The role of Prrx1 in mouse development and vascular biology

PhD Program in Neuroscience

Instituto de Neurociencias – UMH-CSIC

Universidad Miguel Hernández

- 2021 -

Doctoral Thesis presented by

Francisco García Asencio

Thesis Director:

M. Angela Nieto Toledano

INSTITUTO DE NEUROCIENCIAS

Sant Joan d'Alacant, 15th of March 2021

Conventional doctoral thesis

To whom it may concern:

The doctoral thesis entitled "The role of *Prrx1* in mouse development and vascular biology" has been developed by myself, Francisco García Asencio. This thesis is presented in a conventional format. It is based on experimental studies undertaken at the Neuroscience Institute of Alicante during the PhD program in neuroscience of the Miguel Hernández University.

Yours sincerely,

FRANCISCO|GARCIA|ASENCIO

Firmado digitalmente por FRANCISCO|GARCIA|ASENCIO
Fecha: 2021.03.16 15:04:52 +01'00'

Francisco García Asencio

Sant Joan d'Alacant, 15th of March 2021

To whom it may concern,

The doctoral thesis entitled "The role of Prrx1 in mouse development and vascular biology" has been developed by myself, Francisco García Asencio. This thesis includes the following publication: of which I am the fifth author: Fazilaty, H., L. Rago, K. Kass Youssef, O. H. Ocaña, **F. Garcia-Asencio**, A. Arcas, J. Galceran, and M. A. Nieto. 2019. "A Gene Regulatory Network to Control EMT Programs in Development and Disease." *Nature Communications* 10(1):5115

I declare that the publication has not been used and will not be used in any other thesis in agreement with my thesis director Prof. M. Angela Nieto Toledano:

Yours sincerely,

FRANCISCO|GARCIA|ASENCIO
Firmado digitalmente por FRANCISCO|GARCIA|ASENCIO
Fecha: 2021.03.16 15:07:29 +01'00'

NIETO TOLEDANO M.ANGELA - DNI 50694461T
2021.03.16 14:40:01 +01'00'

D. Francisco García Asencio

Prof. M. Angela Nieto Toledano

A quien corresponda:

La Prof. M. Ángela Nieto Toledano, Profesora de Investigación del Consejo Superior de Investigadores Científicas,

AUTORIZA la presentación de la tesis Doctoral titulada “The role of *Prrx1* in mouse development and vascular biology”, Realizada por D. Francisco García Asencio (DNI 74384680N) bajo su dirección y supervisión en el Instituto de Neurociencias de Alicante, centro mixto CSIC-UMH, y que presenta para la obtención del grado de Doctor por la Universidad Miguel Hernández.

Y para que así conste, y a los efectos oportunos, firma el presente certificado en San Juan de Alicante, a 15 de marzo de 2021.

NIETO TOLEDANO
M.ANGELA - DNI
50694461T
2021.03.16 14:40:37 +01'00'

Fdo. M. Ángela Nieto Toledano

Sant Joan d'Alacant, 15 de marzo de 2021

Dña. Elvira De la Peña García, Profesora titular de la Universidad Miguel Hernández (UMH) y Coordinadora del programa de doctorado en Neurociencias del Instituto de Neurociencias de Alicante, centro mixto de la Universidad Miguel Hernández (UMH) y de la Agencia Estatal Consejo Superior de Investigaciones Científicas (CSIC).

CERTIFICA:

Que la Tesis Doctoral titulada "The role of *Prrx1* in mouse development and vascular biology" has sido realizada por D. Francisco García Asencio (DNI 74384680N), bajo la dirección de Prof. M. Ángela Nieto Toledano como directora, y doy mi conformidad para que sea presentada a la Comisión de Doctorado de la Universidad Miguel Hernández.

Y para que conste, a los efectos oportunos, firmo el presente certificado en San Juan de Alicante a 15 de marzo de 2021.

ELVIRA MARIA DE LA PEÑA GARCIA - NIF:21488194F

Firmado digitalmente por ELVIRA MARIA DE LA PEÑA GARCIA - NIF:21488194F
Fecha: 2021.03.16 08:52:56 +01'00'

Fdo. Elvira de la Peña García

Financiación/Subvención/Beca:

Este trabajo de Tesis Doctoral ha sido posible gracias a una beca predoctoral de Formación de Personal Investigador (FPI), financiada por el Ministerio de Economía y Competitividad (MINECO) con referencia BES-2015-073796, un contrato (2015-2019) de titulado superior en actividades técnicas y profesionales financiado por (SEV-2013-0317-02). Además, un contrato (2019-Actualidad) de titulado superior en actividades técnicas y profesionales financiado por el proyecto del Plan Nacional de I+D+i con referencia RTI2018-096501-B-I00.

La investigación ha sido financiada por los siguientes proyectos de investigación:

- Plasticidad y comportamineto celular en desarrollo embrionario y progresión tumoral CELLPLAST–MINECO-BFU2014-53128-R. 2015-2018.
- “Programa Severo Ochoa” para centros de Excelencia al IN– SEV-2013-0317-02 y SEV-2017-0273
- “La complejidad de la plasticidad epitelial: implicacion en fisiologia, patologia y estrategias terapeuticas” – RTI2018-096501-B-I00. 2019-2021.

Acknowledgements

He pasado mucho tiempo como estudiante de doctorado en el Instituto de Neurociencias. Esta etapa no ha sido fácil, y ha requerido de mucho sacrificio. Esta tesis no es únicamente un logro propio, sino que es una combinación de esfuerzos profesionales y personales de muchas personas que he tenido alrededor durante estos años. Por todo ello, me gustaría expresar mi más sincero agradecimiento a estas personas.

En primer lugar, quiero agradecer a la Dra. Ángela Nieto por darme la oportunidad de formar parte de su grupo y trabajar en un proyecto tan interesante como este. Gracias por guiarme a través de mis estudios y al mismo tiempo permitirme la libertad de explorar y convertirme en investigador. Gracias por tu apoyo científico, por enseñarme a pensar de manera crítica y por la confianza depositada en mí. Muchas gracias por todo.

Muchísimas gracias a Joan Galceran, su amplio conocimiento ha sido imprescindible para la generación de las líneas de ratón que me han permitido desarrollar esta tesis. Además, me has aconsejado y ayudado en todo lo que he necesitado durante el recorrido de esta tesis.

También me gustaría dar las gracias a todos los demás miembros del grupo por brindarme un ambiente tan agradable en el laboratorio a lo largo de estos años.

Especialmente, quiero dar las gracias a Khalil, mi supervisor práctico. Gracias por toda la orientación y ayuda durante estos cinco años. Todas tus sugerencias han sido de gran importancia para el desarrollo del contenido de esta tesis. Estoy muy agradecido por todo el tiempo que me has dedicado, y deseo que tengas mucho éxito en todo lo que te propongas.

Muchísimas gracias a Cristina, por su ayuda infinita con la organización de todas las líneas de ratón, que ha sido imprescindible para el desarrollo de esta tesis. Muchas gracias también por todo tu apoyo técnico desde que llegué al laboratorio, y por ser una persona tan alegre y amable.

Muchas gracias también a Hassan, por sus consejos y ayuda durante mi investigación. En especial gracias por ser una persona tan positiva y estar siempre “excelente”, espero que no cambies nunca y que tu carrera científica llegue tan lejos como desees.

Acknowledgements

Muchas gracias a todas las increíbles personas que he conocido durante mi etapa de doctorando, y que se han convertido en grandes amigos. Gracias a Ainara, Noemi, Andrea, Lucía, Aysha, Marta y Fran. Gracias por apoyarme día tras día, y por estar ahí en lo bueno y en lo malo. Haberos conocido es lo mejor que me llevo de esta etapa.

Muchas gracias a Sonia y Auxi por toda la ayuda técnica que me habéis brindado facilitando mucho el trabajo en el laboratorio y solucionándome dudas de cualquier tipo.

También quiero dar las gracias a las últimas incorporaciones del laboratorio, aunque no hayamos tenido aún mucho tiempo para conocernos me habéis ayudado en todo lo que he necesitado. En especial gracias a Nitin, por toda su ayuda con los análisis bioinformáticos.

Gracias también a todos los miembros del laboratorio de la Dra. Berta López. Gracias Berta por la buena predisposición y asesoramiento siempre que lo he necesitado. Muchas gracias, Javi, Marilyn, Fran C, Fran Q y Marta por toda la alegría que aportais al laboratorio.

Quiero agradecer también todas aquellas personas que han contribuido directamente en esta tesis pero que ahora se encuentran persiguiendo sus metas y siguiendo sus propios caminos. Gracias a Óscar, Luciano, Hakan, Diana y Verona. En especial gracias a Sergio y Sandra, con quienes empecé compartiendo esta etapa, os deseo lo mejor en vuestro futuro.

También me gustaría reconocer y agradecer todo el esfuerzo de los trabajadores y trabajadoras del Instituto de Neurociencias, cuya labor es indispensable. En especial gracias a todas las personas que trabajan en el animalario por su esfuerzo manteniendo todos nuestros animales. Muchas gracias a Verona y Giovanna, en el servicio de microscopia, por toda vuestra ayuda técnica. Muchas gracias también a Antonio Caler por toda su ayuda, y por ser una fuente tan grande de entusiasmo.

Quisiera agradecer no solo a las personas que han colaborado directamente en esta tesis, sino también a las personas que han contribuido indirectamente, facilitando un equilibrio personal y profesional.

Me gustaría agradecer a mis amigos de Elche de toda la vida. Gracias a Dani por ayudarme a ver las cosas siempre desde una perspectiva racional. Además, gracias a Carla por estar ahí y por las risas cuando nos podemos ver. Gracias a David, por todo lo que hemos aprendido juntos y por estar siempre ahí para apoyarnos mutuamente.

Acknowledgements

Y a ti, Marta, te estoy enormemente agradecido. Gracias por soportarme y apoyarme cuando más lo he necesitado. Gracias por estar siempre a mi lado y por compartirlo todo conmigo. Gracias por tu sentido del humor, por tu positividad y, por soñar conmigo proyectos comunes. Gracias a ti esta etapa ha sido más fácil. Te quiero.

Por último y más importante, quiero dar las gracias a mi familia, en especial a mis padres y mi hermano. Gracias por vuestro apoyo incondicional. Desde pequeño me habéis enseñado a dar lo mejor de mí en todo lo que hago. Además, mi interés por la ciencia os lo debo a vosotros, por despertar mi curiosidad desde que era niño. Vuestra constante fuente de ánimo, no solo durante mi proyecto de tesis sino también durante toda mi vida, ha sido crucial para poder completar esta tesis.

Finalmente, estoy muy agradecido de haber podido disfrutar de una beca que me ha permitido dedicarme a tiempo completo a mi doctorado, y quiero volver a agradecer todas las personas que he mencionado en las anteriores líneas y que han sido parte importante de esta etapa de mi vida. Gracias por todo

Table of contents

Acknowledgements i

Table of contents v

Abbreviations ix

Abstract 1

Introduction 5

1.1 The Epithelial to Mesenchymal Transition in development and disease 7

1.1.2 The EMT during embryonic development 8

1.2 EMT-TFs in development 11

1.2.1 *Snail* 11

1.2.2 *Prrx1* 13

1.2.3 Twist1 15

1.3 Mural cells in vascular biology 16

1.3.1 Pericyte ontogeny and function 17

1.3.2 Endothelial cell-pericyte interactions 18

1.3.3 Pericytes in cancer, diabetic retinopathy, and dementia 20

1.3.4 Different models to study blood vessel formation 21

Objectives 25

Materials and methods 29

3.1 Experimental animals 31

3.1.1 Mouse handling 31

3.1.2 Mouse transgenic lines 31

3.1.3 Genotyping 34

3.1.4 Tamoxifen administration 35

3.2 Histology 35

3.2.1 Paraffin sections 35

3.2.2 Cryostat sections 35

Table of contents

3.3 Immunofluorescent staining 36
3.3.1 Whole mount immunofluorescence 36
3.3.2 Paraffin sections immunofluorescence 36
3.3.3 OCT sections immunofluorescence 37
3.4 Mouse embryo cartilage staining 38
3.5 Matrigel plug assay 39
3.5.1 Matrigel injection 39
3.5.2 Cell sorting from matrigel plugs 39
3.6 RNA extraction, cDNA synthesis and qPCR 40
3.7 Western blot 41
3.8 Statistical analysis. 42
Results 43
4.1 Genetic interaction between EMT-TFs during mouse development 45
4.1.1 *Prrx1* and *Snail1* have a complementary expression pattern during mouse development 45
4.1.2 Genetic interaction between *Snail1* and *Prrx1* 48
4.1.3 *Prrx1* represses *Snail1* in mouse development 53
4.1.4 *Twist1* and Prrx1 are co-expressed during development and genetically interact during digit morphogenesis. 56
4.1.5 Generation of isoform-specific *Prrx1* mutant mice 59
4.2 The role of Prrx1 in the development and homeostasis of the vasculature 64
4.2.1 PRRX1 is specifically expressed in the pericytes 64
4.2.2 PRRX1 is necessary for the integrity of the vasculature 67
4.2.4 PRRX1 is required for neoangiogenesis in the *in vivo* matrigel plug assay 72
4.2.5 The mechanism of *Prrx1* function in the pericytes 76
Discussion 83
5.1 The relationship between *Prrx1* and *Snail1* during mouse development 85

Table of contents

5.2 *Prrx1* and *Twist1* genetically interact during digit morphogenesis 87

5.3 *Prrx1* isoforms can compensate each other during mouse development 88

5.4 The role of PRRX1 in pericyte function 89

Conclusions 95

References 101

Annex: Publication 123

Abbreviations

AAS: Alveolar air space

AC: Astrocytes

APC: Allophycocyanin

AW: Airway

BA: Branchial arch

BBB: Blood brain barrier

BM: Basement membrane

BMP: Bone morphogenetic protein

ChIP: Chromatin immunoprecipitation

CP: Cartilage Primordium

CNS: Central nervous system

DA: Dorsal aorta

DIG: Digoxigenin

DM: Dermomyotome

EC: Endothelial cell

ECM: Extracellular matrix

EDTA: Ethylenediaminetetraacetic acid

EGF: Epidermal growth factor

EMT: Epithelial to mesenchymal transition

FACS: Fluorescence-activated cell sorting

FB: Fibroblast

FGF: Fibroblast growth factor

Abbreviations

FGFR: Fibroblast growth factor receptor

FNP: Frontonasal process

GRN: Gene regulatory network

HV: Hyaloid vasculature

LB: Limb bud

LPM: Lateral plate mesoderm

MC: Meckel cartilage

MG: Microglia

MMP: Matrix metalloproteinase

MNC: Migratory neural crest

NP: Nasal pit

NT: Neural tube

PHVB: Primary head vein branch

ORF: Open reading frame

OL: Oligodendrocytes

PBS: Phosphate-buffered saline

PBST: Phosphate-buffered saline and Tween

PCR: Polymerase chain reaction

PC: Pericyte

PFA: Paraformaldehyde

PMNC: Premigratory neural crest

PS: Palatal selves

PSM: Pre-somitic mesoderm

Abbreviations

RT-PCR: Reverse-transcription polymerase chain reaction

SC: Sclerotome

SDS: sodium dodecyl sulphate

TF: Transcription factor

TGFβ: Transforming growth factor-β

TNC: Tenascin-C

TP: Tectum posteriorus

VSMC: Vascular Smooth Muscle Cell

Abstract

The epithelial-mesenchymal transition (EMT) endows cells with migratory and invasive properties, and it is crucial for the formation of many tissues and organs during embryonic development. This cellular program is triggered after the activation of transcription factors, referred to as EMT-TFs that belong to several gene families, including *Snail*, *Twist* and *Prrx*. During embryonic development, they are mainly expressed in areas of EMT that produce different mesodermal derivatives, and their study along the years have shown important roles in craniofacial and limb development, among others. EMT-TFs are expressed in distinct or overlapping territories, fulfilling specific or cooperative functions. In the adult, the reactivation of the EMT program contributes to the progression of diseases like fibrosis and cancer. The *Prrx1* gene, identified as an EMT-TF in our lab, produces three differently spliced transcripts leading to three protein isoforms, whose functions remain to be characterized. *Snail1* is a classical EMT-TF and it lies very high in the temporal hierarchy of the EMT both in development and tumour progression, always preceding *Prrx1*. On the other hand, *Twist1* is another classical EMT-TF whose expression pattern is very similar to that of *Prrx1* in development and cancer progression.

We have generated and characterized different transgenic mouse lines that have allowed us to study the function of these EMT-TFs, including *Prrx1* isoforms, and their genetic interactions during mouse development. We reveal interactions between *Prrx1* and *Snail1* in the formation of craniofacial derivatives and between *Prrx1* and *Twist1* in digit morphogenesis. We have observed that *Prrx1* is specifically expressed in mural cells during development and in adult mice and revealed that it is required for the integrity of the vasculature in the retina and the lungs. We also find that it is essential for neoangiogenesis, as assessed in *in vivo* assays. Genome-wide transcriptomic analysis indicates that, at the molecular level, Prrx1 is involved in the regulation of the extracellular matrix and of the immune response in pericytes.

Resumen

La transición epitelio mesénquima (EMT) confiere propiedades migratorias e invasivas a las células, y es crucial para la formación de muchos tejidos y órganos durante el desarrollo embrionario. Este programa celular se acciona por la activación de factores de transcripción (EMT-TFs) que pertenecen a distintas familias de genes, incluyendo *Snail*, *Twist* y *Prrx*. Durante el desarrollo embrionario, se expresan en áreas de EMT que producen diferentes derivados del mesodermo y su estudio a lo largo de los años ha mostrado que son importantes para el desarrollo de las extremidades y del esqueleto craneofacial. Los EMT-TFs se expresan en territorios distintos o solapantes, llevando a cabo funciones específicas o cooperativas. En el adulto, la reactivación de la EMT contribuye a la progresión de enfermedades como la fibrosis y el cáncer. El gen *Prrx1*, que se fue identificado como EMT-TF en nuestro laboratorio, produce tres isoformas por procesamiento alternativo de los transcritos, cuya función no está bien caracterizada. *Snail1* es un EMT-TF clásico, que está muy alto en la jerarquía temporal de la EMT durante el desarrollo embrionario y durante la progresión del cáncer, y su expresión siempre precede a la de *Prrx1*. Por otro lado, *Twist1* es otro EMT-TF clásico cuyo patrón de expresión es muy similar al de *Prrx1* en desarrollo embrionario y cáncer.

Hemos generado y caracterizado diferentes líneas de ratones transgénicos que nos han permitido estudiar la función de estos EMT-TFs, incluyendo las isoformas de *Prrx1*, y sus interacciones génicas durante el desarrollo embrionario de ratón. Hemos descubierto interacciones entre *Prrx1* y *Snail1* en la formación de los derivados craneofaciales y entre *Prrx1* y *Twist1* en la morfogénesis de los dedos. Además, hemos observado que *Prrx1* se expresa específicamente en las células murales durante el desarrollo embrionario y en ratones adultos y hemos observado que *Prrx1* es necesario para la integridad de la vasculatura de la retina y de los pulmones. También hemos descubierto que *Prrx1* es necesario para la neoangiogenesis, evaluada mediante ensayos *in vivo*. Análisis transcriptómico a nivel del genoma completo indican que a nivel molecular *Prrx1* está implicado en la regulación de la matriz extracelular y de la respuesta inmune en los pericitos.

Introduction

1.1 The Epithelial to Mesenchymal Transition in development and disease

During embryonic development, a single cell, the zygote, develops to give rise to an organism, with multiple interconnected organs and tissues, each of which contain different specialized cell types. Cells proliferate, differentiate, migrate, and interact with each other in a finely-tuned manner.

During this process, cells travel and differentiate far away from their initial position. This is known as cell migration, which is crucial for the proper formation of different tissues and organs in the embryo. In the case of epithelial cells, in most of the cases, they need to undergo an epithelial to mesenchymal transition (EMT). The EMT is a cellular process by which epithelial cells lose their epithelial, adherent phenotype, and acquire a motile one. There are several cellular features that determine the epithelial-mesenchymal identity of the cell, which include cell-cell adhesion proteins, cytoskeleton remodelling and apicobasal cell polarity (**Figure 1**). Depending on these cellular features, cells can be in epithelial-mesenchymal hybrid states, where they express a mix of epithelial and mesenchymal traits. In addition, the EMT is reversible, and the reverse process, by which a mesenchymal cell becomes epithelial, is called Mesenchymal to Epithelial Transition (MET) (Nieto et al., 2016; Yang et al., 2020).

The EMT is induced by signals from the microenviroment, which in turn, trigger the activation of transcription factors that repress the epithelial identity and/or induce the mesenchymal program. These are known as EMT Transcription Factors (EMT-TFs), which include the members of several families such as *Snail*, *Zeb*, *Twist* and *Prrx*. These EMT-TFs exert their function during embryonic development (Carver et al., 2001; Martin et al., 1995; Nieto et al., 2016 et al., 1994; Takagi et al., 1998), and can be reactivated in pathologies such as cancer (Cano et al., 2000; Comijn et al., 2001; Hajra et al., 2002; Ocaña et al., 2012a; Yang et al., 2004). In this way, the EMT as a developmental program can be hijacked by cancer cells, especially in carcinomas, to delaminate from the primary tumour and migrate, crucial for metastasis (Brabletz et al., 2018). Since metastases are responsible for over 90% of cancer related deaths, understanding EMT programs during development is essential to understand and ultimately treat cancer.

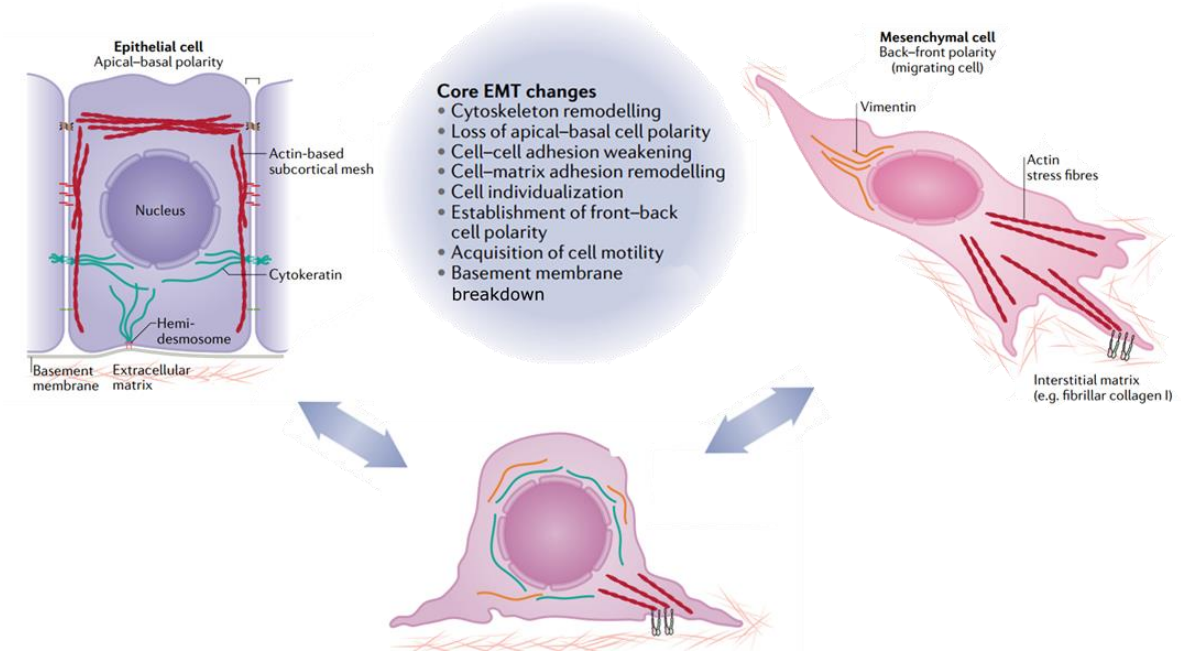


Figure 1 | The epithelial–mesenchymal plasticity. Epithelial cells have a combination of features which include cell to cell contacts, apico-basal polarity, and attachment to an underlying basement membrane. During EMT, these epithelial features are progressively lost. At the same time, cells secrete proteases to break the basement membrane and invade. Cells also remodel their cytoskeleton to acquire motility. Cells can stay in Epithelial/Mesenchymal intermediate states, maintaining combinations of epithelial and mesenchymal features.

1.1.2 The EMT during embryonic development

The EMT is a common event during development, and it can be reactivated in the adult, during wound healing or in pathologies such as cancer and fibrosis. Except for the CNS and the epidermis, all organs and tissues are the result of one or several rounds of EMT/MET occurring throughout ontogeny. The first rounds are called primary EMTs and include those occurring for the formation of the parietal endoderm, and the delamination of cells from the primitive streak at gastrulation and of the neural crest cells (NCCs) from the neural tube (**Figure 2**).

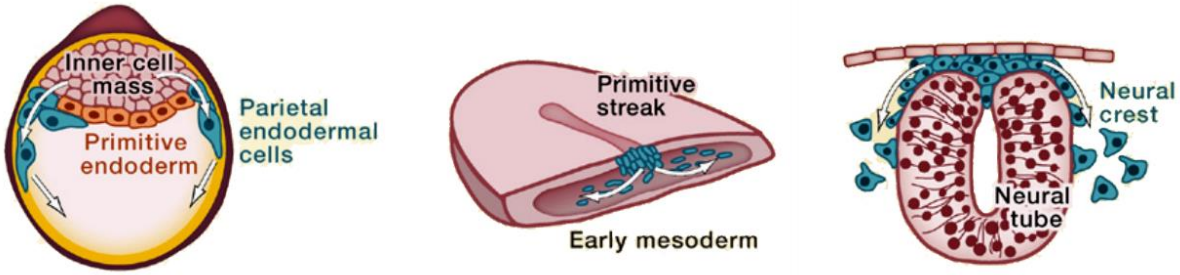


Figure 2| First EMT events during development. Primary EMTs occur early during embryonic development, being the formation of the parietal endoderm the first event of EMT (**left**). Later, gastrulation takes place and cells from the primitive streak undergo EMT to give rise to the precursors of the mesoderm and the endoderm (**middle**). The delamination of NCCs from the dorsal part of the neural tube is considered the last primary EMT (**right**). Adapted from (Adapted from Thiery et al. 2009).

During gastrulation, cells from the medial epiblast undergo an EMT to delaminate from the primitive streak and give rise to the mesoderm and endoderm. At the same time, the ingressing mesoderm instructs the rostral part of the ectoderm to become neural plate, in a process known as neural induction (Stern, 2006). The borders of the neural plate will give rise to the neural crest, which is a vertebrate-specific transient cell population that is induced during the formation of the neural tube (**Figure 3**).

During neurulation, cells at the dorsal part of the neural tube activate a gene regulatory network (GRN) that includes the activation of EMT-TFs such as *Snail1/2* and *Twist1* (Betancur et al., 2010). After the closure of the neural tube, or before that in mammals, cells at the dorsal part undergo EMT and delaminate, forming the migratory neural crest. Eventually, these cells will give rise to the cartilage and bone in the face, including the jaw, to melanocytes and to neurons of the peripheral nervous system. The appearance of the neural crest during evolution has allowed the transition from filter feeding to active predation behaviour in vertebrates (Green et al., 2015).

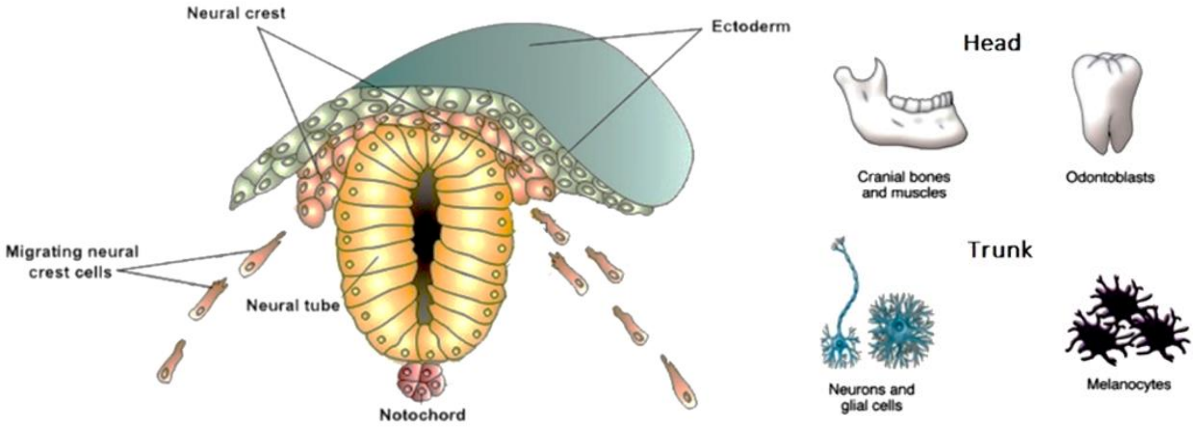


Figure 3| The neural crest and its derivatives. At the dorsal part of the neural tube, pre-migratory neural crest cells undergo EMT, migrate, and then differentiate to give rise to a wide variety of derivatives, such as neurons of the PNS, the craniofacial skeleton, and melanocytes. Adapted from (Kratochwil et al., 2015).

As the mesoderm ingresses through the primitive streak, it is divided into different populations, which form transient epithelial structures, such as the lateral plate mesoderm (LPM) and the paraxial mesoderm. The LPM is divided in somatopleure (dorsal) and splanchnopleure (ventral), which will undergo EMT again, migrate and give rise to different cell populations and tissues. This second round of EMT is termed secondary EMT (**Figure 4**). The posterior paraxial mesoderm is called presomitic mesoderm (PSM), and as it occupies anterior positions, it is subdivided into epithelial blocks called somites. This segmentation of the paraxial mesoderm is controlled by oscillations in gene expression in the PSM (Pourquié, 2003). The genes controlling the generation of the somites include downstream targets of *Notch* and *Snail1/2*. *Snail2* downregulation is required for the formation of the epithelial somites (Dale et al., 2006), which eventually give rise to the segmented body plan of the vertebrates. Somites are divided in sclerotome (ventral) and dermomyotome (dorsal). The sclerotome expresses *Snail1* and undergoes a secondary EMT early after the formation of the somite, and the dermomyotome will undergo EMT later in development.

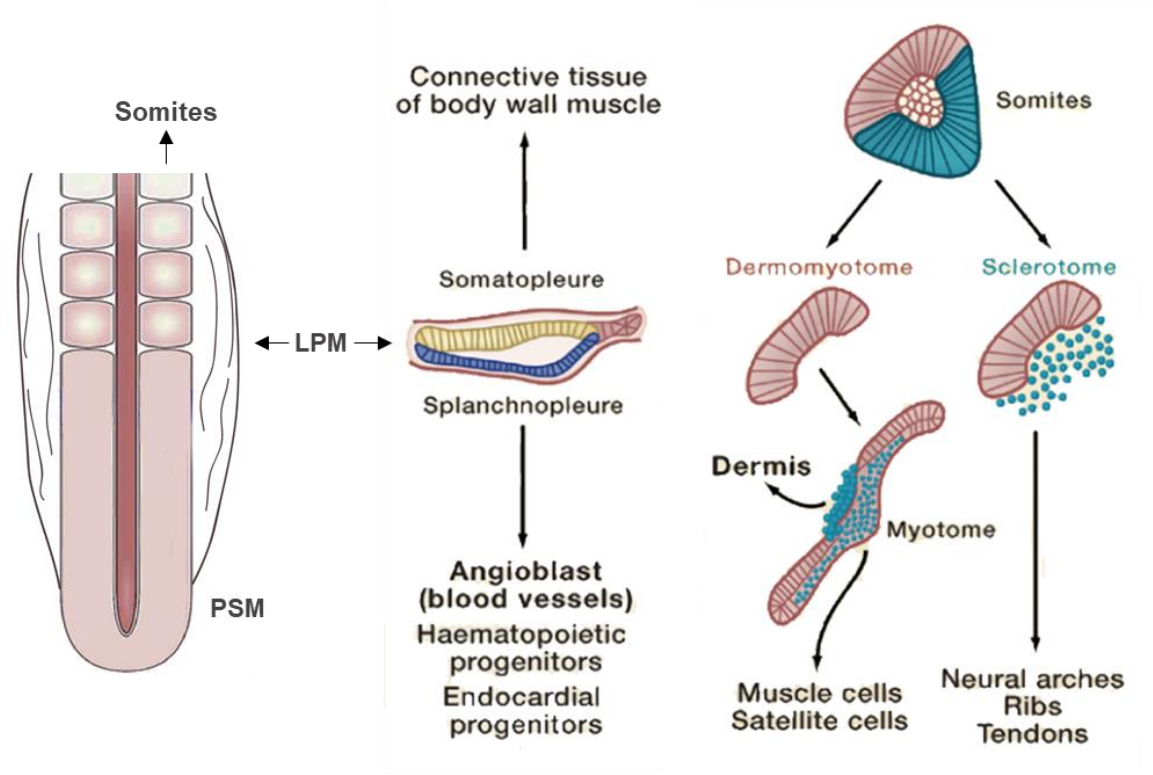


Figure 4| Secondary events of EMT during development. During gastrulation, the mesoderm is subdivided into axial, paraxial, intermediate, and lateral plate mesoderm (LPM). The intermediate and LPM condense into transient epithelial structures that will undergo EMT again. This second round is termed secondary EMT and will give rise to mesenchymal cells that will subsequently migrate and differentiate into specific cell types.

1.2 EMT-TFs in development

During development, different mesenchymal cell populations express unique combinations of EMT-TFs. Their expression pattern is conserved during evolution, although sometimes their territories of expression can be interchanged (Locascio et al., 2002; Ocaña et al., 2017; Sefton et al., 1998), and this is reflected in their developmental phenotypes. In order to understand the function of each EMT-TF during development and disease progression, we must look at the phenotype of the mutant mice for each of these genes.

1.2.1 *Snail*

The *Snail* gene is a zinc-finger transcription factor first identified in *Drosophila*, where it is expressed during the formation of the mesoderm (Boulay et al., 1987), and it is required, together with *Twist,* for gastrulation (Leptin, 1991). The whole genome duplication events

during the evolution at the base of the vertebrate lineage gave rise to the paralogs *Snail1* and *Snail2* genes, coming from the only *Snail* gene present in the non-vertebrate chordates.

Snail1 is expressed in the primitive streak, the ingressing mesoderm, the pre-migratory and migratory neural crest and other mesenchymal cell populations of the mouse embryo (Nieto et al., 1992). *Snail1* represses *E-cadherin* transcription during development and in disease, contributing to the EMT of carcinomas and fibrotic organs (Cano et al., 2000; Grande et al., 2015). Consistent with this, *Snail1* knock-out embryos die during gastrulation (Carver et al., 2001). Interestingly, the mesoderm of *Snail1* knock-out embryos expresses mesodermal identity genes, such as *brachyury,* but shows an epithelial phenotype, confirming the role of *Snail1* as an inducer of EMT and changes in cell behaviour rather than as a fate determinant, as previously thought in *Drosophila*. On the other hand, *Snail2* is required for gastrulation and neural crest formation in chicken embryos (Nieto et al., 1994), but *Snail2* knock-out mice are viable and fertile (Jiang et al., 1998). This difference between mouse and chicken in the role of the members of the *Snail* family is consistent with the reshuffling that occurred during evolution in their expression patterns (Locascio et al., 2002; Sefton et al., 1998).

Mice with the specific deletion of *Snail1* in the migratory neural crest are viable and fertile. However, mice with this same conditional mutation in a *Snail2* knock out background have a much shorter Meckel cartilage and cleft palate (Murray et al., 2007). These same defects are observed in the conditional mutation of other neural crest genes, such as *AP2* and *Tgfbr2* (Brewer et al., 2004; Ito et al., 2003). This indicates that there is a cooperation of *Snail1* and *Snail2* in the formation of the neural crest derivatives. Interestingly, even though both *Snail* genes are mutated in this mouse line, the neural crest can still migrate. Consistent with this, many other genes that have a severe neural crest-related phenotype when mutated in zebrafish, xenopus, or chicken embryos, do not show an equivalent phenotype when the homologous gene is mutated in the mouse. This suggests that the GRN responsible of neural crest migration in mammals might be more robust than in other species (Barriga et al., 2015). In addition, *Snail1* is implicated in bone development and homeostasis (De Frutos et al., 2009; de Frutos et al., 2007). In this context it interacts with *Snail2* and they can compensate for each other's loss (Chen & Gridley, 2013; Tang et al., 2016). Together, these studies indicate that *Snail* genes can cooperate in different contexts.

1.2.2 *Prrx1*

Prrx1 is a homeobox transcription factor which was first identified as a mesenchyme-restricted gene expressed during mouse development (Kern et al., 1992; Cserjesi et al., 1992). *Prrx1* knock-out mice have shorter zeugopods, a malformed Meckel cartilage which results in a shorter jaw, a lack of *tectum posteriorus* resulting in the absence of the supraoccipital bone, and cleft palate among other craniofacial bone malformations (Martin et al., 1995). These mutants die neonatally, most likely from respiratory failure as a consequence of the cleft palate. Interestingly, although *Prrx2* mutants are viable and fertile, *Prrx1* and *Prrx2* double mutants show an intensification of the phenotype observed in *Prrx1* mutants. In addition, they have forelimb and hindlimb polydactyly (Berge et al., 1998), probably as a result of *Fgf8, Shh* and BMP4 misexpression in the double mutant limb (Lu et al., 1999b). This is consistent with the fact that conditional limb-specific *Bmp4* mutations lead to polydactyly (Selever et al., 2004). Mutations in the *Prrx1* gene have been detected in babies with agnathia-otocephaly, showing a very reduced or absent jaw and dying after birth due to respiratory distress (Çelik et al., 2012; Dasouki et al., 2013; Donnelly et al., 2012).

Prrx1 mutant mice also have defects in the development of the vasculature, including defects in the positioning of the great elastic arteries, which are intensified in a *Prrx2* mutant background (Bergwerff et al., 2000). PRRX1 is expressed in the vasculature during lung development (Ihida-Stansbury et al., 2004), and *Prrx1* mutants have defects in the vascularization of the lung (Ihida-Stansbury et al., 2004).

During zebrafish and chicken embryo development, *Prrx1* is asymmetrically expressed in the LPM, with higher levels in the right LPM, inducing a differential left/right EMT that drives the leftward displacement of the posterior pole of the heart (Ocaña et al., 2017). In the mouse, this asymmetric EMT is driven by *Snail1* asymmetric expression. The asymmetric expression of these EMT-TFs is achieved by a wave of asymmetrically expressed miRNAs that are activated by the posterior-to-anterior Nodal wave in the left LPM (Rago et al., 2019).

Prrx1 produces two protein isoforms, resulting after the alternative splicing of exon 4 sequences in its mRNA (Kern et al., 1992). See Figure 5 for Prrx1 protein domains. Both isoforms have the Prrx domain and the homeodomain, located in exon 1 and exon 2, respectively. The isoforms differ at the carboxyl terminus. The longer protein isoform, from

here on PRRX1L (PRRX1A in the bibliography), skips exon 4, which has a premature stop codon, and translates exon 5, which codes for the OAR and AD3 domains. On the other hand, the shorter protein isoform, from here on PRRX1S (PRRX1B in the bibliography), includes all the exons, and stops translation in exon 4 (**Figure 5**).

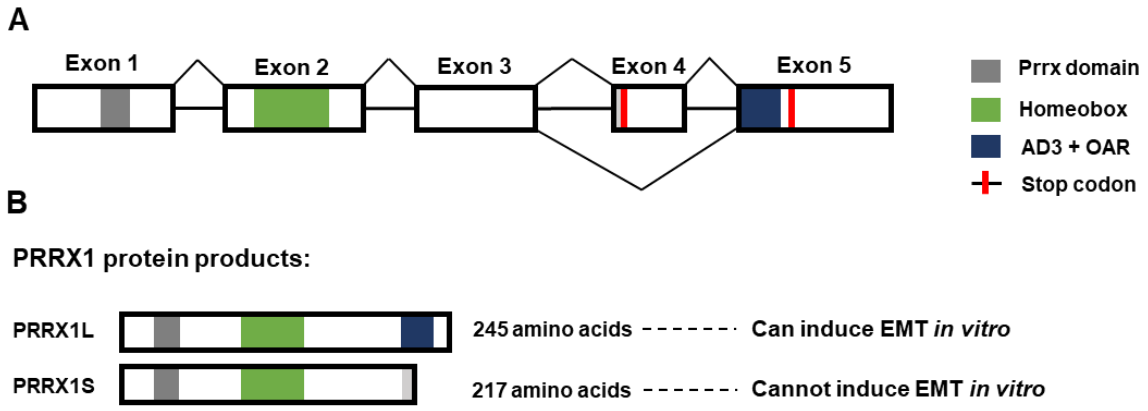


Figure 5| *Prrx1* gene isoforms and protein domains. (A) Schematic representation of the *Prrx1* gene, which has 5 exons. Alternative splicing of exon 4 results in the generation of two different protein products. (B) Schematic representation of the two proteins produced by the *Prrx1* gene. The main difference between both proteins is that PRRX1L includes the AD3 and OAR domains, and PRRX1S does not.

Prrx1 isoforms are conserved in evolution (Braasch et al., 2014), and they are differentially expressed during mouse development (Kern et al., 1992) and in different tissues in the adult (Norris et al., 2000). In addition, they have a substantial difference in their transactivation activity and DNA binding activity in vitro (Norris & Kern, 2001).

Prrx1 is an EMT-TF and can induce EMT and migration in embryos and cancer cells, cooperating with *Twist1*, another EMT-TF commonly co-expressed with *Prrx1*. *Prrx1* downregulation is required for the reversal of the mesenchymal phenotype of disseminated cancer cells and subsequent metastatic colonization (Ocaña et al., 2012). In addition, unpublished data from the lab indicate that PRRX1L can induce EMT *in vitro* but PRRX1S cannot. However, very little is known about the *Prrx1* isoform functions in vivo.

1.2.3 Twist1

Twist1 is a basic helix-loop-helix transcription factor originally described in *Drosophila*, where it is required for mesoderm formation during gastrulation (Thisse et al., 1988). Mouse *Twist1* knock out embryos die at E11.5. They fail to close the anterior region of the neural tube, and show focal haemorrhages, neural crest migration defects, and malformed branchial arches, somites and limb buds (Chen & Behringer, 1995; Soo et al., 2002). Consistent with this, mice with specific deletion of *Twist1* in the migratory neural crest have a drastic phenotype in neural crest-derived tissues, which include the total absence of the frontal bone and the whole frontonasal skeleton (Bildsoe et al., 2009). In addition, mice with *Twist1* deletion in the mesoderm show epithelialization of the head mesenchyme, which also affects neural crest migration (Bildsoe et al., 2013).

Twist1 heterozygous mice are viable and fertile, but they have craniofacial defects that resemble Saethre-Chotzen syndrome in humans. Indeed, patients with Saethre–Chotzen syndrome bear *TWIST1* mutations in heterozygosis. These defects also include hindlimb polydactyly with incomplete penetrance (Bourgeois et al., 1998; Ghouzzi et al., 2000) and craniosynostosis (Carver et al., 2002). Limb-specific deletion of *Twist1* leads to severe limb defects, including polydactyly with complete penetrance, probably caused by defective expression of *Shh* and *Fgf8* (Krawchuk et al., 2010).

Twist1 can also induce EMT in epithelial cells and breast tumours, and plays an important role in invasion and metastasis (Xu et al., 2017; Yang et al., 2004). Similar to *Prrx1*, *Twist1* needs to be downregulated for squamous cell carcinoma to revert EMT and colonize the lung (Tsai et al., 2012).

Interestingly, both *Twist1* and *Prrx1* are important players in the neural crest mesenchymal lineage (Soldatov et al., 2019). In fibroblasts, *Twist1* can directly activate *Prrx1* transcription, while *Prrx1* indirectly activates *Twist1* by directly activating *TNC* transcription. When TWIST1 induces *Prrx1* expression, *Tnc* expression is directly induced by PRRX1 and TNC is secreted and incorporated into the ECM. Fibroblasts become activated when they interact with TNC in the ECM, which leads to an increase in *Twist1* expression. This constitutes a positive feedback loop between *Twist1*, *Prrx1* and TNC in fibroblasts (Yeo et al., 2018). Despite this knowledge, the potential genetic interaction between these two EMT-TFs in mouse development has not yet been explored (Yeo et al., 2018).

1.3 Mural cells in vascular biology

The vasculature is the first functional organ that forms during embryonic development due to its crucial role in providing oxygen and nutrients to every cell in the organism. Blood vessels are composed by two closely associated cell types, endothelial and mural cells. Endothelial cells face the lumen of the vessel, and mural cells tightly wrap around them, sharing a basement membrane and making cell-cell contacts, which include peg-socket junctions, tight junctions, and gap junctions (**Figure 6**).

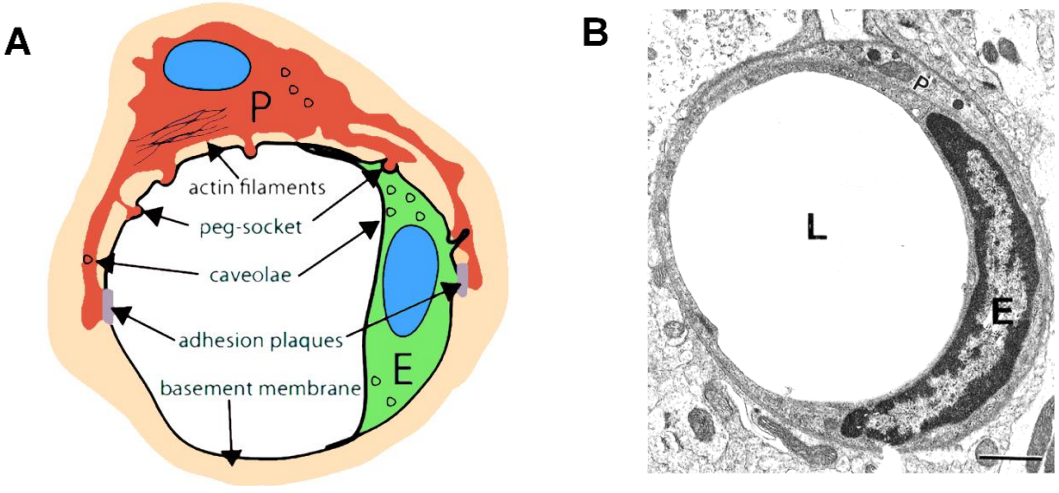


Figure 6| Structure of capillaries. (A) Schematic representation of the transverse section of a capillary, which is composed of endothelial cells that are wrapped and in close contact with pericytes. Endothelial cells and pericytes share a basement membrane, and also make cell-cell contacts in the regions where the basement membrane is not present, adapted from (Armulik et al., 2011). (B) Electron micrograph of a rat sensorymotor cortex capillary, taken from (Melgar et al., 2005). E: Endothelial cell, P: Pericyte, L: Lumen. Scale bar: 1 μm

Mural cells are divided into pericytes and vascular smooth muscle cells (vSMCs). Large vein and arteries are coated by vSMCs and collagen fibres. Arteries are surrounded by multiple layers of vSMCs that allow them to withstand higher blood pressures, whilst veins are sparsely covered by vSMCs. On the other hand, pericytes are associated with smaller vessels, including venules, arterioles, and capillaries (Armulika & Betsholtz, 2009; Bergers & Song, 2005; Gaengel et al., 2009). Pericytes and vSMCs share many molecular markers and embryonic origin, and pericytes can be progenitors of vSMCs (Volz et al., 2015), which

indicates a phenotypic plasticity between the two types of mural cells. Consistent with this, a recent study used single-cell RNA sequencing to uncover the transcriptional profiles of all vascular cell populations of the brain and the lungs and found a gradual phenotypic conversion from pericytes to vSMCs (Vanlandewijck et al., 2018a).

The identification of pericytes is challenging because there is no universal pericyte marker. Morphologically, they have a protruding nucleus and a small cytoplasm with processes that wrap around one or more endothelial cells. The most widely used markers to identify pericytes are *Pdgfrβ* and *NG2*. *Smooth Muscle Actin* (*SMA*) is expressed specifically by vSMCs and therefore, it is used to distinguish pericytes from vSMCs. Other well-known markers include *RGS5*, *Desmin* and *CD13*. Pericytes from different tissues and different developmental stages will express a different combination of markers, although the relationship between the expression of these markers and pericyte function is still not well understood (Armulik et al., 2011). In summary, the identification of pericytes often requires a combination of morphological features, the counterstaining of endothelial cells, and pericyte markers corresponding to the population of pericytes that is being examined.

1.3.1 Pericyte ontogeny and function

The ontogeny of mural cells is very complex because they have multiple embryonic origins, including the neural crest, the lateral plate mesoderm, and the somites (Shen & McCloskey, 2017). Mural cells of the anterior CNS, the face, the carotid arteries and the ascending aorta originate from the neural crest (Etchevers et al., 2001). Mural cells of the limbs, body walls and posterior CNS originate from the sclerotome of the somites (Pouget et al., 2008). In addition, mural cells from one tissue can have multiple developmental origins. Whilst the ascending aorta mural cells originate from the neural crest, the descending aorta is populated with LPM-derived mural cells which become gradually displaced by sclerotome-derived vSMCs (Wasteson et al., 2008). In the adult, pericytes can also originate from various sources, including the bone marrow, and $sca1^{+}$ resident stem progenitor cells (Lamagna & Bergers, 2006; Tigges et al., 2008).

Pericyte function is very diverse and tissue dependent. Their main function is to stabilize blood vessels and control vascular tone and permeability (Hall et al., 2014). However, pericytes are implicated in a multitude of biological processes, including guiding

vasculature growth and remodelling in response to angiogenic stimuli, and orchestrating immune function by secreting cytokines (Bergers & Song, 2005; Kovac et al., 2011; Navarro et al., 2016). Different organs and tissues have dramatically different pericyte density. The ratio between the number of pericytes and endothelial cells ranges from 1:100 in skeletal muscle to 1:1 in the retina (Shepro & Morel, 1993).

The CNS contains the highest density of pericytes in the body, as their function is fundamental for the maintenance of the blood-brain barrier (Armulik et al., 2010). Pericyte ablation from the brain leads to rapid neuron loss due to disruption in the blood-brain barrier, edema, and decrease in blood flow (Nikolakopoulou et al., 2019). They are also crucial players, together with astrocytes, in modulating neurovascular coupling, controlling blood flow based on changes in neural activity (Alarcon-Martinez et al., 2020).

Unlike endothelial cells in the blood-brain barrier, in the liver they are fenestrated to favour the diffusion of metabolites of the portal blood, to allow the processing of toxins by hepatocytes. Hepatic stellate cells are the pericyte equivalent cell population in the liver, where they are in close contact with endothelial cells, and interact with them by secreting extracellular matrix (ECM) and matrix metalloproteases (Sato et al., 2003). During hepatic fibrosis, hepatic stellate cells are one of the major contributors of the excessive accumulation of ECM (Seki & Schwabe, 2015). Interestingly, *Prrx1* expression is induced in activated stellate cells during liver fibrosis in response to PDGF signalling (Gong et al., 2017). In this context, *Prrx1* directly induces the expression of the ECM protein *Col1a1* (Jiang & Stefanovic, 2008).

1.3.2 Endothelial cell-pericyte interactions

The interaction between endothelial cells and pericytes is essential for the growth and remodelling of the vasculature and the control of vascular permeability in different organs (Armulik et al., 2005). This interaction is altered in many diseases such as cancer, diabetic retinopathy and dementia, among others (Park et al., 2017; Yang et al., 2020a; Zhang et al., 2020). Endothelial cells and pericytes interact through several mechanisms, including paracrine signalling, cell-cell contacts, and cell-ECM contacts (**Figure 7**).

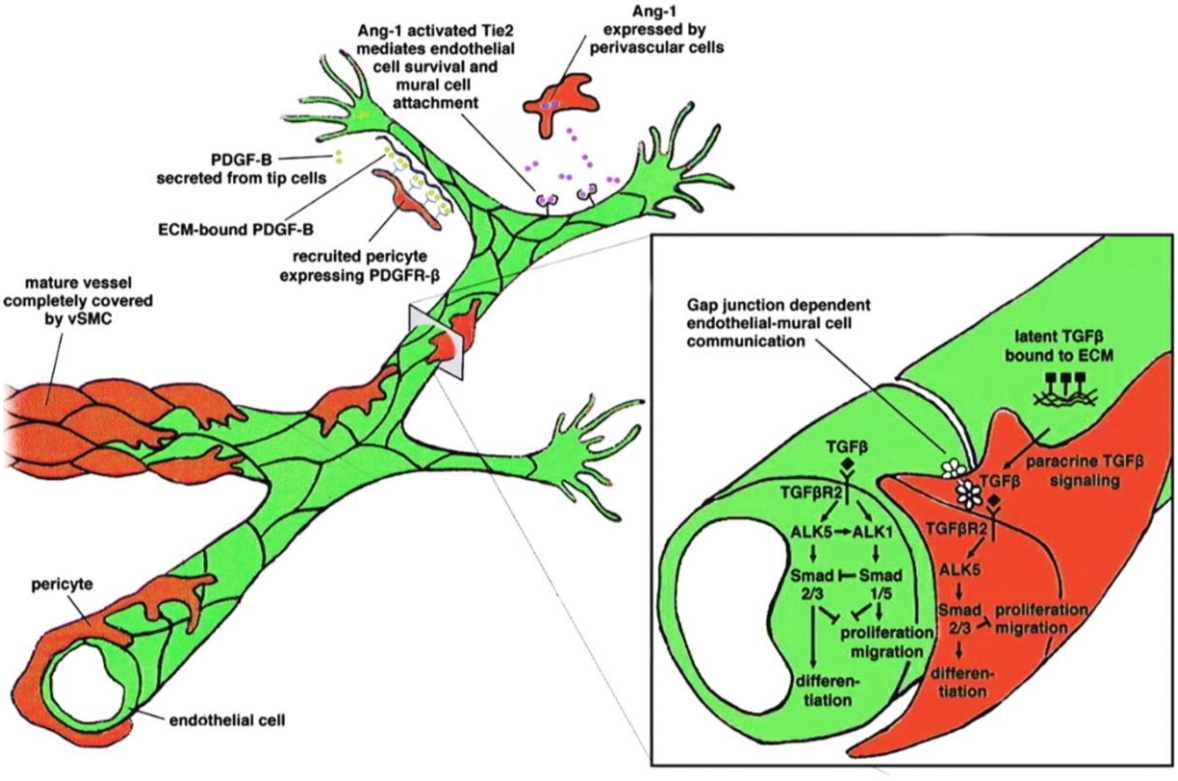


Figure 7| Interaction between pericytes and endothelial cells. Schematic representation of the different interactions between endothelial cells and pericytes. They communicate through paracrine signalling, cell-cell contacts including tight junctions and gap junctions, and cell-ECM contacts. Endothelial tip cells secrete PDGF-B, which serves as a chemoattractant for pericytes expressing the receptor, PDGFRβ. At the same time, pericytes secrete angiopoietin1 (ANG1), which binds to its receptor TIE1 in endothelial cells and stimulates survival. Once in contact, endothelial cells produce TGFβ, that signals to pericytes through gap junctions and can induce different pathways depending on the intensity and duration of the signal. ECM: Extracellular matrix, EC: Endothelial cell. Adapted from (Gaengel et al., 2009).

During angiogenesis, which is the growth of blood vessels from pre-existing vasculature, endothelial cells from angiogenic sprouts secrete PDGFB, which is a chemoattractant for pericytes, which express its receptor, PDGFRβ. Activation of PDGFRβ induces pericyte proliferation and survival. Remarkably, mutant mice for *Pdgfb* and *Pdgfrβ* have the same phenotype, showing perinatal lethality and extensive vascular leakage (Betsholtz, 2004; Lindahl et al., 1997).

On the other hand, pericytes express and secrete Ang1, which activates its receptor, Tie2, in endothelial cells. Activation of Tie2 in turn promotes endothelial cell survival. *Tie2*

knock out mice die around E10.5 and have severe vascular problems and haemorrhage. Interestingly, mice with endothelial cells null for *Tie2* lack pericytes, and *Ang1* mutant mice have a similar but less severe phenotype (Patan, 1998; Suri et al., 1996; Uemura et al., 2002).

TGFβ1 is expressed in endothelial cells and signals to the pericytes through gap junctions and paracrine signalling. *TGFβ1*, *TGFβR2* and *SMAD5* knock-out mutants have a similar phenotype and display defects in yolk sac vasculogenesis (Goumans et al., 2002; Oshima et al., 1996; Yang et al., 1999). Consistent with this, conditional deletion of *Tgfbr2* in hematopoietic cells causes a massive decrease in pericyte coverage in the developing skin vasculature (Yamazaki et al., 2017). In humans, mutations in the *SMAD4* gene, are associated with haemorrhagic hereditary telangiectasia, which is characterized by small haemorrhagic lesions that are the result of leakage of blood from small blood vessels (Gallione et al., 2004).

The interaction between endothelial cells and pericytes stimulates the production of extracellular matrix (ECM) proteins by endothelial cells and pericytes. ECM production is essential for the proper signalling between endothelial cells and pericytes and the maturation and integrity of the vasculature (Stratman & Davis, 2012). Quantitative analysis of the ECM secreted by pericytes using mass spectrometry revealed that collagens are the main ECM protein being secreted, particularly *Col1a1, Col1a2, Col3a1, Col6a1* and *Col6a2.* Consistent with this, hypomorphic mutations in *Col1a1* cause aortic dissection and rupture in adult mice, which die of thoracic haemorrhage around 18 months after birth. Collagen I is also secreted by activated stellate cells during liver fibrosis, in response to increased levels of *Prrx1* (Jiang & Stefanovic, 2008). *Col3a1* mutant mice also have aortic dissection and rupture and mutation of *COL3A1* in humans is associated with Vascular Ehlers-Danlos Syndrome, which is characterized by excessive arterial fragility (Pope et al., 1996).

1.3.3 Pericytes in cancer, diabetic retinopathy, and dementia

Pericyte-endothelial cell interactions are disrupted in cancer. Tumour vessels, deficient in pericytes, poorly organized and with an abnormally loose association with endothelial cells, are leaky (Baluk et al., 2002). Interestingly, there is a negative correlation between pericyte density and tumour metastasis (Yonenaga et al., 2005). As such, mice with a hypomorphic mutation in *Pdgfb* that decreases the recruitment of pericytes to endothelial cells ($Pdgfb^{ret/ret}$) have increased metastatic potential of β pancreatic solid tumour cells (Xian et al., 2006).

Consistent with this, mice with the conditional mutation of *Pdgfb* in the platelets, which are a major source of Pdgfb in the tumour, have a decreased pericyte coverage in pancreatic β solid tumours and show an increase in circulating tumour cells and metastasis (Zhang et al., 2020). Together, these studies indicate that pericytes play a major role during the metastatic cascade.

Diabetic retinopathy is one of the leading causes of blindness and one of the most common complications of diabetes. Interestingly, one of the hallmarks of the non-proliferative phase of diabetic retinopathy is pericyte apoptosis, which leads to capillary leakage and macular edema (Armulik et al., 2011). Mice with a conditional mutation of *Pdgfb* in endothelial cells have reduced pericyte coverage and spontaneously develop diabetic retinopathy (Enge et al., 2002). In addition, hyperglycaemia leads to increased *Prkcd* expression *in vivo* in retinal pericytes, leading to the activation of apoptotic pathways and phosphorylation of PDGFRβ, which desensitises pericytes to PDGFB. Consistent with this, *Prkcd*$^{-/-}$ mice are resistant to hyperglycaemia-induced pericyte death (Geraldes et al., 2009).

Vascular lesions in the brain are very frequent in people over the age of 70, and a big percentage of dementia cases might be linked to cerebrovascular disease. In particular, cerebrovascular disease is thought to be one of the primary triggers of non-genetic Alzheimer's disease (De La Torre, 2004; The Lancet Neurology, 2017). Human subjects with Alzheimer's disease have a very significant reduction in the number of brain pericytes (Sengillo et al., 2014). Consistent with this, mouse models of Alzheimer's disease also show pericyte loss, and mice with a heterozygous mutation in *Pdgfrβ,* which decreases pericyte coverage show an intensification of the disease in the mouse model of Alzheimer's disease (Sagare et al., 2013). Moreover, a recent study has shown that the loss of pericyte coverage during normal ageing leads to increase blood-brain barrier permeability, which in turn leads to neuronal cell death (Yang et al., 2020a).

1.3.4 Different models to study blood vessel formation

The impact of pericytes in cancer, diabetic retinopathy and dementia calls for a need to understand the molecular mechanisms that govern their function and indicates that pericytes represent a therapeutic target to halt the progression of these diseases. With this purpose, animal models are extensively used to investigate the molecular mechanisms of vascular biology.

The postnatal retina

In humans, the retinal vascularization starts at 16 weeks of gestation, and it is fully developed at 40 weeks, right before birth. Premature babies that are born before week 31 are at risk of developing a retinopathy due to relative hyperoxia, that affects the growth of the vasculature. In fact, premature retinopathy is the leading cause of visual impairment and blindness in children (Mechoulam & Pierce, 2003).

In mice, the intraretinal vasculature starts developing after birth in a tightly regulated manner, forming a two-dimensional flat structure during the first week of postnatal development. Angiogenesis starts at the head of the optic nerve around P1, and it spreads reaching the retinal periphery around P8 (**Figure 7**). This is the superficial plexus of the retina, as the intermediate and deep plexus start developing at P7 by vertical growth of the vessels, and are completely mature around P25 (Fruttiger, 2002; Stahl et al., 2010).

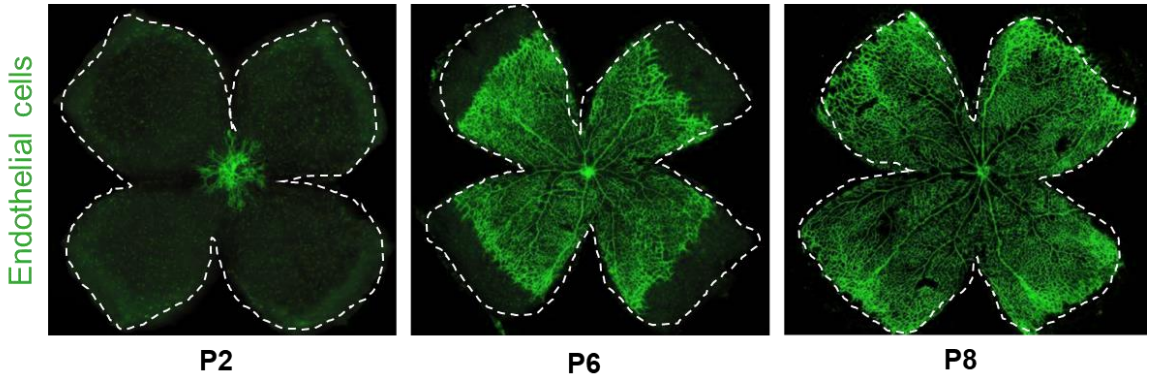


Figure 8| Postnatal retina vascular development. Flat mount of mouse retinas dissected at postnatal day 2, 6 and 8 respectively stained with Isolectin, which labels endothelial cells. Dashed lines represent the borders of the retina. All the endothelial cells with the exception of the endothelial tip cells are covered by pericytes. When the vascular plexus reaches the periphery of the retina, the capillaries start sprouting vertically to form the intermediate and deep plexuses, completed at postnatal day 25. Adapted from (Stahl et al., 2010).

The development of retinal vasculature is a well stablished model to study angiogenesis. Endothelial cell-specific depletion of *Pdgfb* during retinal vascular development leads to pericyte loss, severe haemorrhage, defective organization, and the development of diabetic retinopathy. This indicates that pericytes are essential for the formation of the retinal vasculature and the integrity of the blood-retina barrier (Park et al., 2017). The retina is also widely used as a model for retinopathy of prematurity. For this, P7

pups are exposed to very high concentrations of oxygen for 4 days, and then returned to normal. This leads to the regression of the capillaries in the central part of the retina, which then becomes hypoxic when the pups are under normal concentrations of oxygen. Then, a *Hif1a*-mediated angiogenesis revascularizes the retina. However, the newly vascularized retina shows many hallmarks of retinopathy of prematurity (Ozaki et al., 1999; Smith et al., 1999).

The postnatal retinal vasculature is very easy to observe, especially the superficial plexus, and it is easy to manipulate experimentally, which makes it a very good model to study angiogenesis *in vivo*.

The matrigel plug assay

The matrigel plug assay allows the evaluation of neoangiogenesis *in vivo* (Passaniti et al., 1992). In this model, matrigel is mixed with angiogenic factors such as VEGFA and FGF2 and injected subcutaneously in adult mice. After injection, the matrigel solidifies forming a matrigel plug, which is left implanted in the mouse from 1 to 3 weeks. During this time, neoangiogenesis occurs by migration of vascular progenitors into the matrigel plug, which develop a functional and perfused vascular network 7 days implantation. The assay was initially called “matrigel plug angiogenesis assay”, however, the vasculature of the matrigel plug does not form pre-existing vessels, so the term angiogenesis should not be used.

Different pathway inhibitors or drugs can be mixed into the matrigel to assess their impact on neoangiogenesis (Ponce, 2009). The formation of the vasculature in the matrigel plug can be easily quantified by histological analysis and/or RT-qPCR (Coltrini et al., 2013).

VEGF-A and FGF2 act synergistically to induce neoangiogenesis in matrigel plugs. *FGF2* enhances the expression of *Pdgfrβ* in pericytes, while *VEGF-A* induces the expression of *Pdgfb* in endothelial cells which, in turn, enhances the recruitment of pericytes (Kano et al., 2005). The first steps of neoangiogenesis in the plugs involve the invasion by macrophages and pericyte progenitors. Matrigel plugs implanted in animals that have bone marrow transplants from EGFP+ donors show that most of the pericytes that invade matrigel plugs are bone marrow-derived (Tigges et al., 2008). Five days after implantation, the matrigel plug is invaded by large amounts of $EGFP^+$ pericytes and macrophages. Endothelial cell precursors invade later, and after 7 days, endothelial cells start to be covered by pericytes and the vasculature is formed and perfused (Tigges et al., 2008).

The matrigel plug is a very good model to investigate the molecular mechanisms involved in neoangiogenesis and can be used to dissect the role of the pericytes in this process.

Objectives

Objectives

The epithelial to mesenchymal transition (EMT) is a cellular program that allows epithelial cells to acquire a mesenchymal and migratory phenotype. The activation of EMT is orchestrated by signals from the microenviroment, which in turn activate different EMT-TFs. During development, different cell populations express distinct combinations of EMT-TFs and display a variety of epithelial and mesenchymal features. Among the EMT-TFs we are particularly interested in members of the Snail, Twist and Prrx families. Despite the critical role of the individual EMT-TFs in the vertebrate embryo, which is manifested by the developmental phenotypes of the mutant mice, their potential interactions during embryonic development is not well understood. In addition, *Prrx1* gene transcription leads to the generation of protein isoforms that are conserved in evolution and have different functions *in vitro*, but the evaluation of their function *in vivo* requires the generation of specific mouse mutants for the isoforms that had not been generated yet. Phenotypes associated with *Prrx1* mutants include craniofacial defects, limb abnormalities and vasculature deficiencies (Ihida-Stansbury et al., 2004; Lu et al., 1999b; Martin et al., 1995). While the craniofacial and limb defects have been well characterized, the vascular defects are poorly described. With all of the above, the **general objective** of this work is to better *understand the function of the EMT-TFs during development and getting further insight into the roles of Prrx1*.

Objective 1:

- To generate and characterize different mouse lines to investigate the interaction between different EMT-TFs during development.

Objective 2:

- To generate and characterize mutant mouse lines for the isoforms of *Prrx1* to study their function *in vivo.*

Objective 3:

- To characterize the role of Prrx1 in the formation and maintenance of the vasculature.

Materials and methods

3.1 Experimental animals

3.1.1 Mouse handling

All the animal experimental procedures were conducted in compliance with the European Community Council Directive (2010/63/EU) and Spanish legislation. The experimental protocols were approved in advance by the Animal Welfare Committee at the Institute of Neurosciences, Alicante and the CSIC Ethical Committee.

3.1.2 Mouse transgenic lines

The following mouse lines were used during this thesis:

129S-Prrx1^{tm1Jfm}, abbreviated as *Prrx1* ^{LacZ}:

Conventional constitutive mutant of *Prrx1* gene (M. F. Lu, Cheng, Kern, et al., 1999)**.** A lacZ gene was inserted in-frame at the 5' end of the homeobox domain. A PGK- neomycin cassette was introduced 3' of the lacZ in the reverse transcriptional orientation. Backcrossed to C57 from 129S4/SvJae. Obtained from MMRRC. The phenotype is very similar to the previously described constitutive mutant (Martin et al., 1995), which has an insertion of a neomycin cassette.

FVB-Prrx1$^{em7(e5stop)An}$, abbreviated as *Prrx1* $^{\Delta L}$:

Constitutive mutation of the long isoform of *Prrx1* gene*,* achieved by a CRISPR mediated mutation in exon 5. The mutation was induced using an oligonucleotide and pronuclear injection to induce homologous recombination on FVB mice. A methionine in exon 5 was changed to a STOP codon. This mutation generates a Prrx1 isoform that resembles that of the T (truncated) isoform described in this work. Isoforms S and T are not affected (see validation by Western blot shown in Figure 19). Mouse line generated in the course of this work.

***C57N/129 Prrx1$^{em(e5kie3)An}$*, abbreviated as *Prrx1* $^{\Delta ST}$:**

Constitutive mutation of the short and truncated isoforms of *Prrx1* gene*,* achieved by a CRISPR induced insertion of exon 5 in frame with exon 3. Recombination was performed in ES cells C57Bl6/N/129 using a sgRNA and Cas9-GFP. GFP positive cells were selected and cloned, those that had the insertion of exon5 in exon3 were selected to generate

chimaeras. This mutation generates the Prrx1L isoform instead of the Prrx1S and Prrx1T (see validation by Western blot shown in Figure 19). Mouse line generated in the course of this work.

C57N/129 Prrx1$^{em(e2flox)/An}$, abbreviated as *Prrx1 ^{Flox}*:

In this mouse line, exon 2 of *Prrx1* gene is flanked by loxP sites. It was generated by CRISPR induced homologous recombination, which was performed in C57Bl6/N/129 ES cells, using a sgRNA and Cas9-GFP. GFP positive ESCs were selected and cloned and those that had the two loxP sites and no defects in the region were used to generate chimaeras. The line was backcrossed to FVB and C57Bl6/J and maintained in FVB with Rosa26-tdTomato. When the mice express a Cre recombinase, they will have the exon 2 of *Prrx1* deleted in the tissue that expresses the Cre recombinase. When the exon 2 is removed, the RNA is degraded, and the protein is not produced (see validation by Western blot shown in Figure 21). In this work, this allele was combined with the *UBC-Cre/ERT2* and the *Rosa26-TdTomato* mice. As a result, we have a temporal control of *Prrx1* loss of function and the TdTomato fluorescence is used as a recombination reporter. Mouse line generated in the course of this work.

(Tg(UBC-cre/ERT2)^{1Ejb}, abbreviated as UBC-Cre/ERT2:

This mouse line contains the Cre recombinase fused to the human estrogen receptor ligand binding domain together with the *ubiquitin C* (*UBC*) promoter, which was inserted via transgenesis using a self-inactivating lentiviral vector. Mice display Cre activity in all tissue types upon tamoxifen administration. Allows the temporal control of recombination of any gene of interest. This mouse line was originally generated in Eric Brown's laboratory (Ruzankina et al., 2007).

B6;129S6t(ROSA)26Sor$^{tm9(CAG-td\ Tomato)\ Hze}$, abbreviated as Rosa26-TdTomato:

The *TdTomato* gene is inserted into the *Rosa26* locus with a STOP cassette flanked by loxP sites, which is designed to prevent transcription of *tdTomato*. When the mice express Cre recombinase, they will have the STOP cassette deleted in the tissue that expresses the Cre recombinase, resulting in tdTomato fluorescence. This mouse line was originally generated in Hongkui Zeng's laboratory (Madisen et al., 2010).

H2az2 $^{Tg(Wnt1\text{-}cre)11Rth}$, abbreviated as *Wnt1-Cre*^{tg}

This mouse line, carries a transgene containing the Cre recombinase and the GAL4 transcriptional activator under the *Wnt1* promoter integrated into chromosome 11. It causes a 31 kbp deletion in the *H2afv* gene, but mice are viable and fertile in homozygosis. Expression of the Cre recombinase under the *Wnt1* promoter is directed to the dorsal part of the neural tube and the midbrain. As a result, neural crest cells delaminating from the neural tube will recombine any "floxed" gene of interest (Danielian et al., 1998). This mouse linewas acquired from Jackson laboratory and originally described in (Danielian et al., 1998).

B6.129S6-Snai1$^{tm1.1Stj}$, abbreviated as *Snail1*^{Flox}*:*

In this mouse line, which was originally generated in S. Weiss laboratory (Rowe et al., 2009), who kindly sent it to us. This line has the third exon of *Snail1* flanked by LoxP sites**.** *Snail1* constitutive null mutants die during gastrulation (Carver et al., 2001). This mouse line allows us to bypass the early lethality of *Snail1* mutation and therefore study *Snail1* function in different contexts.

B6;129S-Twist1$^{tm2.1Bhr/J}$, abbreviated as *Twist1*^{Flox}*:*

In this mouse line, exon 1 of *Twist1* is flanked by loxP sites. Exon 1 contains the entire ORF of *Twist1*, and recombination produces a null allele (Chen et al. 2007). *Twist1* constitutive null mutants die at E11.5 (Chen & Behringer, 1995). This mouse line allows us to bypass the embryonic lethality of *Twist1* mutation and study *Twist1* function in different contexts. This mouse line was originally generated in R. Behringer laboratory (Chen et al., 2007) and was kindly provided to us by Dr. Walid D. Fakhouri.

B6;129S-Twist1$^{tm2.2Bhr/J}$, abbreviated as Twist1^{-}

This mouse line has been generated in our lab by inducing early recombination of the *Twist1*^{Flox} allele using the Tat-Cre in 2 cell stage embryos. Homozygous mutants die during development (Chen & Behringer, 1995), so the line is maintained in heterozygosity. Mouse line generated in the course of this work.

3.1.3 Genotyping

The genotyping of the animals used for the maintenance of the colony was performed after weaning, from DNA extracted from ear tissue. The genotyping of the embryos used for the experiments was performed using DNA extracted from the yolk sac. The DNA was extracted from unfixed tissue using the REDExtract-N-Amp Tissue PCR Kit (Sigma), and the PCR was carried out using the primers indicated in Table1 and the KAPA FAST 2x mix.

Table 1. Primers used for genotyping.

TRANSGENIC LINE	AMPLICON LENGTH	SEQUENCE (5' to 3')	DIRECTION
Prrx1 ^{LacZ}	WT: 800 bp	CAT AGA GTA TAA GGG AGT GAG GT	Forward
		TCC CTT CCC TAC AGC ATA TTC C	Reverse
	LacZ: 650 bp	GCT GCA AGG CGA TTA AGT TG	Reverse
Prrx1 $^{\Delta L}$	WT: 242 bp	AGCAGTCTCCTCTTACCCCT	Forward
		CCCCGGCGGATGAAGATATG	Reverse
	ΔL: 326 bp	TCCTTTGTCCCTACAGCGCCATG	Forward
		CATGTGGCAGAATAAGTAGCTTA	Reverse
Prrx1 $^{\Delta ST}$	WT: 650bp	TTGTTGTTGTTCCTCTCCCTCT	Forward
	ΔST: 750bp	GGCAAGAATTCCTGAGTCCA	Reverse
Prrx1 ^{Flox}	WT: 780 bp	TTTCTTTCCCCGTCTTTGGT	Forward
	Flox: 850 bp	AAGGAGATAGCCTTCCCTTCC	Reverse
ROSA26 tdTOMATO	WT: 297 bp	AAG GGA GCT GCA GTG GAG TA	Forward
		CCG AAA ATC TGT GGG AAG TC	Reverse
	tdTomato: 196 bp	GGC ATT AAA GCA GCG TAT CC	Reverse
		CTG TTC CTG TAC GGC ATG G	Forward
UQ-CreERT2	Cre-ERT2: 100 bp	GCGCTCTGGCAGTAAAAACTATC	Forward
		GTGAAACAGCATTGCTGTCACTT	Reverse
Wnt1-Cre	Wnt1-Cre: 358 bp	CCTCTATCGAACAAGCATGCG	Forward
		GCCAATCTATCTGTGACGGC	Reverse
Snail1 ^{Flox}	WT: 324 bp	CTG CCA GGT GGG AAG GAC T	Forward
	LacZ: 424 bp	CAA GGA CAT GCG GGA GAA GGT	Reverse
Twist1 ^{Flox}	WT: 245 bp	GGT TTC CGA CTA GAG GTT TCC	Forward
	Flox: 353 bp	ACT GTC TGG GTC GCT GTT G	Reverse
Twist1 $^{-}$	WT: 245 bp	GAC ACC GGA TCT ATT TGC ATT	Reverse
	Twist1^{-}: 296 bp	CAA TCA GCC ACT GAC AGG AA	Forward
		GGT TTC CGA CTA GAG GTT TCC	Forward

3.1.4 Tamoxifen administration

Tamoxifen (Sigma) was prepared in corn oil with a 5% of EtOH at 10mg/mL for adult mice and 5mg/mL for new-born mice. Adult mice were injected intraperitoneally with 0.1mg/g of body weight. New-born mice were injected intraperitoneally with 0.15mg of tamoxifen (30uL of the 5mg/mL solution) at P3 and P5.

3.2 Histology

The fixation protocol was dependent on the downstream applications. For paraffin sections, the fixation was always at 4ºC overnight with 4% PFA. The same protocol was used for most of the samples processed for cryostat sections. However, small samples were fixed at 4ºC with 4% PFA for 2-6 hours, depending on the side of the sample, to ensure an optimal quality of the immunohistochemical staining.

3.2.1 Paraffin sections

After fixation, the samples were extensively washed with PBS for one day. After overnight washing, the samples were incubated in 70% ethanol in PBS and left overnight. The samples can be saved at -20ºC indefinitely. Then, the samples were incubated in an ascending series of ethanol in PBS (80%, 90%, 100%) for one hour and washed with ethanol 100% three times to ensure that there is no water remaining in the sample. After dehydration, the samples were incubated 2 times for 1 hour in xylol, and then incubated in paraffin. The paraffin was changed after 5 minutes to remove the remaining xylol, and again after 1 hour, and then left overnight in paraffin. The samples were transferred to fresh paraffin in the morning, and then let to solidify in the desired orientation. After solidification, the samples were cut in a microtome (Leica RM2245) to obtain sections 4-10 μm thick.

3.2.2 Cryostat sections

After fixation, the samples were extensively washed with PBS for one day. After overnight washing, the samples were incubated for three days in PBS with 30% sucrose. Then, the samples were embedded in OCT media and placed in the desired orientation before freezing them in dry ice. After this step, the samples can be stored indefinitely at -80ºC. When desired, the samples were cut in a cryostat (Leica CM1850) to obtain sections 10-40 μm thick.

3.3 Immunofluorescent staining

3.3.1 Whole mount immunofluorescence

The fixation protocol was dependent on the size of the sample, smaller samples require shorter fixation times. An optimal fixation time and temperature is crucial for a successful immunofluorescent staining. E8.5 embryos were fixed for 2 hours at 4ºC with PFA, E9.5 embryos were fixed for 6 hours at 4ºC and postnatal eyes were fixed 1 hour at room temperature. Bigger samples were fixed overnight at 4ºC. After fixation, the samples were washed thoroughly with PBST (PBS with 0.1% Twin) to remove all the remaining PFA. Whole-mounted samples were dehydrated in a series of PBT:methanol (25%, 50%, 75% and then 100% two times). Samples can be kept at -20ºC after this step indefinitely. Samples were rehydrated in the reverse order and washed with PBS three times for 10 minutes. Then, the samples were equilibrated in retrieval buffer (TRIS HCl 10mM PH 9, 1 mM EDTA and 0.1% Tween) for 10 minutes and antigen retrieval was performed at 70ºC for 20 minutes. After washing the samples to remove the retrieval buffer, 3 times with PBST, the samples were blocked for 4 hours in blocking solution (1% BSA, 5% NGS, 0.3% Triton in PBS). Samples were subsequently incubated with the primary antibody diluted in blocking buffer overnight. The incubation with the primary antibody varies depending on the sample; E9.5 embryos were incubated for 3 days. The primary antibody was washed extensively, at least 10 times with PBST. The samples were incubated overnight with the secondary antibody diluted in blocking solution. Finally, the samples were washed with PBST at least 5 times. The embryos were embedded in 0.8% low melting temperature agarose in a glass bottom plate for imaging with an Olympus FV1200 confocal microscope. The retinas were flat-mounted on slides with Dako fluorescence mounting medium and imaged with a Leica DMi8 or an Olympus FV1200 confocal microscopes for higher resolution pictures.

3.3.2 Paraffin sections immunofluorescence

The slides containing the paraffin sections were incubated at 60ºC for 40 minutes and then in xylol 2 times for 10 minutes to remove paraffin. Then, they were rehydrated in a series of 0.87% NaCl-mq Water:EtOH (100%, 75%, 50%, 25% and NaCl-Water) for 5 minutes each and washed in PBS. Antigen retrieval was performed in retrieval buffer (TRIS HCl 10mM PH 9, 1 mM EDTA and 0.1% Tween) at 100ºC for 20 minutes. The slides were washed 3 times with PBS and then blocked in blocking solution (1% BSA, 5% NGS, 0.3% Triton in PBS) for 1 hour in a humid chamber. Then, the sections were incubated with the antibody diluted in

the blocking solution overnight in a humid chamber, to avoid the evaporation of the solution. The next day, the slides were washed 3 times for 10 minutes with PBS-T and incubated with the secondary antibody and DAPI diluted in the blocking solution for 1 hour at room temperature. Finally, the slides were washed with PBST 3 times for 10 minutes and mounted with Dako fluorescence mounting medium. The images were acquired with a Leica DMi8 or an Olympus FV1200 confocal microscopes for higher resolution pictures.

3.3.3 OCT sections immunofluorescence

The slides containing OCT sections were thaw at room temperature and let dry for 2 hours at room temperature. Then, they were washed 3 times with PBS and when necessary, antigen retrieval was performed in TRIS HCl 10mM PH 9, 1 mM EDTA and 0.1% Tween at 70ºC for 20 minutes. After retrieval, the sections were cooled down at room temperature and washed 3 times with PBS and blocked in 1% BSA, 5% NGS, 0.3% Triton in PBS for 2 hours in a humid chamber. The antibody was diluted in blocking solution and the incubation was performed overnight at 4ºC in a humid chamber. Then, the slides were washed 5 times for 20 minutes and incubated with the secondary antibody and DAPI diluted in blocking solution for 1 hour at room temperature. Finally, the slides were washed with PBST 5 times for 10 minutes and mounted with Dako fluorescence mounting medium. Images were acquired with a Leica DMi8 or an Olympus FV1200 confocal microscopes for higher resolution pictures. Image analysis and parameters quantification was carried out using the image analysis software ImageJ.

Table 2. Antibodies used for immunostaining

Primary antibody	Species	Conjugated fluorophore	Reference	Company
PRRX1 and PRRX2	Rabbit	-	Ocana et al. 2017	Elly Tanaka's Lab
NG2	Rat	-	MAB6689	RnD systems
NG2 (FACS)	Mouse	APC	FAB2585A	RnD systems
Isolectin-Biotin	-	-	B-1325	Vector labs
CD31	Rat	-		
SNAIL1	Rabbit		3879S	Cell Signalling
SMA	Mouse	-	A2547	Sigma
Ter119	Rat	-	AB_466042	Thermofisher
TWIST1	Mouse	-	ab50887	Abcam
TdTomato	Mouse	-	MA515257	Life technologies
Secondary antibody	**Species**	**Conjugated fluorophore**	**Reference**	**Company**
Anti-Rabbit	Goat	Alexa 488	A11008	Invitrogen
Anti-Rabbit	Goat	Alexa 568	A11011	Invitrogen
Anti-Rabbit	Goat	Alexa 647	A27040	Life technologies
Anti-Rat	Goat	Alexa 488	A11006	Invitrogen
Anti-Rat	Goat	Alexa 568	A11077	Invitrogen
Anti-Rat	Goat	Alexa 647	A21247	Invitrogen
Anti-Mouse	Goat	Alexa 488	A11001	Invitrogen
Anti-Mouse	Goat	Alexa 568	A11004	Invitrogen
Anti-Mouse	Donkey	Alexa 647	A31571	Invitrogen
Streptavidin	-	Alexa 488	S11223	Invitrogen
Streptavidin	-	Alexa 568	S11226	Invitrogen
Streptavidin	-	Alexa 647	S32357	Invitrogen

3.4 Mouse embryo cartilage staining

Embryos were fixed in 4% PFA at 4ºC overnight and then extensively washed with water. For embryos older than E14.5, the skin must be removed to ensure an optimal staining. Then, the embryos were incubated in 35ug/mL Alcian blue 8GX (Sigma) 70% EtOH and 30% glacial Acetic Acid solution handled under the fume hood. The incubation time depends on the embryonic stage. For E14.5 embryos, the optimal incubation time is 24h at room

temperature. Longer incubation times will cause non-specific staining of different tissues. Alcian blue was washed with 100% EtOH multiple times for 3 hours and embryos rehydrated in a series of EtOH:Water (100%, 75%, 50%, 25%) and finally extensively washed with water for 3 hours to ensure there were no traces of EtOH. Next, embryos were incubated in trypsin-borax solution (1% Trypsin, 30% saturated borax solution, 70% mq Water) for different times depending on the embryonic stage. For E14.5 embryos, 2 hours at room temperature is recommended, and over-digestion will result in the loss of the fingers. Embryos were washed in water and then in a 0.5% KOH solution for 1 hour, followed by a series of increasing glycerine concentrations (33%, 66% and then 100%). They can be stored in 100% glycerine indefinitely. Pictures were taken in 66% glycerine solution.

3.5 Matrigel plug assay

3.5.1 Matrigel injection

Matrigel (BD Biosciences) was mixed by pipetting thoroughly on ice and aseptic conditions with 200ng/mL recombinant human VEGF-A (R&D systems, Inc.), 1ug/mL FGF-2 (R&D systems) and 0.1mg/mL Heparin (Sigma). The mix of Matrigel was kept at 4ºC until the time of the injection to avoid solidification. The syringes and needles were kept at -20ºC until the injection time. Mice were anesthetized with isoflurane and then placed on a heating pad. For analgesia, 3ul Buprex/g body weight were injected subcutaneously and the eyes were covered with hydrating gel. The dorsal area was shaved with an electric shaver and then with depilating cream. Then, 500 µl of Matrigel mix were injected subcutaneously in the dorsal region. After injection, the needle must remain in the same position for a few seconds and withdrawn slowly, so that the matrigel solidifies and does not leak out. Matrigel plugs were injected bilaterally and mice placed in a warm chamber until they wake up. After 2 or 3 weeks, mice were sacrificed by cervical dislocation and the matrigel plugs were dissected, split in two parts and fixed overnight at 4ºC with 4% paraformaldehyde. After fixation, one half was incubated for 3 days in 30% sucrose and frozen in OCT and the other half embedded in paraffin.

3.5.2 Cell sorting from matrigel plugs

Matrigel plugs were thoroughly chopped after dissection. Cell Recovery Solution (Corning) was used for the non-enzymatic digestion of matrigel (2mL/g of matrigel) at 4ºC. The chopped pieces of matrigel in cell recovery solution were passed several times through a P1000 pipette tip and incubated at 4ºC together with and 0.2 mg/ml of DNAse in a 6 well

plate for up to 2 hours in an orbital shaker. Every 10 minutes the mix was passed several times through a P1000 to ensure that cells were freed from small pieces of matrigel. Then, complete media was added to the mix to a final volume of 10ml and sequentially passed through 70 μm and 40 μm filters. The cell suspension was centrifuged for 5 minutes at 400g. The media was discarded, and cells resuspended in FACS buffer (PBS, 2 mM EDTA, 25 mM HEPES) together with the desired antibody to label the specific cell population. The cells were incubated with NG2-APC (1:10) and CD31-FITC (1:50) antibodies for 30 minutes at 4ºC. After washing 2 times and spinning down at 400g for 5 minutes, cells were finally resuspended in FACS buffer with DAPI. NG2+ td-Tomato+ cells were FACS sorted (BD FACSARIA III) directly into lysis solution from the Arcturus PicoPure RNA Isolation Kit (Thermofisher). The RNA was extracted using the same kit.

3.6 RNA extraction, cDNA synthesis and qPCR

RNA extraction was performed using the Arcturus ® PicoPure ® RNA Isolation Kit (Thermofisher), following the manufacturer´s instructions. The RNA was eluted in a volume of 13 μl, and 1 μl was used for quantification and quality control using the Bioanalyzer High Sensitivity RNA chip. cDNA synthesis was performed using the Maxima First Strand cDNA Synthesis Kit (ThermoFisher), following the instructions included in the kit. Before RT-qPCR, all the samples were standardized to the same concentration. Then, RT-qPCR were performed using the Fast SYBR Green Mastermix (Applied Biosystems) and the primers listed in **Table 3**. All the primers were tested by performing a standard curve with serial dilutions of the sample of interest, only the primers with efficiency between 90 and 110 were used. Relative levels of expression were calculated using a housekeeping gene. Then, experimental samples where normalized to their respective controls:

$$\Delta Ct = Ct\ (\text{gene of interest}) - Ct\ (\text{housekeeping})$$

$$\Delta\Delta Ct = \Delta Ct\ (\text{experimental sample}) - \Delta Ct\ (\text{control sample})$$

$$\text{Fold change} = 2^{-\Delta\Delta Ct}$$

Table 3. Primers used for qPCR

Gene	Sequence (5' to 3')	Tm	Amplicon size	Direction
Esf2	AGCGAGGACAAAGACAAGGA	59	128bp	Forward
	TACTTCTGTGCAGTGACGGG	60		Reverse
Tbp	CCTTGTACCCTTCACCAATGAC	60.6	119bp	Forward
	ACAGCCAAGATTCACGGTAGA	60.9		Reverse
Cd31	ACGCTGGTGCTCTATGCAAG	62.6	109bp	Forward
	TCAGTTGCTGCCCATTCATCA	62		Reverse
Ng2	AGGACCTAACATTCCGGGTCA	62.4	102bp	Forward
	CTGTGTTGTGGAGGATCTGTATG	60.4		Reverse
Pdgfrβ	CAAGAAGCGGCCATGAATCAG	61.6	114bp	Forward
	CGGCCCTAGTGAGTTGTTGT	61.8		Reverse
CD45	CCCTTCTTCTGCCTCAAAGT	56.7	121bp	Forward
	GTGGATAACACACCTGGATGAT	57.8		Reverse
vWf	CTTCTGTACGCCTCAGCTATG	60	125bp	Forward
	GCCGTTGTAATTCCCACACAAG	62		Reverse
Desmin	GTGGATGCAGCCACTCTAGC	62.6	218bp	Forward
	TTAGCCGCGATGGTCTCATAC	61.7		Reverse
Rgs5	GGGTTGCCTGTGAGAATTACA	60	123bp	Forward
	TGAAGTGGTCAATGTTCACCTCT	61.4		Reverse

3.7 Western blot

The samples (E11.5 mouse embryos or postnatal retinas) were lysed with freshly prepared RIPA buffer right after dissection (150mM NaCl, 25mM TrisHCl pH 7.6, 1% Sodium Deoxycholate, 1% SDS, 1% NP40 and protease Inhibitor Cocktail (Sigma). First, they were chopped with a sharp knife in RIPA buffer, then the samples passed through a P1000 pipette tip multiple times and then through a 25G needle to homogenize the sample. The samples were sonicated for 5 rounds of 15 minutes and then mixed with loading buffer and boiled for 10 minutes at 99ºC. The samples were run in a 13% polyacrylamide gel for 2 hours at 30 mA and 120mV and transferred in CAPS buffer (10mM CAPS-pH10.5) with 15% MetOH for 3 hours at 400mA to a PDVF membrane (Roche). Then, the membrane was washed with TBST (Tris 10mM pH 7.5, 150mM NaCl, 0.1% Tween) and blocked with 5% powder milk in TBST for 3 hours. The primary antibody was diluted in the same blocking solution and incubated with the membrane overnight at 4ºC. Then, the membrane was washed 5 times

for 10 minutes with TBST and incubated with the secondary antibody diluted in blocking solution. Finally, the membrane was washed 3 times for 10 minutes with TBST and developed using Chemiluminescent HRP Substrate (Millipore) and the pictures obtained using the Amersham Imager 680. Quantification of the pictures was carried out using the open-source image analysis software ImageJ.

3.8 Cell culture

The human breast tumour cell line BT-549 was acquired from the ATCC and maintained in F12HAM:DMEM media (1:1) supplemented with 10% of heat inactivated foetal bovine serum (Sigma), 10 µg/mL of insulin (Roche) and 1% Gentamicin (Sigma). Cells were maintained at 37ºC and 5% CO_2, and the media was replaced every 1 to 3 days, depending on the cell density. Cells were passaged when 90% of confluence was reached.

3.9 Chromatin immunoprecipitation (ChIP)

Parental BT-549 cells were fixed at confluency by adding PFA to the media to a final concentration of 1% to the confluent plates for 10 minutes, then the fixation was stopped by adding glycine to a final concentration of 0.125M. Then, the cells were harvested and pooled together from the 4 plates, and the chromatin was isolated using the Pierce™ Magnetic ChIP Kit (Thermo Fisher) following the manufacturers' instructions. The sonication was performed in 15 cycles of 30 second on/off intervals in a Bioruptor® Pico sonication device (Diagenode). Next, the immunoprecipitation and DNA isolation was performed using the same Pierce™ Magnetic ChIP Kit and the anti PRRX1 antibody listed on table 2. The isolated DNA was used for qPCR assays.

3.10 Statistical analysis

Statistical analyses were performed using Microsoft Excel 365 Pro Plus and GraphPad Prism, Version 8. Student's t-test or One-way ANOVA with Bonferroni's multiple comparison test were performed to determine the significancy values. All the values were shown as mean values ± SEM (Standard Error of the Mean). Significant differences between groups were represented as follows: * = P≤0.05, ** = P≤0.01 and *** = P≤0.001.

Results

4.1 Genetic interaction between EMT-TFs during mouse development

During embryonic development, cells migrate from their place of origin to their final destination. To migrate, cells need to undergo EMT, which is triggered by finely-tuned signals. The EMT program is orchestrated by EMT-TFs, which control the genes that in turn lead to the conversion of an epithelial cell into a mesenchymal one. Different cells and tissues activate distinct EMT programs, which require unique combinations of EMT-TFs (Nieto et al., 2016). How do these different EMT-TFs interact during mouse development is still not fully understood.

4.1.1 *Prrx1* and *Snail1* have a complementary expression pattern during mouse development

In mouse embryonic development, *Prrx1* is expressed in mesenchymal cell populations, including the lateral plate mesoderm, the head mesenchyme, and the neural crest, as described using mRNA *in situ hybridization* (Hassan Fazilaty et al., 2019; Ocaña et al., 2012). Using an antibody against PRRX1 (**Table 2**), we have confirmed the expression pattern at the protein level (**Figure 9A**). *Snail1* is also expressed in mesenchymal cell populations, including the presomitic mesoderm, the sclerotome, and the pre-migratory and migratory neural crest populations (Nieto et al., 1992). As for PRRX1, we have confirmed SNAIL1 expression pattern at the protein level (**Figure 9B**). PRRX1 and SNAIL1 show a clearly complementary expression pattern in the posterior part of the early embryo, where SNAIL1 is expressed in the presomitic mesoderm, and PRRX1 is expressed in the lateral plate mesoderm (**Figure 9**, sections 1 and 2). In addition, SNAIL1 is expressed in the pre-migratory neural crest, whereas PRRX1 will be activated later in the migratory neural crest cells and its derivatives, as can be observed in the branchial arches. This is all compatible with recent data from single cell RNAseq analysis revealing that *Prrx1* is expressed in cells that retain the mesenchymal phenotype in the neural crest (Soldatov et al., 2019a). These are the migratory cells before reaching their destination and differentiating into different cell types. Some territories, such as the branchial arches, might seem to co-express *Snail1* and *Prrx1*. However, hierarchical clustering heatmap obtained from the analysis of publicly available single cell RNAseq data from different organs in E9.5 mouse embryos (Dong et al., 2018) (GEO: GSE87038) reveals that *Prrx1* and *Snail1* are rarely co-expressed at the single cell level (**Figure 10A**). Spearman r correlation analysis confirms that there is a significant negative correlation between the expression of *Prrx1* and *Snail1*.

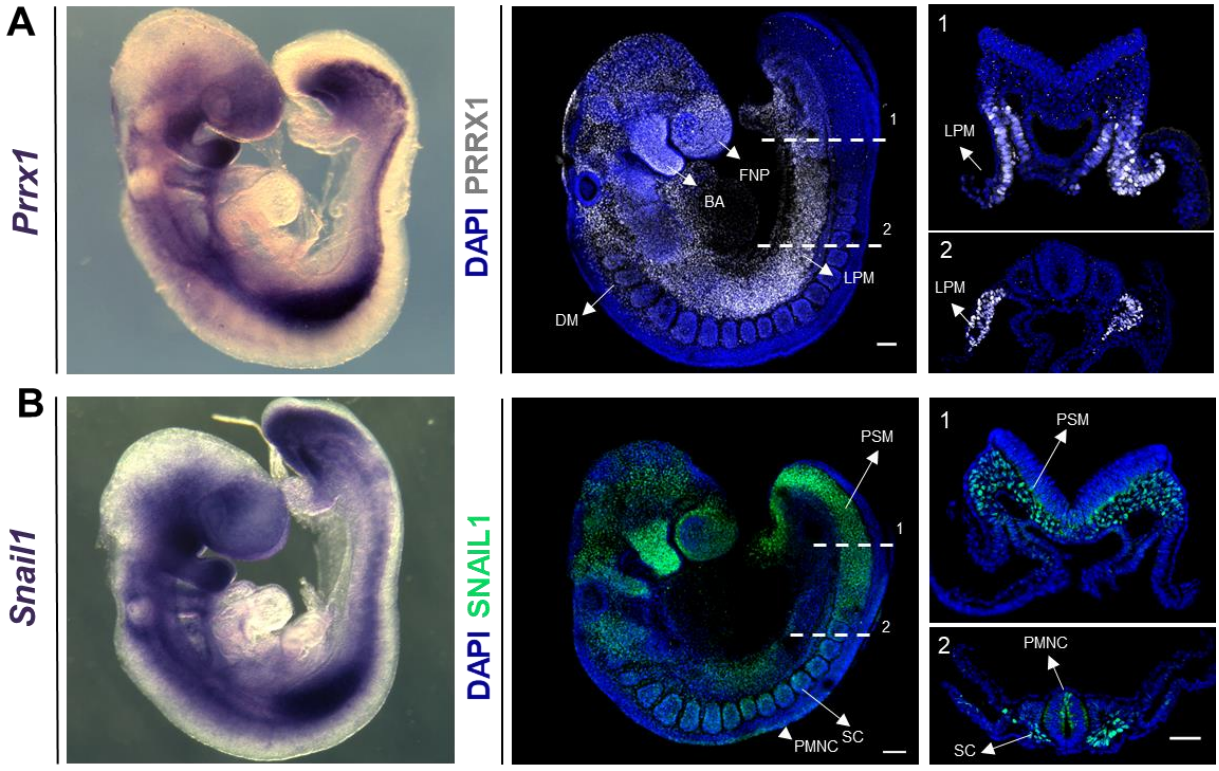


Figure 9| *Prrx1* and *Snail1* are expressed in the neural crest and mesodermal derivatives in a complementary manner during mouse development. (A) Lateral view of an E9.5 mouse embryo showing *Prrx1* expression detected by *in situ* hybridization (left panel). PRRX1 protein expression detected by immunofluorescence. Transverse sections 1 and 2 were taken from the embryo at the position indicated by the dashed lines. Expression is detected in the lateral plate mesoderm, the limb bud, the dermomyotome, the head mesenchyme and the branchial arches. We can confirm that the expression of PRRX1 protein corresponds with the previously described RNA expression pattern. (B) Lateral view of an E9.5 mouse embryo showing *Snail1* expression detected by *in situ* hybridization (left panel). SNAIL1 protein is expressed in the head mesenchyme and the branchial arches. In addition, as observed in sections 1 and 2, SNAIL1 is expressed in the presomitic mesoderm, the sclerotome of the posterior somites and the pre-migratory neural crest. The blue signal corresponds to DAPI staining to label the cell nuclei. All these territories are negative for PRRX1. Scale bar, whole-mounted embryos: 100 μm, sections: 50 μm. LPM: Lateral plate mesoderm; NT: Neural Tube; FNP: Frontonasal Process; BA: Branchial Arch; PSM: Presomitic mesoderm; SC: Sclerotome; PMNC: Premigratory Neural Crest

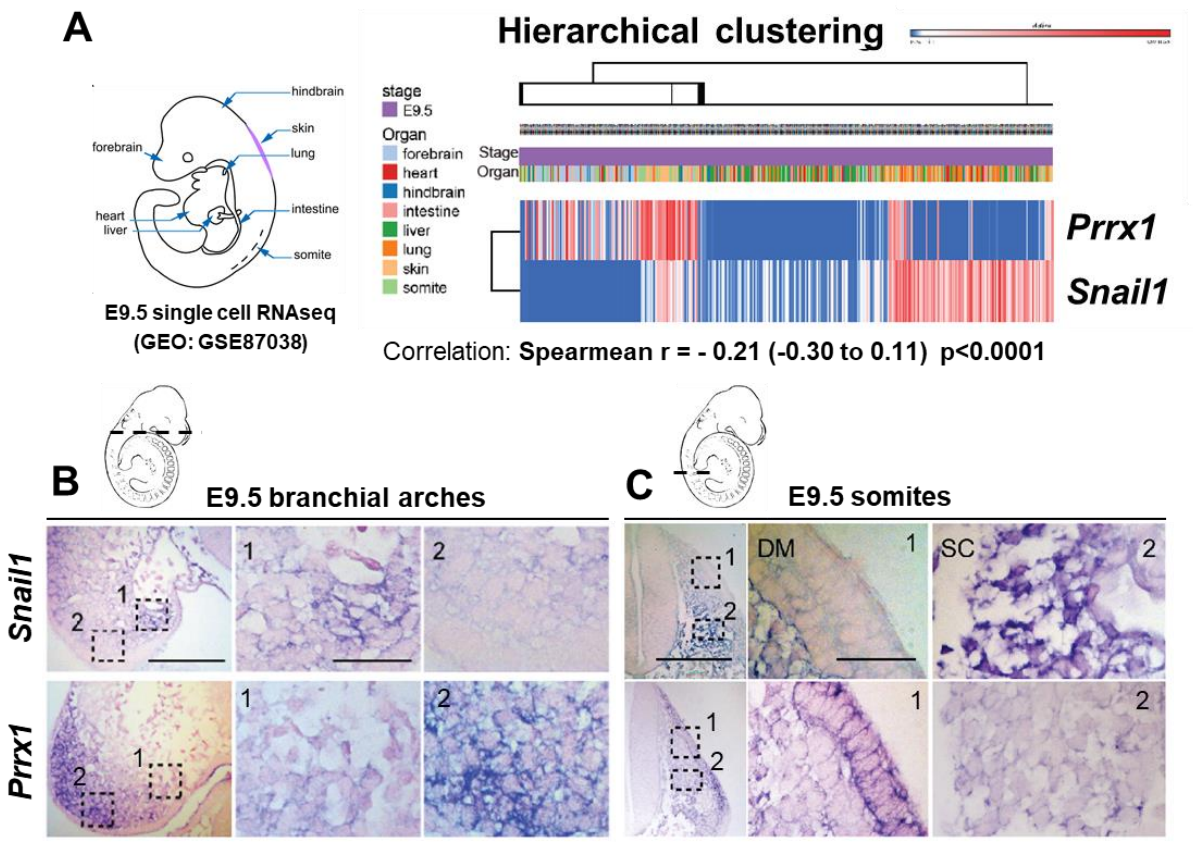


Figure 10| *Snail1* and *Prrx1* show complementary expression patterns. (A) Hierarchical clustering heatmap from publicly available single cell RNA sequencing from different organs and tissues from E9.5 embryos (GEO: GSE87038), showing complementary expression of both transcription factors. (B and C) Transverse sections of E9.5 mouse embryos at the level of the branchial arches and the anterior somites, respectively. Dashed boxes represent the region where the higher magnification pictures (1 and 2) were taken. The expression of *Snail1* and *Prrx1* in these territories is complementary, consistent with the RNAseq data. Scale bars: 100 μm for sections and 25 μm for insets (boxes 1-2). SC: Sclerotome, DM: Dermomyotome.

In situ hybridization and sectioning of E9.5 embryos revealed that although both EMT-TFs are expressed in the branchial arches, the expression pattern is complementary (**Figure 10B**) (Fazilaty *et al*. 2019). In the differentiating somites, *Snail1* is expressed in the sclerotomal mesenchymal cells (**Figure 10C**), cells from the ventral part of the somite that undergo EMT and will ultimately give rise to the vertebrae. At the same time and in a complementary way, *Prrx1* is expressed in the dermomyotome, the dorsal part of the somite that will eventually give rise to muscle cells and satellite cells and the dermis.

4.1.2 Genetic interaction between *Snail1* and *Prrx1*

Snail1 expression precedes *Prrx1* expression in the neural crest and in the mesoderm. In the neural crest, *Snail1* is expressed in the premigratory population and *Prrx1* is activated in the migratory cells when they acquire a mesenchymal phenotype (Soldatov et al., 2019a). In the mesoderm, *Snail1* is first expressed in the primitive streak (Nieto et al., 1992). The ingressing early mesodermal cells will give rise, among others, to the lateral plate mesoderm, which does not express *Snail1* anymore, but expresses high levels of *Prrx1*. *Snail1* expression oscillates in the presomitic mesoderm (Dale et al., 2006), where *Prrx1* is not expressed. Then, *Snail1* is downregulated and these cells undergo mesenchymal to epithelial transition to form the epithelial somites, where *Snail1* will be later reactivated in the sclerotome and *Prrx1* in the dermomyotome as cells prepare to migrate again (**Figure 10B**). When analysed in detail with a *Snail1* reporter, we found that *Prrx1* positive cells derive from cells that had previously expressed *Snail1* (Fazilaty et al., 2019). Altogether, this led us to wonder whether these two EMT-TFs could genetically interact during mouse development.

First of all, we should look at the phenotype of mice mutant for both genes (R. Mani et al., 2008a). Thus, we generated a new mouse line combining one with constitutive loss of *Prrx1* function (M. F. Lu, Cheng, Kern, et al., 1999), with another line bearing the *Snail1* allele with exon 3 floxed (Rowe et al., 2009). We also included a *Wnt1-Cre* transgene, which when combined with *Snail1*-floxed will induce *Snail1* deletion and subsequent loss of function in the neural crest.

Mice with constitutive loss of *Prrx1 (pan-Prrx1* constitutive mutants*)* have an insertion of the *LacZ* gene in exon 2 (**Figure 11A**). *Pan-Prrx1* homozygous mutants have defects in the craniofacial skeleton, shorter limbs, and die at P0 due to a severe cleft palate (M. F. Lu, Cheng, Kern, et al., 1999; Martin et al., 1995). To confirm the absence of PRRX1 protein in the homozygous mutant embryo, we performed western blot for proteins extracted from an E11.5 embryo (**Figure 11B**). In the WB from the WT embryo two bands corresponding to the long (L) and short (S) isoforms are detected. Prrx1L or S are not present in the homozygous mutant $Prrx1^{LacZ/LacZ}$, and the faint band that persists in this sample matches the size of *Prrx2*, which is also recognized by the anti-Prrx1 antibody. This indicates that, as expected, PRRX1 is not expressed in *Pan-Prrx1* homozygous mutants.

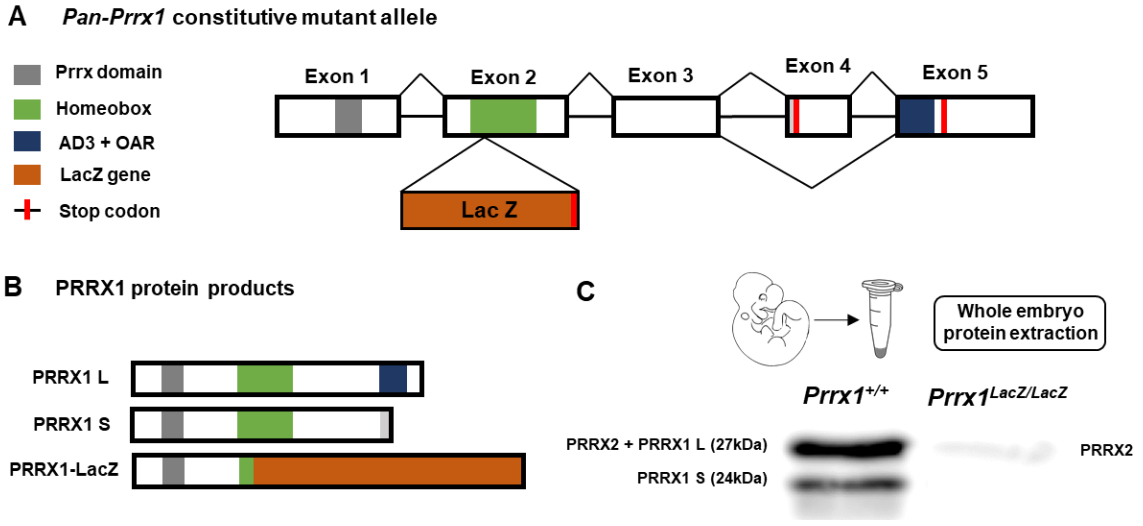


Figure 11. *Prrx1* constitutive mutant mouse model. (A) Schematic representation of the *Prrx1* gene. The *pan-Prrx1* constitutive mutant (129S-Prrx1^{tm1Jfm}) has the *LacZ* gene inserted in frame at the beginning of the sequence encoding the homeodomain. (B) Schematic representation of PRRX1 protein isoforms and that produced by the constitutive mutant. (C) Validation of the constitutive mutant by western blot of whole protein extracts from a WT embryo and a homozygous constitutive mutant. The antibody recognizes both PRRX1 and PRRX2, since PRRX2 has the same size as the PRRX1L isoform, we observe a faint band in

During evolution, *Snail1* and *Snail2* have interchanged expression territories (Locascio et al., 2002; Sefton et al., 1998). In the chicken, *Snail2* is expressed in the primitive streak and pre-migratory neural crest, and it is necessary for gastrulation and neural crest cell delamination (Nieto et al., 1994). In mouse embryonic development, *Snail1* is expressed in the premigratory neural crest and the primitive streak*, and Snail1* constitutive mutant mice die during gastrulation due to a defective EMT in the ingressing mesodermal cells (Carver et al., 2001). To bypass *Snail1* early lethality and study its function in the neural crest, we have bred together the *Snail1*^{Flox} allele (Rowe et al., 2009) bearing *Cre* recombinase expressed under the control of the *Wnt1* promoter (Danielian et al., 1998) and *Rosa26-TdTomato* as a reporter of recombination (**Figure 12A, B**). The *Wnt1* promoter directs *Cre* recombinase expression to the dorsal part of the neural tube, including the premigratory crest and thus, the migratory neural crest cells are labelled with *TdTomato* (**Figure 12C**).

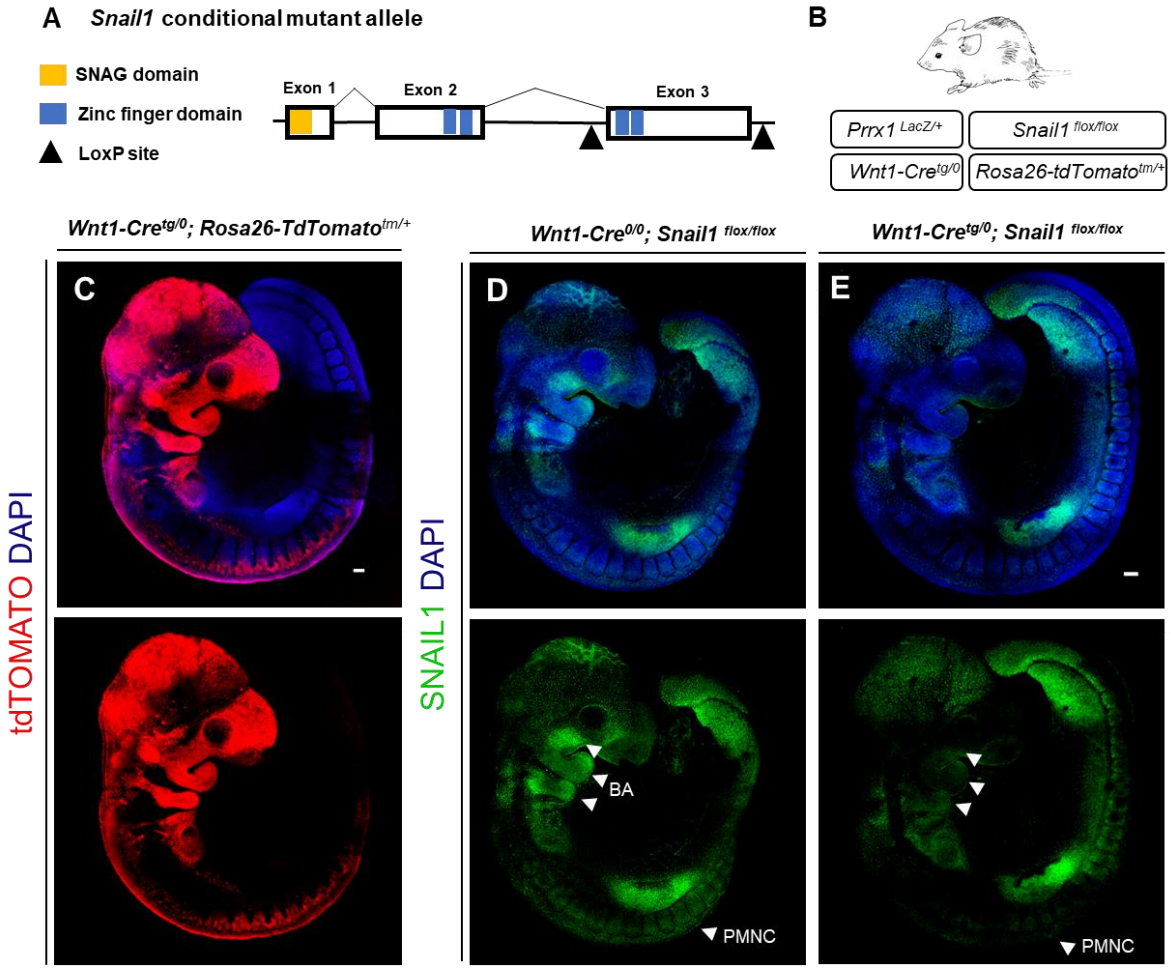


Figure 12. Mouse model for the evaluation of the genetic interaction between *Snail1* and *Prrx1* in the migratory neural crest and its derivatives. (A) Schematic representation of the *Snail1* gene. The floxed allele has loxP sites flanking exon 3. (B) Mouse model used for the evaluation of the genetic interaction between *Snail1* and *Prrx1.* The *Wnt1-Cre* line induces recombination of the floxed alleles in the dorsal part of the neural tube. The *Rosa26-TdTomato* allele has a premature stop codon that is floxed, and once recombination occurs, TdTomato is expressed, and its red fluorescence is used as a reporter for recombination. (C) Transgenic E9.5 embryo bearing the *Wnt1-Cre* and *Rosa26-TdTomato* alleles. The recombination happens in the dorsal part of the neural tube, where the neural crest originates, and TdTomato expression is observed in the migratory neural crest. (D) SNAIL1 expression in a WT embryo. (E) Transgenic embryo bearing the *Wnt1-Cre* allele and the *Snail1* floxed allele in homozygosis. SNAIL1 expression is lost in the neural crest derived territories (arrowheads). Scale bar: 100 μm. BA: branchial arches, PMNC: pre-migratory neural crest.

As expected, *Wnt1-Cre*$^{tg/0}$*; Snail1*$^{flox/flox}$ mice show neural crest-specific SNAIL1 loss, as assessed by immunofluorescence (**Figure 12D, E**). However, the Cre recombinase does not act soon enough to prevent *Snail1* expression in the pre-migratory neural crest. This might be the reason why Wnt1-Cre^{tg}; Snail1$^{flox/flox}$ are viable and fertile (Murray et al., 2007).

We have found that *Wnt1-Cre*$^{tg/0}$*; Snail1*$^{flox/flox}$*; Prrx1*$^{LacZ/+}$ mice are also viable and fertile (**Figure 13A**), although Wnt1-Cre$^{tg/0}$; Snail1$^{flox/flox}$ have a shortened head (**Figure 13B**), which is not mentioned in previous studies using these mice (Murray et al., 2007).

Prrx1 mutants also have craniofacial defects, including a malformed Meckel cartilage, which is derived from the neural crest (Martin et al., 1995). We examined the development of the Meckel cartilage in the double mutants at E14.5 by cartilage staining and observed that the phenotype of the double mutant is more severe than that of the single mutant embryos (**Figure 13C**). This shows that *Snail1* and *Prrx1* genetically interact in the formation of the Meckel cartilage.

Ventral view of the head of DAPI stained embryos support the genetic interaction between *Snail1* and *Prrx1* in the development of the lower jaw, as the phenotype is much more evident in the double than in single mutants (**Figure 13D**).

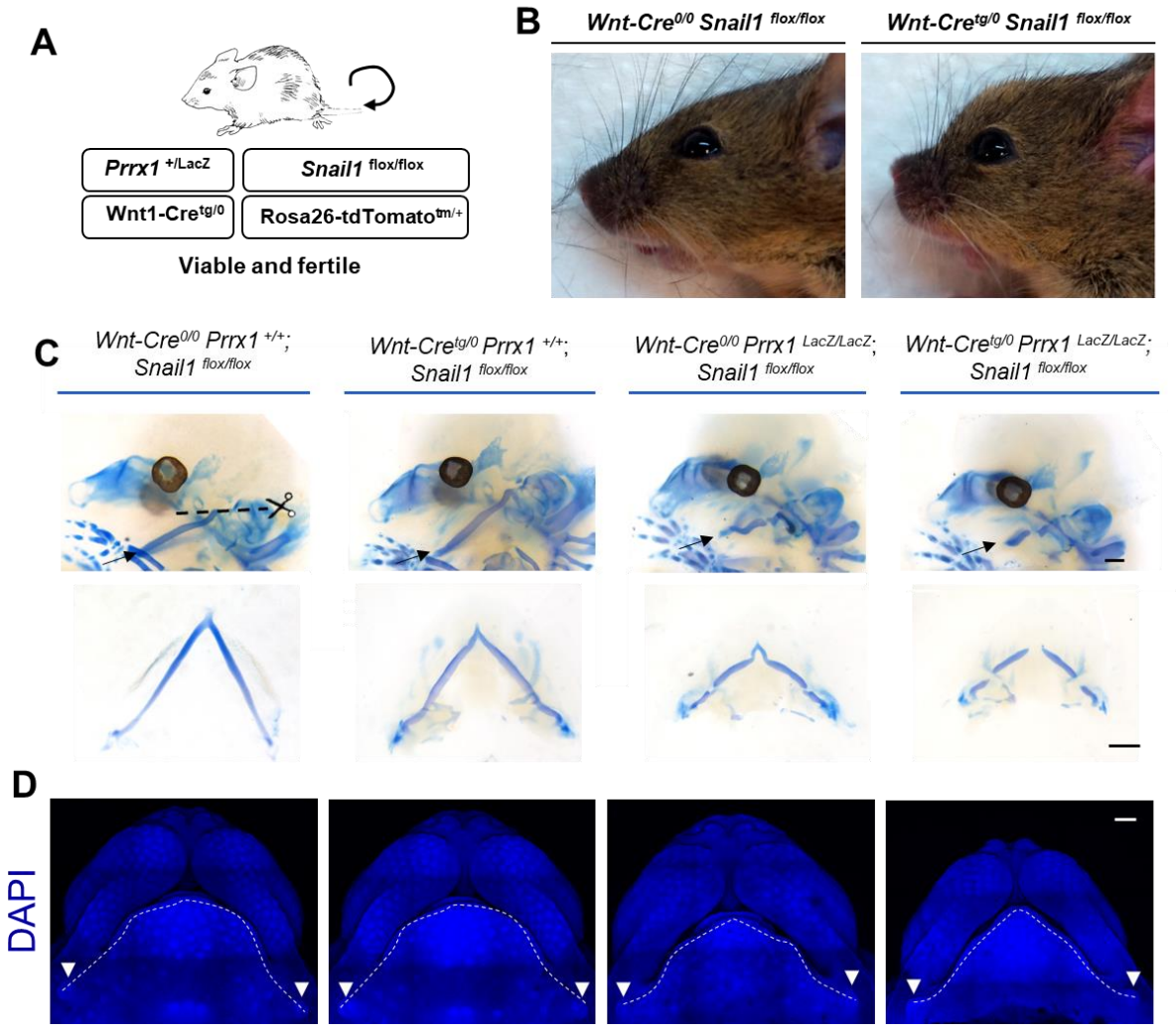


Figure 13. Genetic interaction between *Prrx1* and *Snail1* in the formation of the Meckel cartilage. (A) Schematic representation of the experimental strategy, *Wnt1-Cre*$^{+}$*; Snail1*$^{flox/flox}$*; Prrx1*$^{LacZ/+}$ mice are viable and fertile. (B) *Wnt-Cre*$^{tg/0}$ *Snail1* $^{flox/flox}$ are viable and fertile and have a shortened nose. (C) Cartilage staining of E14.5 mouse embryos, showing the whole head (upper panel) and the dissected Meckel cartilage (lower panel), which is derived from the neural crest. Both *Snail1* cKO and the *Prrx1* pan-constitutive mutant show defects in the formation of the Meckel cartilage, and the double mutant shows an intensification of the phenotype (n = 3). (D) Ventral view of DAPI stained E14.5 embryos. The dashed lines follow the limit of the lower jaw until the condylar process (arrowheads). Scale bar: 1mm.

4.1.3 *Prrx1* represses *Snail1* in mouse development

To assess whether *Prrx1* and *Snail1* regulate each other during mouse development, we have used the E11.5 the *Prrx1* pan-constitutive mutant mice. First, we examined whether the embryonic territories of cells expressing *Prrx1* are visibly changed in the *Prrx1*$^{LacZ/Lac}$ mutant (**Figure 14A**). Next, we measured SNAIL1 levels in `protein extracts from whole E11.5 embryos. We found an increase in SNAIL1 protein levels in *Prrx1*$^{LacZ/LacZ}$ embryos **(Figure 14B**). Quantification of the western blot revealed a 50% increase of SNAIL1 protein levels in these embryos a significant increase considering that the protein is extracted from the whole embryo.

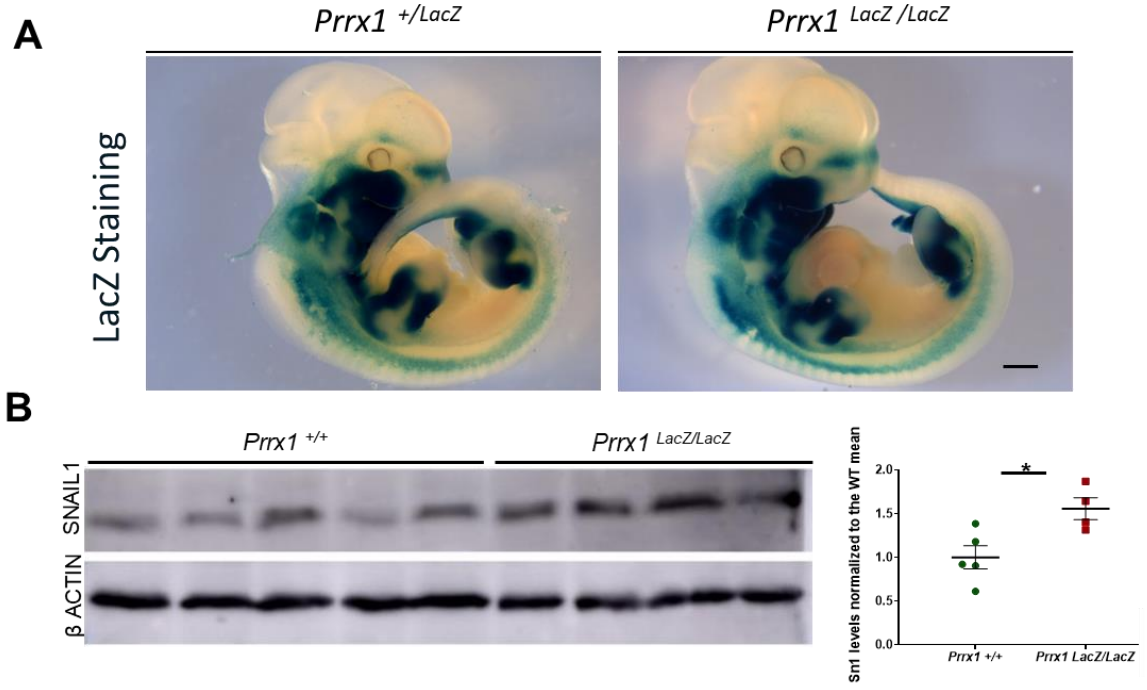


Figure 14. Snail1 expression is increased in the *Prrx1* constitutive mutant. (A) LacZ staining of E11.5 *Prrx1* mutant mouse embryos. (B) Western blot of whole protein extracts from E11.5 mouse embryos. Protein levels were normalized using β-ACTIN as a sample loading control. Quantification shows a significant increase in SNAIL1 protein levels in the *Prrx1* constitutive mutant (n=4). Bars represent mean plus SEM, asterisks indicate significant p-value in t-test (* $p < 0.05$,** $p < 0.01$ and *** $p < 0.001$). Scale bar: 1mm.

We next performed immunofluorescence staining of transversal sections obtained from these embryos (**Figure 15A**) and detected the normal PRRX1 expression in the mesenchyme surrounding the nasal pit in control embryos (**Figure 15B**). In the *Prrx1* mutant embryos, we observed an expansion in the territory of SNAIL1 expression in the mesenchyme surrounding the nasal pit, normally occupied by *Prrx1* expression in a WT embryo (**Figure 15C**). We quantified the intensity of SNAIL1 immunofluorescence signal by calculating intensity of each nuclei the *Prrx1* pan-constitutive mutant compared with those in the WT embryo. We found a statistically significant increase in SNAIL1 expression (**Figure 15D**). This result, together with previous experiments, constituted additional evidence to the finding that PRRX1 and SNAIL1 repress each other's transcription. PRRX1 indirectly attenuates *Snail1* expression through the activation of the *miR-15* family. and SNAIL1 binds to *Prrx1* promoter directly repressing its expression. This mutual repression of EMT-TFs represents a novel gene regulatory network described in the lab (**Figure 15E**) (Fazilaty et al., 2019). *Snail1* is a very early response gene to EMT-inducing signals, probably due to its poised promoter (Lagha et al., 2013). *Prrx1* is activated much later, when cells have already started the epithelial-to mesenchymal transition, and it establishes its own EMT-mode by repressing *Snail1*. This is consistent with the expression pattern that is observed during development for these EMT-TFs and as such, it leads to a sequential activation of two different EMT programs.

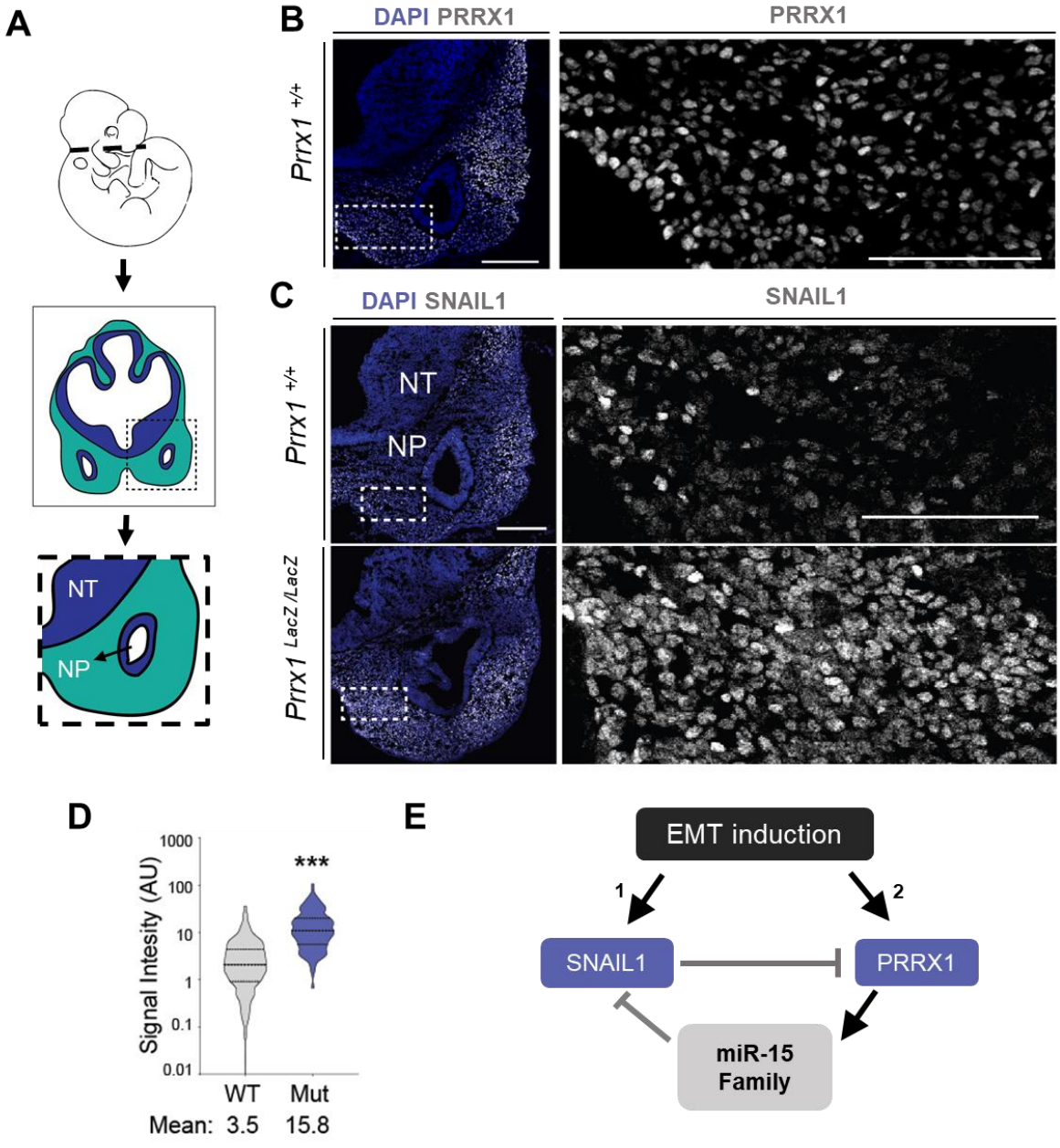


Figure 15. SNAIL1 expression territory is expanded in the *Prrx1* constitutive mutant. (A) Schematic representation of the transverse sections on E11.5 embryos. (B) PRRX1 immunostaining in a transverse section at the level of the nasal pit. (C) SNAIL1 immunostaining in a transverse section of *Prrx1* $^{+/+}$ and *Prrx1* $^{LacZ/LacZ}$ embryos at the level of the nasal pit**.** (D) Quantification of the intensity of SNAIL1 signal in each nuclei of the selected region. SNAIL1 expression territory is expanded in the *Prrx1* constitutive mutant. (E) Schematic representation of the gene regulatory network in which PRRX1 and SNAIL1 repress each other's transcription. Scale bar: 250μm. NP: Nasal pit, NT: Neural tube, AU: arbitrary units.

4.1.4 *Twist1* and Prrx1 are co-expressed during development and genetically interact during digit morphogenesis.

During mouse embryonic development, *Twist1* is expressed in mesenchymal cells of the mesoderm and the neural crest. Unlike *Snail1*, and similar to *Prrx1*, *Twist1* is not expressed in primitive streak or the pre-migratory neural crest, but it is expressed in the ingressing mesoderm and in migratory neural crest cells. We have performed double immunolabeling in E9.5 mouse embryos (**Figure 16**) and observed that they are co-expressed in cells of the limb bud, the lateral plate mesoderm, the migratory neural crest cells and branchial arches.

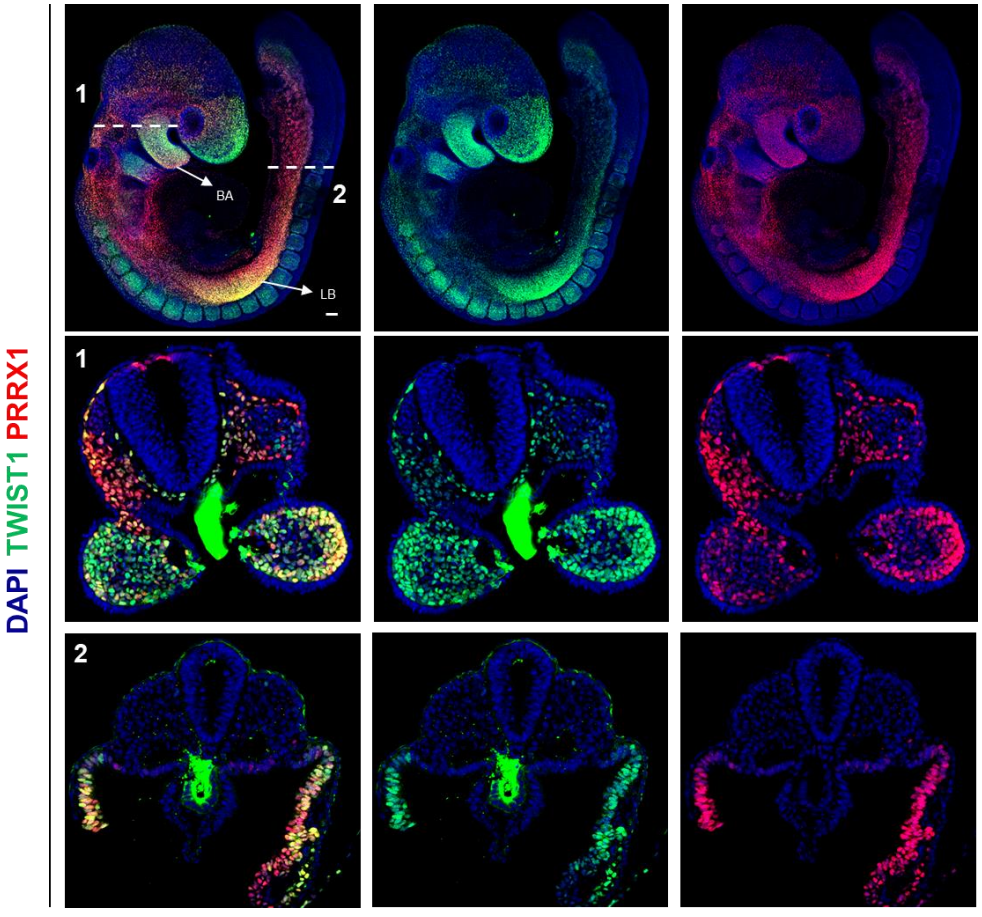


Figure 16| TWIST1 is co-expressed with PRRX1 in early mouse embryos. Immunofluorescence of TWIST1 and PRRX1 in a whole-mounted E9.5 mouse embryo. Both EMT-TFs are expressed in the branchial arches, lateral plate mesoderm, and the limb buds. (1) Section at the level of the first branchial arch, where both EMT-TFs are expressed in the same cells. (2) A section at the level of the tail shows expression of PRRX1 and TWIST1 in the same cells in the lateral plate mesoderm. Scale bar, whole mount: 100 μm, sections: 50μm. BA: branchial arch, LB: limb bud.

In addition, PRRX1 and TWIST1 are expressed in the same territories, and coexpressed in the same cells in the developing limbs of E13 embryos. Interestingly, both transcription factors are coexpressed in the interdigital area mesenchyme, surrounding the mesenchymal condensations that will give rise to the cartilage of the digits (**Figure 17**).

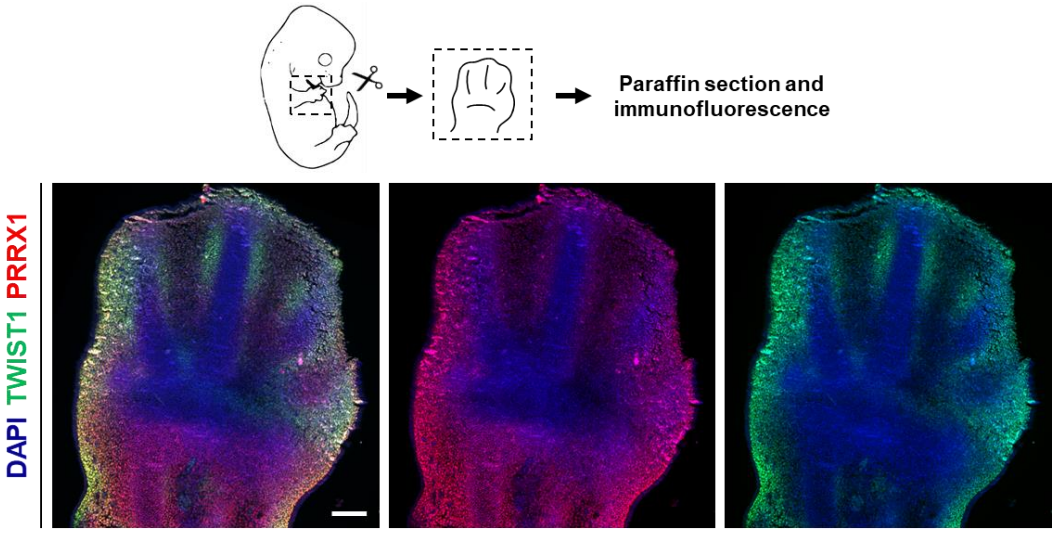


Figure 17| TWIST1 is co-expressed with PRRX1 in the forelimb autopod of E13 embryos. Immunofluorescence of TWIST1 and PRRX1 in E13 forelimbs autopod sections. Both transcription factors are expressed in the mesenchyme of the interdigital area, surrounding the condensations that will give rise to the cartilage of the digits. Scale bar: 100 μm.

As these two EMT-TFs are expressed in the same embryonic territories and cells, we examined whether there could be a genetic interaction between them. We generated a double mutant mouse line, with the combined constitutive mutations for *Twist1* and *Prrx1* (**Figure 18A**) and found that double heterozygous mice are viable and fertile.

In humans, *Twist1* heterozygous mutations are associated with Saethre–Chotzen syndrome. *Twist1* heterozygous mice present similar defects, including 50% penetrance hindlimb polydactyly (Bourgeois et al., 1998; Ghouzzi et al., 2000) and craniosynostosis (Carver et al., 2002). Although *Twist1* heterozygous do not present polydactyly in our genetic background, *Twist1* and *Prrx1* double heterozygous mice show full penetrance polydactyly in the hindlimb (**Figure 18B-C**). In addition, *Prrx1* null mice also bearing a *Twist1* heterozygous mutation, present polydactyly in the hindlimb, which has never been reported in *Twist1* or *Prrx1* single mutants, revealing a genetic interaction between *Prrx1* and *Twist1* in the formation of the digits. Interestingly, *Prrx1* and *Prrx2* double mutants also present

polydactyly (Berge et al., 1998). This suggests a role of *Prrx1* in the formation of the digits that can be compensated either by its paralog, *Prrx2*, or by *Twist1*.

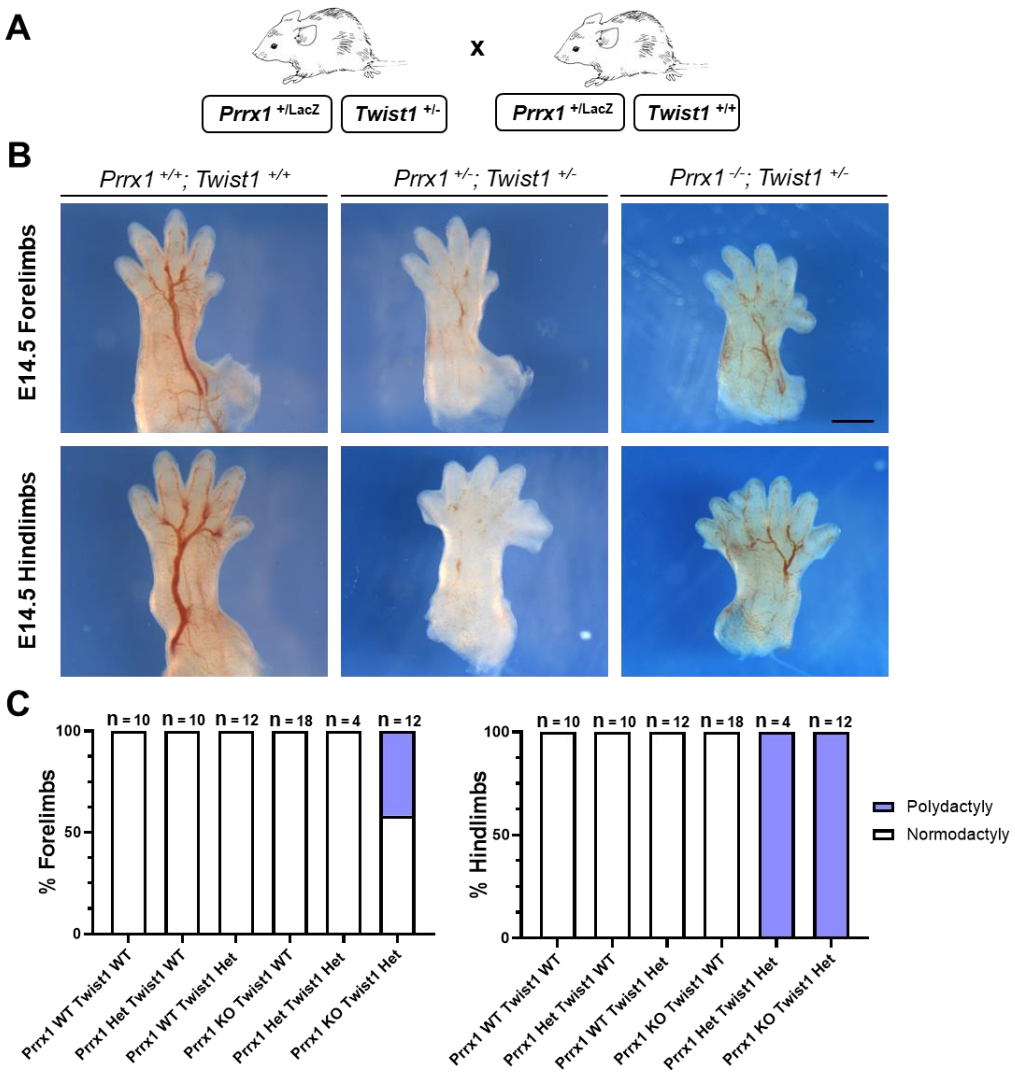


Figure 18| *Prrx1* and *Twist1* genetically interact during digit morphogenesis. (A) Schematic representation of the mattings used to produce the double mutant embryos. *Prrx1* $^{+/-}$*; Twist1* $^{+/-}$ mice are viable and fertile. (B) Forelimbs and hindlimbs dissected from E14.5 embryos. The first panel represents *Prrx1* $^{+/+}$*; Twist1* $^{+/+}$, *Prrx1* $^{+/-}$*; Twist1* $^{+/+}$, *Prrx1* $^{+/+}$*; Twist1* $^{+/-}$ and *Prrx1* $^{-/-}$*; Twist1* $^{+/+}$, which do not have polydactyly. (C) Bar graph representing the penetrance of polydactyly in the mutants from each of the genotypes. Only double mutants show polydactyly, and only *Prrx1* $^{-/-}$*; Twist1* $^{+/-}$, show polydactyly in the forelimbs. This indicates a genetic interaction between *Prrx1* and *Twist1* in the formation of the digits of the forelimb and the hindlimb. Scale bar: 2 mm.

4.1.5 Generation of isoform-specific *Prrx1* mutant mice

PRRX1 has two isoforms generated after alternative splicing of exon 4 in its mRNA (Kern et al. 1992). These isoforms, referred to as *Prrx1L* and *Prrx1S* from here on, contain different transcription regulatory domains (Norris & Kern, 2001a). *Prrx1S* lacks the OAR and AD3 domains, present in *Prrx1L*. In addition to these two previously described isoforms, we have identified a third isoform by western blot (**Figure 19A**), from here on *Prrx1T*. This isoform can be observed by increasing the running time of the western blot, since its size (200aa) is very close to that of *PRRX1S* (217aa), and it is not present in *Prrx1*$^{LacZ/LacZ}$ mutants. *Prrx1T* arises from the retention of intron 3 and gives rise to a truncated protein. Although predicted by Ensembl (transcript ID:ENSMUST00000174397.2; Hunt et al., 2018), had not been identified before. It contains the same domains as *Prrx1S*, and thus, likely has similar functions. Some studies have addressed the study of these isoforms *in vitro* (Norris & Kern, 2001a), but their role *in vivo* remains unknown.

We have generated two new loss of function transgenic mouse lines. For the loss of function of the *Prrx1L* (*Prrx1* $^{\Delta L}$), we have included CRISPR-Cas9-mediated point mutation at the beginning of exon 5, which produces a premature STOP codon. Instead of producing the PRRX1L isoform, these mice will produce a protein similar in size to PRRX1T, also without the OAR and AD3 domains. We have confirmed this using protein extracts from *Prrx1* $^{\Delta L/\Delta L}$ embryos (**Figure 19B**).

To generate a mouse line with the loss of function of both the *Prrx1S* and *Prrx1T* (*Prrx1* $^{\Delta ST}$), we have induced a CRISPR-Cas9 knock in insertion of exon 5 in frame with exon 3 (**Figure 19C**). This mouse line will produce only 3 isoforms which will be identical at the protein level to PRRX1L. We have validated this using protein extracts from E11.5 *Prrx1* $^{\Delta ST/\Delta ST}$ embryos.

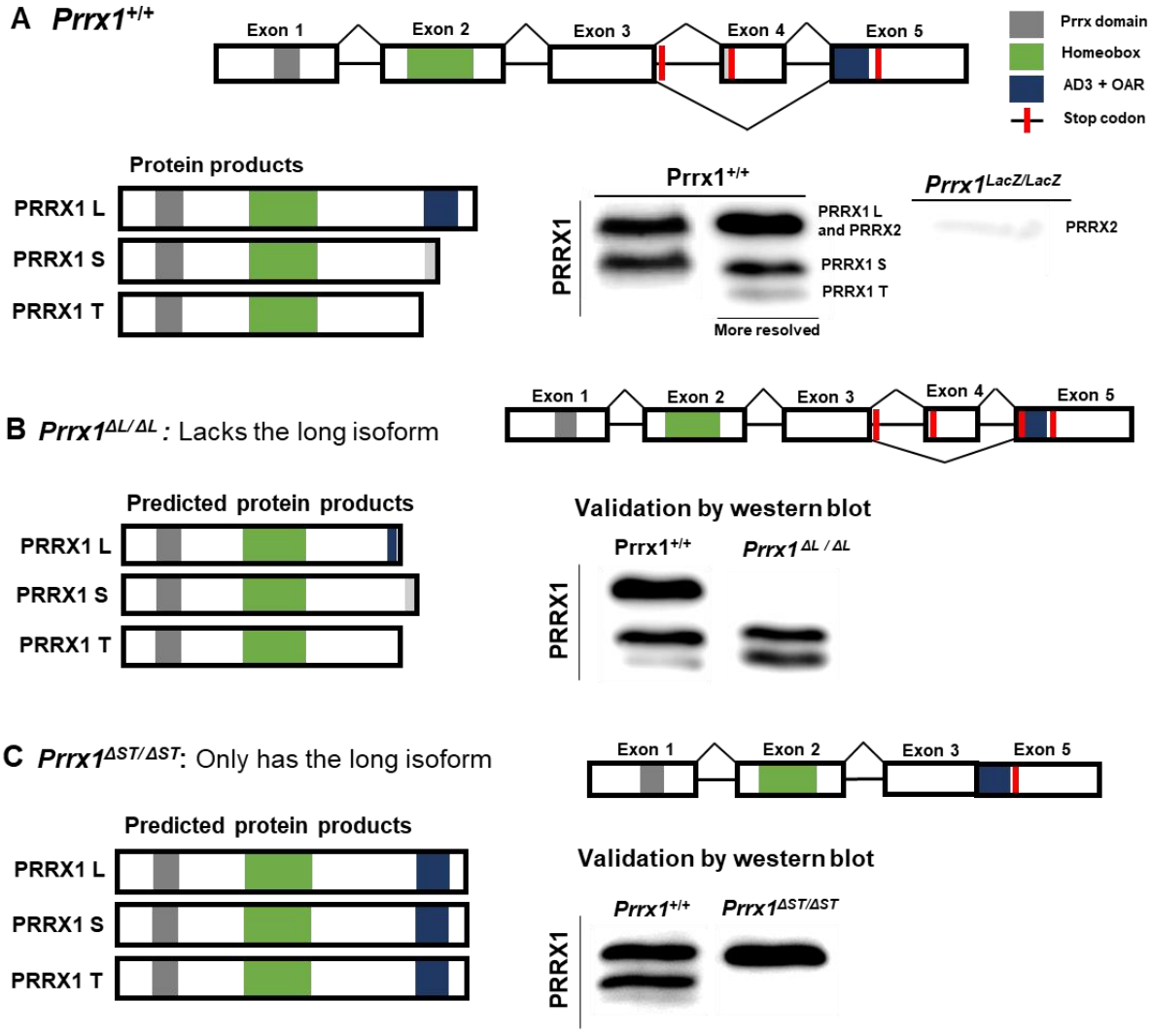


Figure 19| Generation of isoform-specific *Prrx1* mutant mouse models. (A) Schematic representation of the *Prrx1* wildtype locus. Western blot of samples obtained from Prrx1$^{+/+}$ embryos revealed the presence of an additional PRRX1 isoform, Prrx1T. (B) Mouse model bearing a CRISPR-Cas9-mediated point mutation at the beginning of exon 5, generating a STOP codon. This allele does not produce the long isoform. (C) Mouse model with a CRISPR-Cas9-mediated knock-in insertion of exon 5 upstream of exon 3. The STOP codon at the end of the exon 5 ensures that only the long isoform is produced.

To characterize the phenotype of these transgenic lines during mouse development, we compared their phenotype with that of *Prrx1* pan-constitutive mutants. These mutants have a malformed Meckel cartilage, and they lack a *tectum posteriorus*, which is the precursor of the supraoccipital bone, also absent. In addition, they have severe cleft palate,

which is likely the cause of their early postnatal death (Martin et al., 1995). We performed cartilage staining in E14.5 embryos from the isoform-specific mutants and we did not find any obvious defect (**Figure 20A**). Similarly, all these mutants showed normally fused palatal shelves (**Figure 20B**). This indicates that each of the PRRX1 isoforms is enough to fulfil PRRX1 function during craniofacial development.

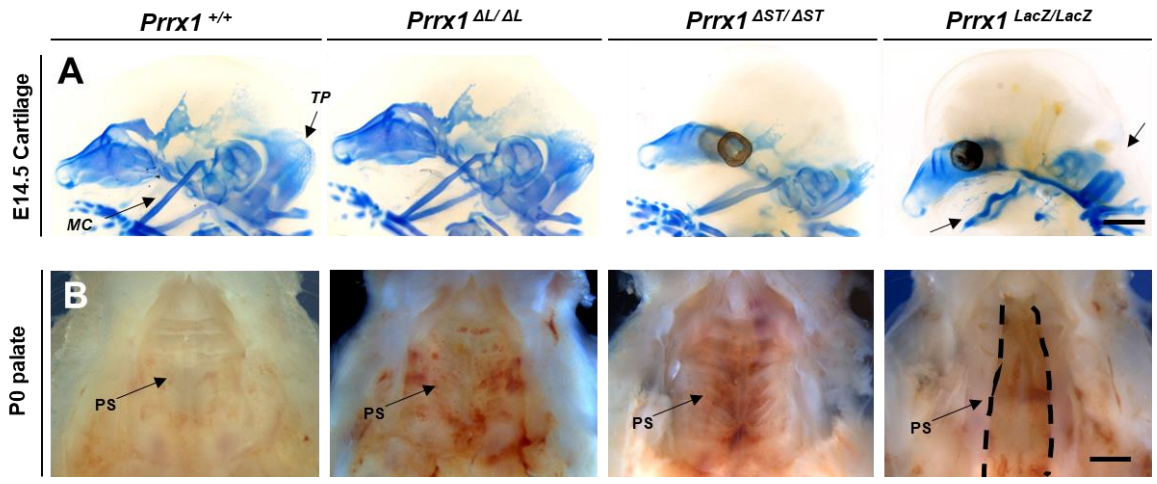


Figure 20| The different PRRX1 isoforms are not essential for the formation of the facial skeleton and the closure of the palate. (A) Cartilage staining of the different mouse models at E14.5. (B) Ventral view of the palate dissected from newborn mice. The conventional mutant shows the described phenotype, including shortened and abnormal Meckel cartilage and absence of the *tectum posteriorus* at E14.5, and cleft palate at P0. The isoform mutants show normal Meckel cartilage and *tectum posteriorus*, and the palate closes properly. Scale bar: 2 mm MC: Meckel cartilage, TP: *Tectum posteriorus* PS: palatal shelves.

Both isoform-specific mutants reach adulthood and are fertile. To study their function after birth and during adulthood, we again compared these mutants with the complete loss of *Prrx1* function. We needed a new conditional mouse model, since the pan-constitutive mutant dies at birth. We generated a new mouse line with loxP sites surrounding *Prrx1* exon 2, from here on referred to as $Prrx1^{flox}$ (**Figure 21A**). We confirmed that the recombination of this new allele leads to the complete loss of function of *Prrx1*, after incubation of 2-cell stage embryos with a cell-permeable version of the Cre recombinase including a nuclear localization sequence (TAT-Cre) (Jo et al., 2001) (**Figure 21B**). Recombination at this very early stage of embryogenesis should phenocopy the pan-constitutive mutant if the recombinant allele is really a null.

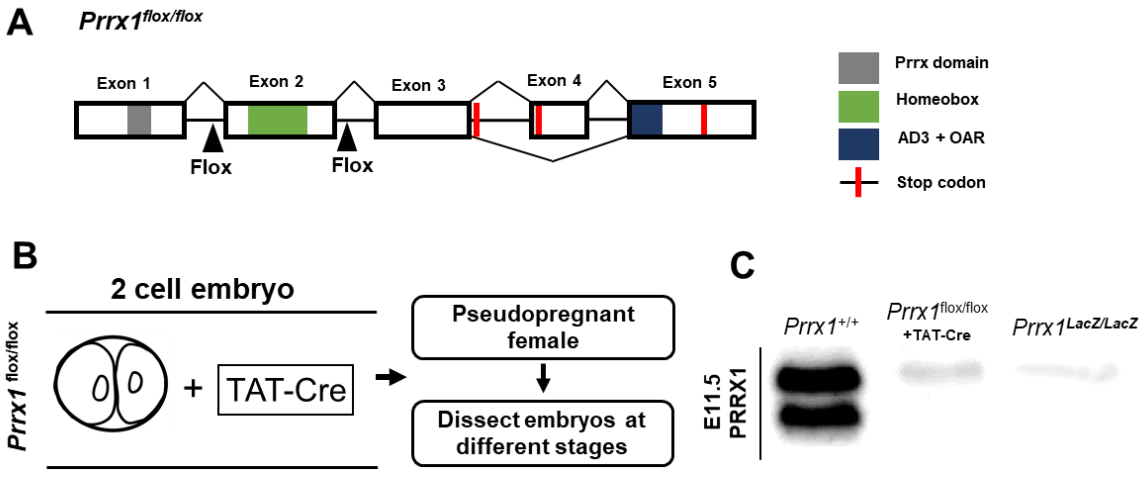


Figure 21| Generation of a *Prrx1* conditional mouse mutant. (A) Schematic representation of the *Prrx1* floxed locus. (B) Strategy for the validation of the *Prrx1* floxed allele. 2-cell stage embryos were treated with a cell-permeable Cre recombinase including a nuclear localization sequence (TAT-Cre). (C) Western blot of protein extracts obtained from *Prrx1*$^{+/+}$, TAT-Cre treated *Prrx1*$^{flox/flox}$, and *Prrx1*$^{LacZ/LacZ}$ E11.5 embryos.

We next dissected E14.5 embryos, performed cartilage staining and found that the *Prrx1*$^{flox/flox}$ TAT-Cre treated embryos have defective Meckel cartilage and *tectum posteriorus* as the *Prrx1* pan constitutive mutants (**Figure 22A**). We also examined the palates of P0 *Prrx1*$^{flox/flox}$ TAT-Cre treated mice and found that they phenocopied the cleft palate defects of the *Prrx1* pan constitutive mutants (**Figure 22B**). In conclusion, the recombination of the *Prrx1*^{flox} allele produces a null allele. To follow the temporal control of *Prrx1* locus recombination, we generated a mouse line that combines the *Prrx1*^{flox} allele with the UBC-cre/ERT2$^{tg/0}$, allowing us to induce recombination by administering tamoxifen and use the *Rosa26-TdTomato* system as a recombination reporter (**Figure 22C**).

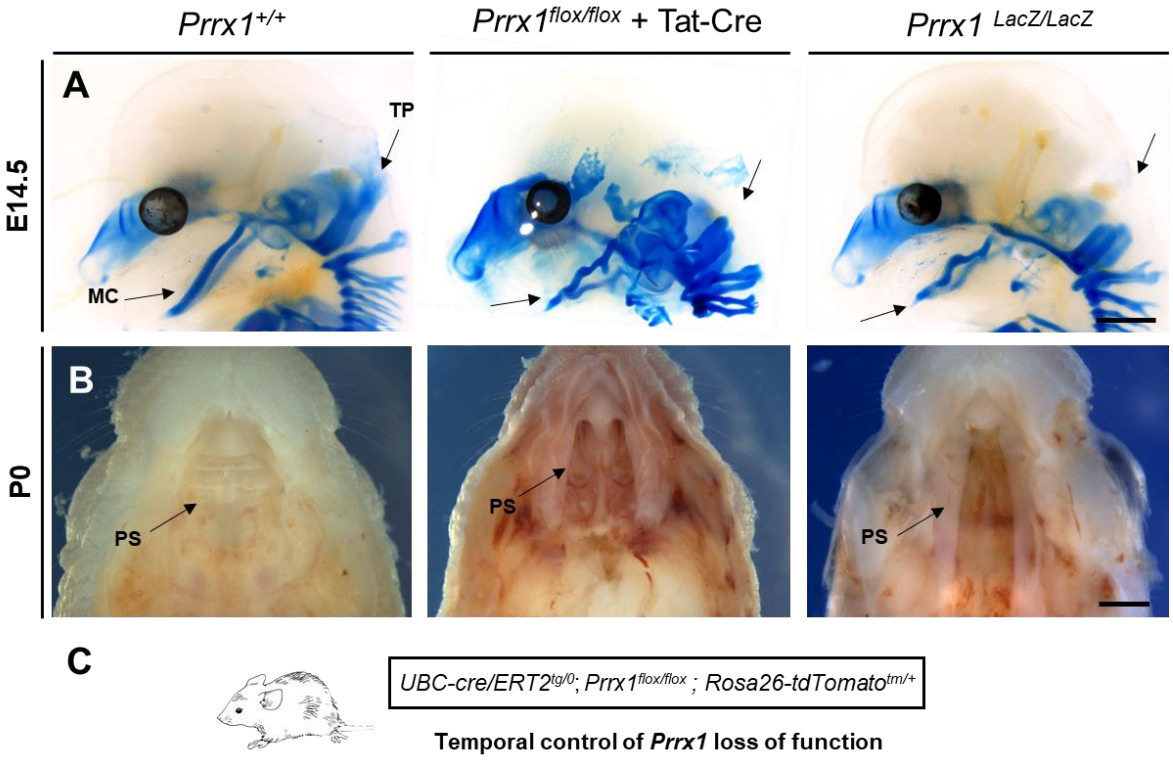


Figure 22| Validation of a *Prrx1* conditional mouse mutant. (A) Cartilage staining of E14.5 embryos. The conditional mutant embryos treated with Tat-Cre show defects in the Meckel cartilage and absence of the *tectum posteriorus*. This phenotype is identical to that described in the *Prrx1* pan-constitutive mutant. (B) Ventral views of the palate dissected from P0 mice. The conditional mutant embryos treated with Tat-Cre show a cleft palate that resembles that observed in the *Prrx1* pan-constitutive mutant. (C) Genotype of the *Prrx1* conditional mouse line. UBC-Cre/ERT2 allows the temporal control of *Prrx1* loss of function. Scale bar: 2mm. MC: Meckel cartilage, TP: *tectum posteriorus*, PS: Palatal shelves.

4.2 The role of Prrx1 in the development and homeostasis of the vasculature

There are evidences in the literature pointing towards a role of *Prrx1* in vascular development (Ihida-Stansbury et al., 2004; Jones et al., 2001). We have observed that embryos dissected at different stages show a variable degree of blood leakage with incomplete penetrance (**Figure 23**), indicating a defect in the development of the vasculature in *Prrx1*$^{LacZ/LacZ}$ embryos. This prompted us to look at the vasculature in the different *Prrx1* mutants.

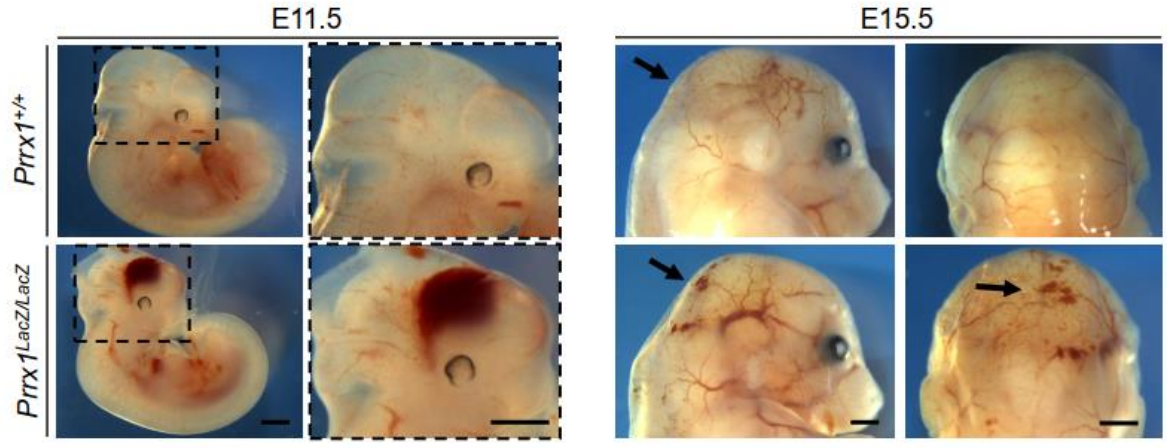


Figure 23| *Prrx1* is necessary for the integrity of the developing vasculature. E11.5 and E15.5 embryos showing blood leakage in the *pan-Prrx1* constitutive mutant. Blood leakage indicates that the vasculature is not properly formed in the *Prrx1* conventional mutant. Scale bar: 2mm

4.2.1 PRRX1 is specifically expressed in the pericytes

To understand the function of *Prrx1* in the development of the vasculature, we first characterized its expression at different developmental stages. In E9.5 embryos, PRRX1 is expressed in the head mesenchyme and the neural crest. We observed that PRRX1 is not expressed by endothelial cells using Isolectin as a marker, but also there were PRRX-positive cells closely associated (**Figure 24A**).

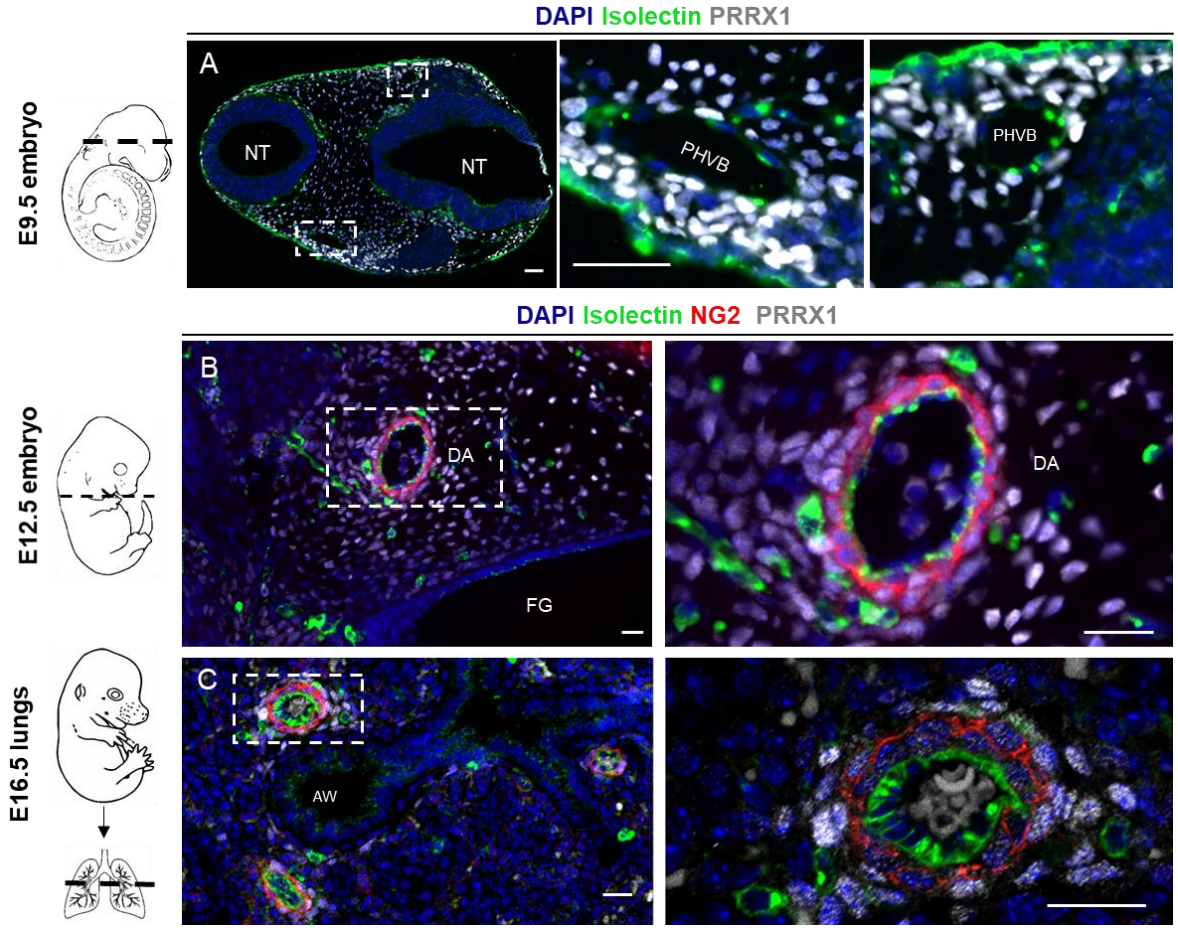


Figure 24. PRRX1 is expressed in the pericytes during mouse development. (A) Transverse section at the level of the head shows expression of PRRX1 in the head mesenchyme. Isolectin is used as a marker of endothelial cells. Although endothelial cells do not express PRRX1, cells closely associated to the endothelial cells do express PRRX1. (B) Transverse section of a E12.5 embryo. PRRX1 is expressed in the pericytes, labelled by NG2 proteoglycan. Higher magnification of the area highlighted shows peri-vascular expression of PRRX1 surrounding the endothelial cells from lateral branches of the dorsal aorta. (C) Section of the E16.5 lung shows expression of PRRX1 in the pericytes and other perivascular cells surrounding the vasculature of the embryonic lung. (E) Higher magnification from the area highlighted in D. Scale bar, A: 50μm. B, C, D and E: 25 μm. AW: Airway, DA: Dorsal Aorta, FG: Foregut NT: Neural tube, PHVB: Primary head vein branch.

Next, we checked E12.5 sections, using NG2 as a marker for pericytes, and we noticed that PRRX1 is expressed in the pericytes and other cells closely associated to the vasculature (**Figure 24B**). Previous studies claimed that PRRX1 was expressed in endothelial cells of E16.5 lungs (Ihida-Stansbury et al., 2004). We examined its expression

in more detail, and we found that, consistent with our previous observations, PRRX1 is not expressed in the endothelial cells of the developing lung, but rather in the surrounding $NG2^+$ pericytes (**Figure 24C**). Pericytes regulate vascular development, and mice lacking pericytes show vascular leakiness and die during development (Hellström et al., 1999; Lindahl et al., 1997; Nikolakopoulou et al., 2019). Thus, PRRX1 expression in the pericytes is consistent with the phenotype observed in *Prrx1* pan constitutive embryos.

To understand the role of *Prrx1* and its isoforms during vascular development, we decided to use the postnatal retina, one of the best known models to study angiogenesis (Fruttiger, 2002; Selvam et al., 2018). The postnatal retina forms a flat vascular bed that starts growing from the optic nerve at P0 and reaches the periphery of the retina at P10. We have used the P6 retina to observe the remodelling of the vasculature and the leading cells at the tip of the vascular sprouts growing towards the periphery. PRRX1 immunofluorescence shows widespread expression in the retina's vasculature (**Figure 25A**), but higher magnification pictures and orthogonal projections reveal that PRRX1 is specifically expressed in the pericytes (**Figure 25B-E**), consistent with our previous observations in the embryo.

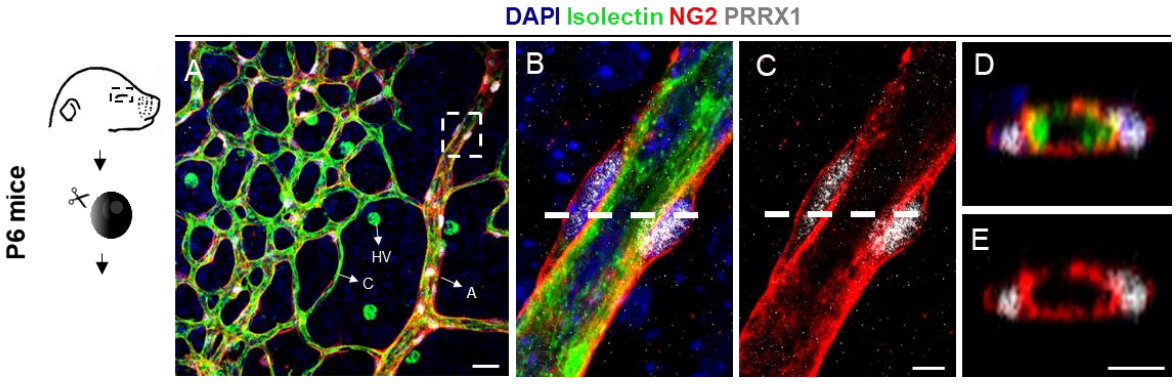


Figure 25. PRRX1 is expressed in postnatal mouse retina pericytes. (A) Flat mount of the postnatal mouse retina, showing widespread expression of PRRX1. (B and C) Higher magnification from area highlighted in A, showing PRRX1 expression. (D and E) Orthogonal projection pictures reveal specific PRRX1 expression in the pericytes. Scale bar: A: 25 μm; (B, C, D and E): 5 μm. C: capillary, HV: hyaloid vasculature, A: artery.

4.2.2 PRRX1 is necessary for the integrity of the vasculature

To induce *Prrx1* loss of function during postnatal retinal vascular development, we administered tamoxifen intraperitoneally in the conditional *Prrx*^{flox}; mutant mice at P3 and P5 and dissected the retina at P6 (**Figure 26A**).

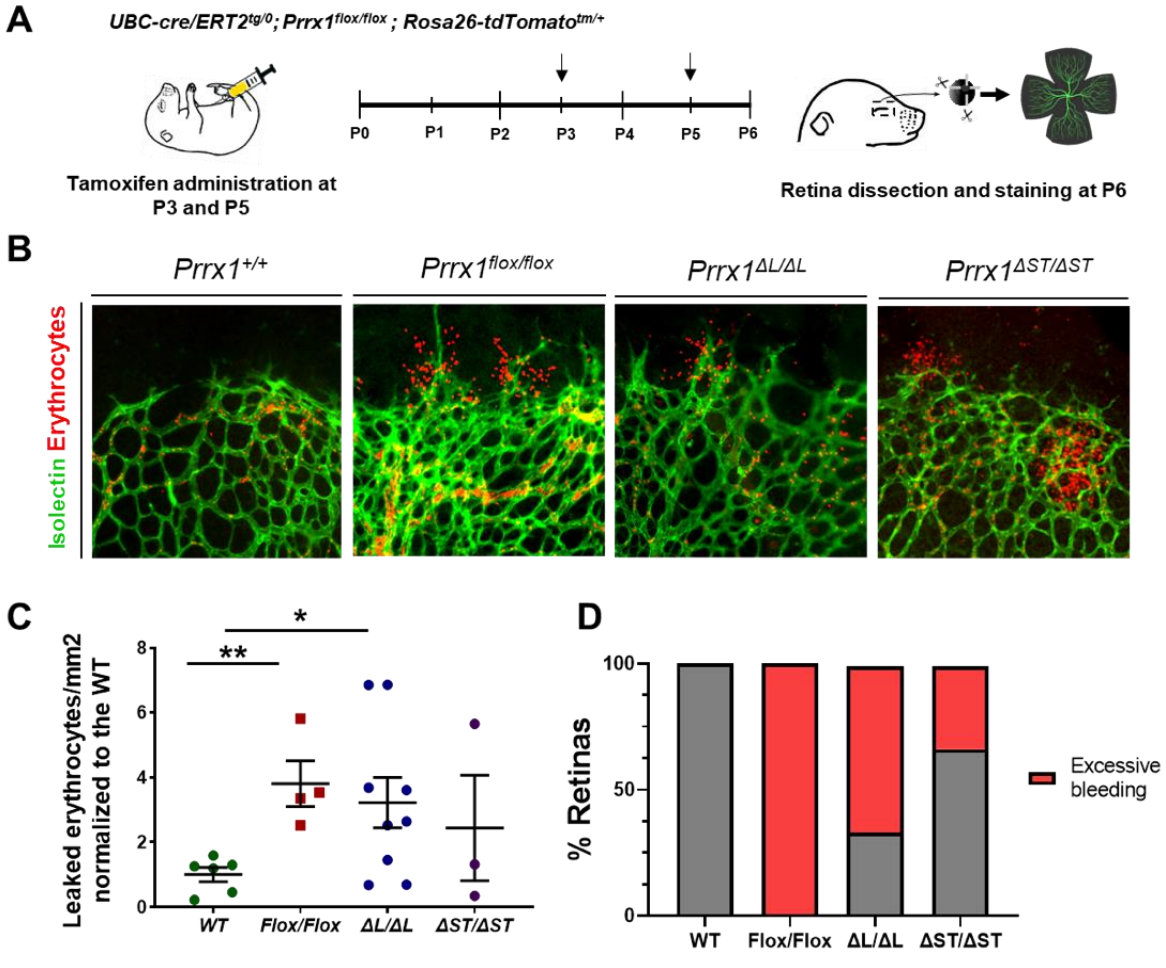


Figure 26| *Prrx1* is required for the integrity of the developing retinal vasculature. (A) Strategy used to study vascular development in the postnatal mouse retina. (B) Flat-mounted retinas from the different mutants, showing erythrocytes and endothelial cells labelled by Ter 119 and isolectin, respectively. (C) Quantification of erythrocytes outside the vasculature at the growing edge. Bars represent mean and SEM, asterisks specify p-value significance in t test (*p < 0.05, **p < 0.01, ***p < 0.001 and **** p<0.0001). (D) Penetrance of the bleeding phenotype in each mouse models. Samples were labelled as excessive bleeding when the number of erythrocytes out of the vessels was higher than twice the average in the wild type.

In these mice, we observed erythrocytes outside the vasculature at the growing edge. These events normally happen with a low frequency in wild type animals, but the leakiness was prominent in the *Prrx1* conditional mutant and also in the isoform mutants (**Figure 26B**), with the latter showing incomplete penetrance (**Figure 26C**).

Interestingly, mice with impaired pericyte recruitment to the developing retina endothelial cells, show a similar kind of haemorrhage (Park et al., 2017). The same phenotype has also been observed in mice with increased VE-cadherin endocytosis, which also show vascular leakiness in the embryo and the lung (Grimsley-Myers et al., 2020). Altogether, our data indicate a defect of vascular integrity in the different *Prrx1* mutants. Since pericytes regulate vascular integrity by interacting with endothelial cells in development and disease (Bergers & Song, 2005), this phenotype is probably arising from a defective interaction between pericytes and endothelial cells.

As previous reports point towards a deficient lung vascularization in *Prrx1* pan-constitutive mutants (Ihida-Stansbury et al., 2004), we wanted to further explore this finding. We dissected lungs from E18.5 *Prrx1* pan-constitutive mutant embryos and did not find any obvious morphological difference with the WT (**Figure 27A**). Contrary to this previous study, we did not find differences in the tissue morphology in sections stained with haematoxylin-eosin (**Figure 27B**), nor differences in the distribution of mural cells or endothelial cells, labelled with SMA and Isolectin, respectively (**Figure 27C-D**).

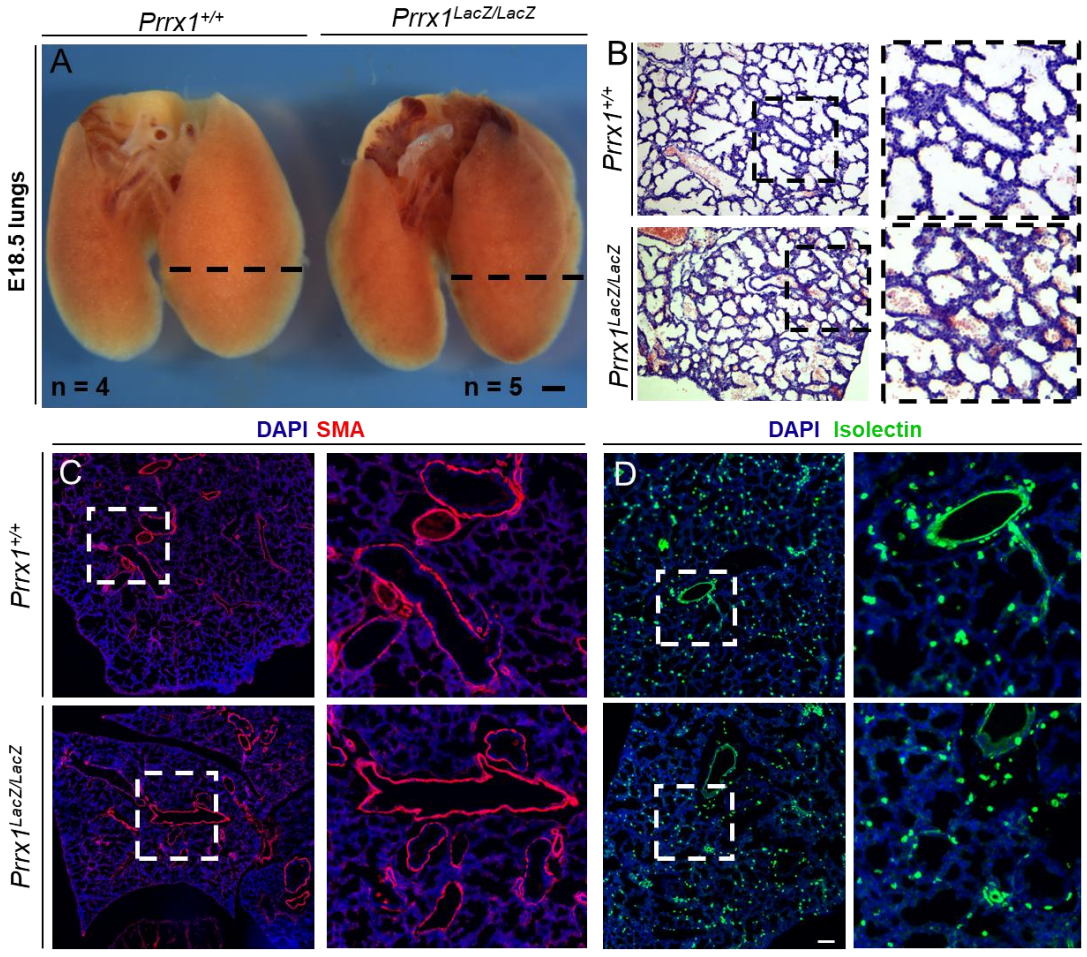


Figure 27| Prrx1 is not required for the formation of the vasculature in the embryonic lung. (A) Bright field picture of the whole-mounted lungs dissected from E18.5 mouse embryos. (B) Haematoxylin and eosin staining of sections obtained from these lungs. Histological analysis reveals no major defects in the pan constitutive *Prrx1* mutant. (C) Mural cell staining using SMA in these sections reveals that the pan-constitutive mutant has a normal coverage of mural cells in the E18.5 lung. (D) Endothelial cell staining using Isolectin shows a similar distribution of endothelial cells in the pan constitutive *Prrx1* mutant and wild type animals. Whole mount morphology was observed in 4 wild type and 5 *Prrx1* mutant littermates. Sections were analyzed for 3 wild type and 2 *Prrx1* mutant samples. Scale bar: A:

We also wanted to investigate whether *Prrx1* and its isoforms played a role in the vasculature in adult mice. We used publicly available single-cell RNA sequencing dataset (Vanlandewijck, et al. 2018a; Vanlandewijck et al. 2018b) and found that in lung and brain, *Prrx1* is expressed in fibroblastic populations, and absent in endothelial cells (**Figure 28A-B**). In the adult brain, *Prrx1* is expressed in pericytes and in venous and arterial smooth muscle cells (SMCs), which are part of the blood brain barrier. In the adult lung, *Prrx1* is

expressed in the cartilage perichondrium, the vascular SMCs, and in the pericytes. In addition, we have checked if PRRX1 is expressed in the adult retina, and we have found specific expression of PRRX1 in the mural cells of the veins (**Figure 28C**), which are part of the blood retinal barrier. We also examined the expression of PRRX1 protein in adult lungs and found specific expression in vascular SMA^{+} cells (**Figure 28D**), confirming the RNA sequencing data and encouraging us to look at the function of *Prrx1* in adult vascular homeostasis.

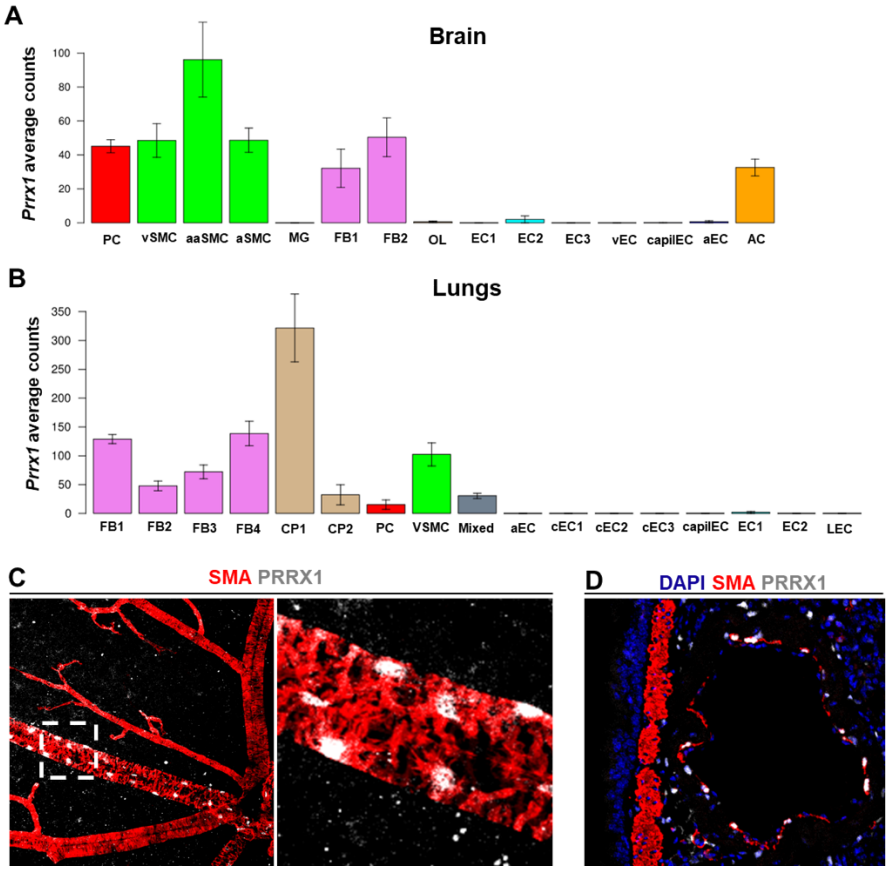


Figure 28| PRRX1 is expressed in mural cells of the blood-brain barrier, the retina, and the lung in adult mice. (A and B) Average reads of *Prrx1* mRNA in vascular and vessel-associated cell types of the mouse adult brain and lung. Data obtained from a publicly available scRNAseq dataset (Vanlandewijck et al. 2018a). (C) PRRX1 expression in the adult retinal vasculature. PRRX1 is specifically expressed in the veins of the retinal adult mural cells. (D) PRRX1 expression in the mural cells of the adult lung. PC: Pericytes; SMC: Smooth muscle cells; MG: Microglia; FB: Vascular fibroblast-like cells; OL: Oligodendrocytes; EC: Endothelial cells; AC: Astrocytes; v: venous; c: capillary; a: arterial; aa: arteriolar; CP: Cartilage perichondrium; L: Lymphatic.

To examine Prrx1 function in the adult lung, we used 8-week conditional mouse mutants (*UBC-cre/ERT2$^{tg/0}$; Prrx1$^{flox/flox}$; Rosa26-tdTomato$^{tm/+}$)* and administered tamoxifen twice a week for four weeks, to induce a complete loss of PRRX1 (**Figure 29A**). We found internal haemorrhage in conditional mutant mice (**Figure 29B, C**) and also in those lacking *Prrx1* isoforms (**Figure 29D, E**).

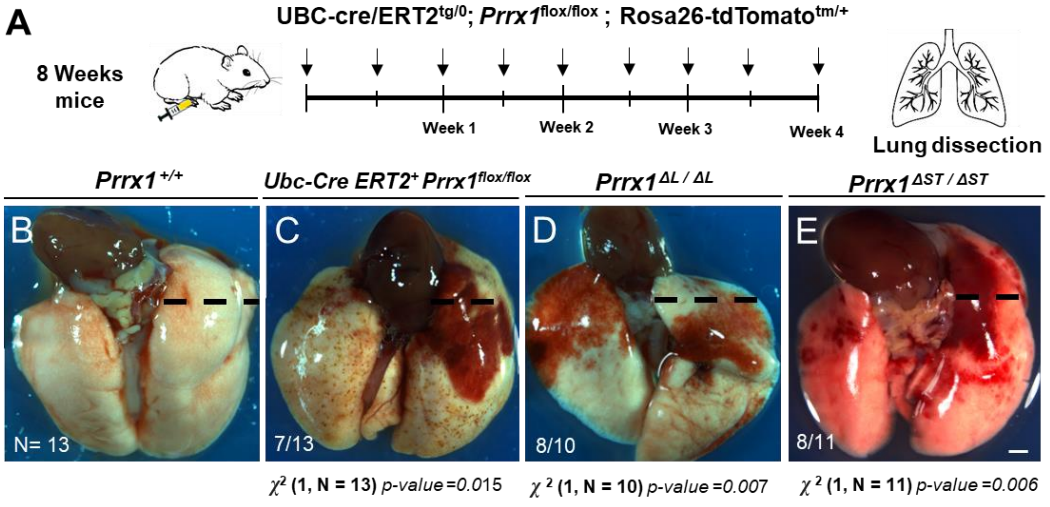
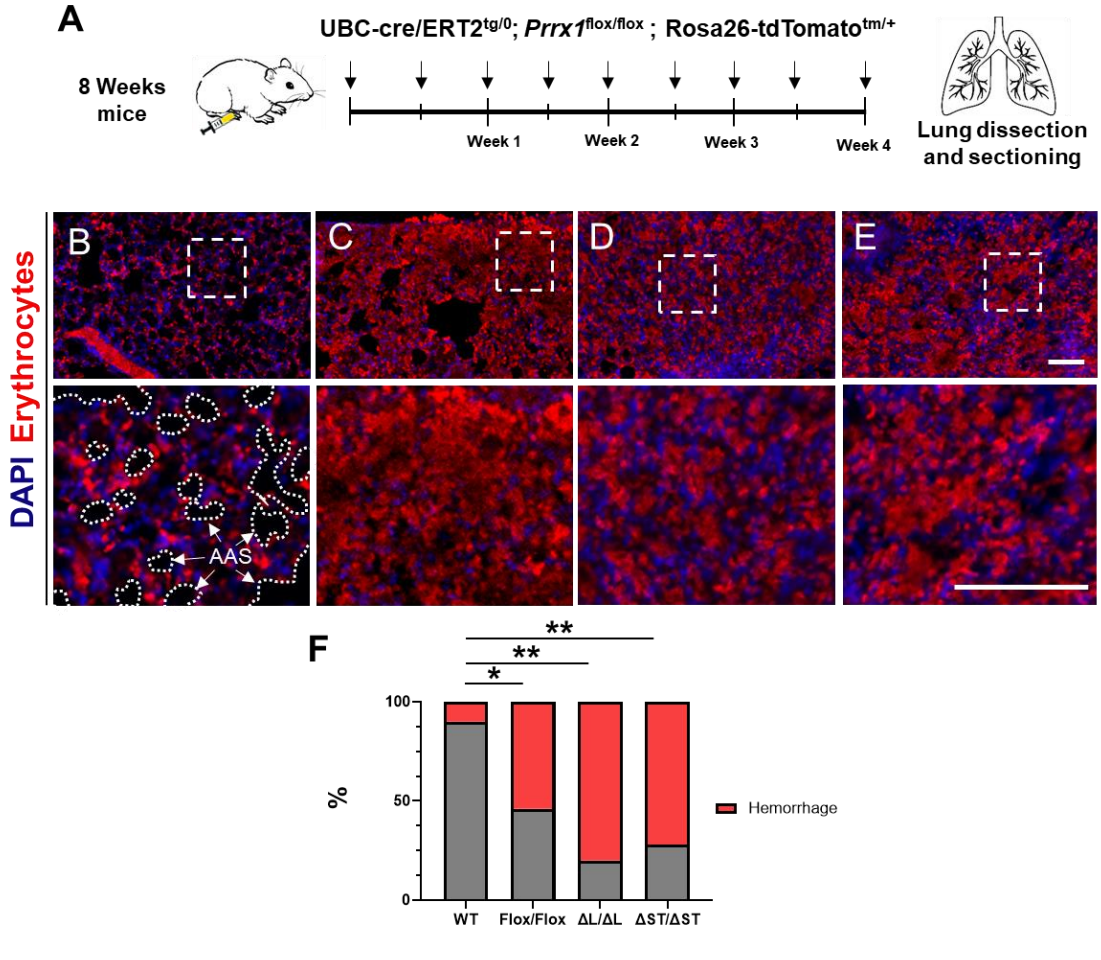


Figure 29| Prrx1 is required for vascular homeostasis of the adult lung. (A) Schematic representation of the tamoxifen administration protocol used in the Prrx1 conditional null mice. (B-E) Bright field picture of the whole mount lungs dissected from adult mice. All Prrx1 mutant adult mice display some degree of pulmonary haemorrhage. Scale bar: 2mm

Sections obtained from these lungs revealed that erythrocytes occupy the alveolar space in the mutants (**Figure30 A-E**). Chi square test of the frequency of this phenotype for each of the mutants revealed that they are statistically significant (**Figure30F**), indicating that PRRX1 is required for vascular homeostasis in the adult lung.



a

Figure 30| Prrx1 mutants show leakiness of blood into the alveolar air space. (A) Schematic representation of the tamoxifen administration protocol used in the Prrx1 conditional null mice. (B-E) Sections taken from these lungs revealed that the erythrocytes in the alveolar air space in the mutants. (F) Frequency distribution of the internal haemorrhage phenotype in the different Prrx1 mutants. Statistics were performed by chi-squared test (*p < 0.05, **p < 0.01, ***p < 0.001). Scale bar: 100um, AAS: Alveolar air space

4.2.4 PRRX1 is required for neoangiogenesis in the *in vivo* matrigel plug assay

After the finding that PRRX1 is required for vascular homeostasis in adult mice, we decided to perform and assay to test its potential role in neoangiogenesis. We used the matrigel plug assay (Passaniti et al., 1992), where Matrigel mixed with the angiogenic factors FGF2 and VEGF is injected subcutaneously into mice, and vascularization examined 1 to 3 weeks later (Kano et al., 2005) (**Figure 31A**). We found that Matrigel plugs that do not contain

angiogenic factors are not vascularized (**Figure 31B**), but 1 week after plug implantation in the presence of angiogenic factors, pericytes, endothelial cells and macrophages invaded the plug (Tigges et al., 2008). Immunofluorescence analyses confirmed the presence of these cell types and also that, as expected from our previous data, $NG2^+$ pericytes express PRRX1, while $CD31^+$ endothelial cells and $CD45^+$ immune cells do not (**Figure 31C-D**). Once we had this model optimised, we used it to evaluate PRRX1 function in neoangiogenesis.

To study the function of PRRX1 in neoangiogenesis in this system, we used the *Prrx1* conditional mouse model (*UBC-cre/ERT2$^{tg/0}$; Prrx1$^{flox/flox}$; Rosa26-tdTomato$^{tm/+}$)* starting tamoxifen administration 1 week before injecting the matrigel plug and continued the administration twice a week until we dissected the plug for analysis (**Figure 32A**). In plugs dissected 3 weeks after implantation, we observed a drastic decrease in the overall number of cells in the plug, and in particular, in the population of pericytes and endothelial cells were compared with plugs implanted in control mice (**Figure 32B-C**). We also examined plugs implanted into mutants lacking the long isoform and found that they show the same phenotype with less severe effects. This indicates that PRRX1 has a role in neoangiogenesis affecting both the amount of pericytes and endothelial cells in matrigel plugs, suggesting both cell autonomous and non-cell autonomous effects in pericytes and endothelial cells, respectively. This is very promising, and given the intricate relationship between these two cell populations (Armulik et al., 2005; Bergers & Song, 2005), we decided to investigate the molecular mechanisms behind both actions.

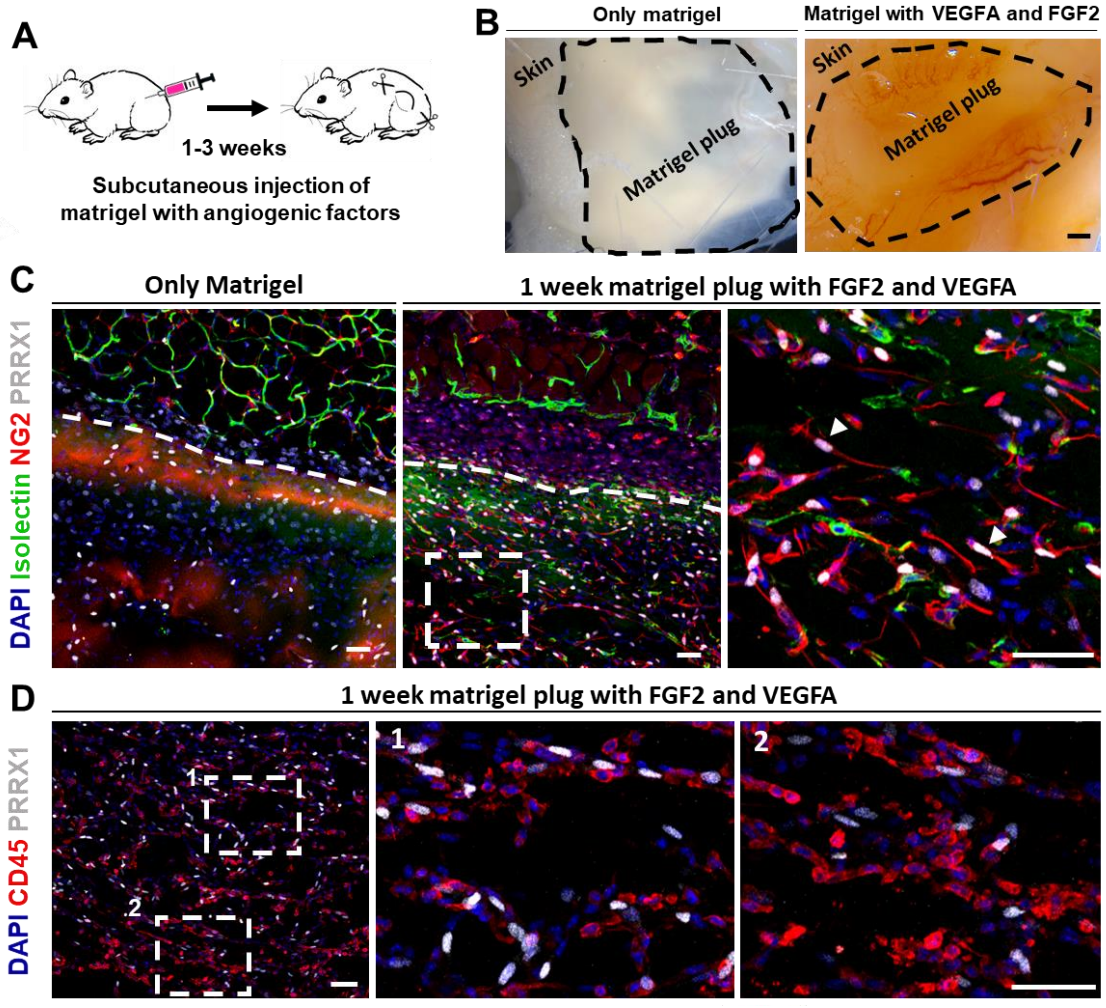


Figure 31| Matrigel plugs as an *in vivo* vasculogenesis assay. (A) Schematic representation of vasculogenesis assay. (B) Matrigel plugs dissected 2 weeks after implantation. Plugs without angiogenic factors (left) are not vascularized, while in those containing the angiogenic factors FGF2 and VEGFA there is *de novo* formation of blood vessels. Scale bar: 2mm (C) Sections of matrigel plugs in the absence (left) or presence of angiogenic factors (right). The dashed line delineates the boundary between the dermis (upper part) and the matrigel plug. The pericytes (NG2+) that invade the plug (arrowheads in the higher magnification picture) express PRRX1. (D) Sections from angiogenic matrigel plugs showing CD45+ immune cells, which do not express PRRX1**.** Scale bar: 50 μm.

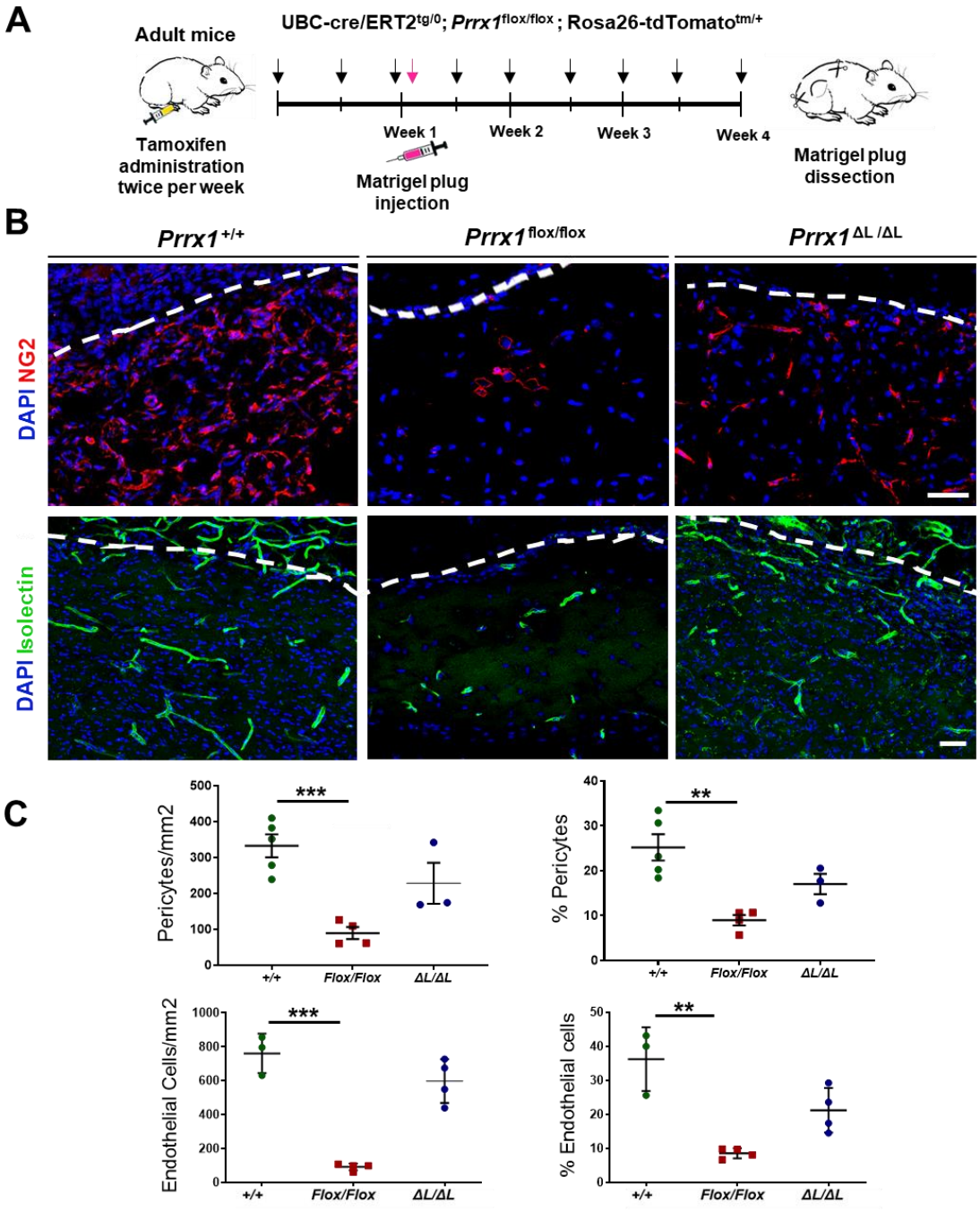


Figure 32| *Prrx1* mutants have less pericytes and endothelial cells invading the matrigel plug. (A) Schematic representation of the matrigel plug neoangiogenesis assay used in *Prrx1* conditional null mice. (B) Sections of the matrigel plugs, with pericytes labelled with NG2 and endothelial cells with Isolectin (C) Number of pericytes or endothelial cells per squared millimetre and their relative percentage with respect to all cells. Bars represent mean and SEM, asterisks indicate p-value in t test (*p < 0.05, **p < 0.01 and ***p < 0.001).

4.2.5 The mechanism of *Prrx1* function in the pericytes

As a transcription factor, PRRX1 performs its function by regulating gene expression. To understand its role in the pericytes, we decided to investigate its potential downstream effectors in this cell population. Thus, we decided to perform a transcriptomic analysis comparing *Prrx1* wild type with *Prrx1* null pericytes. With this purpose, we designed a new protocol to isolate pericytes from the matrigel plugs one week after implantation, disaggregating the matrigel with a non-enzymatic solution, labelling the pericytes with an anti-PDGFRβ antibody conjugated with APC and sorting the positive population by fluorescence-activated cell sorting (FACS; **Figure 33A-B**). qPCR analysis validated the protocol, showing that the isolated cell population was enriched in pericyte markers and had a very low presence of immune and endothelial cell markers (**Figure 33C**) (see next page).

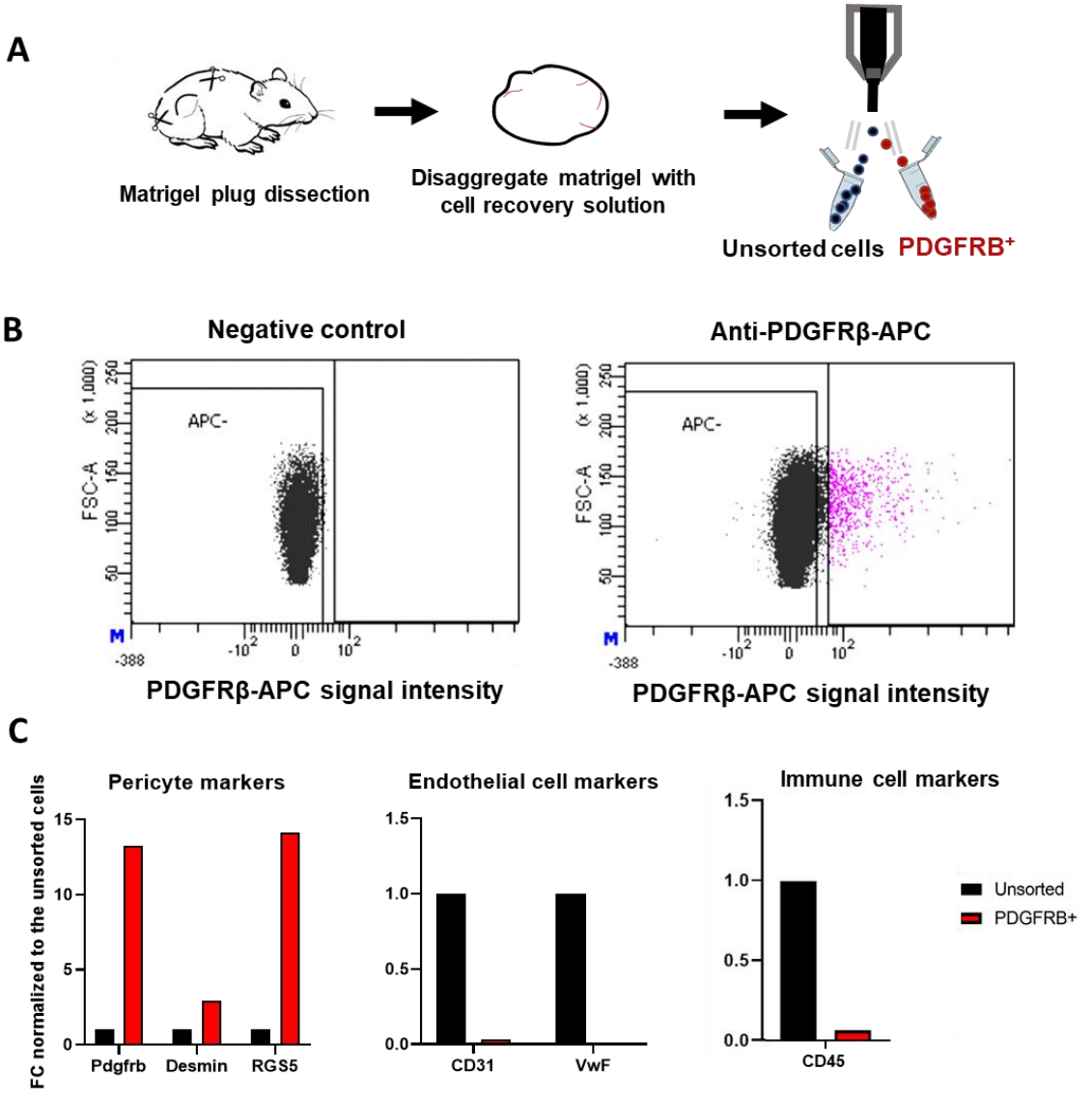


Figure 33| Isolation of pericytes from matrigel plugs. (A) Schematic representation of the strategy used to isolate pericytes from matrigel plugs. (B) FACS graph showing APC signal intensity of unlabelled cells and cells labelled with the anti-PDGFRβ-APC antibody. (C) qPCR for different pericytes, endothelial and immune cell markers used to check the pericyte isolation assay, resulting in a highly enriched population. FC: Fold change. APC: Allophycocyanin.

Next, we used *Prrx1* conditional knock out mice and starting administering tamoxifen 1 week before injecting the matrigel (**Figure 34A**). We used Tomato expression as a reporter of recombination, and selected cells that expressed both PDGFRβ and Tomato (**Figure 34B**).

Interestingly, we found that, in contrast to the situation in 3-weeks plugs, the percentage of pericytes was similar in the plugs implanted in WT or *Prrx1* conditional mutant mice 1 week after plug implantation (**Figure 34C**), suggesting that the same number of pericytes invade the plug than later diie of proliferate less.

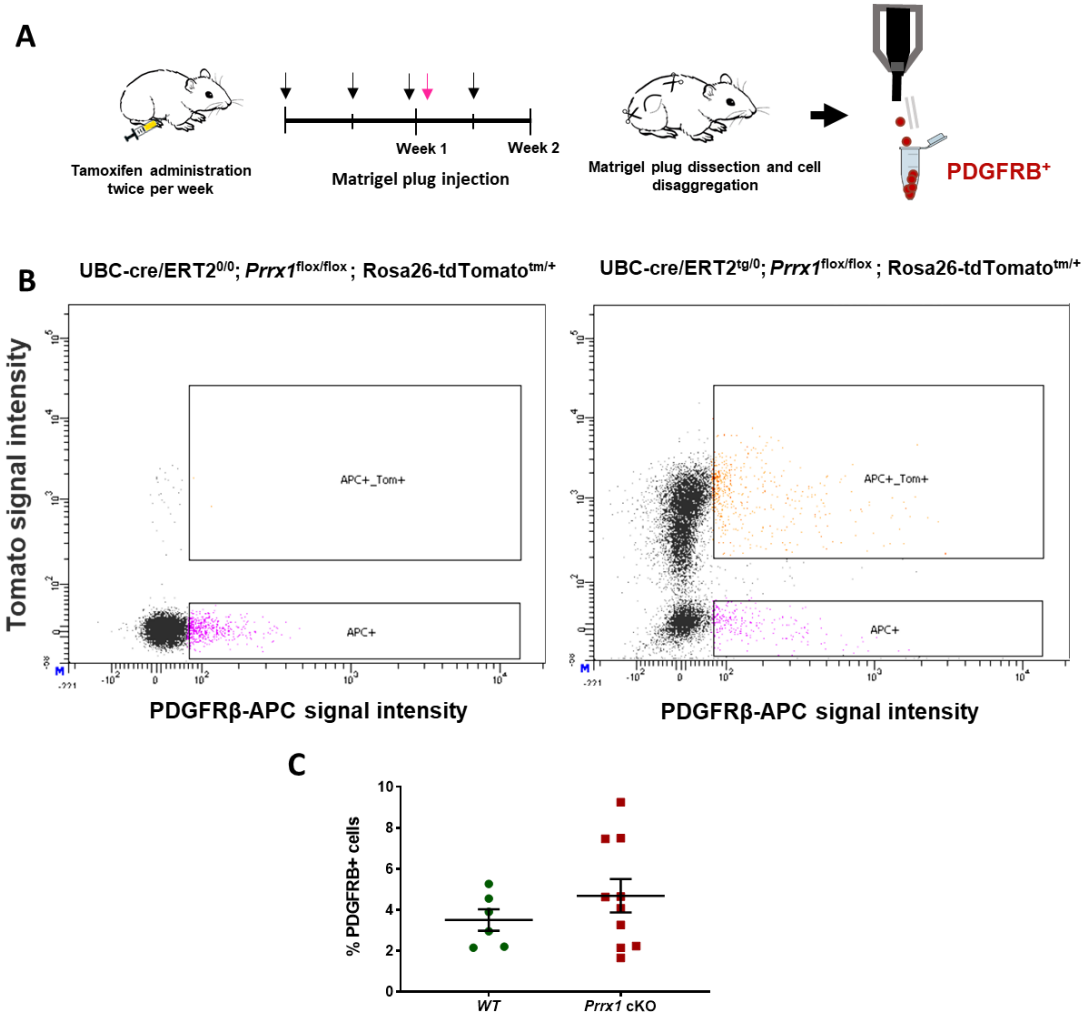
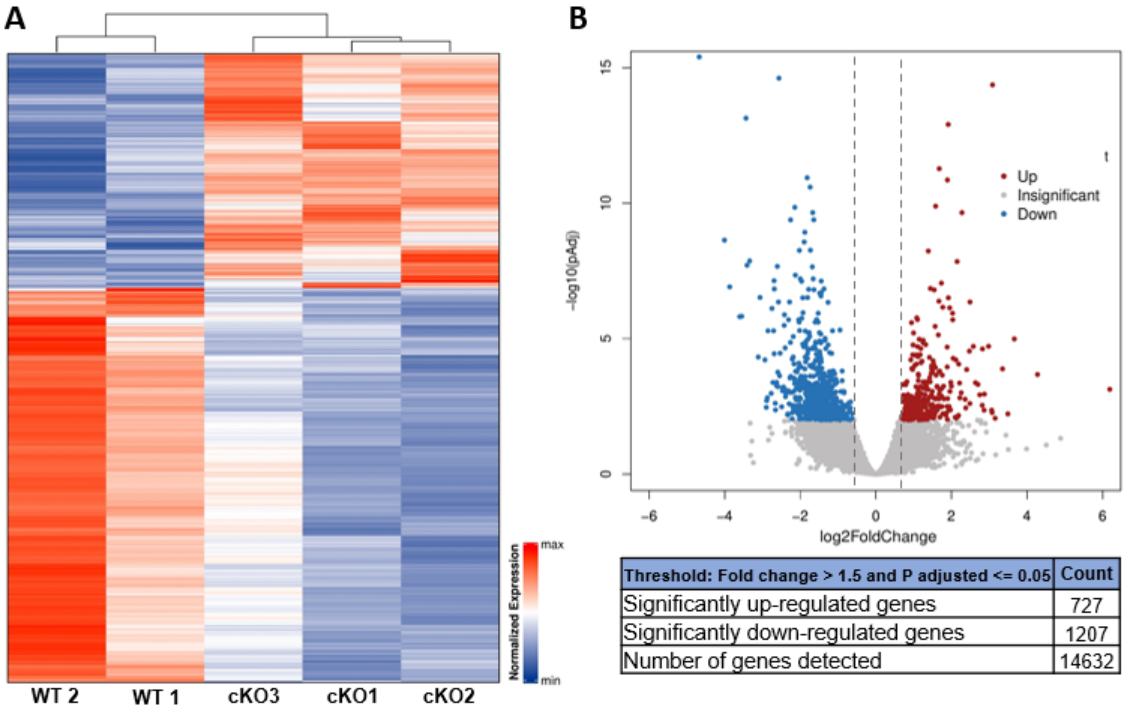


Figure 34| Strategy for the isolation and sequencing of pericytes with PRXX1 loss of function. (A) Schematic representation of the strategy used to isolate pericytes from matrigel plugs. (B) Isolation of pericytes from plugs implanted in *Prrx1* WT and *Prrx1* mutant mice. The selection of PDGFRβ^{+} *Tomato*$^{+}$ cells allows the specific selection of pericytes that have undergone recombination, and therefore, are mutant for *Prrx1*. (C) Percentage of pericytes in all the samples processed with this protocol. There is no evident change in the percentage of pericytes that invade the plug after 1 week. We used these samples for transcriptomic analysis.

We next performed transcriptomic analysis of three *Prrx1* WT and three *Prrx1* cKO samples. After preliminary analysis, one of the WT samples was excluded due to a potential contamination, revealed by its low *Prrx1* and pericyte marker expression. Differential gene expression analysis was performed, and hierarchical clustering reveals that samples from the same group cluster together (**Figure 35A**). We obtained a higher number of significantly downregulated than upregulated genes, consistent with PRRX1 behaving as an activator (**Figure 35B**).



Threshold: Fold change > 1.5 and P adjusted <= 0.05	Count
Significantly up-regulated genes	727
Significantly down-regulated genes	1207
Number of genes detected	14632

Figure 35| Differential gene expression after pericytes RNA-seq analysis. (A) Cluster analysis heatmap from differentially expressed genes. Samples from the same condition behave similarly. (B) Volcano plot of the differentially expressed genes. There are substantially more downregulated genes, consistent with the described function of *Prrx1* as a transcriptional activator.

We next carried out a gene ontology enrichment analysis of the significantly downregulated or upregulated differentially expressed genes (DEGs) using Toppgene (J. Chen et al., 2009). (**Figure 36**).

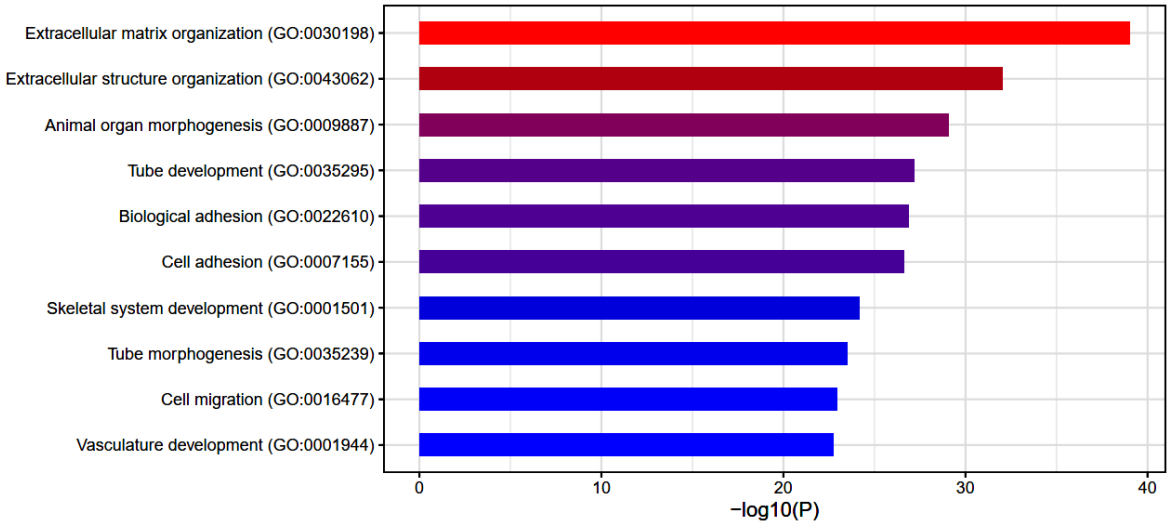


Figure 36|. Gene ontology enrichment analysis of the differentially downregulated genes. Bar plot of the most significant GO terms of the downregulated genes, which are mostly related with ECM composition, cell adhesion, and migration, consistently with previously described *Prrx1* functions.

We found that many genes involved in extracellular matrix organization (GO:0030198), including several collagen and matrix metalloprotease genes were downregulated. This is consistent with the role of *Prrx1* as an EMT-TF (Ocaña et al., 2012) and with previous reports showing that *Prrx1* can activate *collagen I* (*Col1a1*) transcription in hepatic stellate cells, which are liver-specific pericytes (Jiang & Stefanovic, 2008). As such, *Col1a1* is included in our donwnregulated DEGs in pericytes. *Col3a1, Col1a2*, *Col6a1* and *Col6a2*, all highly expressed in pericytes (Brown et al., 2015) are also significantly downregulated in *Prrx1* defective pericytes, *Col3a1* in particular. Furthermore, *Pdrgfb*, known to be crucial for pericyte recruitment, is also significantly downregulated.

In order to validate if *Prrx1* is required for the expression of *Col3a1* and *Col1a1* in the pericytes, we isolated pericytes from matrigel plugs (**Figure 37A**). We compared the expression by RT-qPCR of *Col1a1* and *Col3a1* in the WT and *Prrx1* mutant pericytes and found a significant decrease in the expression of both collagens, indicating that *Prrx1* is necessary to maintain high levels of expression of these collagens in pericytes. (**Figure 37B**). To test if PRRX1 binds to *Col3a1* promoter, we first looked for potential PRRX1 binding sites (TAATKDS) near the human and mouse *COL3A1* promoter region. We found 3 regions that have two or more potential PRRX1 binding sites in the human *COL3A1* locus

(BS1, BS2 and BS3), with BS3 being conserved between human and mouse gene. We also used a region on intron 3 without potential binding sites as a negative control (**Figure 37C**). We carried out the ChIP assay using BT-549 cells, which express high levels of PRRX1, and found significant enrichment of binding to BS3 (**Figure 37D**). Altogether, this data shows that PRRX1 can bind the *Col3a1* promoter region, suggesting that PRRX1 may directly activate the expression of *Col3a1* and *Col1a1* in pericytes.

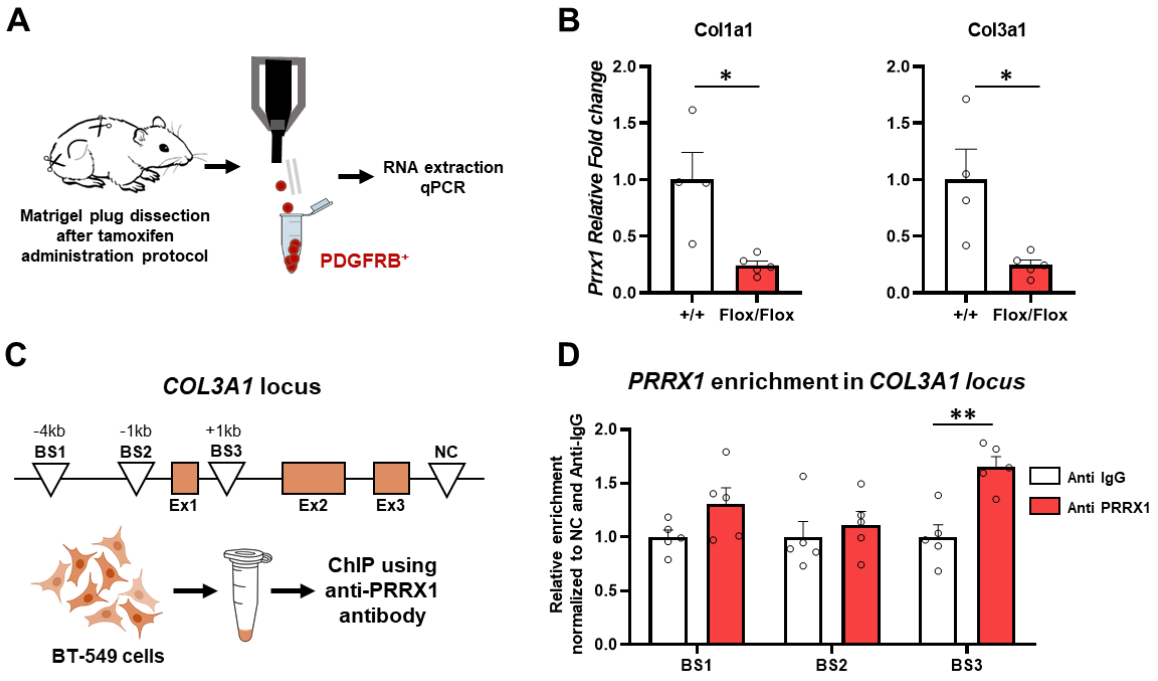


Figure 37| Validation of *Col1a1* and *Col3a1* as downstream targets of *Prrx1* in the pericytes. (A) Samples for the validation of potential targets of PRRX1 in the pericytes using the same protocol as that to obtain samples for RNA sequencing. (B) RT-qPCR for *Col3a1* and *Col1a1* in samples obtained from WT and *Prrx1* conditional mutants. There is a significant decrease in the expression of both collagens in the mutant pericytes. (C) Schematic representation of the human *COL3A1* locus. We found three regions with potential PRRX1 binding sites (TAATKDS) and tested for binding using the ChIP assay in BT-549 cells, with high levels of PRRX1 expression. A region on intron 3 without potential binding sites was used as a negative control (NC). (D) Relative enrichment of PRRX1 binding to the three potential sites, normalized to the negative control region and to the anti-IgG controls. Binding site 3 (BS3) shows a significant enrichment, which indicates t that PRRX1 can bind *Col3a1* promoter, potentially activating its expression.

On the other hand, analysis of the significantly upregulated DEGs shows a robust activation of immune response genes (**Figure 38**). Pericytes are known to secrete cytokines and to express cell surface markers that allow a crosstalk with immune cells (Navarro et al., 2016). Our data suggests that *Prrx1* could mediate a swtich that activates an immunogenic cascade in the perictytes, which could in turn influence vascular development in different contexts (Albini et al., 2018; Bruno et al., 2014).

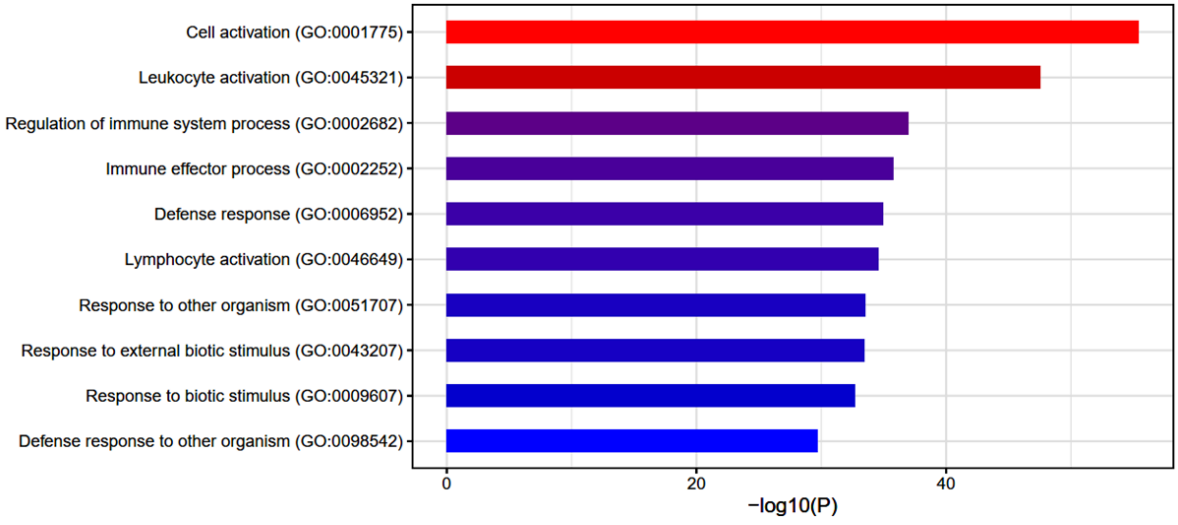


Figure 38|. Gene ontology enrichment analysis of the differentially upregulated genes. Bar plot of the most significant GO terms of the upregulated genes. Most upregulated genes are related with an activation of the immune response.

Discussion

During embryonic development, cells need to move from their place of origin to their destination. To do that, epithelial cells need to undergo EMT, lose their epithelial phenotype and acquire mesenchymal features. After migration, cells frequently need to lose their mesenchymal phenotype and differentiate. For this reason, cell migration during development is a tightly regulated process, controlled by extracellular signals from the cell microenvironment which, in turn, modulate the expression of EMT-TFs. Different cell populations express different combinations of them, leading to a wide variety of cellular phenotypes (Thiery et al., 2009; Yang et al., 2020b). Despite the accumulated knowledge in the last few decades, the genetic interactions between different EMT-TFs in development is still poorly understood. In this work, we decided to systematically assess potential genetic interactions among several EMT-TFs, namely *Snail1, Twist1* and *Prrx1* genes, and study their impact on the development of several EMT-TF expression territories, including neural crest derived craniofacial structures and the limb. In addition, we have generated a series of transgenic mice mutant for different *Prrx1* alleles, allowing the functional analysis of different isoforms and importantly, the role of PRRX1 in adult mice.

5.1 The relationship between *Prrx1* and *Snail1* during mouse development

Both *Prrx1* and *Snail1* are expressed in mesenchymal cell populations during embryonic development, but their expression pattern significantly differs. *Snail1* is expressed earlier, in cell populations that will undergo EMT, including the epiblast cells that ingress through the primitive streak and the premigratory neural crest (Nieto et al., 1992). Both the mesoderm and the migratory neural crest will later express *Prrx1* (Cserjesi et al., 1992; Ocaña et al., 2012; Soldatov et al., 2019). In this work, we have confirmed that this expression pattern is also true at the protein level. In addition, we have generated and validated a mouse model that combines the constitutive mutation of *Prrx1* with the conditional mutation *Snail1* in the neural crest. We have found that a *Snail1* homozygous mutant background enhances the described *Prrx1* mutant phenotype in the Meckel Cartilage. The neural crest-specific homozygous *Snail1* mutation in a *Snail2* null background also generates a phenotype in the Meckel cartilage that is not present in any of the single mutants (Murray et al., 2007). Similarly, double *Prrx1* and *Prrx2* mutants also have a more defective Meckel cartilage than *Prrx1* mutants (Lu et al., 1999a). Interestingly, despite the combined loss of these EMT-TFs (*Snail1/2* or *Prrx1* and *Snail1*), the neural crest still forms and migrates in these mutants.

This can be explained by the spatiotemporal limitation of the driver to generate the neural crest-specific deletions, Wnt1-Cre-induced recombination, which takes place when the pre-migratory neural crest GRN has already been stablished, including the expression of *Snail1*, and in a territory which does not include the whole neural crest population. Thus, depletion of SNAIL1 protein does not happen until the neural crest has delaminated from the neural tube, explaining why the first steps of neural crest formation, including the delamination and the initial migration do not show evident defects. However, these conditional *Snail1* mutants show defects in neural crest derivatives, such as a shortened snout in the single mutant, and the enhanced phenotype in the Meckel cartilage in the *Prrx1* and *Snail1* combined mutant. This reveals a role of *Snail1* in the mouse neural crest, which is consistent with the function of the *Snail* gene family in other vertebrates (Nieto et al., 1994). An analysis of an earlier function of *Snail1* in the neural crest of the mouse will require the use of genetic tools that allow an earlier deletion of *Snail1* without causing mesoderm-derived defects.

Single mutants in lamprey, zebrafish, *Xenopus* or chicken show a much stronger phenotype in neural crest migration or delamination. Thus, a non-excluding possibility for the lack of phenotype in mouse mutants is that the mammalian neural crest GRN is very robust (Barriga et al., 2015) and different EMT-TFs cooperate and can compensate for each other´s loss. This is in agreement with the enhanced phenotypes observed after deleting different combinations of EMT-TFs, as observed previously (Lu et al., 1999a; Lu et al., 1999b; Murray et al., 2007) and in this work.

In the case of *Snail1* and *Prrx1*, we have evidences for this potential compensation. Both are induced by TGFβ as shown in cells in culture (Ocaña et al., 2012), but *Snail1* is an early-response gene. This is consistent with the expression patterns observed, with *Snail1* already expressed in the premigratory neural crest population and Prrx1 only in migratory neural crest cells. Furthermore, in a recent study, we have already shown that neural crest *Prrx1*-positive cells are descendants of *Snail1*-positive cells, which in turn became *Snail1* negative (Fazilaty et al., 2019). Interestingly, we show that *Snail1* levels increase and its expression territory is expanded in *Prrx1* homozygous mutants. This expansion of *Snail1* expression into *Prrx1* territory upon *Prrx1* loss, likely can at least partially compensate due to some overlapping functions as EMT-TFs (Fazilaty et al., 2019; Ocaña et al., 2012, 2017). This is also consistent with the observed weak phenotype of the single *Prrx1* mutant.

On the other hand, the lack of a more obvious phenotype in the *Snail1/ Prrx1* combined mutant can be explained by their complementary expression patterns. This complementarity is conserved in mouse, chicken, zebrafish embryos and cancer cells, and it arises from a GRN whereby they repress each other´s transcription. *Snail1* directly represses *Prrx1* gene expression, while *Prrx1* acts as an activator, inducing the expression of the *miR-15* family, which in turn attenuates *Snail1* expression (Fazilaty et al., 2019).

5.2 *Prrx1* and *Twist1* genetically interact during digit morphogenesis

In contrast to the complementary expression pattern of *Prrx1* and *Snail1*, *Prrx1* and *Twist1* are co-expressed in many embryonic territories and in cancer cells, where they can cooperate (Ocaña et al., 2012; Soldatov et al., 2019). In this work, we show that they are co-expressed in the same cells during early mouse embryonic development, in the first branchial arch, the lateral plate mesoderm and the limb bud. In addition, they are also co-expressed in the mesenchymal cells in the interdigital area of the E13 limb buds. Proliferation and cell death in this area are tightly controlled, as deregulation of these processes can lead to polydactyly (Hernández-Martínez & Covarrubias, 2011).

We have generated a mouse model that combines mutations in *Prrx1* and *Twist1*. Double heterozygous mice are viable and fertile, but they show digit defects, including hindlimb polydactyly with full penetrance, a stronger phenotype than that observed in the single mutants. In addition, homozygous mutants for *Prrx1* combined with *Twist1* mutation in heterozygosis present forelimb polydactyly with incomplete penetrance, which has never been observed in *Twist1* heterozygous mutants. Thus, this is another case of genetic interaction among EMT-TFs.

As EMT-TFs, both *Prrx1* and *Twist1* could have an impact in the proliferation and cell death of these cells (Vega et al., 2004; Thiery et al., 2009). In addition, specific mutation of *Twist1* in the mouse limb impacts the expression of *Shh* and *Fgf8,* essential for digit morphogenesis (Krawchuk et al., 2010). Interestingly, double mutants for *Prrx1* and *Prrx2*, which show polydactyly, also show defective *Shh* and *Fgf8* expression, plus a lack of *Bmp4* expression, directly associated with polydactyly (Selever et al., 2004). *Bmp4* expression has not been explored in *Twist1* mutant limbs, but *Twist1* mutants lack *BMP4* in a different context (Tischfield et al., 2017). Whether the combined mutation of *Prrx1* and *Twist1* results in the alteration of *Bmp4*, *Shh* or *Fgf8*, or in an abnormal proliferation or cell death of the

interdigital mesenchyme cells needs to be assessed to understand the mechanism for the interaction of these EMT-TFs in digit morphogenesis. In addition, it will be interesting to explore the phenotype of *Prrx1* and *Twist1* double mutants, as we have only explored the *Twist1* heterozygous condition. The reason is that *Twist1* homozygous mutants die at E10.5 with severe developmental defects, including an open neural tube and defects in the migration of the head mesoderm and the neural crest (Chen & Behringer, 1995; Soo et al., 2002), and another tissue-specific conditional *Twist1* mutant is required.

5.3 *Prrx1* isoforms can compensate each other during mouse development

Two PRRX1 protein isoforms resulting from alternative splicing of its transcripts, *Prrx1L* and *Prrx1S*, were previously described. They are conserved in evolution and differentially expressed during development and in adult mice (Braasch et al., 2014; Kern et al., 1992; Norris et al., 2000). They contain different transactivation and DNA binding activities as assessed in *in vitro* experiments (Norris & Kern, 2001b), and have differential EMT-inducing capabilities when acting individually in cultured cells (unpublished data). In this work, we have identified an additional *Prrx1* isoform (PRRX1T), which retains intron 3, and gives rise to a truncated protein containing the same protein domains than the *Prrx1S* isoform, although it is 17 amino acids shorter.

To investigate the role of the *Prrx1* isoforms in vivo, we have generated 2 mouse lines, one of them lacking PRRX1L, and the other lacking PRRX1S and PRRX1T. Strikingly, none of these isoform-specific mouse lines reproduces the developmental phenotype described for the constitutive *Prrx1* null mutants, and furthermore, *Prrx1* isoform-specific homozygous mutants are viable and fertile. All of this indicates that *Prrx1* isoforms can compensate for each other's loss during mouse embryonic development. The fact that these mice are viable in a very controlled environment does not mean that they do not have any phenotypic differences with wild type mice. In addition, to understand *Prrx1* functions in adult mice, and dissect the functions of each of the isoforms, we have generated a conditional *Prrx1* loss of function model, that allows us to induce the recombination of *Prrx1* in every cell and tissue in a time-dependent manner by administering tamoxifen. Comparing the phenotype of the conditional *Prrx1* loss of function with the wild type and the isoform-specific mutants, we have found a very interesting phenotype in the maintenance of vascular integrity, as described below.

5.4 The role of PRRX1 in pericyte function

It was already known that *Prrx1* was expressed in activated hepatic stellate cells during liver fibrosis (Gong et al., 2017; Jiang and Stefanovic, 2008) and that *Prrx1* mutants have defects in the architecture of the great arteries and in the formation of the vasculature of the embryonic lung (Bergwerff etl., 2000; Ihida-Stansbury et al., 2004). However, the cellular and molecular basis for these defects were not known.

In this work, we have shown that *Prrx1* is specifically expressed in pericytes during mouse development, in retinal neonatal angiogenesis and during adult homeostasis in different organs. Consistent with this expression pattern, we have observed that constitutive *Prrx1* mouse embryos show a variable degree of haemorrhage during development. However, in contrast to a previous study in new-born mice (Ihida-Stansbury et al., 2004), lungs from *Prrx1* constitutive mutant E18.5 embryos did not show an evident defect in endothelial cells nor vascular smooth muscle cells. *Prrx1* mutant new-born mice have breathing problems due to a severe cleft palate and die of suffocation shortly after (Martin et al., 1995). Endothelial cell death is one of the first events after neonatal asphyxia (Gerosa et al., 2014), compatible the lack of endothelial cells observed by Ihida-Stansbury et al., 2004. In order to investigate the defects in lung vascularization that are a direct consequence of *Prrx1* mutation, we dissected lungs from embryos right before birth, to avoid the damage caused by suffocation. These lungs did not show overt histological defects and there was no evidence of lack of endothelial cells or vascular smooth muscle cells, indicating that *Prrx1* is not required for the vascularization of the embryonic lung, and the phenotype previously observed in new-born mice was surely due to the secondary effects produced by the suffocation.

Conditional mutation of *Prrx1* in postnatal mouse retinas led to an increased leakage of erythrocytes in the growing end of the vasculature. Interestingly, *Prrx1* isoform mutants showed similar defects although at a lower penetrance, suggesting a cooperation among the isoforms. Both conditional *Prrx1* mutants and isoforms-specific constitutive mutations led to pulmonary haemorrhage in adult mice, indicating that *Prrx1* is required for the maintenance of vascular homeostasis in adult lungs. Interestingly, the isoforms-specific mutations result in a higher penetrance of haemorrhage than the conditional loss of function of *Prrx1*. This might be due to the fact that the conditional mutants only lack *Prrx1* for 3-4 weeks whereas *Prrx1* isoform mutants are constitutive, and small vascular integrity defects

have more time to accumulate, from the initial development of the lung to the adulthood. The matrigel plug assay allows us to assess neoangiogenesis in different mouse mutants. Mice lacking the *Prrx1L* isoform mutants show a less severe phenotype, indicating that the S and/or T isoforms can compensate for its loss. Thus, all together, our data indicate that PRRX1 is required for the maintenance of vascular integrity and suggest that the different isoforms cooperate and can compensate each other´s loss.

During angiogenesis, endothelial tip cells at the growing end recruit pericytes through paracrine signalling. When this signalling is defective, the vasculature does not form properly and the integrity is compromised (Hellström et al., 1999), leading to haemorrhages like those we observe in our mutants. As PRRX1 is expressed in pericytes, our data suggest a defective recruitment of mutant pericytes by endothelial cells. To get further insight at the molecular level and to understand how *Prrx1* operates in the pericytes, we performed massive transcriptomic analysis, comparing pericytes isolated from plugs implanted in wild type or *Prrx1* conditional mutant mice. Gene ontology enrichment analysis from the differentially downregulated genes reveals a potential role of *Prrx1* in the regulation of ECM protein expression in pericytes (**Figure 39**). The role of *Prrx1* in the regulation of the transcription of secreted proteins, such as collagens, among others, is consistent with the non-cell autonomous effect that we observe in *Prrx1* mutants.

Our analysis shows that the 5 most expressed ECM proteins by pericytes, collagens *Col1a1, Col1a2, Col3a1, Col6a1 and Col6a2*, are significantly downregulated in our analysis (Brown et al., 2015). Interestingly, *Col1a1,* has already been described to be a direct PRRX1 target in activated stellate cells, which are the pericytes of the liver (Jiang & Stefanovic, 2008), and *Col1a1* hippomorphic mutations cause aortic dissection and rupture, which leads to lethality due to aortic haemorrhage in 18 month old mice (Marjamaa et al., 2006; Rahkonen et al., 2004). In addition to collagen I, *Col3a1* is among the most downregulated ECM genes in our transcriptomic analysis of pericytes from *Prrx1* mutant mice, compatible with the finding that it is among the most highly expressed ECM proteins by pericytes. We have analysed its promoter in different species, and we found *Prrx1* binding sites, one of them conserved from mouse to humans. Our preliminary data indicate that PRRX1 can bind to this site and thus, may directly activate the transcription of the *Col3a1* gene. Interestingly, mutations in *Col3a1* in humans are associated with Vascular Ehlers-Danlos Syndrome, which is characterized by an excessive arterial fragility (Pope et al., 1996; Mizuno et al.,

2013). Heterozygous mutations in *Col3a1* on mice have similarities with Vascular Ehlers-Danlos Syndrome in humans, including spontaneous rupture of the bowel, large or medium arteries, and aortic dissection (Cooper et al., 2010; Pope et al., 1996; L. B. Smith et al., 2011). Together, these studies point towards a role of *Prrx1* in the secretion of the ECM and collagens in particular, essential for the interaction between endothelial cells and pericytes and the maturation and integrity of the vasculature (Stratman & Davis, 2012).

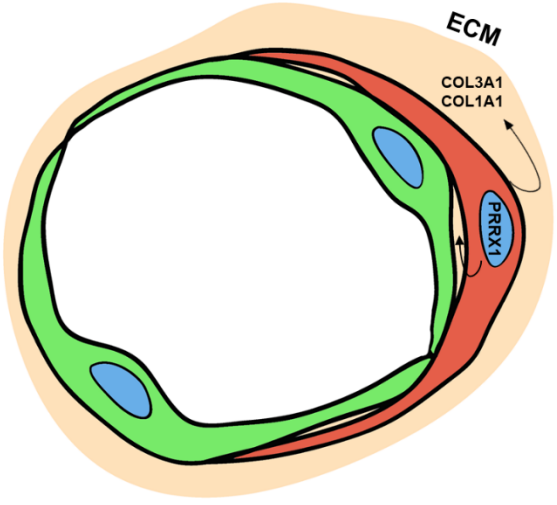


Figure 39. PRRX1 is required for blood vessel integrity. Schematic representation of a capillary. Pericyte (red) wraps around endothelial cells (green) and expresses PRRX1 which activates the transcription genes encoding ECM proteins, including a series of collagen genes such as *Col3a1* and *Col1a1*.

In addition to extracellular matrix proteins, many genes that are crucial for vascular development are significantly downregulated in *Prrx1* mutant pericytes. These include *Pdgfb* and its receptor, *Pdgfrβ*, both of which are known to be crucial for pericyte recruitment (Abramsson et al., 2003; Lindahl et al., 1997), as well as other pericyte markers, such as *Ng2*, and proteins that are essential for the formation of blood vessels, such as *Fgfr2* and *Vegfa*. All of these genes are annotated in the gene ontology term “Vasculature development” (GO:0001944), which is one of the top GO terms in our gene ontology enrichment analysis. Many of these genes have a direct implication in pericyte function, and their missexpression in *Prrx1* mutant pericytes could account for *Prrx1* cell autonomous functions in the pericytes.

The interaction between the endothelial cells and the pericytes is crucial for the formation and the integrity of the vasculature during homeostasis. In this work, we have shown that *Prrx1* is necessary for the integrity of the vasculature in different contexts and for the neovascularization of matrigel plugs, where it has a non-cell autonomous effect over the endothelial cell population. Indeed, by regulating the expression of secreted proteins such as collagens and signalling proteins such as *Pdgfb* and *Vegfa*, *Prrx1* could be exerting non-cell autonomous effects over the endothelial cells which in turn will affect the integrity of the vasculature. The tight control and regulation of the permeability of the vasculature is crucial in many organs. In particular, in the brain, pericytes have an essential role in regulating the permeability and integrity of the blood-brain barrier (BBB) (Armulik et al., 2010; Hall et al., 2014). During normal aging, the integrity of the BBB declines due to defective pericyte function, which eventually leads to neuronal cell death (Yang et al., 2020a). Consistent with this, BBB integrity loss is a landmark in the progression of many cerebrovascular diseases, including dementias such as Alzheimer's disease, the most common form of dementia (Govindpani et al., 2019; Sagare et al., 2013; Sweeney et al., 2018). Furthermore, the increase in life expectancy and the decrease in natality rates is leading to the overall aging of the population worldwide, which in turn leads to a rise in the prevalence of dementias. Despite this, there are very few drugs that are effective in preventing the progression of dementia. Thus, understanding the molecular mechanisms leading to the development of these diseases is essential in order to develop new therapeutical strategies to stop their progression (Cao et al., 2020). In this work, we have shown that *Prrx1* is expressed in the pericytes of the brain, and that *Prrx1* loss of function is essential for the integrity of the blood-retinal barrier. Further experiments will be required in order to understand if *Prrx1* is also involved in the maintenance of the vascular integrity in the brain and explore whether *Prrx1* could play a role in the development of vascular dementias.

In addition to the proposed implication of PRRX1 in blood vessel formation and vascular integrity, our gene ontology analysis from the transcriptome of *Prrx1* mutant pericytes, suggests a potential role in the modulation of the immune response. In normal conditions, pericytes secrete cytokines that can be activated during inflammation and in some diseases, increasing their expression and organizing the immune response (Navarro et al., 2016). Interestingly, most of the upregulated genes in *Prrx1*-deficient pericytes are

cytokines and cell surface markers implicated in mediating the immune response. Some of these genes include chemokines such as *Ccl3* and *Ccl4,* which respectively are the top 6 and 15 most upregulated genes in *Prrx1* mutant pericytes. Both proteins act as chemoattractants for immune cells and are significantly upregulated in brain microvascular pericytes in response to inflammatory stimuli (Kovac et al., 2011). The role of pericytes in immunomodulation is a new avenue that would be extremely interesting to study and, in particular, it can be especially relevant during cancer progression. Pericytes are an important part of the tumour microenviroment, and have been shown to influence metastatic progression, although the mechanisms are not well understood (Navarro et al., 2016; Xian et al., 2006; Zhang et al., 2020). Thus, it will be very interesting to study the impact of *Prrx1* mutant pericytes in the models of cancer that we have available in the lab.

Last but not least, the two proposed functions for PRRX1 can be interconnected. Indeed, immune cells can contribute to vascular development (Bruno et al., 2014). In particular, during cancer progression, angiogenesis and immune regulation are deeply interdependent. Many immunogenic cytokines also have pro-angiogenic effects, and inhibition of the immune response has an impact on angiogenesis in cancer (O'Byrne et al., 2000). This could represent another mechanism by which *Prrx1* has a non-cell autonomous effect on the formation of blood vessels, by regulating the expression of cytokines that in turn affect the growth of the vasculature. The vascular-immune crosstalk can be targeted during cancer progression using a combination of immune checkpoint inhibitors and anti-angiogenic drugs. This combined therapy has passed clinical trials for the treatment of several cancer types and is already being used in cancer patients, highlighting the importance of understanding the molecular mechanisms that control immune regulation in the pericytes (Lee et al., 2020). Thus, it will be very interesting to study the impact of *Prrx1* mutant pericytes in the models of cancer that we have available in the lab, as further experiments are necessary to determine if *Prrx1* mediates a switch that can trigger the immune response mediated by pericytes, and whether this can significantly influence cancer progression.

Conclusions

Conclusions

1. *Snail1* and *Prrx1* are expressed in complementary domains in the developing embryo. SNAIL1 territory is expanded in *Prrx1* mutants, compatible with a recently described GRN in which they repress each other and with a compensation of PRRX loss by SNAIL1. Consistent with this, when both *Prrx1* and *Snail1* are mutated in the neural crest, the phenotype in the Meckel cartilage is enhanced.

2. TWIST1 and PRRX1 are co-expressed during embryonic development and they genetically interact during digit morphogenesis.

3. *Prrx1* isoform mutants are viable and fertile, suggesting that they can compensate each other's functions during mouse development.

4. *Prrx1* is specifically expressed in perivascular cells during development and adult tissue homeostasis. Consistent with this, *Prrx1* conditional mutants have vascular defects leading to leakiness, indicating that PRRX1 is required for the integrity of the vasculature.

5. Transcriptomic analysis of *Prrx1* mutant pericytes reveals a putative role of PRRX1 in (i) the interaction of pericytes with endothelial cells and the formation of the vascular basement membrane, and (ii) in the pericyte-mediated regulation of the immune response.

6. All together, these data reveal:

- A high degree of compensation among EMT-TFs during neural crest and digit morphogenesis in the mouse, providing robustness to the developing embryo.
- PRRX1 as a new player in vascular homeostasis and immunoregulation with important implications in vascular congenital or degenerative diseases and likely, in tumour angiogenesis and immunity.

Conclusiones

1. *Snail1* y *Prrx1* se expresan en territorios complementarios durante el desarrollo embrionario. El territorio de expresión de SNAIL1 se expande en los mutantes de *Prrx1*, lo cual es compatible con un GRN recientemente descrito en el que *Prrx1* y *Snail1* se reprimen mutuamente y con una compensación de la falta de PRRX1 por SNAIL1. Esto es consistente con el hecho de que la mutación conjunta de *Snail1* y *Prrx1* en la cresta neural resulta en la intensificación en el fenotipo en el cartílago de Meckel.

2. TWIST1 y PRRX1 se coexpresan durante el desarrollo embrionario e interaccionan genéticamente durante la morfogénesis de los dedos.

3. Los ratones mutantes de las isoformas de *Prrx1* son viables y fértiles, lo cual sugiere que las isoformas de *Prrx1* pueden compensar mutuamente su función durante el desarrollo embrionario de ratón.

4. *Prrx1* se expresa de manera específica en células perivasculares durante el desarrollo embrionario y durante la homeostasis. De manera consistente, los ratones mutantes de *Prrx1* tienen defectos vasculares que incluyen la pérdida de sangre. Esto indica que PRRX1 es necesario para la integridad de la vasculatura.

5. El análisis transcriptómico de pericitos mutantes para *Prrx1* ha revelado un posible rol de *Prrx1* en (i) la interacción de los pericitos con las células endoteliales y la formación de la membrana basal de los vasos sanguíneos, y (ii) la regulación mediada por los pericitos de la respuesta inmune.

6. En conjunto, estos datos revelan:

- Un alto grado de compensación entre EMT-TFs en la cresta neural y en la morfogénesis de los dedos, lo cual confiere robustez al embrión en desarrollo.

- PRRX1 tiene una función en la homeostasis vascular y regulación de la respuesta inmune, con implicaciones importantes en enfermedades vasculares congénitas o degenerativas, y probablemente en la angiogénesis y respuesta inmune durante el progreso del cáncer.

References

References

Abramsson, A., Lindblom, P., & Betsholtz, C. (2003). Endothelial and nonendothelial sources of PDGF-B regulate pericyte recruitment and influence vascular pattern formation in tumors. *Journal of Clinical Investigation*, *112*(8), 1142–1151. https://doi.org/10.1172/JCI200318549

Alarcon-Martinez, L., Villafranca-Baughman, D., Quintero, H., Kacerovsky, J. B., Dotigny, F., Murai, K. K., Prat, A., Drapeau, P., & Di Polo, A. (2020). Interpericyte tunnelling nanotubes regulate neurovascular coupling. *Nature*, *585*(7823), 91–95. https://doi.org/10.1038/s41586-020-2589-x

Albini, A., Bruno, A., Noonan, D. M., & Mortara, L. (2018). Contribution to tumor angiogenesis from innate immune cells within the tumor microenvironment: Implications for immunotherapy. *Frontiers in Immunology*, *9*(APR). https://doi.org/10.3389/fimmu.2018.00527

Armulik, A., Abramsson, A., & Betsholtz, C. (2005). Endothelial/pericyte interactions. *Circulation Research*, *97*(6), 512–523. https://doi.org/10.1161/01.RES.0000182903.16652.d7

Armulik, A., Genové, G., & Betsholtz, C. (2011). Pericytes: Developmental, Physiological, and Pathological Perspectives, Problems, and Promises. *Developmental Cell*, *21*(2), 193–215. https://doi.org/10.1016/j.devcel.2011.07.001

Armulik, A., Genové, G., Mäe, M., Nisancioglu, M. H., Wallgard, E., Niaudet, C., He, L., Norlin, J., Lindblom, P., Strittmatter, K., Johansson, B. R., & Betsholtz, C. (2010). Pericytes regulate the blood-brain barrier. *Nature*, *468*(7323), 557–561. https://doi.org/10.1038/nature09522

Armulik, A., & Betsholtz, C. (2010). Role of pericytes in vascular biology. *Experimental Approaches to Diabetic Retinopathy*, *20*, 194–202. https://doi.org/10.1159/000262670

Baluk, P., Morikawa, S., Baluk, P., Kaidoh, T., Haskell, A., Jain, R. K., & Mcdonald, D. M. (2002). Abnormalities in Pericytes on Blood Vessels and Endothelial Sprouts in Tumors Abnormalities in Pericytes on Blood Vessels and Endothelial Sprouts in Tumors. *The American Journal of Pathology*, *160*(3), 985–1000. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1867175/pdf/3058.pdf%0Ahttp://www.ncbi.nlm.nih.gov/pubmed/11891196%0Ahttp://www.pubmedcentral.nih.gov/articlerend

er.fcgi?artid=PMC1867175%0Ahttp://dx.doi.org/10.1016/S0002-9440(10)64920-6

Barriga, E. H., Trainor, P. A., Bronner, M., & Mayor, R. (2015). Animal models for studying neural crest development: Is the mouse different? *Development (Cambridge)*, *142*(9), 1555–1560. https://doi.org/10.1242/dev.121590

Berge, D. Ten, Brouwer, A., Korving, J., Martin, J. F., & Meijlink, F. (1998). Prx1 and Prx2 in skeletogenesis: Roles in the craniofacial region, inner ear and limbs. *Development*, *125*(19), 3831–3842.

Bergers, G., & Song, S. (2005). The role of pericytes in blood-vessel formation and maintenance. *Neuro-Oncology*, *7*(4), 452–464. https://doi.org/10.1215/s1152851705000232

Bergwerff, M., Gittenberger-de Groot, A. C., Wisse, L. J., DeRuiter, M. C., Wessels, A., Martin, J. F., Olson, E. N., & Kern, M. J. (2000). Loss of function of the Prx1 and Prx2 hameobox genes alters architecture of the great elastic arteries and ductus arteriosus. *Virchows Archiv*, *436*(1), 12–19. https://doi.org/10.1007/PL00008193

Betancur, P., Bronner-Fraser, M., & Sauka-Spengler, T. (2010). Assembling neural crest regulatory circuits into a gene regulatory network. *Annual Review of Cell and Developmental Biology*, *26*, 581–603. https://doi.org/10.1146/annurev.cellbio.042308.113245

Betsholtz, C. (2004). Insight into the physiological functions of PDGF through genetic studies in mice. *Cytokine and Growth Factor Reviews*, *15*(4), 215–228. https://doi.org/10.1016/j.cytogfr.2004.03.005

Bildsoe, H., Loebel, D. A. F., Jones, V. J., Chen, Y. T., Behringer, R. R., & Tam, P. P. L. (2009). Requirement for Twist1 in frontonasal and skull vault development in the mouse embryo. *Developmental Biology*, *331*(2), 176–188. https://doi.org/10.1016/j.ydbio.2009.04.034

Bildsoe, H., Loebel, D. A. F., Jones, V. J., Hor, A. C. C., Braithwaite, A. W., Chen, Y. T., Behringer, R. R., & Tam, P. P. L. (2013). The mesenchymal architecture of the cranial mesoderm of mouse embryos is disrupted by the loss of Twist1 function. *Developmental Biology*, *374*(2), 295–307. https://doi.org/10.1016/j.ydbio.2012.12.004

References

Boulay, J. L., Dennefeld, C., & Alberga, A. (1987). The Drosophila developmental gene snail encodes a protein with nucleic acid binding fingers. *Nature*, *330*(6146), 395–398. https://doi.org/10.1038/330395a0

Bourgeois, P., Bolcato-Bellemin, A. L., Danse, J. M., Bloch-Zupan, A., Yoshiba, K., Stoetzel, C., & Perrin-Schmitt, F. (1998). The variable expressivity and incomplete penetrance of the twist-null heterozygous mouse phenotype resemble those of human Saethre-Chotzen syndrome. *Human Molecular Genetics*, *7*(6), 945–957. https://doi.org/10.1093/hmg/7.6.945

Braasch, I., Guiguen, Y., Loker, R., Letaw, J. H., Ferrara, A., Bobe, J., & Postlethwait, J. H. (2014). Connectivity of vertebrate genomes: Paired-related homeobox (Prrx) genes in spotted gar, basal teleosts, and tetrapods. *Comparative Biochemistry and Physiology Part - C: Toxicology and Pharmacology*, *163*, 24–36. https://doi.org/10.1016/j.cbpc.2014.01.005

Brabletz, T., Kalluri, R., Nieto, M. A., & Weinberg, R. A. (2018). EMT in cancer. *Nature Reviews Cancer*, *18*(2), 128–134. https://doi.org/10.1038/nrc.2017.118

Brewer, S., Feng, W., Huang, J., Sullivan, S., & Williams, T. (2004). Wnt1-Cre-mediated deletion of AP-2α causes multiple neural crest-related defects. *Developmental Biology*, *267*(1), 135–152. https://doi.org/10.1016/j.ydbio.2003.10.039

Brown, L. A., Sava, P., Garcia, C., & Gonzalez, A. L. (2015). Proteomic Analysis of the Pericyte Derived Extracellular Matrix. *Cellular and Molecular Bioengineering*, *8*(3), 349–363. https://doi.org/10.1007/s12195-015-0408-5

Bruno, A., Pagani, A., Pulze, L., Albini, A., Dallaglio, K., Noonan M., D. M., & Mortara, L. (2014). Orchestration of angiogenesis by immune cells. *Frontiers in Oncology*, *4 JUL*(July), 1–13. https://doi.org/10.3389/fonc.2014.00131

Cano, A., Pérez-Moreno, M. A., Rodrigo, I., Locascio, A., Blanco, M. J., Del Barrio, M. G., Portillo, F., & Nieto, M. A. (2000). The transcription factor Snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nature Cell Biology*, *2*(2), 76–83. https://doi.org/10.1038/35000025

Cao, Q., Tan, C. C., Xu, W., Hu, H., Cao, X. P., Dong, Q., Tan, L., & Yu, J. T. (2020). The

Prevalence of Dementia: A Systematic Review and Meta-Analysis. *Journal of Alzheimer's Disease : JAD*, *73*(3), 1157–1166. https://doi.org/10.3233/JAD-191092

Carver, E. A., Jiang, R., Lan, Y., Oram, K. F., & Gridley, T. (2001). The Mouse Snail Gene Encodes a Key Regulator of the Epithelial-Mesenchymal Transition. *Molecular and Cellular Biology*, *21*(23), 8184–8188. https://doi.org/10.1128/mcb.21.23.8184-8188.2001

Carver, E. A., Oram, K. F., & Gridley, T. (2002). Craniosynostosis in Twist heterozygous mice: A model for Saethre-Chotzen syndrome. *Anatomical Record*, *268*(2), 90–92. https://doi.org/10.1002/ar.10124

Çelik, T., Simsek, P. O., Sozen, T., Ozyuncu, O., Utine, G. E., Talim, B., Yiğit, Ş., Boduroglu, K., & Kamnasaran, D. (2012). PRRX1 is mutated in an otocephalic newborn infant conceived by consanguineous parents. In *Clinical Genetics* (Vol. 81, Issue 3, pp. 294–297). https://doi.org/10.1111/j.1399-0004.2011.01730.x

Chen, J., Bardes, E. E., Aronow, B. J., & Jegga, A. G. (2009). ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Research*, *37*(SUPPL. 2), 305–311. https://doi.org/10.1093/nar/gkp427

Chen, Y., & Gridley, T. (2013). Compensatory regulation of the Snai1 and Snai2 genes during chondrogenesis. *Journal of Bone and Mineral Research*, *28*(6), 1412–1421. https://doi.org/10.1002/jbmr.1871

Chen, Y. T., Akinwunmi, P. O., Jian, M. D., Tam, O. H., & Behringer, R. R. (2007). Generation of a Twist1 conditional null allele in the mouse. *Genesis*, *45*(9), 588–592. https://doi.org/10.1002/dvg.20332

Chen, Z. F., & Behringer, R. R. (1995). Twist Is Required in Head Mesenchyme for Cranial Neural Tube Morphogenesis. *Genes and Development*, *9*(6), 686–699. https://doi.org/10.1101/gad.9.6.686

Coltrini, D., Di Salle, E., Ronca, R., Belleri, M., Testini, C., & Presta, M. (2013). Matrigel plug assay: Evaluation of the angiogenic response by reverse transcription-quantitative PCR. *Angiogenesis*, *16*(2), 469–477. https://doi.org/10.1007/s10456-012-9324-7

Comijn, J., Berx, G., Vermassen, P., Verschueren, K., Van Grunsven, L., Bruyneel, E.,

Mareel, M., Huylebroeck, D., & Van Roy, F. (2001). The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Molecular Cell*, *7*(6), 1267–1278. https://doi.org/10.1016/S1097-2765(01)00260-X

Cooper, T. K., Zhong, Q., Krawczyk, M., Tae, H. J., Müller, G. A., Schubert, R., Myers, L. A., Dietz, H. C., Talan, M. I., & Briest, W. (2010). The Haploinsufficient Col3a1 Mouse as a Model for Vascular Ehlers-Danlos Syndrome. *Veterinary Pathology*, *47*(6), 1028–1039. https://doi.org/10.1177/0300985810374842

Cserjesi, P., Lilly, B., Bryson, L., Wang, Y., Sassoon, D. A., & Olson, E. N. (1992). MHox: A mesodermally restricted homeodomain protein that binds an essential site in the muscle creatine kinase enhancer. *Development*, *115*(4), 1087–1101.

Dale, J. K., Malapert, P., Chal, J., Vilhais-Neto, G., Maroto, M., Johnson, T., Jayasinghe, S., Trainor, P., Herrmann, B., & Pourquié, O. (2006). Oscillations of the snail genes in the presomitic mesoderm coordinate segmental patterning and morphogenesis in vertebrate somitogenesis. *Developmental Cell*, *10*(3), 355–366. https://doi.org/10.1016/j.devcel.2006.02.011

Danielian, P. S., Muccino, D., Rowitch, D. H., Michael, S. K., & McMahon, A. P. (1998). Modification of gene activity in mouse embryos in utero by a tamoxifen-inducible form of Cre recombinase. *Current Biology*, *8*(24), 1323–1326. https://doi.org/10.1016/s0960-9822(07)00562-3

Dasouki, M., Andrews, B., Parimi, P., & Kamnasaran, D. (2013). Recurrent agnathia-otocephaly caused by DNA replication slippage in PRRX1. *American Journal of Medical Genetics, Part A*, *161*(4), 803–808. https://doi.org/10.1002/ajmg.a.35879

De Frutos, C. A., Dacquin, R., Vega, S., Jurdic, P., Machuca-Gayet, I., & Angela Nieto, M. (2009). Snail1 controls bone mass by regulating Runx2 and VDR expression during osteoblast differentiation. *EMBO Journal*, *28*(6), 686–696. https://doi.org/10.1038/emboj.2009.23

De Frutos, C. A., Vega, S., Manzanares, M., Flores, J. M., Huertas, H., Martínez-Frías, M. L., & Nieto, M. A. (2007). Snail1 Is a Transcriptional Effector of FGFR3 Signaling during Chondrogenesis and Achondroplasias. *Developmental Cell*, *13*(6), 872–883. https://doi.org/10.1016/j.devcel.2007.09.016

References

De La Torre, J. C. (2004). Is Alzheimer's disease a neurodegenerative or a vascular disorder? Data, dogma, and dialectics. *Lancet Neurology*, *3*(3), 184–190. https://doi.org/10.1016/S1474-4422(04)00683-0

Delmonte, C., & Capelozzi, V. L. (2001). Morphologic determinants of asphyxia in lungs: A semiquantitative study in forensic autopsies. *American Journal of Forensic Medicine and Pathology*, *22*(2), 139–149. https://doi.org/10.1097/00000433-200106000-00006

Dong, J., Hu, Y., Fan, X., Wu, X., Mao, Y., Hu, B., Guo, H., Wen, L., & Tang, F. (2018). Single-cell RNA-seq analysis unveils a prevalent epithelial/mesenchymal hybrid state during mouse organogenesis. *Genome Biology*, *19*(1), 1–20. https://doi.org/10.1186/s13059-018-1416-2

Donnelly, M., Todd, E., Wheeler, M., Winn, V. D., & Kamnasaran, D. (2012). Prenatal diagnosis and identification of heterozygous frameshift mutation in PRRX1 in an infant with agnathia-otocephaly. *Prenatal Diagnosis*, *32*(9), 903–905. https://doi.org/10.1002/pd.3910

Enge, M., Bjarnegård, M., Gerhardt, H., Gustafsson, E., Kalén, M., Asker, N., Hammes, H. P., Shani, M., Fässler, R., & Betsholtz, C. (2002). Endothelium-specific platelet-derived growth factor-B ablation mimics diabetic retinopathy. *EMBO Journal*, *21*(16), 4307–4316. https://doi.org/10.1093/emboj/cdf418

Etchevers, H. C., Vincent, C., Le Douarin, N. M., & Couly, G. F. (2001). The cephalic neural crest provides pericytes and smooth muscle cells to all blood vessels of the face and forebrain. *Development*, *128*(7), 1059–1068.

Fazilaty, H., Rago, L., Kass Youssef, K., Ocaña, O. H., Garcia-Asencio, F., Arcas, A., Galceran, J., & Nieto, M. A. (2019). A gene regulatory network to control EMT programs in development and disease. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-13091-8

Fruttiger, M. (2002). Development of the mouse retinal vasculature: Angiogenesis versus vasculogenesis. *Investigative Ophthalmology and Visual Science*, *43*(2), 522–527.

Gaengel, K., Genové, G., Armulik, A., & Betsholtz, C. (2009). Endothelial-mural cell signaling in vascular development and angiogenesis. *Arteriosclerosis, Thrombosis, and*

Vascular Biology, *29*(5), 630–638. https://doi.org/10.1161/ATVBAHA.107.161521

Gallione, C. J., Repetto, G. M., Legius, E., Rustgi, A. K., Schelley, S. L., Tejpar, S., Mitchell, G., Drouin, É., Westermann, C. J. J., & Marchuk, D. A. (2004). A combined syndrome of juvenile polyposis and hereditary haemorrhagic telangiectasia associated with mutations in MADH4 (SMAD4). *Lancet*, *363*(9412), 852–859. https://doi.org/10.1016/S0140-6736(04)15732-2

Geraldes, P., Hiraoka-Yamamoto, J., Matsumoto, M., Clermont, A., Leitges, M., Marette, A., Aiello, L. P., Kern, T. S., & King, G. L. (2009). Activation of PKC-and SHP-1 by hyperglycemia causes vascular cell apoptosis and diabetic retinopathy. *Nature Medicine*, *15*(11), 1298–1306. https://doi.org/10.1038/nm.2052

Gerosa, C., Fanni, D., Puddu, M., Locci, G., Obinu, E., Fanos, V., & Faa, G. (2014). Histological markers of neonatal asphyxia: the relevant role of vascular changes. *Journal of Pediatric and Neonatal Individualized Medicine*, *3*(2). https://doi.org/10.7363/03027

Ghouzzi, V. El, Merrer, M. Le, Perrin-schmitt, F., Lajeunie, E., Benit, P., Renier, D., Bourgeois, P., Munnich, A., & Bonaventure, J. (2000). *Saethre-Chotzen syndrome*. *15*(january 1997), 42–46.

Gong, J., Han, J., He, J., Liu, J., Han, P., Wang, Y., Li, M., Li, D., Ding, X., Du, Z., Liao, J., & Tian, D. (2017). Paired related homeobox protein 1 regulates PDGF-induced chemotaxis of hepatic stellate cells in liver fibrosis. *Laboratory Investigation*, *97*(9), 1020–1032. https://doi.org/10.1038/labinvest.2017.65

Goumans, M. J., Valdimarsdottir, G., Itoh, S., Rosendahl, A., Sideras, P., & Ten Dijke, P. (2002). Balancing the activation state of the endothelium via two distinct TGF-β type I receptors. *EMBO Journal*, *21*(7), 1743–1753. https://doi.org/10.1093/emboj/21.7.1743

Govindpani, K., McNamara, L. G., Smith, N. R., Vinnakota, C., Waldvogel, H. J., Faull, R. L., & Kwakowsky, A. (2019). Vascular Dysfunction in Alzheimer's Disease: A Prelude to the Pathological Process or a Consequence of It? *Journal of Clinical Medicine*, *8*(5), 651. https://doi.org/10.3390/jcm8050651

Grande, M. T., Sánchez-Laorden, B., López-Blau, C., De Frutos, C. A., Boutet, A., Arévalo,

M.", Rowe RG, Weiss SJ, López-Novoa JM, & Nieto, M. A. (2015). Snail1-induced partial epithelial-to-mesenchymal transition drives renal fibrosis in mice and can be targeted to reverse established disease. Nature medicine, 21(9), 989-997. https://doi.org/10.1038/nm.3901

Green, S. A., Simoes-Costa, M., & Bronner, M. E. (2015). Evolution of vertebrates as viewed from the crest. *Nature*, *520*(7548), 474–482. https://doi.org/10.1038/nature14436

Grimsley-Myers, C. M., Isaacson, R. H., Cadwell, C. M., Campos, J., Hernandes, M. S., Myers, K. R., Seo, T., Giang, W., Griendling, K. K., & Kowalczyk, A. P. (2020). VE-cadherin endocytosis controls vascular integrity and patterning during development. *Journal of Cell Biology*, *219*(5). https://doi.org/10.1083/jcb.201909081

Hajra, K. M., Chen, D. Y. S., & Fearon, E. R. (2002). The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Research*, *62*(6), 1613–1618.

Hall, C. N., Reynell, C., Gesslein, B., Hamilton, N. B., Mishra, A., Sutherland, B. A., Oâ Farrell, F. M., Buchan, A. M., Lauritzen, M., & Attwell, D. (2014). Capillary pericytes regulate cerebral blood flow in health and disease. *Nature*, *508*(1), 55–60. https://doi.org/10.1038/nature13165

Hellström, M., Kalén, M., Lindahl, P., Abramsson, A., & Betsholtz, C. (1999). Role of PDGF-B and PDGFR-β in recruitment of vascular smooth muscle cells and pericytes during embryonic blood vessel formation in the mouse. *Development*, *126*(14), 3047–3055.

Hernández-Martínez, R., & Covarrubias, L. (2011). Interdigital cell death function and regulation: New insights on an old programmed cell death model. *Development Growth and Differentiation*, *53*(2), 245–258. https://doi.org/10.1111/j.1440-169X.2010.01246.x

Hunt, S. E., McLaren, W., Gil, L., Thormann, A., Schuilenburg, H., Sheppard, D., Parton, A., Armean, I. M., Trevanion, S. J., Flicek, P., & Cunningham, F. (2018). Ensembl variation resources. *Database : The Journal of Biological Databases and Curation*, *2018*(8), 1–12. https://doi.org/10.1093/database/bay119

Ihida-Stansbury, K., McKean, D. M., Gebb, S. A., Martin, J. F., Stevens, T., Nemenoff, R., Akeson, A., Vaughn, J., & Jones, P. L. (2004). Paired-related homeobox gene Prx1 is required for pulmonary vascular development. *Circulation Research*, *94*(11), 1507–

1514. https://doi.org/10.1161/01.RES.0000130656.72424.20

Ito, Y., Yeo, J. Y., Chytil, A., Han, J., Bringas, P., Nakajima, A., Shuler, C. F., Moses, H. L., & Chai, Y. (2003). Conditional inactivation of Tgfbr2 in cranial neural crest causes cleft palate and calvaria defects. *Development*, *130*(21), 5269–5280. https://doi.org/10.1242/dev.00708

Jiang, F., & Stefanovic, B. (2008). Homeobox gene Prx1 is expressed in activated hepatic stellate cells and transactivates collagen α1(I) promoter. *Experimental Biology and Medicine*, *233*(3), 286–296. https://doi.org/10.3181/0707-RM-177

Jiang, R., Lan, Y., Norton, C. R., Sundberg, J. P., & Gridley, T. (1998). The slug gene is not essential for mesoderm or neural crest development in mice. *Developmental Biology*, *198*(2), 277–285. https://doi.org/10.1006/dbio.1998.8909

Jo, D., Nashabi, A., Doxsee, C., Lin, Q., Unutmaz, D., Chen, J., & Ruley, H. E. (2001). Epigenetic regulation of gene structure and function with a cell-permeable Cre recombinase. *Nature Biotechnology*, *19*(10), 929–933. https://doi.org/10.1038/nbt1001-929

Jones, F. S., Meech, R., Edelman, D. B., Oakey, R. J., & Jones, P. L. (2001). Prx1 controls vascular smooth muscle cell proliferation and tenascin-C expression and is upregulated with Prx2 in pulmonary vascular disease. *Circulation Research*, *89*(2), 131–138. https://doi.org/10.1161/hh1401.093582

Kano, M. R., Morishita, Y., Iwata, C., Iwasaka, S., Watabe, T., Ouchi, Y., Miyazono, K., & Miyazawa, K. (2005). VEGF-A and FGF-2 synergistically promote neoangiogenesis through enhancement of endogenous PDGF-B-PDGFRβ signaling. *Journal of Cell Science*, *118*(16), 3759–3768. https://doi.org/10.1242/jcs.02483

Kern, M. J., Witte, D. P., Valerius, M. T., Aronow, B. J., & Potter, S. S. (1992). A novel murine homeobox gene isolated by a tissue specific PCR cloning strategy. *Nucleic Acids Research*, *20*(19), 5189–5195. https://doi.org/10.1093/nar/20.19.5189

Kovac, A., Erickson, M. A., & Banks, W. A. (2011). Brain microvascular pericytes are immunoactive in culture: Cytokine, chemokine, nitric oxide, and LRP-1 expression in response to lipopolysaccharide. *Journal of Neuroinflammation*, *8*(1), 139.

https://doi.org/10.1186/1742-2094-8-139

Kratochwil, C. F., Geissler, L., Irisarri, I., & Meyer, A. (2015). Molecular evolution of the neural crest regulatory network in ray-finned fish. *Genome Biology and Evolution*, *7*(11), 3033–3046. https://doi.org/10.1093/gbe/evv200

Krawchuk, D., Weiner, S. J., Chen, Y. T., Lu, B. C., Costantini, F., Behringer, R. R., & Laufer, E. (2010). Twist1 activity thresholds define multiple functions in limb development. *Developmental Biology*, *347*(1), 133–146. https://doi.org/10.1016/j.ydbio.2010.08.015

Lagha, M., Bothma, J. P., Esposito, E., Ng, S., Stefanik, L., Tsui, C., Johnston, J., Chen, K., Gilmour, D. S., Zeitlinger, J., & Levine, M. S. (2013). Paused Pol II coordinates tissue morphogenesis in the drosophila embryo. *Cell*, *153*(5), 976. https://doi.org/10.1016/j.cell.2013.04.045

Lamagna, C., & Bergers, G. (2006). The bone marrow constitutes a reservoir of pericyte progenitors. *Journal of Leukocyte Biology*, *80*(4), 677–681. https://doi.org/10.1189/jlb.0506309

Lee, W. S., Yang, H., Chon, H. J., & Kim, C. (2020). Combination of anti-angiogenic therapy and immune checkpoint blockade normalizes vascular-immune crosstalk to potentiate cancer immunity. *Experimental and Molecular Medicine*, *52*(9), 1475–1485. https://doi.org/10.1038/s12276-020-00500-y

Leptin, M. (1991). twist and snail as positive and negative regulators during Drosophila mesoderm development. *Genes and Development*, *5*(9), 1568–1576. https://doi.org/10.1101/gad.5.9.1568

Lindahl, P., Johansson, B. R., Levéen, P., & Betsholtz, C. (1997). Pericyte loss and microaneurysm formation in PDGF-B-deficient mice. *Science*, *277*(5323), 242–245. https://doi.org/10.1126/science.277.5323.242

Locascio, A., Manzanares, M., Blanco, M. J., & Nieto, M. A. (2002). Modularity and reshuffling of snail and slug expression during vertebrate evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(26), 16841–16846. https://doi.org/10.1073/pnas.262525399

Lu, M. F., Cheng, H. T., Kern, M. J., Potter, S. S., Tran, B., Diekwisch, T. G. H., & Martin, J.

F. (1999). Prx-1 Functions Cooperatively With Another Paired-Related Homeobox Gene, Prx-2, To Maintain Cell Fates Within the Craniofacial Mesenchyme. *Development*, *126*(3), 495–504. https://doi.org/10.1006/dbio.1998.9116

Lu, M. F., Cheng, H. T., Lacy, A. R., Kern, M. J., Argao, E. A., Potter, S. S., Olson E. N. & Martin, J. F. (1999). Paired-related homeobox genes cooperate in handplate and hindlimb zeugopod morphogenesis. Developmental biology, 205(1), 145-157.

Madisen, L., Zwingman, T. A., Sunkin, S. M., Oh, S. W., Zariwala, H. A., Gu, H., Ng, L. L., Palmiter, R. D., Hawrylycz, M. J., Jones, A. R., Lein, E. S., Zeng, H., Hatim, A., Gu, H., Ng, L. L., Palmiter, R. D., Hawrylycz, M. J., Allan, R., Lein, E. S., & Zeng, H. (2010). A robust and high-throughput Cre Repooting and characterization. *Nat Neurosci*, *13*(1), 133–140. https://doi.org/10.1038/nn.2467.A

Mani, R., St. Onge, R. P., Hartman IV, J. L., Giaever, G., & Roth, F. P. (2008a). Defining genetic interaction. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(9), 3461–3466. https://doi.org/10.1073/pnas.0712255105

Mani, S. A., Guo, W., Liao, M. J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., Brooks, M., Reinhard, F., Zhang, C. C., Shipitsin, M., Campbell, L. L., Polyak, K., Brisken, C., Yang, J., & Weinberg, R. A. (2008b). The Epithelial-Mesenchymal Transition Generates Cells with Properties of Stem Cells. *Cell*, *133*(4), 704–715. https://doi.org/10.1016/j.cell.2008.03.027

Marjamaa, J., Tulamo, R., Abo-Ramadan, U., Hakovirta, H., Frösen, J., Rahkonen, O., Niemelä, M., Bornstein, P., Penttinen, R., & Kangasniemi, M. (2006). Mice with a deletion in the first intron of the Col1a1 gene develop dissection and rupture of aorta in the absence of aneurysms: High-resolution magnetic resonance imaging, at 4.7 T, of the aorta and cerebral arteries. *Magnetic Resonance in Medicine*, *55*(3), 592–597. https://doi.org/10.1002/mrm.20798

Martin, J. F., Bradley, A., & Olson, E. N. (1995). The paired-like homeo box gene MHox is required for early events of skeletogenesis in multiple lineages. *Genes and Development*, *9*(10), 1237–1249. https://doi.org/10.1101/gad.9.10.1237

Mechoulam, H., & Pierce, E. A. (2003). Retinopathy of prematurity: Molecular pathology and therapeutic strategies. *American Journal of PharmacoGenomics*, *3*(4), 261–277.

https://doi.org/10.2165/00129785-200303040-00004

Mizuno, K., Boudko, S., Engel, J., & Bächinger, H. P. (2013). Vascular Ehlers-Danlos syndrome mutations in type III collagen differently stall the triple helical folding. *Journal of Biological Chemistry*, *288*(26), 19166–19176. https://doi.org/10.1074/jbc.M113.462002

Murray, S. A., Oram, K. F., & Gridley, T. (2007). Multiple functions of Snail family genes during palate development in mice. *Development*, *134*(9), 1789–1797. https://doi.org/10.1242/dev.02837

Navarro, R., Compte, M., Álvarez-Vallina, L., & Sanz, L. (2016). Immune regulation by pericytes: Modulating innate and adaptive immunity. *Frontiers in Immunology*, *7*(NOV), 1–10. https://doi.org/10.3389/fimmu.2016.00480

Nieto, M. A., Huang, R. Y.-J., Jackson, R. A. A., & Thiery, J. P. (2016). Emt: 2016. *Cell*, *166*(1), 21–45. https://doi.org/10.1016/j.cell.2016.06.028

Nieto, M. A., Bennett, M. F., Sargent, M. G., & Wilkinson, D. G. (1992). Cloning and developmental expression of Sna, a murine homologue of the Drosophila snail gene. *Development*, *116*(1), 227–237. https://doi.org/10.13039/501100000780

Nieto, M. A., Sargent, M. G., Wilkinson, D. G., & Cooke, J. (1994). Control of cell behavior during vertebrate development by Slug, a zinc finger gene. *Science*, *264*(5160), 835–839. https://doi.org/10.1126/science.7513443

Nikolakopoulou, A. M., Montagne, A., Kisler, K., Dai, Z., Wang, Y., Huuskonen, M. T., Sagare, A. P., Lazic, D., Sweeney, M. D., Kong, P., Wang, M., Owens, N. C., Lawson, E. J., Xie, X., Zhao, Z., & Zlokovic, B. V. (2019). Pericyte loss leads to circulatory failure and pleiotrophin depletion causing neuron loss. *Nature Neuroscience*, *22*(7), 1089–1098. https://doi.org/10.1038/s41593-019-0434-z

Norris, R. A., & Kern, M. J. (2001). The Identification of Prx1 Transcription Regulatory Domains Provides a Mechanism for Unequal Compensation by the Prx1 and Prx2 Loci. *Journal of Biological Chemistry*, *276*(29), 26829–26837. https://doi.org/10.1074/jbc.M100239200

Norris, R. A., Scott, K. K., Moore, C. S., Stetten, G., Brown, C. R., Jabs, E. W., Wulfsberg,

E. A., Yu, J., & Kern, M. J. (2000). Human PRRX1 and PRRX2 genes: Cloning, expression, genomic localization, and exclusion as disease genes for Nager syndrome. *Mammalian Genome*, *11*(11), 1000–1005. https://doi.org/10.1007/s003350010193

O'Byrne, K. J., Dalgleish, A. G., Browning, M. J., Steward, W. P., & Harris, A. L. (2000). The relationship between angiogenesis and the immune response in carcinogenesis and the progression of malignant disease. *European Journal of Cancer*, *36*(2), 151–169. https://doi.org/10.1016/S0959-8049(99)00241-5

Ocaña, O. H., Córcoles, R., Fabra, Á., Moreno-Bueno, G., Acloque, H., Vega, S., Barrallo-Gimeno, A., Cano, A., & Nieto, M. A. (2012). Metastatic Colonization Requires the Repression of the Epithelial-Mesenchymal Transition Inducer Prrx1. *Cancer Cell*, *22*(6), 709–724. https://doi.org/10.1016/j.ccr.2012.10.012

Ocaña, O. H., Coskun, H., Minguillón, C., Murawala, P., Tanaka, E. M., Galcerán, J., Muñoz-Chápuli, R., & Nieto, M. A. (2017). A right-handed signalling pathway drives heart looping in vertebrates. *Nature*, *549*(7670), 86–90. https://doi.org/10.1038/nature23454

Oshima, M., Oshima, H., & Taketo, M. M. (1996). TGF-β receptor type II deficiency results in defects of yolk sac hematopoiesis and vasculogenesis. *Developmental Biology*, *179*(1), 297–302. https://doi.org/10.1006/dbio.1996.0259

Ozaki, H., Yu, A. Y., Della, N., Ozaki, K., Luna, J. D., Yamada, H., Hackett, S. F., Okamoto, N., Zack, D. J., Semenza, G. L., & Campochiaro, P. A. (1999). Hypoxia inducible factor-1α is increased in ischemic retina: Temporal and spatial correlation with VEGF expression. *Investigative Ophthalmology and Visual Science*, *40*(1), 182–189.

Park, D. Y., Lee, J., Kim, J., Kim, K., Hong, S., Han, S., Kubota, Y., Augustin, H. G., Ding, L., Kim, J. W., Kim, H., He, Y., Adams, R. H., & Koh, G. Y. (2017). Plastic roles of pericytes in the blood-retinal barrier. *Nature Communications*, *8*(May), 1–16. https://doi.org/10.1038/ncomms15296.

Passaniti, A., Taylor, R. M., Pili, R., Guo, Y., Long, P. V, Haney, J. A., Pauly, R. R., Grant, D. S., & Martin, G. R. (1992). A simple, quantitative method for assessing angiogenesis and antiangiogenic agents using reconstituted basement membrane, heparin, and fibroblast growth factor. *Laboratory Investigation; a Journal of Technical Methods and Pathology*, *67*(4), 519–528.

References

Patan, S. (1998). TIE1 and TIE2 receptor tyrosine kinases inversely regulate embryonic angiogenesis by the mechanism of intussusceptive microvascular growth. *Microvascular Research*, *56*(1), 1–21. https://doi.org/10.1006/mvre.1998.2081

Ponce, M. L. (2009). Tube formation: an in vitro matrigel angiogenesis assay. In *Methods in molecular biology (Clifton, N.J.)* (Vol. 467). https://doi.org/10.1007/978-1-59745-241-0_10

Pope, F. M., Narcisi, P., Nicholls, A. C., Germaine, D., Pals, G., & Richards, A. J. (1996). COL3A1 mutations cause variable clinical phenotypes including acrogeria and vascular rupture. *British Journal of Dermatology*, *135*(2), 163–181. https://doi.org/10.1111/j.1365-2133.1996.tb01143.x

Pouget, C., Pottin, K., & Jaffredo, T. (2008). Sclerotomal origin of vascular smooth muscle cells and pericytes in the embryo. *Developmental Biology*, *315*(2), 437–447. https://doi.org/10.1016/j.ydbio.2007.12.045

Pourquié, O. (2003). The segmentation clock: Converting embryonic time into spatial pattern. *Science*, *301*(5631), 328–330. https://doi.org/10.1126/science.1085887

Rago, L., Castroviejo, N., Fazilaty, H., Garcia-Asencio, F., Ocaña, O. H., Galcerán, J., & Nieto, M. A. (2019). MicroRNAs Establish the Right-Handed Dominance of the Heart Laterality Pathway in Vertebrates. *Developmental Cell*, *51*(4). https://doi.org/10.1016/j.devcel.2019.09.012

Rahkonen, O., Su, M., Hakovirta, H., Koskivirta, I., Hormuzdi, S. G., Vuorio, E., Bornstein, P., & Penttinen, R. (2004). Mice with a Deletion in the First Intron of the Col1a1 Gene Develop Age-Dependent Aortic Dissection and Rupture. *Circulation Research*, *94*(1), 83–90. https://doi.org/10.1161/01.RES.0000108263.74520.15

Rowe, R. G., Li, X. Y., Hu, Y., Saunders, T. L., Virtanen, I., De Herreros, A. G., Becker, K. F., Ingvarsen, S., Engelholm, L. H., Bommer, G. T., Fearon, E. R., & Weiss, S. J. (2009). Mesenchymal cells reactivate Snail1 expression to drive three-dimensional invasion programs. *Journal of Cell Biology*, *184*(3), 399–408. https://doi.org/10.1083/jcb.200810113

Ruzankina, Y., Pinzon-guzman, C., Asare, A., Ong, T., Cotsarelis, G., Zediak, V. P., Velez,

M., Bhandoola, A., & Brown, E. J. (2007). Deletion of the Developmentally Essential Gene ATR in Adult Mice. *Cell Stem Cell*, *1*(1), 113–126. https://doi.org/10.1016/j.stem.2007.03.002.Deletion

Sagare, A. P., Bell, R. D., Zhao, Z., Ma, Q., Winkler, E. A., Ramanathan, A., & Zlokovic, B. V. (2013). Pericyte loss influences Alzheimer-like neurodegeneration in mice. *Nature Communications*, *4*, 1–14. https://doi.org/10.1038/ncomms3932

Sato, M., Suzuki, S., & Senoo, H. (2003). Hepatic stellate cells: Unique characteristics in cell biology and phenotype. *Cell Structure and Function*, *28*(2), 105–112. https://doi.org/10.1247/csf.28.105

Sefton, M., Sánchez, S., & Nieto, M. A. (1998). Conserved and divergent roles for members of the Snail family of transcription factors in the chick and mouse embryo. *Development*, *125*(16), 3111–3121. https://doi.org/10.13039/501100000780

Seki, E., & Schwabe, R. F. (2015). Hepatic inflammation and fibrosis: Functional links and key pathways. *Hepatology*, *61*(3), 1066–1079. https://doi.org/10.1002/hep.27332

Selever, J., Liu, W., Lu, M. F., Behringer, R. R., & Martin, J. F. (2004). Bmp4 in limb bud mesoderm regulates digit pattern by controlling AER development. *Developmental Biology*, *276*(2), 268–279. https://doi.org/10.1016/j.ydbio.2004.08.024

Selvam, S., Kumar, T., & Fruttiger, M. (2018). Retinal vasculature development in health and disease. *Progress in Retinal and Eye Research*, *63*, 1–19. https://doi.org/10.1016/j.preteyeres.2017.11.001

Sengillo, J. D., Ethan A. Winkler, Corey T. Walker, John S. Sullivan, Mahlon Johnson, and Berislav V. Zlokovic1. (2014). DEFICIENCY IN MURAL VASCULAR CELLS COINCIDES WITH BLOOD-BRAIN BARRIER DISRUPTION IN ALZHEIMER'S DISEASE. *Brain Pathology*, *23*(3), 303–310. https://doi.org/10.1111/bpa.12004.DEFICIENCY

Shen, E. M., & McCloskey, K. E. (2017). Development of Mural Cells: From in Vivo Understanding to in Vitro Recapitulation. *Stem Cells and Development*, *26*(14), 1020–1041. https://doi.org/10.1089/scd.2017.0020

Shepro, D., & Morel, N. M. L. (1993). Pericyte physiology. *The FASEB Journal*, *7*(11), 1031–

1038. https://doi.org/10.1096/fasebj.7.11.8370472

Smith, L. B., Hadoke, P. W. F., Dyer, E., Denvir, M. A., Brownstein, D., Miller, E., Nelson, N., Wells, S., Cheeseman, M., & Greenfield, A. (2011). Haploinsufficiency of the murine Col3a1 locus causes aortic dissection: A novel model of the vascular type of EhlersDanlos syndrome. *Cardiovascular Research*, *90*(1), 182–190. https://doi.org/10.1093/cvr/cvq356

Smith, L. E. H., Shen, W., Perruzzi, C., Soker, S., Kinose, F., Xu, X., Robinson, G., Driver, S., Bischoff, J., Zhang, B., Schaeffer, J. M., & Senger, D. R. (1999). Regulation of vascular endothelial growth factor-dependent retinal neovascularization by insulin-like growth factor-1 receptor. *Nature Medicine*, *5*(12), 1390–1395. https://doi.org/10.1038/70963

Soldatov, R., Kaucka, M., Kastriti, M. E., Petersen, J., Chontorotzea, T., Englmaier, L., Akkuratova, N., Yang, Y., Häring, M., Dyachuk, V., Bock, C., Farlik, M., Piacentino, M. L., Boismoreau, F., Hilscher, M. M., Yokota, C., Qian, X., Nilsson, M., Bronner, M. E., … Adameyko, I. (2019). Spatiotemporal structure of cell fate decisions in murine neural crest. *Science*, *364*(6444). https://doi.org/10.1126/science.aas9536

Soo, K., O'Rourke, M. P., Khoo, P. L., Steiner, K. A., Wong, N., Behringer, R. R., & Tam, P. P. L. (2002). Twist function is required for the morphogenesis of the cephalic neural tube and the differentiation of the cranial neural crest cells in the mouse embryo. *Developmental Biology*, *247*(2), 251–270. https://doi.org/10.1006/dbio.2002.0699

Stahl, A., Connor, K. M., Sapieha, P., Chen, J., Dennison, R. J., Krah, N. M., Seaward, M. R., Willett, K. L., Aderman, C. M., Guerin, K. I., Hua, J., Löfqvist, C., Hellström, A., & Smith, L. E. H. (2010). The mouse retina as an angiogenesis model. *Investigative Ophthalmology and Visual Science*, *51*(6), 2813–2826. https://doi.org/10.1167/iovs.10-5176

Stern, C. D. (2006). Neural induction: 10 years on since the "default model." *Current Opinion in Cell Biology*, *18*(6), 692–697. https://doi.org/10.1016/j.ceb.2006.09.002

Stratman, A. N., & Davis, G. E. (2012). Endothelial Cell-Pericyte Interactions Stimulate Basement Membrane Matrix Assembly. *Microscopy and Microanalysis*, *18*(1), 68–80.

References

Stratman, A. N., & Davis, G. E. (2017). Microscopy Microanalysis Endothelial Cell-Pericyte Interactions Stimulate Basement Membrane Matrix Assembly : Influence on Vascular Tube. *Microscopy and Microanalysis : The Official Journal of Microscopy Society of America, Microbeam Analysis Society, Microscopical Society of Canada*, *18*(1), 68–80. https://doi.org/10.1017/S1431927611012402

Suri, C., Jones, P. F., Patan, S., Bartunkova, S., Maisonpierre, P. C., Davis, S., Sato, T. N., & Yancopoulos, G. D. (1996). Requisite role of angiopoietin-1, a ligand for the TIE2 receptor, during embryonic angiogenesis. *Cell*, *87*(7), 1171–1180. https://doi.org/10.1016/S0092-8674(00)81813-9

Sweeney, M. D., Kisler, K., Montagne, A., Toga, A. W., & Zlokovic, B. V. (2018). The role of brain vasculature in neurodegenerative disorders. *Nature Neuroscience*, *21*(10), 1318–1331. https://doi.org/10.1038/s41593-018-0234-x

Takagi, T., Moribe, H., Kondoh, H., & Higashi, Y. (1998). δEF1, a zinc finger and homeodomain transcription factor, is required for skeleton patterning in multiple lineages. *Development*, *125*(1), 21–31.

Tang, Y., Feinberg, T., Keller, E. T., Li, X. Y., & Weiss, S. J. (2016). Snail/Slug binding interactions with YAP/TAZ control skeletal stem cell self-renewal and differentiation. *Nature Cell Biology*, *18*(9), 917–929. https://doi.org/10.1038/ncb3394

The Lancet Neurology. Editorial. (2017). Vascular disease and neurodegeneration: advancing together. *The Lancet Neurology*, *16*(5), 333. https://doi.org/10.1016/S1474-4422(17)30086-8

Thiery, J. P., Acloque, H., Huang, R. Y. J., & Nieto, M. A. (2009). Epithelial-Mesenchymal Transitions in Development and Disease. *Cell*, *139*(5), 871–890. https://doi.org/10.1016/j.cell.2009.11.007

Thisse, B., Stoetzel, C., Gorostiza-Thisse, C., & Perrin-Schmitt, F. (1988). Sequence of the twist gene and nuclear localization of its protein in endomesodermal cells of early Drosophila embryos. *The EMBO Journal*, *7*(7), 2175–2183. https://doi.org/10.1002/j.1460-2075.1988.tb03056.x

Tigges, U., Hyer, E. G., Scharf, J., & Stallcup, W. B. (2008). FGF2-dependent

neovascularization of subcutaneous Matrigel plugs is initiated by bone marrow-derived pericytes and macrophages. *Development*, *135*(3), 523–532. https://doi.org/10.1242/dev.002071

Tischfield, M. A., Robson, C. D., Gilette, N. M., Chim, S. M., Sofela, F. A., DeLisle, M. M., Gelber, A., Barry, B. J., MacKinnon, S., Dagi, L. R., Nathans, J., & Engle, E. C. (2017). Cerebral Vein Malformations Result from Loss of Twist1 Expression and BMP Signaling from Skull Progenitor Cells and Dura. *Developmental Cell*, *42*(5), 445-461.e5. https://doi.org/10.1016/j.devcel.2017.07.027

Tsai, J. H., Donaher, J. L., Murphy, D. A., Chau, S., & Yang, J. (2012). Spatiotemporal regulation of EMT is Essential for Squamous Cell Carcinoma Metastasis. *Changes*, *29*(6), 997–1003. https://doi.org/10.1016/j.ccr.2012.09.022.Spatiotemporal

Uemura, A., Ogawa, M., Hirashima, M., Fujiwara, T., Koyama, S., Takagi, H., Honda, Y., Wiegand, S. J., Yancopoulos, G. D., & Nishikawa, S.-I. (2002). Recombinant angiopoietin-1 restores higher-order architecture of growing blood vessels in mice in the absence of mural cells. *Journal of Clinical Investigation*, *110*(11), 1619–1628. https://doi.org/10.1172/jci200215621

Vanlandewijck, M., He, L., Mäe, M. A., Andrae, J., Ando, K., Del Gaudio, F., Nahar, K., Lebouvier, T., Laviña, B., Gouveia, L., Sun, Y., Raschperger, E., Räsänen, M., Zarb, Y., Mochizuki, N., Keller, A., Lendahl, U., & Betsholtz, C. (2018a). A molecular atlas of cell types and zonation in the brain vasculature. *Nature*, *554*(7693), 475–480. https://doi.org/10.1038/nature25739

Vanlandewijck, M., Vanlandewijck, M., Lebouvier, T., Mäe, M. A., Nahar, K., & Betsholtz, C. (2018b). Primary isolation of vascular cells from murine brain for single cell sequencing. *Protocol Exchange*, 1–11. https://doi.org/10.1038/protex.2017.159

Volz, K. S., Jacobs, A. H., Chen, H. I., Poduri, A., McKay, A. S., Riordan, D. P., Kofler, N., Kitajewski, J., Weissman, I., & Red-Horse, K. (2015). Pericytes are progenitors for coronary artery smooth muscle. *ELife*, *4*, 1–22. https://doi.org/10.7554/elife.10036

Wasteson, P., Johansson, B. R., Jukkola, T., Breuer, S., Akydsurek, L. M., Partanen, J., & Lindahl, P. (2008). Developmental origin of smooth muscle cells in the descending aorta in mice. *Development*, *135*(10), 1823–1832. https://doi.org/10.1242/dev.020958

References

Xian, X., Håkansson, J., Ståhlberg, A., Lindblom, P., Betsholtz, C., Gerhardt, H., & Semb, H. (2006a). Pericytes limit tumor cell metastasis. *Journal of Clinical Investigation*, *116*(3), 642–651. https://doi.org/10.1172/JCI25705

Xu, Y., Lee, D. K., Feng, Z., Xu, Y., Bu, W., Li, Y., Liao, L., & Xu, J. (2017). Breast tumor cell-specific knockout of Twist1 inhibits cancer cell plasticity, dissemination, and lung metastasis in mice. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(43), 11494–11499. https://doi.org/10.1073/pnas.1618091114

Yamazaki, T., Nalbandian, A., Uchida, Y., Li, W., Arnold, T. D., Kubota, Y., Yamamoto, S., Ema, M., & Mukouyama, Y. suke. (2017). Tissue Myeloid Progenitors Differentiate into Pericytes through TGF-β Signaling in Developing Skin Vasculature. *Cell Reports*, *18*(12), 2991–3004. https://doi.org/10.1016/j.celrep.2017.02.069

Yang, A. C., Stevens, M. Y., Chen, M. B., Lee, D. P., Stähli, D., Gate, D., Contrepois, K., Chen, W., Iram, T., Zhang, L., Vest, R. T., Chaney, A., Lehallier, B., Olsson, N., du Bois, H., Hsieh, R., Cropper, H. C., Berdnik, D., Li, L., … Wyss-Coray, T. (2020a). Physiological blood–brain transport is impaired with age by a shift in transcytosis. *Nature*, *583*(7816), 425–430. https://doi.org/10.1038/s41586-020-2453-z

Yang, J. *et al.* On behalf of the EMT International Association (TEMTIA). (2020b). Guidelines and definitions for research on epithelial–mesenchymal transition. Nature Reviews Molecular Cell Biology, 21(6), 341–352. https://doi.org/10.1038/s41580-020-0237-9

Yang, J., Mani, S. A., Donaher, J. L., Ramaswamy, S., Itzykson, R. A., Come, C., Savagner, P., Gitelman, I., Richardson, A., & Weinberg, R. A. (2004). Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell*, *117*(7), 927–939. https://doi.org/10.1016/j.cell.2004.06.006

Yang, X., Castilla, L. H., Xu, X., Li, C., Gotay, J., Weinstein, M., Liu, P. P., & Deng, C. X. (1999). Angiogenesis defects and mesenchymal apoptosis in mice lacking SMAD5. *Development*, *126*(8), 1571–1580.

Yeo, S. Y., Lee, K. W., Shin, D., An, S., Cho, K. H., & Kim, S. H. (2018). A positive feedback loop bi-stably activates fibroblasts. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-05274-6

References

Yonenaga, Y., Mori, A., Onodera, H., Yasuda, S., Oe, H., Fujimoto, A., Tachibana, T., & Imamura, M. (2005). Absence of smooth muscle actin-positive pericyte coverage of tumor vessels correlates with hematogenous metastasis and prognosis of colorectal cancer patients. *Oncology*, *69*(2), 159–166. https://doi.org/10.1159/000087840

You-Tzung Chen, Peter O. Akinwunmi, Jian Min Deng, Oliver H. Tam, and R. R. B. (2007). Generation of a Twist1 Conditional Null Allele in the Mouse. *Genesis*, *45*(2), 76–82. https://doi.org/10.1002/dvg

Zhang, Y., Cedervall, J., Hamidi, A., Herre, M., Viitaniemi, K., D'Amico, G., Miao, Z., Unnithan, R. V. M., Vaccaro, A., van Hooren, L., Georganaki, M., Thulin, Å., Qiao, Q., Andrae, J., Siegbahn, A., Heldin, C. H., Alitalo, K., Betsholtz, C., Dimberg, A., & Olsson, A. K. (2020). Platelet-Specific PDGFB Ablation Impairs Tumor Vessel Integrity and Promotes Metastasis. *Cancer Research*, *80*(16), 3345–3358. https://doi.org/10.1158/0008-5472.CAN-19-3533

Annex: Publication

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-019-13091-8 OPEN

A gene regulatory network to control EMT programs in development and disease

Hassan Fazilaty [1,2], Luciano Rago [1,3], Khalil Kass Youssef [1], Oscar H. Ocaña [1], Francisco Garcia-Asencio [1], Aida Arcas [1,4], Juan Galceran [1] & M. Angela Nieto [1*]

The Epithelial to Mesenchymal Transition (EMT) regulates cell plasticity during embryonic development and in disease. It is dynamically orchestrated by transcription factors (EMT-TFs), including Snail, Zeb, Twist and Prrx, all activated by TGF-β among other signals. Here we find that Snail1 and Prrx1, which respectively associate with gain or loss of stem-like properties and with bad or good prognosis in cancer patients, are expressed in complementary patterns during vertebrate development and in cancer. We show that this complementarity is established through a feedback loop in which Snail1 directly represses *Prrx1*, and Prrx1, through direct activation of the miR-15 family, attenuates the expression of Snail1. We also describe how this gene regulatory network can establish a hierarchical temporal expression of Snail1 and Prrx1 during EMT and validate its existence in vitro and in vivo, providing a mechanism to switch and select different EMT programs with important implications in development and disease.

[1] Instituto de Neurociencias (CSIC-UMH), Avda. Ramón y Cajal s/n, Sant Joan d', Alacant 03550, Spain. [2] Present address: Institute of Molecular Life Sciences, University of Zurich, 8057 Zurich, Switzerland. [3] Present address: Department of Oncogenomics, Academic Medical Center, Amsterdam, the Netherlands. [4] Present address: Department of Gene Therapy and Regulation of Gene Expression, Center for Applied Medical Research, University of Navarra, Pamplona, Spain. *email: anieto@umh.es

The epithelial to mesenchymal transition (EMT) is a developmental process that can be ectopically reactivated in diseases like cancer. When they undergo EMT, epithelial immotile cells become mesenchymal, acquiring the ability to migrate and invade[1]. During tumor progression, EMT provides invasive and migratory properties to cancer cells[1–4]. This phenotypic transition is governed by extracellular signals that activate a plethora of EMT transcription factors (EMT-TFs) that include those belonging to the Snail, Zeb, Twist, and Prrx families[1]. Different mesenchymal cells of embryonic, healthy adult or pathological origin, express different combinations of EMT-TFs, leading to a tissue-specific EMT-TF code[5], which may influence the overall state, function, and behavior of the cell. In addition to the transition toward the mesenchymal phenotypes, Snail1, Twist1, and Zeb1 can also induce stemness[6,7], while Prrx1 expression is concomitant with the loss of stemness[8,9]. In cancer, highSnail1 expression is associated with malignant phenotype and poor prognosis[10,11], while the high expression of Prrx1 is associated with good prognosis and metastasis-free disease[8]. In the chicken embryo, *PRRX1* and *SNAIL1* are expressed in a complementary manner[8] and in breast cancer Prrx1 expression correlates with that of Twist1 but not Snail1[8]. These differences can be considered as different EMT modes associated with the dominant EMT-TF in a given cellular context[5]. Studying the differences between all these EMT-TFs is important to understand cell plasticity during embryonic development, which can ultimately help to distinguish the key altered cellular and molecular mechanisms in disease.

Combined expression of *SNAIL1* and *PRRX1* covers almost the entire mesenchymal cell population in the chicken embryo[8]. Although there are clear differences in the EMT activated by each factor in development and cancer, the two are activated by the same extracellular signals, the transforming growth factor beta (TGF-β) superfamily[8,12]. Therefore, we want to assess whether there is a crosstalk between Snail1 and Prrx1, by which each factor promotes its own EMT mode, particularly by differential regulation of stemness.

Here, we describe a gene regulatory network (GRN) by which Snail1 directly represses *Prrx1* transcription, and Prrx1, through direct activation of the miR-15 family, attenuates Snail1 expression. We find that Snail1 is a direct target of these microRNAs (miRNAs) among different vertebrate species. miRNAs are short noncoding RNAs that posttranscriptionally regulate their target genes[13], and are crucial players in regulating cell plasticity and EMT[14]. We also find that this GRN triggers an expression switch from Snail1 to Prrx1, with Snail1 being an early response gene to EMT-inducing signals, followed by the activation of Prrx1 that in turn attenuates Snail1 expression. We support our findings by analyses in cultured cells, in vivo in different vertebrate embryos and public databases of cancer patients. We illustrate that this GRN rather than regulating the balance between epithelial and mesenchymal states as the previously described networks involving microRNAs, drives the selection of the EMT mode.

Results

Prrx1 and Snail1 are expressed in complementary patterns. In zebrafish embryos, which bear two paralogs for each gene (*prrx1a* and *prrx1b; snail1a* and *snail1b)* due to the extra duplication in the teleost genome[3,15], we performed RNA in situ hybridization (ISH) and found a complementary expression pattern. In the developing somites where *snail* genes are abundantly expressed, *prrx1* genes expression are restricted to small cell populations where *snail* expression is low or absent (Fig. 1a). Although at 20-somite stage both *snail1* and *prrx1* are expressed in the cranial neural crest (Fig. 1a), transverse sections of double-fluorescent ISH shows that they are also expressed in a complementary manner (Fig. 1b). Single-cell RNA sequencing (scRNA-seq) data from zebrafish embryos at 18 h post fertilization (hpf) (GEO: GSM3067194)[16] provides further evidence for this complementary expression of *snail1a/b* and *prrx1a/b* in the majority of cells, with a significant negative correlation (Fig. 1c, Supplementary Fig. 1a). This is compatible with our previous findings in the chicken embryo[8] (Fig. 1d).

To assess whether this complementarity was also conserved in mammals, we characterized the expression of Snail1 and Prrx1 in the mouse. At E8.5, embryos manifest a complementary expression pattern for *Snail1* and *Prrx1* at whole-mount level, particularly evident in the mesodermal populations along the medio-lateral axis and in the somites (Fig. 1e). Transverse sections of anterior regions (Fig. 1e) confirm this complementarity in the neural crest populations, where *Snail1* is expressed in the premigratory crest (arrow) and both are expressed in migratory neural crest populations (arrowheads) in a distinct manner. In more posterior tissues (Figs. 1e and dashed line2), the complementarity is evident in the mesodermal cells, where *Snail1* is highly expressed in the pre-somitic mesoderm (arrow) while *Prrx1* is expressed in the lateral plate mesoderm (LPM, arrowhead). Expression is also complementary at E9.5 mouse embryos in the somites and as in the chick, *Snail1* is highly expressed in immature somites where *Prrx1* expression is not evident. Also, LPM cells are positive for *Prrx1* expression but not *Snail1* (Supplementary Fig. 1b). Analysis of transverse sections at the level of the branchial arches, also showed complementarity, where in more internal cell populations *Snail1* was highly expressed concomitant with low *Prrx1* expression, while more ventral cell populations showed high expression of *Prrx1* and lower of *Snail1* (Supplementary Fig. 1c). A similarly complementary pattern was also evident in the mature somites. *Snail1* is highly expressed in the sclerotome (SC) and *Prrx1* in the dermomyotome (DM) (Supplementary Fig. 1d). The analysis of scRNA-seq from E9.5 embryos (GEO: GSE87038)[17] clearly shows this complementary expression at the single-cell level (Fig. 1f). These data confirm that *Snail1* and *Prrx1* are expressed in a conserved dynamic and complementary manner in cells undergoing EMT during vertebrate development.

To determine whether this complementary expression pattern holds true in pathology, we first looked at the expression levels of SNAIL1 and PRRX1 in human breast cancer cell lines[18], and did not find cells with both *SNAIL1* and *PRRX1* high. Supplementary Figure 1e shows examples of cell lines with high-SNAIL1 expression (SUM149PT); cells with intermediate/low levels of both (MDA436), and cells with high-PRRX1 expression (BT549). These three cell lines are derived from triple-negative/basal type breast tumors, and manifest mesenchymal properties[19], including migratory and invasive behaviors[8,20]. However, SUM149PT and MDA436 cells are metastatic and show cancer stem cell-like properties[21], while BT549 cells need to lose PRRX1 expression to acquire both stem cell properties and metastatic potential[8]. Thus, the comparison of the three cell lines for *SNAIL1* and *PRRX1* expression can provide insights into their potential regulation. Signal intensity analyses for SNAIL1 and PRRX1 proteins show a negative correlation in SUM149PT and BT549. Interestingly, MDA436 expresses both SNAIL1 and PRRX1 at similar intermediate levels (Supplementary Fig. 1f), altogether suggesting that SNAIL1 or PRRX1 cannot be co-expressed at high levels.

Next, we examined databases from breast cancer patients (GEO: GSE75688)[22], and found that *SNAIL1* and *PRRX1* depict a complementary expression pattern with a significant negative correlation (Fig. 1g). In addition, in a larger scRNA-seq dataset from head and neck carcinoma patients (GEO: GSE103322)[23], *SNAIL1* and *PRRX1* expressions were also complementary and

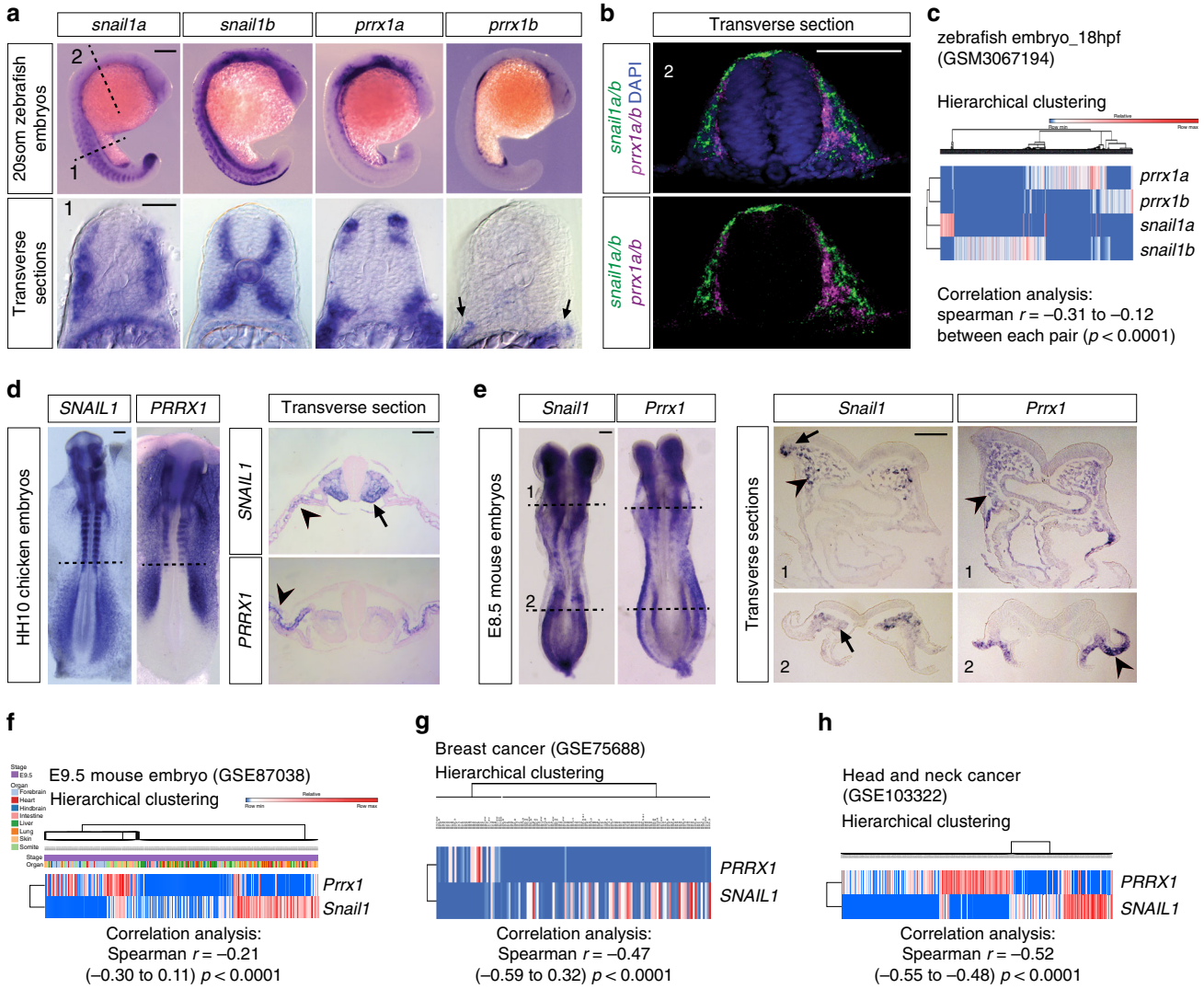


Fig. 1 Snail1 and Prrx1 complementary expression in development and disease. **a** Lateral view of 20-somite zebrafish embryos showing *snail1a*, *snail1b*, *prrx1a* and *prrx1b* expression in whole-mount (top) and transverse sections (1), showing complementary patterns in somites. **b** Transverse section of a zebrafish embryo in the cranial neural crest region showing complementary expression of *snail1a/b* (green) and *prrx1a/b* (red) taken at the level indicated by (2) in (**a**) with or without DAPI staining (nuclei). **c** Heatmap showing hierarchical clustering of scRNA-seq data from 18 hpf zebrafish embryos, from public database GEO: GSM3067194, with significant negative correlations between gene pairs (detailed in Supplementary Fig. 1a). **d** Dorsal view of HH10 chicken embryos showing *PRRX1* and *SNAIL1* expression in whole-mount and transverse sections at the level indicated by dashed lines, showing complementary patterns for *SNAIL1* and *PRRX1* In the somites (arrow) and in the LPM (splanchnopleura and somatopleura, respectively; arrowheads). **e** Expression of *Snail1* and *Prrx1* in dorsal views of E8.5 mouse embryos. Transverse sections of E8.5 embryos from the regions indicated by dashed lines (anterior and posterior, 1 and 2, respectively), showing complementary expression of *Snail1* and *Prrx1* in premigratory (1, arrow), and migratory (1, arrowheads) neural crest (PNC and MNC, respectively) and mesodermal populations including presomitic mesoderm (2, arrow) and lateral plate mesoderm (2, arrowhead). **f** Hierarchical clustering heatmaps of scRNA-seq data from E9,5 mouse embryos for *Snail1* and *Prrx1* expression (GEO: GSE87038), showing a significant negative correlation. This dataset contains cells from different embryonic tissues, shown in different colors. **g**, **h** Heatmaps of hierarchical clusterings of single-cell RNA sequencing data from public datasets of breast (GEO: GSE75688) and head and neck carcinoma (GEO: GSE103322), showing mutually exclusive expression of Snail1 and Prrx1. The color scale shown in (**f**) is also valid for (**g**) and (**h**), representing the normalized values for the number of reads. Scale bars: 250 µm for whole mounts and 100 µm for sections. Statistical analyses for scRNA-seq data has been done using Spearman *r* correlation test, scRNA-seq single-cell RNA sequencing, vs. versus, hpf hours post fertilization. Source data are provided as a Source Data file

negatively correlated, with some small cell populations expressing both (Fig. 1h). Notably, in all the analyzed scRNA-seq datasets, extremely few cells were found to express both EMT-TFs at high levels (Supplementary Table 1).

Taken together, these findings suggest that there may be a regulatory mechanism between these two EMT-TFs that is conserved not only during vertebrate development but also in pathological EMTs. Therefore, we next wanted to understand the molecular basis of the complementary expression of Snail1 and Prrx1. Considering that they are TF and there have been previous examples of mutually regulated TF pairs during embryonic development[24,25], we set up to study this possibility for Snail1 and Prrx1. To test this hypothesis, we first examined the response in terms of gene expression after acute downregulation of either SNAIL1 or PRRX1 in the same cell context. We used MDA436 cells, which express both TFs at low/intermediate levels, and found that SNAIL1 knockdown (KD) using short interfering RNA (siRNA) resulted in an increased expression of *PRRX1* and

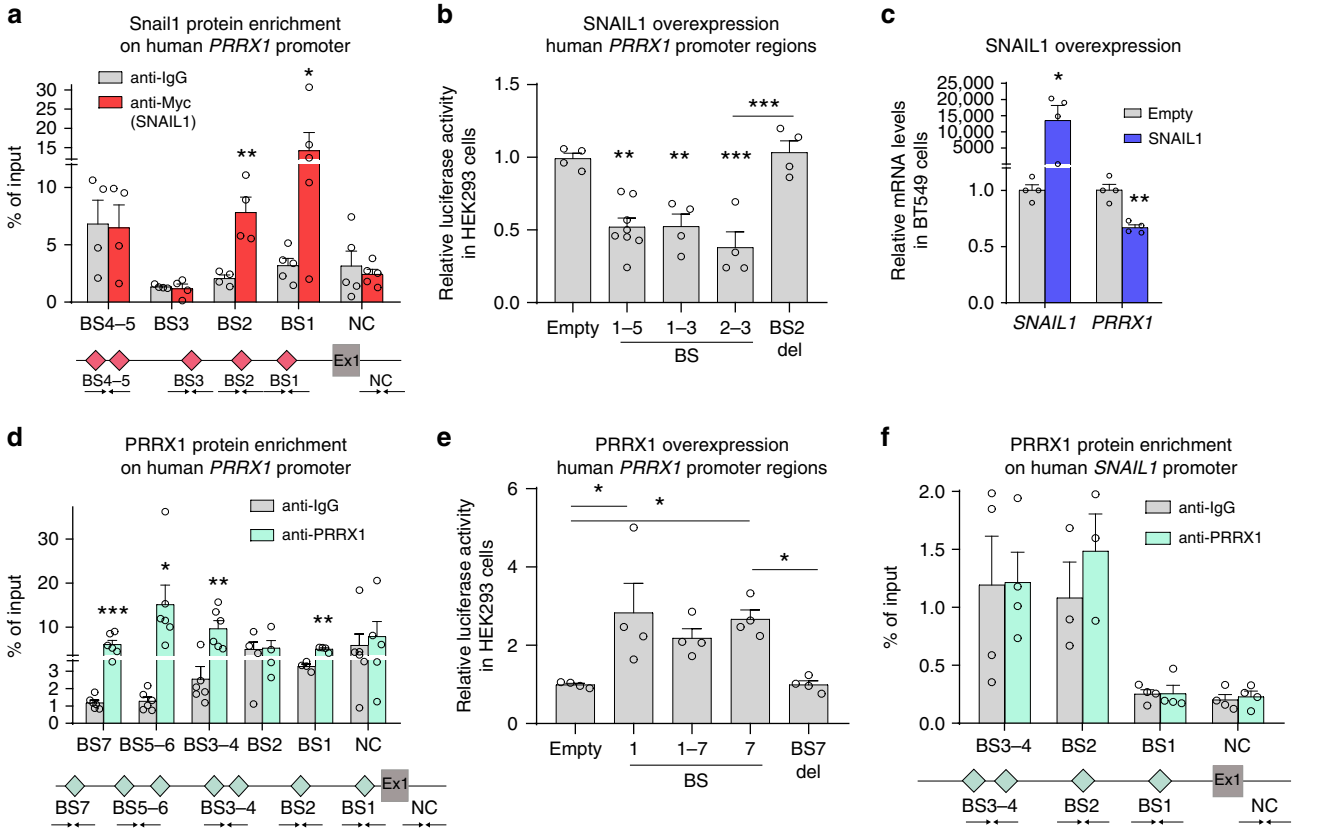


Fig. 2 SNAIL1 and PRRX1 act antagonistically on *PRRX1* promoter. **a** Snail1 enrichment in the human *PRRX1* promoter shown by ChIP assay in BT549 cells, using anti Myc antibody (for Snail1-Myc overexpression). A schematic map is shown; red diamonds represent SNAIL1 potential binding sites (E-boxes; CANNTG) on *PRRX1* promoter. (BS1: -1689, BS2: −4102, BS3: −6753, BS4: −7277, BS5: −7318 and NC: +104560). Ex1: *PRRX1* exon 1 ($n = 4$). **b** Activity of different regions of the human *PRRX1* promoter after SNAIL1 transfection assessed by luciferase assays in HEK293 cells ($n = 4$, except for BS1-5 for which $n = 8$). **c** qPCR assay showing downregulation of *PRRX1* transcription upon SNAIL1 transfection in BT549 cells ($n = 4$). **d** PRRX1 directly binds to its own promoter, as assessed by ChIP assays in BT549 cells using a PRRX1 specific antibody. A schematic map is shown; cyan diamonds represent PRRX1 potential binding sites (TAATKDS) on its own promoter. (BS7: −6875, BS6: −5579, BS5: −5202, BS4: −4147, BS3: −4096, BS2: −1653, BS1: −1197 and NC: +104560). Ex1: *PRRX1* exon 1 ($n = 6$, except for BS1-2 for which $n = 4$). **e** Activity of different regions of the human *PRRX1* promoter after PRRX1 overexpression in luciferase assays in HEK293 cells ($n = 4$). **f** Lack of enrichment for PRRX1 binding to human *SNAIL1* promoter assessed by ChIP assays in BT549 cells (BS4-1: −1753, −1621, −954, −198, and NC: +11799). Ex1: *SNAIL1* exon 1 ($n = 4$). Locations of red and cyan diamonds represent distances between BSs and the promoter. Arrows represent primers used for qPCR amplification. BS biding site, NC negative control region, which does not contain potential BSs. Symbols in binding sites are as follows. K: T/G, D: G/A/T, S: G/C and N: G/A/T/C. Bars represent mean plus standard error of the mean (SEM), indicated (n) represent number of independent experiments as biological replicates and asterisks indicate significant p value in t test compared to the control in each test for (**a**), (**e**), (**d**), (**f**) and ANOVA with Bonferroni's multiple comparison test for (**b**), (**e**) ($^*p < 0.05$, $^{**}p < 0.01$ and $^{***}p < 0.001$). Source data are provided as a Source Data file

similarly, PRRX1 KD resulted in *SNAIL1* upregulation (Supplementary Fig. 1g). These observations prompted us to investigate the molecular relationship between Snail1 and Prrx1.

Snail1 directly represses *Prrx1*. As Snail1 has been described as a potent transcriptional repressor[26], we examined whether it could directly target *Prrx1* transcription. We looked for consensus Snail1 E-boxes[27] within the mouse and human *Prrx1* promoters and putative enhancer regions, and found several predicted binding sites (BS). To narrow down the number of BSs we used ENCODE data from different cell contexts[28] using UCSC genome browser (https://genome-euro.ucsc.edu/index.html), and focused on the regions that were conserved and positive for DNase-I hypersensitivity and H3K27Ac marks, suggesting an active chromatin region. We performed chromatin immunoprecipitation (ChIP) assay in NIH3T3 mouse fibroblast cells transiently overexpressing Snail1, and found several regions with significant enrichment for Snail1 protein occupancy in the mouse *Prrx1* promoter; BS1–3 (Supplementary Fig. 2a), which results in repression of promoter activity showed using dual luciferase reporter assay (Supplementary Fig. 2b). We then analyzed the human *PRRX1* promoter and performed ChIP in BT549 cells transiently overexpressing Snail1, and found at least two sites with significant enrichment for its protein occupancy, BS1–2 (Fig. 2a). Snail1 binding leads to repression of the promoter activity, which is abolished upon deletion of BS2 (Fig. 2b). In agreement with this, *PRRX1* was downregulated after SNAIL1 transfection in BT549 (Fig. 2c). Altogether, these results indicate that Snail1 directly binds to, and represses *Prrx1* promoter activity both in mouse and human cells.

Prrx1 induces its own expression. When we searched for Snail1 E-boxes in both mouse and human *Prrx1* promoters, we also found several predicted Prrx1 BSs (TAATKDS)[29] (Fig. 2d and Supplementary Fig. 2c). Thus, we next examined whether Prrx1 could regulate its own transcription. We used above-mentioned strategy to narrow down the potential BSs. We performed ChIP for endogenous chromatin bound Prrx1 protein in NIH3T3, and

found two regions with enrichment for Prrx1 occupancy in its own promoter (BS3–4, Supplementary Fig. 2c). In human BT549 cells, endogenous PRRX1 protein occupancy was significantly enriched in several BS (Fig. 2d). Luciferase assay for both mouse and human sequences show an activation of the *Prrx1* promoter upon Prrx1 overexpression (Fig. 2e and Supplementary Fig. 2d) and this activation no longer occurs when BS7 is deleted (Fig. 2e).

Next, we wanted to know whether there was a reciprocal regulation between Snail1 and Prrx1, and we examined Prrx1 protein binding in the *Snail1* promoter. Although we found putative Prrx1 BSs within the *Snail1* proximal promoter that met our criteria, ChIP in both mouse (Supplementary Fig. 2e) and human cells (Fig. 2f) failed to show Prrx1 protein binding enrichment in those sites. In addition, activity assay for *Snail1* promoter shows no significant change when Prrx1 is overexpressed (Supplementary Fig. 2f). Altogether, these results indicate that Prrx1 is able to bind to its own promoter and enhance its own expression, but it does not bind to *Snail1* promoter, suggesting that Prrx1 is not a direct *Snail1* transcriptional repressor. In fact, Prrx1 has been described as a transcriptional activator[30], which suggests a putative indirect regulation of Snail1 expression by Prrx1.

Prrx1 directly induces expression of *miR-15* family members. To investigate downstream targets of PRRX1, we used MDA-MB-231 breast cancer cells which express low level of PRRX1 to generate a stable cell line where PRRX1 was ectopically overexpressed (MDA231-PRRX1), and performed comparative microarray analyses (GEO: GSE138078). *SNAIL1* was downregulated after PRRX1 overexpression, reinforcing the idea of a mutual negative regulation. Interestingly, we found many miRNAs that were upregulated in MDA231-PRRX1 cells (Fig. 3a and Supplementary Data 1). As miRNA are potent regulators of gene expression[13], we hypothesized that Prrx1 may repress Snail1 via recruiting miRNAs. We selected for further analyses miR-424 and miR-503 that are members of miR-15 family (hereafter referred as Mir-15-P1d and Mir-15-P2d, as recently suggested[31]; Supplementary Table 2) because they were predicted to target *Snail1*, are conserved in different vertebrates, have a relatively well known promoter, and most importantly because their studied functions as regulators of stemness, invasion, migration and cell proliferation[32–34], are compatible with those described for PRRX1[8]. We validated the microarray data by qPCR analyses (Fig. 3b).

To test whether these miRNAs could be directly induced by PRRX1, we examined the expression of their precursor miRNAs (Pre-miRNAs) upon transient overexpression of mouse Prrx1 in MDA231 cells. To do so, we generated an inducible lentiviral system (Tet-ON) including P2A peptide and overexpressed both Prrx1 and nuclear yellow fluorescent protein (nYFP)-P2A-Prrx1 (MDA231 conditional Prrx1; MDA231cP) upon doxycycline (dox) treatment. Forty-eight hour after dox induction, YFP-positive cells were FACS sorted and subjected to qPCR, showing that, as expected, Prrx1 was overexpressed and also that *Pre-miRNAs* were significantly upregulated (Fig. 3c and Supplementary Fig. 3a), suggesting that they may be direct Prrx1 targets.

As miRNA molecules of the same family share the seed sequence (nucleotides 2–7) and they can potentially target the same mRNAs, we examined whether the expression of other miR-15 family (miR-15f) members (Supplementary Fig. 3b) also correlated with that of PRRX1. Thus, we performed transient KD of PRRX1 in BT549 cells with a specific short-hairpin RNA[8] (shRNA) vector containing nYFP. We then FACS-sorted positive cells, and checked pre-miRNAs expression for an early transcriptional response. We observed that *PRRX1* transcript was significantly reduced, and also that all the members of the miR-15f were significantly downregulated upon PRRX1 KD (Fig. 3d). We also confirmed that this was the case in HEK293 cells (Supplementary Fig. 3c).

To assess whether PRRX1 directly binds the *miR-15f* promoters, we performed ChIP assays in BT459 on the human *Mir-15-P1/2d* promoter. We found enrichment for PRRX1 binding in at least one putative BS, indicating that PRRX1 can directly bind to the *Mir-15-P1/2d* promoter (Fig. 3e). Also, PRRX1 overexpression increased the activity of the *Mir-15-P1/2d* promoter, and the deletion of PRRX1 potential BS prevented this activation (Fig. 3f). On the other hand, KD of PRRX1 in BT549 cells showed a reduction in promoter activity, confirming that PRRX1 enhances *Mir-15-P1/2d* transcription (Supplementary Fig. 3d). We also performed ChIP and Luciferase assays in mouse NIH-3T3 cells and confirmed that this direct binding and activation of the *Mir-15-P1/2d* promoter is conserved between human and mouse (Supplementary Fig. 3e, f). Furthermore, ChIP assays for Prrx1 in the promoter regions of other members of the miR-15f in human and mouse, show that Prrx1 can directly bind and likely activate other members of the family (Supplementary Fig. 3g–i). Altogether, these experiments indicate that PRRX1 is able to directly bind to the promoter, and activate the transcription of *miR-15f* members.

To elucidate in vivo the relationship between Prrx1 and miR-15f members we studied the expression of the miRNAs by ISH using specific DIG-labeled LNA (locked nucleic acid) probes to detect their mature versions in mouse embryos. At E8.5, expression can be detected in similar territories to those of *Prrx1*, e.g., MNC cells (Fig. 3g), suggesting that Prrx1 induces the expression of these miRNAs in those territories. Altogether these results indicate that miR-15f members are directly induced by Prrx1 in vitro and in vivo.

Prrx1 indirectly attenuates *Snail1* expression through *miR-15*. We validated our microarray data (Fig. 3a) by qPCR, and confirmed that in MDA231-PRRX1cells, *SNAIL1* is significantly downregulated (Supplementary Fig. 4a). To assess whether this was an early response, we examined *SNAIL1* expression in the inducible Prrx1 system. We analyzed MDA231-cPrrx1 cells after 48hs of dox induction and found that *SNAIL1* was downregulated, confirming the microarray data also in a short-term experiment (Fig. 4a).

The observed downregulation of *SNAIL1* after constitutive and short-term induction of PRRX1 expression together with the fact that Prrx1 directly induces the transcription of *miR-15f* members that were predicted to target *Snail1* (Supplementary Fig. 4b), suggested that Prrx1 could be attenuating Snail1 expression through the activation of miR-15f members. Importantly, the miRNA responsive element (MRE) in *Snail1* 3′ untranslated region (UTR) is conserved in all vertebrates (Supplementary Fig. 4c), and the in silico prediction for binding of human, mouse, chicken, and zebrafish miR-15f members to the corresponding *Snail1* 3′ UTRs suggest a strong hybridization not only in the seed region but also in the 3′ end of the miRNAs (Supplementary Fig. 4b, d–f). Thus, we examined whether *Snail1* could be directly targeted by miR-15f. We first overexpressed *Mir-15-P1/2d* in MDA231, and found a significant downregulation of *SNAIL1*. As a control, we overexpressed a seed-mutated version of the miRNAs (*Mir-15-P1/2d*-mut), unable to bind to *SNAIL1* 3′ UTR (Fig. 4b). Conversely, we used specific miRNA inhibitors/LNAs to KD *Mir-15-P1/2d* in BT549, which express high levels of PRRX1 and miR-15, and that resulted in upregulation of *SNAIL1* (Fig. 4c). To confirm that this downregulation was through direct binding of the miRNAs to the 3'UTR, we performed Luciferase assays and observed that overexpression of *Mir-15-P1/*

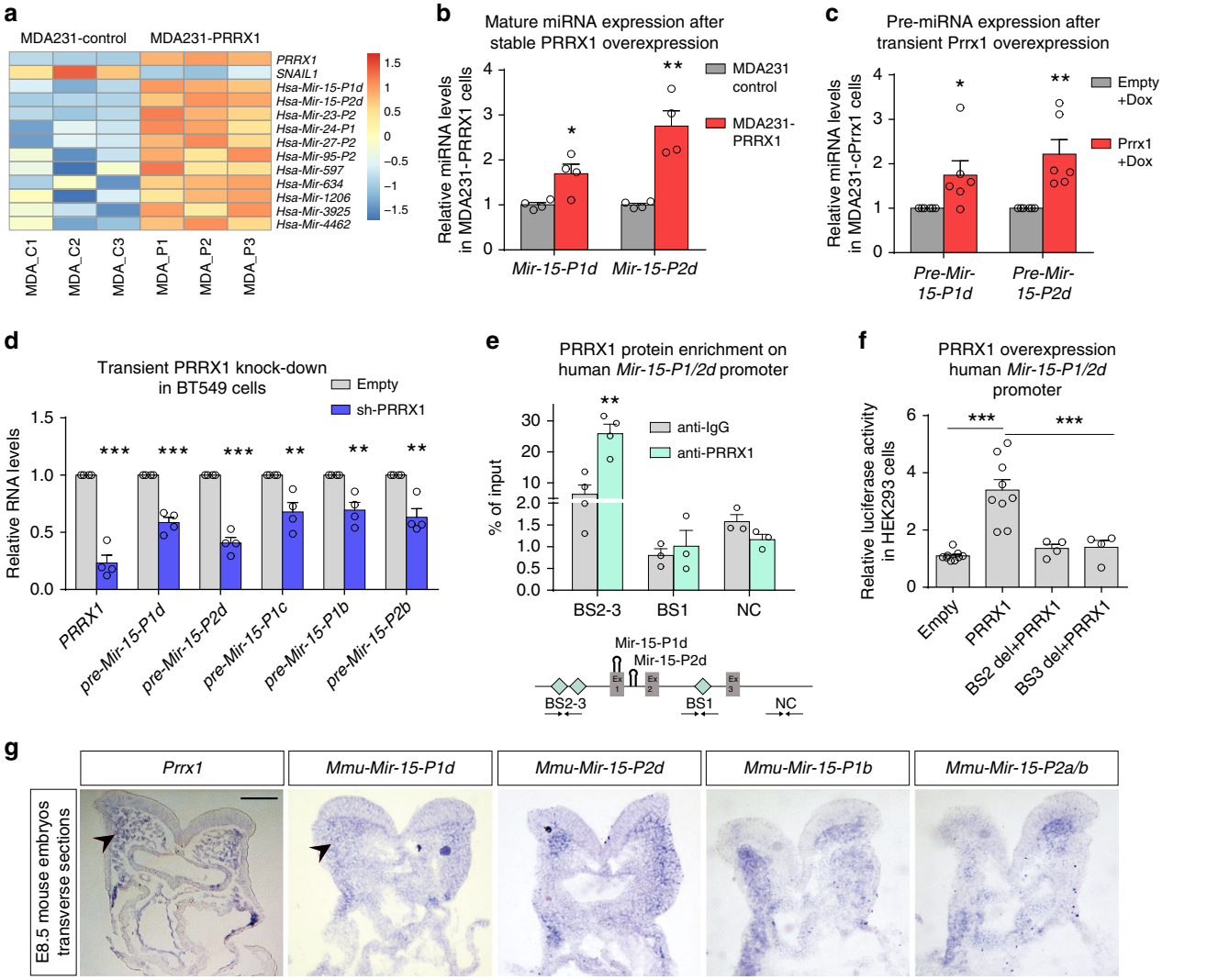


Fig. 3 Prrx1 directly induces the expression of *miR-15* family. **a** Heatmap showing Robust Multi-array Average (RMA) normalized probe intensity values of *PRRX1*, *SNAIL1*, and selected miRNAs in MDA231 cells, with three control samples (MDA_C1 to C3) and three samples in which PRRX1 was overexpressed (MDA_P1 to P3). The intensities mapped as color scale show normalized fold change with respect to the average ($n = 3$). **b** Validation by Taq-Man qPCR of *Mir-15-P1d* and *Mir-15-P2d* upregulation after PRRX1 stable overexpression in MDA231 cells lines, using specific probes to detect the mature miRNAs ($n = 4$). **c** *Mir-15-P1d_pre* and *Mir-15-P2d_pre* upregulation upon conditional (Dox-mediated) Prrx1 overexpression in MDA231 cells ($n = 6$). **d** qPCR assay showing downregulation of *PRRX1* and *premiRNAs* transcription upon transient knockdown (KD) of PRRX1 in BT549 cells. *PRRX1* shRNA plus YFP transfected cells were sorted after 4 days ($n = 4$). **e** PRRX1 directly binds to the human *Mir-15-P1/2d* promoter as shown by ChIP assay in BT549 cells using a PRRX1 specific antibody. Cyan diamonds in the schematic map represent distances between PRRX1 potential BSs and the promoter potential motifs, TAATKDS, on *Mir-15-P1/2d* promoter. (BS1: −293, BS2: −949, BS3: −1451 and NC: +1893). Ex1-3 represent exons of the host long noncoding RNA *MIR503HG*. Arrows represent primer sets used for ChIP detection. ($n = 3$ except for BS2-3 for which $n = 4$). **f** Activation of human *Mir-15-P1/2d* promoter by PRRX1 overexpression shown by dual luciferase assay in HEK293 cells. This activation is abolished upon deletion of the PRRX1 binding sites in *Mir-15-P1/2d* promoter (del1/2 + PRRX1) ($n = 4$ except for deletions for which $n = 4$). **g** Transverse sections of the cranial region of E8.5 embryos showing the expression of *Prrx1*, *Mir-15-P1d*, *Mir-15-P2d*, *Mir-15-P1b*, *and Mir-15-P2a/b* in similar regions (arrowheads). Scale bar: 100 µm. Dox doxycycline, sh short hairpin RNA (shRNA), BS binding site, del deletion. Bars represent mean plus SEM, indicated (*n*) represent number of independent experiments as biological replicates and asterisks indicate significant *p* value in *t* test for (**b**–**e**) and one-way ANOVA with Bonferroni's multiple comparison test for (**f**). (*$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$). Source data are provided as a Source Data file

2d repressed human *SNAIL1* 3′ UTR activity. This repression was abolished when we used either the seed-mutant miRNAs or a MRE-mutated version of the *SNAIL1* 3′ UTR (Fig. 4d). Moreover, we checked other miR-15f members and found that *Mir-15-P1/2b* behaves similarly (Supplementary Fig. 4g). This regulation is conserved in mouse, chicken and zebrafish, as overexpression of *miR-15f* members can repress *Snail1* 3′ UTR from those species (Supplementary Fig. 4h–j), indicating that *Snail1* mRNA is a bona fide target of miR-15f in vertebrates.

To test whether this regulation operates in vivo, we knocked-down *miR-15f* in zebrafish embryos by injecting a sponge RNA[35] containing partially complementary sequences to all *miR-15f* members (Supplementary Fig. 4k). Zebrafish embryos were injected at 1–2 cell stage and then collected at 20-somite stage. qPCR from pools of embryos confirmed that *snail1a/b* transcripts were more abundant in those injected with sponge (Fig. 4e). ISH analysis showed that in one third (30/96) of the injected embryos, *snail1a/b* expression was increased (Fig. 4f). Transverse sections

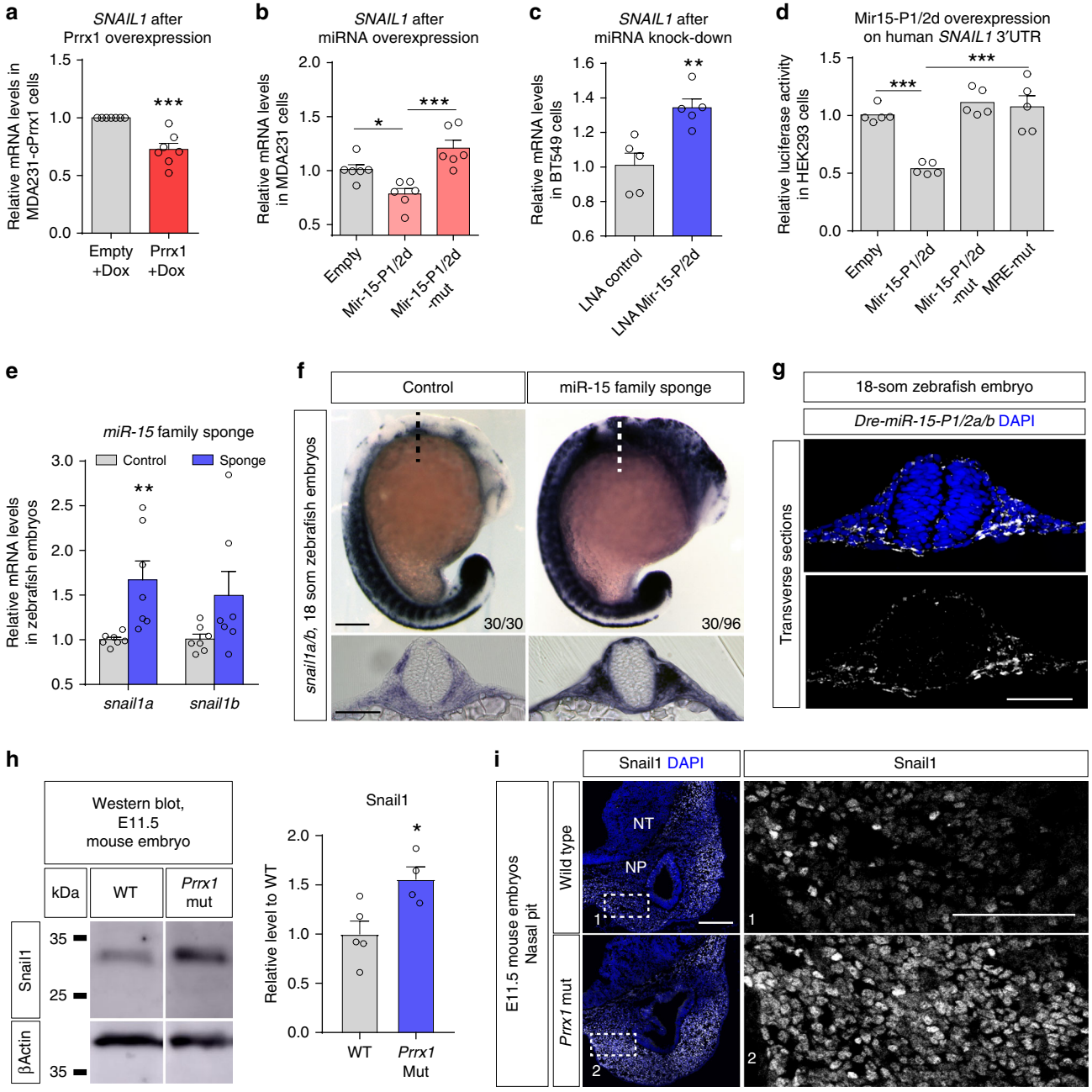


Fig. 4 Prrx1 attenuates *Snail1* expression through the activation of *miR-15* family members. **a** qPCR assay showing the downregulation of *SNAIL1* upon conditional overexpression of Prrx1 in MDA231, induced by doxycycline after 48 h ($n = 7$). **b** qPCR assay showing downregulation of *SNAIL1* upon transient overexpression of *Mir-15-P1/2d* in MDA231 cells, using seed mutated miRNAs as control (*Mir-15-P1/2d-mut*) ($n = 6$). **c** qPCR assay showing upregulation of *SNAIL1* upon transient KD of in *Mir-15-P1/2d* in BT549 cells using miRNA LNA inhibitors ($n = 5$). **d** Repression of human *SNAIL1* 3′ UTR after *Mir-15-P1/2d* overexpression, using seed mutated miRNAs (*Mir-15-P1/2d-mut*) or MRE mutant *SNAIL1* 3-UTR (*MRE-mut*) as control ($n = 5$). **e** qPCR analysis shows upregulation of *snail1a/b* expression in the sponge-injected (*miR-15* family knocked-down) zebrafish embryos compared to controls ($n = 6$). **f** 20-somite control or miR-15 family sponge-injected zebrafish embryos showing combined *snail1a/b* expression by in situ hybridization. Transverse sections taken at the levels indicated by the dashed lines. Scale bar: 250 µm. **g** Fluorescent miRNA in situ hybridization to detect mature miRNAs in transverse sections of 20-somite zebrafish embryo, showing the expression of *Dre-miR-15-P1/2a/b* in neural crest derived cell populations. Scale bar: 250 µm. **h** Western blot for Snail1 and β-actin from E11.5 control and *Prrx1* mutant embryos, including quantification of intensity of Snail1 signal ($n = 5$ WT, 4 KO). **i** Snail1 IF in the nasal pit regions of E11.5 WT and mutant embryos. Scale bar: 250 µm for NP sections and 100 µm for insets (boxes 1 and 2). MRE miRNA responsive element, LNA locked nucleic acid, WT wild type, Mut mutant, NT neural tube, NP nasal pit. Quantifications are performed for one section of WT or mutant embryos, and the increase and expansion is observed in $n = 2/2$ mutant embryos compared to $n = 3/3$ different E11.5 WT. Bars represent mean plus SEM, indicated (n) represent number of independent experiments as biological replicates and asterisks indicate significant p value in t test for a, c, e, h and one-way ANOVA with Bonferroni's multiple comparison test for (**b**), (**d**) ($^*p < 0.05$, $^{**}p < 0.01$ and $^{***}p < 0.001$). Source data are provided as a Source Data file

confirmed the drastic increase in *snail1* expression in the territories of endogenous *miR-15* expression (Fig. 4f, g). Some embryos showed ectopic *snail1* positive cells that migrated beyond their normal position (Supplementary Fig. 4l, asterisk), as may be expected for a Snail gain of function. These findings are compatible with *miR-15f* attenuating *snail1* expression in zebrafish, confirming the results previously observed in cultured cells and living embryos.

To assess whether the loss of Prrx1 protein also affects Snail1 levels in the mouse, we performed western blot analysis in *Prrx1* mutant embryos (bearing a LacZ knock-in in the *Prrx1* locus (*Prrx1*^{tm1Jfm})[36]) (Supplementary Fig. 5a), and found a significant increase of Snail1 protein levels compared with WT embryos (Fig. 4h and Supplementary Fig. 9). Snail1 IF in cryo-sections from E11.5 embryos (Fig. 4i and Supplementary Fig. 5b, d) showed an increased expression and expansion of Snail1 territories in the head of mutant embryos, compatible with a mutual Prrx1/Snail1 repression (Supplementary Fig. 5b–e and Fig. 4i). Altogether, these findings are compatible with our in vitro analyses and indicate that Prrx1 indirectly attenuates the expression of Snail1 through the induction of miR-15f, a mechanism that is conserved in vertebrates.

***SNAIL1* and *PRRX1-miR-15* correlation with patients' prognosis.** As shown previously in breast and in lung squamous cell carcinoma patients, high expression of *SNAIL1* in tumors correlate with poor prognosis[8,11], while high *PRRX1* correlates with good prognosis[8]. We confirmed this notion analyzing breast cancer patients' overall survival (OS) from a different dataset[37] containing different clinical conditions and subtypes, including tumoral grade, ER/PR/HER2, lymph-node status, etc. (Total) (Fig. 5a). These correlations are even stronger when we analyze patients with lymph-node positive status (Fig. 5b). We also found that in terms of patients' OS the expression of *miR-15f* members have similar profile to that of *PRRX1*, i.e., high expression of *miR-15f* members correlate with better OS in total and lymph-node positive patients' groups (Fig. 5a, b).

Breast cancer is highly heterogeneous, including several different subtypes. Among all subtypes, triple-negative breast cancer (TNBC) is classified as basal-like[38] and has been associated with EMT features and mesenchymal phenotype, also manifesting worst prognosis and lowest response to therapy[39,40]. Therefore, we examined the expression of *SNAIL1/PRRX1* and *miR-15f* members in patients with basal/TNBC subtype. Similar to the results obtained with total patients, high expression of *SNAIL1* significantly correlated with poor OS, while *PRRX1* high expression was concomitant with more favorable OS (Supplementary Fig. 6). Also, three members of the *miR-15f*, out of the four that were shown in total patients' group, followed *PRRX1* expression trend (Supplementary Fig. 6). Altogether, these results indicate that similar to *PRRX1*, when *miR-15* are expressed at high levels in cancer patients, they are associated with better prognosis, even in patients who have the cancer spread to their lymph nodes and also in basal/TNBC subtype which comprises the most unfavorable disease.

Snail1 and Prrx1 are sequentially expressed during EMT. Snail1 seems to be the first EMT-TF to be expressed in regions of the embryo that will undergo an EMT process, including the primitive streak and the neural crest[41]. *Snail* genes (*Snail1* in the mouse and *Snail2* in the chick) are expressed in the PNC before delamination from the neural tube and continue to be expressed in early migratory cells but their expression decays along the migratory routes[41,42]. In contrast, Prrx1 is expressed in migratory crest subpopulations (MNC; ref. [8] and this work). The expression of these two EMT-TFs follows a temporal order in the migration of the neural crest, and thus, likely reflects a hierarchy of activation during the EMT process that drives neural crest delamination and migration. To test this idea, and taking into account that both Snail and Prrx are activated by the TGF-β superfamily members bone morphogenetic proteins (BMPs) during embryonic development[8,12], we first treated developing zebrafish embryos from 4-somite stage with BMP and examined the levels of *snail1, prrx1* and *miR-15f* members by qPCR. *snail1b* was quickly upregulated, while *prrx1a/b* and *miRNAs* were induced later, concomitant with downregulation of *snail* genes (Fig. 6a).

To test this hypothesis in other species, we implanted BMP-coated beads close to somites in developing chicken embryos, and monitored the expression of these two EMT-TFs at different time points. *SNAIL1* was quickly upregulated 1 h after bead implantation, whereas *PRRX1* upregulation was observed after 5 h (Fig. 6b), supporting the notion of sequential activation of these EMT-TFs. In line with our findings in zebrafish, *SNAIL1* transcripts were reduced in the somites close to the bead after longer BMP treatment, while *PRRX1* was still upregulated in similar somite regions (Fig. 6b). We also found that the peak of SNAIL1 protein expression occurred after 5 h, and that of PRRX1 protein after 10 h (Supplementary Fig. 7a, b). This suggests that although *SNAIL1* transcription is quickly upregulated, the increase in protein level takes longer, allowing a slower *PRRX1* activation by BMP, that will gradually outcompete SNAIL1 expression.

To examine whether this sequential activation was also conserved in the mouse, we used a transgenic model containing a *downstream enhancer* of *Snail1* (DES) that drives the expression of enhanced green fluorescent protein (EGFP) to its endogenous territories, recapitulating the vast majority of *Snail1* expression in the developing embryo (*Snail1* (DES)-EGFP reporter mouse; Supplementary Fig. 8). The half-life of the GFP protein is very long[43] whereas Snail1 protein is very unstable[44], thus, this model can provide a short-time lineage tracing system. In transverse sections of E8.5 embryos, we observed that PNC and early MNC cells expressed Snail1 and GFP, while late MNC cells are negative for Snail1 protein but positive for both GFP and Prrx1 (Fig. 6c). This observation is compatible with the idea that Snail1 and Prrx1 are expressed in a sequential manner, as Prrx1 positive cells seem to be descendants of Snail positive cells (double-labeled cells in Figs. 6c box4). Altogether, these results suggest that during embryonic development, cells undergo EMT by activating Snail1 first, and then attenuation of Snail1 is indirectly mediated by Prrx1 though the activation of the miR-15f, all being part of a GRN in which miR-15f may coordinate the transition from Snail1- to Prrx1-mediated EMT programs (Fig. 6d).

Discussion

Epithelial cells transition to a variety of mesenchymal states to fulfill different roles in different contexts. The existence of diverse mesenchymal cells allows for a high degree of cell heterogeneity in terms of potential fates and therefore, functions. They are all called mesenchymal based on morphology, the expression of particular markers (although there is no universal marker to define them), and very often, their ability to migrate. Differences arise from the tissue of origin as these phenotypic transitions are governed by extracellular signals that activate a plethora of EMT-TFs.

Here, we show that Snail1 and Prrx1 are expressed in a complementary manner during vertebrate development and in cancer patients. Considering that all EMT-TFs are activated by the same extracellular signals, this strongly suggests that Snail1 and Prrx1 may be part of a mutual regulatory network. In this work, we describe such a mechanism. We find that Snail1 and Prrx1 behave

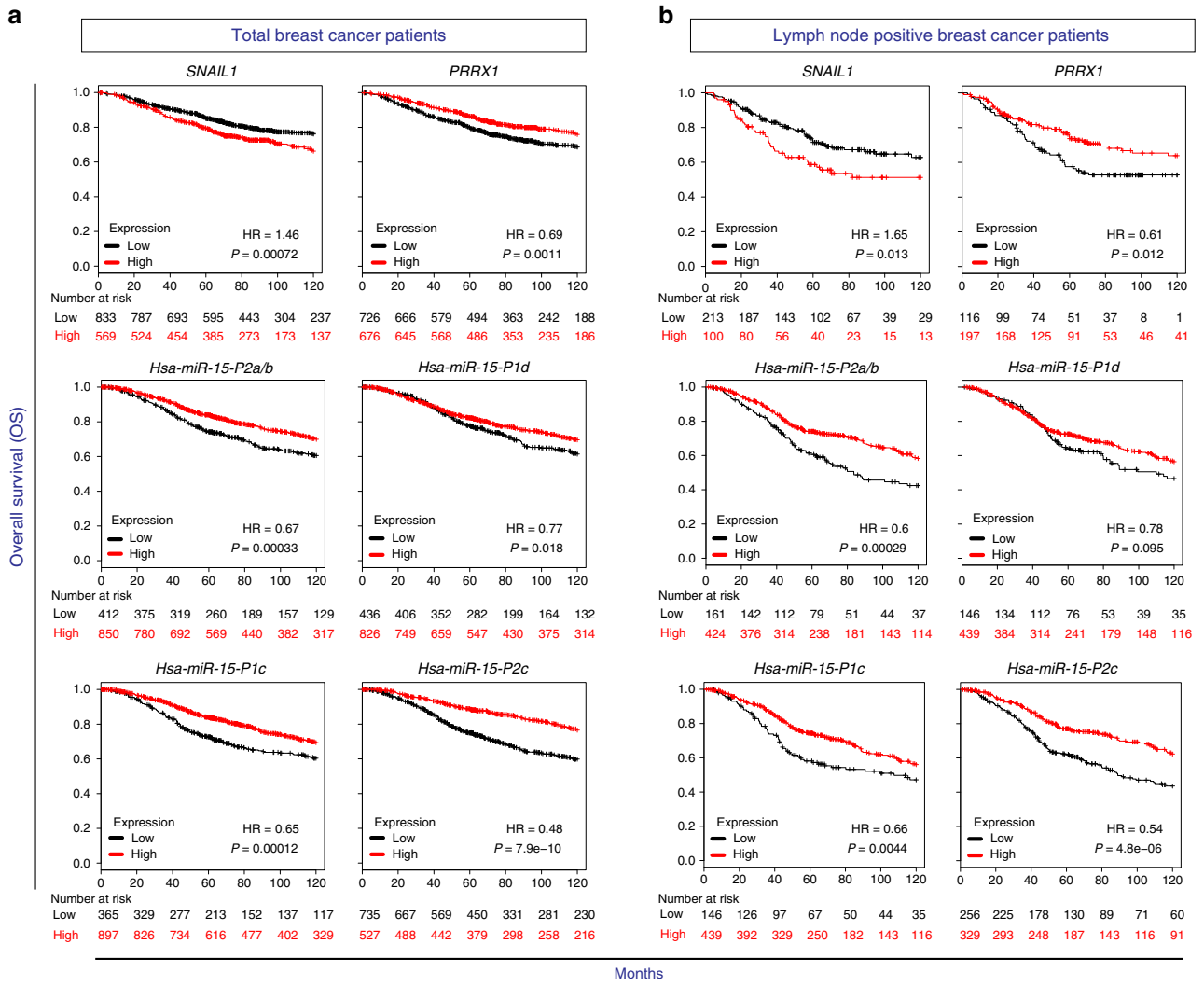


Fig. 5 Relationship between the expression of *SNAIL1*, *PRRX1*, and *miR-15* family members and overall survival in breast cancer patients. Kaplan–Meier overall survival (OS) plots from breast cancer patients with lymph-node positive status, showing that high expression of *SNAIL1* correlates with low survival, while *PRRX1* high expression correlates with a better survival. The expression of *miR-15* family members follow a similar trend as that of *PRRX1*. Hazard ratio (HR) and logarithmic ranked *p* Value (longrank *P*) were analyzed to infer the significance of the differences. Numbers below each graph represent number of patients at risk in any given time (months), black for low expression and red for high expression of each gene/miRNA. The cut-off is automatically calculated based on the best performing threshold

as mutual repressors, albeit they use different mechanisms. Although in some contexts Snail1 can function as an activator[45], it is mostly a strong transcriptional repressor[26,46], and this is the mechanism that it uses to repress *Prrx1* expression. Prrx1, in turn, is a transcriptional activator[30] and accordingly, here we show that it represses Snail1 indirectly through the activation of a repressor, miR-15f. As miR-15f is composed of eight members in four different clusters, this provides a certain degree of variability in the repression, e.g., one or several family members can be used in specific cell contexts to exert a particular level of repression.

Following the developmental time in a particular tissue, it is possible to infer potential regulatory scenarios in which the mutual repression of these TFs could be integrated. For instance, in the mouse, Snail1 is expressed in the PNC and Prrx1 is expressed in subpopulations of MNC. During somite formation, Snail1 is expressed in the precursors of the whole somite and, later on in the mature somites, it is expressed in the SC and Prrx1 in the DM, also showing complementary expression. This implies that Prrx1 is activated in some descendants of Snail1 expressing cells. Altogether, this points to the successive upregulation of EMT-TFs during the development of the neural crest and the somites, providing a hierarchical temporal order in the activation of these two TFs. As such, we provide evidence of this temporal hierarchy both in vitro and in vivo, with Snail1 being activated at earlier time points in response to an EMT-inducing signal.

Interestingly, the inducing signal is the same for Snail1 and Prrx1 (the TGF-β superfamily[8,12] and this work), and thus, the question is why they are not simultaneously activated when the signal is available in the embryo or in cancer cells. One explanation is that Snail1 bears a poised promoter[47], explaining why it usually is an early response gene activated shortly after TGF-β administration[48]. Snail1 fast activation may set an inhibitory scenario for Prrx1 upregulation, which may change when the duration of the TGF-β signal accumulates sufficient stimulus to induce Prrx1. This establishes a temporal order of predominant EMT programs, firstly mediated by Snail1, a strong epithelial repressor needed for cells to detach from their neighbors. Subsequently, Prrx1, a strong mesenchymal inducer, is activated and in turn, it will promote the acquisition and maintenance of robust mesenchymal features for cells to migrate to their destination.

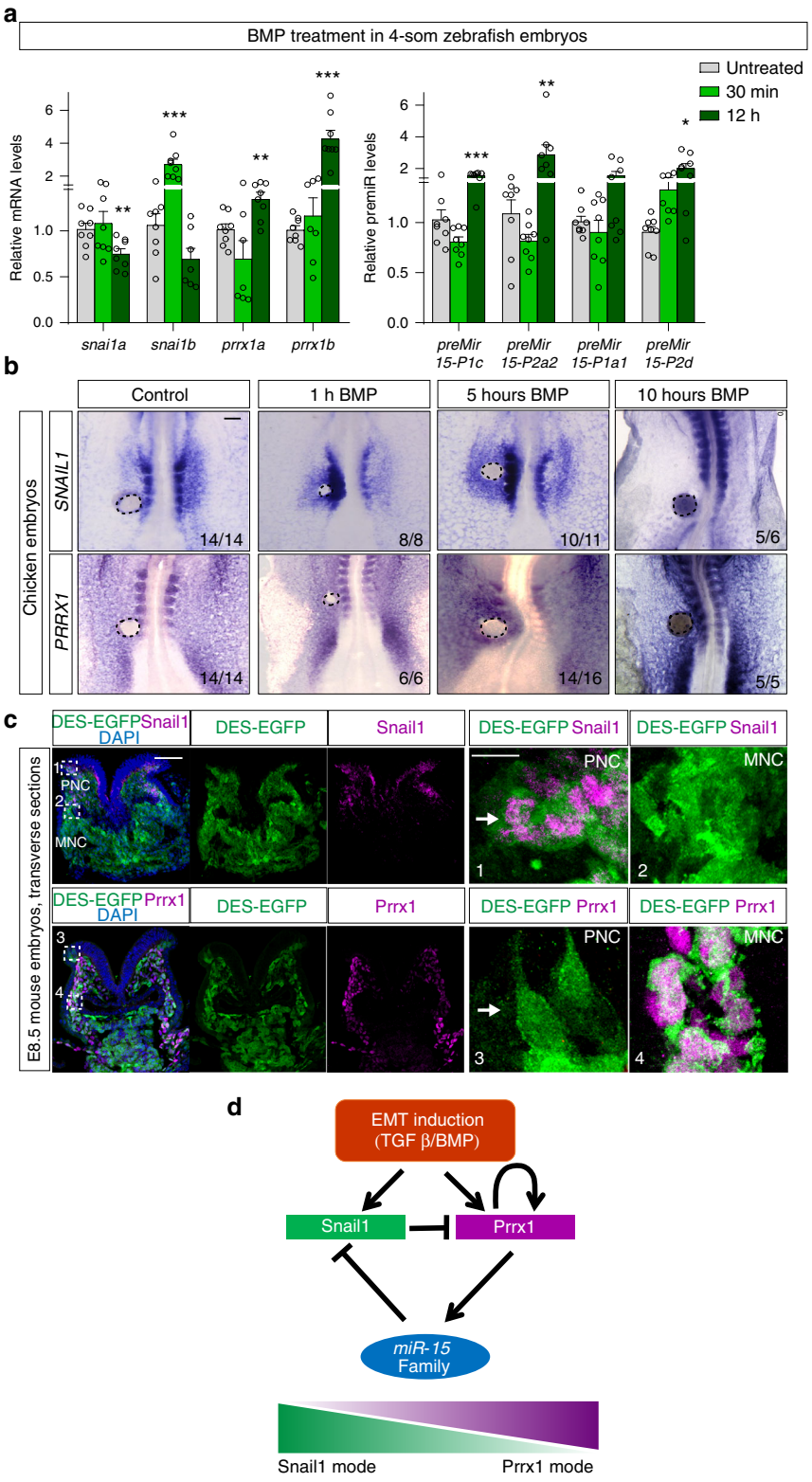


Fig. 6 Snail1 and Prrx1 are sequentially expressed during EMT. **a** qPCR assay showing the expression of *prrx1a/b* and *snail1a/b* (left) and *premiRNAs* (right) in zebrafish embryos after different treatments with BMP ($n = 8$). Asterisks indicate significant p value in one-way ANOVA with Bonferroni's multiple comparison test compared to the control. (*$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$). Source data are provided as a Source Data file. **b** Dorsal view of HH9 chicken embryos showing *SNAIL1* and *PRRX1* expression in control embryos and at different times after BMP-soaked bead implantation. **c** Transverse sections of E8.5 DES-EGFP reporter mouse embryos showing IFs for GFP and Snail1 (upper panel) or Prrx1 (lower panel). Insets show higher magnification pictures for GFP and Snail1 (1 and 2), or for GFP and Prrx1 (3 and 4). **d** Schematic model of the gene regulatory network. DES downstream enhancer of Snail1, PNC premigratory neural crest shown by arrows, MNC migratory neural crest. Scale bars: 100 μm. Scale bars: 100 μm for sections and 10 μm for insets (boxes 3–4)

Cells do not only express these two TFs but also members of the other EMT-TF families in a context-dependent combination that will determine their final phenotype. Nonetheless, Snail1 and Prrx1, as shown here, are mutually exclusive, thereby promoting their own mode of EMT. We also describe in this work that Prrx1 further reinforces its own program by activating its own promoter. Interestingly, Snail1 represses its promoter, thereby modulating its own expression[49]. This can also contribute to the establishment of the temporal hierarchy of EMT-TF activation, promoting Prrx1 upregulation, as we observe during embryonic development. Altogether, these mutual regulations and the response to the inducing signal establish a GRN that controls the EMT programs.

Snail1 and Prrx1 EMT programs differ in some important aspects. As such, EMT has been associated with the acquisition of stem cell properties[6] but Prrx1 represses stemness[8,9]. This impinges into the plasticity required in both embryonic and cancer cells to revert to a more epithelial phenotype at the site of their final destination either to differentiate into different organs during development or to form macrometastases during cancer progression (see refs. 8 and 50–52). Prrx1 needs to be downregulated in cancer cells to promote the reversion, and its downregulation implies an increase in stem cell properties that are also essential for metastatic growth. This explains why there is a positive correlation between high levels of Prrx1 and better prognosis in cancer patients[8]. Those cells may be stuck at a very mesenchymal state, lacking the plasticity required to colonize and restart proliferation in distant organs. In addition, as we have found that miR-15f members are direct targets of Prrx1, Prrx1 high levels will lead to miR-15 activation, resulting in Snail1 attenuation, promoting a decrease in stem cell properties and therefore in metastatic abilities. In agreement with that, we find that miR-15 expression also correlates with better survival rate in patients. However, we cannot exclude that other regulatory circuits impinging on Snail1 expression also play a role, including the well described Snail1/miR-34 and miR-200 reciprocal feedback loops[14].

Importantly, the relationship between EMT and stemness has been mainly studied in adult mammary gland and in cancer cells[6], but these differences in conferring stem-like cell properties may also have a role during embryonic development. Snail1 expressing cells might be in more stem state, as they can give rise to cells with a variety of fates. For instance, Snail1 is expressed early in primitive streak and the premigratory neural crest, both of which later give rise to a large variety of cell types[53]. As such, it has been proposed that the premigratory neural crest may still retain characteristics of the pluripotent blastula stage[54]. These multipotent territories do not express Prrx1, which is rather expressed in cells that have been already committed to particular lineages such as specific mesodermal or crest subpopulations. This is also compatible with the already discussed temporal hierarchy in the activation of these two EMT-TFs that occurs in parallel with commitment to different fates.

As much as Snail1 and Prrx1 trigger different EMT programs[8], they have important common functional roles inducing the transition towards a mesenchymal state, activating invasion and migration and attenuating proliferation. Snail1 induces migration and invasion in many different contexts through activation of a variety of factors including matrix metalloproteinases that enable cells to break the basal lamina and delaminate[55,56], and attenuates proliferation by repressing the *cyclin D* transcription[57]. Prrx1 also decreases cell proliferation and induces invasiveness[8,58]. Here, we show that some roles of Prrx1 are mediated by miR-15, such as promoting its own EMT mode by attenuating Snail1 expression. Interestingly, some miR-15f members have previously been associated with regulating cell proliferation, migration and invasion[32–34], suggesting that in addition to Snail1, they have other targets that impinge into particular characteristics of the EMT process.

Overall, in this work, we describe a GRN that operates in development and in disease to regulate the activation of different EMT programs, providing a superimposed degree of controlled heterogeneity required for the EMT, a crucial and complex process, fundamental for organ formation and cancer progression.

Methods

Cell culture. MDA-MB-231 (MDA231), MDA436, BT549, and SUM149PT human breast tumor cell lines, HEK293 human embryonic kidney and NIH3T3 mouse fibroblast cells were purchased from the ATCC (Virginia, USA) and ASTERAND BIOSCIENCE. MDA231, MDA436, and BT549 human tumor cell lines were cultured in DMEM:F12 HAM media (1:1), and SUM149PT cells in F12 HAM media, supplemented with 10% heat inactivated fetal bovine serum (FBS) (Sigma), 10 µg/ml insulin (Roche), 1% Gentamicin (Sigma), and 1% amphotericin (Sigma). HEK293 cells were cultured in DMEM supplemented with 10% heat inactivated FBS (Sigma) and 1% Gentamicin (Sigma) 1% amphotericin (Sigma). NIH3T3 cells were cultured in DMEM supplemented with 10% heat inactivated calf new-born serum (Sigma), 1% amphotericin (Sigma) and 1% Gentamicin (Sigma). Cells were kept at 37 °C and 5% CO_2, and the media was replaced every 2/3 days. HEK293 cells were passaged when they reached 80% confluency 1:10 every 48 h, while BT549 and MDA231, MDA436, and SUM149PT cells were passaged when reached 80% confluency 1:5 every 72 h. Cells were discarded after up to five consecutive passages and replaced by freshly thawed stocks. All cell lines were tested and confirmed negative for mycoplasma on a monthly bases at the host institution. All the cell lines were authenticated using STR profile by the Genetic Analysis Service at Miguel Hernandez University, Spain.

Microarray analyses. MDA231 human breast cancer cell line was transduced using a lentiviral system to stably express control GFP (MDA231-Control, MDA231-C) and PRRX1 (MDA231-PRRX1, MDA231-P)[8]. Three independent samples per condition from MDA231-C and MDA231-P cells, were hybridized to Human Gene 2.0 ST expression arrays (Affymetrix). Microarray data were analyzed using R and Bioconductor[59]. Expression data was preprocessed and normalized using the RMA algorithm[60], and differential expression analysis was performed using limma[61] *t* test to extract statistically significant changes between *PRRX1* and control samples. Accordingly, the genes that have a log2 fold change > 1.5 and corrected p value < 0.05 (Benjamini–Hochberg FDR) were considered as being differentially expressed.

Plasmid constructs and interfering RNAs. For overexpression studies, human PRRX1 ORF (ENSG00000116132) was cloned in the pBABE expression vector, and mouse Prrx1 ORF (ENSMUSG00000026586) was cloned in the pCDNA3.1 expression vector. Lentiviral vectors were generated via subcloning nuclear yellow fluorescent protein (nYFP) followed by P2A and mouse Prrx1-L open reading frame into pTRIPZ construct. P2A is a short peptide derived from porcine teschovirus-1 which mediates production of equimolar levels of Prrx1 protein sequence plus the nYFP reporter[62]. For RNA interference in BT549 cells, nYFP under hPGK promoter was subcloned into pLKO.1 construct containing PRRX1 specific shRNA (Dharmacon, RHS3979–201751764, Mature antisense: TTATTGGCTAGCATGGCTCTC)[8]. For RNA interference in MDA436 cells, specific siRNAs were used for SNAIL1 and PRRX1 (SNAIL1 siRNA (antisense): UGGCACUGGUACUUCUUGACAUCUGtt and PRRX1 siRNA (antisense): GAUGACUAAAUGGUAUUCCtt). PRRX1 siRNA was obtained from Silencer® predesigned (Ambion). For miRNA overexpression studies, approximately 150 base pairs up- and downstream of the stem–loop sequences of each miRNA pair (obtained from www.mirbase.org) were cloned into the pcDNA3.1 vector. 3' UTRs of mouse, human, chicken and zebrafish *Snail1* were cloned downstream of Firefly luciferase ORF in pGL3 basic vector. Promoters of human and mouse *Mir-15-P1/2d* as well as human and mouse *Prrx1* plus human *PRRX1* upstream (enhancer-like; containing BS7) region (−7046 to −6439) were cloned upstream of the Firefly luciferase ORF in pGL3 basic vector, which contains a minimal promoter, considering cases to analyze DNA segments lacking promoter activity. For Snail1 reporter assay, mouse *Snail1* promoter cloned in pXP1 luciferase construct[63] was used. For mutagenesis, primers were designed using the online tool https://www.agilent.com/store/primerDesignProgram.jsp. PCR was performed using Pwo polymerase following manufacturer's instructions, and PCR products were treated with DpnI restriction enzyme, then transformed in bacteria to be amplified, extracted using Mini-prep and then sequenced to confirm the mutation/deletions. All constructs were generated using the primers described in Supplementary Table 3.

Transfection of plasmids and interfering RNAs. BT549 cells were transfected with either control pLKO.1 vector or PRRX1 specific shRNA[8] using Lipofectamin 3000 (Invitrogen) following the manufacturer's instructions. 500,000 cells were

seeded in 10 cm plates 24 h prior transfection. Five days post transfection YFP-positive cells were fluorescence-activated cell sorting (FACS) (BD FACSARIA III, USA) sorted directly collected into lysis solution from *mir*Vana™ miRNA Isolation Kit and subjected to RNA extraction. For LNA transfection, 250,000 BT549 cells were seeded in 6-well plates 24 h prior transfection, and 48 h after transfection cells were lysed for RNA extraction. HEK293 cells were transfected using the same procedure. MDA231 cells were transfected with pCDNA3.1 vectors for miRNA overexpression using Lipofectamin 3000 (Invitrogen) following the manufacturer's instructions. 200,000 cells were seeded in 6-well plates 24 h prior transfection, and 48 h after transfection cells were lysed for RNA extraction.

Dual luciferase reporter assay. HEK-293 and BT549 cells were cultured in 24-well plate 10 h before transfection. Cells were seeded in triplicate and transfected with 100 ng of the constructs containing sequences for promoters or 3'UTRs cloned in pGL3P vector, 50 ng of the pRL control vector (Renilla luciferase) and 200 ng of the pcDNA3.1 plasmid subcloned with different transgenes for pre-miRNA or Prrx1/Snail1 overexpression, using the Lipofectamin 2000 (Invitrogen). The media were refreshed 12 h after transfection. After 48 h, the cells were lysed using 5× lysis buffer (Promega), and the enzymatic activity was analyzed after adding Luciferase Assay and Stop&Glo reagents (Promega), to measure Firefly and Renilla successively, using Sirius Luminometer (Berthold, Germany).

Lentiviral infection. MDA231 cells were infected by pTRIPZ inducible lentiviral system containing nuclear nYFP-P2A-Prrx1 or control vectors with turbo RFP. The pool of infected cells was selected by treatment with puromycine for 48 h. Infected cells were seeded in 10 cm dishes, subjected to 2µg/ml doxycycline to induce the Tet-ON promoter (tetracycline-controlled transcriptional activation that is induced upon presence of antibiotic tetracycline or one of its derivatives, doxycycline) and FACS-sorted after 48 h for either YFP or RFP positive signals for Prrx1 overexpressing or control cells, respectively. Cells were collected immediately after sorting in lysis solution from *mir*Vana™ miRNA Isolation Kit (Invitrogen).

Total RNA extraction cDNA synthesis and qPCR analysis. For gene expression assays, total RNA was extracted using the illustra RNAspin Mini isolation kit (GE Healthcare), following the manufacturers' instructions. For Reverse transcription cDNA synthesis, oligo $(dT)_{18}$ and random hexamer primers with the Maxima First Strand cDNA Synthesis kit (Thermo Scientific) were used, following the manufacturers' instructions. Primers used are listed in Supplementary Table 3. Quantitative RT-PCR was performed using Fast SYBR Green Mastermix (Applied Biosystems), in a Step One Plus machine (Applied Biosystems) according to the manufacturers' instructions. Relative levels of expression were calculated using the comparative Ct method normalized to the internal control *TBP* housekeeping gene.

For mature miRNA expression assays, total RNA enriched for small RNAs was extracted using *mir*Vana™ miRNA Isolation Kit (Invitrogen), following the manufacturers' instructions. Reverse transcription reactions were performed using TaqMan® MicroRNA Assay (Applied Biosystem) according to the manufacturers' instructions. For quantitative real-time retro-transcriptase PCR (qPCR), specific probes from TaqMan® MicroRNA Assay (Applied Biosystem) were used (Catalog # 4427975, hsa-miR-424-5p Assay ID: 000604 and hsa-miR-503–5p Assay ID: 001048). TaqMan® Universal Master Mix II, no UNG (Applied Biosystems) was used, in a Step One Plus machine (Applied Biosystems) according to the manufacturers' instructions. Relative levels of expression were calculated using the comparative Ct method normalized to the internal control U6 snRNA.

ChIP assay. Parental BT549 cells, BT549 cells transfected with SNAIL1-MYC and NIH-3T3 cells were used for ChIP assays for each indicated experiment. Cells were collected when at 80% confluency from 10 cm culture dish after trypsinization, and fixed in 1% PFA 15 min, and subsequently quenched with 0.125 M glycine for 15 min. Next, cells were lysed in 300 µl of lysis buffer (1% sodium dodecyl sulfate (SDS), 10 mM EDTA 50 mM Tris-HCl pH 8, protease inhibitor cocktail) on ice for 10 min. Lysates were sonicated four times for 15 min each in Bioruptor (H, 30" on/30" off). Sonicated lysates were then diluted to 5 ml with dilution solution (0.01% SDS, 1% Tx100, 2 mM EDTA, 20 mM Tris-HCl pH 8, 150 mM NaCl, protease inhibitor cocktail). Aliquots were de-crosslinked and ran in agarose gel to check for proper fragmentation. Samples were used for antibody incubation overnight at 4 °C in a rotating rotor. IgG was used as control for normalization and Histone H3 was used, at least in one of the replicates, to ensure efficient precipitation. Information and dilution for antibodies are listed in Table 1. Protein A coupled magnetic beads (Biorad) were blocked overnight with 0.05% bovine serum albumin (BSA), 2 µg/ml of salmon sperm DNA in dilution solution. Antibody and control fractions were then incubated with the magnetic beads for 4 h at 4 °C in a rotating rotor, and subsequently washed with Wash buffers (WB) 1 (0.1% SDS, 1% Tx100, 2 mM Tris-HCl pH 8, 150 mM NaCl), WB 2 (0.1% SDS, 1% Tx100, 2 mM Tris-HCl pH 8, 500 mM NaCl), WB 3 (1% NP40, 1% sodium deoxycholate, 10 mM Tris-HCl pH 8, 1 mM EDTA, 0.25 mM LiCl), and finally with WB4 (10 mM Tris-HCl pH8 and 1 mM EDTA). 10% CHELEX was added to the samples and then de-crosslinked at 95 °C for 10 min, Proteinase K treated (2 µg/ml) at 55 °C for 30 min and inactivated at 95 °C for 10 min. Supernatant was used for direct qPCR reaction (2.5 µl).

RNA in situ hybridization. Whole-mount RNA ISH[64] was carried out using digoxigenin-labeled probes synthesized from mouse, chicken and zebrafish Snail1 and Prrx1[12,41]. Mouse, chicken and zebrafish embryos were fixed in 4% PFA-DEPC O/N. Zebrafish embryos were then dechorionated in cold 4% PFA-DEPC. Embryos were washed in PBS 0.1% Tween 20 (PBS-T). Next, embryos were dehydrated through a series of increasing methanol concentrations in PBS-T (25, 50, 75, and 100%) and kept O/N at −20 °C. Then, they were rehydrated through methanol:PBS-T in reverse order and washed in PBS-T at the end. For fluorescent ISH, embryos were incubated in 1% hydrogen peroxide (H_2O_2) for 10 min, and then washed with PBS-T. Depending on species and developmental stage, embryos were treated with 10 µg/ml proteinase K in PBS-T between 3 and 6 min at room temperature. Then they were refixed with 4% PFA-DEPC, and washed. Embryos were then incubated with pre-hybridization solution (50% formamide, 5× SSC, 2% Boehringer blocking powder, 0.1% Tween 20, 50 µg/ml heparin, 1 mg/ml t-RNA, 1 mM EDTA, 0.1% CHAPS) at 60 °C, and then O/N upon refreshing prehybridization solution. Embryos were either used the next day for ISH or stored at −20° C.

Prehybridized embryos were incubated with 1µg/ml of DIG or FLUO probes O/N at 60° C. The next day, embryos were washed several times first with 2X SSC, 0.1% CHAPS and then 0.2X SSC, 0.1% CHAPS, and then with KTBT washing buffer (50 mM Tris-HCl pH7.5, 150 mM NaCl, 10 mM KCl, 0.1% Triton X-100 in H2O). After the washes, embryos were incubated in blocking solution (15% sheep serum, 0.7% Boehringer blocking solution, 0.1% Triton X-100 in KTBT) for 3 h at 4 °C. Then embryos were incubated with antibodies (1/1000 anti-DIG-AP, 1/500 anti-DIG-POD or anti-FLUO-POD, depending on the experiment) in blocking solution O/N at 4 °C. The whole next day, embryos were washed many times in KTBT buffer and kept O/N in KTBT at 4 °C.

For bright field ISH (chemical development of signal) embryos were washed in NTMT (100 mM Tris-HCl pH9.5, 59 mM $MgCl_2$, 100 mM NaCl, 0.1% Tween-20, 1 mM levamisole in H_2O) buffer three times prior to developing the signal. Embryos were then incubated with NTMT containing freshly added 3 µl NBT and 2.6 µl BCIP per 1 ml (developing solution), in the dark at RT until the color reaction develops. After obtaining the desired signal level in positive tissues, embryos were washed several times in KTBT and left O/N at 4 °C. After hybridization embryos were fixed in 4% PFA, washed in PBS and imaged in a Leica M125 dissecting scope with a Leica DFC 7000T digital camera (Leica, Wetzlar, Germany). Some embryos were embedded in paraffin or gelatin, and sections were obtained at 7 or 30 µm, respectively. Sections were photographed under a Leica DMR microscope (Leica, Wetzlar, Germany).

For double-fluorescent ISH, DIG, and FLUO labeled probes were mixed and added to the hybridization solution. After the washing steps with SSC buffer and blocking, embryos were incubated with the first antibody (anti-DIG-POD or anti-FLUO-POD). The next day, after two times washing with KTBT, embryos were incubated with Amplification solution (TSA® fluorescein detection kit, PerkinElmer) for 2 min to adjust the pH. Cy3 or FITC were added to the Amplification solution for developing red or green colors, respectively. The embryos were then incubated in the mix for 45 min in the dark at room temperature. Next, the samples were washed 5 times in KTBT, and incubated in 2% H_2O_2 for 2 h, then washed 5 times in KTBT. Embryos were then incubated in blocking solution for 3 h at 4 °C, prior to adding the second antibody (anti-DIG-POD or anti-FLUO-POD) in which they were incubated O/N at 4 °C. Information and dilution for antibodies are listed in Table 1. The next day, fluorescent developing procedures were repeated for the other color, plus stained with DAPI. After washing in KTBT, embryos were imaged with Olympus FV1200 confocal microscope and subjected to mosaic merge using Image J software, for pictures of whole-mounted embryos. Some embryos were embedded in 4% low-melting agarose and sectioned using a Leica VT1000S vibratome at 200 µm and subjected to confocal microscope imaging.

miRNA in situ hybridization. For mature miRNA ISH the mouse miR-322 (Mir-15-P1d), miR-503 (Mir-15-P2d), miR-15b (Mir-15-P1b), miR-16 (Mir-15-P2a/b) and scramble LNA probes, which were 5′–3′ DIG labeled, were purchased from Exiqon. ISH was performed on 10 µm frozen embryo sections, using the EDC method[65]. Dried cryo-sections were treated with 20 mg/ml proteinase K (pH7.4) in Tris-buffered saline (TBS) 1× (for 10× TBS, 69.6 g Tris, 87.6 g NaCl + 800 ml water, adjust pH to 7.6 and add water up to 1 L) for 20 min at RT, then washed 2 times in TBS and fixed in 4% PFA for 10 min at room temperature. Then, sections were washed once with TBS 0.2% glycine for 5 min, and washed twice with TBS.

Samples were then incubated twice for 10 min in Imidazole 0.13 M buffer (for 160 ml of buffer, 1.6 ml of 1-Methylimidazole was added to 130mlof water, pH was adjusted to 8 with HCl, and then 16 ml of NaCl 3 M was added and then water to final volume). Sections were fixed in 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) solution (176 µl of EDC (Sigma) was added to 10 ml of Imidazole buffer, then pH was readjusted to 8 by adding HCl) for 3 h at room temperature. Samples were washed once with TBS 0.2% glycine, and twice with TBS. Next, freshly prepared 0.1 M Triethanolamine (TEA), 05% acetic anhydride was added for 30 min at room temperature. After washing twice with TBS, the sections were prehybridized with hybrid-mix solution (50% formamide, 5× SSC, 5× Denhardt's solution (Applichem), 250 µg/ml yeast tRNA (Sigma), 500 µg/ml salmon sperm DNA (Sigma), 2% (w/v) Blocking Reagent (Roche), 0.1% 3-

Table 1 Antibodies

Antibody	Concentration	Species and type	Provider (Cat #)
Prrx1 (for ChIP)	1:200	Rabbit polyclonal	Sigma (HPA051084)
IgG	1:1000	Rabbit	Diagenode (C15410206)
Myc tag	1:500	Goat pAb	Abcam (ab9132)
Prrx1 (for IF)	1:100	Tanaka lab[12]	
Snail1 (for IF in embryos and western blot)	1:50	Rabbit monoclonal	Cell Signaling (C15D3, #3879)
GFP	1:500	Chicken polyclonal	Aveslab (2BScientific, GFP-1020)
DIG-AP	1:1000	Fab fragments, sheep polyclonal	Roche (11093274910)
DIG-POD	1:500	Fab fragments, sheep polyclonal	Roche (11207733910)
FLUO-POD	1:500	Fab fragments, sheep polyclonal	Roche (11426346910)
β-actin	1:2000	Rabbit polyclonal	Abcam (ab8227)
Alexa Fluor 488	1:500	Goat anti-rabbit	Invitrogen (A11008)
Alexa Fluor 568	1:500	Goat anti-rabbit	Invitrogen (A11011)
Alexa Fluor 488	1:500	Goat anti-chicken	Life technologies (A11039)
Alexa Fluor 568	1:500	Goat anti-rat	Invitrogen (A11077)
Snail1 (for IF in cells)	1:50	Rat monoclonal	Cell signaling (SN9H2, #4719)
Histone H3	1:500	Rabbit polyclonal	Abcam (ab1791)

[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPs) (Sigma), 0.5% Tween) for 2 h at room temperature.

For hybridization, 4 pmol of DIG-labeled LNA probes were diluted in of hybrid-mix solution for each slide, and covered with Parafilm M (Sigma). The slides were kept in a sealed humidified chamber for at least 16 h at a temperature 20 °C below the melting temperature of the miRNA–LNA probes. Next day, sections were washed twice for 30 min in washing solution (50% formamide, 1X SSC, 0.1% Tween 20) at hybridization temperature, and then washed once with 0.2% SSC for 15 min at room temperature plus one wash with TBS 0.1% Tween 20. Slides were then incubated with 3% hydrogen peroxide, 0.1% Tween 20 in TBS for 30 min, and then washed 3 timed in TBS 0.1% Tween 20 at room temperature. Subsequently, sections were blocked by adding 0.5% Blocking Reagent (Roche), 10% sheep serum, 0.1% Tween 20 for 1 h at room temperature. Then, samples were incubated with anti-DIG antibody in blocking solution O/N at 4 °C. Information and dilution for antibodies are listed in Table 1. The next day, sections were washed 5 times with TBS 0.1% Tween 20, and incubated in NTMT buffer 3 times. Slides were next immersed in chambers containing developing solution (details in RNA ISH) at 37 °C until the color reaction developed. The sections were then washed several times in TBS and fixed in PFA before being photographed under a Leica DMR microscope (Leica, Wetzlar, Germany).

Generation of the *Tg(hsp68-GFP-DES1)54An* mouse transgenic line. The *Tg (hsp68-GFP-DES1)*^{54An} line was generated after pronuclear injection of the vector-free reporter construct into FVB/NJ mouse blastocysts. The reporter construct contains mouse *hsp68* (HSPA1A) minimal promoter (−872 to +1) driving eGFP followed by the SNAIL1 DES1 region (Downstream Enhancer SNAIL1). DES1 region contains the last 335 bp of SNAIL1 exon 3 and 1396 bp downstream of the gene. This region was cloned by PCR with the following primers: mmDES1-F 5′ GGATCCGCAGGGTGGTTACTGGACAC 3′ and mmDES1-R 5′ TGTCGACTCCTCCTCCCTCTCTGGAAT 3′. Founders were screened for the presence of GFP and they were mated to C57BL/6J mice to establish a line. Germ line transmission was tested both by PCR and GFP expression in E9.5 embryos. Two independent founders showed the same expression pattern and one of them (line 54) was backcrossed to C57BL/6J mice and used for further analysis.

Mouse, chicken, and zebrafish embryo sections. C57BL/6J wild-type mice were used. Embryos were staged as embryonic day (E) according to days post coitum. Wild-type Zebrafish strain AB were maintained at 28 °C under standard conditions, and the embryos were staged using zebrafish standard staging system[66]. Fertilized hen eggs were incubated in an humidified incubator at 37 °C, and embryos were staged based on Hamburger-Hamilton system[67].

For ISH to detect mRNA (see protocol above), mouse and chicken embryos were dehydrated through a series of increasing methanol:PBS-T series (25, 50, 75, and 100%) and twice in butanol, then embedded in paraffin O/N. Sectioning was performed in a Leica RM2245 microtome at 7 µm thickness.

For ISH to detect mature miRNA, mouse embryos were fixed in 4% PFA-DEPC (Diethyl pyrocarbonate) O/N. The next day embryos were washed twice with PBS, 0.1% Tween 20 before being embedded in 15% sucrose. Upon sinking, embryos were embedded in 30% sucrose, and after sinking they were kept in fresh 30% sucrose O/N. Embryos were then kept in a 1:1 mix of 30% sucrose:OCT for 30 min while rolling, before embedding in OCT. Embedded embryos were kept on dry ice and transferred to −80 °C before sectioning. OCT-embedded embryos were cryosectioned in a Slee MNT cryotome at 10 µm, dried for 2 h at room temperature before either being directly used for ISH (see protocol below) or stored at −80 °C.

Zebrafish embryos that were subjected to double-fluorescent ISH (see protocol below) were directly embedded in 4% low-melting agarose and sectioned using a Leica VT1000S vibratome at 100 µm, mounted using Dako fluorescent mounting medium and subjected to Confocal microscope imaging.

We affirm to have complied with all relevant ethical regulations for animal testing and research. All animal procedures were conducted in compliance with the European Community Council Directive (2010/63/EU) and Spanish legislation. The protocols were approved by the CSIC Ethical Committee and the Animal Welfare Committee at the Institute of Neurosciences, Alicante.

miR-15 family sponge generation and injection in zebrafish. We designed miR-15 sponge as inhibitors that can quench all members of the family[35], with the following sequence, ATCGTCGTAGTCATACCAAAAGCAATTCC (Supplementary Fig. 4k), in three consecutive repeats and cloned in pCS2 construct. Complementary sequences were chosen again conserved sequences among different family members, and few spacer nucleotides in between serve make bulges that can be recognized by the cell degradation machinery[35]. After linearization of the vector to synthesize the sponge transcripts followed by SV40 poly A, RNA was synthesized using the mMESSAGE mMACHINE SP6 Kit (Ambion) following manufacturer's indications prior injection in zebrafish.

For zebrafish embryo injections we first titrated the non-toxic effective dose of synthesized miR-15 family sponges to be injected safely. 100 ng of sponges were injected in the yolk of 1–2 cell stage embryos and incubated at 25 °C. At 20-somite stage, embryos were collected and fixed in 4% paraformaldehyde overnight at 4 °C for ISH or collected and lysed immediately for RNA extraction.

Immunofluorescent (IF) staining. Whole-mount mouse and chicken embryos were fixed in 4% PFA for 2 h at 4 °C. Then, embryos were dehydrated in a series of PBS 1% Triton x100 (PBS-T):methanol proportions (25, 50, 75, and 100%), then kept in 100% methanol O/N at −20 °C. Embryos where then rehydrated in the reverse order of PBS-T:methanol proportions. Antigen retrieval was performed by treating the mouse embryos with either 150 mM Tris-EDTA pH 9.0 (for Prrx1) or 10 mM Sodium Citrate pH 6.0 (for Snail1)-1% Triton x100 buffer at 70 °C for 20 min, following three washes with PBS-T. Embryos were blocked with 5% NGS 1% BSA 1% Tx-100 for 5 h or O/N at 4 °C and then incubated with the primary antibodies 48 h at 4 °C. After several washes for several hours in PBS-T, the embryos were incubated with the secondary antibodies and DAPI O/N at 4 °C. After washing the secondary antibody with PBS-T, embryos were subjected to imaging. Pictures of whole-mount embryos were taken with an Olympus FV1200 confocal microscope with 10/20× objective and then subjected to mosaic merge using Image J software. Information and dilution for antibodies are listed in Table 1.

For cell lines IF, cells were cultured and treated on cover-slips in 6-well plates and collected at corresponding time points, fixed with 4% PFA for 15 min at room temperature, and washed three times with PBS. After washing fixed cells were either directly subjected to IF or stored at 4 °C. Cells were blocked with 5% NGS 1% BSA 0.2% Triton x-100 1 hour at room temperature and incubated with the primary antibodies O/N at 4 °C. After washing three times with PBS, cells were incubated with the secondary antibodies and DAPI 1 h at room temperature. Information and dilution for antibodies are listed in Table 1. After washing the secondary antibody with PBS, cells were photographed. Pictures were taken with Leica DMi8 microscope and HAMAMTSU C11440 digital camera.

BMP treatment on zebrafish and chicken embryos. Zebrafish embryos were collected and dechorionated at around 4-somite stage and treated with 200 µg/ml

of Human recombinant BMP4 protein (R&D Systems) which was added to the water, and kept at 25 °C. Embryos were collected after 30 min or 12 h and subjected to RNA extraction –qPCR. Collected embryos at different time points were subjected to ISH and IF for *SNAIL1* and *PRRX1*. Chicken embryos were collected and cultured for 30 min prior BMP-soaked bead implantation. Human recombinant BMP2 protein (R&D Systems) was loaded onto heparin acrylic beads at a concentration of 0.2 mg/ml by soaking for 3 h at room temperature, and they were introduced into one half of the lateral plate mesoderm of chick embryos, close to the developing somites, at stage HH7, which were harvested at different time points (between 30 min to 10 h). Beads loaded with PBS were used as controls.

Western blot. E11.5 mouse embryos were dissected and lysed directly in RIPA buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% NP40, 1% sodium deoxycholate, 1% SDS and Proteinase inhibitor cocktail). Homozygous LacZ-Knock-in Prrx1 mutant embryos were identified among littermates by PCR genotyping and used for further analyses, together with wild-type embryos. Lysates were then homogenized followed by heating, sonication (6 rounds of 15 min, 30 s ON/30 s OFF, high). Samples were then loaded and subjected Coomassie Blue staining to determine the quality and concentration of the samples. Next, protein samples were loaded and migrated in 12% acrylamide gel and then transferred to PVDF western blotting membrane (Roche). The membranes were blocked with 10% milk and incubated with anti-Snail1 antibody overnight. Information and dilution for antibodies are listed in Table 1. The next day, membranes were extensively washed and incubated with secondary Peroxidase goat anti-Rabbit antibody, washed and developed using Chemiluminescent HRP Substrate (ImmobilonTM Western, Millipore). Pictures were obtained using Amersham Imager 680 (GE Healthcare). β-actin (Rabbit polyclonal, Abcam) was used as housekeeping protein for normalization. The uncropped and unprocessed scans of the blot are supplied in Supplementary Fig. 9.

In silico analyses. Single-cell RNA-seq public data**:** Processed datasets of publicly available data for single-cell RNA-seq were downloaded from NCBI Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). Data included single-cell RNA-seq from developing zebrafish and mouse embryos and different cancer types (Supplementary Table 1). Among analyzed single cells, the ones with no value for both *Snail1* and *Prrx1* were excluded. Hierarchical clustering was performed using GENE-E (version 3.0.215, Broad Institute, Inc.). Values were subjected to correlation analyses using Prism (GraphPad Softwares, Version 6.01, 2012) calculated based on Spearman *r*.

Metanalysis of oncogenomic data from breast cancer patients. To assess the putative correlations between the expression of *SNAIL1, PRRX1* and mature *miR-15* family members with OS in all subtypes ($n = 1403$), lymph-node positive patients ($n = 313$) or basal subtype ($n = 241$), all available data sets (E-TABM-43, GSE16716, GSE18728, GSE20194, GSE20271, GSE31448, GSE32646, GSE41998, GSE6532, GSE20711, GSE7390, GSE21653, GSE31519, GSE5327, GSE17907, E-MTAB-365, GSE37946, GSE2034, GSE2990, GSE17705, GSE1456, GSE12093, GSE9195, GSE45255, GSE20685, GSE12276, GSE2603, GSE16391, GSE42568, GSE11121, GSE3494, GSE16446, GSE4611, GSE26971, and GSE19615) were analyzed. Kaplan–Meier plots were generated using http://kmplot.com[37,68]. The patient samples were grouped as either high or low for the expression of the genes of interest, and the auto-select best cut-off was chosen for computation over the entire data set.

Statistical and data analysis. Images were prepared using Adobe Photoshop and Adobe Illustrator CS6. Sample size was estimated using GPower 3.1, and values were set at $p = 0.05$ and beta = 0.8. All statistical analyses were performed using Microsoft Excel 2013 and Prism (GraphPad Softwares, Version 6.01, 2012). For reporter assays and qRT–PCR experiments, the corresponding treatments were compared with controls using Student's two-tailored *t* test or One-way ANOVA with Bonferroni's multiple comparison test. Spearman r was used for correlation test of single-cell RNA-seq data. All bar graphs represent mean + SEM (Standard Error of the Mean). Statistical significances were as follows: $^{*}P \le 0.05$, $^{**}P \le 0.01$ and $^{***}P \le 0.001$.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The microarray data generated in this study, as well as public scRNA-seq data datasets analyzed during the current study are available in the Gene Expression Omnibus (GEO) repository under the following accession numbers that are also listed in Supplementary Table 1. Microarray in MDA231 cells: GSE138078, Zebrafish embryos: GSM3067194, mouse embryo: GSE87038, head and neck cancer patients: GSE103322, and breast cancer patients: GSE75688. The source data underlying Fig. 1c, f–h, 2a–f, 3b–f, 4a–e and h, and 6a, and Supplementary Figs. 1f, g, 2a–f, 3a and c–i, 4a and g–j, 5c and e, and 7b are provided as a Source Data file.

Received: 23 May 2019; Accepted: 18 October 2019;
Published online: 11 November 2019

References

1. Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. Emt: 2016. *Cell* **166**, 21–45 (2016).
2. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
3. Nieto, M. A. Epithelial plasticity: a common theme in embryonic and cancer cells. *Science* **342**, 1234850 (2013).
4. Lu, W. & Kang, Y. Epithelial-Mesenchymal Plasticity in Cancer Progression and Metastasis. *Dev. Cell* **49**, 361–374 (2019).
5. Nieto, M. A. Context-specific roles of EMT programmes in cancer cell dissemination. *Nat. Cell Biol.* **19**, 416–418 (2017).
6. Mani, S. A. et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704–715 (2008).
7. Wellner, U. et al. The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nat. Cell Biol.* **11**, 1487–1495 (2009).
8. Ocana, O. H. et al. Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. *Cancer Cell* **22**, 709–724 (2012).
9. Shi, L. et al. A SIRT1-centered circuitry regulates breast cancer stemness and metastasis. *Oncogene* **37**, 6299–6315 (2018).
10. Barrallo-Gimeno, A. & Nieto, M. A. The Snail genes as inducers of cell movement and survival: implications in development and cancer. *Development* **132**, 3151–3161 (2005).
11. Ye, X. et al. Distinct EMT programs control normal mammary stem cells and tumour-initiating cells. *Nature* **525**, 256–260 (2015).
12. Ocana, O. H. et al. A right-handed signalling pathway drives heart looping in vertebrates. *Nature* **549**, 86–90 (2017).
13. Bartel, D. P. Metazoan MicroRNAs. *Cell* **173**, 20–51 (2018).
14. Diaz-Lopez, A., Moreno-Bueno, G. & Cano, A. Role of microRNA in epithelial to mesenchymal transition and metastasis and clinical perspectives. *Cancer Manag. Res.* **6**, 205–216 (2014).
15. Jaillon, O. et al. Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* **431**, 946–957 (2004).
16. Wagner, D. E. et al. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981–987 (2018).
17. Dong, J. et al. Single-cell RNA-seq analysis unveils a prevalent epithelial/mesenchymal hybrid state during mouse organogenesis. *Genome Biol.* **19**, 31 (2018).
18. Klijn, C. et al. A comprehensive transcriptional portrait of human cancer cell lines. *Nat. Biotechnol.* **33**, 306–312 (2015).
19. Blick, T. et al. Epithelial mesenchymal transition traits in human breast cancer cell lines. *Clin. Exp. Metastasis* **25**, 629–642 (2008).
20. Sun Y. et al. Resveratrol inhibits the migration and metastasis of MDA-MB-231 human breast cancer by reversing TGF-beta1-induced epithelial-mesenchymal transition. *Molecules* **24** pii: E1131 (2019).
21. Manuel Iglesias, J. et al. Mammosphere formation in breast carcinoma cell lines depends upon expression of E-cadherin. *PLoS ONE* **8**, e77281 (2013).
22. Chung, W. et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat. Commun.* **8**, 15081 (2017).
23. Puram, S. V. et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell* **171**, 1611–1624 e1624 (2017).
24. Acloque, H. et al. Reciprocal repression between Sox3 and snail transcription factors defines embryonic territories at gastrulation. *Dev. Cell* **21**, 546–558 (2011).
25. Acloque, H., Ocana, O. H. & Nieto, M. A. Mutual exclusion of transcription factors and cell behaviour in the definition of vertebrate embryonic territories. *Curr. Opin. Genet. Dev.* **22**, 308–314 (2012).
26. Nieto, M. A. The snail superfamily of zinc-finger transcription factors. *Nat. Rev. Mol. Cell Biol.* **3**, 155–166 (2002).
27. Villarejo, A., Cortes-Cabrera, A., Molina-Ortiz, P., Portillo, F. & Cano, A. Differential role of Snail1 and Snail2 zinc fingers in E-cadherin repression and epithelial to mesenchymal transition. *J. Biol. Chem.* **289**, 930–941 (2014).
28. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
29. Xie, Z., Hu, S., Blackshaw, S., Zhu, H. & Qian, J. hPDI: a database of experimental human protein-DNA interactions. *Bioinformatics* **26**, 287–289 (2010).

30. Grueneberg, D. A., Natesan, S., Alexandre, C. & Gilman, M. Z. Human and Drosophila homeodomain proteins that enhance the DNA-binding activity of serum response factor. *Science* **257**, 1089–1095 (1992).
31. Fromm, B. et al. A uniform system for the annotation of vertebrate microRNA genes and the evolution of the human microRNAome. *Annu Rev. Genet* **49**, 213–242 (2015).
32. Drasin, D. J. et al. TWIST1-induced miR-424 reversibly drives mesenchymal programming while inhibiting tumor initiation. *Cancer Res.* **75**, 1908–1921 (2015).
33. Li, Y. et al. Metastatic heterogeneity of breast cancer cells is associated with expression of a heterogeneous TGFbeta-activating miR424-503 gene cluster. *Cancer Res.* **74**, 6107–6118 (2014).
34. Liu, Q. et al. miR-16 family induces cell cycle arrest by regulating multiple cell cycle genes. *Nucleic Acids Res.* **36**, 5391–5404 (2008).
35. Ebert, M. S., Neilson, J. R. & Sharp, P. A. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nat. Methods* **4**, 721–726 (2007).
36. Lu, M. F. et al. prx-1 functions cooperatively with another paired-related homeobox gene, prx-2, to maintain cell fates within the craniofacial mesenchyme. *Development* **126**, 495–504 (1999).
37. Gyorffy, B. et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res. Treat.* **123**, 725–731 (2010).
38. Denkert, C., Liedtke, C., Tutt, A. & von Minckwitz, G. Molecular alterations in triple-negative breast cancer-the road to new treatment strategies. *Lancet* **389**, 2430–2442 (2017).
39. Polyak, K. Heterogeneity in breast cancer. *J. Clin. Invest.* **121**, 3786–3788 (2011).
40. Sarrio, D. et al. Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype. *Cancer Res.* **68**, 989–997 (2008).
41. Sefton, M., Sanchez, S. & Nieto, M. A. Conserved and divergent roles for members of the Snail family of transcription factors in the chick and mouse embryo. *Development* **125**, 3111–3121 (1998).
42. Del Barrio, M. G. & Nieto, M. A. Relative expression of Slug, RhoB, and HNK-1 in the cranial neural crest of the early chicken embryo. *Dev. Dyn.* **229**, 136–139 (2004).
43. Corish, P. & Tyler-Smith, C. Attenuation of green fluorescent protein half-life in mammalian cells. *Protein Eng.* **12**, 1035–1040 (1999).
44. de Herreros, A. G., Peiro, S., Nassour, M. & Savagner, P. Snail family regulation and epithelial mesenchymal transitions in breast cancer progression. *J. Mammary Gland Biol. Neoplasia* **15**, 135–147 (2010).
45. Hsu, D. S. et al. Acetylation of snail modulates the cytokinome of cancer cells to enhance the recruitment of macrophages. *Cancer Cell* **26**, 534–548 (2014).
46. Cano, A. et al. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* **2**, 76–83 (2000).
47. Lagha, M. et al. Paused Pol II coordinates tissue morphogenesis in the Drosophila embryo. *Cell* **153**, 976–987 (2013).
48. Peinado, H., Quintanilla, M. & Cano, A. Transforming growth factor beta-1 induces snail transcription factor in epithelial cell lines: mechanisms for epithelial mesenchymal transitions. *J. Biol. Chem.* **278**, 21113–21123 (2003).
49. Peiro, S. et al. Snail1 transcriptional repressor binds to its own promoter and controls its expression. *Nucleic Acids Res.* **34**, 2077–2084 (2006).
50. Tsai, J. H., Donaher, J. L., Murphy, D. A., Chau, S. & Yang, J. Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell* **22**, 725–736 (2012).
51. Celia-Terrassa, T. et al. Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells. *J. Clin. Invest.* **122**, 1849–1868 (2012).
52. Beerling, E. et al. Plasticity between epithelial and mesenchymal states unlinks EMT from metastasis-enhancing stem cell capacity. *Cell Rep.* **14**, 2281–2288 (2016).
53. Thiery, J. P., Acloque, H., Huang, R. Y. & Nieto, M. A. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871–890 (2009).
54. Buitrago-Delgado, E., Nordin, K., Rao, A., Geary, L. & LaBonne, C. NEURODEVELOPMENT. Shared regulatory programs suggest retention of blastula-stage potential in neural crest cells. *Science* **348**, 1332–1335 (2015).
55. De Craene, B. et al. The transcription factor snail induces tumor cell invasion through modulation of the epithelial cell differentiation program. *Cancer Res.* **65**, 6237–6244 (2005).
56. Miyoshi, A. et al. Snail and SIP1 increase cancer invasion by upregulating MMP family in hepatocellular carcinoma cells. *Br. J. Cancer* **90**, 1265–1273 (2004).
57. Vega, S. et al. Snail blocks the cell cycle and confers resistance to cell death. *Genes Dev.* **18**, 1131–1143 (2004).
58. Sugiyama, M. et al. Paired related homeobox 1 is associated with the invasive properties of glioblastoma cells. *Oncol. Rep.* **33**, 1123–1130 (2015).
59. Gentleman, R. C. et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).
60. Irizarry, R. A. et al. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* **31**, e15 (2003).
61. Smyth G. K. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. In: *Statistical Applications in Ge- netics and Molecular Biology*. **3** (Springer, 2004).
62. Kim, J. H. et al. High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* **6**, e18556 (2011).
63. Bolos, V. et al. The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J. Cell Sci.* **116**, 499–511 (2003).
64. Nieto, M. A., Patel, K. & Wilkinson, D. G. In situ hybridization analysis of chick embryos in whole mount and tissue sections. *Methods Cell Biol.* **51**, 219–235 (1996).
65. Pena, J. T. et al. miRNA in situ hybridization in formaldehyde and EDC-fixed tissues. *Nat. Methods* **6**, 139–141 (2009).
66. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253–310 (1995).
67. Hamburger, V. & Hamilton, H. L. A series of normal stages in the development of the chick embryo. *J. Morphol.* **88**, 49–92 (1951).
68. Lanczky, A. et al. miRpower: a web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients. *Breast Cancer Res. Treat.* **160**, 439–446 (2016).

Acknowledgements

We thank Sonia Vega for her help and support in managing cell lines, Verona Villar Cerviño for advice with imaging analysis, Cristina López Blau, Diana Abad Bataller and Sandra Moreno Valverde for helpful technical support, and other lab members for continuous and helpful discussions. Also we thank Antonio Caler Escribano for technical help in the FACS/Omics facility. This work was supported by grants from the Spanish Ministries of Economy and Competitiveness (MINECO BFU2014-53128-R), of Science, Innovation and Universities (MICIU RTI2018-096501-B-I00), Generalitat Valenciana (2017/150) and the European Research Council (ERC AdG 322694) to M.A.N., who also acknowledges financial support from the Spanish State Research Agency, through the "Severo Ochoa" Program for Centres of Excellence in R&D (SEV-2017–0273). H.F. was recipient of PhD student scholarships Santiago Grisolia from Generalitat Valenciana (GRISOLIA/2014/004). L.R. was holder of a Juan de la Cierva Formación postdoctoral scholarship and F.G.A. is recipient of a PhD student scholarship (FPI), both from MINECO.

Author contributions

H.F. performed the majority of experiments, analyzed and interpreted the data, and wrote the paper. L.R. injected and treated the zebrafish embryos, did the ChIP assay in Supplementary Fig. 2c, e and contributed to ISHs in Figs. 1, 3, and 5. K.K.Y. contributed to conditional in vitro gain of function of Prrx1 (Fig. 3c) and expression analyses in Supplementary Fig. 1e, plus knockdown experiments in Supplementary Fig. 1g. O.H.O. performed the MicroArray experiments (Fig. 3a), BMP treatment of chicken embryos and ISHs in Fig. 1d. F.G.A. performed western-blots in Fig. 5i and LacZ staining (Supplementary Fig. 6a). A.A. analyzed the MicroArray (Fig. 3a) and did some in silico analysis. J.G. conceived, designed, and generated the mouse transgenic line. M.A.N. conceived the project, designed experiments, interpreted the data, wrote the paper, and secured funding.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-019-13091-8.

Correspondence and requests for materials should be addressed to M.A.N.

Peer review information *Nature Communications* thanks Gregory Goodall and other, anonymous, reviewers for their contributions to the peer review information. Peer review reports are available.

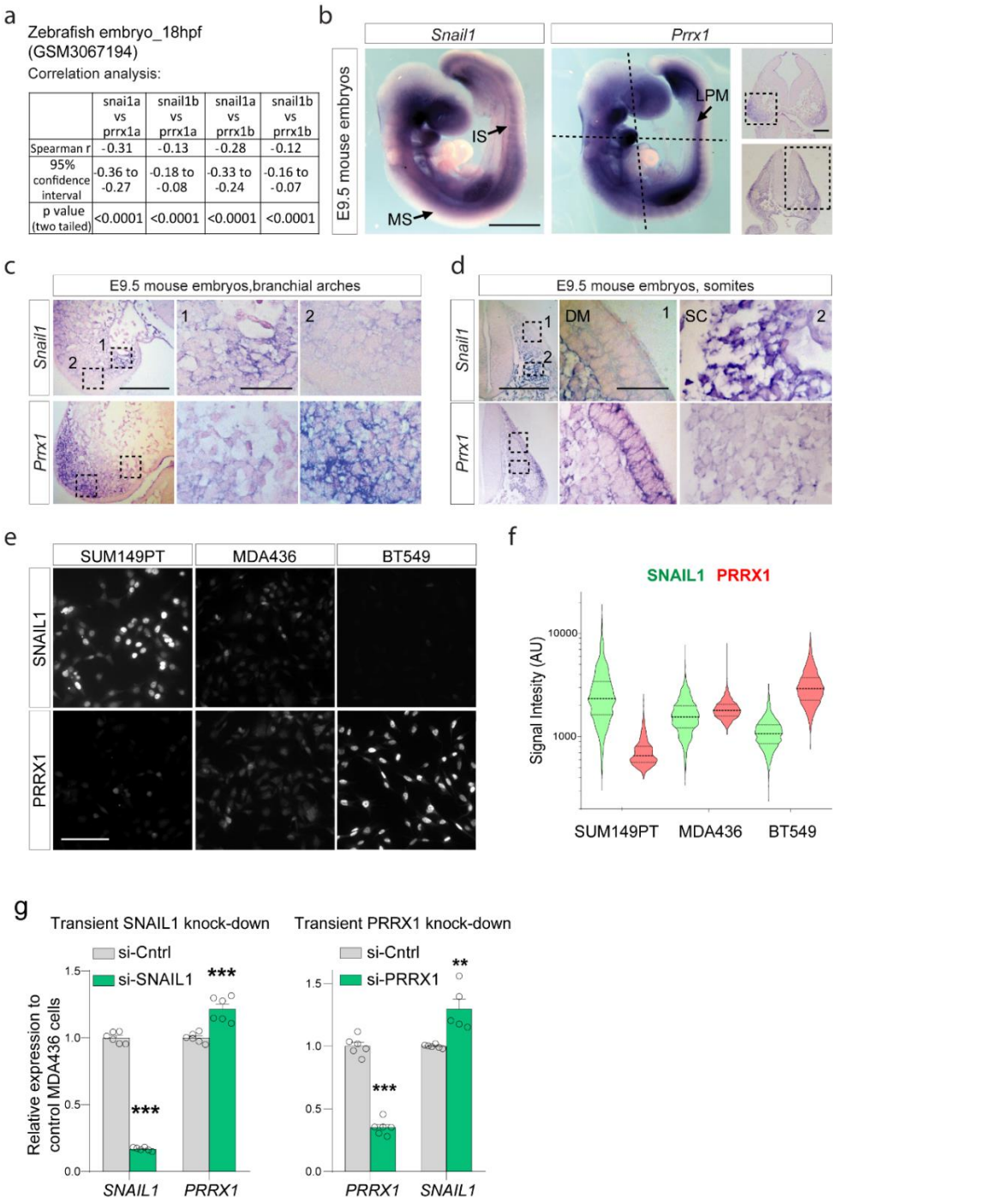
Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary material

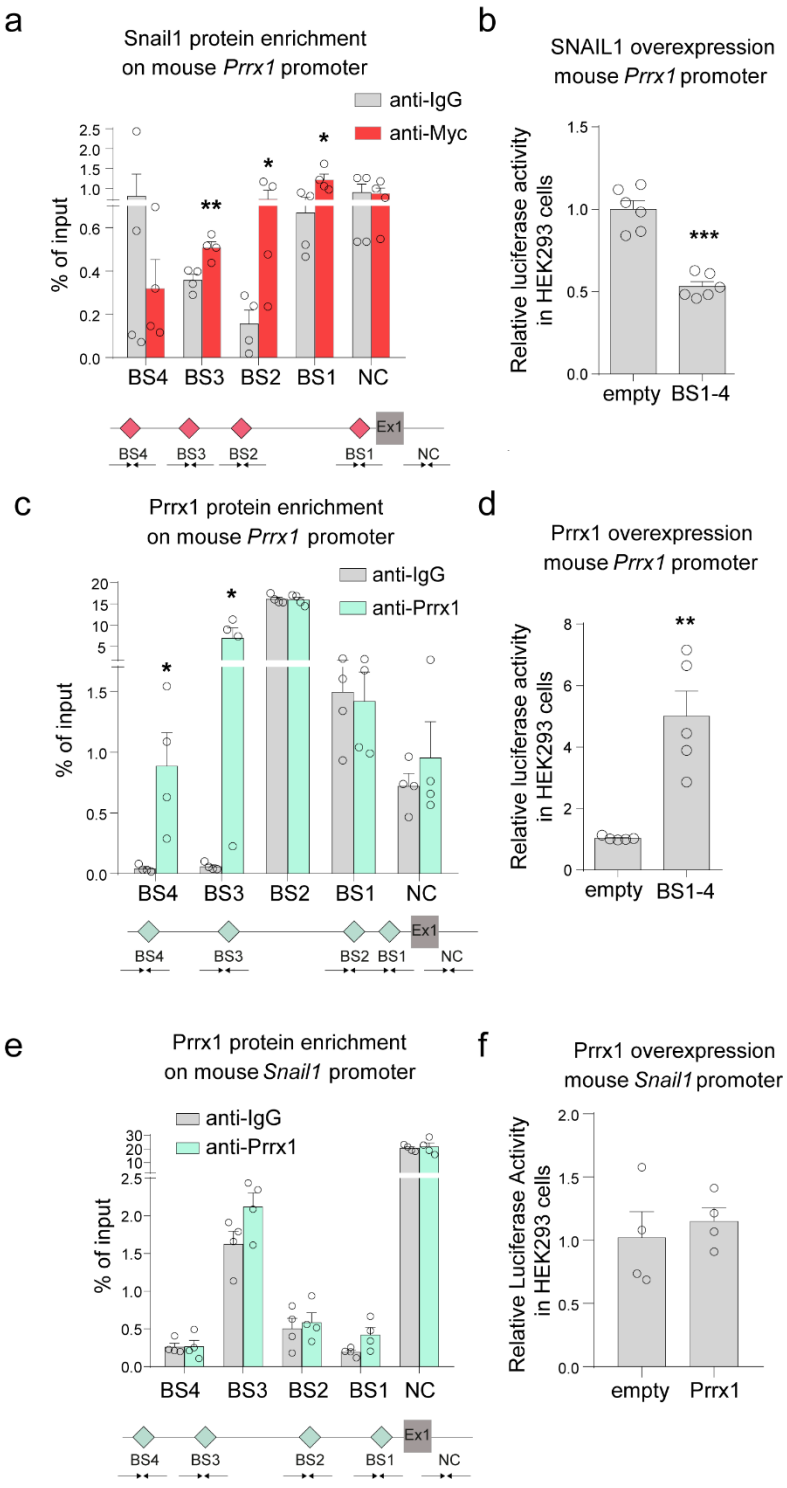
A Gene Regulatory Network to Control EMT Programs in Development and Disease

Fazilaty et al.

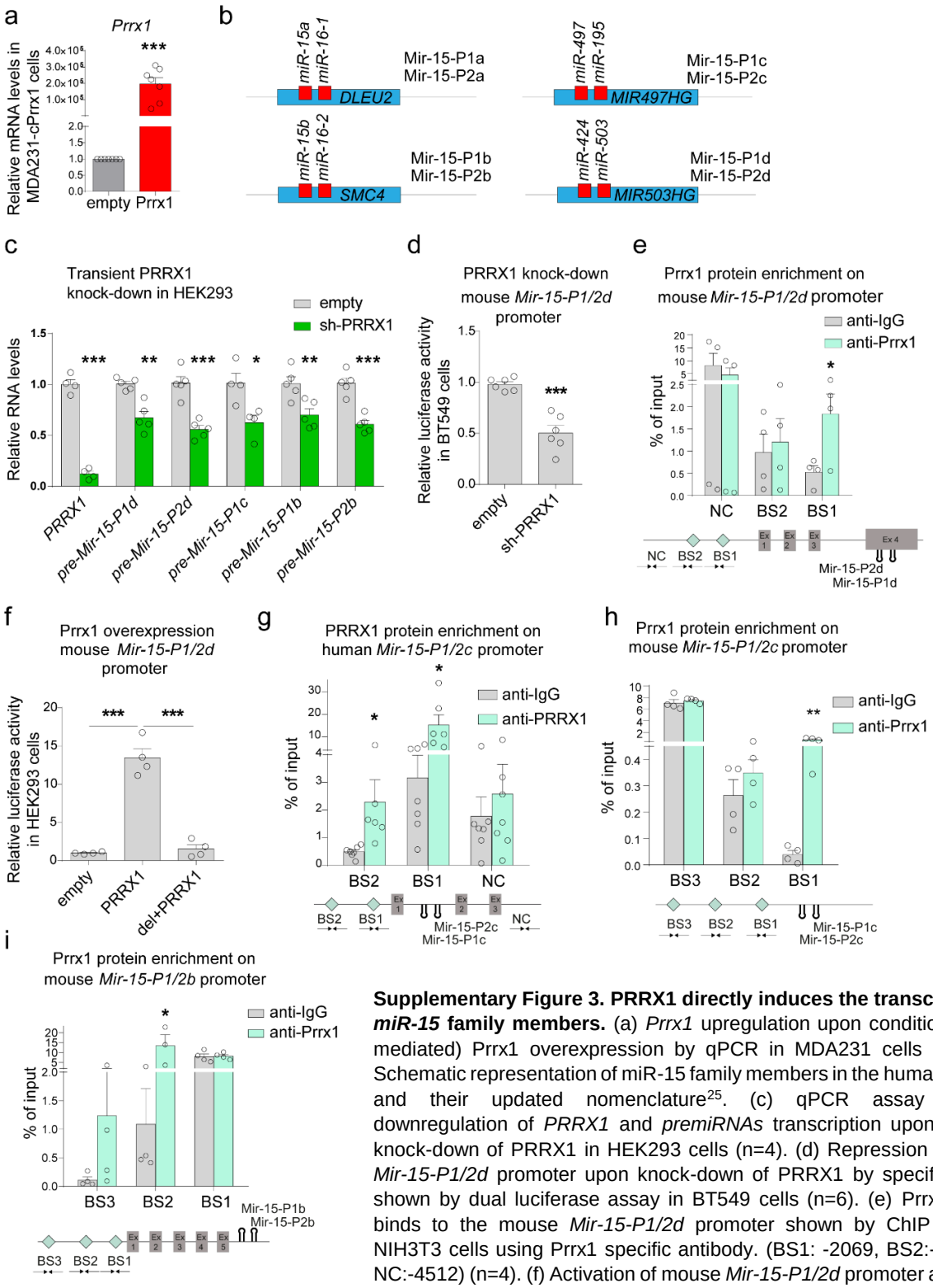


	snai1a vs prrx1a	snail1b vs prrx1a	snail1a vs prrx1b	snail1b vs prrx1b
Spearman r	-0.31	-0.13	-0.28	-0.12
95% confidence interval	-0.36 to -0.27	-0.18 to -0.08	-0.33 to -0.24	-0.16 to -0.07
p value (two tailed)	<0.0001	<0.0001	<0.0001	<0.0001

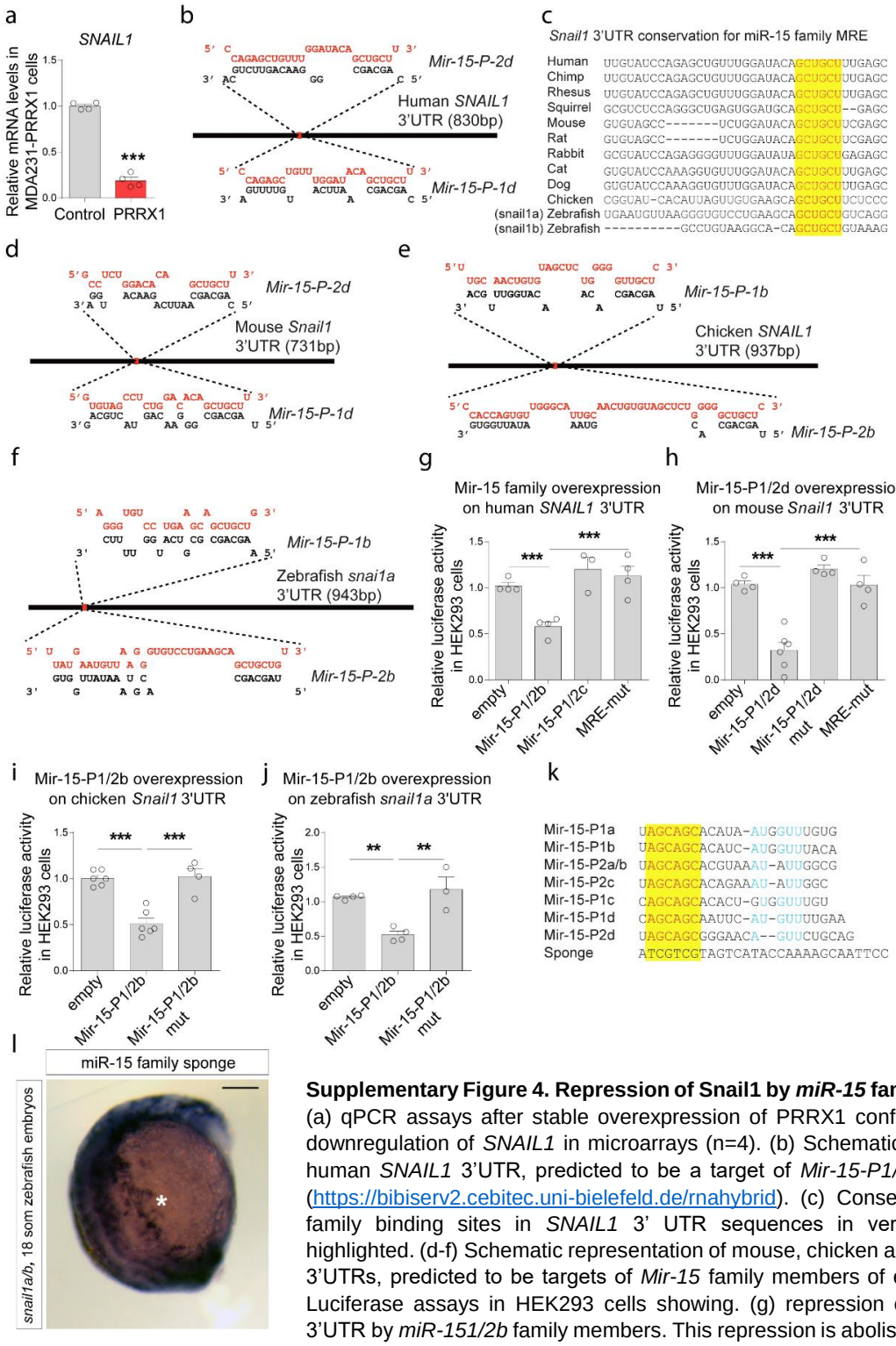
Supplementary Figure 1. Snail1 and Prrx1 complementary expression in development and disease. (a) Correlation analyses between each pair of *snai1a*/b and *prrx1a/b* genes extracted from scRNA-seq data (GSM3067194) of 18hpf zebrafish embryos. (b) Expression of *Snail1* and *Prrx1* in lateral views of E9.5 mouse embryos. Scale bars: 250 μm for whole mounts and 100 μm for sections. (c-d) Transverse sections of the embryos shown in (b) to better observe the branchial arches (c) and the somites (d), together with higher power pictures of the boxed areas (1-2). Complementary expression is observed in the branchial arches and the somites. Scale bars: 100 μm for sections and 25 μm for insets (boxes 1-2). (e) Double IF for SNAIL1 and PRRX1 in three breast cancer cell lines, SUM149PT, MDA436 and BT549, with different levels for each protein. Scale bar: 100 μm. (f) Signal intensity analyses for SNAIL1 and PRRX1 in the three cell lines show a gradual reduction of SNAIL1 concomitant with a gradual increase for PRRX1 levels from SUM to BT cells. (g) qPCR assay showing upregulation of *PRRX1* (left) and *SNAIL1* (right) transcription upon transient knock-down of SNAIL1 or PRRX1, respectively, in MDA436 cells. hpf: hours post fertilization; IS: immature somites; MS: mature somites; LMP: lateral plate mesoderm; DM: dermomyotome; SC: sclerotome; AU: arbitrary units; si: short interfering RNA (siRNA). Bars represent mean plus SEM, n=6 independent experiments as biological replicates and asterisks indicate significant p-value in t-test (** $p < 0.01$ and *** $p < 0.001$). Source data are provided as a Source Data file.



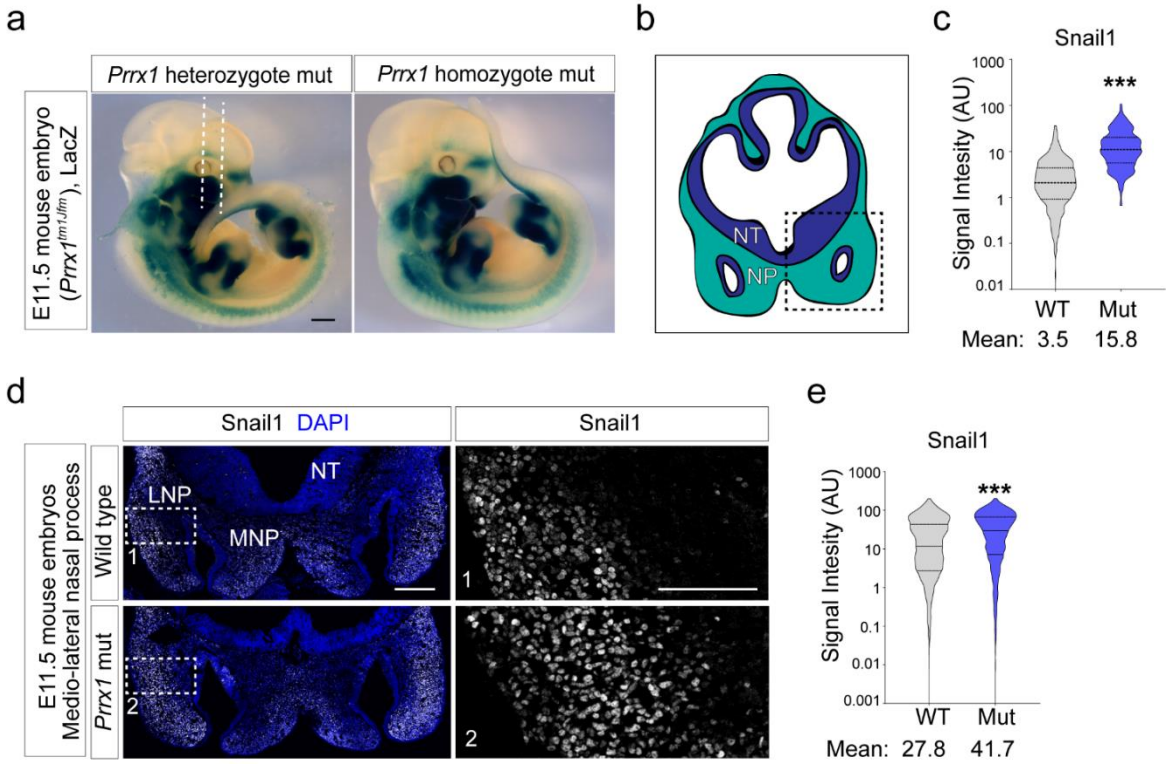
Supplementary Figure 2. Snail1 and Prrx1 act in an antagonistic manner on *Prrx1* promoter in the mouse. (a) Snail1 enrichment in the mouse *Prrx1* promoter shown by ChIP assay in NIH3T3 cells using anti Myc antibody (for Snail1-Myc overexpression). A schematic map is shown, where red diamonds represent Snail1 potential binding sites (E-boxes; CANNTG) on *Prrx1* promoter. (BS1: -265, BS2: -2100, BS3: -3028, BS4: -4074 and NC: +2099) Ex1: Prrx1 exon 1 (n=4). (b) Decrease in mouse *Prrx1* promoter activity after SNAIL1 transfection in HEK293 cells (n=6). (c) Prrx1 directly binds to its own promoter, as assessed by ChIP assays in NIH3T3 cells using Prrx1 specific antibody. A schematic map is shown, where cyan diamonds represent Prrx1 potential binding sites (TAATKDS) on its own promoter. (BS1: -114, BS2: -662, BS3: -2113, BS4: -4053 and NC: +2099) Ex1: Prrx1 exon 1 (n=4). (d) Activation of mouse *Prrx1* promoter after Prrx1 overexpression in luciferase assays in HEK293 cells (n=5). (e) Lack of enrichment for Prrx1 binding to mouse *Snail1* promoter in ChIP assays in NIH3T3 cells with a Prrx1 antibody. (BS1-4: -3110, -2827, -2667, -1714 and NC: +302). Cyan diamonds represent Prrx1 potential binding sites on *Snail1* promoter (n=4). (f) Luciferase assays in HEK293 cells showing the lack of significant difference in the activity of mouse *Snail1* promoter after Prrx1 overexpression. Arrows represent primers used for qPCR amplification. Locations of red and cyan diamonds represent distances between BSs and the promoter. Ex1: *Snail1* exon 1 (n=4). BS: biding site and NC: negative control region, which does not contain potential BS. Symbols in binding sites are as follows. K: T/G, D: G/A/T, S: G/C and N: G/A/T/C. Bars represent mean plus standard error of the mean (SEM), indicated (n) represent number of independent experiments as biological replicates and asterisks indicate significant p-value in t-test (* $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$). Source data are provided as a Source Data file.



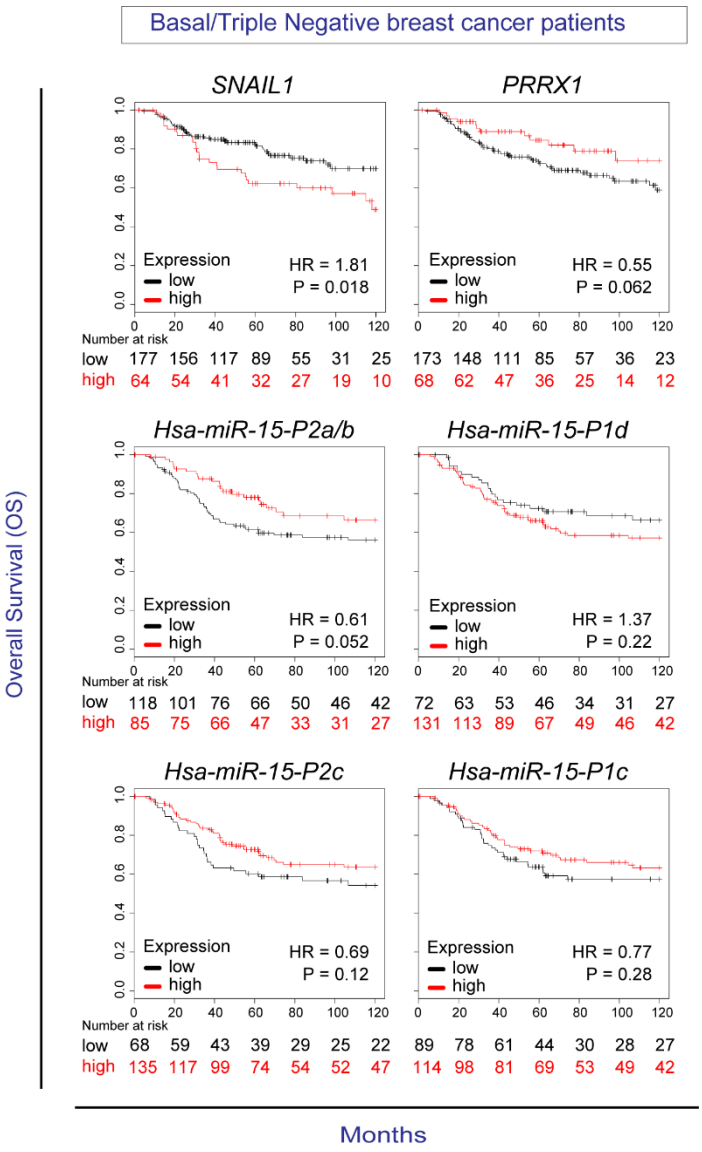
Supplementary Figure 3. PRRX1 directly induces the transcription of *miR-15* family members. (a) *Prrx1* upregulation upon conditional (Dox-mediated) Prrx1 overexpression by qPCR in MDA231 cells (n=7). (b) Schematic representation of miR-15 family members in the human genome and their updated nomenclature[25]. (c) qPCR assay showing downregulation of *PRRX1* and *premiRNAs* transcription upon transient knock-down of PRRX1 in HEK293 cells (n=4). (d) Repression of human *Mir-15-P1/2d* promoter upon knock-down of PRRX1 by specific shRNA shown by dual luciferase assay in BT549 cells (n=6). (e) Prrx1 directly binds to the mouse *Mir-15-P1/2d* promoter shown by ChIP assay in NIH3T3 cells using Prrx1 specific antibody. (BS1: -2069, BS2:-2972 and NC:-4512) (n=4). (f) Activation of mouse *Mir-15-P1/2d* promoter after Prrx1 transfection in HEK293 cells. This activation is abolished upon deletion of the Prrx1 binding site in *Mir-15-P1/2d* promoter (n=4). (g-i) Prrx1 directly binds to the human *Mir-15-P1/2c* (BS1: -214, BS2: -2134 and NC: +5141) (n=6), mouse *Mir-15-P1/2c* (BS1:-627) (n=4) and *Mir-15-P1/2b* (BS1: -999 and BS2: -1386) (n=4) promoters shown by ChIP assay in BT549 and NIH3T3 cells, respectively, using Prrx1 specific antibody. Locations of cyan diamonds represent distances between BSs and the promoter. cPrrx1: conditional Prrx1 overexpression; BS: biding site and NC: negative control region, which does not contain potential BS. Bars represent mean plus standard error of the mean (SEM), indicated (n) represent number of independent experiments as biological replicates and asterisks indicate significant p-value in t-test for all except f for which one-way ANOVA with Bonferroni's multiple comparison test is performed (* $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$). Source data are provided as a Source Data file.



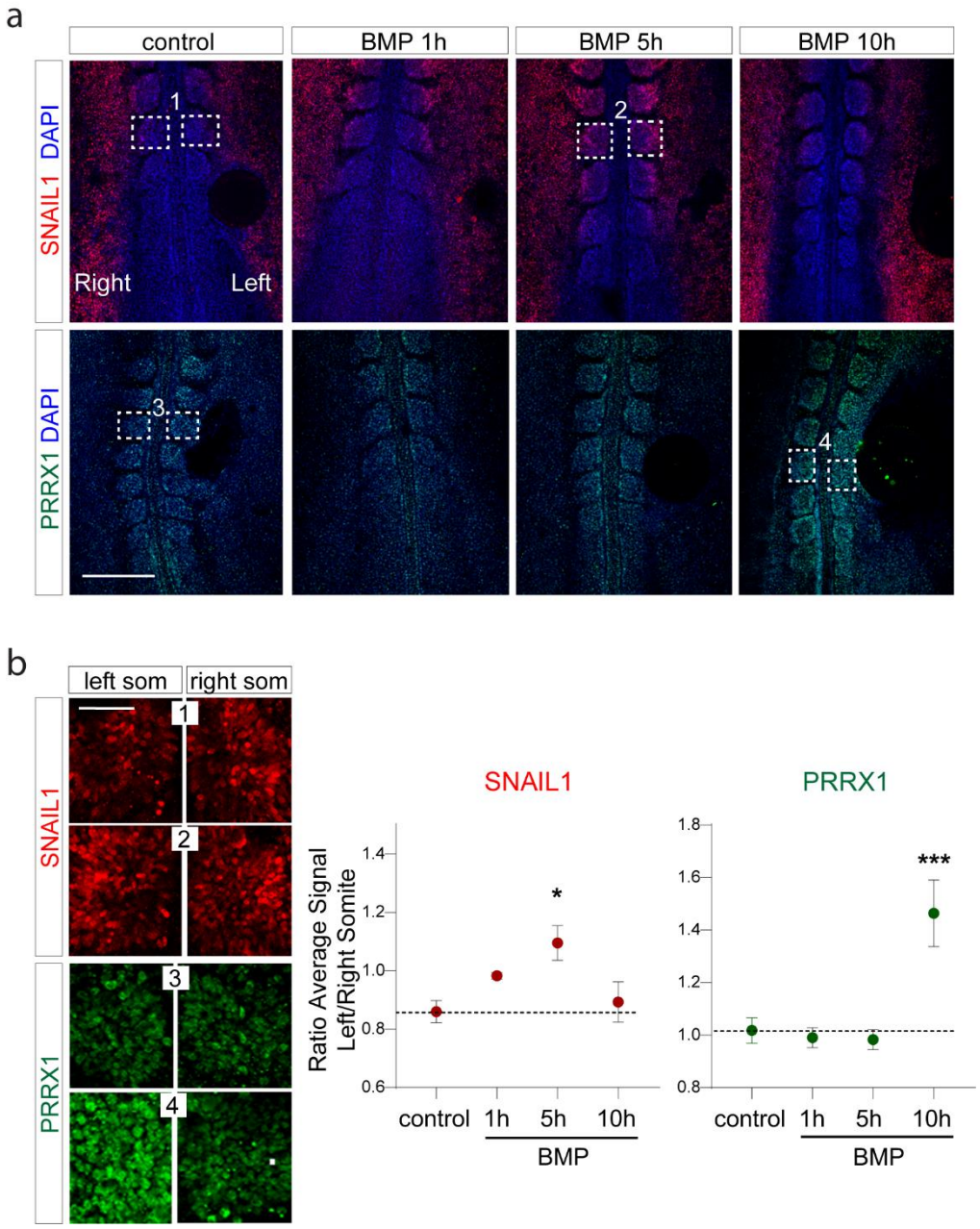
Supplementary Figure 4. Repression of Snail1 by *miR-15* family is conserved. (a) qPCR assays after stable overexpression of PRRX1 confirms the observed downregulation of *SNAIL1* in microarrays (n=4). (b) Schematic representation of human *SNAIL1* 3'UTR, predicted to be a target of *Mir-15-P1/2d*, by RNAhybrid (https://bibiserv2.cebitec.uni-bielefeld.de/rnahybrid). (c) Conservation of miR-15 family binding sites in *SNAIL1* 3' UTR sequences in vertebrates. MRE is highlighted. (d-f) Schematic representation of mouse, chicken and zebrafish *Snai1* 3'UTRs, predicted to be targets of *Mir-15* family members of each species. (g-j) Luciferase assays in HEK293 cells showing. (g) repression of human *SNAIL1* 3'UTR by *miR-151/2b* family members. This repression is abolished upon mutation of *MRE* in *SNAIL1* 3'UTR (n=4). (h-j) Similar experiments to those shown in (d), but using either mouse or chicken *Snail1* or zebrafish *snail1a* 3'UTR regions, respectively. These repressions are abolished upon mutations in the miRNAs seed sequence (n=4 except for i that n=6). (k) Sequence alignment for members of the miR-15 family, and the partially complementary sequence of the Sponge used. The seed sequence is highlighted in yellow box and blue nucleotides indicate additional conserved residues within family members. MRE: miRNA responsive element. (l) miR-15 family sponge-injected zebrafish embryos showing *snail1a/b* expression by *in situ* hybridization. Note the ectopic *snail1* positive cells that migrated beyond their normal position shown by asterisk. This very extreme phenotype is observed in a small percentage of embryos. Scale bar: 250 μm. Bars represent mean plus standard error of the mean (SEM), indicated (n) represent number of independent experiments as biological replicates and asterisks indicate significant p-value in t-test for a and one-way ANOVA with Bonferroni's multiple comparison test for the rest (* $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$). Source data are provided as a Source Data file.



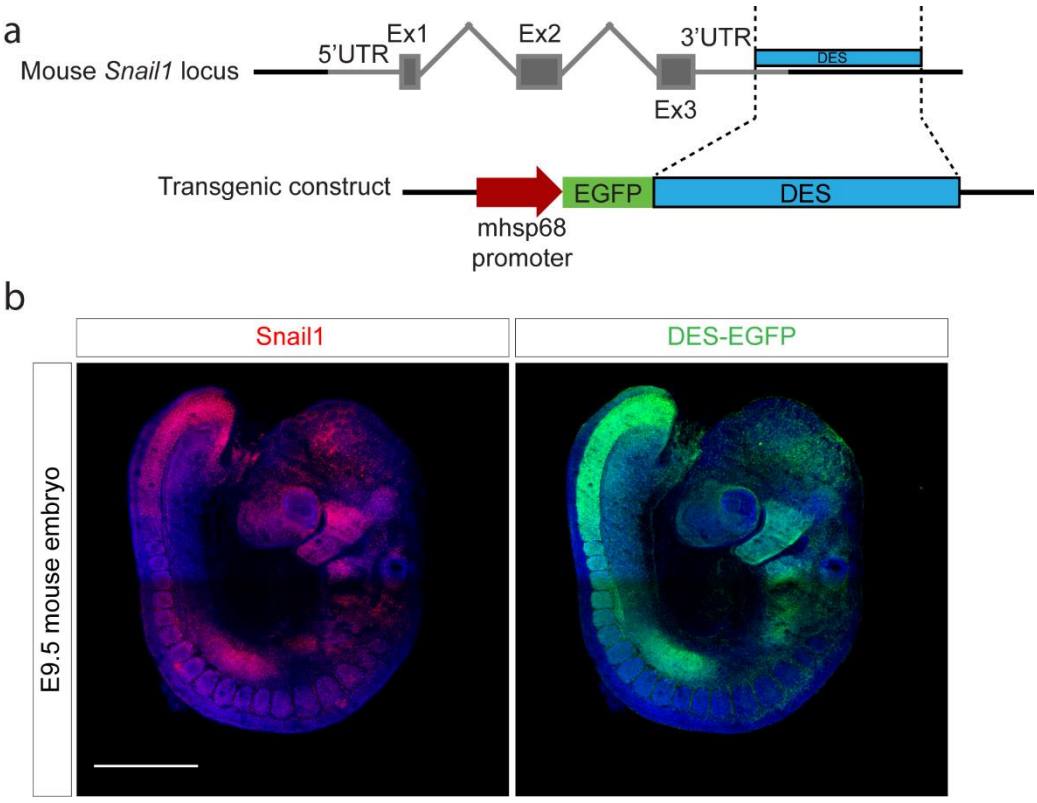
Supplementary Figure 5. Repression of Snail1 through Prrx1-induced miR-15 family in the mouse. (a) LacZ staining of E11.5 hetero or homozygous *Prrx1* mutant embryos, (129*S-Prrx1tm1Jfm)*[1]. Scale bar: 1mm. (b) Schematic representation of transverse sections of cranial regions of E11.5 embryos at the level of the nasal pit. (c) Signal intensity for Snail1 is measured by calculating the average signal in the nuclei of all cells in the selected regions of nasal pit of Snail1 IF in Figure 4i. (d) IF in the nasal process region of E11.5 WT and mutant embryos show increased levels as well as expansion of Snail1 expression domains. Scale bar: 250 μm for sections and 100 μm for insets (boxes 1 and 2). (e) Signal intensity for Snail1 is measured calculating the average signal in the nuclei of all cells in the selected regions of lateral nasal process in d. WT: wild type; Mut: mutant; NT: neural tube, NP: nasal pit; MNP: medial nasal process, LNP: lateral nasal process; AU: arbitrary units. Asterisks indicate significant p-value in t-test (* $p < 0.05$ and *** $p < 0.001$). Quantifications are performed for one section of WT or mutant embryos, and the increase and expansion is observed in n=2/2 mutant embryos compared to n=3/3 different E11.5 WT. Source data are provided as a Source Data file.



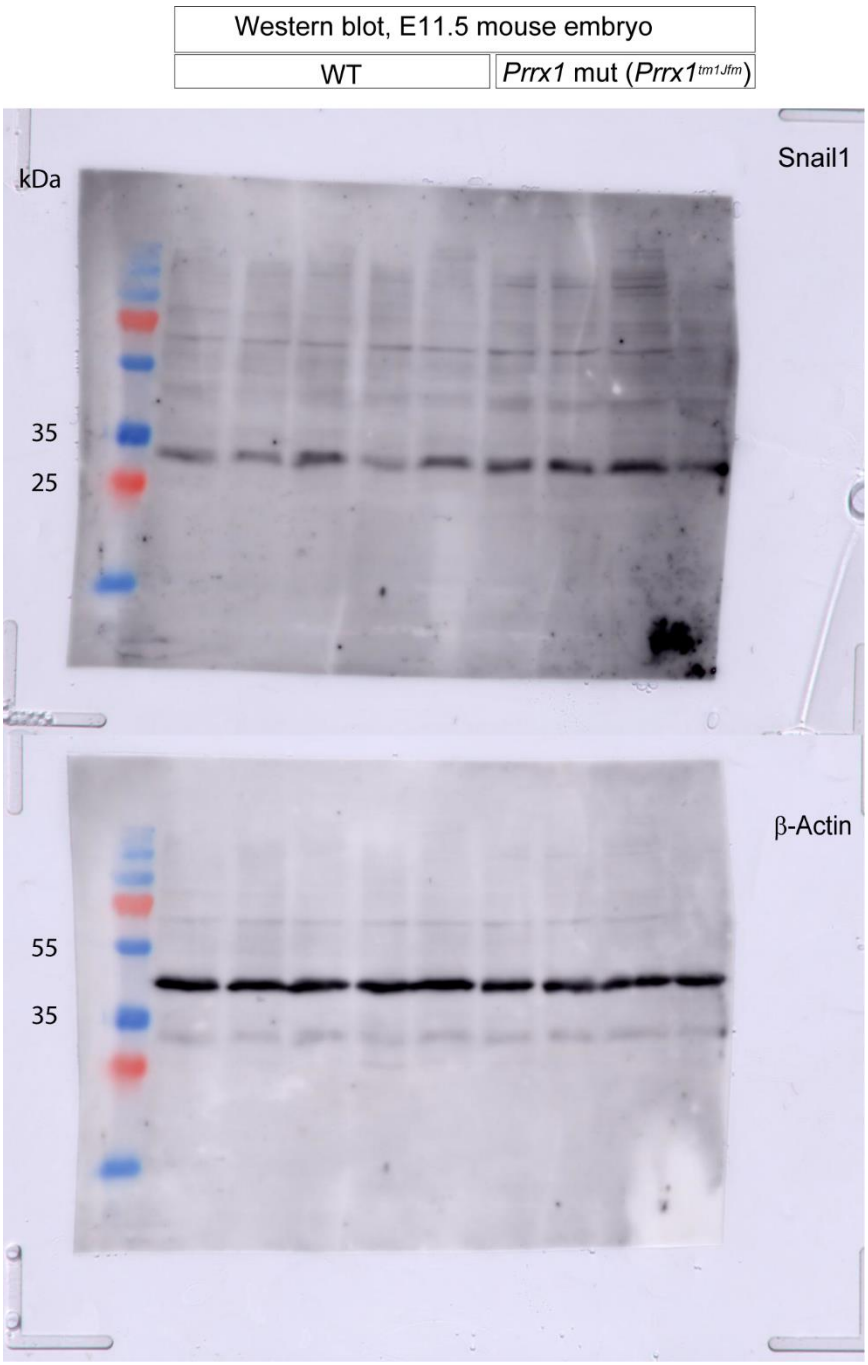
Supplementary Figure 6. Expression of miR-15 family members correlates with Prrx1 in survival of triple-negative breast cancer patients. Kaplan-Meier overall survival (OS) plots from basal/triple-negative breast cancer patients showing that high expression of *SNAIL1* correlates with low survival, while *PRRX1* high expression correlates with a better survival. Expression of *miR-15* family members follow a similar trend as that of *PRRX1*. Hazard ratio (HR) and logarithmic ranked p Value (longrank P) were analyzed to infer the significance of the differences. Numbers below each graph represent number of patients at risk in any given time (months), black for low expression and red for high expression of each gene/miRNA. The cut-off is automatically calculated based on the best performing threshold.



Supplementary Figure 7. Sequential upregulation of Snail1 and Prrx1 upon BMP treatment in zebrafish and chicken embryos. (a) Ventral view of chicken embryos showing SNAIL1 or PRRX1 protein expression by IF in control embryos and at different times after BMP-soaked bead implantation. Insets are shown at the levels of left and right somites in the vicinity of bead. Scale bar: 100 μm. (b) Signal intensity for SNAIL1 or PRRX1 protein expressions are measured by calculating the average signal in 3/4 somites in each side of each embryo, and the average ratio of left/right is represented for each time point. Numbers indicate the region of the inset in a, shown as dashed boxes. Total of n=6-8 somites from n=2 embryos were quantified for each condition at each time point. The value from each somite is considered as separate for the statistical test. Scale bar: 50 μm. Som: somite. Asterisks indicate significant p-value in One-way ANOVA with Bonferroni's multiple comparison test compared to the control (* $p < 0.05$ and *** $p < 0.001$). Source data are provided as a Source Data file.



Supplementary Figure 8. DES-GFP transgenic mouse recapitulated Snail1 endogenous expression. (a) Representation of the construct used to generate the transgenic DES-EGFP Snail1 line and the localization of the DES enhancer in the mouse *Snail1* locus. (b) Double IF for Snail1 and GFP in E9.5 transgenic mouse embryo showing endogenous Snail1 and EGFP expression corresponding to Snail1. Scale bar: 250 μm. UTR: untranslated region; Ex: exon; DES: downstream enhancer of Snail1; EGFP: enhanced green fluorescent protein; mhsp68 promoter: mouse heatshock protein 68 minimal promoter.



Supplementary Figure 9. Uncropped and unprocessed western blot for Snail1 and β-actin from E11.5 control and Prrx1 mutant embryos.

Supplementary Table 1, GEO datasets

GEO acc. #	Sample scRNA-seq	Total cell #	# of cells	# of Snail1 single + cells	# of Prrx1 single + cells	# of cells Snail1 high and Prrx1 low/-	# of cells Snail1 low/- and Prrx1 high	# of cells Snail1 high and Prrx1 high	Ref.
GSE87038	E9.5 mouse embryos	764	440	219 (49%)	87 (19%)	245 (55%)	143 (32%)	54 (12%)	2
GSE103322	18 head and neck cancer patients	5902	1718	480 (28%)	1005 (58%)	564 (32%)	1031 (60%)	123 (7%)	3
GSE75688	11 breast cancer patients	551	138	96 (69%)	22 (16%)	103 (74%)	28 (20%)	7 (5%)	4
GSM3067194	18hpf zebrafish embryos	6962	1620	snail1a: 582 (36%) snail1b: 98 (6%)	prrx1a: 376 (23%) prrx1b: 171 (10%)	snail1a: 787 (48%) snail1b: 116 (7%)	prrx1a: 627 (38%) prrx1b: 332 (20%)	prrx1a: 52 (1.6%) prrx1b: 2 (0.1%)	5

Supplementary Table2, miR-15 family nomenclature and accession numbers	
miR-15 family	MIPF0000006
Hsa-Mir-15-P1d (hsa-miR-424)	MI0001446
Hsa-Mir-15-P2d (hsa-miR-503)	MI0003188
Hsa-Mir-15-P1b (hsa-miR-15b)	MI0000438
Hsa-Mir-15-P2b (hsa-miR-16)	MI0000115
Hsa-Mir-15-P2c (hsa-miR-195)	MI0000489
Hsa-Mir-15-P1c (hsa-miR-497)	MI0003138

Supplementary Table 3. Primer sequences	
Human miR424-503-promoter-F (Cloning)	GACTGCTAGCTTCAATATCATTCTCACAAATACAAAATAG
Human miR424-503-promoter-R (Cloning)	GATCCTCGAGTGGAACAACGTAGTGGGTGA
Human miR424-503-F (Cloning)	GACTGGATCCGGCTTCCTTCAGTCATCCAG
Human miR424-503-R (Cloning)	GACTCTCGAGTCTACCTGAGCAGGGAAAGG
h-424-mut-F	GATCCCCCTTCATTGACTCCGAGGGGATACTATTATAATTCATGTTTTGAAGTGTTCTAAATGGTTC
h-424-mut-R	GAACCATTTAGAACACTTCAAAACATGAATTATAATAGTATCCCCTCGGAGTCAATGAAGGGGGATC
h-503-mut-F	GTGCCCGCGCTCAGCCGTGCCCTTATTATGGGAACAGTTCTGCAGTGA
h-503-mut-R	TCACTGCAGAACTGTTCCCATAATAAGGGCACGGCTGAGCGCGGGCAC
miR-15b-16-2-cloning-F	GACTGGATCCACTAAAGCTTGAAAGAGTGTTCTTCTGT
miR-15b-16-2-cloning-R	GACTCTCGAGATAAACAAAAGGGACAGATTATCAAAAG
QPCR HsPrrx1 FW	CTGATGCTTTTGTGCGAGAA
QPCR HsPrrx1 RV	ACTTGGCTCTTCGGTTCTGA
hSNAI1-F	GCTGCAGGACTCTAATCCAGA
hSNAI1-R	ATCTCCGGAGGTGGGATG
QPCR HsPrrx1 FW	CTGATGCTTTTGTGCGAGAA
QPCR HsPrrx1 RV	ACTTGGCTCTTCGGTTCTGA
premir424-F	TTGACTCCGAGGGGATACAG
premir424-R	GACCCCACCTTCTACCTTCC
premir503-F	GCGAGTCGAGGAGAGACG
premir503-R	GAACGGCAGTCCCAGACTTA
hsa-TBP-F	CGGCTGTTTAACTTCGCTTC
hsa-TBP-R	CACACGCCAAGAAACAGTGA
h-PRRX1-del1-F	TTATTCGTCTACCTTCAGAAGATTCTCTTCCACACTAATTCG
h-PRRX1-del1-F	CGAATTAGTGTGGAAGAGAATCTTCTGAAGGTAGACGAATAA
h-PRRX1-del2-F	TTCAGAGTGGGTTTTTTTTTTAATGCAATATTGTATTCAAACAAAAAGAGGG
h-PRRX1-del2-F	CCCTCTTTTTGTTTGAATACAATATTGCATTAAAAAAAAAACCCACTCTGAA
h-PRRX1-del3-F	GGAGAAAGTTGTGGAGTTTTACATGAACATTTGCTAAACATGTTTT
h-PRRX1-del3-F	AAAACATGTTTAGCAAATGTTCATGTAAAACTCCACAACTTTCTCC
miR-15b-16-2-cloning-F	GACTGGATCCACTAAAGCTTGAAAGAGTGTTCTTCTGT
miR-15b-16-2-cloning-R	GACTCTCGAGATAAACAAAAGGGACAGATTATCAAAAG
mmu-miR,15bmut-F	GCTGAGTCCTGTCTTTTGGAACCTTAAAGTACTGTTCGTCGACATCATGGTTTACATACTACAG

mmu-miR,15bmut-R	CTGTAGTATGTAAACCATGATGTCGACGAACAGTACTTTAAGGTTCCAAAAG ACAGGACTCAGC
mmu-miR,16mut-F	GTATTATGTTTGGATATCTGACATGCTTGTTCCACTCTTCGTCGACGTAAATA TTGGCGTAGTGAAATAAAT
mmu-miR,16mut-R	ATTTATTTCACTACGCCAATATTTACGTCGACGAAGAGTGGAACAAGCATGT CAGATATCCAAACATAATAC
m497-PxB-F	CTAAAAACAGGCCTTATTGACTACA
m497-PxB-R	GATGTGGCTTTACTTTGTGAAGATT
m497-up-F	AGGATATTAGAATTGGGGAGGTTAGT
m497-up-R	TGAGTCCAATACATTAGAAAGAGGAGT
m497-dw-F	ACTTGCTCGTACAGGTTGTATATGTTCT
m497-dw-R	CTTGGGAAGACTAATAGGAGTTAAAGTG
miR-195mut-F	GTTGCCCACACCCAACTCTCCTGGCTCTTCGTCGACAGAAATATTGGCATGG
miR-195mut-R	CCATGCCAATATTTCTGTCGACGAAGAGCCAGGAGAGTTGGGTGTGGGCAA C
miR-497mut-F	CAGTCCTGCCCCCGCCCCTCGTCGACACTGTGGTTTGTACGG
miR-497mut-F	CCGTACAAACCACAGTGTCGACGAGGGGCGGGGGCAGGACTG
m195-497-clon-F2	GACTGGATCCCTAAACTACTTTTGCTGGTTCCTGATT
m195-497-clon-R2	GACTCTCGAGGACTTCTGTGTGATGGACATTTTTATAC
SNAI1,3',MRE503-F	TTTGTATCCAGAGCTGTTTGGATACACGACGATTGAGCTACAGGACAAAGGC TGACAG
SNAI1,3',MRE503-F	CTGTCAGCCTTTGTCCTGTAGCTCAATCGTCGTGTATCCAAACAGCTCTGGAT ACAAA
h503P-del1-F	AAAAAAAAACATGTTTAGCAAATGTATCATGTAAAACTCCACAACTTTCTCC
h503P-del1-R	GGAGAAAGTTGTGGAGTTTTACATGATACATTTGCTAAACATGTTTTTTTTT
h503P-del2-F	CCCTCTTTTTGTTTGAATACAATATTAGCATTAAAAAAAAAACCCACTCTGA
h503P-del2-R	TCAGAGTGGGTTTTTTTTTTAATGCTAATATTGTATTCAAACAAAAAGAGGG
hPRRX1-P-F2	ATTAGTAAATGTGTGCGGTGTTTTTCTTG
hPRRX1-P-R3	CTTTGTAGGGAATACAAGAAAATTGGAG
hPRRX1-P-R4	ATTAGTAAATGTGTGCGGTGTTTTTCTTG
497P-PXB1-F	AGGAGGGAGTGACTTCCAAAA
497P-PXB1-R	AAATTTGGGGTCCTCAGATACC
497P-PXB2-F	CTACCCAGATGTCTTTGGAGGT
497P-PXB2-R	CCGAAACAAAATATGAGGGTGT
497P-dw-F	TTATTGAGATACGGGACACAGC
497P-dw-R	CTCCAGCCCCTCCTCTATTTA
SMC4P-dw-F	AGCAAGACCCCATCTCTACAAA
SMC4P-dw-R	AGTGCACACGTAAAAGGACTGA
SMC4i-PXB1-2-F	TTGGATATGGTGGGAATCTTGT
SMC4i-PXB1-2-R	TCAATTGAACATGCACACAAAC
SMC4P-PXB3-F	GAACTTATTTCTTCCTGTGGGGTA
SMC4P-PXB3-R	CCTGGAAAAGACTGGGTTACAG
PRRX1P-dw-F	GGACTTACAGACATTCCCTTGG

PRRX1P-dw-R	GGCCTGAGAAACAAATAGATGG
PRRX1P-PXB1-F	CTCCCTTTCTCTCTAACTCTGATGTTG
PRRX1P-PXB1-R	GCCAACATCAGAGTTAGAGAGAAAG
PRRX1P-PXB2-F	TTGATTGCCTGCATTCTTACTT
PRRX1P-PXB2-R	TCAATGAGTCCAAAATGTAAACC
PRXP-B1-F2	GTACTTTTCTCATTGCCCTAGTCCT
PRXP-B1-R2	GTAAGCATATTAAGGCTATTTTTGGTTC
PRXP-B1-F3	AAGAAGGAGATTGTGATGGAGAAAG
PRXP-B1-R3	TTTTCCTTCTCAGTTGGATCAAGAA
PRXP-up-F	CCTAGCACCTGAGAAACATACTGATAAC
PRXP-up-R	ACTTAGCCTATAAATCAGAGAATACTTGT
PRXP-PXB3-4-F	AGACTATAACAGAGCAGATCAATTACCA
PRXP-PXB3-4-R	TCCCATATTTTAATGTTTTTAGGAGTCT
PRXP-PXB5-6-F	TTTATGAGGTGAACGCATTATCAG
PRXP-PXB5-6-R	TCTTGCTTGTTTAAGTCTTTGTGG
SNAI-B-1-2-R	CAGTTAGTAGTCCAGGTGTCTTGC
PRRX1P-F01Xho	GACTCTCGAGCCTTTATTCTCTTTCTTAATCCTCTCAA
PRX-B1-R2Nco	GACTCCATGGGTAAGCATATTAAGGCTATTTTTGGTTC
PRXP-CpG-R	GACTCCATGGAGGTTTCACCCTTAAAGGACATTC
PRXP-COM-R	GACTCCATGGAGGTGACTGACGGAGAAGTTCTTT
m503-F2	GACTGCTAGCCCCTCCCAAATTATTCCAAA
SNAI1-PXB3-5-F	CAACAGTTCTTTACTTTGCTTGGTT
SNAI1-PXB3-5-R	ACAGGGAAGGATTAAAACACCTAAG
SNAI1-PXB6-F	GACTTTGGCTTTTACTCTGAGACAG
SNAI1-PXB6-R	ACAGAGGCAGTAAGCAGTCATTAAG
SNAI1-PXB7-F	TCCCTTTGCATTGTAATTATCTGTT
SNAI1-PXB7-R	GTTCCTTCCCTTATCCAGTGTTTAC
SNAI1-dw-F	GCCCAAATTGTCAGTTTCATAAATA
SNAI1-dw-R	ACAAAGATGTAAACCAAGATCTCCA
EGFP_fwd	CCGACCTCTCTCCCCAGGGGGATCCGTGAGCAAGGGCGAGGAG
EGFP_rev	TAGTAGCTCCGCTTCCCTTGTACAGCTCGTCCATGC
PuroR_fwd	GCTGTACAAGGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGG AGACGTGGAGGAGAACCCTGGACCTATGACCGAGTACAAGCCCACGG
PuroR_rev	GTCATTGGTCTTAAAGGTACCTCAGGCACCGGGCTTGCG
hUbC_promoter-F0	TCTTTCCAGAGAGCGGAACA
YFP-BamHI-F	TTAAGGATCCGCTCGTTTAGTGAACCGTCAG
YFP-KpnI+stop-R	TTAAGGTACCTACTTGTACAGCTCGTCCATGC
hPGK promoter-F	GTAGTGTGGGCCCTGTTCCT
3'LTR-R	TCGTTGGGAGTGAATTAGCC
hPRX1-enh-F	GACTGCTAGCGGTTTACTACTGGGGCAAGAGTT
hPRX1-enh-R	GACTCTCGAGAATATGGCAGAGAGGGTAAAAGG
hPRX1-enh-del-F	caaatatatatggtaggtgccacattatttgagagatgggtaag

hPRX1-enh-del-R	cttacccatctctcaaataatgtggcacctaccatatatatttg
hPRX-P-SN-BS4-del-F	cataatcagagtgtggatccaatctaattgctaatagaaaaaaaggaa
hPRX-P-SN-BS4-del-R	ttccttttttctattagcaattagattggatccacactctgattatg

References

1. Lu, M.F. *et al.* prx-1 functions cooperatively with another paired-related homeobox gene, prx-2, to maintain cell fates within the craniofacial mesenchyme. *Development* **126**, 495-504 (1999).
2. Dong, J. *et al.* Single-cell RNA-seq analysis unveils a prevalent epithelial/mesenchymal hybrid state during mouse organogenesis. *Genome biology* **19**, 31 (2018).
3. Puram, S.V. *et al.* Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* **171**, 1611-1624 e24 (2017).
4. Chung, W. *et al.* Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nature communications* **8**, 15081 (2017).
5. Wagner, D.E. *et al.* Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* (2018).