

substrate for Flcn GAP activity to drive cytoplasmic sequestration of Tfe3 and differentiation.

To directly address the role of the lysosome in this process we applied lysosome perturbing drugs. Addition of the v-ATPase inhibitors BafilomycinA (BafA) and ConcanamycinA (ConA), the lysosomotropic compound Chloroquine, or Vacuolin-1, which impairs delivery and maturation of lysosomal enzymes [20], indeed impaired exit from self-renewal (Extended Data Fig. 3h,i). Since BafA, ConA and Chloroquine also triggered dose-dependent cell lethality (data not shown) we focussed our further efforts on Vacuolin-1. Treatment with Vacuolin-1 induced nuclear Tfe3 translocation (Extended Data Fig. 3j), while expression of LTS-RagC-GDP or knockout of Tfe3 (Extended Data Fig. 1f,g) rescued Vacuolin-1-elicited resistance to differentiation (Fig. 3e,f). Despite altered morphology in Vacuolin-1-treated cells, LAMP1-marked lysosomes still colocalized with RagC (Extended Data Fig. 3j), indicating an intact, but likely inactive Ragulator-Rag GTPase complex. This suggests that lysosomes not only provide a membranous interaction surface for assembling Rag GTPase signalling, but that lysosome function *per se* is required upstream of the RagC/D nucleotide state and Tfe3 to drive exit from pluripotency.

To explore how signalling at the lysosomal membrane regulates Tfe3, we determined the Tfe3 interactome in ESCs using label-free quantitative mass spectrometry (Fig. 3g,h, Supplementary Table 2). Tfe3 in RagC/D-GDP overexpressing ESCs specifically interacted with the Ragulator-Rag protein complex, with members of the lysosomal v-ATPase, and with 14-3-3 protein-family members (Fig. 3h). Pulldown of Tfeb by the RagA/B-GTP - RagC/D-GDP heterodimer in human cell lines has been reported before [21], but Tfe3 failed to interact with any RagB nucleotide loading mutant in ESCs (data not shown), demonstrating RagC/D specificity. Association of Tfe3 with Ragulator, Rag GTPases and v-ATPase proteins was undetectable in wildtype ESCs (Fig. 3g), suggesting transient and nucleotide-dependent binding that induces stable cytoplasmic sequestration, likely via further posttranslational modifications. To test the necessity of this interaction in the context of the endogenous Rag nucleotide cycle, we expressed in Tfe3 KO ESCs inducible Tfe3 alleles with mutations in residues homologous to the ones required for interaction of Tfeb with the active RagA/B-C/D heterodimer [21]. In contrast to wildtype Tfe3, Tfe3(S111A, R112A) and Tfe3(Q118A, L119A) localized to the nucleus and impaired exit from self-renewal (Fig. 3i,j), indicating that association of Tfe3

with endogenous RagC/D-GDP is required for nuclear exclusion and ESC differentiation. 14-3-3 proteins recognize phosphorylated Ser and Thr residues, and 14-3-3 binding and Ser phosphorylation have previously been implicated in controlling the subcellular localization of Tfe/Mitf family members [21,22]. We therefore generated Tfe3 alleles with Ala mutations in respective Ser residues. This revealed that Tfe3(S245A) [23] and Tfe3(S320A) [24] were partially, and Tfe3(S245,S320A) exclusively nuclear when compared to wildtype Tfe3 and Tfe3(10xA) [25] (Fig. 3i). Single and double Tfe3 mutants also impaired exit from self-renewal (Fig. 3j), suggesting that S245 and S320 phosphorylation, downstream of RagC/D-GDP, is necessary for 14-3-3 protein-mediated cytoplasmic Tfe3 retention. This requirement for, at least, two posttranslational modifications is consistent with multiple additional levels of regulation.

We therefore tested mechanisms described to modulate Tfe/Mitf transcription factors, including protein kinases C (PKCs) [26], Mcoln1-dependent regulation of the phosphatase Calcineurin [27], and the mechanistic target of Rapamycin (mTOR) complex 1 (mTORC1) [24]. Presence of the pan-PKC inhibitor Gö6983 caused nuclear Tfe3 enrichment and impaired differentiation (Extended Data Fig. 4a,b) [28], but these defects were insensitive to overexpression of LTS-RagC-GDP, demonstrating additional Rag GTPase-independent mechanisms such as NFκB regulation [28]. Knockdown of Mcoln1, the calcineurin catalytic subunit isoform beta (Ppp3cb), mTOR and the mTORC1 subunit Raptor, or treatment with the mTOR inhibitors AZD8055 and Rapamycin (Rapa) did not cause retention of self-renewal or revert resistance to differentiation in the absence of Lamtor1 (Extended Data Fig. 4c-e). Developmental progression of ESCs is therefore independent of Calcineurin or mTORC1 activation. The latter is surprising since mTORC1, via phosphorylation of S320, promotes cytoplasmic Tfe3 translocation upon release from amino acid starvation in somatic cell lines, and since Flcn, Ragulator and Rag GTPases are required for recruitment and activation of mTORC1 in this context [22]. This suggests decoupling of lysosomal signalling and mTORC1 activity in ESCs. Consistently, phosphorylation of the mTORC1-targets S6 kinase (S6K) and 4EBP1 was not significantly altered upon knockout of Flcn and Lamtor1 (Extended Data Fig. 4f) or overexpression of LTS-RagC mutants (data not shown), and transcriptional deregulation in Flcn and Lamtor1 KO ESCs did not correlate with mTOR-inhibition [12] (Extended Data Fig. 2c). Furthermore, removal of essential amino acids induced nuclear Tfe3 localization, but resistance to differentiation in these conditions was not reverted by LTS-

RagC-GDP expression or in Tfe3 KO ESCs (Extended Data Fig. 4g,h). Taken together these observations suggest a RagC/D-dependent, but mTORC1 activation-independent mechanism driving ESC differentiation. We also did not identify any mTORC1-activator, but, surprisingly, only the mTORC1-inhibitors Tsc1 and Tsc2 in our primary screen (Fig. 1b). Loss of Tsc2 by siRNA transfection or in KO ESCs impaired differentiation upstream of hyperactive mTORC1 (Extended Data Fig. 4d,e,i,j) [4], nuclear Tfe3 and RagC/D-nucleotide loading (Fig. 1d, Extended Data Fig. 4k-m). Tsc2-dependent mTORC1 inhibition and Flcn-Ragulator-function therefore converge on RagC/D-GDP to promote exit from self-renewal. This is an unexpected observation and we therefore speculated that hyperactive mTORC1 and ectopic nuclear Tfe3 may exhibit similar phenotypes in other developmental systems.

Tsc1/2 and Flcn loss of function mutations, hyperactive mTOR alleles and deletion of Tfe3 exons1-3 by translocation to distinct partner genes are frequently detected in renal cell carcinoma [29,30]. The molecular consequences of these Tfe3 alterations are poorly understood [29], but our findings predict them to behave as nuclear gain of function alleles. We therefore conditionally expressed corresponding Tfe3 deletions in Tfe3 KO ESCs (Fig. 4a). Inclusion of exon3 (exon3-end) recapitulated cytoplasmic localization of wildtype Tfe3, while removal (exon4-end) induced nuclear translocation and impaired differentiation (Fig. 4b,c), suggesting that it is the absence of exon3 and not the presence of fusion partners that underlies oncogenic activity of chromosomal Tfe3 rearrangements. Surprisingly, transcript sequencing of gain of function Tfe3 alleles recovered in our screen (Fig. 1b,c, Extended Data Fig. 1a) revealed small in-frame-deletions frequently involving T180-P185 and skipping of exon4 rather than exon3 (data not shown). Nuclear localization of Tfe3Δexon4 protein and resistance to differentiation was validated in a Tfe3 exon4 skipping mutant ESC clone generated by an independent gRNA (Extended Data Fig. 1f, 5a,b) and in Tfe3 KO cells harbouring inducible Tfe3Δexon4 transgenes (Fig. 4a-c). Alterations in exon4 have, however, not been associated with disease as of now.

Mosaic mTOR gain of function alleles also cause pigmentary mosaicism [31], but mutations in Flcn or Tfe3 have not been associated with this phenotype before. Pigmentary mosaicism and hypomelanosis of Ito (HI, MIM 300337) are unspecific terms encompassing a heterogeneous group of disorders characterized by hypopigmented

whorls and streaks along Blaschko's lines and variable extracutaneous features affecting the musculoskeletal and nervous systems [32]. In a program designed to identify underlying genetic causalities in a cohort of patients recruited for HI using whole-exome sequencing we identified four *de novo* germline (female patient 1: Q119P, female patient 2: Q201P, female patient 3: T187M, female patient 5: P186L) and one *de novo* mosaic (male patient 4: T187R) mutations in *TFE3* (Fig. 4d, Supplementary Clinical Data, Supplementary Table 3). Patient phenotypes included hypopigmentation on Blaschko's lines, severe intellectual disability (ID), coarse facial features, frontonasal dysplasia, obesity, epilepsy and growth retardation, suggesting a pleiotropic developmental disorder. All five *TFE3* variants affect conserved amino acids (Fig. S5c), and the missense variants are *in silico* predicted as pathogenic (Supplementary Table 3). The Q119P mutation resides in exon3 and affects an amino acid implicated in Rag interaction (Fig. 3i, j)[21], while the remaining four mutations (P186L, T187M, T187R, and Q201P) induce amino acid changes in exon4. Nuclear localization and resistance to differentiation were validated in Tfe3 KO ESCs expressing murine Tfe3 alleles (Q118P and P185L) (Fig. 4e,f). Secondary structure analysis of Tfe3 using I-Tasser [33] predicts residues 110-215, encoded by exon3 and exon4, to form a domain of two stable α-helices (Extended Data Fig. 5c). The Q->P mutations at positions 119 and 201 are likely disrupting helix formation. The two helices are connected by a short loop and an induced α-helix, and contain the highly-conserved residues P186 and T187 that may be crucial for Rag C/D interaction. These observations suggest that exon3 and exon4 form a Rag binding fold whose structural integrity is indispensable for cytoplasmic Tfe3 inactivation and, consequentially, ESC differentiation, human development and kidney tumor suppression.

Here we show that lysosomal Rag GTPases drive ESC differentiation by inactivating Tfe3 through cytoplasmic redistribution. Interference with this regulation through mutations that produce inactivation-resistant, ectopic nuclear, Tfe3 proteins are associated with pleiotropic developmental defects in humans. These findings reveal an important role of lysosomal signalling in controlling embryonic cell fate transitions.

Lysosomes are crucial for cellular homeostasis, and their function is under transcriptional feedback control by Tfe/Mitf transcription factors [23,34]. Absence of amino acids upon cellular starvation locks the heterodimeric RagA/B-C/D-complex in a guanine nucleotide

loading conformation unable to recruit and activate mTORC1 at the lysosomal membrane. This, in turn, leads to nuclear translocation of Tfe/Mitf proteins and transcription of metabolic target genes [22] (Extended Data Fig. 5d). Exit from ESC self-renewal, in contrast, depends on GDP-loading of RagC/D by Flcn, but not on the RagA/B nucleotide status, mTOR activity nor presence of Tfeb or Mitf, making it mechanistically distinct from amino acid sensing. Consistent with such an alternative pathway, hyperactivation of mTORC1 in Tsc2 KO cells phenocopies Flcn and Ragulator mutants upstream of RagC/D-GDP and Tfe3. Loss of Tsc2 and Flcn also predispose patients for renal carcinogenesis [35] and induce nuclear Tfe protein translocation in cancer cell lines [25,36]. Furthermore, patients with Tfe3 mutations present classical features of hyperactive mTORC1 pathway-related disorders [37] (epilepsy, frequently associated with cortical malformations) (Supplementary Table 3). Ragulator is the major guanine nucleotide exchange factor (GEF) and, consequently, activator of RagA/B [14]. However, ectopic lysosomal RagC in the absence of a functional Ragulator complex is sufficient for cytoplasmic Tfe3 localization and differentiation of ESCs. Furthermore, RagA/B is dominantly contributing to mTORC1 binding and activation compared to RagC/D [38,39], but Tfe3 interacts specifically with RagC/D-GDP. We therefore speculate that hyperactive mTORC1 competes with Tfe3 for binding to the Rag heterodimer. Mechanistically, we have identified a cytoplasmic inactivation domain within Tfe3 that contains all five identified patient mutations and is frequently deleted in renal cell carcinoma associated Tfe3 translocations. We predict this domain to form a RagC/D-GDP binding fold and note that it overlaps with a conserved amino acid stretch specific to Tfe/Mitf transcription factors (Fig. 4a). Tfe3 is the predominant family member expressed in ESCs, but overexpression of a constitutive nuclear Tfeb allele is sufficient to impair exit from self-renewal (data not shown), thus showing functional redundancy between Tfe/Mitf family members and directly explaining the Tfe3 specificity observed in ESCs.

Our findings show that lysosome activity controls ESC differentiation through RagC/D and Tfe3. Genes deregulated by Myc-inhibition, which mimics biosynthetic dormancy in diapause [11], also overlap with a subset of lysosomal Tfe3-targets in ESCs (Extended Data Fig. 2e). We therefore speculate that lysosomal signalling acts as a metabolic differentiation checkpoint that anticipates energy requirements faced during developmental progression. Alternatively, and not mutually exclusive, lysosome activity

may guide the remodelling of differentiation-specific gene regulatory networks. The latter would be consistent with the dynamic regulation of Tfe3 subcellular localization and lysosomal content during blastocyst development [3,4,6,40]. Noteworthy, Tfe3-regulation discriminates lysosome activity from other metabolic inputs such as the tricarboxylic acid cycle that permissively provides substrates for genome-wide chromatin modifications [41]. Importantly, nuclear Tfe3 antagonizes the downregulation of lysosome regulators during differentiation, which is congruent with general metabolic feedback regulation by Tfe/Mitf transcription factors [22] and is, likely, reflected in lysosomal storage disorder related phenotypes in most individuals harbouring Tfe3 mutations (coarse facial features, umbilical hernia, postnatal growth retardation, obesity, hepatomegaly and hypoglycemia). Since Tfe3 acts as a transcriptional activator, we surmise that the concomitant suppression of developmental programs is consequential to direct metabolic targets. This suggests that Tfe3-dependent metabolic remodelling is important for steady state homeostasis and instructive to exit from pluripotency.

Subcellular Tfe3 distribution is associated with developmental progression of human pluripotent stem cells [5,7]. It, thus, might seem surprising that ectopic nuclear Tfe3 is not lethal in patients. This, however, is probably a consequence of mosaicism, since Tfe3 is an X chromosome linked gene: Genetic mosaicism by postzygotic *TFE3* mutation in the only male patient, likely functional mosaicism by random X chromosome inactivation in three female patients (demonstrated for one patient) and skewed X inactivation in one patient without hypopigmentation (patient 5). These mosaicisms are thus in line with the necessity for Tfe3 inactivation in early human development. Clinically, cutaneous pigmentary mosaicism, coarse features and unspecific syndromic ID are recognizable, and broadening of the phenotypic spectrum by the identification of more patients appears likely. In addition to regulating embryogenesis, mosaic *TFE3* mutations may thus be considered causative to ID.

In summary, our findings suggest that cellular metabolism instructs developmental progression in ESCs by a lysosomal Flcn-Ragulator-RagC/D-Tfe3 signalling axis. Despite shared molecular players this mechanism is distinct to amino acid dependent mTORC1-regulation, indicating repurposing of an ancient metabolic sensor conserved in yeast [18] for controlling transcriptional remodelling during cell fate transition. The identification of Tfe3 alleles insensitive to this lysosome-mediated inactivation in a novel

human developmental disorder demonstrates functional conservation and importance to disease, but also identifies opportunities for targeted intervention.

Acknowledgments

We would like to thank the patients and families who participated in this study. We also thank the University of Burgundy Centre de Calcul (CCuB) for technical support and management of the informatics platform. We are grateful to S. Dessus-Babus, K. Jacobeit and T. Roloff (FMI) for processing sequencing samples; L. Gelman (FMI) for imaging assistance; H. Gut (FMI) for performing structural modelling; M. Flemr (FMI) for advice on genome editing; J.Chao (FMI), A.Smith (University of Cambridge), F.Stewart (University of Dresden), K.Yusa (Wellcome Trust Sanger Institute, Hinxton), and M.Leeb (University of Vienna) for providing reagents; and S. Gasser, P. Liberali, A. Peters, D. Schuebeler, N. Thomae (FMI) and D. Teis (University of Innsbruck) for comments on the manuscript. This work was funded by an EMBO long term fellowship (to D.O.), the Programme Hospitalier de Recherche Clinique (PHRC) National (to P.V.), the Regional Council of Burgundy and the Centre Hospitalo-Universitaire de Dijon (to L.F.) (PARI 2015), and the Novartis Research Foundation (to J.B.).

Author contribution

F.V., M.R. M.A., L.D.J., J.S-O. and T.J. performed experiments. D.M., M.S., Y.D., P.K., J.T., K.L.I.VG. and S.VK. performed bioinformatic analysis. M.Z. and D.O. executed the CRIPSR/Cas9 screen. D.H. performed mass spectrometry. P.V., J-M.D.B., F.D., D.G., N.H., K.D.L., C.T.R., M.VH. recruited and evaluated the patients. S.S. established quantitative gRNA sequencing. L.F., J-B.R. and J.B. supervised the work. D.L. and J.B. wrote the paper.

Conflict of interest statement

The authors declare no conflict of interest.

References

1. Martello, G. & Smith, A. The nature of embryonic stem cells. *Annu. Rev. Cell Dev. Biol.* **30,** 647–675 (2014).
2. Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M. D. C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat. Biotechnol.* **32,** 267–273 (2014).
3. Kalkan, T. *et al.* Tracking the embryonic stem cell transition from ground state pluripotency. *Development* **144,** 1221–1234 (2017).
4. Betschinger, J. *et al.* Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell* **153,** 335–347 (2013).
5. Takashima, Y. *et al.* Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158,** 1254–1269 (2014).
6. Geula, S. *et al.* Stem cells. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science* **347,** 1002–1006 (2015).
7. Gafni, O. *et al.* Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504,** 282–286 (2013).
8. Westerman, B. A. *et al.* A genome-wide RNAi screen in mouse embryonic stem cells identifies Mp1 as a key mediator of differentiation. *J. Exp. Med.* **208,** 2675–2689 (2011).
9. Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. & Saitou, M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146,** 519–532 (2011).
10. Boroviak, T. *et al.* Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. *Dev. Cell* **35,** 366–382 (2015).
11. Scognamiglio, R. *et al.* Myc Depletion Induces a Pluripotent Dormant State Mimicking Diapause. *Cell* **164,** 668–680 (2016).
12. Bulut-Karslioglu, A. *et al.* Inhibition of mTOR induces a paused pluripotent state. *Nature* (2016). doi:10.1038/nature20578
13. Buecker, C. *et al.* Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* **14,** 838–853 (2014).
14. Bar-Peled, L., Schweitzer, L. D., Zoncu, R. & Sabatini, D. M. Ragulator is a GEF for the rag GTPases that signal amino acid levels to mTORC1. *Cell* **150,** 1196–1208 (2012).
15. Bar-Peled, L. & Sabatini, D. M. Regulation of mTORC1 by amino acids. *Trends Cell Biol.* **24,** 400–406 (2014).
16. Tsun, Z.-Y. *et al.* The folliculin tumor suppressor is a GAP for the RagC/D GTPases that signal amino acid levels to mTORC1. *Mol. Cell* **52,** 495–505 (2013).
17. Tsun, Z.-Y. *et al.* The Folliculin Tumor Suppressor Is a GAP for the RagC/D GTPases That Signal Amino Acid Levels to mTORC1. *Mol. Cell* (2013). doi:10.1016/j.molcel.2013.09.016
18. Péli-Gulli, M.-P., Sardu, A., Panchaud, N., Raucci, S. & De Virgilio, C. Amino Acids Stimulate TORC1 through Lst4-Lst7, a GTPase-Activating Protein Complex for the Rag Family GTPase Gtr2. *Cell Rep* **13,** 1–7 (2015).
19. Nada, S. *et al.* The novel lipid raft adaptor p18 controls endosome dynamics by anchoring the MEK-ERK pathway to late endosomes. *EMBO J.* **28,** 477–489 (2009).
20. Sano, O. *et al.* Vacuolin-1 inhibits autophagy by impairing lysosomal maturation via PIKfyve inhibition. *FEBS Lett.* **590,** 1576–1585 (2016).
21. Martina, J. A. & Puertollano, R. Rag GTPases mediate amino acid-dependent recruitment of TFEB and MITF to lysosomes. *J. Cell Biol.* **200,** 475–491 (2013).

22. Raben, N. & Puertollano, R. TFEB and TFE3: Linking Lysosomes to Cellular Adaptation to Stress. *Annu. Rev. Cell Dev. Biol.* **32,** 255–278 (2016).
23. Settembre, C. *et al.* A lysosome-to-nucleus signalling mechanism senses and regulates the lysosome via mTOR and TFEB. *EMBO J.* **31,** 1095–1108 (2012).
24. Roczniak-Ferguson, A. *et al.* The transcription factor TFEB links mTORC1 signaling to transcriptional control of lysosome homeostasis. *Sci Signal* **5,** ra42–ra42 (2012).
25. Peña-Llopis, S. *et al.* Regulation of TFEB and V-ATPases by mTORC1. *EMBO J.* **30,** 3242–3258 (2011).
26. Li, Y. *et al.* Protein kinase C controls lysosome biogenesis independently of mTORC1. *Nat. Cell Biol.* **18,** 1065–1077 (2016).
27. Medina, D. L. *et al.* Lysosomal calcium signalling regulates autophagy through calcineurin and TFEB. *Nat. Cell Biol.* **17,** 288–299 (2015).
28. Dutta, D. *et al.* Self-renewal versus lineage commitment of embryonic stem cells: protein kinase C signaling shifts the balance. *Stem Cells* **29,** 618–628 (2011).
29. Kauffman, E. C. *et al.* Molecular genetics and cellular features of TFE3 and TFEB fusion kidney cancers. *Nat Rev Urol* **11,** 465–475 (2014).
30. Grabiner, B. C. *et al.* A diverse array of cancer-associated MTOR mutations are hyperactivating and can predict rapamycin sensitivity. *Cancer Discov* **4,** 554–563 (2014).
31. Mirzaa, G. M. *et al.* Association of MTOR Mutations With Developmental Brain Disorders, Including Megalencephaly, Focal Cortical Dysplasia, and Pigmentary Mosaicism. *JAMA Neurol* **73,** 836–845 (2016).
32. Sybert, V. P. Hypomelanosis of Ito: a description, not a diagnosis. *J. Invest. Dermatol.* **103,** 141S–143S (1994).
33. Yang, J. *et al.* The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* **12,** 7–8 (2015).
34. Martina, J. A. *et al.* The nutrient-responsive transcription factor TFE3 promotes autophagy, lysosomal biogenesis, and clearance of cellular debris. *Sci Signal* **7,** ra9–ra9 (2014).
35. Linehan, W. M., Srinivasan, R. & Schmidt, L. S. The genetic basis of kidney cancer: a metabolic disease. *Nat Rev Urol* **7,** 277–285 (2010).
36. Petit, C. S., Roczniak-Ferguson, A. & Ferguson, S. M. Recruitment of folliculin to lysosomes supports the amino acid-dependent activation of Rag GTPases. *J. Cell Biol.* **202,** 1107–1122 (2013).
37. Parrini, E., Conti, V., Dobyns, W. B. & Guerrini, R. Genetic Basis of Brain Malformations. *Mol Syndromol* **7,** 220–233 (2016).
38. Sancak, Y. *et al.* The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science* **320,** 1496–1501 (2008).
39. Kim, E., Goraksha-Hicks, P., Li, L., Neufeld, T. P. & Guan, K.-L. Regulation of TORC1 by Rag GTPases in nutrient response. *Nat. Cell Biol.* **10,** 935–945 (2008).
40. Tsukamoto, S. *et al.* Functional analysis of lysosomes during mouse preimplantation embryo development. *J. Reprod. Dev.* **59,** 33–39 (2013).
41. Carey, B. W., Finley, L. W. S., Cross, J. R., Allis, C. D. & Thompson, C. B. Intracellular α-ketoglutarate maintains the pluripotency of embryonic stem cells. *Nature* **518,** 413–416 (2015).

Villegas et al., Figure 1: Identification of differentiation drivers in a CRISPR/Cas9 screen.

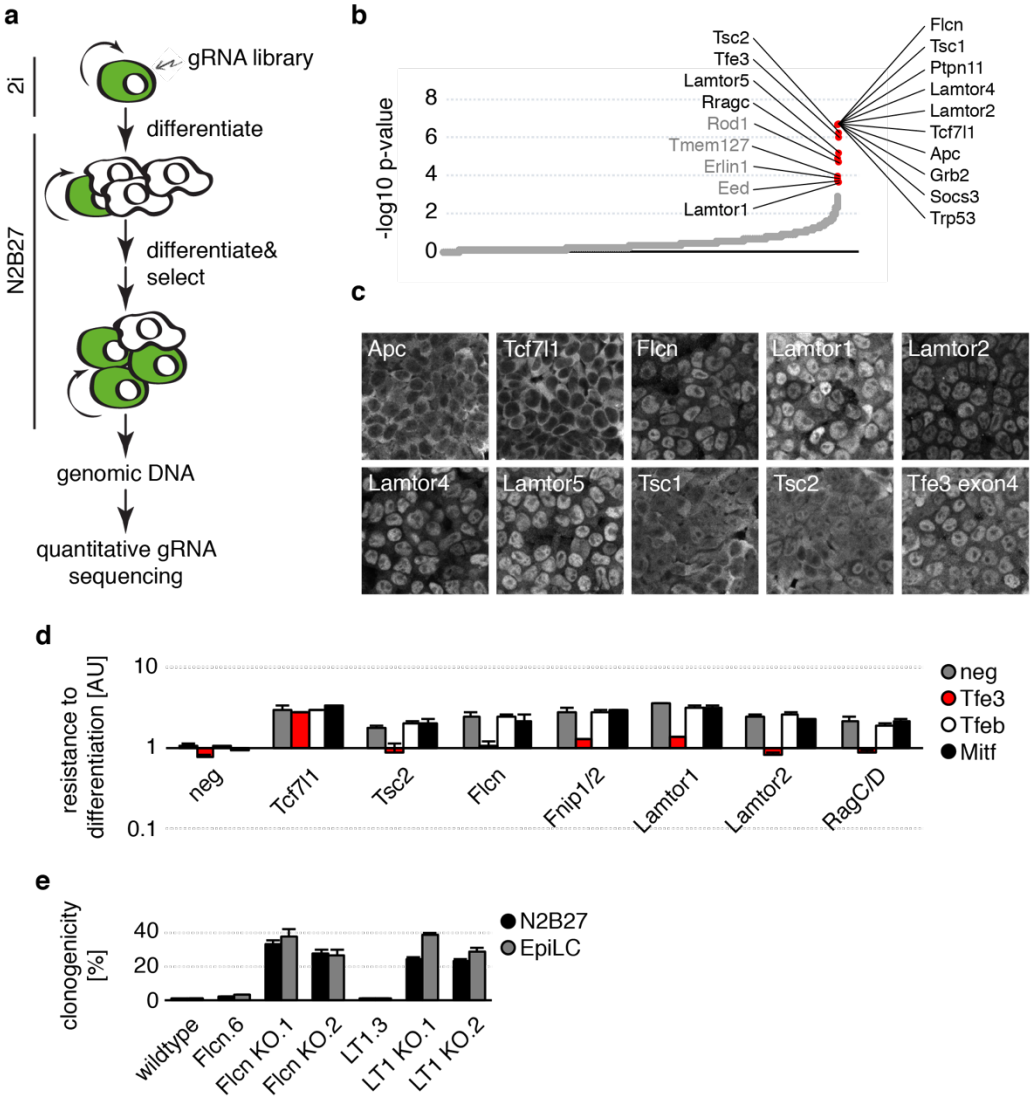


Figure 1: Identification of differentiation drivers in a CRISPR/Cas9 screen.

(a) Schematic outline of the screening procedure. Green indicates Rex1 expression.

(b) Screen results ordered by statistical significance. Genes labeled with red were retested, of which genes annotated in black were validated.

(c) Tfe3 localization in RGd2 cells transiently transfected with Cas9 and gRNAs targeting indicated genes after 3-5 passages upon 2i withdrawal. Note that control cells are expectedly lost during the procedure due to terminal differentiation.

(d) Resistance to differentiation of Oct4GIP ESCs transfected with indicated siRNA combinations after 3 days (d) of 2i withdrawal. Results were normalized to negative (neg) siRNA cells. Average and standard deviation (SD) of 2 technical replicates.

(e) Retention of self-renewal in 2i after 2d of differentiation in indicated conditions. Average and SD of 3 biological replicates.

Villegas et al., Figure 2: Transcriptional programs controlled by Flcn-Lamtor1-Tfe3.

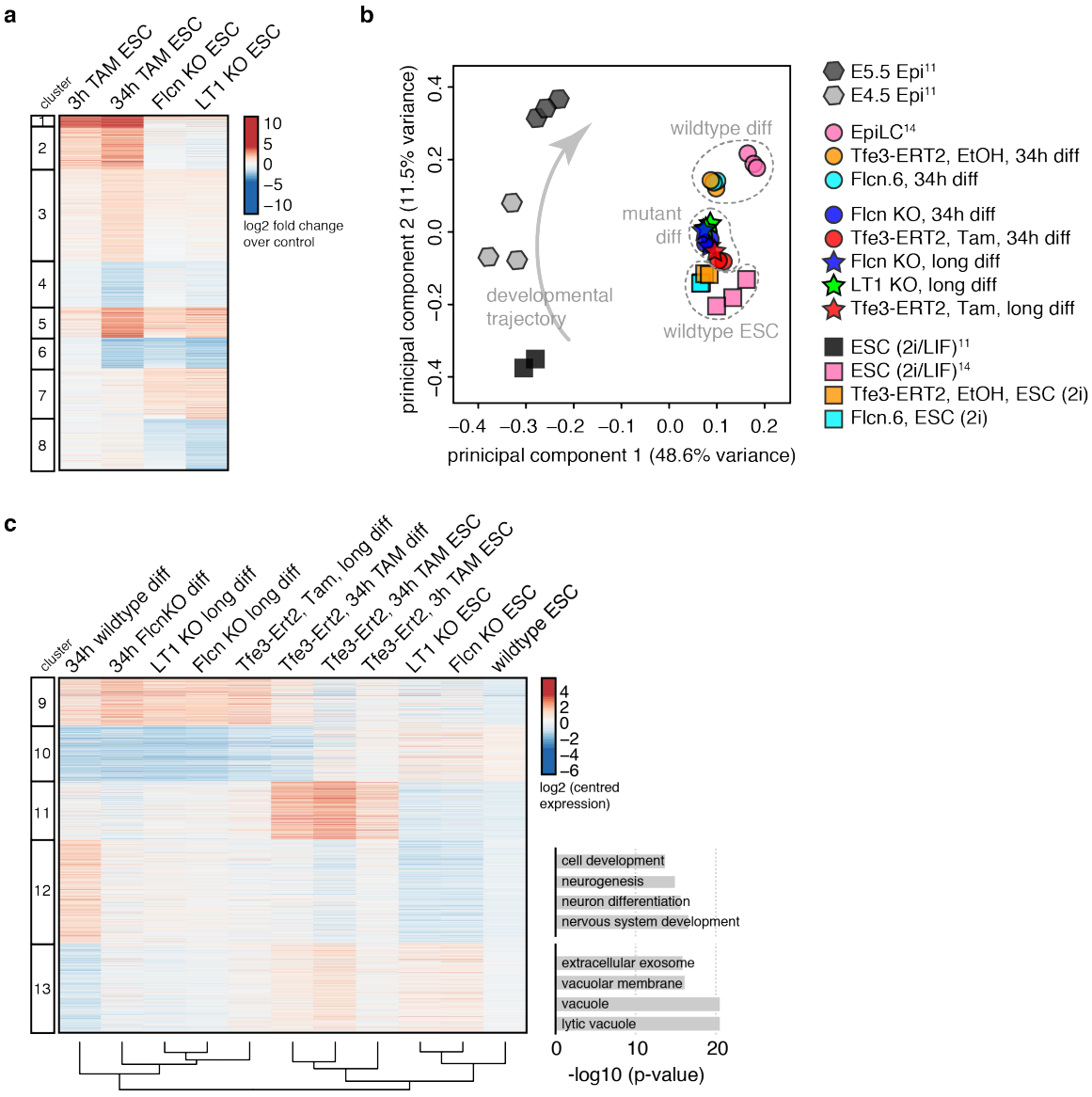


Figure 2: Transcriptional programs controlled by Flcn-Lamtor1-Tfe3.

(a) k-means clustering of log2 fold changes of transcripts significantly deregulated in ESCs of indicated genotypes.

(b) PCA analysis of indicated samples. Note that PC1 distinguishes experimental variations between studies.

(c) k-means clustering of centred expression in ESCs and during differentiation of indicated genotypes, and top five GO term enrichments of gene clusters 12 and 13. Samples (columns) were ordered by hierarchical clustering (dendrogram shown below

the heatmap). Transcripts dynamically regulated during wildtype ESC differentiation and specifically deregulated during differentiation of Tfe3-ERT2 Tam-treated and Flcn KO cells were considered.

Villegas et al., Figure 3: Mechanistic dissection of lysosomal signalling in ESCs.

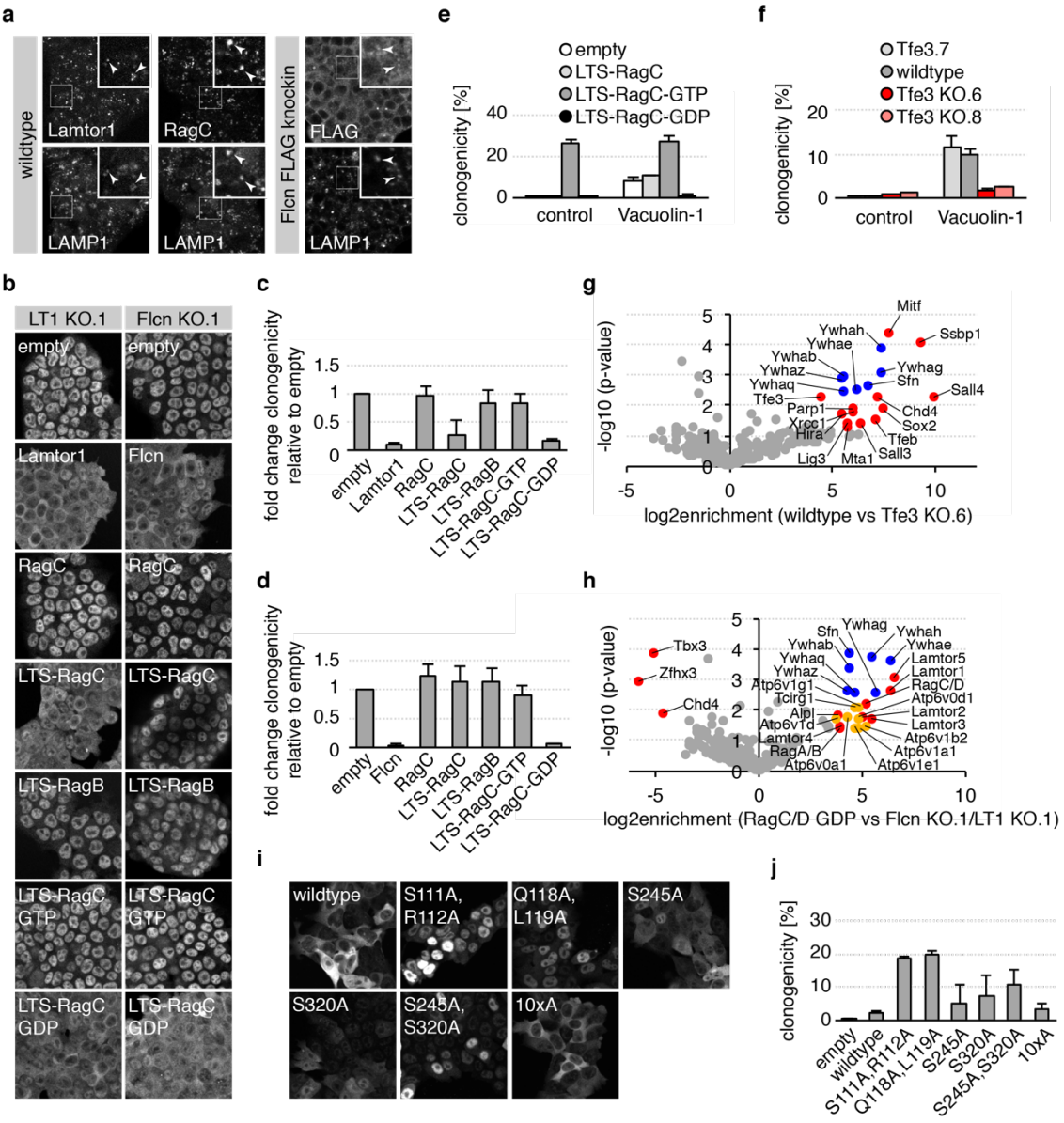


Figure 3: Mechanistic dissection of lysosomal signalling in ESCs.

(a) Colocalization of Lamtor1, RagC and tagged Flcn with LAMP1 (arrowheads) in ESCs.

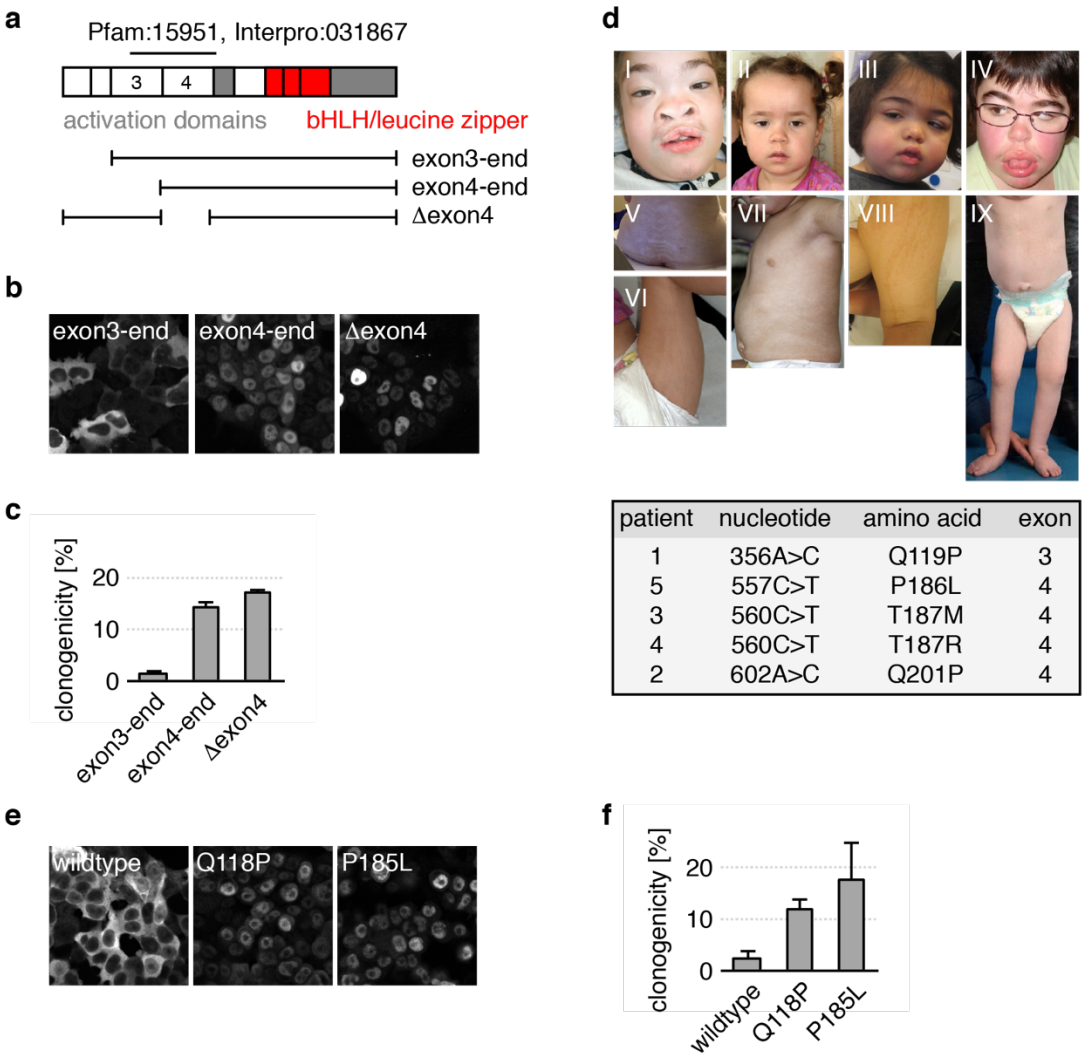
(b-d) Tfe3 localization (b) and retention of self-renewal after 3d of 2i withdrawal in LT1 KO.1 (c) and Flcn KO.1 (d) cells expressing indicated Rag transgenes. Average and SD of 2 biological replicates.

(e,f) Retention of self-renewal in wildtype cells expressing indicated Rag transgenes (e) and indicated cell lines (f) differentiated for 3d in the absence or presence of 100nM Vacuolin-1. Average and SD of 2 technical replicates.

(g,h) Tfe3 interactome in ESCs. Red denotes highly significant enriched proteins, of which blue marks 14-3-3 protein family members and yellow v-ATPase subunits. Quantification is based on 3 biological replicates.

(i,j) Tfe3 localization (i) and retention of self-renewal after 3d of 2i withdrawal (j) in Tfe3 KO.6 cell lines conditionally expressing indicated Tfe3 alleles. Average and SD of 2 biological replicates.

Villegas et al., Figure 4: Disease-associated nuclear Tfe3 gain of function alleles.



patient	nucleotide	amino acid	exon
1	356A>C	Q119P	3
5	557C>T	P186L	4
3	560C>T	T187M	4
4	560C>T	T187R	4
2	602A>C	Q201P	4

Figure 4: Disease-associated nuclear Tfe3 gain of function alleles.

(a) Tfe3 domain and exon structure, and Tfe3 deletion constructs used in (b) and (c). Exons 3 and 4, and the Tfe/Mitf transcription factor family specific homology stretch are indicated.

(b,c,e,f) Tfe3 localization (b,e) and retention of self-renewal after 3d of 2i withdrawal (c,f) in Tfe3 KO.6 cell lines conditionally expressing indicated Tfe3 alleles. Average and SD of 2 biological replicates.

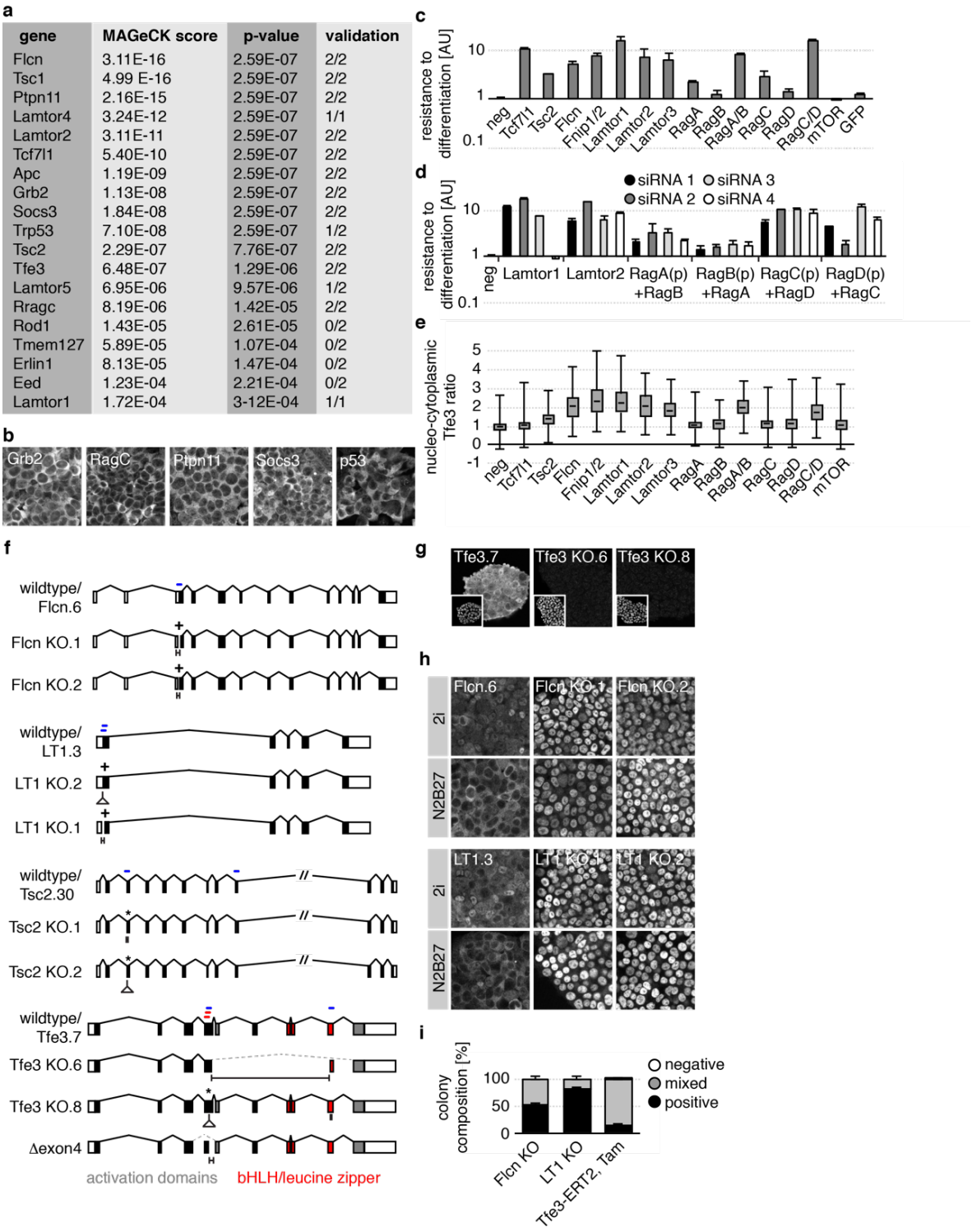
(d) Clinical phenotype of patients and molecular details of corresponding *TFE3* mutations. I-IV: Facial features of patient 1 (I), 2 (II), 3 (III) and 5 (IV). Note the coarse features, facial hypertrichosis and full cheeks common to the individuals. V-VIII:

Cutaneous phenotype of patient 1 (V), 2 (VI,VII) and 3 (VIII). Note the whorls and streak hypopigmentation along Blaschko's line. IX: Body asymmetry, umbilical hernia in Patient 4.

Villegas et al., Extended Data Figure 1

a

gene	MAGeCK score	p-value	validation
Flcn	3.11E-16	2.59E-07	2/2
Tsc1	4.99 E-16	2.59E-07	2/2
Ptpn11	2.16E-15	2.59E-07	2/2
Lamtor4	3.24E-12	2.59E-07	1/1
Lamtor2	3.11E-11	2.59E-07	2/2
Tcf7l1	5.40E-10	2.59E-07	2/2
Apc	1.19E-09	2.59E-07	2/2
Grb2	1.13E-08	2.59E-07	2/2
Socs3	1.84E-08	2.59E-07	2/2
Trp53	7.10E-08	2.59E-07	1/2
Tsc2	2.29E-07	7.76E-07	2/2
Tfe3	6.48E-07	1.29E-06	2/2
Lamtor5	6.95E-06	9.57E-06	1/2
Rragc	8.19E-06	1.42E-05	2/2
Rod1	1.43E-05	2.61E-05	0/2
Tmem127	5.89E-05	1.07E-04	0/2
Erlin1	8.13E-05	1.47E-04	0/2
Eed	1.23E-04	2.21E-04	0/2
Lamtor1	1.72E-04	3-12E-04	1/1



Extended Data Figure 1: Characteristics of loss of function in lysosomal signalling components.

(a) MAGeCK scores and p-values of top 19 screen hits, including validation (phenotypically scoring gRNAs/all gRNAs tested).

(b) Tfe3 localization in long term differentiated RGd2 cells transiently transfected with Cas9 and gRNAs targeting indicated genes.

(c) Resistance to differentiation of Oct4GIP ESCs transfected with indicated siRNAs after 3d of 2i withdrawal. Results were normalized to negative (neg) siRNA cells. Average and SD of 2 technical replicates.

(d) Deconvolution of siRNA pools in O4GIP ESC commitment assay after 3d of differentiation. Results were normalized to neg siRNA cells. Average and SD of 2 technical replicates. (p) denotes pool of 4 siRNAs. Note that the Lamtor3 siRNA pool has been deconvoluted before[4].

(e) Nucleo-cytoplasmic Tfe3 ratios in O4GIP ESCs transfected with indicated siRNAs.

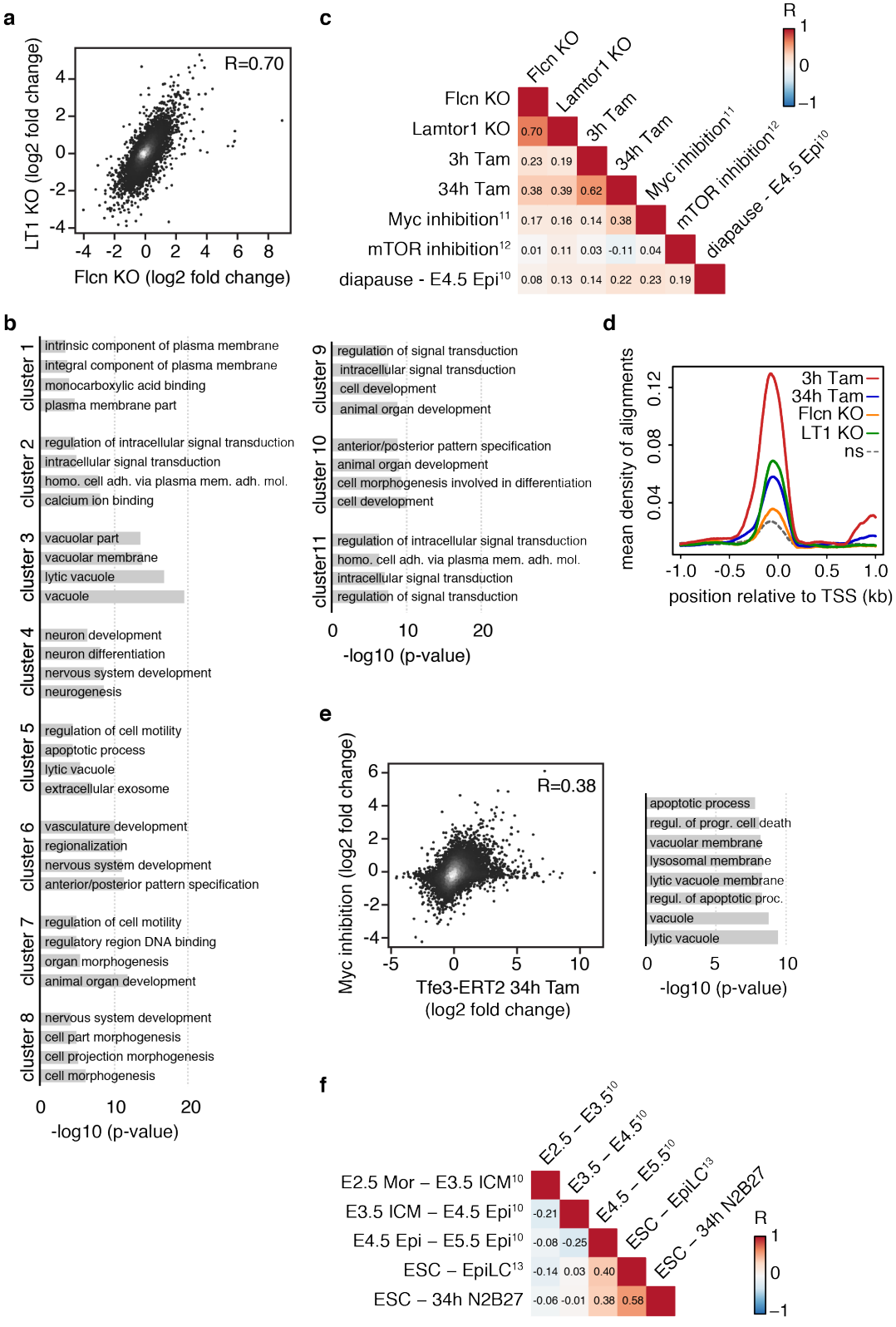
(f) CRISPR/Cas9 mutants generated in this study. Location of gRNAs is indicated with blue bars. Red bars indicate location of Tfe3 gRNAs recovered in the primary screen. (+) indicates mutation of ATG Start codons, (*) generation of in frame STOP codons and dashed lines exon skipping.

(g) Absence of Tfe3 immunoreactivity in Tfe3 KO ESC clones. Insets are DNA staining of the field of view.

(h) Tfe3 localization in 2i and after 24h of 2i withdrawal in indicated cell lines.

(i) Alkaline phosphatase composition of colonies derived from indicated long term differentiated cell lines subjected to clonal self-renewal in 2i.

Villegas et al., Extended Data Figure 2



Extended Data Figure 2: Analysis of RNA sequencing data.

(a) Scatter plot of log2fold changes compared to wildtype controls in Flcn and LT1 KO ESCs.

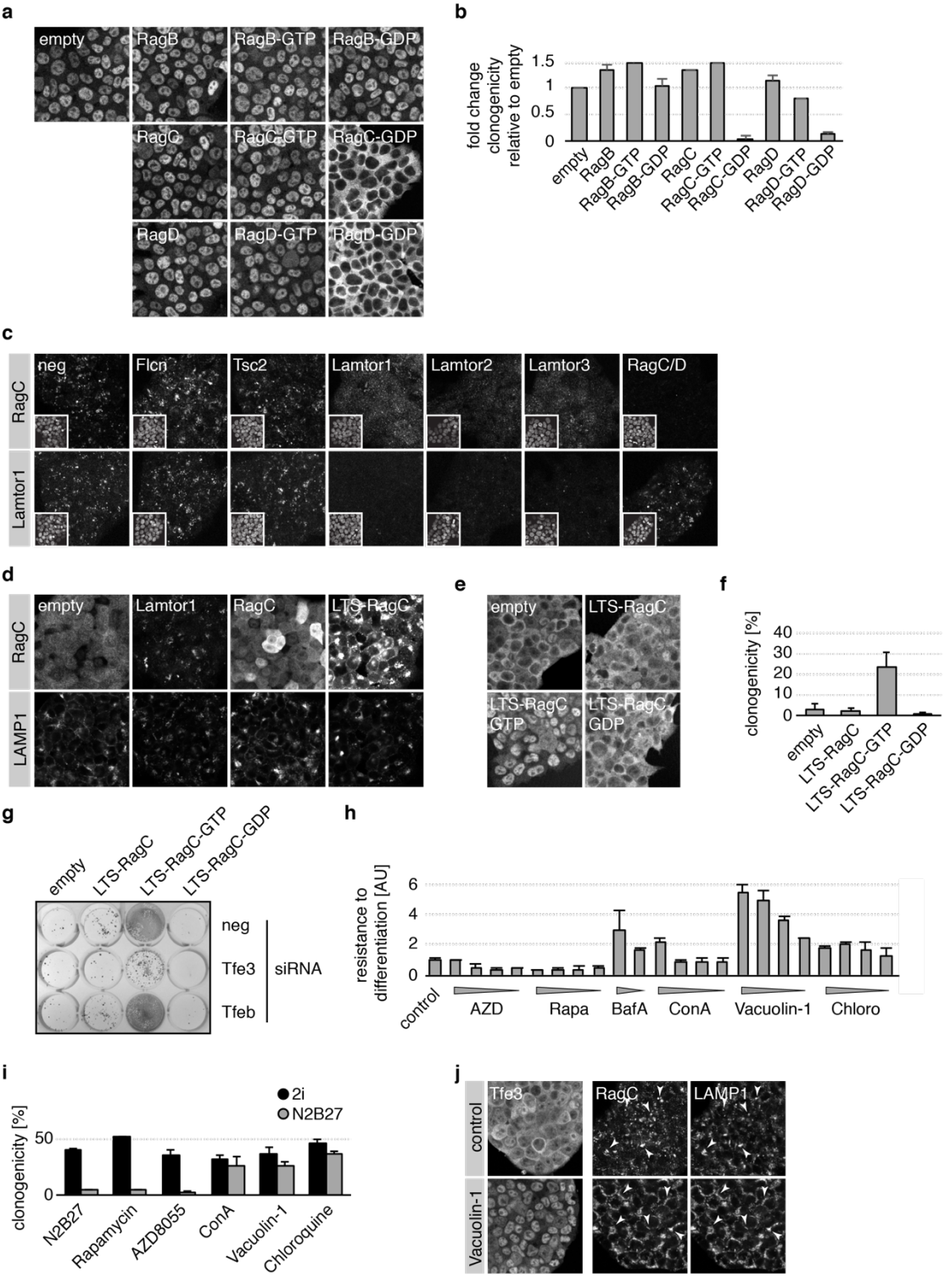
(b) Top five gene ontology (GO) term enrichments in gene clusters 1-11 (Fig. 2a,c). Clusters 1-6 contain early and late Tfe3 target genes after 34h Tam induction in Tfe3-ERT2 ESCs. Clusters 3-6 contain genes congruently altered in Flcn and Lamtor1 KO ESCs, while clusters 7 and 8 are specific to KO ESCs, suggesting the presence of indirect targets and adaption to Tfe3 activation in established KO ESC lines. Clusters 9 and 10 contain genes whose regulation during differentiation is largely unaffected by genetic perturbations. Genes in cluster 11 are specifically upregulated in acute but not long-term differentiated Tam-treated Tfe3-ERT2 or KO cells and are unchanged during wildtype differentiation. This cluster of genes behaves similar to cluster 2, also with respect to GO terms enrichment, and likely reflects transiently induced Tfe3-targets.

(c, f) Pearson's correlation coefficients of pairwise comparisons between log2fold changes of indicated samples. Adaption to Tfe3 activation is reflected in a higher correlation of transcriptional changes in Flcn and Lamtor1 KO cells with 34h compared to 3h Tam treated Tfe3-ERT2 ESCs (c). Changes in differentiating wildtype cells 34h after 2i withdrawal correlate with those reported in the ESC to EpiLC and the E4.5 to E5.5 epiblast (Epi) transition (f). Inner cell mass (ICM), morula (Mor).

(d) Mean density of Tfe3 ChIP-Seq alignments in wildtype ESCs at transcriptional start sites (TSSs) of upregulated genes upon indicated genetic perturbations. Consistent with indirect Tfe3 targets, TSSs of genes induced after 3h Tam treatment had a higher average Tfe3 occupancy when compared to upregulated genes in other samples. (ns) indicates non-significantly changed control transcripts.

(e) Scatter plot of log2fold changes compared to respective wildtype controls in 34h Tam-treated Tfe3-ERT2 ESCs and Myc inhibitor treated ESCs. GO term enrichment identifies an enrichment of lysosome function in genes significantly deregulated in both conditions.

Villegas et al., Extended Data Figure 3



Extended Data Figure 3: Functional analysis of lysosomal signalling.

(a,b) Tfe3 localization (a) and retention of self-renewal after 3d of 2i withdrawal (b) in Flcn KO.1 cells expressing indicated Rag transgenes. Average and SD of 2 technical replicates.

(c) Subcellular localization of RagC and Lamtor1 in O4GIP ESCs transfected with indicated siRNAs. Insets are DNA staining of the field of view. Note cytoplasmic delocalization of RagC upon knockdown of Lamtor1, 2 and 3 but not Flcn or Tsc2.

(d) Subcellular localization of RagC and LAMP1 in LT1 KO.1 ESCs expressing indicated constructs. Note rescue of lysosomal RagC localization upon Lamtor1 expression, and cytoplasmic and lysosomal RagC localization of overexpressed RagC and LTS-RagC, respectively.

(e,f) Tfe3 localization (e) and retention of self-renewal after 3d of 2i withdrawal (f) in wildtype cells expressing indicated RagC transgenes. Average and SD of 2 biological replicates.

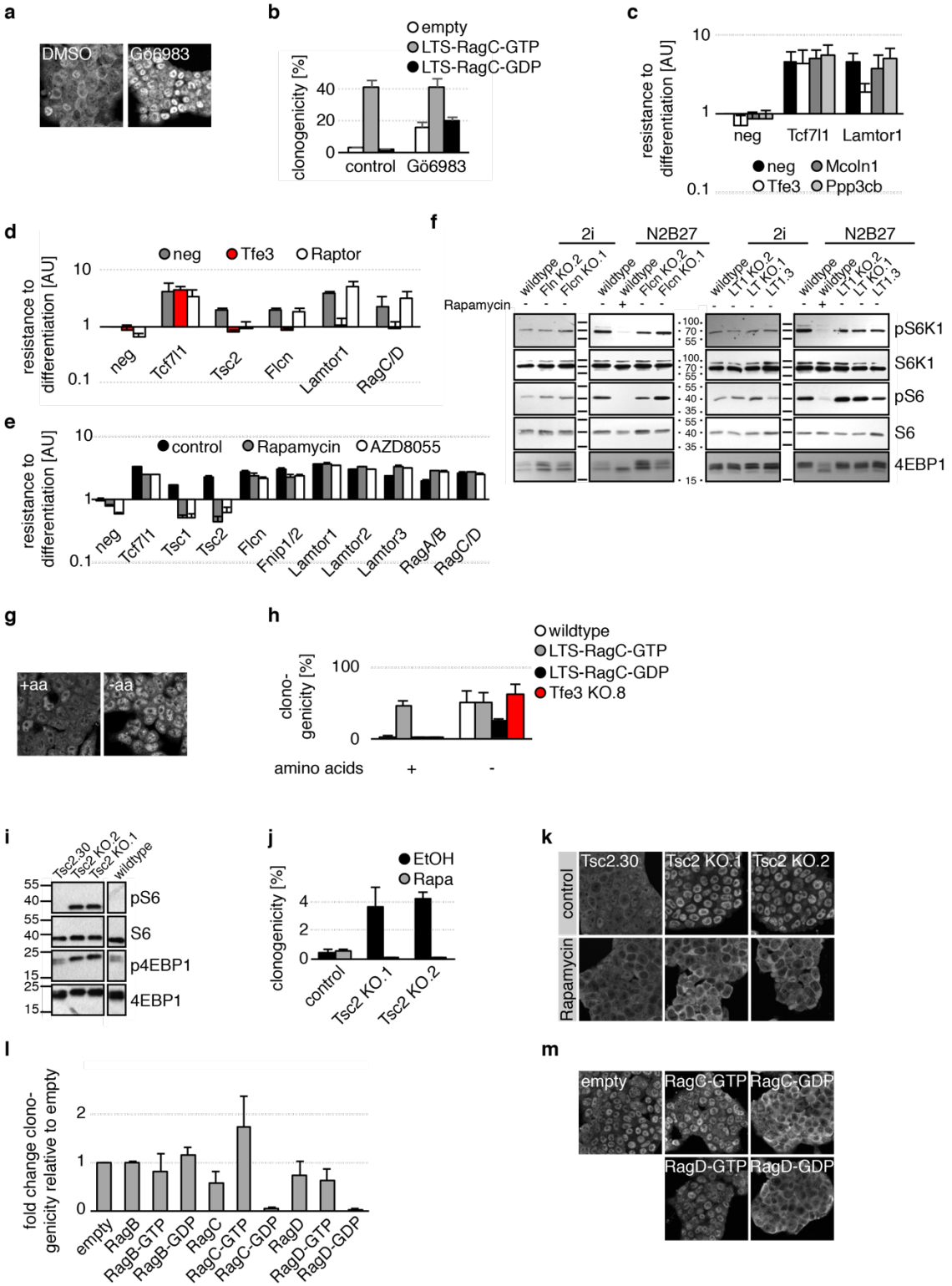
(g) Alkaline phosphatase staining of wildtype cells expressing indicated RagC transgenes and transfected with indicated siRNAs replated after 3d of differentiation in 2i.

(h) Resistance to differentiation of Oct4GIP ESCs differentiated for 3d in the presence of AZD8055 (AZD, 200nM, 100nM, 50nM, 25nM), Rapamycin (Rapa, 80nM, 40nM, 20nM, 10nM), BafA (2.5nM, 1.25nM), ConA (20nM, 10nM, 5nM, 2.5nM), Vacuolin-1 (400nM, 200nM, 100nM, 50nM) and Chloroquine (20μM, 10μM, 5μM, 2.5μM)). Results were normalized to control cells. Average and SD of 2 technical replicates.

(i) Self-renewal of wildtype cells in 2i after 24h treatment with 40nM Rapamycin, 100nM AZD8055, 20nM ConA, 100nM Vacuolin-1 and 10μM Chloroquine in the presence or absence of 2i. Average and SD of two technical replicates.

(j) Tfe3, RagC and LAMP1 localization in ESCs treated for 24h with 100nM Vacuolin-1. Arrowheads indicate colocalization of RagC and LAMP1.

Villegas et al., Extended Data Figure 4



Extended Data Figure 4: mTORC1 hyperactivation, but not inhibition, impairs differentiation.

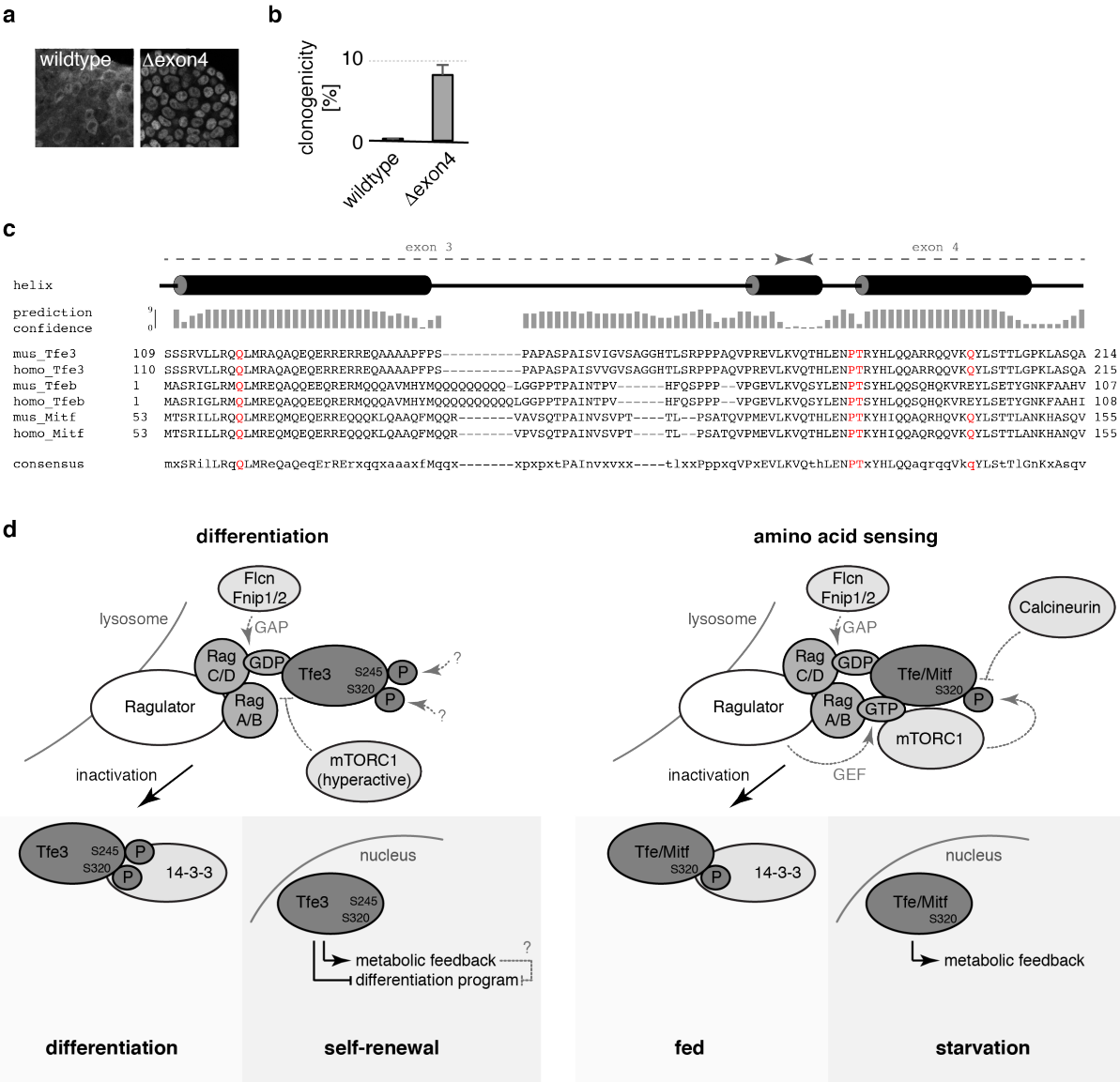
(a,g,k,m) Tfe3 subcellular localization in wildtype ESCs treated for 24h with 2.5μM Gö6983 (a), in wildtype ESCs cultured for 2d with and without amino acids (aa) (g), in indicated cell lines in the presence of 20nM Rapamycin (k), and in Tsc2 KO.1 ESCs expressing indicated Rag constructs (m).

(b,h,j,l) Retention of self-renewal after 3d in N2B27 (b,j,l) or 2d in indicated nutrient conditions (h) of wildtype cells (h) or Tsc2KO.1 cells (j,l) expressing indicated transgenes (b,h,l) or Tfe3KO.8 cells (h) in the presence of 2.5μM Gö6983 (b) or 20nM Rapamycin (j). Average and SD of 2 technical (b), 2 biological (j,l) or 3 biological (h) replicates.

(c,d,e) Resistance to differentiation in O4GIP ESCs transfected with indicated siRNA combinations and treated with 20nM Rapamycin and 50nM AZD8055 (e) after 3d of 2i withdrawal. Results were normalized to neg siRNA cells. Average and SD of 2 technical replicates.

(f,i) Western blot analysis of mTORC1 targets in ESCs and 48h after 2i withdrawal (f) and in indicated ESC lines (i). 20nM Rapamycin was added where indicated.

Villegas et al., Extended Data Figure 5



Extended Data Figure 5: Secondary structure prediction of the Tfe3 inactivation domain and model.

(a,b) Tfe3 localization (a) and retention of self-renewal after 3d of 2i withdrawal (b) in Tfe3Δexon 4 cells. Average and SD of 2 technical replicates.

(c) Amino acid sequence alignment of the exon3-4 boundary in mouse and human Tfe/Mitf family members. I-Tasser secondary structure prediction with confidence values is indicated. Patient mutation homologues are highlighted in red.

(d) Model summarizing lysosomal signalling in ESC differentiation and comparison with mTORC1-dependent amino acid sensing.

Supplementary Table 1: Results of CRISPR/Cas9 screen, normalized read counts, data illustrated in Fig. 1b, 2, Extended Data Fig. 1a and 2.

Supplementary Table 2: Quantitative mass spectrometry, data illustrated in Fig. 3g and h.

Supplementary Table 3: Clinical and molecular features of the six patients with a *TFE3* mutation. Characteristics, inheritance, frequency in the public database ExAC and prediction scores regarding pathogenicity (Polyphen http://genetics.bwh.harvard.edu/pph2/, Grantham and CADD scores) of mutations.

Supplementary Table 4: Oligonucleotide Sequences, siRNAs, sequencing datasets.

Extended Clinical Data: Description of clinical history and phenotypes of patients with *TFE3* mutations.

We identified two germline *de novo* missense variants (NM_006521.4:c.356A>C p.(Q119P) and NM_006521.3:c.602A>C; p.(Q201P) in the X-linked gene *TFE3* in two unrelated females (Patient 1 and 2), who shared a similar phenotype of severe intellectual disability (ID) and hypopigmentation on Blaschko's lines (clinical pictures are summarized in Table S3). The eldest patient had in addition coarse facial features, frontonasal dysplasia, obesity, epilepsy and growth retardation. We thus activated datasharing through GeneMatcher [1], and recruited two additional individuals with pigmentary mosaicism and a *de novo* missense *TFE3* variant. One was a 3 years old female with severe ID, hypopigmentation, epilepsy and coarse facial features, who harbored a germline *TFE3* variant (NM_006521.5:c.560C>T, p.T187M). The other one was a 5 years old male with severe ID, growth retardation, epilepsy, hypopigmentation and a neonatal history of cholestasis, hepatomegaly and hypoglycemia, in whom WES identified a *de novo* mosaic *TFE3* variant (NM_006521.5:c.560C>G, p.T187R), supported by around 65% of sequencing reads in the blood. Finally, we identified a fifth female adult patient with a *de novo* germline missense *TFE3* variant (NM_006521.4:c.557C>T p.P186L) in our local diagnostic WES cohort. She presented with a similar phenotype of severe ID, epilepsy, growth retardation, obesity, and coarse facial features with facial hypertrichosis, but did not have pigmentation anomaly; she had been suspected of a storage disorder. Complementary DNA analysis for patient 1, 2 and 5 revealed that the allele carrying the *TFE3* mutation was supported by respectively 89%, 91% and 95% of the reads. X-inactivation studies showed a random X-inactivation (29-69%) in the non-cultured fibroblasts of the first patient; and a biased X-inactivation (85-15%) in the fibroblasts of the fifth patient.

The first patient was the first child of non-related parents. The family history was unremarkable except for a labial pit in her mother and a unilateral cleft lip and palate in her maternal grandfather. Antenatal karyotype was performed due to high serum markers, which showed a normal result. She was born at 34 weeks of gestation, with normal birth weight (30^{th} centile) and length (60^{th} centile). Examination revealed bilateral cleft lip and palate, hypo-pigmented areas on the back along Blaschko's lines, and a bilateral hip dislocation. Because of the cleft lip and palate in the context of the family history, Van

Der Woude syndrome was suspected; *IRF6* testing by Sanger sequencing was negative. Early milestones were reported as normal until 7 months, when she developed generalized epilepsy and psychomotor regression. When she was referred to a clinical geneticist, at 14 years, she had severe intellectual disability. She could not walk independently, had no speech. She had growth retardation, with length at -4.5 S.D. and obesity (body mass index (BMI) +3 S.D.). OFC was average. Orthopedic examination showed slightly asymmetric and short lower limbs, flat feet and hyperlordosis. Facial dysmorphism was striking with bifid nose, hypertelorism, hypertrichosis and coarse features. Echocardiography demonstrated asymptomatic aortic insufficiency.

The second patient was the first child born from unrelated parents. The mother has deafness as well as her two brothers. Pregnancy was marked by the discovery of ureteral dilatation at 5 months of gestation in a context of maternal gestational diabetes. She was born at 36 weeks of gestation, with fetal macrosomia (weight 99th p., length 90th p. occipitofrontal circonference 95th p.). She presented with neonatal hypotonia and feeding difficulties. She sat at 12 months of age. At last examination at 4 years and 3 months of age she was unable to stand and walk alone, had no speech. She presented sleep disturbance and hand stereotypies. Her weight was +1.5 SD, length was +2.5 SD with a parental target size at +1.2 SD and her OFC was +1 SD. BMI was 15.4 (mean). Brain MRI, skeletal X-rays, eye and auditive exams were normal.

The third female was referred for suspicion of incontinentia pigmenti. She was the first child of nonrelated healthy parents. Pregnancy was complicated by oligohydramnios and third trimester intra-uterine growth retardation. She was born at 41 weeks of gestation with a low birthweight on the 2nd percentile, and was admitted for hypoglycemia in the first two weeks of life. At 6 months she was noted to have delayed milestones and a poor eye contact. She developed tonic seizures from 18 months. Brain MRI brain was normal, EEG showed an abnormal pattern with epileptic encephalopathy and multifocal spikes. She had chronic upper airway infections in the winter, periods during which the parents reported she did not make any developmental progress or even regressed. At examination at 3 years old, she could roll over but could not walk or grasp objects; she made sounds but no words. She had linear and whorled hyperpigmentation all over the body except on the face, palms and soles. Growth parameters were in the normal ranges. She was noted to have umbilical hernia, strabismus and a single palmar crease. The facial features were remarkable by synophris, hypertrichosis, full cheeks and lips, flat nasal bridge,

hypertelorism and thick eyebrows. In the context of the pigmentary changes, chromosomal mosaicism was considered, and a skin biopsy was performed, showing a mosaic translocation t(X;11) with a microdeletion of Xp22.3 in 1 out of 2 cell-cultures. However, as the deletion contained no genes of which haploinsuffiency was known to cause developmental delay and seizures, trio-WES was initiated on the blood, after parental consents.

The fourth patient was a male, from unrelated parents with no family history. The pregnancy was unremarkable. He was born at term by caesarean section for fetal distress, with normal growth parameters and Apgar score at 8 and 9. He was admitted in the neonatal period for seizures on the first day of life, neonatal hypoglycemia, transient renal insufficiency, bronchomalacia, and cholestasis. Examination revealed hepatomegaly, hypotonia, umbilical and inguinal hernias. He had severely delayed milestones. When he was referred at 5 years old, he could not walk or speak, and had just developed epilepsy and hearing loss. He had pigmentation anomalies along Blaschko's lines on trunk and back, asymmetric lower limbs with flat feet, growth retardation on -4.5 SD. Facial features were coarse, he had a flat nasal bridge.

The fifth female patient had been included in the local diagnostic WES cohort because of severe ID and suspicion of storage disorder, despite normal results of extensive metabolic workup. She was the first child of unrelated parents, born at 39 weeks of gestation following an uneventful pregnancy. Birth weight was on 5th centile, length on 10th centile, OFC on 10th centile. Neonatal examination showed bilateral clubfeet. Her first year of life was marked by recurrent infections and by rapid weight gain. Psychomotor development was normal until 9 months; she then made very slow progress, walked at 40 months, and then went in a school for children with special needs. She developed generalized epilepsy from 10 years old with tonic-clonic seizures. Behavioural issues were noted from teenage, treated by antipsychotic drug. On the last examination at 20 years old, she could speak a few words. Her weight was on +4 S.D., length -4 S.D, BMI 30. She had coarse facial features, hypertelorism, flat nasal bridge, anteverted nares, full mouth, thick gums, protruding tongue, facial hypertrichosis and convergent strabismus. Orthopedic examination showed hyperlordosis, limitation in elbow extension, genu valgum, flat feet, small hands and feet. Bone age on the X-ray, as well as the spine, were normal. Blood glucose level was normal, and there was no hepatomegaly on the abdominal ultrasound.

1. Sobreira, N., Schiettecatte, F., Valle, D. & Hamosh, A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum. Mutat.* **36,** 928–930 (2015).

Materials and Methods

Cell Culture

ESCs (RGd2 containing a Rex1:GFPd2-IRES-Blasticidin reporter [1] (gift from Martin Leeb, Vienna), Oct4GIP [2], E14 (gift from Martin Leeb, Vienna)) were cultured on plastic coated with gelatin or laminin (Sigma) in N2B27 medium (DMEM/F12 (Life Technologies), Neurobasal (Gibco) supplemented with N2 (homemade) and B-27 Serum-Free Supplement (Gibco), 1mM sodium pyruvate (Gibco), 2mM L-glutamine (Gibco), 0.1 mM 2-mercaptoethanol (Sigma)), and where indicated with 2i (3μM CHIR99021 and 1μM PD0325901 (Steward lab, Dresden)), knockout serum replacement (Gibco), Rapamycin (Cell Signaling), AZD8055 (Santa Cruz), BafilomycinA (Millipore), ConcanamycinA (Santa Cruz), Vacuolin-1 (Santa Cruz), Chloroquine (Sigma), Gö6983 (2.5μM, Sigma), Myc inhibitor (10058-F4, Sigma), 4-hydroxytamoxifen (0.1μM, Sigma), Doxycycline (0.1μg/ml, Sigma), LIF, FGF and ActivinA (Smith lab, Cambridge). For alkaline phosphatase assays (Sigma), cells were grown on laminin-coated plates, fixed, and stained according to the manufacturer's instruction. For differentiation, ESCs were plated on gelatin-coated plates at 1.5 x 10^4 cells/cm^2 and, the following day, 2i was withdrawn. For EpiLC differentiation [3], knockout serum replacement (1%), FGF (12ng/ml) and ActivinA (20ng/ml) were additionally added. For starvation experiments, ESCs grown on plastic or glass plates were incubated in R-N2B27 (amino-acid free RPMI (homemade), supplemented with B27 and N2, and 3.8mg/ml glucose), supplemented with 2i and amino acids (Sigma; Glycine: 24.375 mg/l, L-Alanine: 3.225 mg/l, L-Arginine Hydrochloride: 115.75 mg/l, L-Asparagine: 4.165 mg/l, L-Histidine: 36.74 mg/l, L-Isoleucine: 79.735 mg/l, L-Leucine: 82.025 mg/l, L-Lysine: 118.625 mg/l, L-Methionine: 23.62 mg/l, L-Phenylalanine: 50.74 mg/l, L-Proline: 12.505 mg/l, L-Serine: 34.125 mg/l, L-Threonine: 74.225 mg/l, L-Tryptophan: 12.51 mg/l, L-Tyrosine: 63.895 mg/l, L-Valine: 59.925 mg/l, L-Glutamic acid: 3.675 mg/l, L-Cysteine: 8.78 mg/l, L-Cystine: 15.645 mg/l, L-Aspartic acid: 3.325 mg/l) as indicated for 2d and differentiation was induced by 2i removal. siRNA (Supplementary Table 4) and plasmid transfections, and exit from self-renewal of O4GIP and RGd2 ESCs was performed and quantified as described before [2].

Immunostaining

Cell were seeded on laminin-coated 96well glass plates (Greiner Bio-One), fixation and immunostaining performed as described [2], and images were captured using a LSM 710 scanning head confocal microscope (Zeiss). Primary antibodies used were anti-Flag (Sigma, F1804, 1:1000), anti-Lamp1 (DSHB, 1D4B, 1:500), anti-Lamtor1 (Cell signaling, 8975S, 1:500), anti-RagC (Cell signaling, 9480S, 1:100) and anti-Tfe3 (Sigma, HPA023881, 1:1000). Quantification of nucleo-cytoplasmic Tfe3 ratios using CellProfiler (Broad Institute) was performed as described [2].

CRISPR/Cas9 screen

Lentiviral gRNA libraries were generated in HEK293 cells. Stable Cas9 expressing female RGd2 ESC (gift from Martin Leeb, Vienna) clones were derived by transfecting with pPB-LR5.1-EF1a-hph2ACas9 (derived from pPB-LR5.1-EF1a-puro2ACas9**,** gift of Kosuke Yusa, Hinxton) and $7.2x10^6$ cells of two independent clones were infected with the lentiviral gRNA library at a 20 fold coverage and multiplicity of infection of 0.25. Stable integrations were selected for 6d in the presence of 0.1 μg/ml Puromycin and 10^7 cells for each clone were plated on Laminin-coated dishes in the absence of 2i, while $2x10^7$ were used for generation of input sequencing libraries. After 3d of differentiation, $2x10^7$ cells were passaged in N2B27 containing 1μg/ml Blasticidin to select for Rex1 expression, and cells at passage 4 and 5 taken for generation of experimental sequencing libraries. Uninfected control cells were lost after the second passage. Genomic DNA was isolated using Proteinase K (Roche) digestion and phenol-chloroform extraction (Sigma). For gRNA sequencing, we developed an approach using unique molecular identifiers (UMIs) for accurate quantification. Biological replicates were divided into aliquots of 500ng genomic DNA (input samples: 8μg total; selected samples: 2μg total). An oligo nucleotide containing UMIs and the 3' part of the Illumina P5 sequence (5'-TCCCTACACGACGCTCTTCCGATCTN(5-9)TCTTGTGGAAAGGACGAAACACC-3') was integrated using Phusion DNA polymerase (Thermo Fisher Scientific) in a single round of denaturation, annealing (62°C) and extension. Then 2μl of exonuclease I (NEB) were added and incubated at 37°C for 45min, followed by heat inactivation and purification using AMPure XP beads (Agencourt) with a 1:1 ratio. gRNAs were amplified using Phusion DNA polymerase for 25 cycles (Fwd: 5'-TCCCTACACGACGCTCTTCCGATCT-3'; Rev: 5'-

GTTCAGACGTGTGCTCTTCCGATCTCTTGTGTAGCGCCAAGTGCC-3';
annealing temperature of 64°C)) and purified using AMPure XP beads (1:1 ratio, elution in 22μl). 5μl were used in an Illumina indexing PCR reaction using Phusion DNA polymerase (6cycles) and NEBnext multiplex oligos set 1 (NEB). Samples were purified using AMPure XP beads (1:1 ratio) and sequenced on one lane of HiSeq2500 (high output mode, 50bp single-end reads). UMI and gRNA sequences were extracted from individual reads by searching for exact matches to the vector backbone sequence at expected offsets (more than 76% of all reads). About 95% of extracted gRNA sequences corresponded to one of the expected sequences from the gRNA library [4], and the number of UMIs was counted for each gRNA. Counts of UMIs were 2.8 to 7.7-fold lower than read counts, indicating a low to moderate level of PCR duplication. As read versus UMI counts displayed a linear relationship, no correction for saturation of highly abundant gRNAs was applied. gRNAs with at least one non-zero count in any sample were selected for further analysis (84,471 gRNAs, 96.6% of the library). Counts from individual replicates for each of the six libraries (2 input, 4 experimental) were summed and analyzed by MAGeCK (version 0.5.4) [5] using parameters "--norm-method total --gene-test-fdr-threshold 0.25 --adjust-method fdr --sort-criteria pos --remove-zero none" to identify gRNAs and genes significantly enriched in the experimental compared to the input libraries.

Genome editing

For C-terminal 3xFLAG Flcn tagging, the TALEN was assembled using Golden Gate TALEN cloning kit [6] (Addgene plasmid 1000000024) into acceptor vectors SV40-ELD and SV40-KKR [7] (FlcnC_F-Q3ELD: HD HD NN NI NI NN HD HD HD HD NI HD NI NN HD NG, FlcnC_R-Q3KKR: HD HD NI NN NI NI NN NN NI NN NG NG HD NG HD NN NN NI NN NG). E14 ESCs were transiently transfected with a recombination reporter [7], TALENs and Flcn tagging template (gtccaccgtccgaagccccacagctacagagtcacggagcGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGtgactccgagaactcctctggaaggtggtgtacagacca) and derivative clones were genotype for successful recombination. CRISPR/Cas9 genome editing was performed by transient transfection of hCas9 and U6-gRNA plasmids [8] (Addgene plasmids 41815 and 41824) into male RGd2 ESCs. Target gRNA sequences are detailed in Supplementary Table 4. For KO

ESC clone generation, a dsRed expression plasmid was cotransfected and single dsRed expressing cells were deposited into 96well plates 2d later. For screen validation, transfected cell pools were differentiated by 2i withdrawal after 2d and serially passaged in N2B27, 1μg/ml Blasticidin.

Molecular biology

Coding sequences were cloned from ESC cDNA and recombined into pDONR221 using Gateway technology (Invitrogen). The Lamtor1 LTS (amino acids 1-39) was inserted by In-Fusion cloning (Clontech) and point mutations (Rag-GDP (RagB: T54N, RagC: S75N, RagD: S77L), Rag-GTP (RagB: Q99L, RagC:Q120L, RagD:Q121L), otherwise indicated) were introduced by polymerase chain reaction. In Tfe3(10xA), S545,S550,S551,S553,S557,S560,S561,S564,S565 and S567 are mutated to alanine residues. Expression destination vectors were pPB-CAG-DEST-pgk-hph [2], pPB-TRE-DEST-rTA-HSV-neo and pPB-TRE-DEST-rTA-pgk-hph. Gene expression analysis was performed using Trizol (Thermo Fisher) RNA extraction, SuperScriptIII (Invitrogen) and TaqMan Universal Master Mix (Applied Biosystem) with the Universal Probe Library (UPL, Roche). An endogenous control (GAPDH, Applied Biosystems) was used for normalizing expression. Primers and UPL probes are detailed in Supplementary Table 4.

Protein methods

Immunoprecipitations using Tfe3 antibodies of three biological replicates of Tfe3.7, LT1.3 and wildtype empty vector control ESCs, and Tfe3KO.6, LT1 KO.1, Flcn KO.1 and wildtype ESCs expressing RagC-GDP and RagD-GDP were performed as described before [2] using DynaBeads (ThermoFisher). Proteins were digested on the beads first with 0.2μg LysC (WAKO) for 6h at 37°C in 2.5M Guanidin-HCl, containing 20mM EPPS pH8.5, 10mM CAA and 5mM TCEP in a total volume of 6ul. Then, samples were diluted with 18ul 50mM HEPES pH 8.5 containing 0.2μg modified porcine Trypsin (Promega) and cleaved overnight at 37°C. Another 0.2μg Trypsin was added and the cleavage was continued for 4 hours. The generated peptides were acidified with 1μl of 20% TFA and analyzed by capillary liquid chromatography tandem mass spectrometry with an EASY-nLC 1000 using the two column set up (Thermo Scientific). The peptides were loaded with 0.1% formic acid, 2% acetonitrile in H2O onto a peptide trap (Acclaim PepMap 100, 75um x 2cm, C18, 3μm, 100Å) at a constant pressure of 800 bar. Peptides were

separated, at a flow rate of 150 nl/min with a linear gradient of 2–6% buffer B in buffer A in 3 minutes followed by an linear increase from 6 to 22% in 40 minutes, 22-28% in 9 min, 28-36% in 8min, 36-80% in 1 min and the column was finally washed for 14 min at 80% B (Buffer A: 0.1% formic acid, buffer B: 0.1% formic acid in acetonitrile) on a 50um x 15cm ES801 C18, 2μm, 100Å column mounted on a DPV ion source (New Objective) connected to a Orbitrap Fusion (Thermo Scientific). The data were acquired using 120000 resolution for the peptide measurements in the Orbitrap and a top T (3s) method with HCD fragmentation for each precursor and fragment measurement in the LTQ according the recommendation of the manufacturer (Thermo Scientific).

Protein identification and relative quantification of the proteins was done with MaxQuant version 1.5.3.8 using Andromeda as search engine [9] and label free quantification [10] as described in[11]. The mouse subset of the UniProt version 2015_01 combined with the contaminant DB from MaxQuant was searched and the protein and peptide FDR were set to 0.01. Only proteins identified with at least three peptides were plotted in Fig. 3g and h. Cell lysates for Western blotting were generated in RIPA buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% Tx-100, 0.1% SDS). Primary antibodies were 4EBP1 (Cell Signaling, 9452S, 1:100), S6 (Cell Signaling, 2217, 1:100), S6K1 (Cell Signaling, 9202S, 1:100), pS6 (S235/236) (Cell Signaling, 2211, 1:100), p4EBP1 (S65) (Cell Signaling, 13443S, 1:100) and pS6K1 (T389) (Cell Signaling, 9205, 1:100).

Microarray data analysis

Raw data (CEL files) were downloaded (all publically available sequencing data used in this study are indicated in Supplementary Table 4), background corrected and normalized using the rma function from the Bioconductor package oligo [12] (package version 1.40.0). A single probeset was selected per gene using annotation from the Bioconductor package mouse4302.db (package version 3.2.3), selecting the probeset with the largest interquartile range for each gene. Differentially expressed genes between Myc inhibition and DMSO samples were identified using the Bioconductor package limma [13] (package version 3.32.0) with default parameters. Genes were considered as significantly deregulated with a minimum absolute fold-change of 2 and a p-value smaller than 0.01.

Next generation sequencing data analysis

Cellular RNA from ESCs, cells 34h after 2i withdrawal, and long-term differentiated cells sorted for Rex1 expression was purified using RNAeasy (Qiagen), strand-specific RNA-seq libraries were generated using TruSeq mRNA Library preparation kit (Illumina), and libraries sequenced on an Illumina HiSeq2500 machine (50bp single-end reads). Reads were aligned using qAlign from the Bioconductor package QuasR [14] (package version 1.16.0) to mouse GRCm38/mm10 genome with default parameters except for splicedAlignment=TRUE, which performs spliced alignments by internally running SpliceMap [15]. EpiLC data were sequenced as 36bp reads [16], and therefore no spliced alignment could be performed. For *in vivo* embryo data [17] pre-existing alignments to mouse GRCm38/mm10 genome downloaded from ArrayExpress (E-MTAB-2958) were used. Alignments were quantified with qCount from the Bioconductor package QuasR [14] (package version 1.16.0) for known UCSC genes obtained from the TxDb.Mmusculus.UCSC.mm10.knownGene package (package version 3.4.0) using default parameters (Supplementary Table 1).

For identification of genes specifically regulated in either ESCs or during differentiation, we used generalized linear models. Only genes with at least 3 counts per million in at least two biological samples were considered (ESCs: 12,478 genes; differentiation: 12,709 genes). Statistically significantly deregulated genes were identified using edgeR [18] (package version 3.18.0) and these genes were fitted to five separate generalized linear models for different sets of samples:

a) ~ *genotype*: Flcn KO or LT1 KO ESCs compared to respective control Flcn.6 and LT1.3 ESCs.

b) ~ *treatment*: 3h Tam or 34h Tam-treated Tfe3-ERT2 ESCs compared to 3h or 34h ethanol-treated Tfe3-ERT2 ESCs.

c) ~ *time*: Flcn.6 and Ethanol-treated Tfe3-ERT2 ESCs and 34h differentiated progeny

d) ~ *time* + *treatment*: 34h Tam-treated differentiated Tfe3-ERT2 cells compared to ethanol-treated Tfe3-ERT2 ESCs and 34h differentiated progeny. *treatment* therefore identifies differentiation-specific nuclear Tfe3 targets.

e) ~*time* + *genotype* + *time:genotype*: Flcn KO.1 ESCs and differentiated progeny compared to Flcn.6 ESCs and differentiated progeny. *time:genotype* therefore identifies differentiation-specific Flcn targets.

Raw P values were corrected for multiple testing by calculating false discovery rates (FDR) [19]. Significant effects in each contrast were defined as changes with a minimum absolute fold-change of 2 and a FDR less than 0.01.

For visualization in Fig. 2a, calculations of RNA log2 fold changes from edgeR were used and only significantly deregulated genes in at least one ESC condition (*genotype*, model (a) and *treatment*, model (b)) were considered (total: 2,483 genes). For visualization in Fig. 2c, significantly deregulated genes in at least one differentiation condition (*time*, model (c), *treatment*, model (d) and *time:genotype*, model (e)) were considered (total: 3,370 genes). Read counts were normalized (dividing by the total number of aligned reads and multiplying with minimal library size, and adding a pseudocount of 8) and log2 transformed. Centred expression values were generated by calculating differences in log-space of each gene in each sample to its average over all samples. For log2 fold change comparisons shown in Extended Data Fig. 2a, c and e, ESC genes detected in this study (12,478 genes) and annotated in other datasets (total: 11,770 genes) were considered. For the pairwise comparison shown in Extended Data Fig. 2f, only differentiation genes detected in this study (12,709 genes) and annotated in other datasets (total: 12,698 genes) were considered. Log2 fold changes were calculated using log2-transformed normalized read counts as described above, averaging biological replicates, and calculating differences (corresponding to fold-changes in linear space) between conditions. Analyses of enriched gene sets (Fig. 2c, Extended Data Fig. 2b and e) were performed using DAVID for GO terms of biological processes, cellular components and molecular functions [20,21]. Principal component analysis (Fig. 2b) was performed on normalized, log2-transformed and mean-centred expression data in R using the prcomp function.

ChIP-seq data analysis

Reads (Supplementary Table 1) were aligned using qAlign from the Bioconductor package QuasR [14] (package version 1.16.0) to mouse GRCm38/mm10 genome with default parameters. A set of non-redundant transcription start sites was created by randomly selecting for each gene a single start site. For meta-gene profiles, reads of 1kb up- and downstream for non-redundant transcription start sites were counted using qProfile from the Bioconductor package QuasR [14] (package version 1.16.0). The average profiles for statistically significantly upregulated genes in ESCs (3h Tam: 368 genes; 34h

Tam: 1,131 genes; Flcn KO: 420 genes; LT1 KO: 622 genes) as well as for a set of non-significantly changing genes (absolute log2 fold-changes <1; P Value: >0.1; ns: 1,882 genes) were plotted.

Patients

The study ascertained five affected individuals and their healthy parents. We obtained written informed consent from all subjects or their legal representatives. Genomic DNA from fresh skin, cultured skin fibroblasts, and blood samples was extracted using the Gentra Puregene Blood and Tissue Extraction Kit (Qiagen). Genomic DNA integrity and quantity was assessed by agarose gel electrophoresis, NanoDrop spectrophotometry, and Qubit fluorometry (Thermo Fisher).

Whole exome sequencing (WES)

We performed 26 trios WES on patient's affected skin-derived DNA and parental blood sample. Exome capture and sequencing for patients 1, 2 and 5 were performed at Integragen (Evry, 209 France) from 1 μg of genomic DNA per individual using the Agilent SureSelect Human All Exon V5. Libraries were sequenced on a HiSeq platform (Illumina) using paired-end 75-bp reads. Sequences were aligned to the human genome reference sequence (GRCh37/hg19 212 build of UCSC Genome Browser), and single-nucleotide variants and small insertions/deletions were systematically detected as previously described [22]. Candidate *de novo* mutational events were identified by focusing on protein-altering and splice-site changes: (1) supported by at least three reads and 10% of total reads in the patient; (2) absent in both parents, as defined by variant reads representing less than 5% of total reads; (3) at base-pair positions covered by at least four reads in the entire trio; and (4) present at a frequency less than 1% in dbSNP (build 147) and 0.1% in the Exome Aggregation Consortium (ExAC Browser, exac.broadinstitute.org).

WES for Patient 3 was enriched using the SureSelect XT Human All Exon V6 kit (Agilent) and sequenced (paired end 150bp) on a NextSeq500 sequencing system (Illumina) at a mean exome depth of 120X. The exome was defined as all coding exons of UCSC and Ensembl +/- 20bp intron flanks. This coverage was >99,5% for known intellectual disability genes, and 96.3% of the exome was covered at least 15X. Reads were aligned to hg19 using BWA (BWA-MEM v0.7.5a) and variants were called using

the GATK haplotype caller (v3.4.46). Detected variants were annotated, filtered and prioritized using the Bench NGS Lab platform (Cartagenia, Leuven, Belgium), and confirmed by Sanger sequencing.

TFE3 sequencing

Regions of interest were amplified using custom intronic primers and long-range polymerase chain reactions with the PrimeSTAR GXL DNA Polymerase (Takara Bio, Saint-Germain-en-Laye, France). Polymerase chain reaction amplicons were pooled, purified, and quantified from each affected individual. Sequencing libraries were prepared using the Nextera XT DNA Sample Preparation kit (Illumina, Paris, France). Paired-end sequencing of 150-bp reads was performed on a MiSeq platform using 300-cycle reagent kits (v2; Illumina, Paris, France).

Complementary DNA analysis

Complementary DNA was sequenced from the first patient's fresh skin, and second and fifth patient's fibroblasts. Total RNA was isolated with TRIzol Reagent (Life Technologies) from fibroblasts, according to the manufacturer's instructions. 1μg RNA was transcribed into cDNA with the QuantiTect Reverse Transcription Kit (Qiagen). Using PCR primers positioned in exons 2 and 6 in *TFE3* gene, complementary DNA was amplified by polymerase chain reaction (PCR) using Prime STAR GXL kit (Takara) according to the manufacturer's protocol. The libraries were prepared with the Nextera XT DNA Sample Preparation Kit (Illumina). Generated libraries were sequenced on a MiSeq instrument (Illumina) according to the manufacturer's recommendations for paired-end 150 bp reads.

Determination of the X-chromosome inactivation pattern at the *HUMARA* locus

X-chromosome inactivation (XCI) pattern was estimated at the *HUMARA* locus as described previously [22] on DNA extracted from uncultured fibroblasts. Fluorescent PCR products were analyzed by capillary electrophoresis on an ABI3130XL genetic analyzer (Applied Biosystems) and peak areas were generated with the GeneMapper software (Applied Biosystems). Skewing was defined as greater than 80% of one X allele active.

1. Wray, J. *et al*. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat. Cell Biol*. **13,** 838–845 (2011).
2. Betschinger, J. *et al*. Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell* **153,** 335–347 (2013).
3. Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. & Saitou, M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146,** 519–532 (2011).
4. Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M. D. C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat. Biotechnol*. **32,** 267–273 (2014).
5. Li, W. *et al*. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol*. **15,** 554 (2014).
6. Cermak, T. *et al*. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*. **39,** e82–e82 (2011).
7. Flemr, M. & Bühler, M. Single-Step Generation of Conditional Knockout Mouse Embryonic Stem Cells. *Cell Rep* **12,** 709–716 (2015).
8. Mali, P. *et al*. RNA-guided human genome engineering via Cas9. *Science* **339,** 823–826 (2013).
9. Cox, J. *et al*. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res*. **10,** 1794–1805 (2011).
10. Cox, J. *et al*. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell Proteomics* **13,** 2513–2526 (2014).
11. Hubner, N. C. *et al*. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J. Cell Biol*. **189,** 739–754 (2010).
12. Carvalho, B. S. & Irizarry, R. A. A framework for oligonucleotide microarray preprocessing. *Bioinformatics* **26,** 2363–2367 (2010).
13. Ritchie, M. E. *et al*. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. **43,** e47–e47 (2015).
14. Gaidatzis, D., Lerch, A., Hahne, F. & Stadler, M. B. QuasR: quantification and annotation of short reads in R. *Bioinformatics* **31,** 1130–1132 (2015).
15. Au, K. F., Jiang, H., Lin, L., Xing, Y. & Wong, W. H. Detection of splice junctions from paired-end RNA-seq data by SpliceMap. *Nucleic Acids Res*. **38,** 4570–4578 (2010).
16. Buecker, C. *et al*. Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* **14,** 838–853 (2014).
17. Boroviak, T. *et al*. Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. *Dev. Cell* **35,** 366–382 (2015).
18. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. **11,** R25 (2010).
19. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B* **57,** 289–300 (1995).
20. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. **37,** 1–13 (2009).

21. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4,** 44–57 (2009).
22. Allen, R. C., Zoghbi, H. Y., Moseley, A. B., Rosenblatt, H. M. & Belmont, J. W. Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am. J. Hum. Genet.* **51,** 1229–1239 (1992).

5 Addendum

The main results are presented in the manuscript submitted for publication. Additional data acquired during my PhD project besides the manuscript will be documented within the following part.

5.1 A genome-wide genetic screen has identified components of the amino-acid sensing required for ESC differentiation

Exit from self-renewal in ESC can be induced *in vitro* by the removal of the 2i and is followed by the initiation of differentiation. We have performed a genome-wide loss-of-function (LOF) screen that has recovered genes previously identified in the pluripotency network (e.g. Tcf3) serving as a positive control. Importantly, the screen recapitulated previous finding that Folliculin (Flcn) is key for the exit of self-renewal state (Betschinger et al., 2013). Interestingly, among the validated hits, we could find four out of the five Lamtor proteins that constitute the ragulator protein complex, a principal component of the amino-acid sensing machinery (Sancak and Sabatini, 2009).

However, this pathway requires several additional proteins that we did not identify in the screen such as GATOR proteins. Lysosomal recruitment of GATOR through KICSTOR protein complex regulates the activity of the Rag GTPases in the amino-acid sensing machinery (Wolfson et al., 2017). To test the importance of GATOR complex for ESC differentiation, we depleted the GATOR1 and GATOR2 sub-units, Depdc5 and Mios respectively, by siRNA knockdown. To assess the role of these proteins in ESC differentiation, we compared the strength of differentiation defect upon knockdown of the selected genes by measuring the resistance to differentiation (Fig. 5). Whereas the knockdown of Flcn, Ragulator subunits and Rag GTPases compromised the differentiation efficiency, the downregulation of Depdc5, Mios5 and mTOR had no effect on ESCs differentiation.

This supports the global screen findings that some components of the amino-acid sensing mechanism play a role in the exit from pluripotency, but not the entire pathway.

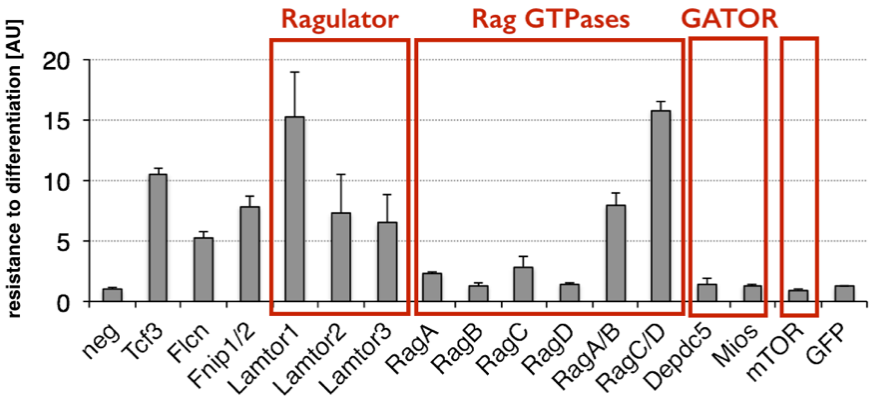


Figure 5: Loss of ragulator and Rag C/D GTPases impair ESC differentiation

Resistance to differentiation of Oct4GIP ESCs transfected with indicated siRNA combinations after 3 days of 2i withdrawal. Results were normalized to negative siRNA cells. Average and standard deviation of 2 technical replicates.

5.2 Flcn is recruited to the lysosome upon starvation in ESCs

Although Flcn was shown to be essential for developmental progression, little is known about its molecular function in the control of ESC differentiation (Betschinger et al., 2013); (Baba et al., 2006). Flcn has been reported to be a cytoplasmic protein, recruited to the lysosome membrane upon amino-acid starvation (Tsun et al., 2013); (Petit et al., 2013); (Péli-Gulli et al., 2015). In order to elucidate Flcn function in ESCs, we sought to gain mechanistic insights into its action. Because of the lack of Flcn antibody compatible with immunofluorescence techniques, we generated a functional Flag-tagged Flcn allele in ESC by using TALEN genome editing system. Since the absence of functional Flcn protein leads to nuclear accumulation of Tfe3, we verified that the subcellular localization of Tfe3 in the Flag-Flcn-knock-in ESCs was the same as in the WT cells (data not shown). Immunofluorescence analysis of the endogenously tagged protein revealed that Flcn is mainly distributed in the cytoplasm (see manuscript, Fig. 3a). To test whether the lysosomal recruitment of Flcn in ESCs also occurs in an amino-acid dependent manner, nutrient starvation was induced by amino-acids withdrawal from the culture conditions. After starvation for 50 minutes, we could observe colocalization of Flcn and the lysosomal protein marker Lamp1 (Fig.6). Upon re-addition of amino-acids in the medium, Flcn appeared released from the lysosome

and became evenly distributed within the cytoplasm. These observations indicate that Flcn is recruited to the lysosome upon nutrient starvation in ESC, as it has been reported for other mammalian cell lines (Petit et al., 2013) (Tsun et al., 2013).

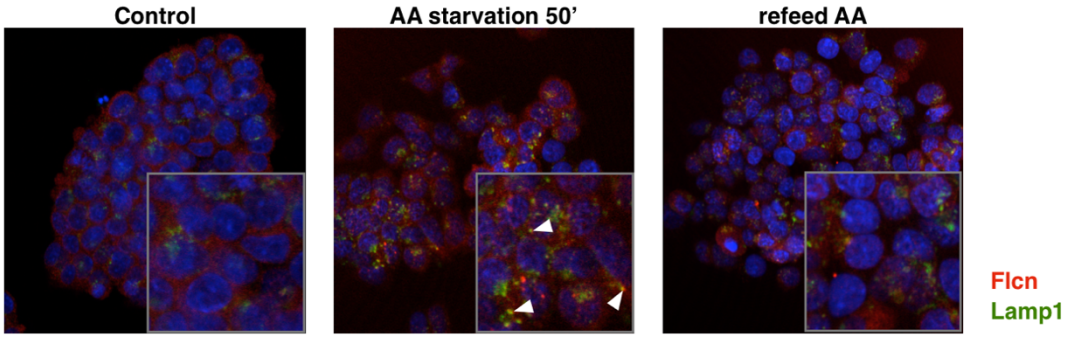


Figure 6: Flcn recruitment to the lysosome in ESC occurs in response to amino-acid starvation

Immunofluorescence staining during amino-acid (AA) starvation. Upon AA withdrawal for 50 minutes, Flcn colocalizes with the lysosomal protein Lamp1. Starved ESCs refed with AA for 10 minutes show an even distribution of Flcn in the cytoplasm.

5.3 Flcn interacts with Fnip1/2 as well as Gabarap in ESCs

Flcn has two described interacting partners, Fnip1 and Fnip2 (Fnip1/2) and similar to Flcn LOF, loss of Fnip1/2 leads to differentiation block in ESCs. To identify specific molecular partners of Flcn in ESCs, we made use of the Flcn knock-in ESC line previously described. Flag-tagged immunoprecipitates of Flcn from cellular lysates of ESCs as well as differentiated cells for 48 hours were analyzed by mass-spectrometry.

Similar to previous findings in other cell types, Flcn interacted with Fnip1/2 (Fig. 7a). Additionally, we identified GABA(A) receptor-associated protein (Gabarap) as an interactor of Flcn in ESCs. Gabarap regulates autophagy and binding of Gabarap to Flcn is regulated by the autophagy kinase Ulk1 (Dunlop et al., 2014). Gabarap and Flcn interaction is further discussed in the section part 5.4.2. To identify differential binding partners of Flcn, we sought to investigate if changes in Flcn interactome occur upon differentiation. We compared the protein interactome of Flcn in the pluripotent state and after two days of differentiation (Fig. 7). Differential

comparison of both interactome did not show any significant changes in binding partners of Flcn.

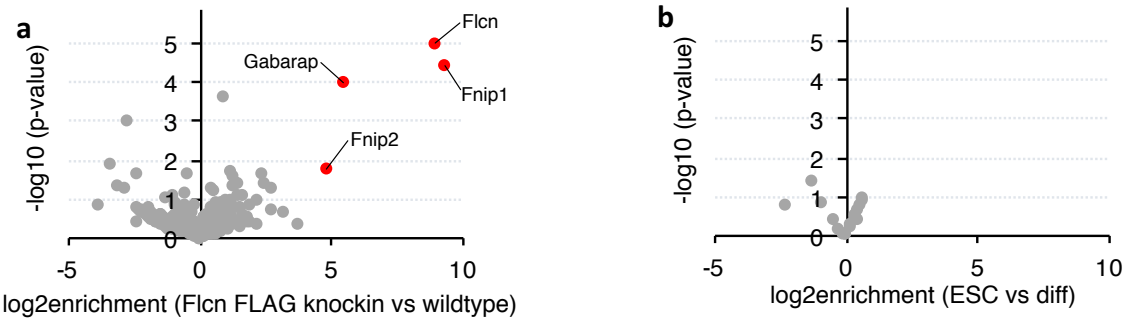


Figure 7: Flcn interactome and differential binding between 2days differentiated cells and ESCs

Label-free quantitative mass spectrometry analysis of Flag-immunoprecipitates from ESCs (a) and differentiating cells (b). Analysis revealed binding of Flcn to Fnip1 and Fnip2, and Gabarap with no differential binding between ESC and differentiated cells. Quantification is based on three biological replicates.

5.4 Metabolic control of ESC differentiation

5.4.1 Nutrient availability controls cell fate transition in ESCs

Our previous observations from the screen led us to hypothesize that key regulators of the amino-acid sensing machinery are required for ESC differentiation. However, whether this mechanism is functional in ESC remains to be determined. We addressed this question in ESCs by performing amino-acid starvation followed by amino-acid stimulation (Fig. 8). As a read out we measured phosphorylation levels of 4EBP1, a direct target of mTORC1, since it has been shown in other cell types that mTORC1 activity can be regulated by nutrient availability. After 50 minutes of amino-acid starvation, mTORC1 activity is down-regulated in ESCs while it gets up-regulated in response to amino- acid stimulation after starvation for 50 minutes.

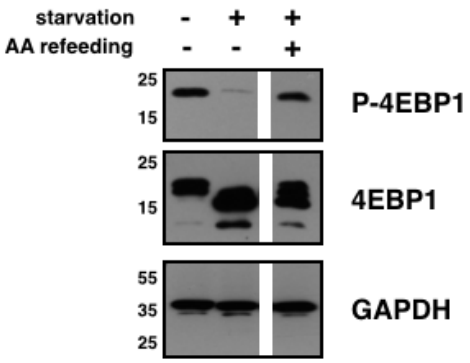


Figure 8: Amino-acid starvation in ESCs down-regulates mTORC1 activity

Western-blot analysis. Amino-acid (AA) starvation for 50 minutes induces down-regulation of mTORC1 activity as seen by the reduced phosphorylation level of the mTORC1 target 4EBP1. AA stimulation after starvation re-activate mTORC1 activity.

Although ESC differentiation does not involve any nutrients alterations, we sought to investigate if amino-acids act in the control of ESC differentiation. To find whether a single or a group of amino-acids is particularly involved in the control of ESC differentiation, we tested if its omission impairs ESC differentiation. To do so we developed a cell culture medium based on the commercially available RPMI medium called R-N2B27 (RPMI based medium that recapitulates the N2B27 medium usually used) that allows the modulation of nutrient composition. Since ESC metabolism depends largely on glutamine and its absence does not impair ESC differentiation (data not showed), this amino-acid was added to the medium by default. As we have shown previously, amino-acid starvation in ESC promoted nuclear enrichment of Tfe3 and impaired differentiation (see manuscript, extended data Fig. 4g-4h).

To determine if differentiation impairment occurs upstream of the transcription factor Tfe3, we conducted the experiment in parallel in WT and Tfe3 KO ESCs. Since the two cell lines express a GFP driven by pluripotency reporter gene (Rex1), we monitored decrease of GFP intensity by flow cytometry as a measure for initiation of ESCs differentiation. Removal of essential amino- acids (EAA) or non-essential amino- acids (NEAA) led to a differentiation defect (Fig. 9a). Removal of individual amino-acids, essential or not, impaired differentiation, independently of Tfe3. A representative sample is shown in the figure bellow.

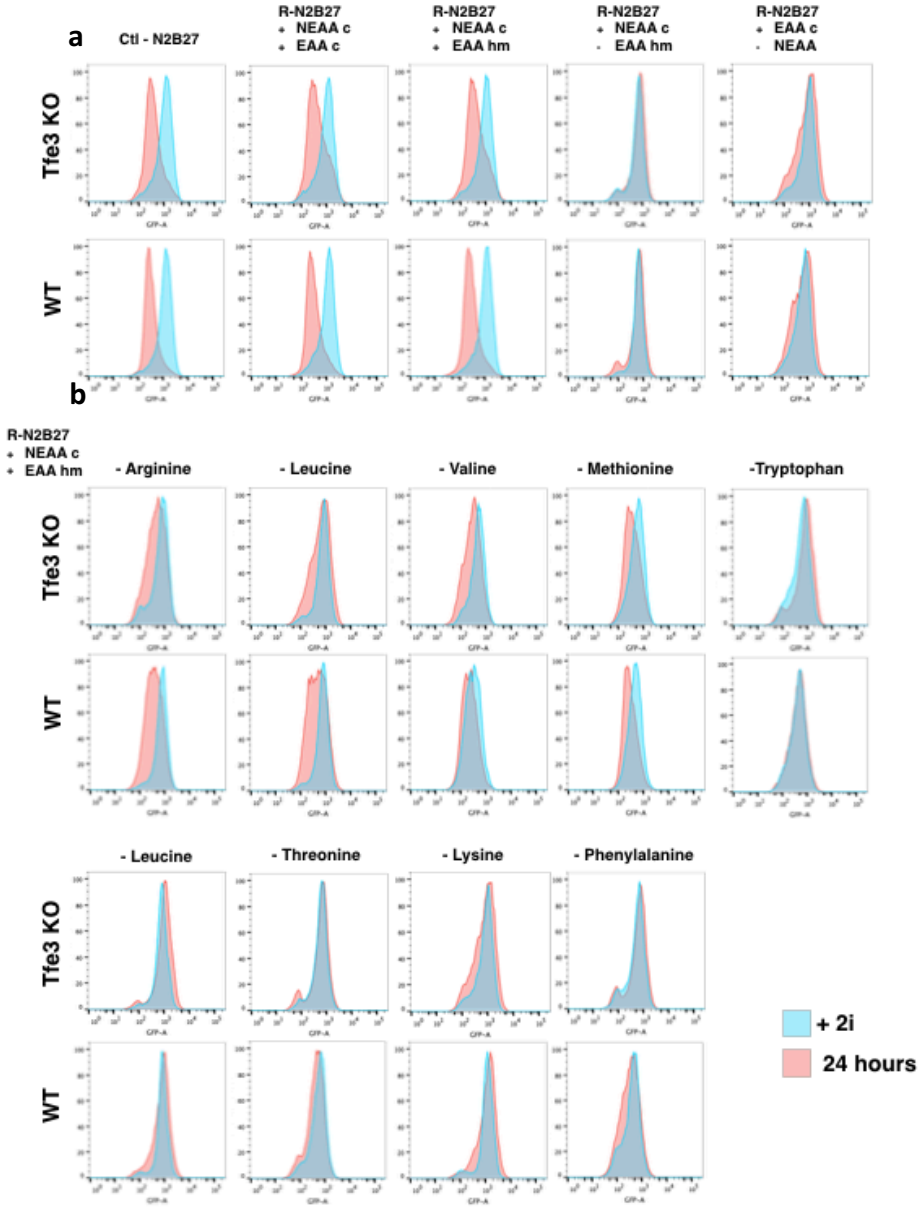


Figure 9: Amino- acid starvation impairs ESC differentiation

Flow cytometry profiles of WT or Tfe3 KO Rex1GFPd2 ESCs. Upon differentiation, GFP intensity decreases. Pluripotent ESCs (blue) and differentiated cells for 24 hours (red). a) In the RPMI based medium (R-N2B27), in presence of essential and non-essential amino-acids (EAA and NEAA, respectively), ESCs differentiate normally compared to the commercial medium (N2B27). Removal of the EEA or the NEAA impairs differentiation of WT and Tfe3 KO ESCs. b) Removal of a single amino-acid residue impairs differentiation in WT and Tfe3 KO ESCs.

5.4.2 Autophagy is not required for the exit of pluripotency

Cellular stresses like starvation often lead to autophagy, a degradative mechanism ensuring cellular homeostasis and cell survival. Autophagy is principally

orchestrated by the lysosome, the key organelle that performs the enzymatic digestion of endogenous macromolecules in order to generate elementary molecules essential for the maintenance of basic anabolic reactions. We have already addressed the role of the lysosome *per se* for ESCs differentiation (see manuscript) and we concluded that the lysosomal functions are required upstream of the nucleotide loading state of RagC/D GTPase and Tfe3, in order to drive exit from pluripotency. During somatic cell reprogramming into pluripotent stem cell, it has been shown that autophagy is induced at an early stage in this process (Mizushima and Levine, 2010). In line with acting in the autophagy pathway, loss of Flcn has been described to impair autophagy (Dunlop et al., 2014).

Additionally, we have identified Gabarap as a Flcn interacting partner in ESCs. Through its binding to Flcn, Gabarap regulates autophagy and the binding of Gabarap to Flcn is regulated by the autophagy related kinase Ulk1 (Dunlop et al., 2014). The transcription factor Tfeb, another MiTF family member, has also been implicated in the regulation of autophagy (Settembre et al., 2012).

Based on these observations, we tested the role of autophagy in the exit from self-renewal of ESCs. Down regulation of key genes involved in autophagy did not impair ESC differentiation nor Tfe3 subcellular localization therefore suggesting that autophagy is not required for the exit from pluripotency in ESC (Fig. 10).

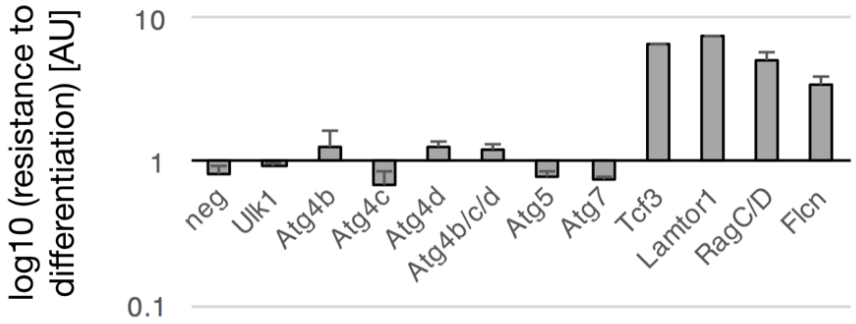


Figure 10: Inactivation of autophagy does not impair ESC differentiation

Resistance to differentiation of Oct4GIP ESCs transfected with indicated siRNAs after 3d of 2i withdrawal. Results were normalized to negative siRNA cells. Average and SD of 2 technical replicates.

5.4.3 mTORC1 activity in Flcn and Lamtor 1 KO ESCs

In our loss of function screen, we identified genes that are essential for ESC exit from self-renewal upon long term exposure to differentiation stimuli. Some of these genes have been described to act as activators or inhibitors of mTORC1 in the amino- acid sensing machinery (Fig. 3). In addition, loss of function of Flcn, Lamtor1,2 or 3, and by extension the Ragulator, decrease mTORC1 activity. Wether these findings can be recapitulated in ESCs remained to be addressed. We further investigated mTORC1 activity in WT, Flcn and Lamtor1 KO ESCs, as well as upon initiation of ESC differentiation. In ESC culture conditions, ESCs display a low basal activity of mTORC1. Neither Flcn nor Lamtor1 KO ESCs showed changes of mTORC1 activity level compared to WT ESCs. Interestingly, upon initiation of differentiation, mTORC1 activity increased in WT ESC as well as in Flcn and Lamtor1 KO ESCs despite the two KO lines being impaired for differentiation (see manuscript, extend data Fig. 4f). Since the differentiation defect of Flcn and Lamtor1 KO ESCs can be rescued by expression of a recombinant version of RagC GTPase in a GDP bound form, we aimed to determine if ESCs overexpressing RagC preferentially bound to GDP or GTP differentially modulate mTORC1 activity. GDP and GTP nucleotide-bound forms of RagC have been described as active and inactive form in regulating mTORC1 activity, respectively. ESCs expressing RagC preferentially bound to GDP or GTP displayed similar level of mTORC1 activity compared to control conditions or ESC expressing a wild-type RagC transgene (Fig. 11). This observation therefore suggests an alternative mode of regulation of mTORC1 activity in ESC.

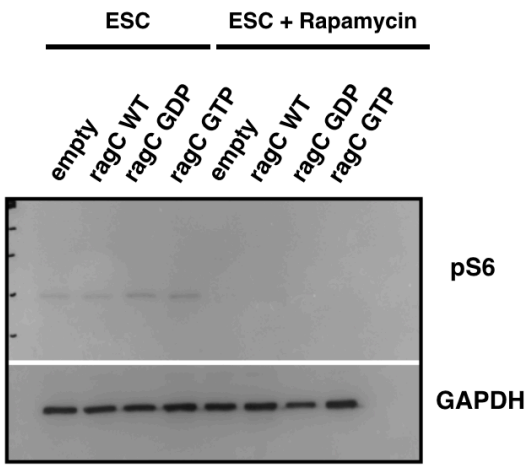


Figure 11: RagC GTPase GDP or GTP bound form do not regulate mTORC1 activity in ESC

Western-blot analysis. Over-expression of RagC GTPase either bound to GDP or GTP do not increase mTORC1 activity measured by phosphorylation level of the ribosomal protein S6. Rapamycin treatment is used as a control that completely inhibits mTORC1 activity.

Additionally, we have investigated mTORC1 localization in ESCs expressing RagC GDP or GTP bound forms. In control ESCs expressing an empty vector, mTORC1 is cytoplasmic although it colocalizes with the lysosomal marker Lamp1. Overexpression of RagC GDP or GTP nucleotide bound form does not lead to any changes in mTORC1 localization.

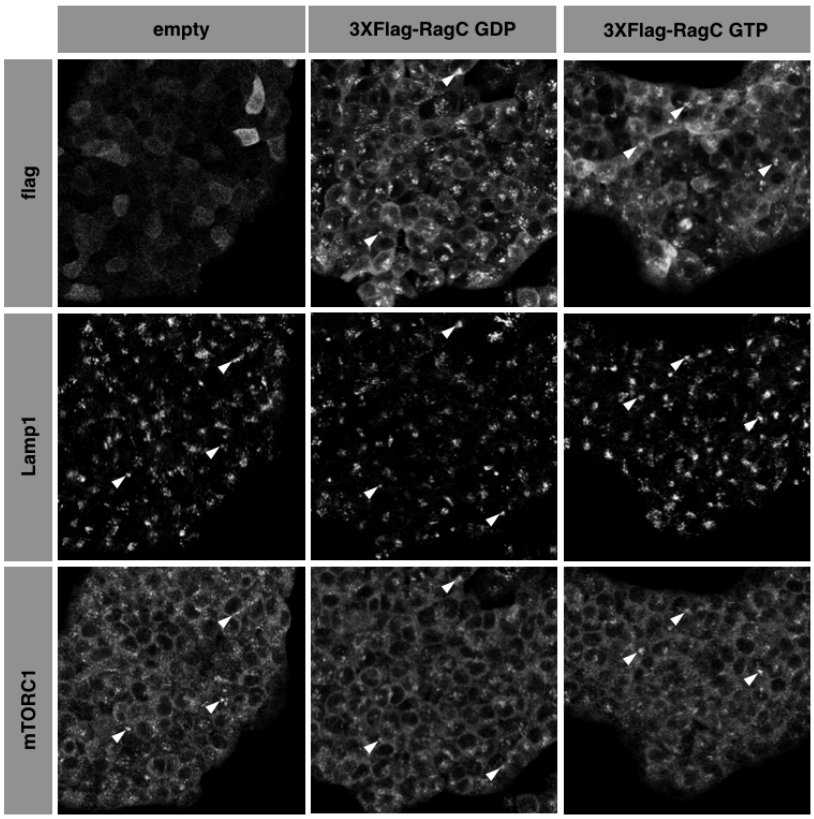


Figure 12: mTORC1 localization in ESCs does not change upon expression of RagC GDP nor GTP

Immunofluorescence staining. Overexpression of Flag-RagC is confirmed using a Flag antibody. Both RagC GDP and GTP colocalize with the lysosomal protein Lamp1. mTORC1 immunostaining do not display significant sub-cellular localization changes compared to control (empty).

5.4.4 Hyper-activation of mTORC1 upon Tsc2 LOF impairs ESC differentiation, downstream of RagC/D GTPases

Loss of Tsc2 in ESC results in a differentiation defect accompanied by a nuclear enrichment of the transcription factor Tfe3. In contrast to the loss of Flcn or Lamtor1 in ESCs, Tsc2 LOF phenotype can be reverted by rapamycin treatment and is consequently mTORC1 dependent. Therefore, Tsc2 KO ESCs lines have been generated where an mTORC1 upregulation was observed (see manuscript, extend data Fig. 4i). Studies have shown that Tsc2 mediated mTORC1 regulation via Tsc2 occurs via lysosomal recruitment and this subcellular localization is regulated by cellular stresses like nutrient starvation (Demetriades et al., 1AD). Indeed, mTORC1 needs to be translocated to the lysosome where it gets activated by Rheb GTPase.

We studied mTORC1 localization changes upon Tsc2 LOF compared to WT ESCs by immunofluorescence (Fig. 13). Although Tsc2 LOF strongly induces mTORC1 activity, significant changes in mTORC1 localization could not be observed. Since overexpression of RagC GDP bound form reverted Tsc2 LOF phenotypes, we repeated the same immunofluorescence experiment in both cell lines. Overexpression RagC GTPases in GDP or GTP nucleotide loading state did not change mTORC1 localization.

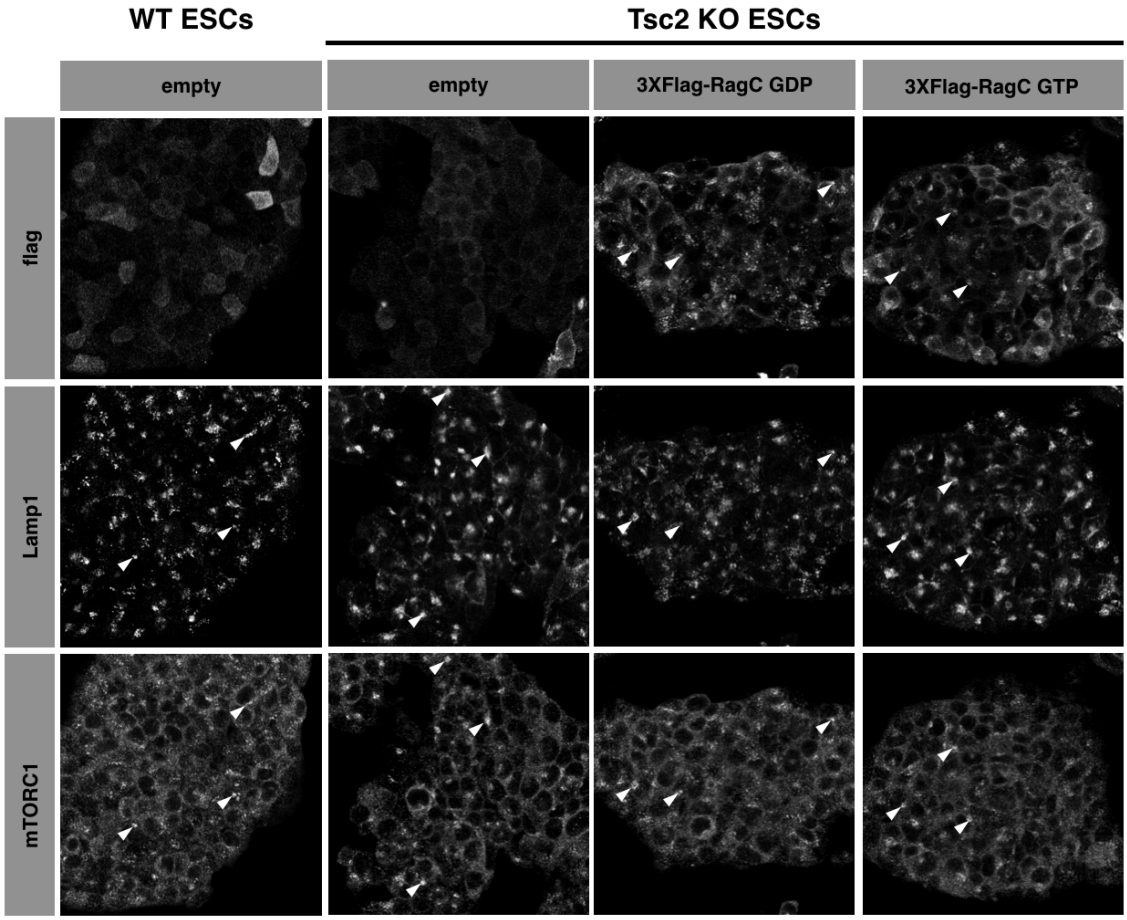


Figure 13: mTORC1 localization in Tsc2 KO ESCs is unchanged upon RagC GDP or GTP expression

Immunofluorescence staining. Overexpression of Flag-RagC in confirmed by using a Flag antibody. Both RagC GDP and GTP colocalize with the lysosomal protein Lamp1. mTORC1 immunostaining do not display significant sub-cellular localization changes in Tsc2 KO ESCs compared to control WT ESCs (empty), nor upon over-expression of any of the RagC bound form.

5.4.5 Tfe3 preferentially binds RagC GDP bound form

Since the nucleotide loading state of RagC/D controlled by Flcn is crucial for ESCs differentiation, we sought to investigate the differential binding partners of RagC GDP and RagC GTP bound form.

We aimed to identify the differential binding partners of RagC GTPases whether they are preferentially bound to GDP or GTP. We found that Tfe3 preferentially binds RagC /D GTPases and this interaction was reported to localize at the lysosomal surface. It was also shown that the binding preferentially occurs when Rag C/D is in the GDP bound form and we could confirm this finding by immunoprecipitation of flag-tagged RagC transgenes (GDP or GTP bound form) and by confocal microscopy. Nevertheless, this interaction might be transient and only a small fraction of Tfe3 is able to bind RagC/D GTPase at the lysosome. We aimed to capture the transient interaction by inducing it while overexpressing Rag's GDP/GTP bound and consequently immunoprecipitated endogenous Tfe3 (Fig. 14).

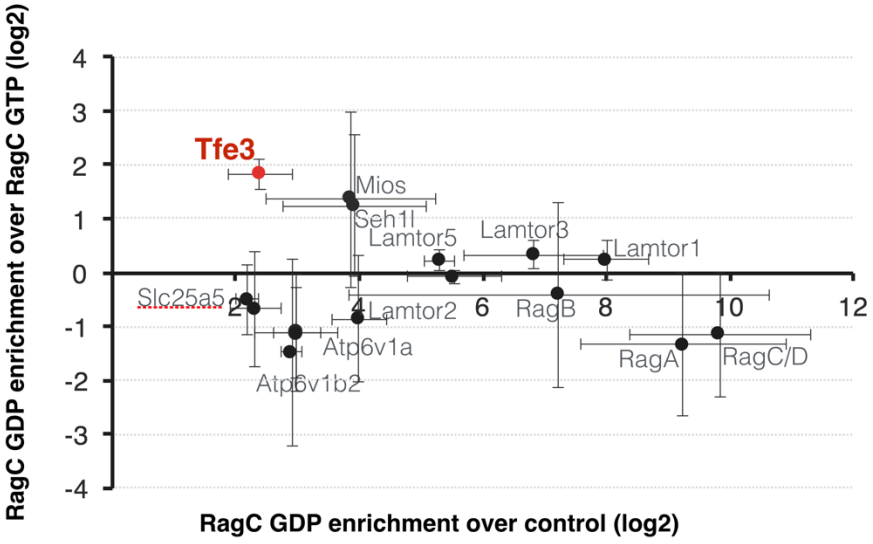


Figure 14: Tfe3 preferentially binds RagC GDP bound form

Tfe3 interactome and differential binding partners of cytoplasmic and nuclear Tfe3. On the x-axis are displayed proteins enriched in RagC-GDP purifications and on the y-axis the ratio of binding to GDP over GTP locked RagC. Known RagC interactors are recovered including the Rag heterodimer as well as Ragulator subunits without significant enrichment in the GDP-bound version. In contrast, Tfe3 is identified as specific binding protein of RagC GDP.

5.5 Tfe3 sub-cellular localization dictates ESC's fate

In ESC, Tfe3 is localized in both cytoplasmic and nuclear compartments. Upon differentiation, Tfe3 is excluded from the nucleus and remains cytoplasmic.

We aimed to understand how this mechanism is regulated. We sought to identify Tfe3 post-translational modifications as well as specific binding partners in WT and Flcn KO ESCs. We made use of an inducible tetracycline-controlled transcriptional activation (TetOn) system to express a Flag-tagged Tfe3 transgene upon Doxycycline (Dox) treatment (Fig. 15a). This overexpression system allows for enrichment of the starting material. After Dox treatment of transfected WT and Flcn KO ESCs, expression of transgenic Tfe3 was controlled by immunofluorescence using a Flag antibody. The observed subcellular localization of the transgenic protein followed the distribution of endogenous Tfe3 in the respective genetic background (Fig. 15b).

To further identify specific partners and differential post-translational modifications of Tfe3 regulating its differential subcellular localization, Flag-tagged Tfe3 immunoprecipitates were purified after 48 hours of Dox treatment and analyzed by mass spectrometry (MS).

Although Tfe3 interacting partners were not different between WT and Flcn KO ESC, MS analysis allowed for the detection of differential post-translational modifications. Phosphorylation of a serine residue in position 320 (P-S320) was significantly enriched when Tfe3 was immunoprecipitated from WT ESCs. P-S320 post-translational modification has been described as an mTORC1 phosphorylation target site, required for the binding with the cytoplasmic protein 14-3-3. Nevertheless, this phosphorylation site is not sufficient to explain the phenotype.

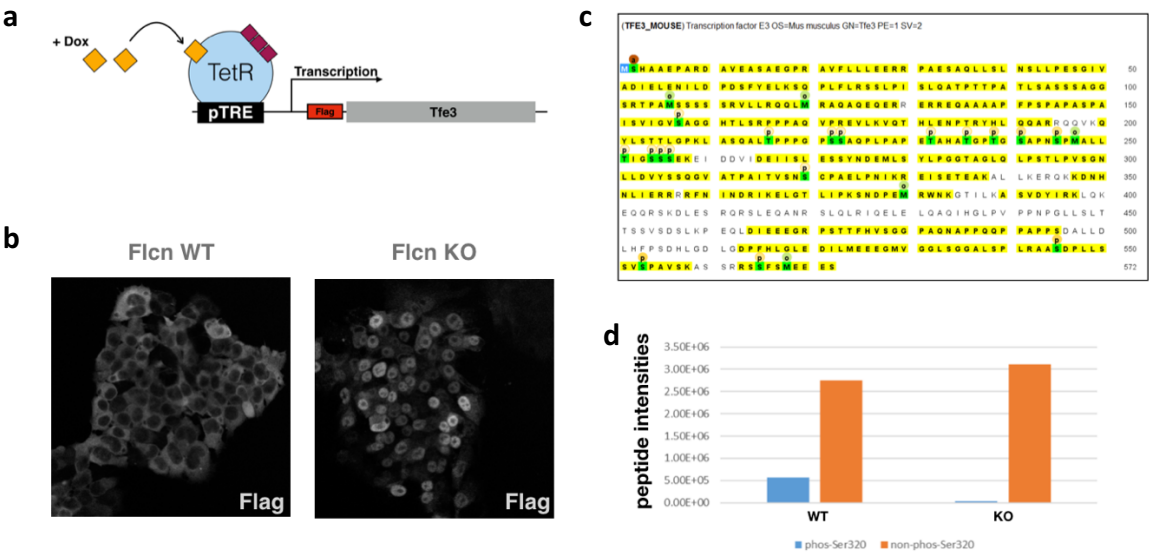


Figure 15: Tfe3 serine 320 is phosphorylated in the cytoplasm

a) schematic of the inducible Flag-tagged Tfe3 transgene b) immunofluorescence using a Flg antibody shows the subcellular localization of the Tfe3 transgene upon doxycycline treatment in WT and Flcn KO ESCs c) peptide sequence coverage of the transgenic Tfe3 after mass spectrometry analysis d) peptide count intensity shows increased phosphorylation of the Serine residue in position 320 of Tfe3 immunoprecipitated from WT ESCs.

Because of possible drawbacks linked to this artificial system, we made use of an antibody targeting the endogenous Tfe3 protein for further experiments where we concluded that Tfe3 binds preferentially RagC/D in a GDP bound conformation (see manuscript, Fig. 3h).

We further hypothesized that the Flcn-RagC-ragulator-Tfe3 axis is acting in parallel with the TSC2-RagC-Tfe3 axis. We therefore sought to investigate the differential migration profile of Tfe3 in TSC2 KO ESCs compared to WT ESCs, as well as upon over expression of RagC/D GDP or GTP bound form in both genetic backgrounds. As we previously described, Tfe3 preferentially interacts with RagC/D-GDP bound form. Nevertheless, we could not observe any difference in the migration profile of Tfe3 between the immunoprecipitates from WT or TSC2 KO ESCs.

Even though Tfe3 is nuclear in TSC2 KO ESCs, the migration profile of Tfe3 in the input fraction is unchanged compared to WT ESCs (Fig. 16). Upon overexpression of Rag C/D GDP bound form, we observed the same migration profile as for WT ESCs expressing the same recombinant protein.

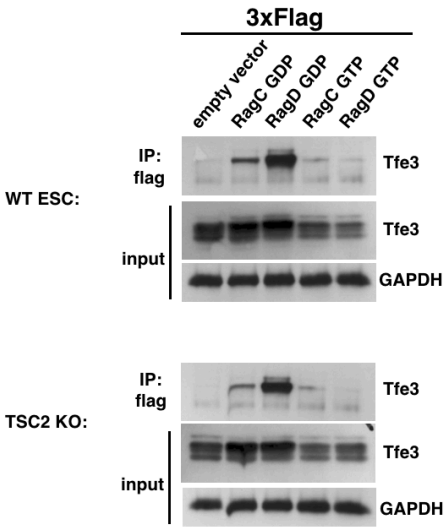


Figure 16: Tfe3 binds RagC/D GDP in WT and Tsc2 KO ESCs

Western blot analysis of Tfe3 immunoprecipitates from WT ESCs (upper part) and Tsc2 KO ESCs (lower part). In both genetic background, Tfe3 migration pattern display 2 bands on the WB; a fast migrating band and a higher band (empty vector line). Over-expression of RagC/D GTPases in GDP bound form induces Tfe3 binding in WT and TSC2 KO ESCs.

6 General discussion and conclusions

Cell fate transition, or how a cell type A undergoes a series of changes to become a cell type B, has been one of the greatest interest in developmental biology over the last decade. A perfect model suited to address cell fate transition related questions are embryonic stem cells (ESCs) since they are pluripotent and can differentiate toward all the lineages as *in vitro* differentiation recapitulates *in vivo* developmental progression. Previous work has identified novel genes required for exit from the ESC state during differentiation (Betschinger et al., 2013), amongst others identifying the tumor suppressor Folliculin (Flcn), acting upstream of the transcription factor Tfe3. Nevertheless, the regulatory framework in which these proteins are acting in ESCs remains unclear.

6.1 A CRISPR/Cas9 screen to identify novel regulator genes acting in the Flcn-Tfe3 pathway

To identify novel regulator genes acting in the Flcn-Tfe3 pathway and controlling exit from pluripotency, we screened for ESC self-renewal and maintenance of the pluripotency transcription factor Rex1 in differentiation conditions using a genome-wide CRISPR/Cas9 gRNA library. The specific goal of this screen was to identify genes whose lack block ESCs differentiation during long-term exposure to differentiation stimuli. Although the biological question addressed here is rather similar to previous work conducted using siRNA, the latter had limitations since ESCs were transfected only transiently. In the present study, the CRISPR/Cas9 screen offers the advantage to screen for genes allowing long-term maintenance of pluripotency. The screen had successfully recovered the tumor suppressor folliculin, which is associated with the Birt-Hogg-Dube (BHD) syndrome and required for ESC exit from pluripotency upstream of the cytoplasmic inactivation of the basic helix-loop-helix (bHLH) transcription factor Tfe3 (Betschinger et al., 2013); (Hong et al., 2010). Of the obtained hits, 15 genes whose absence lead to long-term resistance to differentiation were validated in a secondary screen that also tested for ectopic nuclear localization of Tfe3 in the absence of each of the 15

genes. This has identified Flcn, Lamtor1,2,4 and 5, Tsc1 and 2, and also recovered gain-of function alleles of Tfe3 itself. The small GTPase RagC belongs to the validated hit of the genetic screen but we did not observe obvious changes of Tfe3 subcellular localization upon RagC loss of function. However, neither RagA or B nor the RagC homologue RagD were identified in the screen. Why the screen did not identify RagD although redundant functions with RagC remains unclear.

6.2 Lamtor1 and Flcn knock-out ESCs are pluripotent and impaired for differentiation

Flcn and Lamtor1 null embryos show embryonic development failure around the gastrulation stage, consistent with being important for ESC differentiation *in vitro* (Baba et al., 2006); (Nada et al., 2009). However, their role in ESCs remains poorly understood. To get insights why Flcn and Lamtor1 are be so crucial for ESC differentiation, we generated the corresponding null ESC lines in an isogenic background using gRNA targeting. Different genomic regions of each gene were targeted and two independent knock-out (KO) cell lines per gene were generated. Phenotypic characterization of the corresponding clones showed that Flcn and Lamtor1 loss-of-function (LOF) induced a complete differentiation block in ESCs upon 2i removal as well as in the EpiLC inductive differentiation regime (Fgf, KSR and ActivinA). We therefore concluded that Flcn and Lamtor1 KO ESCs are not able to exit the pluripotent cell state which is consistent with the developmental arrest observed *in vivo* (see manuscript for details). To further study the role of Flcn and Lamtor1 during early development progression, additional *in vivo* experiments will be required.

6.3 The tumor suppressor Flcn is required for ESC differentiation upstream of RagC/D GTPases

Flcn is a cytoplasmic protein that can be recruited to the lysosome upon amino-acid starvation (see 2.3.2 and 2.3.3). The literature suggests that Flcn meets and

activates RagC/D through its GTPase Activating Protein (GAP) activity at the lysosome (Tsun et al., 2013); (Petit et al., 2013). In this work, we have shown that lysosomal RagC/D GDP is essential for ESC differentiation and the genetic manipulations suggest that it is mediated through the GAP activity of Flcn.

In comparison to Lamtor1 KO ESCs that can escape self-renewal by expression of a RagC transgene targeted to the lysosome (LTS-RagC), Flcn KO ESCs failed to do so. It suggests that the differentiation defect observed here is a direct consequence of the absence of RagC/D GDP bound form, the downstream product of Flcn. Indeed, the differentiation defect of Flcn KO ESCs could be bypassed by the expression of a RagC transgene locked in GDP bound conformation, therefore demonstrating that the GAP activity of Flcn is crucial for this process. Interestingly, over expression of RagC GDP leads to lysosomal enrichment of Tfe3. Surprisingly, LTS-RagC-GTP dominantly induced nuclear Tfe3 localization and Tfe3-dependent differentiation impairment in wildtype ESCs. Taken together our results suggest that the lysosomal recruitment of RagC serves as a substrate for Flcn GAP activity to drive cytoplasmic sequestration of Tfe3 and differentiation. This can be summarized on the following cartoon (Fig. 17).

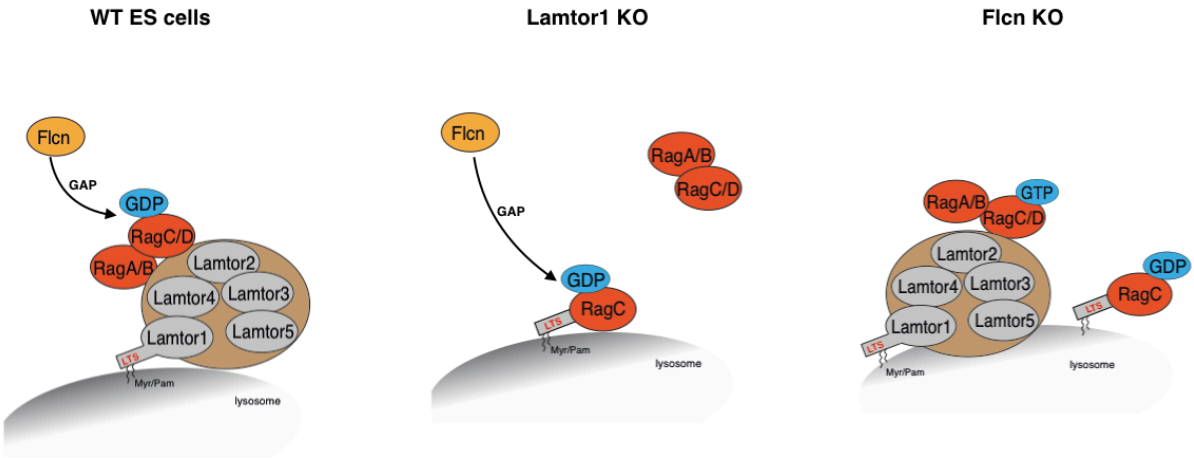


Figure 17: Lysosomal recruitment and nucleotide loading state of RagC/D GTPase downstream of Flcn are crucial for ESC differentiation.

In WT ESC, the ragulator is localized at the lysosomal compartment and allows for Flcn GAP activity. In Lamtor1 KO ESC, absence of a functional ragulator leads to cytoplasmic RagC/D GTPases and impairs ESC commitment upon exposure to differentiation stimuli. Targeted RagC to the lysosome in this genetic background rescues differentiation impairment, presumably because Flcn can act as a GAP. In Flcn KO ESCs, the ragulator is functional but RagC/D is locked

in GTP bound form. Targeted RagC in GDP bound form to the lysosome rescues differentiation block of Flcn KO ESC.

In the context of amino-acid sensing, the role of Flcn is not yet well understood and it might act in conveying a metabolic related information from the extracellular environment into the intracellular milieu. However, it remains unclear whether Flcn itself acts as a metabolic sensor in the amino-acid sensing machinery. As shown in other cell types, Flcn and its binding partners Fnip1 and 2 control the subcellular distribution of the transcription factor Tfeb. Flcn deficient ESCs are impaired for differentiation and exhibit nuclear enrichment of Tfe3. Fnip1 interacts with the C-terminal part of Flcn (Baba et al., 2006) and as Fnip1 was recovered by immunoprecipitation of endogenous flag-tagged Flcn in ESC, it indicates that the Flag peptide added in C-terminal position of the Flcn genetic locus by TALEN genome editing does not perturb the interaction with Fnip1. Further mass spectrometry analysis of flag-tagged Flcn immunoprecipitates in ESCs as well as after two days of differentiation revealed binding of Flcn to Fnip1 and Fnip2. The interaction with Fnip1 has been reported to promote interaction with the nutrient sensor AMP-activated Protein Kinase (AMPK) that responds to low nutrient and energy availability leading to the inhibition of mTORC1 (Baba et al., 2006). In ESCs, we did not identify any interaction between Flcn and AMPK.

Consistent with a previous study, we have found in ESCs that Gabarap interacts with Flcn and consequently tested if autophagy pathway is involved in the control of ESC differentiation. Knock down of additional genes involved in the autophagy pathway did not impair ESC differentiation (Dunlop et al., 2014).

It was shown in human cells that Lamtor3 interacts with Flcn/Fnip2 complex (Rebsamen et al., 2015). However, we did not identify any interaction between ragulator sub-unit and Flcn.

Differential interactome analysis of ESCs compared to differentiated cells did not show any enrichment of state specific interactors suggesting that Flcn interactome remains unchanged upon differentiation. To investigate the role of Flcn in ESCs and specifically during the exit from pluripotency, we overexpressed different RagC transgenes (WT, GDP or GTP bound form) targeted to the lysosome. Here we aimed to observe any changes in Flcn subcellular localization that would suggest a

preferential binding toward a specific nucleotide loading state of RagC. Using our experimental set up, we were not able to determine any significant changes. Moreover, Flcn subcellular localization remains unchanged upon initiation of differentiation. In fact, mass spectrometric analysis of endogenous flag-tagged Flcn did not identify any post-translational modification that could explain a dynamic activity during differentiation.

Additionally, immunoprecipitation of flag-tagged RagC GDP or GTP bound form coupled to mass spectrometry analysis revealed that the transcription factor Tfe3 preferentially binds RagC GDP but not Flcn. Hence, the role of Flcn in ESC during differentiation remains not fully understood.

6.4 Towards an alternate lysosomal signaling pathway

Since the lysosomal functions *per se* have never been reported to act in the control of the exit from pluripotency, we investigated if perturbation of lysosomal activity could affect ESC differentiation. We have shown that perturbation of lysosomal functions impairs ESC differentiation and this can be rescued by lysosomal targeted expression of RagC GDP. Interestingly, Tfe3 LOF ESCs can differentiate in the presence of drugs perturbing lysosomal functions (see manuscript, Fig. 3e – 3f). By opposition to lysosomal functions that are required for cellular survival, we have assessed in the present work the role of the lysosomal functions in the precise framework that is the short time window of the exit from self-renewal in ESCs. Whether lysosomal perturbations will affect the following steps of early development remain to beunderstood.

The amino-acid sensing mechanism is orchestrated at the lysosomal compartment and lysosomal functions are required for this process as it has been shown that the lysosomal V-ATPase acts as a GEF for Rag GTPases RagA/B in response to amino-acid availability within the lysosomal lumen. This is important for the lysosomal recruitment and activation of mTORC1 in response to amino-acid stimili. In ESCs, this pathway might not be strictly identically regulated and could act as a metabolic checkpoint allowing ESCs to undergo differentiation, or not, since ESCs differentiation is a very demanding energetic process and ESCs have to make sure

that nutrients are available for a proper developmental progression.

Consistent with the ragulator acting in ESC differentiation, Lamtor3 protein, also called MP1, has been identified in a genetic screen performed by using siRNA in ESCs in a setup addressing another biological question (Westerman et al., 2011). In this study, the authors identified Lamtor3 as a factor important for differentiation. At that time, the ragulator was not yet characterized and Lamtor proteins could not have been replaced in the lysosomal amino-acid sensing machinery. This study already provided evidence that the ragulator acts in the control of ESC differentiation. In addition, the authors showed that downregulation of Lamtor3 in ESC affects the phosphorylation levels of Erk and therefore argued that the differentiation inhibitory effect of knockdown of Lamtor3 is not likely to be a result of total ablation of Erk signaling but rather a result of modulation of the Erk signaling network that leads to ESC differentiation. These interpretations led the authors to hypothesize that the mechanism might occur through endosomal signaling since Mp1 localizes to late endosomes in ESCs. Interestingly and in line with this hypothesis, Hgs, a component of the ESCRT-0 complex that initiates sorting of ubiquitinylated receptors into the multivesicular body pathway, has been identified in a loss of function genetic screen as a gene required for ESC differentiation (Betschinger et al., 2013). However, neither knockdown of the facultative ESCRT-0 Hgs interactor Stam nor selected ESCRT-I, -II or –III subunits impair ESC differentiation. Although we did not find Hgs in the CRISPR/Cas9 screen, this should be kept in mind as an additional finding supporting the idea that a lysosomal pathway contributes to exit from pluripotency.

We have shown that perturbation of lysosomal activity induced nuclear Tfe3 enrichment and RagC-GDP- and Tfe3-dependent differentiation defects, suggesting that lysosomal activity controls ESC differentiation.

Together, our data repurposed a conserved mechanism of lysosomal amino-acid sensing that plays a fundamental role during early developmental stage.

6.5 Metabolic control of ESC pluripotency

Metabolic profiles vary from one cell type to another and a current notion in the literature indicates that the metabolic switch occurring during cell fate transition is probably the cause and not the consequence of this process (Shyh-Chang and Ng, 2017). As a metabolic shift has been described from oxidative phosphorylation toward anaerobic glycolysis from the pre- to post-implantation epiblast ESCs, a metabolic shift in the opposite direction has also been observed during reprogramming (Panopoulos et al., 2011). We have identified the lysosome acting in the regulation of ESCs differentiation and in addition to previous findings, we describe this degradative organelle as an integrative platform that regulates and secures developmental progression of ESCs *in vitro*. The lysosomal signaling pathway is conceptually interesting since this organelle has key functions in the regulation of cellular metabolic reactions and is highly involved in the control of cellular growth.

The tumor suppressor Flcn is required for ESC differentiation. In the amino-acid sensing pathway, Flcn is a GAP for RagC/D and its recruitment to the lysosome occurs during starvation conditions. We hypothesized that Flcn is involved in a pathway that links metabolism and pluripotency.

This hypothesis suggests that Flcn acts as a sensor in this pathway and cytoplasmic localization of Flcn is secured by the presence of a signaling molecule, which, upon starvation, should not be longer present in the cytoplasm. Although such molecule has not been identified yet, we further aimed to understand if amino-acids have a role in the control of ESC differentiation. We observed that removal of essential or non-essential amino-acids impairs exit from self-renewal in ESCs. However, we did not specifically investigate if amino-acid starvation impairs cellular proliferation. If so, it would suggest that starvation induces a proliferation defect, therefore rendering ESCs insensitive to differentiation stimuli. While we were aiming to differentiate ESCs under starvation stress condition, an interesting observation occurred when only valine residue was omitted from the medium. Exclusive omission of valine residue from the medium led to a complete inhibition of proliferation associated with an important cell death (data not shown). This observation has already been reported for both human and murine hematopoietic

stem cells where valine has been reported to be essential for cellular growth (Taya et al., 2016). Removal of essential amino-acids induced a starvation stress response accompanied by nuclear Tfe3 enrichment and resistance to differentiation in ESCs. Surprisingly, we could observe a mild rescue upon overexpression of LTS-RagC GDP during starvation (see manuscript, extend data Fig. 4h). To address the question if this rescue could be due to a possible effect of RagC-GDP over expression that induces a priming-like effect in ESCs, further investigations must be undertaken, first by assessing the expression levels of pluripotency associated genes. Nevertheless, the reversion of the phenotype was not recapitulated in Tfe3 KO ESCs. It should be kept in mind that overexpression of any nucleotide bound form of RagC/D GTPase does not regulate mTORC1 activity in ESCs. Other studies have also reported that expression of the dominant negative forms of RagGTPases (RagC/D-GTP) abolishes the amino-acid sensing machinery and cells expressing RagC/D GTP remain insensitive to amino-acid stimulation following a starvation period and display Tfeb, a family member of Tfe3, exclusively in the nucleus (Settembre et al., 2012). Interestingly, overexpression of LTS-RagC GTP in ESCs led to accumulation of Tfe3 in the nuclear compartment and impaired ESCs differentiation, downstream of Tfe3.

Finding a nutrient sensor components of the lysosomal pathway involving Flcn should help to understand in more details the metabolic regulation of ESC self-renewal upstream of Tfe3. Our work illustrates the general notion present in the literature about the metabolic control of cell fate transition in ESCs. Indeed, while we were aiming to identify if amino-acid starvation impairs ESC differentiation, we could not identify a unique amino-acid whose omission impaired differentiation in a Tfe3 dependent manner. Taken together these observations suggest a Flcn-RagC/D-ragulator-Tfe3 dependent mechanism driving ESC differentiation.

6.6 mTOR activity is not required for ESCs differentiation

Studies have shown that deregulation of mTOR activity often leads to dramatic consequences and mTOR activation is a common feature of diseases as seen in many

different types of cancer or neurological diseases. Therefore, understanding the cause of these deregulations remains an outstanding challenge.

While we were aiming to identify genes whose absence impairs ESCs differentiation, we identified the tumor suppressor Flcn as well as the Lamtor proteins that form the ragulator protein complex. Together, Flcn and the ragulator activate mTORC1 in an amino-acid dependent manner in somatic cells. Nevertheless, we have shown in ESCs that mTORC1 activity is independent of Flcn and Lamtor1. Additionally, chemical or genetic inhibition of mTORC1 in ESCs did not induce differentiation defect nor nuclear Tfe3 enrichment suggesting that mTORC1 activity is dispensable for ESCs differentiation. Surprisingly, removal of the two inhibitors sustaining ESC pluripotency *in vitro* is accompanied by an increase of mTORC1 activity in Flcn and Lamtor1 KO ESCs although these two cell lines do not differentiate. If this increased mTORC1 activity occurs in response to MAPK/Erk signaling pathway activation remains to be determined.

Published transcriptomics data from mTOR-inhibited ESCs do not correlate with transcriptional alterations in Lamtor1 and Flcn KO ESCs (Bulut-Karslioglu et al., 2016). Interestingly, hyper-activation of mTORC1 upon loss of TSC2 phenocopy Flcn loss of function phenotypes in ESCs and feature nuclear Tfe3 localization. This phenotype can be reverted by rapamycin treatment as well as the over expression of RagC-GDP or Tfe3 LOF.

During nutrient starvation in somatic cells, mTORC1 activity has been reported to be downregulated. We could also confirm this finding in ESCs. The overexpression of RagC GDP bound form has been described to activate mTORC1 activity in somatic cells. However, it appears not to be the case in ESCs. This indicates a tuned regulation of mTORC1 pathway that might differ from one cell type to another, suggesting that minor alterations of this pathway lead to phenotypical variations. Although overexpression of RagC GDP or GTP doesn't perturb mTORC1 activity, Tfe3 subcellular localization does change accordingly. Forced lysosomal expression of Rag GTP redirect Tfe3 to the nucleus and strongly enhanced the maintenance of pluripotency.

The Tfe/MiTF transcription factor family has been linked to metabolic response in different cell types. In ESCs, the nucleocytoplasmic distribution of Tfe3 coincide

with what has been reported for Tfeb. Tfeb has been shown to be a main player in starvation where upon starvation, it translocates to the nucleus (Settembre et al., 2012); (Settembre et al., 2013). In addition, nutrient starvation as well as lysosomal perturbations induced nuclear translocation of Tfe3 in ESCs as it was shown to be the case in other somatic cell type.

Taken together, the work presented here advances our knowledge about a new route for the control of exit from pluripotency in ESC that takes place at the lysosomal compartment and provides insight into the metabolic regulation of this process.

7 References

Baba, M., Furihata, M., Hong, S.B., Tessarollo, L., Haines, D.C., Southon, E., Patel, V., Igarashi, P., Alvord, W.G., Leighty, R., et al. (2008). Kidney-Targeted Birt-Hogg-Dube Gene Inactivation in a Mouse Model: Erk1/2 and Akt-mTOR Activation, Cell Hyperproliferation, and Polycystic Kidneys. JNCI Journal of the National Cancer Institute *100*, 140–154.

Baba, M., Hong, S.-B., Sharma, N., Warren, M.B., Nickerson, M.L., Iwamatsu, A., Esposito, D., Gillette, W.K., Hopkins, R.F., Hartley, J.L., et al. (2006). Folliculin encoded by the BHD gene interacts with a binding protein, FNIP1, and AMPK, and is involved in AMPK and mTOR signaling. Proc Natl Acad Sci USA *103*, 15552–15557.

Bar-Peled, L., Schweitzer, L.D., Zoncu, R., and Sabatini, D.M. (2012). Ragulator Is a GEF for the Rag GTPases that Signal Amino Acid Levels to mTORC1. Cell *150*, 1196–1208.

Baral, V., Chaoui, A., Watanabe, Y., Goossens, M., Attie-Bitach, T., Marlin, S., Pingault, V., and Bondurand, N. (2012). Screening of MITF and SOX10 Regulatory Regions in Waardenburg Syndrome Type 2. PLoS ONE *7*, e41927–e41928.

Barbehenn, E.K., Wales, R.G., and Lowry, O.H. (1978). Measurement of metabolites in single preimplantation embryos; a new means to study metabolic control in early embryos. J Embryol Exp Morphol *43*, 29–46.

Betschinger, J., Nichols, J., Dietmann, S., Corrin, P.D., Paddison, P.J., and Smith, A. (2013). Exit from Pluripotency Is Gated by Intracellular Redistribution of the bHLH Transcription Factor Tfe3. Cell *153*, 335–347.

Brinster, R.L., and Troike, D.E. (1979). Requirements for blastocyst development in vitro. J. Anim. Sci. *49 Suppl 2*, 26–34.

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of Enhancer Patterns in Transition from Naive to Primed Pluripotency. Stem Cell *14*, 838–853.

Bulut-Karslioglu, A., Biechele, S., Jin, H., Macrae, T.A., Hejna, M., Gertsenstein, M., Song, J.S., and Ramalho-Santos, M. (2016). Inhibition of mTOR induces a paused pluripotent state. Nature *540*, 119–123.

Carbognin, E., Betto, R.M., Soriano, M.E., Smith, A.G., and Martello, G. (2016). Stat3 promotes mitochondrial transcription and oxidative respiration during maintenance and induction of naive pluripotency. Embo J *35*, 618–634.

Carey, B.W., Finley, L.W.S., Cross, J.R., Allis, C.D., and Thompson, C.B. (2014). Intracellular α-ketoglutarate maintains the pluripotency of embryonic stem cells. Nature 1–12.

Cash, T.P., Gruber, J.J., Hartman, T.R., Henske, E.P., and Simon, M.C. (2011). Loss of the Birt-Hogg-Dube tumor suppressor results in apoptotic resistance due to aberrant TGFbeta;-mediated transcription. Oncogene *30*, 2534–2546.

Cho, Y.-H., Han, K.-M., Kim, D., Lee, J., Lee, S.-H., Choi, K.-W., Kim, J., and Han, Y.-M. (2014).

Autophagy Regulates Homeostasis of Pluripotency-Associated Proteins in hESCs. Stem Cells *32*, 424–435.

Davidson, K.C., Mason, E.A., and Pera, M.F. (2015). The pluripotent state in mouse and human. Development *142*, 3090–3099.

Demetriades, C., Plescher, M., and Teleman, A.A. (1AD). Lysosomal recruitment of TSC2 is a universal response to cellular stress. Nature Communications 1–14.

Du, H., Sheriff, S., Bezerra, J., Leonova, T., and Grabowski, G.A. (1998). Molecular and Enzymatic Analyses of Lysosomal Acid Lipase in Cholesteryl Ester Storage Disease. Molecular Genetics and Metabolism *64*, 126–134.

Dunlop, E.A., Seifan, S., Claessens, T., Behrends, C., Kamps, M.A., Rozycka, E., Kemp, A.J., Nookala, R.K., Blenis, J., Coull, B.J., et al. (2014). FLCN, a novel autophagy component, interacts with GABARAP and is regulated by ULK1 phosphorylation. Autophagy *10*, 1749–1760.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Folmes, C.D.L., Nelson, T.J., Martinez-Fernandez, A., Arrell, D.K., Lindor, J.Z., Dzeja, P.P., Ikeda, Y., Perez-Terzic, C., and Terzic, A. (2011). Somatic Oxidative Bioenergetics Transitions into Pluripotency-Dependent Glycolysis to Facilitate Nuclear Reprogramming. Cell Metabolism *14*, 264–271.

Hackett, J.A., and Surani, M.A. (2014). Regulatory Principles of Pluripotency: From the Ground State Up. Stem Cell *15*, 416–430.

Hasumi, H., Baba, M., Hasumi, Y., Huang, Y., Oh, H., Hughes, R.M., Klein, M.E., Takikita, S., Nagashima, K., Schmidt, L.S., et al. (2012). Regulation of Mitochondrial Oxidative Metabolism by Tumor Suppressor FLCN. JNCI Journal of the National Cancer Institute *104*, 1750–1764.

Hasumi, Y., Baba, M., Ajima, R., Hasumi, H., Valera, V.A., Klein, M.E., Haines, D.C., Merino, M.J., Hong, S.B., Yamaguchi, T.P., et al. (2009). Homozygous loss of BHD causes early embryonic lethality and kidney tumor development with activation of mTORC1 and mTORC2. Proc Natl Acad Sci USA *106*, 18722–18727.

Hong, S.-B., Oh, H., Valera, V.A., Baba, M., Schmidt, L.S., and Linehan, W.M. (2010). Inactivation of the FLCN Tumor Suppressor Gene Induces TFE3 Transcriptional Activity by Increasing Its Nuclear Localization. PLoS ONE *5*, e15793–12.

Hudon, V., Sabourin, S., Dydensborg, A.B., Kottis, V., Ghazi, A., Paquet, M., Crosby, K., Pomerleau, V., Uetani, N., and Pause, A. (2010). Renal tumour suppressor function of the Birt-Hogg-Dubé syndrome gene product folliculin. Journal of Medical Genetics *47*, 182–189.

Ito, K., and Suda, T. (2014). Metabolic requirements for the maintenance of self-renewing stem cells. Nat Rev Mol Cell Biol *15*, 243–256.

Kauffman, E.C., Ricketts, C.J., Rais-Bahrami, S., Yang, Y., Merino, M.J., Bottaro, D.P., Srinivasan, R., and Linehan, W.M. (2014). Molecular genetics and cellular features of TFE3 and TFEB fusion kidney cancers. Nature Publishing Group *11*, 465–475.

Laplante, M., and Sabatini, D.M. (2009). mTOR signaling at a glance. Journal of Cell Science *122*, 3589–3594.

Leeb, M., Dietmann, S., Paramor, M., Niwa, H., and Smith, A. (2014). Genetic Exploration of the Exit from Self-Renewal Using Haploid Embryonic Stem Cells. Stem Cell *14*, 385–393.

Levy, C., Khaled, M., and Fisher, D.E. (2006). MITF: master regulator of melanocyte development and melanoma oncogene. Trends in Molecular Medicine *12*, 406–414.

Linehan, W.M., Srinivasan, R., and Schmidt, L.S. (2010). The genetic basis of kidney cancer: a metabolic disease. Nature Publishing Group *7*, 277–285.

Marsboom, G., Zhang, G.-F., Pohl-Avila, N., Zhang, Y., Yuan, Y., Kang, H., Hao, B., Brunengraber, H., Malik, A.B., and Rehman, J. (2016). Glutamine Metabolism Regulates the Pluripotency Transcription Factor OCT4. CellReports 1–11.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci USA *78*, 7634–7638.

Martina, J.A., Diab, H.I., Lishu, L., Jeong-A, L., Patange, S., Raben, N., and Puertollano, R. (2014). The Nutrient-Responsive Transcription Factor TFE3 Promotes Autophagy, Lysosomal Biogenesis, and Clearance of Cellular Debris. Science Signaling *7*, ra9–ra9.

Martina, J.A., and Puertollano, R. (2013). Rag GTPases mediate amino acid–dependent recruitment of TFEB and MITF to lysosomes. J Cell Biol *200*, 475–491.

Medina, D.L., Di Paola, S., Peluso, I., Armani, A., De Stefani, D., Venditti, R., Montefusco, S., Scotto-Rosato, A., Prezioso, C., Forrester, A., et al. (2015). Lysosomal calcium signalling regulates autophagy through calcineurin and TFEB. Nat Cell Biol *17*, 288–299.

Mindell, J.A. (2012). Lysosomal Acidification Mechanisms. Annu. Rev. Physiol. *74*, 69–86.

Mizushima, N., and Levine, B. (2010). Autophagy in mammalian development and differentiation. Nature Publishing Group *12*, 823–830.

Moussaieff, A., Rouleau, M., Kitsberg, D., Cohen, M., Levy, G., Barasch, D., Nemirovski, A., Shen-Orr, S., Laevsky, I., Amit, M., et al. (2015). Glycolysis-Mediated Changes in Acetyl-CoA and Histone Acetylation Control the Early Differentiation of Embryonic Stem Cells. Cell Metabolism *21*, 392–402.

Nada, S., Hondo, A., Kasai, A., Koike, M., Saito, K., Uchiyama, Y., and Okada, M. (2009). The novel lipid raft adaptor p18 controls endosome dynamics by anchoring the MEK–ERK pathway to late endosomes. Embo J *28*, 477–489.

Nada, S., Mori, S., Takahashi, Y., and Okada, M. (2014). p18/LAMTOR1: A Late Endosome/

Lysosome-Specific Anchor Protein for the mTORC1/MAPK Signaling Pathway (Elsevier Inc.).

Panopoulos, A.D., Yanes, O., Ruiz, S., Kida, Y.S., Diep, D., Tautenhahn, R., Herrerías, A., Batchelder, E.M., Plongthongkum, N., Lutz, M., et al. (2011). The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming. Nature Publishing Group *22*, 168–177.

Petit, C.S., Roczniak-Ferguson, A., and Ferguson, S.M. (2013). Recruitment of folliculin to lysosomes supports the amino acid–dependent activation of Rag GTPases. J Cell Biol *202*, 1107–1122.

Péli-Gulli, M.-P., Sardu, A., Panchaud, N., Raucci, S., and De Virgilio, C. (2015). Amino Acids Stimulate TORC1 through Lst4-Lst7, a GTPase-Activating Protein Complex for the Rag Family GTPase Gtr2. CellReports *13*, 1–7.

Rebsamen, M., Pochini, L., Stasyk, T., de Araújo, M.E.G., Galluccio, M., Kandasamy, R.K., Snijder, B., Fauster, A., Rudashevskaya, E.L., Bruckner, M., et al. (2015). SLC38A9 is a component of the lysosomal amino acid sensing machinery that controls mTORC1. Nature 1–18.

Rehli, M., Lichanska, A., Cassady, A.I., Ostrowski, M.C., and Hume, D.A. (1999). TFEC is a macrophage-restricted member of the microphthalmia-TFE subfamily of basic helix-loop-helix leucine zipper transcription factors. J. Immunol. *162*, 1559–1565.

Sampath, P., Pritchard, D.K., Pabon, L., Reinecke, H., Schwartz, S.M., Morris, D.R., and Murry, C.E. (2008). A Hierarchical Network Controls Protein Translation during Murine Embryonic Stem Cell Self-Renewal and Differentiation. Cell Stem Cell *2*, 448–460.

Sancak, Y., Peterson, T.R., Shaul, Y.D., Lindquist, R.A., Thoreen, C.C., Bar-Peled, L., and Sabatini, D.M. (2008). The Rag GTPases Bind Raptor and Mediate Amino Acid Signaling to mTORC1. Science *320*, 1496–1501.

Sancak, Y., and Sabatini, D.M. (2009). Rag proteins regulate amino-acid-induced mTORC1 signalling. Biochm. Soc. Trans. *37*, 289–290.

Sancak, Y., Bar-Peled, L., Zoncu, R., Markhard, A.L., Nada, S., and Sabatini, D.M. (2010). Ragulator-Rag Complex Targets mTORC1 to the Lysosomal Surface and Is Necessary for Its Activation by Amino Acids. Cell *141*, 290–303.

Sardiello, M., Palmieri, M., di Ronza, A., Medina, D.L., Valenza, M., Gennarino, V.A., Di Malta, C., Donaudy, F., Embrione, V., Polishchuk, R.S., et al. (2009). A Gene Network Regulating Lysosomal Biogenesis and Function. Science 1–6.

Saxton, R.A., and Sabatini, D.M. (2017). mTOR Signaling in Growth, Metabolism, and Disease. Cell *168*, 960–976.

Settembre, C., Fraldi, A., Medina, D.L., and Ballabio, A. (2013). Signals from the lysosome: a control centre for cellular clearance and energy metabolism. Nature Publishing Group *14*, 283–296.

Settembre, C., Zoncu, R., Medina, D.L., Vetrini, F., Erdin, S., Erdin, S., Huynh, T., Ferron, M., Karsenty, G., Vellard, M.C., et al. (2012). A lysosome-to-nucleus signalling mechanism senses and regulates the lysosome via mTOR and TFEB. Embo J *31*, 1095–1108.

Shen, K., Choe, A., and Sabatini, D.M. (2017). Intersubunit Crosstalk in the Rag GTPase Heterodimer Enables mTORC1 to Respond Rapidly to Amino Acid Availability. Molecular Cell 1–23.

Shyh-Chang, N., and Ng, H.-H. (2017). The metabolic programming of stem cells. Genes Dev. *31*, 336–346.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature *336*, 688–690.

Sperber, H., Mathieu, J., Wang, Y., Ferreccio, A., Hesson, J., Xu, Z., Fischer, K.A., Devi, A., Detraux, D., Gu, H., et al. (2015). The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition. Nat Cell Biol *17*, 1523–1535.

Steingrimsson, E., Tessarollo, L., Pathak, B., Hou, L., Arnheiter, H., Copeland, N.G., and Jenkins, N.A. (2002). Mitf and Tfe3, two members of the Mitf-Tfe family of bHLH-Zip transcription factors, have important but functionally redundant roles in osteoclast development. Proc Natl Acad Sci USA *99*, 4477–4482.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676.

Taya, Y., Ota, Y., Wilkinson, A.C., Kanazawa, A., Watarai, H., Kasai, M., Nakauchi, H., and Yamazaki, S. (2016). Depleting dietary valine permits nonmyeloablative mouse hematopoietic stem cell transplantation. Science *354*, 1152–1155.

Teis, D., Taub, N., Kurzbauer, R., Hilber, D., de Araujo, M.E., Erlacher, M., Offterdinger, M., Villunger, A., Geley, S., Bohn, G., et al. (2006). p14-MP1-MEK1 signaling regulates endosomal traffic and cellular proliferation during tissue homeostasis. J Cell Biol *175*, 861–868.

Tsun, Z.-Y., Bar-Peled, L., Chantranupong, L., Zoncu, R., Wang, T., Kim, C., Spooner, E., and Sabatini, D.M. (2013). The Folliculin Tumor Suppressor Is a GAP for the RagC/D GTPases That Signal Amino Acid Levels to mTORC1. Molecular Cell *52*, 495–505.

Van Blerkom, J. (2009). Mitochondria in early mammalian development. Seminars in Cell & Developmental Biology *20*, 354–364.

Vander Heiden, M.G., Cantley, L.C., and Thompson, C.B. (2009). Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. Science *324*, 1029–1033.

Wada, S., Neinast, M., Jang, C., Ibrahim, Y.H., Lee, G., Babu, A., Li, J., Hoshino, A., Rowe, G.C., Rhee, J., et al. (2016). The tumor suppressor FLCN mediates an alternate mTOR pathway

to regulate browning of adipose tissue. Genes Dev. *30*, 2551–2564.

Watanabe, A., Takeda, K., Ploplis, B., and Tachibana, M. (1998). Epistatic relationship between Waardenburg Syndrome genes MITF and PAX3. Nature Genetics *18*, 283–286.

Westerman, B.A., Braat, A.K., Taub, N., Potman, M., Vissers, J.H.A., Blom, M., Verhoeven, E., Stoop, H., Gillis, A., Velds, A., et al. (2011). A genome-wide RNAi screen in mouse embryonic stem cells identifies Mp1 as a key mediator of differentiation. J Exp Med *208*, 2675–2689.

Wolfson, R.L., Chantranupong, L., Wyant, G.A., Gu, X., Orozco, J.M., Shen, K., Condon, K.J., Petri, S., Kedir, J., Scaria, S.M., et al. (2017). KICSTOR recruits GATOR1 to the lysosome and is necessary for nutrients to regulate mTORC1. Nature *543*, 438–442.

Yang, S.-H., Kalkan, T., Morrisroe, C., Smith, A., and Sharrocks, A.D. (2012). A Genome-Wide RNAi Screen Reveals MAP Kinase Phosphatases as Key ERK Pathway Regulators during Embryonic Stem Cell Differentiation. PLoS Genet *8*, e1003112–e1003115.

Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519–523.

Yoneshima, E., Okamoto, K., Sakai, E., Nishishita, K., Yoshida, N., and Tsukuba, T. (2015). The Transcription Factor EB (TFEB) Regulates Osteoblast Differentiation Through ATF4/CHOP-Dependent Pathway. J. Cell. Physiol. *231*, 1321–1333.

Yue, Z., Jin, S., Yang, C., Levine, A.J., and Heintz, N. (2003). Beclin 1, an autophagy gene essential for early embryonic development, is a haploinsufficient tumor suppressor. Proc Natl Acad Sci USA *100*, 15077–15082.

Zhang, J., Ratanasirintrawoot, S., Chandrasekaran, S., Wu, Z., Ficarro, S.B., Yu, C., Ross, C.A., Cacchiarelli, D., Xia, Q., Seligson, M., et al. (2016). LIN28 Regulates Stem Cell Metabolism and Conversion to Primed Pluripotency. Stem Cell *19*, 66–80.

Zhou, W., Choi, M., Margineantu, D., Margaretha, L., Hesson, J., Cavanaugh, C., Blau, C.A., Horwitz, M.S., Hockenbery, D., Ware, C., et al. (2012). HIF1α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. Embo J *31*, 2103–2116.

Zhu, S., Li, W., Zhou, H., Wei, W., Ambasudhan, R., Lin, T., Kim, J., Zhang, K., and Ding, S. (2010). Reprogramming of Human Primary Somatic Cells by OCT4 and Chemical Compounds. Stem Cell *7*, 651–655.

Zoncu, R., Bar-Peled, L., Efeyan, A., Wang, S., Sancak, Y., and Sabatini, D.M. (2011). mTORC1 Senses Lysosomal Amino Acids Through an Inside-Out Mechanism That Requires the Vacuolar H+-ATPase. Science *334*, 678–683.

8 Curriculum vitae

EDUCATION

Friedrich Miescher Institute, Basel – Switzerland

07/2014 – present
International PhD Program
Thesis title: "Lysosomal signaling is required for embryonic stem cell differentiation and human development"

Claude Bernard University, Lyon – France

09/2011 – 07/2014
MSc in Genetic, Cellular Biology and Pathology
Thesis title: "Understanding the enzymatic function of aspartoacylase in lipid physiology and metabolic diseases using murine models

Claude Bernard University, Lyon – France

09/2009 – 06/2011
BSc in Biochemistry

RESEARCH EXPERIENCE

Master thesis – Max Planck Institute (MPI-IE), Freiburg – Germany

10/2013 – 06/2014
Laboratory of Dr. John Andrew Pospisilik
Subject: "Understanding the enzymatic function of aspartoacylase in lipid physiology and metabolic diseases using murine models"

Internship – Max Planck Institute (MPI-IE), Freiburg – Germany

07/2012 – 12/2012
Laboratory of Dr. John Andrew Pospisilik
Research field: Epigenetic control of metabolic diseases

Master internship – Netherland Cancer Institute (NKI), Amsterdam – The Netherlands

03/2012 – 05/2012
Laboratory of Dr. Anastassis Perrakis
Subject: "Interaction studies between Polo-like kinase proteins and peptides with pharmaceutical potential using biophysical techniques"

Bachelor Internship - Netherland Cancer Institute (NKI), Amsterdam – The Netherlands

03/2011 – 09/2011
Laboratory of Dr. Anastassis Perrakis
Subject: "Purification and structural study of four proteins (Cdc20, Mad2, BubR1 and Bub3) constituting the spindle assembly checkpoint"

Bachelor internship – Institut National de la Santé et de la Recherche Médicale (INSERM Unit 842 – Neuro-oncology and neuro-inflammation), Lyon – France

04/2010 – 06/2010
Laboratory of Prof. Dr. Jérome Honnorat
Subject: "Vascularization study of two different glioblastoma models by xenograft using a nude murine model"

BTS internship – Laboratory of virology and molecular biology Hospices Civils de Lyon, Lyon – France

09/2006 – 12/2007
Laboratory of Prof. Dr. Bruno Lina
Project: "Establishment of a new technique for Epstein Barr virus quantification by qPCR from patient samples (blood or cerebrospinal fluid)"

SCIENTIFIC PRESENTATIONS

-Talk Joint PhD retreat with the Gulbekian
27th – 31st May 2017, Lisbon, Portugal

-Talk Basel stem cell network
18th May 2017, Basel, Switzerland

-Talk FMI annual meeting
16th Septembre 2016, Basel, Switzerland