

N° d'ordre :XXXXX

THESE DE L'UNIVERSITE DE LYON

Délivrée par

UNIVERSITE CLAUDE BERNARD LYON 1

ECOLE DOCTORALE BMIC (BIOLOGIE MOLECULAIRE INTEGREE ET COGNITIVE)

DIPLOME DE DOCTORAT

Soutenue publiquement le 18 décembre 2008

Par Irène AKSOY

*Identification et caractérisation fonctionnelle
des gènes cibles de la voie LIF/STAT3 impliqués dans le
contrôle de l'autorenouvellement des cellules souches
embryonnaires*

Directeur de thèse : Pierre SAVATIER

MEMBRES DU JURY

Mme DUSANTER-FOURT Isabelle : Rapporteur

Mme COULOMBEL Laure : Rapporteur

M SAMARUT Jacques : Président du jury

M VIVILLE Stéphane : Examinateur

M BOURILLOT Pierre-Yves : Examinateur

M SAVATIER Pierre : Directeur de thèse

<u>TABLE DES MATIERES.</u>	1
<u>Liste des abréviations.</u>	6
<u>Liste des figures.</u>	8
<u>I – INTRODUCTION.</u>	9
<u>1 - Origine et propriétés des cellules souches embryonnaires.</u>	11
<u>1.1. Définition.</u>	11
<u>1.2. Dérivation.</u>	11
<u>1.3. Propriétés des cellules ES.</u>	13
<u>1.3.1. Pluripotence.</u>	13
<u>1.3.2. Autorenouvellement.</u>	18
<u>1.3.3. Tumorigénérité.</u>	18
<u>1.4. Distinction entre les cellules ES et les cellules souches de l'embryon.</u>	18
<u>2 - La voie LIF/gp130/STAT3 : un rôle clef dans l'autorenouvellement des cellules ES.</u>	19
<u>2.1. La voie de signalisation LIF/gp130.</u>	19
<u>2.1.1. Le LIF.</u>	19
<u>2.1.2. Les récepteurs aux cytokines de la famille du LIF.</u>	20
<u>2.1.3. Les protéines kinases cytoplasmiques.</u>	23
<u>2.2. Les différentes voies de signalisation activées par le LIF dans les cellules ES.</u>	25
<u>2.2.1. La voie LIF/gp130/STAT3.</u>	25
<u>2.2.2 La voie SHP2/ERK.</u>	35
<u>2.2.3 La voie PI3Kinase.</u>	39
<u>2.2.4 La voie de signalisation Src.</u>	41

<u>3 - La voie BMP4/Smad/Id2.</u>	42
<u>3.1. BMP4 inhibe la différenciation neurale des cellules ES.</u>	44
<u>3.2. Le facteur BMP4 inhibe la différenciation en activant l'expression du gène Id2.</u>	44
<u>4 - Le facteur de transcription Oct4.</u>	45
<u>4.1. Profil d'expression du facteur Oct4 au cours du développement embryonnaire.</u>	45
<u>4.2. Rôle du facteur Oct4 dans le contrôle de l'autorenouvellement des cellules ES.</u>	47
<u>4.3. Le co-facteur de Oct4 : Sox2.</u>	47
<u>5 - Le facteur de transcription Nanog.</u>	48
<u>5.1. Identification du gène Nanog.</u>	48
<u>5.2. Expression du gène Nanog et rôle dans l'autorenouvellement des cellules souches embryonnaires <i>in vivo</i> et <i>in vitro</i>.</u>	48
<u>5.3. Relations fonctionnelles entre les facteurs Nanog, STAT3, Oct4, Sox2 et Sall4.</u>	49
<u>5.3.1 Nanog et STAT3.</u>	49
<u>5.3.2 Nanog, Oct4 et Sox2.</u>	49
<u>5.3.3 Nanog et Sall4.</u>	52
<u>6 - Les sérine/thréonine kinases Pim1, Pim2 et Pim3.</u>	54
<u>6.1. Identification.</u>	54
<u>6.2. Profil d'expression des kinases Pim.</u>	54
<u>6.3. Les kinases Pim répondent aux voies de signalisation activées par les cytokines.</u>	56
<u>6.3.1 L'expression des kinases Pim est activée par les cytokines.</u>	56
<u>6.3.2 La régulation de Pim1 par STAT3.</u>	56
<u>6.4. La fonction des kinases Pim.</u>	58
<u>6.4.1 La régulation du cycle cellulaire.</u>	58
<u>6.4.2 L'apoptose.</u>	59

<u>7 - Les Krüppel-Like Factors ou Klf.</u>	60
<u>7.1. Structure des facteurs de transcription Klf.</u>	62
<u>7.2. Identification des Krüppel-like facteurs Klf4 et Klf5.</u>	62
<u>7.3. In vivo, Klf4 et Klf5 présentent des profils d'expression antagonistes.</u>	64
<u>7.3.1 Profil d'expression de Klf4 dans l'embryon.</u>	64
<u>7.3.2 L'expression de Klf5 varie au cours du développement embryonnaire.</u>	65
<u>7.3.3 Klf5 présente un profil d'expression spécifique dans les tissus adultes.</u>	65
<u>7.4. Activité transcriptionnelle des facteurs de transcription Klf4 et Klf5.</u>	66
<u>7.4.1 Klf4 et Klf5 régulent tous les deux l'expression de Klf4.</u>	66
<u>7.4.2 Klf4 et Klf5 ont des effets antagonistes sur l'expression des gènes cibles.</u>	66
<u>7.5. Les rôles opposés de Klf4 et Klf5 dans la régulation de la prolifération cellulaire.</u>	67
<u>7.5.1 Klf4 : son rôle dans l'inhibition de la prolifération.</u>	67
<u>7.5.2. Contrairement à Klf4, Klf5 est un activateur la prolifération des cellules.</u>	69
<u>7.6. Klf4 et Klf5 dans les cellules ES.</u>	70
<u>7.6.1 Le rôle de Klf4 dans l'autorenouvellement des cellules ES.</u>	70
<u>7.6.2 Klf2, Klf4 et Klf5 est nécessaire à l'autorenouvellement des cellules ES.</u>	71
<u>7.6.3 Les facteurs de transcription KLFs et leur rôle dans la reprogrammation.</u>	71
<u>II – RESULTATS</u>	73
<u>1-Introduction.</u>	74
<u>2- Article n°1.</u>	80
<u>3- Article n°2.</u>	92

<i>1 - Identification des gènes cibles de la voie LIF/gp130/STAT3.</i>	140
<i>1.1. Validation du modèle cellulaire.</i>	140
<i>1.2. Exhaustivité de l'analyse.</i>	142
<i>1.3. Identification des gènes cibles du LIF, indépendants de STAT3.</i>	144
<i>2 - Les gènes cibles de STAT3 inhibent la différenciation des cellules ES.</i>	144
<i>2.1. Caractérisation préliminaire des gènes.</i>	144
<i>2.2. La kinase SGK (Serum and Glucocorticoid inducible Kinase).</i>	145
<i>2.3. Les sérine/thréonine kinases Pim1 et Pim3.</i>	147
<i>2.3.1. Régulation de l'expression des kinases Pim1 et Pim3 par STAT3 et Nanog</i>	147
<i>2.3.2. Les kinases Pim1 et Pim3 inhibent la différenciation des cellules ES.</i>	148
<i>2.3.3. Les kinases Pim1 et Pim3 régulent la prolifération et l'apoptose dans les cellules ES.</i>	148
<i>2.4. Klf4 et Klf5, en coopération avec Sall4 : le lien entre la voie LIF/STAT3 et le réseau de régulation transcriptionnelle Nanog/Oct4/Sox2.</i>	149
<i>3 - Les gènes cibles de STAT3 inhibent la différenciation des cellules ES de souris vers le lignage endodermique et/ou mésodermique.</i>	151
<i>3.1. Les gènes cibles de STAT3 constituent trois familles distinctes inhibant la différenciation vers le mésoderme et/ou vers l'endoderme.</i>	151
<i>3.2. Klf4 et Klf5 exercent des effets opposés sur la différenciation en mésoderme et en endoderme.</i>	153
<i>3.3. Klf4 et Klf5 inhibent l'expression de régulateurs de la différenciation vers le mésoderme et l'endoderme.</i>	154
<i>3.4. L'inactivation de Klf4 induit-elle la différenciation en endoderme extra-embryonnaire ou en endoderme définitif ?</i>	156
<i>IV – CONCLUSION GENERALE</i>	158

LISTE DES ABBREVIATIONS

μm	micromètre
nM	nanomolaire
4'OH-T	4'-Hydroxy-Tamoxifène
ADNc	ADN complémentaire
BMP	Bone Morphogenetic Protein
BrdU	BromodésoxyUridine
CDK	Cyclin Dependant Kinase
CNTF	Ciliary Neurotrophic Factor
CT	Cardiotrophin
EB	Embryoïd Body
EC	Embryonic carcinomal
ER	Estrogen Receptor
ERK	Extracellular Regulated MAP Kinase
ES	Embryonic Stem
GCSF	Granulocyte Colony Stimulating Factor
Grb2	Growth factor receptor binding protein 2
GSK3	Glycogen Synthase Kinase 3
Hck	Hemopoietic cell kinase
HLH	Helix Loop Helix
ICM	Inner Cell Mass
Id	Inhibitor of differentiation
IL	Interleukin
JAK	Janus Kinase

Klf	Krüppel-like factor
LBD	Ligand Binding Domain
LIF	Leukemia Inhibitory Factor
MEF	Mouse Embryonic Fibroblast
MEK	Mitogen activated protein ERK Kinase
OSM	Oncostatin-M
PDK	Phosphoinositide-Dependent Kinase
PI3K	Phosphatidyl Inositol 3 Kinase
Pim	Proviral integration
PKB	Protein Kinase B
PTEN	Phosphatase and Tensin Homolog
R	Receptor
SH2	Srv homology domain 2
SHP	SH2 containing protein tyrosine Phosphatase
sILR	soluble Interleukin Receptor
siRNA	short interfering RNA
shRNA	short hairpin RNA
Smad	MAD homologue
Socs	Suppressor of cytokine signaling
Src	Rous sarcoma oncogene
STAT	Signal Transducer and Activator of Transcription
TAD	Transcription Activating Domain
TGF	Transforming Growth Factor

LISTE DES FIGURES

Figure 1 : Propriétés des cellules ES	10
Figure 2 : Etablissement de lignées de cellules ES de souris	12
Figure 3 : Stades précoce	s du développement embryonnaire de la souris
et origine des lignages embryonnaires	14
Figure 4 : Différenciation <i>in vitro</i> des cellules ES murines	16
Figure 5 : Les complexes de récepteur aux cytokines apparentées au LIF	21
Figure 6 : Structure de la protéine STAT3	24
Figure 7 : La voie de signalisation LIF/gp130/STAT3	26
Figure 8 : Représentation schématique des récepteurs utilisés dans les expériences de substitution des tyrosines de la partie intracytoplasmique de gp130	28
Figure 9 : La voie de signalisation SHP2/ERK	34
Figure 10 : La voie PI3 Kinase	38
Figure 11 : La voie de signalisation BMP4/Smad/Id2	43
Figure 12 : Les facteurs de transcription intervenant dans le contrôle de l'autorenouvellement des cellules ES	46
Figure 13 : Réseau de régulation transcriptionnelle mis en place par Nanog, Oct4 et Sox2 dans les cellules souches embryonnaires	50
Figure 14 : Régulation de la sérine/thréonine kinase Pim1	55
Figure 15 : Le rôle des kinases Pim dans la régulation du cycle cellulaire	57
Figure 16 : Le rôle des membres de la famille des Krüppel-like factors	61
Figure 17 : Expression de Klf4 et Klf5 dans l'intestin et l'épiderme	63
Figure 18 : Rôles de Klf4 et Klf5 dans la régulation du cycle cellulaire	68
Figure 19 : Test d'autorenouvellement	143
Figure 20 : Modèle de régulation de la kinase SGK par les voies LIF/STAT3 et LIF/PI3K	146
Figure 21 : Le LIF alimente le réseau de pluripotence Nanog/Oct4/Sox2	150
Figure 22 : Profil d'expression de Oct4, Nanog, Klf4 et Klf5 au cours de la différenciation	150
Figure 23 : Les trois familles de gènes cibles de STAT3	152
Figure 24 : Modèle d'action de Klf4 pour inhiber la différenciation des cellules ES en endoderme	155

INTRODUCTION

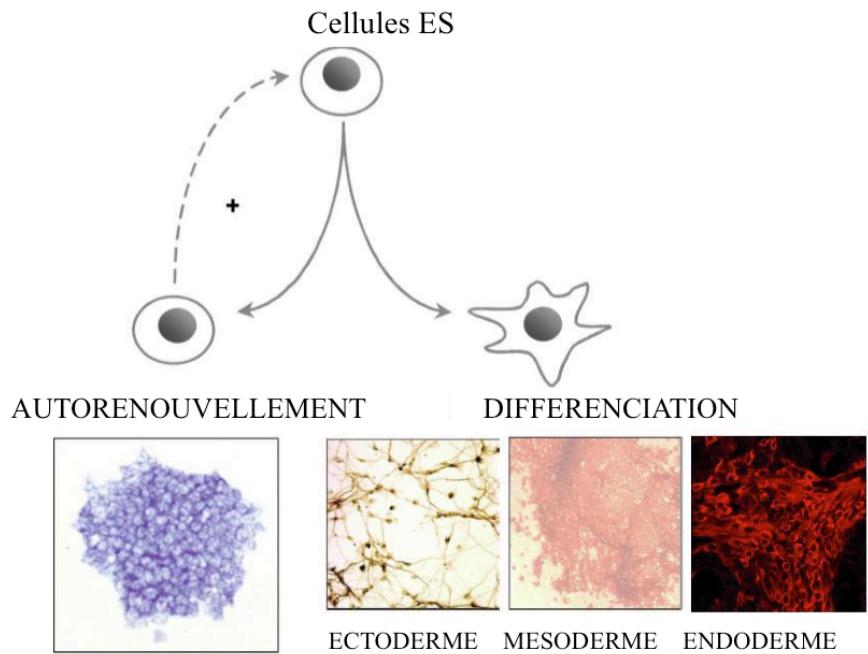


Figure 1 : Propriétés des cellules ES

Les cellules souches embryonnaires (ES) possèdent deux propriétés fondamentales : (i) elles sont pluripotentes ; elles sont capables de se différencier dans les trois lignages à l'origine des tissus composant l'organisme adulte, l'ectoderme, le mésoderme et l'endoderme ; (ii) elles sont immortelles ; elles se multiplient de façon illimitée tout en conservant leur propriété de pluripotence, on parle d'autorenouvellement permanent.

1. Origine et propriétés des cellules souches embryonnaires.

1.1. Définition.

Les cellules souches embryonnaires (cellules ES) sont obtenues à partir des cellules de la Masse Cellulaire Interne ou « Inner Cell Mass » (ICM) du blastocyste (Evans and Kaufman, 1981) (Martin, 1981). L'ICM est à l'origine de l'ensemble des tissus constitutifs de l'organisme adulte y compris les cellules germinales. Les cellules ES possèdent deux propriétés fondamentales (Figure 1) :

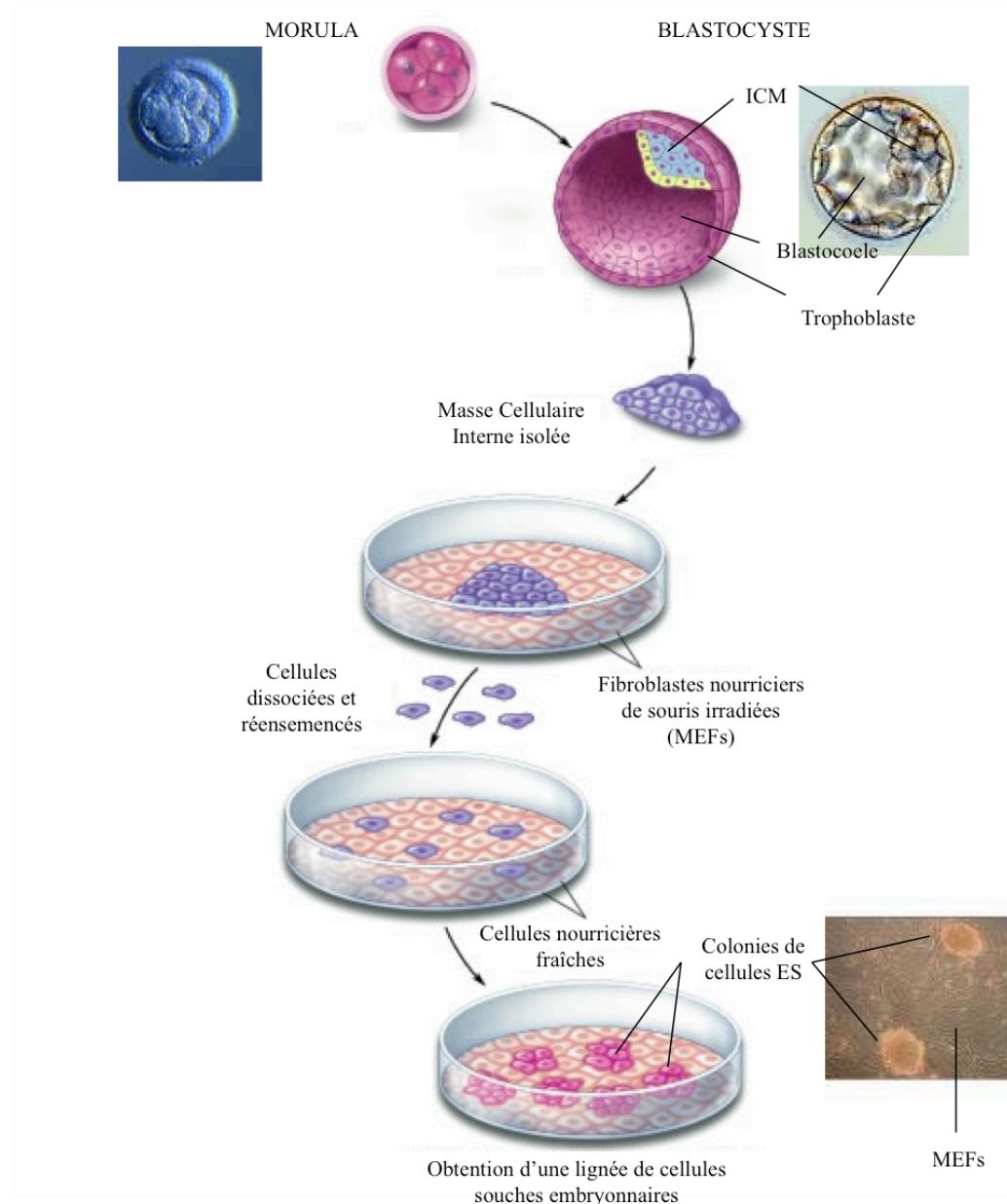
- la pluripotence : elles sont capables de se différencier *in vitro* dans les trois lignages à l'origine des tissus composant l'organisme adulte, l'ectoderme, le mésoderme et l'endoderme.

- l'autorenouvellement permanent : elles sont immortelles et peuvent se multiplier de façon illimitée tout en conservant leur propriété de pluripotence.

1.2. Déivation.

La recherche sur les cellules souches embryonnaires a débuté dans les années 1970 lorsque les premières lignées de cellules souches embryonnaires carcinomales ou cellules EC furent dérivées à partir de tumeurs testiculaires de souris (Kahan and Ephrussi, 1970) (Jakob et al., 1973). Ces tumeurs appelées tératocarcinomes sont constituées de types cellulaires différenciés très variés, associés à des cellules souches pluripotentes (les cellules EC). Ces cellules sont capables de s'autorenouveler mais également de se différencier dans un très grand nombre de types cellulaires.

De nombreuses études ont montré que les cellules EC présentaient de nombreux points communs avec les cellules pluripotentes de l'ICM. En effet, des embryons de souris à un stade précoce de développement implantés dans des souris adultes sont capables, comme les cellules EC, de former des tératocarcinomes (Solter et al., 1970) (Stevens, 1970). En fait, seuls les embryons au stade prégastrulation contenant toujours l'épiblaste peuvent induire la formation de tératocarcinomes après implantation dans des souris. Ce sont donc les cellules de l'épiblaste qui sont à l'origine de la formation des tératocarcinomes dans l'embryon (Diwan and Stevens, 1976).



Adapté de « Report on Stem Cells », 2001, NIH, USA
<http://stemcells.nih.gov/info/scireport>

Figure 2 : Etablissement de lignées de cellules ES de souris

Les cellules souches embryonnaires (ES) de souris sont isolées à partir de la masse cellulaire interne (ICM) d'un embryon au stade blastocyste. L'ICM est ensuite mis en culture sur une matrice nourricière constituée de fibroblastes embryonnaires de souris, les MEFs.

Les cellules EC, comme les cellules de l'ICM, après implantation dans un blastocyste, contribuent à la formation de tous les tissus qui constituent l'embryon (Brinster, 1974). Une fois intégrées dans le blastocyste elles sont donc capables de répondre aux signaux envoyés par le micro-environnement embryonnaire et se comportent exactement comme les cellules de l'épiblaste. Quant aux cellules de l'épiblaste, quand elles sont introduites dans un environnement ectopique, elles prolifèrent activement et peuvent être à l'origine de la formation de tératocarcinomes. Ainsi, prolifération active et tumorigénicité sont des propriétés intrinsèques des cellules pluripotentes de l'embryon précoce.

La caractérisation des cellules EC *in vitro*, et notamment de leurs conditions de culture, a abouti à l'isolement en 1981 des premières lignées de cellules ES de souris. En effet, la co-culture des cellules pluripotentes EC avec des fibroblastes a permis le maintien de leur pluripotence (Martin and Evans, 1975). Au début des années 1980, deux équipes ont réussi à obtenir une lignée de cellules ES de souris en isolant l'ICM d'un embryon et en la mettant en culture sur une matrice nourricière constituée de fibroblastes embryonnaires de souris, les MEFs (Mouse Embryonic Fibroblasts) (Evans and Kaufman, 1981) (Martin, 1981) (Figure 2). Les cellules ES ont de nombreux points communs avec les cellules EC de par leur morphologie, leur prolifération, l'expression de marqueurs spécifiques et leur capacité à former des tératocarcinomes. Elles sont capables de former des tumeurs constituées de cellules de type endodermique, mésodermique et ectodermique (Smith, 2001). Depuis, de nombreuses lignées de cellules ES ont pu être isolées à partir d'autres espèces en utilisant la même technique que pour les embryons de souris (Doetschman et al., 1988) (Graves and Moreadith, 1993) (Iannaccone et al., 1994) (Pain et al., 1996) (Pau and Wolf, 2004) (Suemori et al., 2001) (Thomson et al., 1996). En 1998, Thomson et ses collègues ont isolé la première lignée de cellules ES humaines, permettant ainsi l'espoir de les utiliser en thérapie cellulaire (Thomson et al., 1998).

1.3. Propriétés des cellules ES.

1.3.1. La pluripotence.

1.3.1.1. In vivo.

La pluripotence des cellules ES a été démontrée pour la première fois *in vivo* chez la souris, en réimplantant des cellules ES dans la cavité d'un blastocyste hôte (Bradley et al., 1984). Les cellules ES ainsi ré-implantées se sont intégrées dans l'ICM de l'embryon

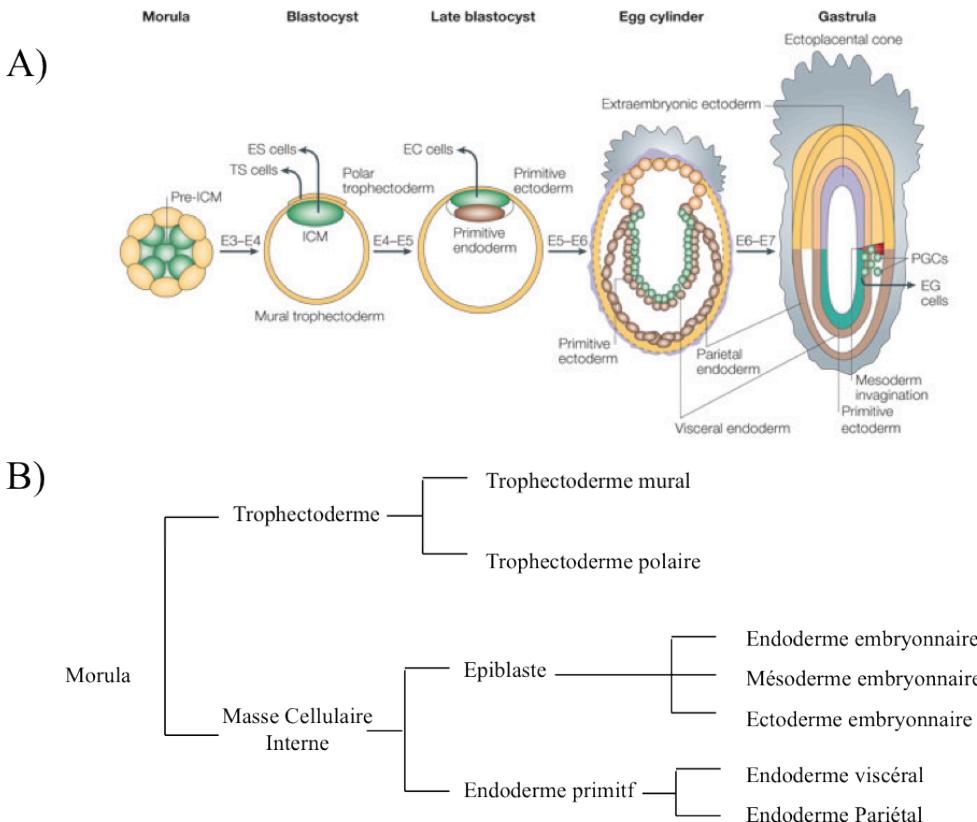


Figure 3 : Stades précoce du développement embryonnaire de la souris et origine des lignages embryonnaires

A) Développement précoce de l'embryon de souris. Les cellules pluripotentes de l'embryon sont colorées en vert. Les dérivés de l'endoderme primitif sont en marron, ceux du trophectoderme polaire en gris. Le trophectoderme et les dérivés du trophectoderme mural sont en jaune.

De gauche à droite :

Morula (E 2.5) : à ce stade, l'embryon est composé de cellules très compactes au centre qui forment le pre-ICM (Inner Cell Mass).

Blastocyste (E3-E4) : Au stade blastocyste l'embryon est composé d'une cavité (le blastocoèle) et le pre-ICM qui devient l'ICM.

Blastocyste tardif (E4-E5) : L'embryon continue à se développer et l'ICM se différencie en ectoderme et en endoderme primitifs.

« Egg cylinder » (E5-E6) : A ce stade, l'endoderme primitif donne naissance à l'endoderme pariétal et viscéral. L'endoderme viscéral reste en contact avec l'épiblaste alors que l'endoderme pariétal va se former à la surface du trophectoderme mural.

Gastrula (E6-E7) : L'embryon commence la gastrulation. A ce stade, les feuillets se mettent en place avec le feuillet mésodermique qui se situe entre les feuillets ectodermique et endodermique. Les cellules germinales primordiales se forment également durant ce stade.

E : jour embryonnaire

B) Schéma récapitulatif de l'origine des différents lignages embryonnaires chez la souris

receveur et se sont développées normalement. Elles ont contribué à la formation de tous les tissus de la souris chimère née de cette expérience et ont également colonisé la lignée germinale. Ces résultats confirment que les cellules ES, dérivées de l'ICM d'embryons murins et cultivées *in vitro*, sont bien pluripotentes et qu'elles peuvent répondre *in vivo* de façon appropriée aux facteurs de régulation du développement embryonnaire, aussi bien pour la prolifération, la différenciation et la migration. Elles conservent ainsi toutes les caractéristiques des cellules de l'ICM dont elles sont issues.

Parce que les cellules ES ont la capacité, une fois réimplantées dans l'embryon, de coloniser la lignée germinale, et de permettre ainsi la transmission de leur génotype à la descendance de la souris chimère, elles sont utilisées pour introduire des modifications génétiques chez la souris et pour en étudier les effets *in vivo*. Pour cela, les cellules ES sont manipulées génétiquement *in vitro* afin d'invalider ou de remplacer l'un des deux allèles d'un gène cible grâce à la technique de recombinaison homologue (techniques de « knock-out » et de « knock-in ») (Thomas and Capecchi, 1987). Les cellules ES ainsi modifiées génétiquement sont ensuite réimplantées dans des embryons receveurs. Du fait de la transmission de l'allèle modifié dans la lignée germinale des souris chimères (Thompson et al., 1989) il est ainsi possible d'obtenir, parmi la descendance, des souris transgéniques hétérozygotes. L'analyse de ces animaux transgéniques hétérozygotes ou homozygotes (après croisement des hétérozygotes entre eux) a permis de déterminer au cours des deux dernières décennies la fonction *in vivo* de nombreux gènes au cours du développement embryonnaire et aussi chez l'adulte.

1.3.1.2. In vitro.

Au cours du développement embryonnaire, l'ICM du blastocyste donne naissance d'une part à l'endoderme primitif, à l'origine de la formation de l'endoderme extra-embryonnaire (pariétal et viscéral), et d'autre part à l'ectoderme primitif ou épiblaste (Figure 3). Au moment de la gastrulation, l'ectoderme primitif évolue en trois couches primaires ou feuillets embryonnaires que sont l'endoderme, le mésoderme et l'ectoderme définitif. L'émergence *in vivo* de ces trois feuillets embryonnaires marque le début de l'organogénèse.

En 1985, Doetschman et ses collègues décrivaient pour la première fois un modèle *in vitro* de ce processus de développement embryonnaire, basé sur la différenciation des cellules ES de souris (Doetschman et al., 1985). Dans des conditions particulières de culture, il est possible d'amorcer la différenciation *in vitro* des cellules ES. Cette différenciation requiert

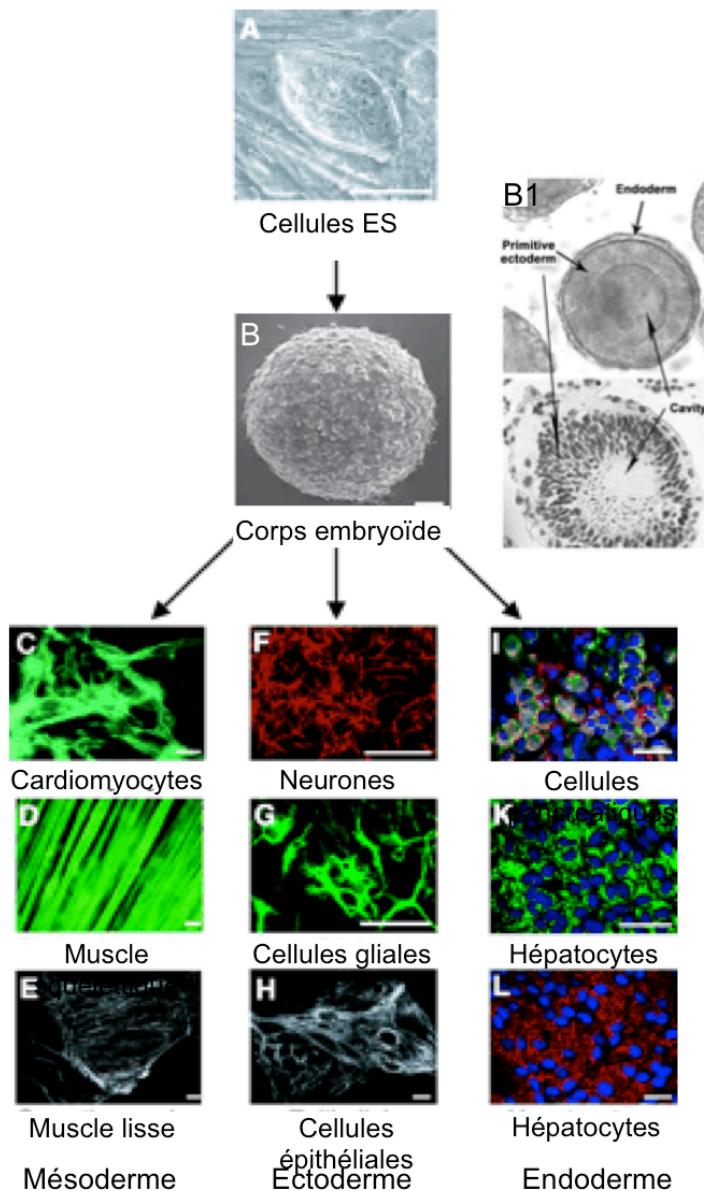


Figure 4 : Différenciation *in vitro* des cellules ES murines

Les cellules souches embryonnaires (ES) de souris indifférenciées (A), sont dissociées et cultivées en conditions non adhérentes, afin de former des structures tridimensionnelles, les corps embryoïdes ou EBs (B). Dans les EBs, les cellules se différencient en cellules de l'endoderme qui forme une couche périphérique, et en ectoderme primitif qui forme une couche interne et entoure une cavité centrale (B1, photographie et coupe en paraffine d'un EB au 6ème jour, d'après Mfopou et al., 2005). Une fois les EBs mis en culture en conditions adhérentes, les cellules émergeantes poursuivent leur différenciation en cellules dérivant des trois feuillets embryonnaires : endoderme, mésoderme et ectoderme (C à L). Echelles : 0,5µm (H), 20µm (C,D,E), 30µm (K,L), 50 µm (B, G), 100µm (A,F) et 250µm (B1). Adapté de Wobus and Boheler, 2005.

d'une part la suppression des agents permettant leur maintien à l'état indifférencié (le LIF ou Leukemia Inhibitory Factor et/ou les MEFs (cellules nourricières sécrétant du LIF)), et d'autre part leur culture sur des supports non adhérents. Dans ces conditions, les cellules ES en suspension s'organisent spontanément en agrégats ou structures tridimensionnelles appelés corps embryoïdes (EB pour Embryoïd Body) (Figure 4). Les cellules ES dans les corps embryoïdes suivent un programme de différenciation mimant la gastrulation et les étapes précoce de l'organogénèse. Durant les deux à quatre premiers jours, une couche de cellules de type endodermique se forme à la périphérie du corps embryoïde, défini à ce stade comme un corps embryoïde « simple ». A partir du quatrième jour de culture, une cavité centrale apparaît, aboutissant à une nouvelle structure nommée corps embryoïde « cystique ». Le développement du corps embryoïde cystique rappelle celui de l'embryon post-implantatoire, du fait de la formation d'une double couche de cellules entourant la cavité, constituée de cellules ectodermiques sur la face interne et de cellules endodermiques sur la face externe. La différenciation du corps embryoïde se poursuit par l'apparition de cellules aux caractéristiques mésodermiques. Ainsi, la différenciation des cellules ES *via* le modèle des corps embryoïdes démontre que les processus précoce du développement embryonnaire peuvent être reproduits *in vitro*. La technique de formation de corps embryoïdes a depuis été largement étudiée chez la souris. Plusieurs méthodes ont été développées telles que la technique des « gouttes » (Wobus et al., 1991) qui est la plus couramment utilisée et qui est connue pour être la plus fidèle au développement *in vivo* ; la culture en suspension liquide dans des boîtes de pétri non adhérente (Doetschman et al., 1985) ; ou l'utilisation d'un milieu semi-solide de méthylcellulose (Wiles and Keller, 1991).

La pluripotence *in vitro* ne se limite pas à l'émergence de cellules endodermiques, mésodermiques et ectodermiques dans les corps embryoïdes. Un nombre considérable d'études ont été menées chez la souris pour démontrer le potentiel de différenciation *in vitro* des cellules ES en cellules hautement spécialisées. De manière générale, trois grands types d'approches expérimentales sont utilisés pour la différenciation *in vitro* des cellules ES. La première consiste en la formation de corps embryoïdes puis en leur mise en culture en conditions adhérentes. Dans la seconde méthode, les cellules ES sont cultivées directement sur des cellules stromales et leur différenciation se fait directement en contact avec celles-ci (Nakano et al., 1994). Enfin, le troisième protocole implique la culture des cellules ES en monocouche sur une matrice extracellulaire de protéines (Nishikawa et al., 1998). Grâce à l'utilisation de ces trois méthodes et de milieux de culture adaptés, il est ainsi possible d'obtenir *in vitro* des cellules différenciées hautement spécialisées.

1.3.2. Autorenouvellement.

Les cellules ES se distinguent des autres cellules par une capacité de prolifération *in vitro* très importante. Elles se divisent activement de façon symétrique (Zwaka and Thomson, 2005) et sans se différencier. Elles peuvent donc être maintenues en culture *in vitro* indéfiniment et ceci tout en conservant leur propriété de pluripotence : on parle d'autorenouvellement permanent.

Les cellules ES s'autorenouvellent indéfiniment sans altération de leur phénotype ni de leur caryotype euploïde, et se distinguent ainsi des cellules tumorales d'origine cancéreuse.

1.3.3. Tumorigénérité.

Les cellules ES présentent des propriétés de prolifération qui leur sont propres (Burdon et al., 2002). Elles sont dérivées sans nécessiter d'aucun agent immortalisant, elles n'entrent pas en quiescence, elles ne deviennent pas sénescantes et peuvent proliférer sans limites apparentes. Elles peuvent également se multiplier en absence de sérum et ne sont pas sujettes à l'inhibition de contact. Les cellules transformées présentent également toutes ces caractéristiques, les cellules ES sont donc tumorigènes. En effet, après injection dans des souris immuno-déficientes elles forment des tératocarcinomes composés de cellules des trois feuillets embryonnaires. Ceci est un phénomène remarquable des cellules ES du fait qu'elles sont génotypiquement normales et qu'elles peuvent générer des adultes sains après leur implantation dans des blastocystes. Elles se distinguent ainsi des cellules tumorales d'origine cancéreuse, qui se caractérisent par des mutations au niveau des gènes codant pour des proto-oncogènes et des gènes suppresseurs de tumeurs.

Cette tumorigénérité des cellules ES est un inconvénient majeur à leur utilisation en thérapie cellulaire. En effet, une population de cellules différenciées à partir de cellules ES contient souvent une infime proportion de cellules toujours indifférenciées pouvant être à l'origine d'un cancer.

1.4. Distinction entre les cellules ES et les cellules souches de l'embryon.

Les cellules ES sont dérivées à partir des cellules souches d'un embryon précoce. Cependant, les cellules souches de l'embryon n'existent que de façon transitoire et se différencient pendant le développement, alors que les cellules ES peuvent être cultivées *in*

vitro indéfiniment sans aucune différenciation. Les cellules ES ne sont donc pas strictement identiques aux cellules souches de l'embryon, même si elles partagent de nombreuses propriétés comme la pluripotence, la tumorigénéité et l'expression de marqueurs spécifiques. Les cellules ES doivent être considérées comme une adaptation en culture des cellules souches existantes dans l'embryon précoce.

En conclusion, les cellules souches embryonnaires (cellules ES) ont pour origine l'embryon au stade blastocyste. Ce sont des cellules indifférenciées, c'est-à-dire non spécialisées, qui se distinguent fondamentalement des autres types cellulaires en répondant à trois grands critères :

- (i) l'autorenouvellement : capacité à proliférer indéfiniment à l'état indifférencié sans altération du phénotype ni du caryotype et tout en conservant la propriété de pluripotence ;
- (ii) la pluripotence : capacité de se différencier *in vitro* en de nombreux types cellulaires représentatifs des trois feuillets embryonnaires (ectoderme, endoderme et mésoderme) et à former *in vivo* toutes les lignées cellulaires de l'embryon, y compris la lignée germinale, lorsqu'elles sont réimplantées dans un embryon précoce receveur ;
- (iii) la tumorigénéité : capacité à former des tumeurs (tératocarcinomes) lorsqu'elles sont réimplantées à l'état indifférencié dans un organisme adulte.

Nous nous intéresserons plus particulièrement aux cellules ES de souris puisqu'elles font l'objet de notre étude expérimentale.

La voie LIF/gp130/STAT3 : un rôle clé dans l'autorenouvellement des cellules ES.

2.1. La voie de signalisation LIF/gp130.

2.1.1. Le LIF.

Les cellules ES dépendent de cytokines pour leur maintien en autorenouvellement. Initialement, les cellules ES de souris dérivées des cellules de l'ICM du blastocyste nécessitaient, pour le maintien de leur pluripotence, d'être cultivées en présence d'une couche

unicellulaire de fibroblastes embryonnaires de souris (cellules nourricières ou feeders) (Evans and Kaufman, 1981) (Martin, 1981). L'isolement ultérieur d'une glycoprotéine soluble sécrétée par ces fibroblastes embryonnaires, le LIF, a permis la culture de certaines lignées de cellules ES sans MEFs (Smith et al., 1988). Il a été démontré par la suite que des lignées de cellules ES pouvaient être isolées à partir de blastocystes uniquement en présence de LIF, puis maintenues en culture en présence de cette même cytokine. Les cellules obtenues par cette méthode contribuent, après injection dans des blastocystes de souris, à la formation des tissus somatiques et germinaux des embryons chimères ; elles sont donc bien pluripotentes (Nichols et al., 1990). Les cellules ES isolées et cultivées en présence de LIF présentent donc les mêmes propriétés d'autorenouvellement, *in vitro*, et de pluripotence, *in vivo*, que les cellules ES isolées et cultivées sur cellules nourricières. En absence de LIF, les cellules ES se différencient en une population hétérogène de cellules essentiellement endodermiques et mésodermiques (Smith et al., 1988). De même, des cellules nourricières préparées à partir de fibroblastes embryonnaires dont le gène du LIF a été inactivé se révèlent incapables d'inhiber la différenciation des cellules ES (Stewart et al., 1992). Le LIF joue donc un rôle central dans la stimulation de l'autorenouvellement et l'inhibition de la différenciation des cellules ES de souris *in vitro*.

Le LIF appartient à la famille des cytokines de type IL6 comprenant l'IL-6 (Interleukine 6), l'IL-11 (Interleukine 11), le CNTF (Ciliary Neurotrophic Factor), l'OSM (Oncostatin M) et le CT-1 (Cardiotrophin 1). Les cytokines OSM, CNTF et CT-1 peuvent également inhiber la différenciation des cellules ES (Conover et al., 1993) (Pennica et al., 1995) (Rose et al., 1994) (Yoshida et al., 1994). En revanche, l'IL-6 et l'IL-11 ne peuvent pas entretenir l'autorenouvellement des cellules ES car leur récepteur spécifique n'est pas exprimé dans ces cellules. Toutefois, si une forme soluble du récepteur de l'IL-6 (sIL6R) ou de l'IL-11 (sIL11R) est ajoutée, ces cytokines peuvent alors inhiber la différenciation des cellules souches (Nichols et al., 1994) (Yoshida et al., 1994).

2.1.2. Les récepteurs aux cytokines de la famille du LIF.

2.1.2.1. Le rôle central du transducteur gp130.

Toutes les cytokines de la famille IL6 utilisent le récepteur gp130 pour la transduction de leurs signaux (Davis et al., 1993) (Gearing et al., 1992) (Pennica et al., 1995) (Yin et al.,

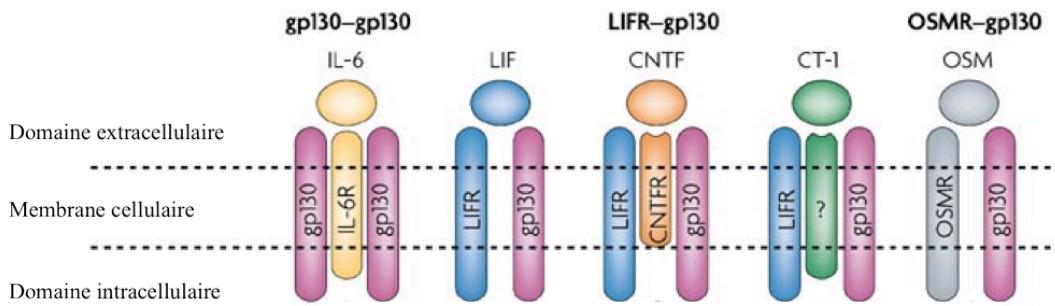


Figure 5 : Les complexes de récepteur aux cytokines apparentés au LIF

L'IL-6 utilise des homodimères de gp130 associées à une sous-unité IL-6R. Le CNTF et le CT-1 utilisent des hétérodimères composés de gp130 et de LIFR. L'OSM utilise, quant à lui, un hétérodimère composé de gp130 et d'OSMR mais il peut aussi utiliser alternativement le récepteur LIFR/gp130.

1993) (Yoshida et al., 1994). Le gp130 fait partie de la superfamille des récepteurs aux cytokines (Bazan, 1990) dont les membres sont dépourvus d'activité tyrosine kinase intrinsèque au niveau de leur région intracellulaire. Les cytokines de type IL6 utilisent des combinaisons variées de complexes de récepteurs contenant toujours la sous-unité gp130 associée à d'autres sous-unités plus ou moins spécifiques du ligand (Figure 5) :

- le LIF, l'OSM et le CT-1 utilisent des récepteurs hétérodimériques constitués de la sous-unité gp130 et du récepteur LIFR. L'OSM peut également utiliser un autre récepteur hétérodimérique comprenant toujours la sous-unité gp130 associé au récepteur spécifique de l'OSM, OSMR.
- Le CNTF se lie à une sous-unité récepitrice spécifique, le CNTFR, puis au complexe hétérodimérique gp130/LIFR.
- Les récepteurs de l'IL-6 et de l'IL-11 sont constitués des récepteurs spécifiques IL6R et IL11R liés à des homodimères de gp130, et ceci sans participation du récepteur LIFR.

Les parties cytoplasmiques des récepteurs spécifiques IL6R, IL11R et CNTFR ne sont pas requises pour la signalisation car ces récepteurs peuvent être remplacés par leurs analogues solubles. Cependant, leur liaison avec le ligand augmente l'affinité de ce dernier pour le récepteur dimérique (Davis et al., 1993) (Yin et al., 1993) (Yoshida et al., 1994). En revanche, la partie intracellulaire des récepteurs gp130, OSMR et LIFR est indispensable pour la transduction du signal.

2.1.2.2. Intérêt du récepteur GCSFR dans les cellules ES.

Le récepteur au GCSF (Granulocyte Colony Stimulating Factor) (GCSFR) appartient à la classe I de la superfamille des récepteurs aux cytokines et présente des régions d'homologie avec les récepteurs gp130 et LIFR ce qui suggère un mécanisme d'action commun à ces trois récepteurs. La fixation du ligand GCSF induit l'homodimérisation des récepteurs GCSFR (Gearing et al., 1991).

Niwa et ses collègues (Niwa et al., 1998) ont ainsi construit des récepteurs chimériques composés du domaine extracellulaire du récepteur au GCSF lié aux régions transmembranaire et cytoplasmique de gp130. Ces récepteurs chimériques ont pour intérêt d'éliminer l'interférence des chaînes LIFR et gp130 endogènes, en exploitant le fait que le GCSFR n'est pas exprimé dans les cellules ES murines. En permettant d'évaluer le rôle de la

partie intracellulaire de gp130 et de LIFR dans l'autorenouvellement des cellules ES, ces expériences ont démontré que le récepteur chimérique GCSFR/gp130 est capable de maintenir l'autorenouvellement des cellules ES en réponse au GCSF et ceci en l'absence de LIF. En revanche, la signalisation par le récepteur GCSFR/LIFR ne permet pas de maintenir les cellules en autorenouvellement. Le récepteur gp130, l'élément commun de ces différents couples récepteur:ligand, apparaît donc comme un élément déterminant dans la voie de signalisation du LIF et donc de l'autorenouvellement des cellules ES.

2.1.3. Les protéines kinases cytoplasmiques.

Les récepteurs aux cytokines sont dépourvus d'activité tyrosine kinase dans leur partie intracellulaire. Ils peuvent toutefois être phosphorylés sur leurs résidus tyrosines grâce à des tyrosine kinases cytoplasmiques auxquelles ils s'associent. Ces kinases cytoplasmiques sont au moins de deux types : les Janus Kinases ou JAKs et les Hemopoietic Cell Kinases ou Hcks.

2.1.3.1. Activation des kinases JAK.

Les JAKs sont une famille de tyrosine kinases cytoplasmiques composé de trois membres : JAK1, JAK2 et Tyk2 (Tyrosine Kinase 2) (Stahl et al., 1994). Les kinases JAK sont constitutivement associées aux récepteurs des cytokines et c'est après homo- ou hétérodimérisation du récepteur que les kinases sont activées par un phénomène de transphosphorylation. Elles phosphorylent alors les deux sous-unités du récepteur sur des tyrosines localisées dans le domaine intracellulaire.

Les trois membres de la famille des kinases JAK, JAK1, JAK2 et Tyk2, sont associés au récepteur gp130 et sont donc activés dans les cellules ES. Cependant, il apparaît que JAK1 joue un rôle plus prépondérant que les autres dans l'autorenouvellement des cellules ES. En effet, une lignée de cellules ES exprimant de façon stable une construction anti-sens inducible de JAK1 a besoin d'une quantité plus importante de LIF pour maintenir son état indifférencié (Ernst et al., 1996). Des études effectuées sur des lignées de cellules ES mutantes pour JAK1 et sur des souris déficientes pour le gène *Jak1* montrent également que cette kinase joue un rôle important dans la transduction du signal (Guschin et al., 1995) (Rodig et al., 1998).

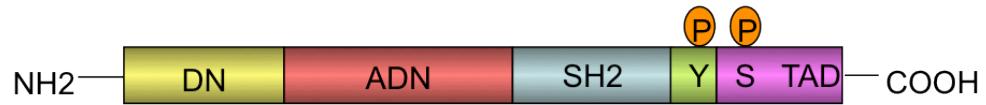


Figure 6 : Structure de la protéine STAT3

DN : domaine N-terminal ; ADN : domaine de fixation à l'ADN ; SH2 : « Src homology domain » se liant à des phosphotyrosines ; Y : tyrosine 705 phosphorylable ; S : sérine 727 phosphorylable ; P : phosphate PO₃²⁻ ; TAD : « Transcription Activating Domain »

2.1.3.2. Activation des kinases Hck.

En plus des JAKs, les protéines cytoplasmiques à activité tyrosine kinase de la famille Src (Rous sarcoma oncogene), dont fait partie la protéine Hck, sont également impliquées dans la signalisation induite par le LIF (Ernst et al., 1994). Le rôle des membres de la famille Src dans la transduction du signal par gp130 sera discuté dans le paragraphe 2.2.4. *La voie de signalisation Src.*

2.2. Les différentes voies de signalisation activées par le LIF dans les cellules ES.

2.2.1. La voie LIF/gp130/STAT3.

2.2.1.1. Le facteur de transcription STAT3.

2.2.1.1.1. Structure.

Les facteurs de transcription de la famille STAT (Signal Transducer and Activator of Transcription) sont impliqués dans la voie de signalisation de nombreuses cytokines et facteurs de croissance. Sept gènes codant pour les STATs ont été identifiés chez les mammifères (Copeland et al., 1995). Ces gènes codent pour des protéines de 750 à 850 acides aminés. Elles ont en commun (Figure 6) (Zhong et al., 1994) (Horvath et al., 1995) (Wen et al., 1995):

- un domaine N-terminal très conservé
- un domaine central de fixation à l'ADN
- un domaine SH2 (Src Homology Domain 2) permettant la liaison des protéines STAT à des phosphotyrosines
- une tyrosine conservée, adjacente au domaine SH2
- un domaine TAD (Transcription Activating Domain) à l'extrémité C-terminale de la protéine
- une sérine conservée dans le domaine TAD (Heinrich et al., 1998)

Parmi les facteurs STAT, STAT3 joue un rôle déterminant dans l'inhibition de la différenciation des cellules ES induite par le LIF (Boeuf et al., 1997).

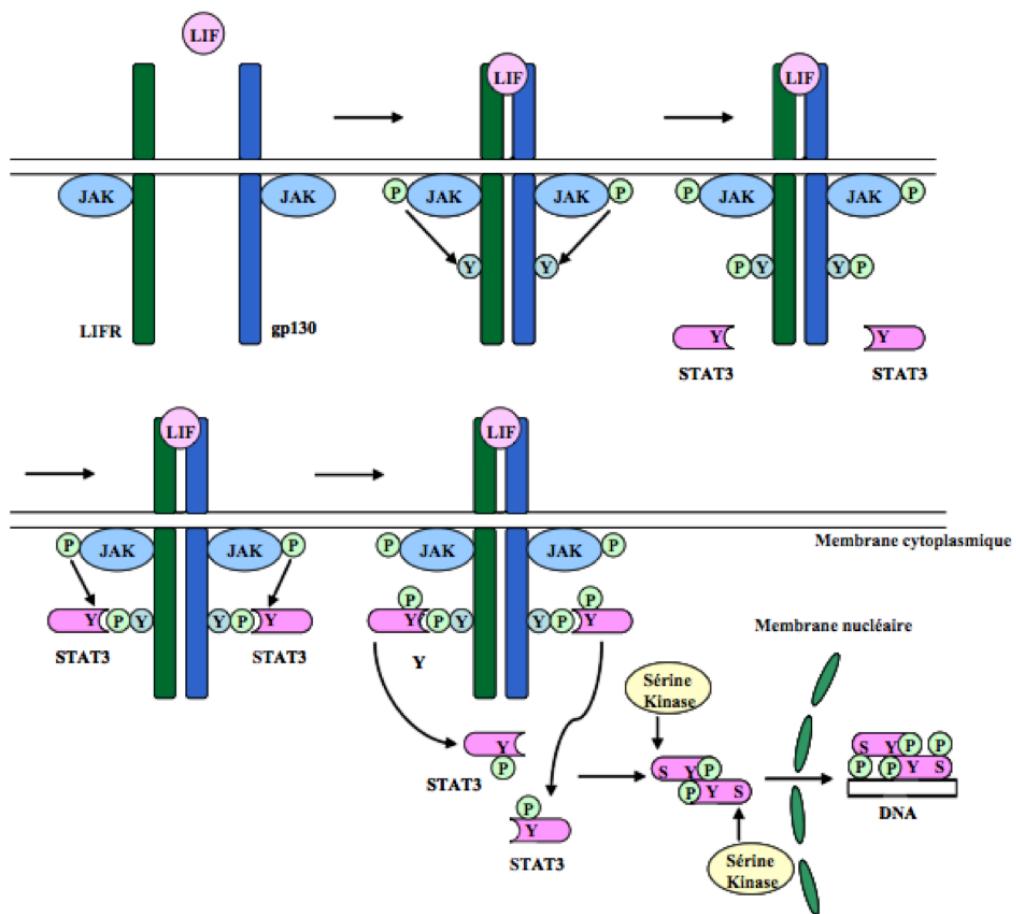


Figure 7 : La voie de signalisation LIF/gp130/STAT3

Le LIF se lie à un récepteur constitué de deux sous-unités : gp130 et LIFR (récepteur spécifique au LIF). Les kinases JAK sont constitutivement associées aux deux sous-unités. L'hétérodimérisation du récepteur induit l'activation des kinases JAK, qui se phosphorylent mutuellement et phosphorylent les deux sous-unités du récepteur sur des tyrosines localisées dans la partie cytoplasmique. Cette phosphorylation induit le recrutement du facteur de transcription STAT3 qui, à son tour, est phosphorylé par les kinases JAK. La phosphorylation de STAT3 induit son homodimérisation par interaction entre les domaines SH2 et les phosphotyrosines. Les dimères de STAT3 sont ensuite phosphorylés sur des résidus séries par des séries kinases cytoplasmiques. Les dimères de STAT3 phosphorylés sont ensuite transloqués dans le noyau, où ils régulent la transcription de gènes cibles responsables du maintien des cellules ES en autorenouvellement.

P : groupement phosphate

Y : résidu tyrosine

2.2.1.1.2. Mécanisme d'action.

La présence d'un domaine SH2 est exceptionnelle dans la famille des facteurs de transcription et confère aux facteurs STAT la propriété de transmettre le signal de la membrane vers le noyau sans autre intermédiaire. En effet, le domaine SH2 permet aux facteurs STAT de se lier aux phosphotyrosines du récepteur aux cytokines ; le facteur STAT3 est recruté au niveau de motifs YXXQ phosphorylés présents dans la partie intracellulaire du récepteur gp130 (Stahl et al., 1995). STAT3 est alors à son tour phosphorylé par les kinases JAK sur la tyrosine Y705 adjacente au domaine SH2 (Figure 7). La phosphorylation de cette tyrosine est nécessaire pour l'activation des STATs. En effet, après phosphorylation, les monomères de STAT3 se dissocient du récepteur et forment des homodimères via l'interaction entre le domaine SH2 d'un monomère et la phosphotyrosine Y705 de l'autre et vice versa. Les dimères de STAT3 sont alors transloqués dans le noyau où ils se lient à l'ADN et régulent ainsi la transcription de gènes cibles (Darnell, 1997).

2.2.1.2. Le facteur de transcription STAT3 joue un rôle central dans l'autorenouvellement des cellules ES de souris.

Plusieurs études complémentaires ont permis de démontrer que le facteur STAT3, lorsqu'il est activé par le LIF, joue un rôle déterminant dans le maintien des cellules ES de souris en autorenouvellement.

2.2.1.2.1. STAT3 est nécessaire à l'autorenouvellement des cellules ES.

Le rôle de STAT3 est donc de transmettre le message du LIF à partir de la membrane cytoplasmique vers le noyau où il se lie, au niveau de l'ADN, à des éléments spécifiques de différents promoteurs.

Au niveau de la membrane, STAT3 se lie à quatre motifs YXXQ présents dans la partie intracellulaire des récepteurs gp130 et LIF, situés au niveau des tyrosines Y126, Y173, Y265 et Y275 ; Y126 étant la tyrosine la plus proche de la membrane

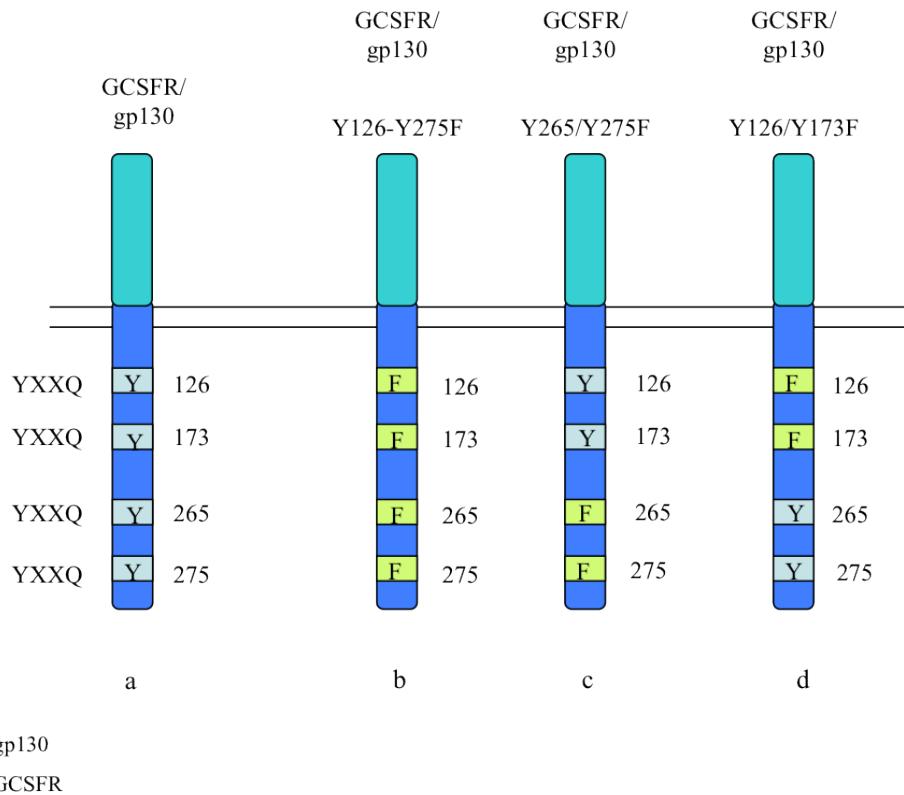


Figure 8 : Représentation schématique des récepteurs utilisés dans les expériences de substitution des tyrosines de la partie intracytoplasmique de gp130 décrites par Niwa et al., 1998

Tous les récepteurs représentés sont chimériques et constitués du domaine extracellulaire de GCSFR et du domaine intracellulaire de gp130.

- (a) Récepteur chimérique GCSFR/gp130 intégral comportant quatre tyrosines incluses dans un motif YXXQ : Y126, Y173, Y265, Y275 dans sa partie cytoplasmique.
- (b) Récepteur chimérique GCSFR/gp130 Y126-Y275F mutant comportant une substitution des quatre tyrosines en phénylalanine Y126F, Y173F, Y265F, Y275F
- (c) Récepteur chimérique GCSFR/gp130 Y265/275F mutant comportant une substitution des deux tyrosines adjacentes à la partie C-terminale en phénylalanine Y265F, Y275F
- (d) Récepteur chimérique GCSFR/gp130 Y126/173F mutant comportant une substitution des deux tyrosines proches de la membrane en phénylalanine Y126F, Y173F

Y : résidu tyrosine , F : résidu phénylalanine

(Figure 8) (Stahl et al., 1995). Des expériences réalisées par Niwa et ses collègues (Niwa et al., 1998) avec des récepteurs chimériques constitués du GCSFR et de la sous-unité gp130 intacte ou portant des substitutions dans la partie carboxy-terminale ont permis de montrer que les quatre tyrosines ne sont pas équivalentes pour la phosphorylation de STAT3. La substitution dans le récepteur chimérique des quatre tyrosines responsables de l'activation de STAT3 en phénylalanine (Y126F/Y173F/Y265F/Y275F) inhibe la phosphorylation de ce facteur et la capacité du récepteur à prévenir la différenciation des cellules ES après stimulation par du GCSF (Figure 8b). La substitution des deux tyrosines adjacentes à la partie C-terminale, Y265 et Y275, en phénylalanine (Y265F/Y275F) entraîne également la différenciation des cellules ES et la diminution de la phosphorylation de STAT3 (Figure 8c) alors que ceci n'est pas observé lors de la substitution des deux tyrosines proches de la membrane Y126 et Y173 (Y126F/Y173F) (Figure 8d). Les deux motifs YXXQ de l'extrémité C-terminale (Y265 et Y275) sont donc nécessaires à l'activation de STAT3 et au maintien du phénotype indifférencié des cellules ES.

Dans les différentes expériences réalisées avec des récepteurs chimériques, la différenciation des cellules ES induite par des mutations ou des délétions de ces récepteurs est corrélée à une diminution de la phosphorylation et, par conséquent, de l'activité de STAT3 (Stahl et al., 1995) (Niwa et al., 1998) (Ernst et al., 1999). Afin d'étudier plus en détail l'importance de STAT3 dans le maintien du phénotype indifférencié des cellules ES, un mutant STAT3F dans lequel la tyrosine Y705, responsable de la dimérisation et de la translocation de ce facteur de transcription, est remplacée par une phénylalanine a été utilisé. La mutation de cette tyrosine empêche la phosphorylation de STAT3 par les kinases JAK. Elle confère ainsi au facteur STAT3 la propriété de mutant dominant négatif car il se fixe de façon irréversible au récepteur et empêche le recrutement du facteur STAT3 endogène (Minami et al., 1996). La surexpression de ce mutant dominant négatif induit la différenciation des cellules ES de souris et ceci même en présence de LIF (Boeuf et al., 1997) (Niwa et al., 1998). L'effet de STAT3F peut être supprimé par la surexpression dans les cellules d'un ADN complémentaire (ADNc) codant pour le facteur STAT3 sauvage, alors que la surexpression d'un autre facteur STAT, STAT1, n'a aucun effet sur les conséquences de l'expression de STAT3F (Niwa et al., 1998). Ces résultats montrent que STAT3 est nécessaire à l'autorenouvellement des cellules ES de souris et que cette fonction lui est spécifique. En fait, STAT1 et STAT3 sont fortement apparentés. En effet, STAT1 peut être activé par le LIF dans les cellules ES à des niveaux cependant plus faibles (Starr et al., 1997),

et former ainsi des hétérodimères avec STAT3 (Darnell, 1997) (Darnell et al., 1997). Cependant, il ne semble pas que STAT1 joue un rôle important dans l'autorenouvellement des cellules ES ; en effet, des cellules ES STAT1^{-/-} présentent un phénotype indifférencié (Durbin et al., 1996).

D'autres stratégies ont été utilisées pour pallier à la difficulté de propagation des cellules ES après expression du mutant STAT3F, comme l'utilisation d'un système d'expression inducible tet-off (tet pour tétracycline). Dans ce système, l'expression de STAT3F est sous le contrôle transcriptionnel du promoteur tet-O (résultant de la fusion entre le répresseur Tet et VP16) qui est lui-même inhibé en présence de tétracycline. Les cellules ES cultivées en présence de tétracycline, dans le milieu de culture, sont maintenues en autorenouvellement dû à l'inhibition de l'expression du mutant STAT3F. La suppression de la tétracycline du milieu de culture entraîne la différenciation des cellules ES ainsi que la diminution de la liaison de STAT3 à l'ADN, évaluée par des expériences de retard de migration sur gel (Niwa et al., 1998). Toutes ces expériences démontrent que le facteur de transcription STAT3 est nécessaire à l'autorenouvellement des cellules ES.

2.2.1.2.2. STAT3 est-il suffisant pour l'autorenouvellement des cellules ES ?

Pour démontrer que STAT3 est suffisant pour l'autorenouvellement des cellules ES, Matsuda et ses collègues (Matsuda et al., 1999) ont généré un mutant conditionnel de STAT3 formé par la protéine STAT3 fusionnée au domaine de fixation du ligand (LBD : ligand binding domain) du récepteur à l'œstrogène (ER : estrogen receptor). Ce domaine LBD a été muté pour ne plus fixer les oestrogènes endogènes. En revanche, il peut fixer le ligand synthétique 4'-hydroxy-tamoxifène (4'OHT). En présence du ligand, le facteur STAT3-ER chimérique s'homodimérisé via son domaine ER puis il est transloqué dans le noyau où il active les gènes cibles. Cette activation est indépendante de la phosphorylation de la tyrosine Y705 par les kinases JAK. Dans ce système, l'activité du facteur de transcription STAT3-ER est indépendante de la voie LIF/gp130. Des cellules ES de souris exprimant STAT3-ER peuvent être maintenues dans un état indifférencié en absence de LIF et en présence de 4'OHT. Après injection dans des blastocystes, les cellules contribuent au développement de tous les tissus embryonnaires avec une efficacité comparable à celle des cellules ES cultivées en présence de LIF. L'activation du facteur STAT3 est donc capable de remplacer complètement l'activation du récepteur par le LIF.

Il est important de noter que les expériences de Matsuda et ses collègues (Matsuda et al., 1999) ont été réalisées en présence de sérum, et à forte densité cellulaire ce qui est connu

comme favorisant l'action de facteurs paracrines. En effet, il a été montré qu'en absence de sérum, les cellules ES se différencient et cela même en présence de LIF et donc de STAT3 actif (Ying et al., 2003). Un ou plusieurs autres facteurs présents dans le sérum agissent donc en synergie avec le LIF pour maintenir les cellules ES en autorenouvellement et inhiber leur différenciation. STAT3 est donc suffisant pour le maintien en autorenouvellement des cellules ES mais sous certaines conditions de culture, c'est-à-dire en présence de sérum. Des travaux récents montrent que c'est la voie de signalisation BMP4/Smad (Bone Morphogenetic Protein 4/Mad Homologue) et le facteur HLH (helix-loop-helix) Id2 (Inhibitor of differentiation 2) qui sont impliqués dans ce processus (Ying et al., 2003) (cf. chapitre 3. *La voie BMP4/Smad/Id2*).

2.2.1.3. *Les gènes cibles de STAT3.*

En dépit du rôle central joué par STAT3 dans le contrôle de l'autorenouvellement des cellules ES, on ne connaît à ce jour que très peu les gènes cibles qu'il régule, gènes dont l'activité serait responsable de l'inhibition de la différenciation.

Les trois gènes *junB*, *c-fos* et *c-jun*, des régulateurs de la transition G1/S du cycle cellulaire, ont été les premiers à être identifiés comme des cibles directes du facteur STAT3 dans les cellules ES (Matsuda et al., 1999) (Duval et al., 2000). L'inactivation de ces gènes ne semble pas altérer l'équilibre autorenouvellement-différenciation des cellules ES de souris, indiquant qu'aucun d'entre eux n'est un élément déterminant de l'action du facteur STAT3 dans ces cellules (Field et al., 1992) (Hilberg and Wagner, 1992) (Schorpp-Kistner et al., 1999).

Ces dernières années de nouveaux gènes cibles de STAT3 ayant un rôle dans l'inhibition de la différenciation des cellules ES ont été identifiés :

- (i) *myc* (myelocytomatosis oncogene) : dont la surexpression sous forme d'un mutant hormono-dépendant Myc-ER inhibe la différenciation des cellules ES induite par la privation de LIF, alors que l'expression d'une forme tronquée de la protéine induit la différenciation des cellules et ceci même en présence de LIF et donc de STAT3 actif (Cartwright et al., 2005),
- (ii) *Socs-3* (Suppressor of cytokine signaling 3) : Socs-3 est un inhibiteur des protéines kinases JAK, et agit comme un signal de rétrocontrôle négatif au niveau du récepteur gp130

(Starr and Hilton, 1999). L'inactivation de Socs-3 entraîne la différenciation des cellules ES. En fait, ce signal de rétrocontrôle négatif semble promouvoir le signal envoyé par les protéines STAT dans les cellules ES, et ceci en inhibant la voie SHP2/ERK induite par le LIF mais dont le rôle est d'induire la différenciation des cellules (Duval et al., 2000) (Forrai et al., 2006). D'autres études semblent désigner Socs-3 comme étant une protéine jouant un rôle dans la survie cellulaire (Duval et al., 2006),

(iii) *Klf4* (Krüppel-like factor 4) : la surexpression de *Klf4* rend les cellules résistantes à la différenciation induite par la privation de LIF (Li et al., 2005),

(iv) *Dax1* (DSS-AHC critical region on the X chromosome, gene 1) et *Eed* (Embryonic ectoderm development) : *Dax1* ou NrOb1 (Nuclear receptor subfamily 0, group B, member 1) est un récepteur nucléaire orphelin et il agit comme un répresseur transcriptionnel. Les cellules ES dans lesquelles l'expression de *Dax1* est inhibée, par la technique d'interférence à l'ARN, se différencient au bout de 24h (Niakan and McCabe, 2005). Un autre gène cible de STAT3 a été découvert très récemment : *Eed* (Ura et al., 2008). *Eed* appartient au complexe répresseur Polycomb 2 (PRC2) qui induit des modifications au niveau de la chromatine, en particulier la méthylation de l'histone H3 et de la Lysine 27. Dans les cellules ES ces modifications se font au niveau de gènes couplés à la différenciation, leur méthylation induit l'inhibition de leur expression (Boyer et al., 2006) (Azuara et al., 2006) (Lee et al., 2006) (Pasini et al., 2007). L'inhibition de l'expression de *Eed* dans les cellules ES induit leur différenciation (Ura et al., 2008).

De nombreux gènes cibles de STAT3 ont, en revanche, été identifiés dans d'autres types cellulaires. Ces gènes sont, pour la plupart, des gènes codant pour des facteurs antiapoptotiques (Bcl-xL, Bcl-2, Mcl-1) ou des gènes codant pour des régulateurs du cycle cellulaire tels que la cycline D1, la protéine Myc, la séro-thréonine kinase Pim1 (Proviral integration site 1), ou l'inhibiteur de kinases p21^{cip1} (Bowman et al., 2000). Le facteur STAT3 joue donc un rôle de régulateur du cycle cellulaire dans les cellules différenciées et un rôle d'inhibition de la différenciation dans les cellules ES de souris. Cette double fonction peut nous amener à penser que la régulation du cycle cellulaire, via la voie LIF/gp130/STAT3, est l'un des mécanismes impliqués dans le contrôle des cellules ES en autorenouvellement.

2.2.1.4. La voie de signalisation LIF/gp130/STAT3 *in vivo*.

Les expériences réalisées sur les cellules souches embryonnaires *in vitro* ont permis de mettre en évidence le rôle crucial de la voie LIF/gp130/STAT3 dans le maintien en autorenouvellement des cellules ES. Cependant, la voie de signalisation LIF/gp130/STAT3 ne semble jouer aucun rôle dans le contrôle de l'autorenouvellement des cellules souches embryonnaires *in vivo*, c'est-à-dire dans les cellules de l'épiblaste. En effet, bien que le LIF, le LIFR et gp130 soient exprimés tous les trois dans le blastocyste (Nichols et al., 1996), aucun d'entre eux n'est nécessaire à l'autorenouvellement des cellules *in vivo*. Les embryons de souris « knock-out » LIF^{-/-}, LIFR^{-/-} ou gp130^{-/-} se développent normalement au cours du développement péri- et post-implantatoire précoce, après la gastrulation, période pendant laquelle les cellules souches pluripotentes du bouton embryonnaire se multiplient activement (Li et al., 1995) (Yoshida et al., 1996). Il en est de même pour les embryons STAT3^{-/-} qui se développent jusqu'au stade de la gastrulation et meurent vers le 8^{ème} jour de gestation après formation de l'épiblaste (Takeda et al., 1997). Les cellules souches embryonnaires prolifèrent donc normalement en l'absence de la voie LIF/gp130/STAT3 active. D'autres voies de signalisation, distinctes de la voie LIF/gp130/STAT3, interviennent alors *in vivo* pour maintenir la pluripotence et l'autorenouvellement des cellules souches embryonnaires.

La voie de signalisation LIF/gp130/STAT3 joue néanmoins un rôle physiologique important durant la diapause. La diapause consiste en un retard d'implantation des embryons dont le développement s'arrête au stade blastocyste, dans l'attente de conditions favorables à la nidation et à la poursuite de l'embryogénèse. La diapause s'installe chez les femelles au cours de la lactation, ou après ovariectomie expérimentale, conduisant à un déficit en œstrogène. Cet état peut durer jusqu'à quatre semaines. L'implantation de l'embryon dans l'utérus d'une souris receveuse (qui n'est pas en cours de lactation) ou l'apport d'œstrogènes dans des conditions expérimentales entraîne la nidation et la reprise du développement normal des embryons. En revanche, les embryons gp130^{-/-} sont incapables de continuer leur développement après la diapause, en raison de l'absence de la formation de l'épiblaste (Nichols et al., 2001). Il est intéressant de noter que les premières lignées de cellules ES ont été dérivées à partir d'embryons en diapause et que la fréquence de dérivation de cellules ES dépend du retard d'implantation (Brook and Gardner, 1997) (Evans and Kaufman, 1981). La protéine gp130 joue donc un rôle important dans le maintien de la pluripotence des cellules de l'ICM pendant la diapause.

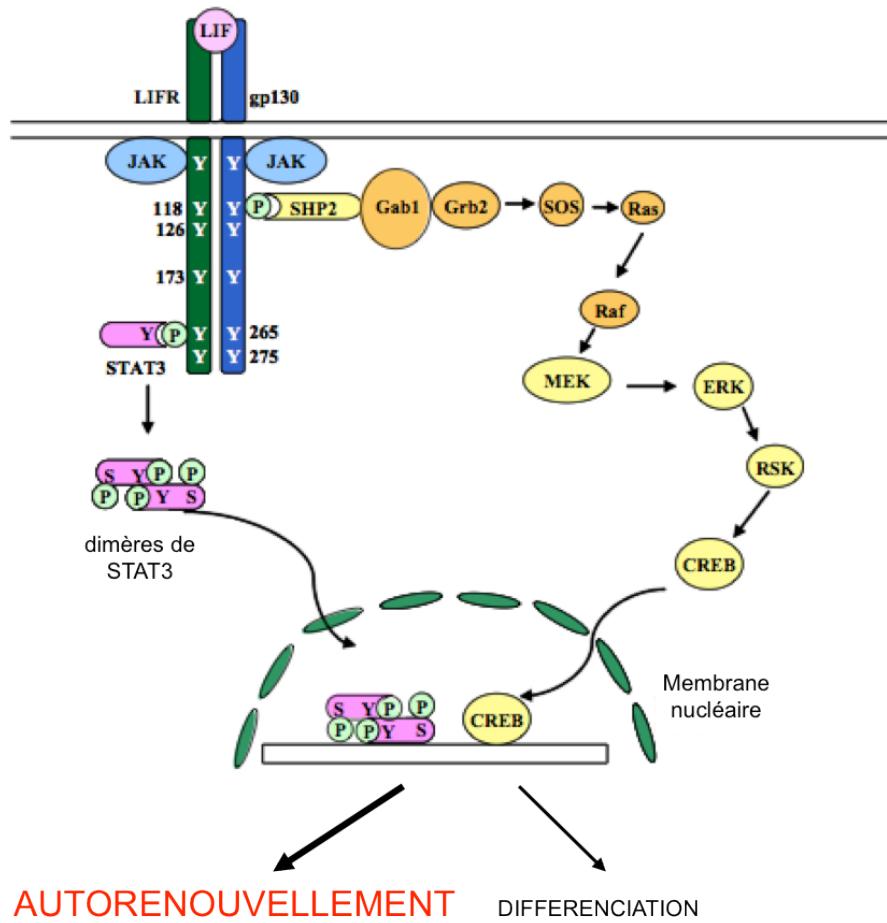


Figure 9 : La voie de signalisation SHP2/ERK

Suite à la liaison du LIF sur le récepteur hétérodimérique LIFR/gp130, la protéine phosphatase SHP2 est recrutée au niveau de la tyrosine Y118 du récepteur ce qui entraîne son activation. Cette dernière recrute les complexes Grb2/SOS via la protéine adaptatrice Gab1 et active ainsi la cascade de signalisations Ras/Raf/MEK/ERK. Puis suite à cette activation, RSK puis CREB sont à leur tour activés. CREB est ensuite transloqué dans le noyau et active ainsi la transcription de gènes responsables de la différenciation, contrairement à STAT3 qui va activer la transcription de gènes nécessaires à l'autorenouvellement des cellules ES. Dans les conditions optimales de culture l'équilibre entre les deux voies est nettement déplacé vers la voie de l'autorenouvellement.

P : groupement phosphate

Y : résidu tyrosine

2.2.2. La voie SHP2/ERK.

Parallèlement à l'activation du facteur STAT3, l'hétérodimérisation du récepteur LIFR avec le transducteur gp130 permet l'activation d'une autre voie de signalisation : la voie SHP2/ERK (SH2 containing protein tyrosine Phosphatase 2 / Extracellular Regulated MAP Kinase). La voie de signalisation SHP2/ERK est impliquée dans de nombreux mécanismes cellulaires comme le contrôle de la croissance, la prolifération ou la régulation du cycle cellulaire dans les cellules somatiques (Lloyd, 1998). L'activation de cette voie au niveau des récepteurs induit une cascade de phosphorylation faisant intervenir de nombreuses molécules de signalisation intermédiaires.

2.2.2.1. Régulation de la voie SHP2/ERK par le LIF.

SHP2 est une tyrosine phosphatase cytoplasmique recrutée par gp130. Au niveau du récepteur gp130, SHP2 se lie par l'intermédiaire de son domaine SH2 à une tyrosine en position 118 (Y118) qui n'est pas située dans les motifs consensus YXXQ utilisés par STAT3 (Fukada et al., 1996). En effet, la mutation de cette tyrosine du récepteur chimérique GCSFR-gp130 en phénylalanine : GR-gp(Y118F) ne permet plus le recrutement de SHP2 au niveau du récepteur après stimulation des cellules par le GCSF (Burdon et al., 1999a).

Le recrutement de SHP2, au niveau du récepteur intact, induit sa phosphorylation au niveau de deux tyrosines localisées dans la partie C-terminale. L'une de ces deux tyrosines, Y452 constitue alors un site de fixation pour la protéine Grb2 (Growth factor receptor bound protein 2). Grb2 recrute ensuite la protéine SOS (Son of Sevenless homologue) qui va activer la protéine Ras (Rat sarcoma virus oncogene) (Li et al., 1994). Cette activation va entraîner une cascade de phosphorylation impliquant les protéines Raf, MEK (Mitogen activated protein ERK Kinase) et ERK (Figure 9).

La phosphorylation des kinases ERK est la conséquence du recrutement de SHP2 sur le récepteur gp130. Elles sont des gènes cibles de la voie LIF (Boeuf et al., 1997) (Ernst et al., 1996). En effet, la mutation de la tyrosine Y118 en phénylalanine au niveau des récepteurs chimériques GSCFR-gp130 transfectés dans des cellules ES empêche la phosphorylation des protéines ERK (Burdon et al., 1999b) après stimulation par du GCSF. L'activation des protéines ERK dépend donc de la transduction du signal par la tyrosine Y118.

2.2.2.2. Rôle de la voie de signalisation SHP2/ERK dans l'autorenouvellement des cellules ES.

De nombreuses études ont permis de montrer que la voie SHP2/ERK joue un rôle négatif dans la régulation de l'autorenouvellement des cellules ES.

2.2.2.2.1. Rôle de SHP2.

Des cellules exprimant le récepteur chimérique GR-gp(Y118F), qui n'engage donc plus la protéine SHP2, sont capables de s'autorenouveler en présence de GCSF. A l'inverse, les cellules exprimant le récepteur chimérique GR-gp(Y126-275F), qui ne permet plus le recrutement de STAT3 au niveau du récepteur, se différencient lorsque le LIF est remplacé par du GSCF (Burdon et al., 1999b) (Niwa et al., 1998). De plus, les cellules ES exprimant le récepteur chimérique GR-gp(Y118F) requièrent des concentrations de GCSF 1000 fois plus faibles que celles requises pour l'autorenouvellement des cellules transfectées par le récepteur chimérique intact dans lesquelles SHP2 se fixe au niveau du récepteur gp130. La protéine SHP2 joue donc un rôle d'inhibiteur de la transduction du signal médié par gp130 dans les cellules ES. Il semble donc que l'activité phosphatase de SHP2 soit impliquée dans cette inhibition. En effet, la transfection de cellules ES par des ADNc codant pour une enzyme SHP2 inactive (où une cystéine indispensable à l'activité catalytique a été remplacée par une sérine), entraîne une augmentation du pourcentage de colonies de cellules indifférenciées (Burdon et al., 1999a).

Par ailleurs, une délétion des 65 acides aminés N-terminaux de SHP2 faisant partie du domaine de fixation au récepteur a été obtenue par recombinaison homologue. Les cellules homozygotes pour cette mutation ont besoin d'une concentration quatre fois moins importante de LIF par rapport à des cellules ES exprimant des protéines SHP2 fonctionnelles pour rester indifférenciées. Ces cellules présentent en outre une meilleure capacité à générer des corps embryoïdes secondaires. Les corps embryoïdes primaires sont obtenus à partir de cellules ES alors que les corps embryoïdes secondaires sont obtenus à partir des cellules issues de corps embryoïdes primaires dissociés. Les corps embryoïdes secondaires reflètent la proportion de cellules ES indifférenciées restant dans les corps embryoïdes primaires. En effet, seules des cellules ES indifférenciées sont capables de générer ces structures. La forte capacité des cellules mutantes pour SHP2 à générer des corps embryoïdes secondaires reflète le rôle important de ce facteur dans la différenciation des cellules ES.

2.2.2.2. Le rôle des protéines ERK.

L'utilisation d'inhibiteurs chimiques et des expériences de « knock-out » chez la souris ont permis de mettre en évidence le rôle des protéines ERK dans l'autorenouvellement des cellules.

L'inhibition de la phosphorylation des protéines ERK (ERK1 et ERK2 dans les cellules ES) par le composé chimique PD098059, un inhibiteur spécifique des MEKs, ne bloque pas l'autorenouvellement des cellules ES (Figure 9). Au contraire, cette inhibition stimule l'autorenouvellement des cellules et ceci de manière dose-dépendante. Les cellules ES cultivées en présence de l'inhibiteur PD098059 sont donc capables de coloniser l'embryon après leur injection dans des blastocystes de souris et contribuent à la formation de la lignée germinale des souris issues de ces embryons. La capacité de ces cellules à former des chimères prouve qu'elles sont pluripotentes et permet de conclure que l'activité des protéines ERK n'est pas nécessaire au maintien de la pluripotence des cellules ES (Burdon et al., 1999a).

Des souris « knock-out » ERK1^{-/-} et ERK2^{-/-} ont été générées afin d'analyser le rôle des ces deux protéines chez l'embryon. Les souris ERK1^{-/-} sont viables, fertiles et de taille normale. Elles présentent uniquement un défaut au niveau de la maturation et de la prolifération des thymocytes (Pages et al., 1999). Par contre, le phénotype des souris ERK2^{-/-} est plus sévère. En effet, ces souris meurent au jour 8.5 du développement embryonnaire. Chez ces embryons, le cône ectoplacental et l'ectoderme extra-embryonnaire ne se forment pas, mais l'épiblaste lui se développe normalement (Saba-El-Leil et al., 2003).

L'ensemble de ces données confirme l'hypothèse selon laquelle les protéines ERK et plus généralement la voie SHP2/ERK ne sont pas impliquées dans le contrôle de l'autorenouvellement des cellules ES de souris. Au contraire, cette voie est importante dans la différenciation des cellules ES puisque les cellules homozygotes pour la mutation SHP2 sont incapables de se différencier *in vitro* et donnent des corps embryoides contenant une proportion importante de cellules souches indifférenciées (Qu and Feng, 1998). De plus, l'expression d'une enzyme SHP2 inactive dans les cellules ES inhibe leur différenciation (Burdon et al., 1999a). Ainsi, la voie de signalisation induite par le LIF stimule à la fois l'autorenouvellement (via STAT3) et la différenciation (via ERK1 et ERK2) des cellules ES de souris. Dans les conditions optimales de culture, l'équilibre entre les deux voies est nettement déplacé vers la voie de l'autorenouvellement.

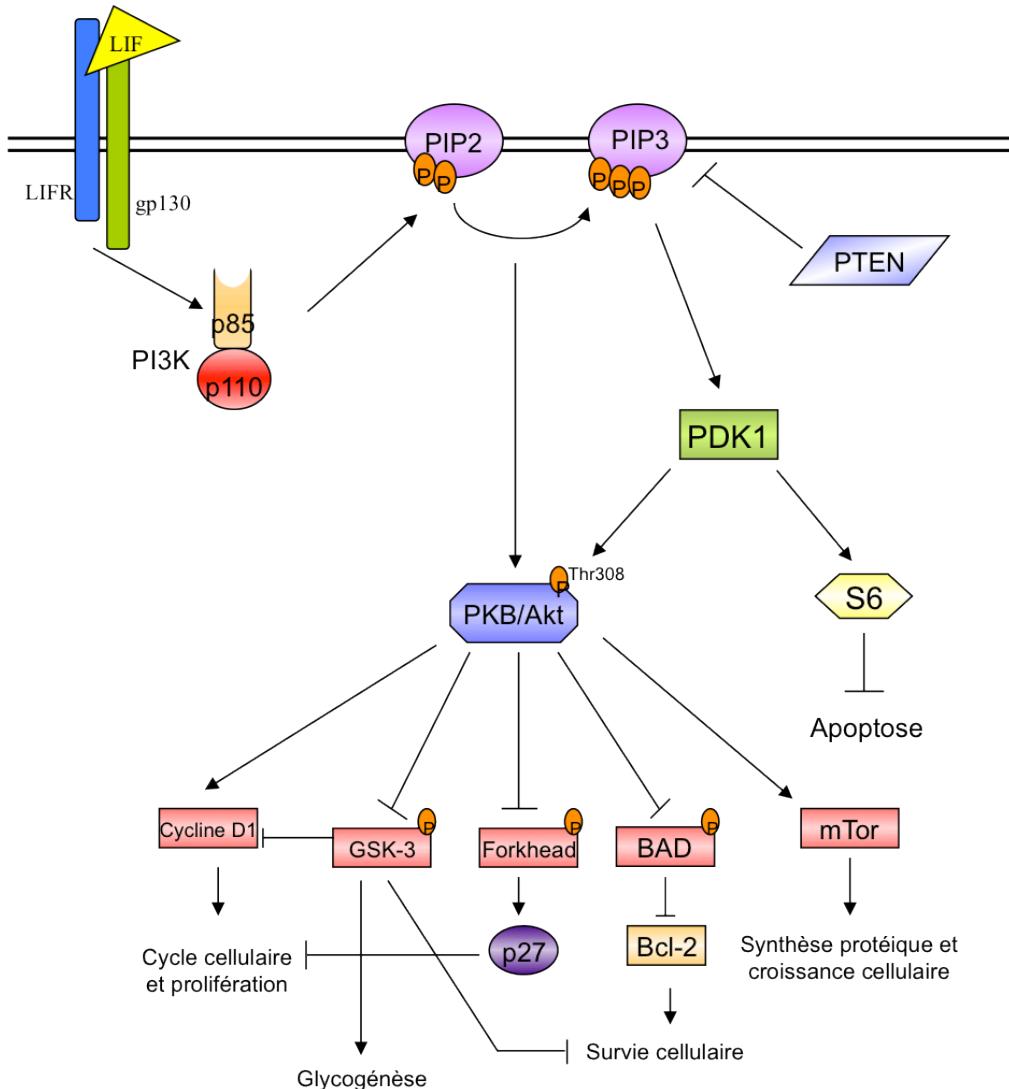


Figure 10 : La voie PI3Kinase

La liaison du LIF sur le récepteur hétérodimérique LIFR/gp130 induit l'activation de la PI3K, composée de deux sous-unités p85 et p110, et la synthèse du PIP3 (Phosphatidylinositol (3,4,5) triphosphate). Cela permet le recrutement de nombreux transducteurs dont les sérine/thréonine kinases PDK1 (Phosphoinositide-Dependent-Kinase 1) et PKB/Akt (Protein Kinase B). La kinase PDK1 phosphoryle PKB/Akt au niveau de la thréonine Thr308 et active également la kinase S6 qui joue un rôle dans l'inhibition de l'apoptose. La protéine PKB/Akt active va ensuite phosphoryler de nombreux substrats dont la GSK-3, Forkhead, BAD et les inhiber. PKB/Akt va également activer l'expression de la Cycline D1 et de mTor et ainsi promouvoir la prolifération cellulaire. Le suppresseur de tumeur PTEN déphosphoryle les 3-phosphoinositides et inhibe cette voie. La voie PI3Kinase module ainsi l'activité de nombreuses protéines régulatrices de l'apoptose, du cycle cellulaire et du métabolisme cellulaire.

2.2.3. La voie PI3Kinase.

La voie PI3K (Phosphatidyl-Inositol 3 Kinase) est impliquée dans la régulation de la prolifération et la survie d'un grand nombre de types cellulaires (Vanhaesebroeck et al., 2001).

La PI3K est composée d'une sous-unité régulatrice p85 et d'une sous-unité catalytique p110. L'activation de la PI3K par le récepteur LIFR/gp130 induit la synthèse *de novo* du PIP3 (Phosphatidyl-Inositol(3,4,5)triphosphate) au niveau de la membrane cellulaire. L'accumulation de ce lipide au niveau de la membrane permet le recrutement de nombreux transducteurs dont les sérine/thréonine kinases PDK1 (Phosphoinositide-Dependent-Kinase 1) et PKB/AKT (Protein Kinase B/Thymoma viral proto-oncogene 1). La kinase PDK1 phosphoryle PKB/AKT au niveau de la thréonine Thr308. La protéine PKB/AKT active va ensuite phosphoryler de nombreux substrats localisés au niveau de la membrane, du cytoplasme ou du noyau des cellules et ainsi moduler l'activité de protéines régulatrices de l'apoptose, du cycle cellulaire et du métabolisme cellulaire (Figure 10). Le suppresseur de tumeur PTEN (Phosphatase and Tensin homolog), qui déphosphoryle les 3-Phosphoinositides est un régulateur négatif de cette voie (Vanhaesebroeck et al., 2001).

2.2.3.1. Activation de la voie PI3K par le LIF.

Nous avons vu que le LIF active non seulement la voie STAT3 mais également la voie SHP2/ERK dans les cellules ES. En plus de ces deux voies, le LIF régule également la voie PI3K. La fixation du LIF sur son récepteur induit la phosphorylation de trois effecteurs de la voie PI3K qui sont les protéines AKT, GSK3 β (Glycogen Synthase Kinase 3 β) et S6. Le LY294002 est un inhibiteur de la PI3K dont l'effet est réversible (Vlahos et al., 1994). Lorsque les cellules sont traitées, juste avant leur stimulation par du LIF, avec cet agent la phosphorylation d'AKT, GSK3 β et S6 est inhibée.

L'utilisation d'un système inductible tet-off pour l'expression dans les cellules ES d'un mutant dominant-négatif de la sous-unité p85 de la PI3K (Δ p85) a permis de mettre en évidence l'activation de la PI3K en réponse au LIF. Dans ce système, en absence de tétracycline, les cellules expriment le mutant Δ p85 qui ne fixe plus la sous-unité p110 et agit donc comme un inhibiteur compétitif de la sous-unité catalytique de la PI3K endogène. L'expression de Δ p85 entraîne une diminution de la phosphorylation des protéines

PKB/AKT , GSK3 β et S6 : la conséquence d'une sensibilité diminuée au LIF (Paling et al., 2004).

2.2.3.2. Les gènes cibles de la voie PI3K.

2.2.3.2.1. PKB/AKT.

L'une des cibles de la voie PI3K est la sérine/thréonine kinase AKT. La protéine AKT est impliquée dans de nombreux mécanismes tels que la régulation du cycle cellulaire, l'apoptose, l'adhésion, la migration, le métabolisme et la tumorigénicité dans les cellules somatiques (Yang et al., 2004). Récemment, AKT a été montrée comme jouant un rôle dans l'autorenouvellement des cellules ES de souris. Des cellules surexprimant une forme myristoylée de l'AKT, lui conférant ainsi une activité kinase très importante, peuvent être maintenues en autorenouvellement et ceci même en l'absence de LIF (Watanabe et al., 2006). Ces cellules conservent par ailleurs leur propriété de pluripotence jusqu'à 11 jours sans LIF. En effet, elles ont la capacité de former des tératomes composés d'une très grande variété de cellules différenciées après injection dans des souris immunodéprimées. Ces cellules contribuent également au développement de l'embryon après transfert dans des blastocystes, elles sont donc pluripotentes.

2.2.3.2.2. Nanog.

Une étude récente à laquelle nous avons participé (Storm et al., 2007) a mis en évidence le rôle de la voie PI3K dans la régulation de l'expression du gène de pluripotence *Nanog*. En effet, l'inhibition de la PI3K, par utilisation de l'inhibiteur LY294002 ou par la surexpression dans les cellules ES du mutant $\Delta p85$, induit une diminution de l'expression de *Nanog* dans ces cellules. L'expression d'un autre gène de pluripotence, *Oct4*, quant à elle ne diminue pas après inhibition de la voie PI3K. L'expression de *Nanog* dans les cellules ES semble donc être régulée en partie par la voie PI3K. Dans cette étude, des cellules ES de souris surexprimant le mutant hormono-dépendant *Nanog-ER*^{T2} ont été cultivées en présence de LY294002 et en présence de LIF. L'analyse du pourcentage de colonies alcaline phosphatase positives (représentatif de l'état indifférencié des cellules) a permis de conclure que l'inhibition de la PI3K induit la différenciation des cellules ES. La voie PI3Kinase

permettrait ainsi le maintien en autorenouvellement des cellules ES en activant l'expression d'un gène majeur de la régulation de la pluripotence : *Nanog*.

2.2.4. La voie de signalisation Src.

2.2.4.1. La tyrosine kinase cYes.

Les protéines tyrosine kinases de la famille Src régulent de nombreux processus comme la division cellulaire, l'adhésion, la différenciation et la survie cellulaire (Brown and Cooper, 1996) (Kefalas et al., 1995) (Summy and Gallick, 2003). Le rôle de ces protéines Src dans les cellules ES n'est encore pas élucidé. Des études menées par Anneren et ses collègues (Anneren et al., 2004) ont permis d'identifier la tyrosine kinase cYes et de mettre en évidence son rôle dans les cellules ES. La protéine cYes présente une forte activité kinase dans les cellules ES de souris indifférenciées qui diminue au cours de la différenciation induite par la formation de corps embryoïdes. La phosphorylation de la kinase cYes est activée par le LIF et également par le sérum. Cependant, elle ne semble pas être un élément de la voie STAT3, de la voie PI3K ou de la voie SHP2/ERK, comme le montre des expériences réalisées avec un inhibiteur spécifique de cYes, le SU6656 (Blake et al., 2000). C'est pour cette raison que la voie Src est considérée comme une voie de signalisation à part, également activée par le LIF. Des cellules ES cultivées en présence du SU6656, se différencient et ceci même en présence de fortes concentrations de LIF dans le milieu. De plus, l'effet de cet inhibiteur est dose dépendant, puisque plus on augmente la quantité de SU6656 plus l'activité alcaline phosphatase de ces cellules diminue et plus l'expression de *Oct4* est inhibée. Pour confirmer ces résultats, les auteurs ont transfecté des cellules ES de souris cultivées en présence de LIF avec un siRNA spécifique de *cYes*. Les cellules ainsi transfectées se différencient, ce qui se traduit par une diminution du niveau d'expression de *Nanog* et une augmentation du niveau d'expression de *GCNF* (Germ Cell Nuclear Factor), un inhibiteur d'*Oct4* dans les cellules différencierées (Fuhrmann et al., 2001). De plus, les cellules dont l'expression de *cYes* est inhibée, ne peuvent plus former de colonies indifférenciées dans le test d'autorenouvellement (Anneren et al., 2004). Ce test, très sensible, permet de mesurer le degré d'autorenouvellement d'une cellule ES en ensemencant les cellules à une très faible densité dans le but d'éliminer tout facteur paracrine qui pourrait contribuer à leur maintien en autorenouvellement.

La tyrosine kinase cYes semble donc jouer un rôle important dans le maintien en autorenouvellement des cellules ES de souris, puisque son inactivation induit leur différenciation.

2.2.4.2. La tyrosine kinase Hck.

Un autre membre de la famille Src, la protéine Hck, est également impliqué dans l'autorenouvellement des cellules ES de souris. L'activité de Hck est contrôlée négativement par la phosphorylation d'un résidu tyrosine localisé dans la partie C-terminale de la protéine. En effet, le remplacement du résidu conservé Y499 en phénylalanine augmente l'activité tyrosine kinase de Hck d'un facteur 10 à 20. La transfection des cellules ES par l'ADNc du mutant Y499F de Hck diminue d'un facteur 15 la concentration en LIF requise pour maintenir leur phénotype indifférencié (Ernst et al., 1994). Cependant, la surexpression de ce mutant constitutivement actif de Hck n'est pas suffisante pour maintenir les cellules ES dans un état indifférencié en absence de LIF. En conclusion, bien que ces cellules soient plus sensibles à la stimulation par le LIF, le LIF reste toujours nécessaire au maintien de leur état indifférencié.

La voie BMP4/Smad/Id2.

L'activité de la voie de signalisation LIF/gp130/STAT3 permet de maintenir les cellules ES de souris en autorenouvellement uniquement en présence de sérum dans le milieu de culture. En absence de sérum, les cellules ES se différencient spontanément en cellules présentant les caractéristiques du neurectoderme et cela même en présence de LIF (Ying et al., 2003). Un ou plusieurs autres facteurs présents dans le sérum agissent donc en synergie avec le LIF pour maintenir les cellules ES en autorenouvellement et inhiber leur différenciation dans la voie neurectodermique. Des travaux ont montré que la voie de signalisation BMP4/Smad et le facteur HLH Id2 sont vraisemblablement impliqués dans ce processus (Ying et al., 2003).

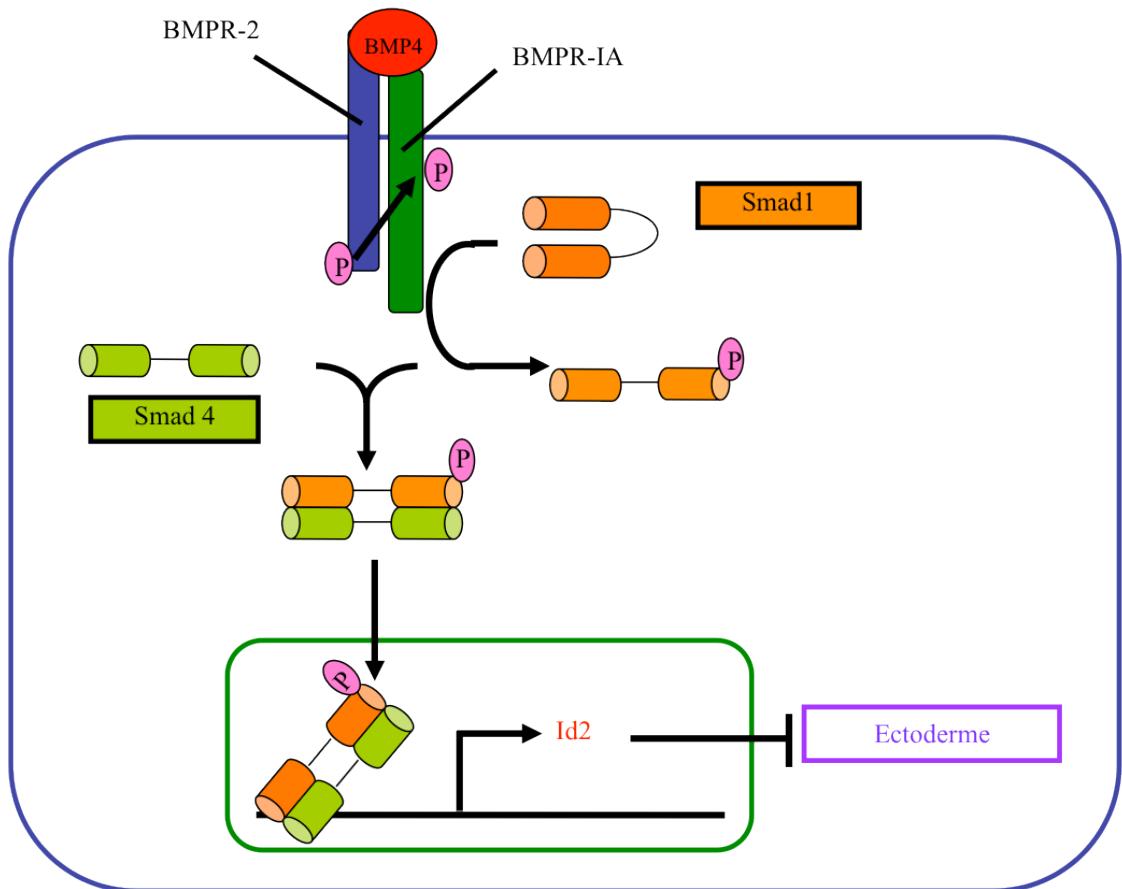


Figure 11 : La voie de signalisation BMP4/Smad/Id2

Le facteur de croissance BMP4 se fixe au niveau du récepteur BMPR-2 qui s'hétérodimérisé ensuite avec le récepteur BMPR-IA. BMPR-IA, activé, phosphoryle alors Smad1 qui va s'hétérodimériser avec Smad-4. Les dimères Smad1/Smad4 vont alors migrer dans le noyau et stimuler l'expression de gènes cibles dont Id2. C'est un facteur de transcription de type HLH (helix-loop-helix) qui agit comme répresseur des facteurs de transcription b-HLH (basic helix-loop-helix). Il va inhiber ainsi la différenciation des cellules ES dans la voie neurectodermique. P : groupement phosphate

3.1. Le facteur de croissance BMP4 inhibe la différenciation neurale des cellules ES.

Le facteur de croissance BMP4 appartient à la superfamille des Transforming Growth Factors- β (TGF- β). BMP4 se fixe au niveau du récepteur BMPR-2 qui s'hétérodimérisé ensuite avec le récepteur BMPR-IA (Miyazawa et al., 2002) (Figure 11). Une fois activé, BMPR-IA phosphoryle et active un transducteur intracellulaire, Smad1, qui s'hétérodimérisé ensuite avec Smad4. Les dimères Smad1/Smad4 migrent alors dans le noyau où ils stimulent l'expression des gènes cibles (Miyazawa et al., 2002).

Initialement les BMPs ont été décrits *in vivo* comme des inducteurs de la formation des tissus osseux et cartilagineux (Wozney and Rosen, 1998). Depuis, de nombreux autres rôles biologiques leur ont été attribués, notamment celui de facteur anti-neural dans les phases précoce du développement embryonnaire (Wilson et al., 2001). Ce rôle d'inhibiteur de la différenciation neurale a été également observé sur les cellules ES de souris. En effet, l'ajout de BMP4 dans le milieu de culture leur permet de s'autorenouveler sous l'action du LIF et en absence de sérum (Ying et al., 2003). La suppression du BMP4 induit leur différenciation spontanée en neurectoderme (Ying et al., 2003). BMP4 apparaît donc comme un inhibiteur de la différenciation neurale des cellules ES de souris.

Il est important de noter que le LIF et le BMP4 exercent des actions complémentaires. Le LIF inhibe la différenciation en mésoderme et en endoderme (Smith et al., 1988) alors que le BMP4 inhibe la différenciation neurale (Ying et al., 2003). Aucun de ces deux facteurs ne peut compenser l'absence de l'autre.

3.2. Le facteur BMP4 inhibe la différenciation en activant l'expression du gène Id2.

Les facteurs de la famille Id inhibent la différenciation cellulaire et activent la prolifération cellulaire (Miyazawa et al., 2002). Ce sont des facteurs de transcription de type HLH (helix-loop-helix). Ils agissent comme répresseurs des facteurs de transcription bHLH (basic helix-loop-helix) (Miyazawa et al., 2002). Parmi les gènes de la famille Id, le gène *Id2* est fortement exprimé dans les cellules ES de souris et son expression décroît après différenciation. En outre, l'expression des gènes *Id* est activée par BMP4 dans de nombreuses lignées cellulaires, y compris les cellules ES de souris (Ying et al., 2003).

Le facteur *Id2* est-il un médiateur de l'action de BMP4 dans les cellules ES ? Ying et ses collègues ont montré que des cellules ES surexprimant le gène *Id2* peuvent être maintenues en autorenouvellement seulement en présence de LIF, sans ajout de sérum ou de

BMP4 dans le milieu de culture (Ying et al., 2003). Ainsi, la surexpression du facteur Id2 est suffisante pour inhiber la différenciation neurale des cellules ES. En revanche, la surexpression du facteur Id2 ne permet pas d'inhiber la différenciation induite par la privation de LIF, ce qui confirme que la voie LIF/gp130/STAT3 et la voie BMP4/Smad/Id2 constituent deux voies indépendantes et non redondantes.

En plus des deux voies de signalisation décrites précédemment, deux autres facteurs de transcription, Oct4 et Nanog, sont également impliqués dans le maintien des cellules ES en autorenouvellement. Ces facteurs de transcription ne sont régulés par aucune voie de signalisation connue. Dans l'état actuel des connaissances, ils sont considérés comme des facteurs de transcription intrinsèques, c'est-à-dire constitutivement actifs.

Le facteur de transcription Oct4.

4.1. Profil d'expression du facteur Oct4 au cours du développement embryonnaire.

Le facteur Oct4 appartient à la famille des facteurs de transcription à domaine POU pouvant se fixer sur l'ADN au niveau de séquences octamériques riches en AT (Pan et al., 2002). Chez la souris, l'expression de *Oct4* est d'abord détectée dans l'ovocyte et dans l'embryon pré-implantatoire jusqu'au stade morula (Pan et al., 2002). Son expression se restreint ensuite aux cellules de l'ICM, puis à celles de l'ectoderme primitif de l'embryon post-implantatoire. Son expression s'éteint alors dans les trois feuillets embryonnaires (ectoderme, mésoderme et endoderme). Elle est néanmoins maintenue dans les cellules germinales primordiales. *Oct4* est également exprimé dans les cellules ES. Son expression diminue rapidement après différenciation (Pan et al., 2002). Ce profil d'expression, spécifique des cellules toti- et pluripotentes, suggère que le facteur de transcription Oct4 joue un rôle important dans le maintien de la pluripotence cellulaire.

L'inactivation du gène Oct4 induit une létalité embryonnaire précoce : l'embryon pré-implantatoire se développe jusqu'au stade de la morula ; au-delà, il dégénère car les cellules fondatrices de l'ICM sont incapables de s'autorenouveler. Elles se différencient alors en trophectoderme (Rosner et al., 1990). On pense que l'embryon est capable de se développer jusqu'au stade morula en utilisant le stock d'ARN et de protéines d'origine maternelle et

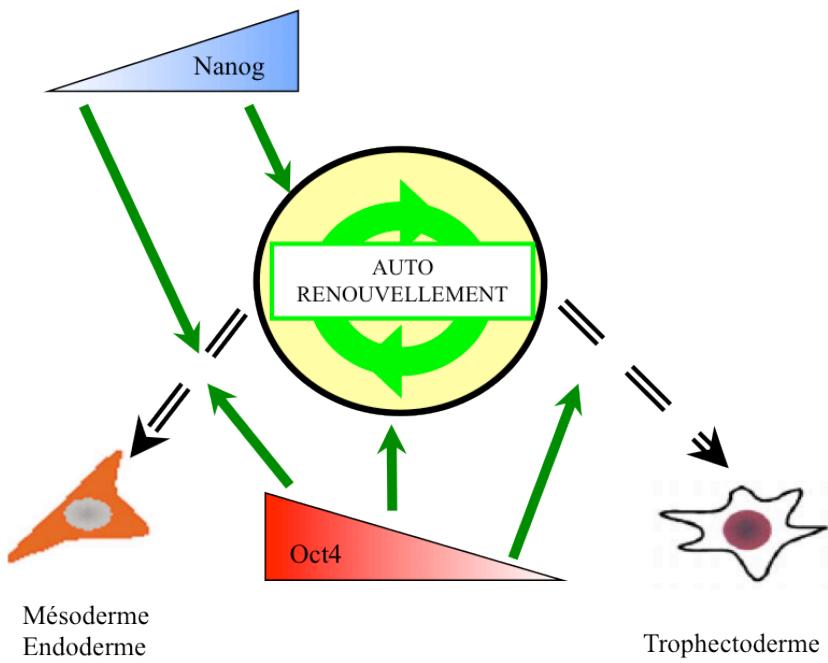


Figure 12 : Les facteurs de transcription intervenant dans le contrôle de l'autorenouvellement des cellules ES

Oct4 : Dans des cellules ES en autorenouvellement, le niveau d'expression d'*Oct4* doit être maintenu à l'intérieur d'une fenêtre très étroite. Une diminution de 50% induit la différenciation des cellules en trophectoderme. Une augmentation d'un facteur 2 entraîne leur différenciation en endoderme et mésoderme.

Nanog : Nanog est un facteur de transcription nécessaire au maintien en autorenouvellement des cellules ES. Son inactivation induit la différenciation des cellules ES en endoderme et mésoderme.

qu'en l'absence de transcrits Oct4 d'origine zygotique, le développement normal cesse lors de la transition morula/blastocyste.

4.2. Rôle du facteur Oct4 dans le contrôle de l'autorenouvellement des cellules ES.

Afin d'étudier plus finement le rôle du gène *Oct4* dans l'autorenouvellement des cellules ES de souris *in vitro*, Niwa et ses collègues ont construit une lignée de cellules dans laquelle le gène *Oct4* est inactivé et remplacé par un transgène dont l'expression est activable ou réprimable par la présence ou l'absence de tétracycline dans le milieu (système d'expression conditionnelle tet-off) (Niwa et al., 2000). Ce système permet de contrôler très précisément le niveau d'expression du gène *Oct4* dans les cellules en fonction de la concentration d'inducteur utilisée. Grâce à ce système d'expression, les auteurs ont pu montrer que l'autorenouvellement des cellules ES nécessite le maintien du niveau d'expression de la protéine Oct4 à l'intérieur d'une fenêtre étroite de concentration. Une diminution de moitié de *Oct4* induit la différenciation des cellules en trophectoderme. De façon plus inattendue, une augmentation du niveau d'expression de *Oct4* d'un facteur deux induit également une différenciation, dans ce cas dans la voie de l'endoderme et du mésoderme (Figure 12). Ainsi, un niveau d'expression très précis est nécessaire au maintien des cellules ES en autorenouvellement. Une élévation ou une baisse de ce niveau induisent des programmes de différenciation divergents.

4.3. Le co-facteur de Oct4 : Sox2.

Le facteur Oct4 peut activer ou réprimer la transcription en interagissant avec le co-facteur Sox2 (SRY-related HMG box 2) (Pesce and Scholer, 2001). L'inactivation de ce co-facteur induit un phénotype similaire à celui observé pour l'inactivation du gène *Oct4* lui-même, c'est-à-dire l'arrêt du développement embryonnaire au stade blastocyste. Des cellules ES de génotype *Sox2^{+/−}* peuvent être isolées à partir d'embryons hétérozygotes mais il est impossible d'inactiver le second allèle (Hanna et al., 2002) (Avilion et al., 2003). On peut donc en conclure que le co-facteur Sox2 joue un rôle essentiel dans le maintien des cellules ES de souris en autorenouvellement.

Le facteur de transcription Nanog.

5.1. Identification du gène Nanog.

Deux laboratoires ont identifié un facteur de transcription appelé Nanog. Mitsui et ses collègues ont réalisé un séquençage exhaustif des transcrits différentiellement exprimés dans les cellules ES et les cellules différencierées (Mitsui et al., 2003). Ils ont ainsi identifié plusieurs séquences fortement représentées dans les cellules ES. L'analyse fonctionnelle de ces gènes a permis d'identifier le gène *Nanog* dont la surexpression inhibe la différenciation des cellules ES de souris induite par la privation de LIF. Chambers et ses collègues ont utilisé une approche de clonage fonctionnel (Chambers et al., 2003). Ils ont mis au point un système d'expression épisomal qu'ils ont utilisé pour transfacter les cellules ES avec une banque d'ADNc. Ils ont ensuite sélectionné les clones capables de s'autorenouveler en l'absence de LIF. La caractérisation d'ADNc a permis d'identifier un nouveau facteur de transcription à homéodomaine appelé Nanog.

5.2. Expression du gène Nanog et rôle dans l'autorenouvellement des cellules souches embryonnaires *in vivo* et *in vitro*.

Au cours du développement embryonnaire chez la souris, l'expression du gène *Nanog* apparaît dans les cellules internes de la morula à l'origine de l'ICM (Chambers et al., 2003). Les transcrits *Nanog* sont ensuite détectés dans l'ICM au stade blastocyste puis dans l'épiblaste lors de l'implantation. L'expression de *Nanog* disparaît rapidement après l'implantation. Le développement des embryons *Nanog*^{-/-} s'interrompt avant l'implantation. Les cellules fondatrices de l'ICM se différencient en endoderme pariétal et l'embryon arrête son développement (Mitsui et al., 2003).

Nanog est également exprimé dans les cellules ES de souris (Mitsui et al., 2003) (Chambers et al., 2003). Son expression s'éteint lors de la différenciation. L'inactivation du gène *Nanog* induit la différenciation spontanée des cellules ES en cellules de l'endoderme extra-embryonnaire (Mitsui et al., 2003) (Figure 12).

Ces données mettent en évidence le rôle du gène *Nanog* dans le contrôle de l'autorenouvellement des cellules de l'ICM *in vivo* et des cellules ES *in vitro*.

5.3. Relations fonctionnelles entre les facteurs Nanog, STAT3, Oct4, Sox2 et Sall4.

5.3.1. Nanog et STAT3.

La surexpression de *Nanog* permet de stimuler l'autorenouvellement des cellules ES de souris en l'absence de LIF, suggérant l'existence d'une redondance fonctionnelle entre *Nanog* et *STAT3* (Mitsui et al., 2003) (Chambers et al., 2003). Cette redondance est néanmoins trompeuse. Les cellules ES de souris requièrent à la fois l'expression du gène *Nanog* endogène et l'activité de la voie LIF/gp130/STAT3 pour s'autorenouveler efficacement (Chambers et al., 2003). La privation de LIF ou l'inactivation de *Nanog* est suffisante pour induire leur différenciation. Seule une très forte surexpression de *Nanog* parvient à suppléer, quoiqu'imparfaitement, l'absence de LIF. Il semble donc que *STAT3* et *Nanog* activent l'expression de gènes cibles différents. L'activité de l'ensemble de ces gènes cibles est nécessaire au maintien des cellules ES en autorenouvellement. *Nanog* serait également capable d'activer les gènes cibles de *STAT3* mais seulement dans des conditions particulières de surexpression expérimentale (Chambers et al., 2003). Notons aussi que l'expression de *Nanog* n'est pas régulée par l'activité de la voie LIF/gp130/STAT3 (Chambers et al., 2003). Le gène *Nanog* ne serait donc pas un gène cible de *STAT3* dans les cellules ES. En conclusion, il est impossible de parler de véritable redondance fonctionnelle entre *Nanog* et *STAT3*, même si *Nanog* est capable de suppléer la fonction de *STAT3* dans certaines conditions expérimentales extrêmes.

Une étude très récente, réalisée par Torres and Watt, montre que *Nanog* coopérerait en fait avec *STAT3*, dans des conditions normales de culture, pour maintenir la pluripotence des cellules ES. En effet, *Nanog* et *STAT3* se fixeraient sur les régions régulatrices de mêmes gènes (Torres and Watt, 2008).

5.3.2. Nanog, Oct4 et Sox2.

Les facteurs de transcription *Nanog*, *Oct4* et *Sox2* sont considérés comme les gènes « maîtres » conférant aux cellules ES leur identité. Ces facteurs ont un profil d'expression qui leur est unique et jouent un rôle essentiel au cours du développement embryonnaire précoce (Avilion et al., 2003) (Chambers et al., 2003) (Mitsui et al., 2003) (Scholer et al., 1990) (Nichols et al., 1990) (Hart et al., 2004) (Lee et al., 2004).

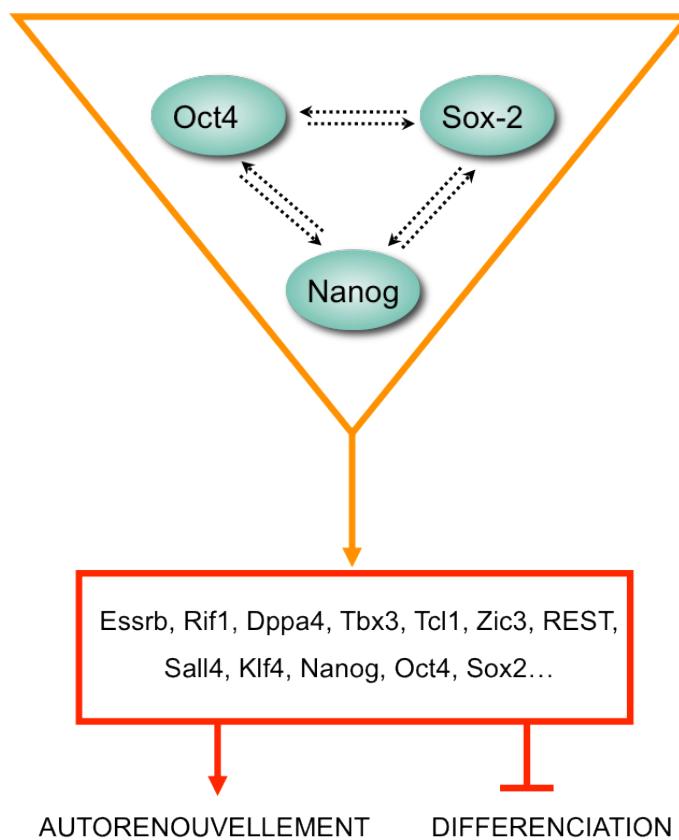


Figure 13 : Réseau de régulation transcriptionnelle mis en place par Nanog, Oct4 et Sox2 dans les cellules souches embryonnaires

Les facteurs de transcription Nanog, Oct4 et Sox2 permettent le maintien de la pluripotence des cellules ES en régulant l'expression de gènes cibles communs qui jouent un rôle important dans le maintien de l'autorenouvellement et l'inhibition de la différenciation. En effet, leur inactivation induit la différenciation des cellules ES alors que leur surexpression l'inhibe. Ce réseau de régulation reste encore bien incomplet.

Afin de mieux comprendre le rôle de ces facteurs clés dans le maintien de la pluripotence et dans l'autorenouvellement et l'inhibition de la différenciation des cellules ES, deux groupes ont réalisé des expériences d'immunoprécipitation de la chromatine (ChIP) à partir de cellules ES de souris et humaines (Loh et al., 2006) (Boyer et al., 2005) afin d'identifier les gènes cibles de Nanog, Oct4 et Sox2.

Ces études ont montré que Nanog, Oct4 et Sox2 se fixent sur les éléments régulateurs des mêmes gènes. En effet, Loh et ses collègues ont montré que, dans les cellules ES de souris, Nanog se fixe sur le promoteur de 44,5% des gènes cibles de Oct4 soit 345 gènes. D'après l'étude de Boyer et ses collègues, dans les cellules ES humaines 353 gènes possèdent des sites de fixation pour NANOG, OCT4 et SOX2. Ces observations suggèrent que les trois facteurs de transcription Nanog, Oct4 et Sox2 collaborent pour réguler des gènes cibles communs (Boyer et al., 2005) (Loh et al., 2006) et forment ainsi un réseau de régulation transcriptionnelle.

Cependant, les expériences de ChIP ne permettent pas de déterminer la présence d'une régulation fonctionnelle au niveau transcriptionnel. Afin de déterminer quels sont les gènes fonctionnellement régulés par Nanog, Oct4 et Sox2, Loh et ses collègues ont réalisé des expériences d'inhibition spécifique de l'expression de *Nanog* ou de *Oct4* par interférence de l'ARN et analysé, par la technique des puces à ADN, les variations des niveaux d'expression de leurs gènes cibles (Loh et al., 2006). Ils ont ainsi identifié dans les cellules ES de souris 77 gènes régulés par Nanog et par Oct4, via des sites de fixation pour ces facteurs.

Parmi les gènes cibles de Nanog, Oct4 et Sox2, trois sont particulièrement intéressants à savoir *Nanog*, *Oct4* et *Sox2* eux-mêmes (Boyer et al., 2005) (Loh et al., 2006). De nombreuses études avaient déjà mis en évidence la régulation de *Nanog*, *Oct4* et *Sox2* par le complexe Oct4/Sox2 (Catena et al., 2004) (Kuroda et al., 2005) (Okumura-Nakanishi et al., 2005) (Rodda et al., 2005). Les expériences de Boyer et ses collègues et de Loh et ses collègues ont permis de mettre en évidence le rôle joué par Nanog en coopération avec Oct4 et Sox2. Ainsi, l'expression et la fonction de ces trois facteurs « maîtres » sont fortement liés les uns aux autres. En régulant leur propre gène Nanog, Oct4 et Sox2 forment une boucle d'autorégulation qui permet de maintenir la stabilité du phénotype indifférencié des cellules ES ainsi que leur pluripotence (Figure 13).

De nombreuses études fonctionnelles ont permis de mettre en évidence qu'une partie des gènes présentant des sites de fixation pour Nanog, Oct4 et Sox2, en dehors de *Nanog*, *Oct4* et *Sox2* eux-mêmes dont on sait qu'ils sont importants, est directement impliquée dans le contrôle de l'autorenouvellement et la pluripotence des cellules souches embryonnaires. En

effet, (i) l'inactivation des gènes *Esrrb* (Estrogen-related receptor β) et *Rif1* (Rap1 interacting factor 1 homolog), par la technique d'interférence à l'ARN, induit la différenciation des cellules ES (Loh et al., 2006) (Ivanova et al., 2006) ; (ii) Dppa4 (Developmental pluripotency-associated 4) inhibe la différenciation des cellules ES en ectoderme primitif (Masaki et al., 2007) ; (iii) l'inhibition de l'expression de *Tbx3* (T-box 3) et *Tcl1* (T-cell lymphoma breakpoint 1) induit également la différenciation des cellules ES (Ivanova et al., 2006), quant à leur surexpression elle inhibe leur différenciation dans le lignage mésodermique pour *Tbx3* et dans les trois lignages endoderme, mésoderme et ectoderme pour *Tcl1* ; (iv) l'inactivation de *Zic3* (Zinc finger protein of the cerebellum) induit la différenciation des cellules ES en cellules de type endodermique (Lim et al., 2007) ; (v) le répresseur neuronal REST (RE1-silencing transcription factor) quant à lui permettrait le maintien en autorenouvellement des cellules ES en inhibant le micro-ARN mir-21 dont l'expression induit la différenciation des cellules ES (Singh et al., 2008).

Nous avons cité plus haut les gènes qui sont les plus récurrents lorsque l'on compare les nombreux travaux qui ont été réalisés dans le but d'identifier les gènes qui sont primordiaux pour la pluripotence des cellules ES (Boyer et al., 2005) (Loh et al., 2006) (Ivanova et al., 2006) (Matoba et al., 2006) (Sharov et al., 2008) (Kim et al., 2008).

Nanog, *Oct4* et *Sox2* permettent donc le maintien des cellules ES en autorenouvellement et leur pluripotence en régulant l'expression de gènes communs (Figure 13) qui jouent un rôle important dans le maintien de la pluripotence puisque leur inactivation induit la différenciation des cellules ES alors que leur surexpression inhibe leur différenciation. Ce réseau de régulation transcriptionnelle reste cependant encore bien incomplet et de nombreux éléments s'y ajoutent régulièrement.

5.3.3. *Nanog et Sall4.*

Sall4 appartient à la famille des protéines Spalt-like. Ce gène est fortement exprimé dans l'ICM de l'embryon précoce de souris (Yoshikawa et al., 2006) ainsi que dans les cellules EC (Kohlhase et al., 2002). L'inactivation du gène *Sall4* dans les cellules ES induit une diminution de l'expression des gènes caractéristiques des cellules indifférenciées et une augmentation de l'expression de marqueurs de différenciation. *Sall4* est donc important pour le maintien des cellules ES en autorenouvellement (Wu et al., 2006). C'est en essayant d'identifier les facteurs qui interagissent avec *Nanog*, par une technique de purification par

affinité, que Wu et ses collègues ont mis en évidence le facteur Sall4. Des expériences de co-immunoprécipitation ont confirmé l'interaction entre Nanog et Sall4 (Wu et al., 2006).

Nanog et Sall4 se fixent réciproquement sur les régions régulatrices de leur propre gène. En étudiant 20 autres sites de fixation à l'ADN de Nanog, mis en évidence dans l'étude de Loh et al. (Loh et al., 2006), ils ont découvert que Sall4 se fixe sur les mêmes sites de régulation. Sall4 se fixe entre autre sur les régions régulatrices de *Tbx3*, *Essrb*, *Sox2* mais également sur *Oct4* dont il régule l'expression dans les cellules ES afin de maintenir leur propriété de pluripotence (Zhang et al., 2006).

Des études génomiques à grande échelle ont donc permis de mettre en évidence un réseau de régulation transcriptionnelle dont les principaux acteurs sont Nanog, Oct4 et Sox2. L'analyse du transcriptome de ces facteurs a permis de mettre en évidence un grand nombre de gènes en commun dont le rôle est d'inhiber la différenciation des cellules ES. Ce réseau constitue en lui-même une boucle d'auto-régulation puisque Nanog, Oct4 et Sox2 régulent chacun leur propre expression ainsi que l'expression des deux autres facteurs, ce qui permet d'entretenir le système et de maintenir la stabilité du phénotype indifférencié ainsi que la pluripotence des cellules ES.

Le contrôle de l'autorenouvellement, de l'inhibition de la différenciation et du maintien de la pluripotence des cellules ES repose sur de nombreux éléments : (i) des mécanismes extrinsèques comme la voie de signalisation LIF/gp130/STAT3 dont l'inactivation induit la différenciation des cellules ES dans les lignages endomésodermiques (Smith et al., 1988) et la voie BMP4/Smad2/Id2 dont l'inactivation, au contraire, induit la différenciation des cellules ES de souris dans le lignage neurectodermique (Ying et al., 2003) ; (ii) mais également des mécanismes intrinsèques comme le facteur de transcription Nanog dont l'inhibition induit la différenciation des cellules dans le lignage mésodermique (Mitsui et al., 2003) et Oct4 dont l'inactivation induit la différenciation en cellules trophoblastiques (Niwa et al., 2000) (Figure 12). Ces observations nous amènent à considérer l'autorenouvellement des cellules ES comme le résultat de l'inhibition de tous les programmes de différenciation possibles, grâce à l'action conjointe de plusieurs voies de signalisation et de facteurs de transcription.

Les sérine/thréonine kinases Pim1, Pim2 et Pim3.

6.1. Identification.

Pim1, Pim2 et Pim3 font partie d'une famille de sérine/thréonine kinases dont les membres sont très conservés chez les vertébrés. Le gène *Pim1* a été identifié comme un site fréquent d'insertion proviral du virus leucémogène murin de Moloney ou MoMuLV (Moloney Murine Leukemia Virus) dans des expériences réalisées afin d'identifier de nouveaux gènes impliqués dans la tumorigénèse (Allen and Berns, 1996). La recherche de séquences homologues à Pim1 a permis de mettre en évidence l'existence de deux autres kinases : Pim2 et Pim3 (Allen and Berns, 1996) (Baytel et al., 1998) (Feldman et al., 1998) (van der Lugt et al., 1995).

Pim1, *Pim2* et *Pim3* présentent une forte homologie aussi bien dans leur séquence que dans leur structure, ce qui leur confère des rôles souvent identiques et donc redondants (Mikkers et al., 2004). En effet, des études ont montré que Pim2 et Pim3 pouvaient remplacer Pim1, entre autre, dans le processus de lymphomagénèse induit par le virus MoMuLV, mais également dans de nombreux autres processus comme la régulation du cycle cellulaire ou la survie des cellules (van der Lugt et al., 1995) (Allen et al., 1997) (Mikkers et al., 2002) (Yan et al., 2003) (Li et al., 2006) (Morishita et al., 2008).

6.2. Profil d'expression des kinases Pim.

Les kinases Pim sont exprimées dans un grand nombre de tissus embryonnaires. A elle trois elles couvrent presque tous les tissus de l'embryon. Pim1 est plus fortement présente dans le thymus et les testicules (Selten et al., 1985), Pim2 dans le cerveau et dans le thymus (Allen et al., 1997) et enfin Pim3 dans les reins (Feldman et al., 1998). Ces kinases sont également exprimées dans de nombreuses lignées cellulaires dérivées de tumeurs humaines comme les cellules HeLa, les cellules K562 issues de patients atteints de leucémie myeloïde chronique ou d'autres lignées issues de tumeurs lymphoïdes ou prostatiques (Bachmann and Moroy, 2005) (Eichmann et al., 2000).

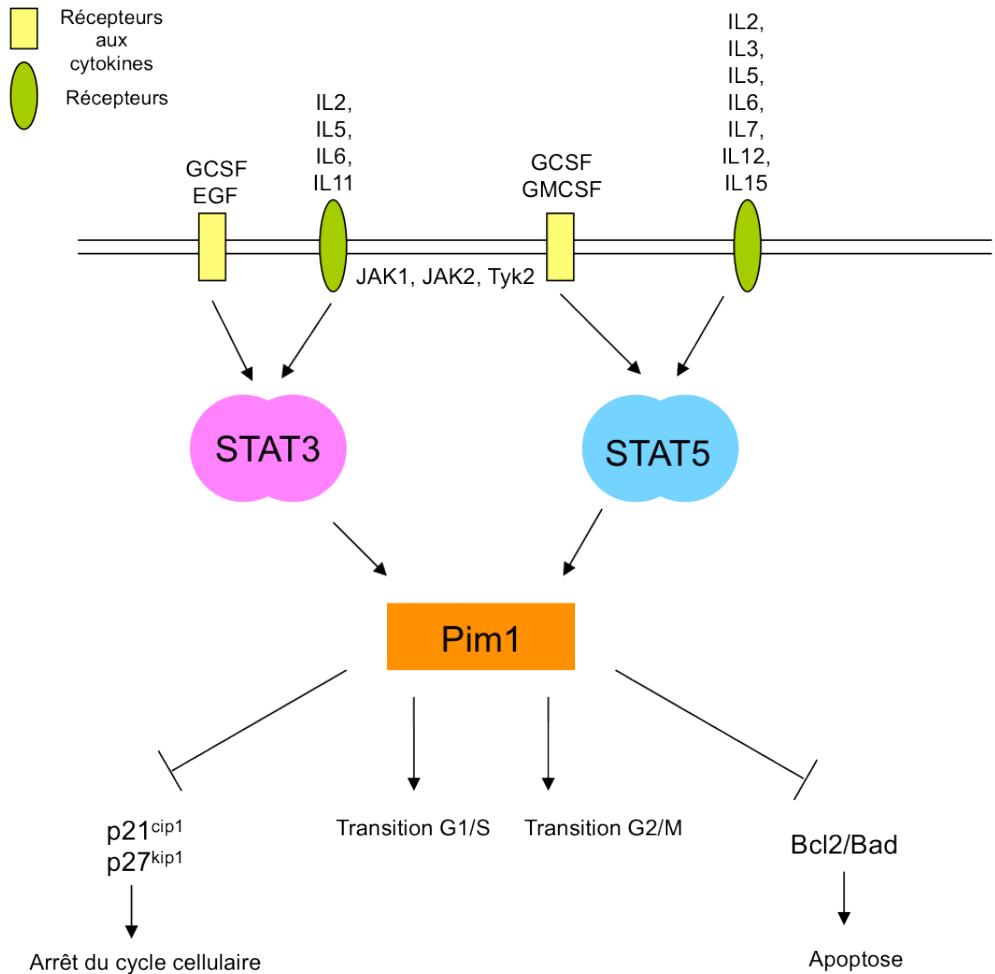


Figure 14 : Régulation de la séro-thréonine kinase Pim1

L'expression de *Pim1* est directement régulée par les facteurs de transcription STAT3 et STAT5. Les protéines STAT sont activées par les kinases JAK (Janus Kinase) après fixation sur les récepteurs aux cytokines d'Interleukines ou de Facteurs de croissance. Pim1 phosphoryle alors des protéines cibles qui interviennent dans la régulation du cycle cellulaire et l'apoptose pour les activer ou les inhiber.

6.3. Les membres de la famille Pim répondent aux voies de signalisation activées par des cytokines.

6.3.1. L'expression des kinases Pim est activée par un grand nombre de cytokines.

Pim1 a été décrit comme étant l'effecteur de plusieurs voies de signalisation dépendantes des cytokines pour leur activation. Son expression, comme celle de *Pim2*, peut être induite par un grand nombre de cytokines telles que l'IL-2, l'IL-3, le GCSF, l'IL-4, l'IL-5, l'IL-6, l'IL-7, l'érythropoïétine, la prolactine et bien d'autres encore (Allen et al., 1997) (Dautry et al., 1988) (Lilly et al., 1992) (Temple et al., 2001) (Telerman et al., 1988) (Borg et al., 1999) (Figure 14). Deux médiateurs de ces voies de signalisation dépendantes des cytokines sont particulièrement importants pour l'expression de *Pim1*. Il s'agit des facteurs de transcription STAT3 et STAT5 qui régulent l'expression de *Pim1* en se fixant directement sur son promoteur (Losman et al., 1999) (Shirogane et al., 1999).

6.3.2. La régulation de Pim1 par STAT3.

La sous-unité gp130 est nécessaire à la transduction du signal envoyé par les cytokines de la famille IL6. Afin d'identifier les gènes activés par cette sous-unité réceptrice, des cellules de la lignée BAF/B03 ont été utilisées. Ces cellules ont été modifiées de façon à exprimer à leur surface un récepteur chimérique constitué de la partie extracellulaire du récepteur GCSFR et de la région transmembranaire et cytoplasmique de gp130. *Pim1* a été identifié comme étant l'un des gènes cibles de la signalisation gp130 ; en effet, après stimulation du récepteur avec du GCSF le niveau d'expression de *Pim1* augmente (Shirogane et al., 1999).

Lorsque la sous-unité gp130 est engagée, de nombreuses molécules s'associent à elle afin de permettre la transduction du signal telles que les kinases JAK, les protéines SHP2/ERK et STAT3. Ainsi, des cellules exprimant différents mutants de gp130 ont été générées dont (i) le mutant G68 qui permet l'activation des kinases JAK mais pas de la voie SHP2/ERK ni de STAT3, (ii) le mutant G133F2 dans lequel la tyrosine Y759 a été remplacée par une phénylalanine (Y759F) bloquant ainsi l'activation de la voie SHP2/ERK mais pas de la voie STAT3, et enfin (iii) le mutant G133F3-Y767F qui lui ne peut pas activer STAT3 mais qui active la voie SHP2/ERK. Après stimulation de ces différentes lignées cellulaires, seules les cellules avec le récepteur G133F2 (qui n'active donc que STAT3) voient leur niveau d'expression de *Pim1* augmenter (Shirogane et al., 1999). STAT3 régule donc

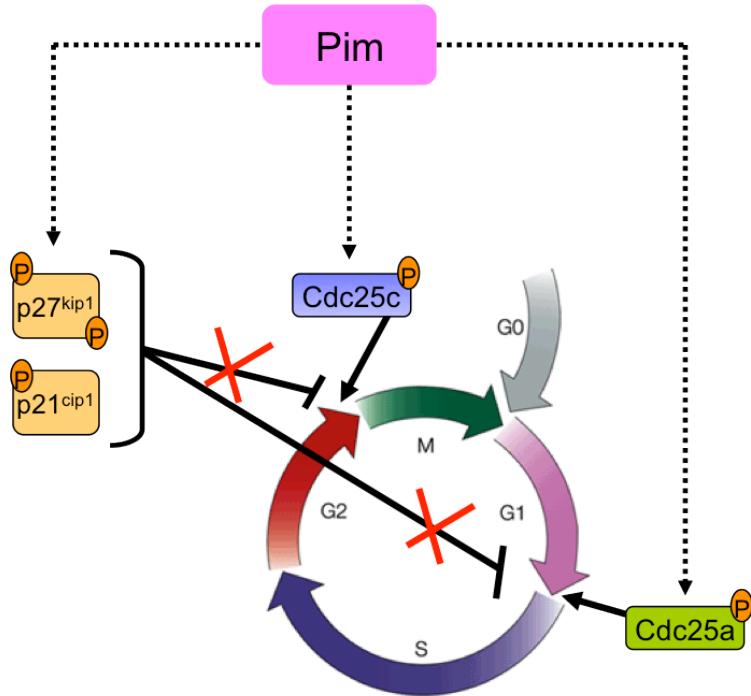


Figure 15 : Le rôle des kinases Pim dans la régulation du cycle cellulaire

Les kinases Pim régulent le cycle cellulaire en contrôlant les phases de transition G1/S et G2/M. La phosphorylation par les kinases Pim de Cdc25a induit son activation et promeut la transition de la phase G1 à la phase S et permet ainsi la progression des cellules au cours du cycle. Les kinases Pim phosphorylent également la phosphatase Cdc25c dont l'activation facilite la transition G2/M du cycle. Les protéines p27^{kip1} et p21^{cip1}, deux inhibiteurs majeurs de la transition G2/M et G1/S, sont phosphorylées par les kinases Pim ce qui induit la dégradation de p27^{kip1} par la machinerie protéasomique et la translocation de p21^{cip1} du noyau vers le cytoplasme. Ainsi, les cellules pourront transiter de la phase G2 à la phase M ainsi que de la phase G1 à la phase S..

l'expression de *Pim1*. Ce résultat a été confirmé par d'autres méthodes, notamment l'utilisation d'un mutant STAT3 dominant-négatif (STAT3-DN). Les cellules exprimant STAT3-DN ne peuvent plus induire l'expression de *Pim1* après stimulation. STAT3 régule donc l'expression de *Pim1* et ceci en se fixant directement sur son promoteur : *Pim1* est un gène cible direct de STAT3 (Shirogane et al., 1999).

6.4. La fonction des kinases Pim.

Les souris triple « knock-out » pour Pim1, Pim2 et Pim3 sont viables mais présentent une importante réduction de leur taille de l'ordre de 30%, ce qui suggère un rôle de ces kinases dans la prolifération cellulaire, ainsi qu'une diminution de leur durée de vie (Mikkers et al., 2004).

6.4.1. La régulation du cycle cellulaire.

6.4.1.1. *Cdc25a, Cdc25c.*

Les études de Losman et ses collègues en 1999 ont été les premières à mettre en évidence le rôle de Pim1 dans la régulation du cycle cellulaire. En effet, elles ont montré que Pim-1 se fixe sur la phosphatase Cdc25a et la phosphoryle. Cdc25a est un régulateur du cycle cellulaire qui intervient durant la phase G1. Pim1, en phosphorylant, Cdc25a induit une augmentation importante de l'activité phosphatase de ce dernier, ce qui a pour effet de faire transiter les cellules de la phase G1 à la phase S de réPLICATION (Figure 15).

Une autre protéine, la phosphatase Cdc25c, a également été identifiée comme étant un substrat de la kinase Pim1 (Bachmann et al., 2006). Celle-ci joue un rôle important dans la transition de la phase G2 à la phase M de mitose du cycle cellulaire (Figure 15).

Pim1 joue donc un rôle important dans la régulation du cycle cellulaire et plus particulièrement au niveau des transitions G1/S et G2/M.

6.4.1.2. *p27^{kip1}.*

L'inhibiteur des Cyclines Dépendantes Kinases (CDK) : p27^{kip1} est une molécule clé dans la régulation du cycle cellulaire. Son inactivation permet la progression de la cellule à

travers la phase G1 puis à travers la transition G1/S (Fujita et al., 2002) (Fujita et al., 2003) (Figure 15).

Les kinases Pim1, Pim2 et Pim3 se fixent directement sur la protéine p27^{kip1} et induisent sa phosphorylation sur les thréonines T157 et T198 (Morishita et al., 2008). La protéine p27^{kip1} ainsi phosphorylée est dégradée par la machinerie protéasomique et permet donc la progression des cellules dans le cycle cellulaire. Les kinases Pim inhibent également l'expression de *p27^{kip1}* au niveau transcriptionnel. Elles phosphorylent pour cela les facteurs de transcription Forkhead (FoxO) qui sont connus pour stimuler l'expression de *p27^{kip1}* (Brunet et al., 1999) (Brunet et al., 2001). Suite à cette phosphorylation, les facteurs FoxO sont inactivés et l'expression de *p27^{kip1}* est donc inhibée.

L'inhibition de p27^{kip1}, par les kinases Pim, aussi bien au niveau post-traductionnel qu'au niveau transcriptionnel, constitue l'un des mécanismes les plus importants dans la progression des cellules dans le cycle cellulaire.

6.4.1.3. *p21^{cip1}*.

Une étude réalisée par Zhang et ses collègues a mis en évidence le rôle de Pim1 dans la régulation d'un autre inhibiteur des CDKs : p21^{cip1} (Zhang et al., 2007). Lorsque p21^{cip1}, est situé dans le noyau, il a pour rôle de bloquer le cycle cellulaire en inhibant la transition de la phase G1 à la phase S. Pim1, en phosphorylant p21^{cip1}, induit la translocation de ce dernier du noyau vers le cytoplasme et promeut ainsi la prolifération cellulaire (Zhang et al., 2007).

6.4.2. *L'apoptose.*

La surexpression de *Pim1* dans la lignée FDC-P1 de progéniteurs monocytaires, dépendante de l'interleukine 3 (IL-3) pour leur prolifération, inhibe l'apoptose induite par privation d'IL-3 dans le milieu (Lilly and Kraft, 1997). Pim1 semble donc avoir un rôle anti-apoptotique en plus de celui de régulateur du cycle cellulaire. Mais comment Pim1 joue-t-il ce rôle ?

Le proto-oncogène c-Myc joue un rôle important dans la régulation de la croissance cellulaire et de la prolifération (Classon et al., 1987) (Resnitzky and Kimchi, 1991) (Iguchi-Ariga et al., 1987). Cependant, des études ont montré que la surexpression de *c-Myc* pouvait induire l'apoptose dans des cellules privées de sérum ou de facteurs de croissance (Askew et

al., 1991) (Evan et al., 1992) (Shi et al., 1992). En effet, la surexpression de *c-Myc* dans des fibroblastes de rat cultivés en absence de sérum induit une augmentation de plus de 15 fois du nombre de cellules mortes (Wagner et al., 1993). La co-expression avec *c-Myc* d'un autre proto-oncogène, *Bcl-2* (B-cell leukemia/lymphoma 2), bloque l'apoptose (Bissonnette et al., 1992) (Wagner et al., 1993). En effet, *Bcl-2* a été décrit dans de nombreuses études comme jouant un rôle anti-apoptotique dans de multiples types cellulaires (Vaux et al., 1988) (Borzillo et al., 1992).

Des études dans différents types cellulaires ont montré que la surexpression de *Pim1* induisait une augmentation de l'expression de *Bcl-2* (Sakai and Kraft, 1997) (Krumenacker et al., 1998). En fait, *Pim1* régule *Bcl-2* en inhibant la protéine pro-apoptotique *Bad* (*Bcl2-associated agonist of cell death*) (Aho et al., 2004), ce qui induit une augmentation de l'activité de *Bcl-2* et promeut donc la survie cellulaire (Figure 15). *Pim2* et *Pim3*, au même titre que *Pim1*, régulent l'activation de *Bad* et jouent donc un rôle anti-apoptotique (Yan et al., 2003) (Li et al., 2006).

De forts niveaux d'expression des kinases Pim ainsi que des mutations au niveau de leurs gènes respectifs ont été observés dans de nombreux cancers. Ces sérine/thréonine kinases pourraient ainsi être des cibles idéales dans la recherche de molécules pharmacologiques à but thérapeutique (Amaravadi and Thompson, 2005). Elles ont un rôle important à la fois dans la prolifération des cellules et dans leur survie. Elles régulent le cycle cellulaire au niveau des points de contrôles des phases de transition G1/S et G2/M et permettent la survie des cellules en inhibant des facteurs pro-apoptotiques.

Les Krüppel-Like Factors ou Klfs.

Les membres de la famille des Krüppel-like factors (Klfs) sont des facteurs de transcription présentant une forte homologie au niveau des sites de fixation à l'ADN avec le gène de segmentation *Krüppel* identifié chez la Drosophile, d'où leur appellation. Ces facteurs de transcription sont impliqués dans un grand nombre de processus biologiques tels que la prolifération, l'apoptose, la différenciation et le développement. Le premier membre de cette famille, *Klf1* ou Erythroid Krüppel-like Factor (EKLF) a été cloné en 1993 par Miller et ses collègues (Miller and Bieker, 1993). A ce jour, 17 membres de la famille des KLFs ont été identifiés (*Klf1* à *Klf17*) (Suske et al., 2005).

Krüppel-like	Autre désignation	Profil d'expression	Année du knock-out	Phénotype après knock-out	Autres rôles fonctionnels	Références clés
Klf1	Eklf (Erythroid KLF)	Cellules érythrocytaires Mastocytes	1995	β -thalassémie létale à E14.5	Différenciation érythrocytaire	Perkins et al., 1995
Klf2	Lklf (Lung KLF)	Poumon, Erythrocytes, Vaisseaux sanguins, Cellules lymphoïdes, Tissu adipeux blanc	1997	Déficience cardiaque létale de E12.5 à E14.5	Différenciation érythrocytaire, Différenciation lymphocytes T, Migration cellulaire, Adipogénèse	Kuo et al., 1997, Wani et al., 1998
Klf4	Gklf (Gut KLF)	Peau, Côlon, Yeux, Cellules lymphoïdes	1999	Déficience dans la barrière cutanée : mort après la naissance dû à une déshydratation	Différenciation des cellules épithéliales coloniques	Segre et al., 1999, Katz et al., 2002
Klf5	Iklf (Intestinal KLF), BTEB2 (Basic transcription element binding 2)	Vaisseau sanguins, Côlon, Tissu adipeux blanc	2002	Mort embryonnaire précoce avant E8.5	Développement cardiovasculaire, développement du côlon, Adipogénèse	Shindo et al., 2002, Oishi et al., 2005
Klf6		Ubiquitaire	2006	Déficience de l'érythropoïèse et de la vascularisation du sac vitellin létale à E12.5		Matsumoto et al., 2006
Klf7	Uklf (Ubiquitous KLF)	Ubiquitaire, forte expression dans les tissus neurosensuels	2005	Défauts neurologiques sévères conduisant à la mort 2 jours après la naissance		Laub et al., 2005
Klf9	Bteb1	Expression générale, forte expression dans le cervelet et l'hippocampe	2003	Viable mais développement de l'utérus anormal et problème d'implantation embryonnaire	Défauts légers de comportement	Morita et al., 2003, Simmen et al., 2004
Klf10	Tieg1 (TGFB inducible early growth response)	Expression générale	2005	Viable mais défauts d'ossification, problème dans la formation du squelette	Défauts cardiaques	Subramaniam et al., 2005
Klf11	Fklf (embryonic/fetal β -like globin gene activating KLF), Tieg2	Expression générale	2005	Viable, aucun phénotype observé		Song et al., 2005
Klf13	Fklf2, Bteb3, Rflat-1 (Rantes factor of late activated T lymphocytes-1)	Expression générale dont les cellules lymphoïdes et cardiaques	2007	Viable mais défaut dans la différenciation des lymphocytes T	Développement cardiaque	Zhou et al., 2007
Klf15	Kklf (Kidney KLF)	Expression générale	2007	Viable mais hypersensibilité à une hypertrophie cardiaque	Adipogénèse	Fisch et al., 2007

Figure 16 : Le rôle des différents membres de la famille des Krüppel-like factors

Ce tableau récapitule le rôle des principaux membres de la famille des Klf en indiquant leur profil d'expression et le phénotype des souris « knock-out » obtenues pour chacun de ces facteurs.

Nous allons nous intéresser plus particulièrement à deux membres de cette famille : les facteurs Klf4 (ou GKlf : Gut enriched Klf) et Klf5 (ou Iklf : Intestinal enriched Klf), car ils font l'objet d'une partie importante de notre étude expérimentale.

7.1. Structure des facteurs de transcription Klf.

Les membres de la famille des Krüppel-like factors sont caractérisés par la présence, dans leur séquence protéique, de trois domaines à doigt de zinc Cys2/His2 fortement conservés. Ces domaines sont localisés à l'extrémité C-terminale des protéines et permettent ainsi la fixation des Klfs au niveau de régions riches en GC et de motifs CACCC de l'ADN (Kaczynski et al., 2003) (Suske et al., 2005). Les doigts de zinc sont séparés par une séquence spécifique de sept acides aminés TGEKP(Y/F)X très conservée (Dang et al., 2000). En dehors de ces régions à doigt de zinc, il y a très peu d'homologie entre les différents membres de la famille des Klfs (Kaczynski et al., 2003). Cette faible homologie permet ainsi à chaque membre d'avoir des fonctions moléculaires et biologiques bien spécifiques (Figure 16).

7.2. Identification des Krüppel-like factors Klf4 et Klf5.

La séquence codante de Klf4 a été isolée à partir d'une banque d'ADNc, issue de cellules NIH3T3 (lignée de fibroblastes embryonnaires murine), criblée avec une sonde ADN correspondant à la séquence d'un motif à doigt de zinc du facteur de transcription Egr1 (Early Growth Response protein 1) (Lanahan et al., 1992). La protéine Klf4 est composée de 483 acides aminés et présente une masse moléculaire de 53kDa. Sa séquence protéique comporte deux sites de localisation nucléaire lui permettant ainsi, une fois transloquée dans le noyau, d'activer ou de réprimer l'expression de ses gènes cibles (Shields and Yang, 1997).

Le facteur de transcription Klf5 a été identifié sur la base de sa forte homologie avec un autre membre de la famille des Klfs : Klf2 ou LKlf (Lung enriched KLF) par Conkright et ses collègues en 1999 (Conkright et al., 1999). La séquence codante de Klf5 est composée de 446 acides aminés, et est à 88% identique à son homologue humain BTEB2 (Basic Transcription Element Binding Protein 2).

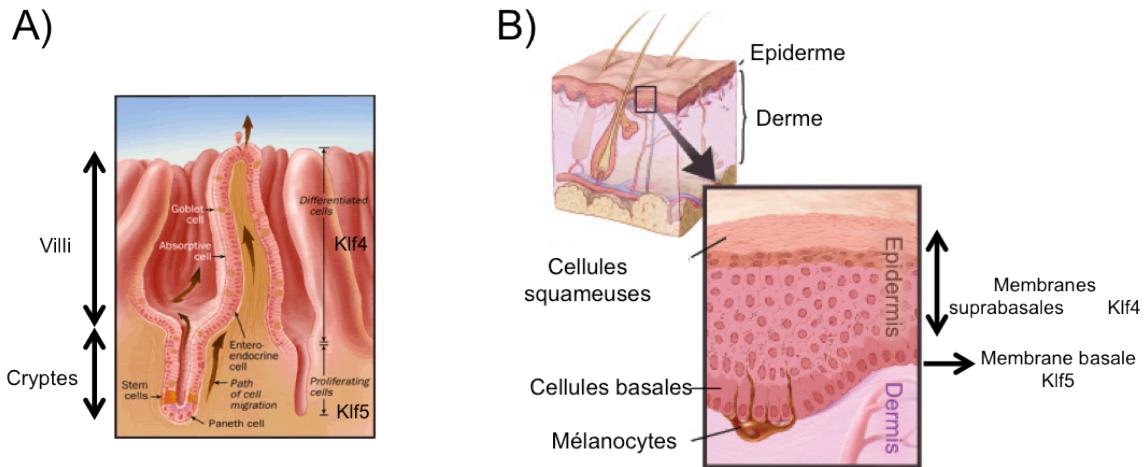


Figure 17 : Expression de Klf4 et Klf5 dans l'intestin et l'épiderme

- A) Klf4 est exprimée au niveau des cellules différencierées de l'épithélium intestinal alors que Klf5 est exprimée au niveau de la base des cryptes composées de cellules indifférencierées (stem cells).
- B) Klf4 est exprimée au niveau des cellules épithéliales de la peau et plus particulièrement au niveau des cellules différencierées de la partie supérieure. Klf5 est exprimée au niveau de la membrane basale de l'épiderme composée de cellules indifférencierées.

7.3. In vivo, Klf4 et Klf5 présentent des profils d'expression antagonistes.

7.3.1. Profil d'expression de Klf4 dans l'embryon.

7.3.1.1. Expression de Klf4 dans les cellules en différenciation terminale de l'épithélium intestinal.

Dans l'embryon, *Klf4* est spécifiquement exprimé tout au long du tractus intestinal, au niveau des testicules et des poumons. Des expériences d'hybridation *in situ* ont permis de montrer que *Klf4* est plus particulièrement exprimé au niveau des cryptes de l'épithélium intestinal dans la membrane suprabasale où sont situées les cellules différenciées (Shields et al., 1996) (Figure 17). Ces cellules ont pour origine la base des cryptes constituées de cellules indifférenciées qui se divisent activement mais sans encore exprimer *Klf4* à ce stade. Une partie de ces cellules migrent régulièrement à la surface de l'épithélium tout en se différenciant et l'expriment alors (Gordon et al., 1992). Une étude réalisée par Ton-That et ses collègues (Ton-That et al., 1997) a mis en évidence l'expression spécifique de *Klf4* entre les jours 15 et 18 du développement embryonnaire, au moment où les cellules épithéliales de la paroi du côlon commencent à se différencier. *Klf4* semble donc jouer un rôle important dans la différenciation terminale des cellules épithéliales au cours du développement.

7.3.1.2. Klf4 et les cellules épithéliales de la peau.

L'expression de *Klf4* a également été mise en évidence dans les cellules épithéliales de la peau, et comme pour l'épithélium intestinal, plus particulièrement au niveau des cellules différenciées de la partie supérieure de l'épiderme. Les embryons de souris « knock-out » *Klf4* ^{-/-} arrivent à terme mais meurent seulement deux semaines après leur naissance. Leur décès est lié à une déficience de la barrière cutanée qui ne joue plus son rôle protecteur (Segre et al., 1999). Au contraire, dans les embryons de souris surexprimant *Klf4* au niveau des cellules de la membrane basale de l'épiderme la barrière cutanée se forment au jour 15.5 du développement, c'est-à-dire un jour plus tôt que les souris contrôles (Jaubert et al., 2003). Chez ces souris transgéniques, l'épiderme est caractérisé par une couche cornée plus épaisse et une expression plus importante de marqueurs de la différenciation terminale comme la filagrine et la loricrine. Une expérience d'incorporation de BrdU (Bromo-désoxyuridine) a été réalisée chez ces souris. La révélation de cet analogue avec un anticorps spécifique permet de

marquer les cellules proliférantes (Jaubert et al., 2003). Une diminution de 40% du nombre de cellules en prolifération au niveau de la membrane basale de l'épiderme a été observée chez les souris transgéniques par rapport aux souris contrôles. Klf4 aurait donc un rôle d'inhibiteur de la prolifération cellulaire.

Dans l'épithélium intestinal, comme celui de la peau, Klf4 semble promouvoir la différenciation terminale des cellules épithéliales (Ton-That et al., 1997) (Segre et al., 1999) tout en inhibant leur prolifération.

7.3.2. L'expression de *Klf5* varie au cours du développement embryonnaire.

Des expériences de northern-blot ont été réalisées, dans lesquels une sonde spécifique de *Klf5* a été hybride sur des extraits d'ARN issus d'embryons de souris à différents stades de gestation (7, 11, 15 et 17 jours). Ces études ont montré que *Klf5* est fortement exprimé au 7^{ème} jour de gestation puis disparaît au 11^{ème} jour, pour enfin réapparaître à J15 et J17 mais à des niveaux beaucoup plus faibles. D'autres membres de la famille des Klfs comme Klf1 et Klf2 présentent des modifications au niveau de leur profil d'expression au cours du développement, mais chaque facteur présente un profil d'expression temporel qui lui est unique (Garrett-Sinha et al., 1996) (Anderson et al., 1995). Les embryons de souris « knock-out » pour *Klf5* meurent vers le 8^{ème} jour de gestation après la formation de l'épiblaste (Shindo et al., 2002).

7.3.3. *Klf5* présente un profil d'expression spécifique dans les tissus adultes.

Dans les tissus adultes, *Klf5* est fortement exprimé dans le conduit gastro-intestinal et notamment au niveau de l'estomac, du petit intestin et du côlon. *Klf5* est également présent, à des niveaux plus faibles, au niveau de l'utérus, du placenta, de la peau, des poumons et des testicules. Comme les autres membres de la famille des Klfs, l'expression de *Klf5* est restreinte à des tissus bien spécifiques, avec tout de même une expression secondaire plus faible dans d'autres tissus. Des études plus précises d'hybridation *in situ* ont permis de mettre en évidence un profil d'expression spécifique de *Klf5* au niveau de l'épithélium intestinal. En effet, aucune expression de *Klf5* n'a été détectée au niveau de la couche musculaire des différents organes. *Klf5* est exprimé au niveau de la base des cryptes qui sont composées de cellules indifférenciées, contrairement à *Klf4* qui est exprimé dans la partie supérieure des cryptes épithéliales (Shields et al., 1996).

7.4. Activité transcriptionnelle des facteurs de transcription Klf4 et Klf5.

7.4.1. Les facteurs de transcription Klf4 et Klf5 régulent tous les deux l'expression de Klf4.

Klf4 et Klf5 ont été décrits comme ayant des rôles antagonistes dans la régulation de l'expression de certains gènes cibles, bien que ces deux facteurs se fixent au niveau d'éléments cis-régulateurs présentant une très forte homologie, voire même des régions strictement identiques (Dhang et al., 2002). L'exemple parfait est le gène codant pour Klf4 lui-même. En effet, il a été montré que Klf4 induit l'activation de sa propre expression, alors que Klf5 inhibe l'expression de *Klf4*. Les deux facteurs Klf4 et Klf5 interagissent dans ce cas précis avec exactement les mêmes éléments cis-régulateurs mais ont des rôles complètement opposés (Dang et al., 2002). Klf5 inhibe l'effet activateur de Klf4 sur le promoteur de *Klf4*, quant à Klf4 il inhibe l'effet répresseur de Klf5. Il semblerait que ces effets antagonistes soient la conséquence d'un phénomène de compétition physique entre ces deux protéines au niveau des sites de régulation du promoteur de *Klf4* (Dhang et al., 2002).

7.4.2. Klf4 et Klf5 ont des effets antagonistes sur l'expression d'autres gènes cibles.

Ces phénomènes de compétition entre Klf4 et Klf5 ont par la suite été mis en évidence pour la régulation transcriptionnelle d'autres gènes cibles comme le gène codant pour l' α -actine (Liu et al., 2003) ou pour la laminine 1 (Piccinni et al., 2004). Dans le cas de l' α -actine, Klf4 inhibe son expression alors que Klf5 l'active. Les auteurs ont réalisé une expérience simple dans laquelle des cellules NIH3T3 ont été transfectées avec (i) un plasmide rapporteur dans lequel le gène de la luciférase est sous le contrôle du promoteur de l' α -actine, (ii) un plasmide codant pour Klf4 et (iii) des concentrations croissantes d'un plasmide codant pour Klf5. En mesurant l'activité de la luciférase dans ces cellules, ils ont montré que plus l'expression de *Klf5* augmente plus l'activité luciférase augmente elle aussi. En variant les rapports Klf5/Klf4, le niveau d'expression de la luciférase varie également, ce qui est conforme avec l'idée d'une compétition physique entre ces deux facteurs et ceci au niveau du même site de fixation à l'ADN.

Les facteurs de transcription Klf4 et Klf5 se fixent donc sur les mêmes éléments de régulation (régions de l'ADN riche en GC, séquences consensus CACCC) mais exercent des rôles opposés en ce qui concerne la régulation transcriptionnelle de leurs gènes cibles.

7.5. Les rôles opposés de Klf4 et Klf5 dans la régulation de la prolifération cellulaire.

7.5.1. Klf4 : son rôle dans l'inhibition de la prolifération.

7.5.1.1. Le niveau d'expression de Klf4 est augmenté dans les cellules quiescentes.

Nous avons vu que *Klf4* est fortement et spécifiquement exprimé au niveau des cellules épithéliales post-mitotiques et complètement différenciées de l'épithélium intestinal (Shields et al., 1996). Cette observation a conduit Shields et ses collègues à étudier *Klf4* dans des systèmes de culture *in vitro*. Ils ont ainsi montré que les cellules NIH3T3, privées de sérum puis stimulées à nouveau avec du sérum frais pour induire leur prolifération, voient leur niveau d'expression de *Klf4* diminuer fortement. Au contraire, le niveau d'expression de *Klf4* augmente significativement dans les mêmes cellules dont la prolifération a été inhibée par privation de sérum ou inhibition de contact. De plus, la surexpression de *Klf4* dans des cellules proliférantes inhibe leur capacité à synthétiser de l'ADN et donc à se diviser (Shields et al., 1996). Le niveau d'expression de *Klf4* est élevé dans des cellules quiescentes alors qu'il est faible ou nul dans des cellules proliférant de manière active. Le facteur de transcription *Klf4* semble donc jouer un rôle important dans la prolifération cellulaire.

7.5.1.2. Klf4 régule l'expression de gènes jouant un rôle dans le cycle cellulaire.

7.5.1.2.1. Analyse transcriptionnelle des gènes cibles de Klf4.

En 2003, Chen et ses collègues (Chen et al., 2003) ont analysé par la technique de puces à ADN le transcriptome de *Klf4*. Ces études ont été réalisées sur des cellules EcR/RKO-Klf4 (issus de cancer du côlon humain) dans lesquelles l'expression de *Klf4* est sous contrôle d'un promoteur inducible résultant de la fusion entre le domaine de fixation du récepteur à l'ecdysone (EcR) et le domaine de transactivation VP16 (No et al., 1996). Dans ce système, l'addition de ponastérone dans le milieu active l'expression de *Klf4*. Ces études ont

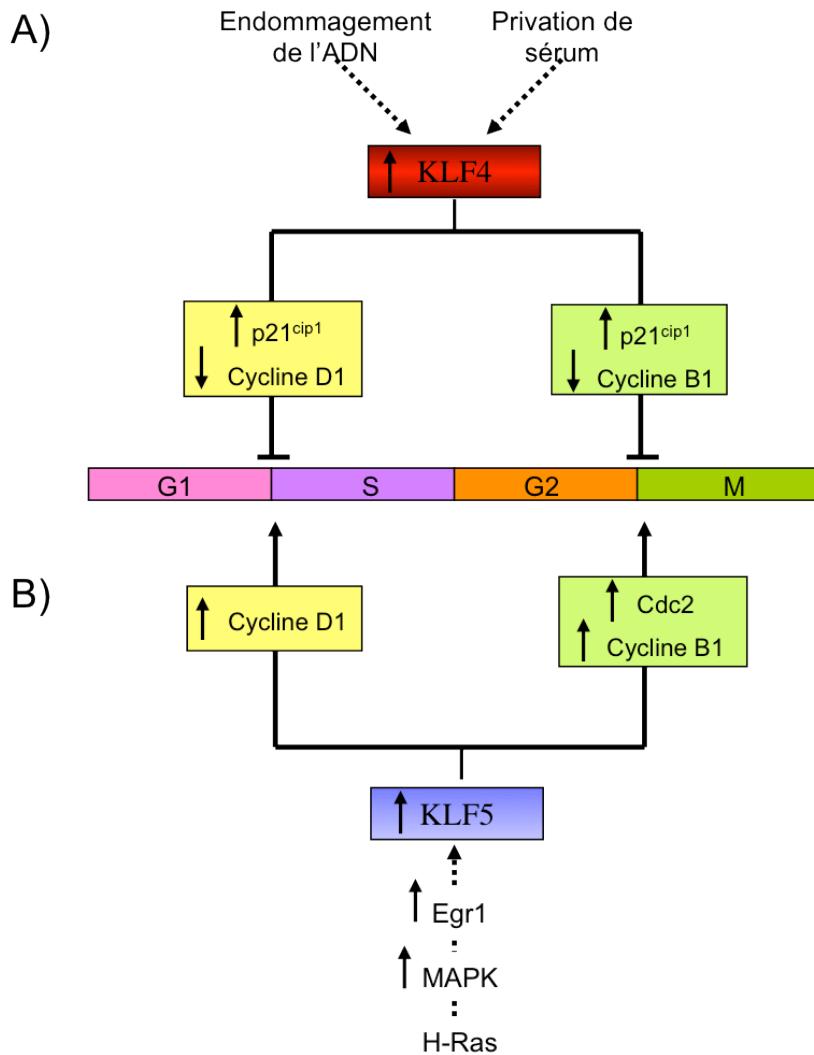


Figure 18 : Rôles de Klf4 et Klf5 dans la régulation du cycle cellulaire

- A) Suite à un endommagement de l'ADN ou à une privation de sérum, le niveau d'expression de *Klf4* augmente dans les cellules. *Klf4* va alors activer l'expression de *p21^{cip1}* qui est un inhibiteur de la transition G1/S et G2/M du cycle cellulaire. *Klf4* va aussi inhiber l'expression de la *cycline D1* et bloquer la transition G1/S, ainsi que l'expression de la *cycline B1* et bloquer la transition G2/M du cycle cellulaire.
- B) L'oncogène H-Ras va induire une cascade d'activation de l'expression de la *MAPK* puis de *Egr1* et enfin de *Klf5*. L'augmentation de l'expression de *Klf5* va induire une augmentation de l'expression de la *cycline D1*, qui va stimuler la transition G1/S, et de *Cdc2* et de la *cycline B1* qui vont stimuler la transition G2/M du cycle cellulaire.

montré que 24h après l'induction de Klf4 l'expression de nombreux gènes codant pour des inhibiteurs du cycle cellulaire (*p57*, *p21^{cip1}*) est activée alors que celle de gènes activateurs du cycle (*cycline D1*, *complexes CDC2/CDK1*) est inhibée.

Cette étude transcriptionnelle met ainsi en évidence le rôle de Klf4 dans la régulation du cycle cellulaire et notamment dans l'arrêt de la croissance.

7.5.1.2.2. Le rôle de Klf4 dans la régulation du cycle cellulaire.

L'exposition de cellules à des agents induisant des dommages au niveau de l'ADN, tel que le méthyl-méthanesulfonate (MMS) (Zhang et al., 2000) ou une irradiation à des rayons gamma, entraîne une activation de l'expression du gène codant pour la protéine p53. Dans ces cellules le niveau d'expression de *Klf4* ainsi que de *p21^{cip1}* (Zhang et al., 2000) est augmenté. En fait, Klf4 et p53 forment un complexe et Klf4, en se fixant directement sur le promoteur de *p21^{cip1}*, va permettre à p53 de se fixer sur ce dernier et de réguler ainsi sa transcription (Zhang et al., 2000). L'activation de l'expression de *p21^{cip1}* suite à des dommages causés au niveau de l'ADN a pour conséquence l'arrêt du cycle cellulaire au niveau de la transition G1/S et de la transition G2/M. Klf4 a été montré comme nécessaire et suffisant dans la régulation de ces points de transition, puisque sans lui p53 ne pourrait induire l'expression de *p21^{cip1}* (Yoon et al., 2003) (Yoon and Yang, 2004). D'autres études ont mis en évidence que Klf4, en plus de son rôle dans l'activation de *p21^{cip1}*, inhibe l'expression de la *cycline D1* (Shie et al., 2000) et de la *cycline B1* (Yoon et al., 2004) respectivement impliquées dans la transition G1/S et G2/M du cycle cellulaire. Klf4 exercerait ce rôle en se fixant directement sur le promoteur de ses deux gènes cibles.

Toutes ces études montrent donc que Klf4 joue un rôle crucial dans la régulation du cycle cellulaire par l'activation ou l'inhibition de l'expression de gènes régulateurs du cycle. Le rôle de Klf4 dans le cycle cellulaire des cellules somatiques est résumé dans la Figure 18.

7.5.2. Contrairement à Klf4, Klf5 est un activateur la prolifération cellulaire.

7.5.2.1. Klf5 est fortement exprimé dans des cellules proliférant activement.

Contrairement à Klf4, de nombreuses études ont mis en évidence que Klf5 a un rôle d'activateur de la prolifération cellulaire. En effet, le niveau d'expression de *Klf5* est très faible dans les cellules NIH3T3 devenues quiescentes suite à une privation de sérum, et il

augmente très fortement après l'addition de sérum dans le milieu (Sun et al., 2001). L'expression constitutive de *Klf5* dans des cellules NIH3T3 ou des cellules épithéliales de l'intestin (Sun et al., 2001) (Chanchevalap et al., 2004) induit une augmentation importante de la vitesse de prolifération de ces cellules, ce qui entraîne leur transformation.

7.5.2.2. Rôle de *Klf5* dans la transformation cellulaire.

Le niveau d'expression de *Klf5* est fortement augmenté dans des cellules NIH3T3 transformées suite à la surexpression de l'oncogène H-Ras (Harvey rat sarcoma virus oncogene), ce qui fait de *Klf5* l'un des médiateurs principaux de l'effet oncogénique de ce dernier (Nandan et al., 2004). La Figure 18 illustre le processus de transformation d'une cellule, médié par *Klf5*, suite à la surexpression de *H-Ras*. Dans ces cellules, H-Ras va activer l'expression de la *MAPK* (Mitogen Activated Protein Kinase) qui va elle-même activer l'expression de *Egr1* qui va stimuler l'expression de *Klf5*. *Klf5* active ensuite l'expression de la *cycline D1*, mais aussi de la *cycline B1* et de *Cdc2*, ce qui a pour effet de transformer la cellule.

Les deux facteurs de transcription *Klf4* et *Klf5* présentent de nombreuses homologies au niveau de leur structure et notamment de leur site de fixation à l'ADN. *Klf4* et *Klf5* sont tous les deux exprimés dans l'épithélium intestinal, mais ils présentent un profil d'expression exclusif au niveau cellulaire. De même, ces deux facteurs régulent les mêmes gènes cibles, mais de façon antagoniste, ce qui entraîne des effets contraires sur la régulation de nombreux mécanismes cellulaires.

7.6. *Klf4* et *Klf5* dans les cellules ES.

7.6.1. Le rôle de *Klf4* dans l'autorenouvellement des cellules ES.

La première étude ayant mis en évidence le rôle de *Klf4* dans les cellules ES est celle de Li et ses collègues en 2005 (Li et al., 2005). Cette étude caractérise *Klf4* comme un gène cible de la voie LIF. Deux analyses ont ainsi été réalisées à partir de cellules ES de souris surexprimant *Klf4* et induites à se différencier par formation de corps embryoïdes :

- (i) l'analyse du niveau d'expression du marqueur de pluripotence *Oct4* : les corps embryoïdes au 6^{ème} jour de différenciation expriment toujours *Oct4* à un niveau élevé, alors que dans les cellules contrôles son niveau d'expression diminue à partir du 3^{ème} jour,
- (ii) la capacité des cellules à former des corps embryoïdes secondaires : les cellules surexprimant *Klf4* sont capables de générer un très grand nombre de corps embryoïdes secondaires, alors que les cellules contrôles n'en génèrent que très peu.

Cette étude démontre donc que *Klf4* joue un rôle dans l'autorenouvellement des cellules ES de souris et dans l'inhibition de leur différenciation.

7.6.2. Le trio *Klf2/Klf4/Klf5* est nécessaire à l'autorenouvellement des cellules ES.

Dans une étude très récente, Jiang et ses collègues (Jiang et al., 2008) ont mis en évidence que *Klf4*, *Klf5* et *Klf2*, régulent l'expression du facteur de pluripotence *Nanog* et sont donc nécessaires au maintien en autorenouvellement des cellules ES. Ces trois facteurs *Klfs* se fixent au niveau de la même région régulatrice du promoteur de *Nanog* et peuvent donc chacun stimuler son expression. Ainsi, seule l'inhibition simultanée de ces trois facteurs, grâce à la technique d'interférence à l'ARN, induit la différenciation des cellules ES et une diminution de l'expression de *Nanog* (Jiang et al., 2008). En effet, l'inhibition de l'expression de l'un ou de deux facteurs simultanément n'est pas suffisante pour induire la différenciation des cellules ES et de réguler le niveau d'expression de *Nanog*.

Selon cette étude *Klf2*, *Klf4* et *Klf5* sont nécessaires au maintien en autorenouvellement des cellules ES. Ils activent de façon redondante l'expression de *Nanog* pour maintenir les cellules ES en autorenouvellement. Ainsi, la présence d'un seul de ces trois facteurs serait suffisante pour assurer cette fonction.

7.6.3. Les facteurs *Klfs* et leur rôle dans la reprogrammation des cellules somatiques en cellules pluripotentes iPS (induced Pluripotent Stem cells).

L'équipe de Yamanaka en 2006 (Takahashi and Yamanaka, 2006) a réussi à établir des cellules pluripotentes, nommées iPS (induced Pluripotent Cells), suite à la surexpression dans des cellules somatiques murines et humaines de *Oct4*, *Sox2*, *c-Myc* et *Klf4*, quatre facteurs spécifiquement exprimés dans les cellules ES indifférenciées. Ce processus de reprogrammation de cellules somatiques (complètement différenciées) en cellules

pluripotentes est une découverte majeure dans l'espoir de thérapies cellulaires. En effet, elle évite le recours controversé à des embryons et permet d'obtenir des cellules pluripotentes ayant le même patrimoine génétique que les cellules somatiques utilisées. Cette découverte conforte l'espoir de futures thérapies réparatrices de tissus ou d'organes endommagés, sans risque de rejet des cellules greffées.

Lorsque seuls les trois facteurs Oct4, Sox2 et c-Myc sont surexprimés, les auteurs n'arrivent pas à obtenir de lignées iPS. Klf4 est donc crucial pour la reprogrammation des cellules somatiques. Cependant, il peut être remplacé par un autre membre de la famille des Klf : Klf2, avec la même efficacité. Même lorsque Klf4 est remplacé par Klf5, il est toujours possible d'obtenir des cellules iPS, mais avec une efficacité plus faible (Nakagawa et al., 2008).

En conclusion, les facteurs de transcription Klf2, Klf4 et Klf5 jouent un rôle important dans la maintien en autorenouvellement des cellules ES en régulant l'expression d'un gène clé de la pluripotence *Nanog*. Ces trois facteurs semblent également indispensables dans la reprogrammation des cellules somatiques en cellules pluripotentes induites.

RÉSULTATS

1- *Introduction*

Les cellules souches embryonnaires (cellules ES) sont obtenues à partir des cellules de la Masse Cellulaire Interne ou Inner Cell Mass (ICM) du blastocyste. Les cellules ES sont pluripotentes : elles sont capables de se différencier *in vitro* dans les trois lignages à l'origine des tissus composant l'organisme adulte, l'ectoderme, le mésoderme et l'endoderme. Les cellules ES sont également immortelles. Elles peuvent se multiplier de façon illimitée tout en conservant leur propriété de pluripotence. On parle d'autorenouvellement permanent (Smith, 2001).

Pour se maintenir en autorenouvellement, les cellules ES de souris sont dépendantes d'une cytokine, le Leukemia Inhibitory Factor ou LIF. Le LIF se fixe sur un récepteur de faible affinité LIFR, qui s'hétérodimérisé avec la sous-unité gp130, co-récepteur commun à toutes les cytokines de la famille de l'IL-6. Les deux sous-unités du récepteur interagissent individuellement, et de manière constitutive, avec un membre de la famille des kinases JAK (Janus Kinase). L'hétérodimérisation LIFR:gp130 induit l'activation des kinases JAK qui phosphorylent alors les deux sous-unités du récepteur sur des tyrosines localisées dans le domaine intracellulaire. Cette phosphorylation induit le recrutement du facteur de transcription STAT3 (Signal Transducer and Activator of Transcription). STAT3 est à son tour phosphorylé par les kinases JAK sur la tyrosine Y705, induisant son homodimérisation et sa translocation nucléaire (Burdon et al., 1999). La fixation du LIF sur son récepteur recrute et active également d'autres facteurs comme la Phosphatidyl-Inositol 3 Kinase (PI3K) (Paling et al., 2004) et la tyrosine kinase Yes de la famille Src (Anneren et al., 2004). Tous ces facteurs jouent un rôle dans le maintien des cellules ES en autorenouvellement. Cependant STAT3 semble avoir un rôle prédominant et déterminant dans l'inhibition de la différenciation des cellules ES induite par le LIF. En effet, la surexpression d'un mutant dominant négatif du facteur STAT3 dans les cellules ES induit leur différenciation en mésoderme et en endoderme (Bœuf et al., 1997 ; Niwa et al., 1998) ; à l'inverse, la surexpression d'un mutant constitutivement actif promeut l'autorenouvellement des cellules ES en l'absence de LIF exogène (Matsuda et al., 1999). Ainsi, le facteur de transcription STAT3 est nécessaire et suffisant à l'autorenouvellement des cellules ES *in vitro* (Burdon et al., 2002).

Parallèlement à l'activation du facteur STAT3, l'hétérodimérisation du récepteur LIFR avec le transducteur gp130 permet le recrutement et l'activation de la tyrosine phosphatase SHP2. Cette dernière recrute les complexes Grb2/SOS via la protéine adaptatrice Gab1 et active ainsi la cascade de signalisation Ras/Raf/MEK/ERK. Cependant, l'activation des protéines ERK par le récepteur LIFR:gp130 n'est pas nécessaire à la croissance et à l'autorenouvellement des cellules ES. Au contraire, son inhibition stimule

l'autorenouvellement des cellules ES et réduit leur dépendance vis-à-vis du LIF (Burdon et al., 1999). Ainsi, la voie de signalisation induite par LIFR:gp130 stimule à la fois l'autorenouvellement (via STAT3) et la différenciation (via ERK1 et ERK2) des cellules ES de souris. Dans les conditions optimales de culture, l'équilibre entre les deux voies est nettement déplacé vers la voie de l'autorenouvellement.

En dépit du rôle central joué par le facteur STAT3 dans le contrôle de l'autorenouvellement des cellules ES de souris, on ne connaît pas ses gènes cibles, dont l'activité serait responsable de l'inhibition de la différenciation. Les gènes *junB* et *c-fos*, des régulateurs de la transition G1/S du cycle cellulaire, ainsi que *socs-3*, un inhibiteur des JAKs participant au rétro-contrôle négatif de l'activité du récepteur gp130, ont été identifiés comme des cibles directes du facteur STAT3 dans les cellules ES (Matsuda et al., 1999 ; Duval et al., 2000). L'inactivation des gènes *junB* et *c-fos* ne semble pas altérer l'équilibre autorenouvellement-différenciation des cellules ES de souris, ce qui indique que ni l'un ni l'autre ne sont des éléments déterminants de l'action du facteur STAT3 dans ces cellules (Schorpp-Kistner et al., 1999). Seuls les gènes cibles *Klf4*, *c-myc*, *Dax1* et *Eed* ont été identifiés comme des cibles importantes pour le maintien des cellules ES en autorenouvellement (Li et al., 2005 ; Cartwright et al., 2005 ; Sun et al., 2008 ; Ura et al., 2008).

L'identification de l'ensemble des gènes activés par le facteur STAT3 dans les cellules ES constitue donc une étape cruciale vers la compréhension des mécanismes de contrôle de l'autorenouvellement. Pour identifier ces gènes, il est d'abord nécessaire de contrôler expérimentalement l'activité du facteur STAT3 afin de comparer les transcriptomes des cellules ES en présence et en absence du facteur STAT3 actif. La stratégie la plus simple consisterait en une comparaison des transcriptomes des cellules ES avant et après stimulation par le LIF. Cependant, comme nous l'avons indiqué précédemment, la stimulation du récepteur LIFR:gp130 active non seulement le facteur STAT3 mais aussi les voies de signalisation PI3K, Yes et ERK via la phosphatase SHP2. La stimulation du récepteur LIFR:gp130 par le LIF induit ou réprime vraisemblablement l'expression de nombreux gènes autres que les gènes cibles de STAT3. Dans le but d'identifier spécifiquement les gènes cibles du facteur STAT3, il est nécessaire de contrôler l'activité de ce facteur sans altérer l'activité du récepteur LIFR:gp130. Pour ce faire, nous avons utilisé un mutant hormono-dépendant du facteur STAT3, STAT3-ER^{T2}. STAT3-ER^{T2} résulte de la fusion de la protéine STAT3 au domaine de fixation du ligand du récepteur de l'oestrogène humain (ER) muté (Indra et al., 1999). L'activité de cette protéine de fusion est régulée par le 4'Hydroxy-Tamoxifène

(4’OHT). En absence du ligand, STAT3-ER^{T2} est séquestré dans le cytoplasme sous forme inactive. En présence du ligand, STAT3-ER^{T2} est transloqué dans le noyau, où il forme des homodimères via le domaine ER^{T2}, et active ses gènes cibles (Matsuda et al., 1999). L’activation du facteur STAT3-ER^{T2} par le 4’OHT ne requiert pas la phosphorylation de la tyrosine Y705 par le récepteur. L’intérêt du système STAT3-ER^{T2} est de permettre un contrôle direct de l’activité STAT3 par le 4’OHT en l’absence de stimulation du récepteur par le LIF.

Dans un premier temps, nous avons d’abord fabriqué une lignée de cellules ES exprimant le facteur STAT3-ER^{T2} à l’aide d’un système d’expression épisomal utilisant l’antigène T du virus du polyome (Aubert et al., 2002). Nous avons montré que cette lignée de cellules ES n’est plus dépendante de l’activation du récepteur par le LIF, mais strictement dépendantes du 4’OHT, pour s’autorenouvellement. Dans un deuxième temps, cette lignée de cellules ES exprimant STAT3-ER^{T2} a été utilisée pour réaliser une analyse comparée des transcriptomes en présence et en absence de 4’OHT. Nous avons ainsi identifié 58 nouveaux gènes cibles de la voie LIF/STAT3 par la technique des puces à ADN Affymétrix. Dans un troisième temps, nous avons réalisé une analyse fonctionnelle sur 23 d’entre eux afin de déterminer leur fonction dans le contrôle de l’autorenouvellement des cellules ES et l’inhibition de la différenciation. Ces gènes ont été sélectionnés sur la base de leur fonction dans les autres types cellulaires et de leur profil d’expression au cours de la différenciation en corps embryoïdes. Parmi ces 23 gènes la majorité codent pour des facteurs de transcription (Klf4, Klf5, Zfp36, Zfp36l1, Smad7, Spry2, Sall4, Gbx2), d’autres codent pour des kinases (Pim1, Pim3 et Sgk), des régulateurs de la signalisation (Rgs16, Dact1), une cycline (Cnnm1), des molécules d’adhésion (Ocln, Icam1), un gène de réponse au stress (Ier3), une protéine de la matrice extracellulaire (Cyr61), une déadénylase (Cern4l), une sulfatase (Sulf1), une protéine du cytosquelette (Vim) ainsi que deux RIKENs (5230400M03Rik, 1700051K13Rik). Afin d’analyser la fonction de ces 23 gènes dans le maintien en autorenouvellement et l’inhibition de la différenciation des cellules ES, leur expression a été inactivée grâce à la technique d’interférence à l’ARN. Nous avons ainsi pu montrer que 22 des 23 gènes cibles de STAT3, codant pour des facteurs de transcription, des kinases, des régulateurs du cycle cellulaire, jouent un rôle dans le maintien de la pluripotence des cellules ES de souris. En effet, l’inactivation de l’expression de ces 22 gènes exerce un effet négatif sur l’autorenouvellement des cellules ES et induit leur différenciation spontanée en présence de LIF. L’analyse de marqueurs de la différenciation en mésoderme et en endoderme nous a permis d’identifier trois familles de gènes : (i) les gènes impliqués dans l’inhibition de la

différenciation en endoderme, (ii) les gènes impliqués dans l'inhibition de la différenciation en mésoderme, et (iii) les gènes impliqués dans l'inhibition de la différenciation dans les deux lignages. Parmi ces 22 gènes, nous avons identifié *Pim1* et *Pim3*, deux gènes codant pour les sérine/thréonine kinases de la famille PIM (Mikkers et al., 2004). Nous avons analysé de manière plus précise la fonction de ces deux gènes. Nous avons ainsi démontré que ces deux facteurs jouent un rôle important dans le contrôle de l'autorenouvellement et l'inhibition de la différenciation des cellules ES de souris. En effet, l'inactivation de l'expression de ces gènes induit la différenciation des cellules ES alors que leur surexpression permet aux cellules de résister temporairement à la différenciation induite par privation de LIF (Article N°1).

Nous avons également identifié deux gènes cibles de STAT3 particulièrement intéressants : *Klf4* et *Klf5*. *Klf4* et *Klf5* font partie de la famille des facteurs de transcription Krüppel-like (Suske et al., 2005). Avec *Klf2*, *Klf4* et *Klf5* forment un réseau de régulation contribuant au maintien des cellules ES en autorenouvellement grâce à l'activation du gène de pluripotence *Nanog* (Jiang et al., 2008). *Klf4* est également nécessaire à la reprogrammation des cellules somatiques en cellules souches pluripotentes induites (iPS) (Okita et al., 2007), sa surexpression retarde la différenciation des cellules ES en corps embryoides (Li et al., 2005), et il agit comme co-facteur de Oct4 et Sox2 pour activer l'expression de gènes cibles (Nakatake et al., 2006). Récemment, une étude a décrit le rôle de *Klf5* dans le maintien des cellules ES à l'état indifférencié (Parisi et al., 2008).

Nous avons observé que l'atténuation de l'expression de *Klf4* et *Klf5* par interférence à l'ARN induit une augmentation importante de la différenciation spontanée des cellules ES. De plus, l'expression de ces deux gènes diminue très fortement après seulement 24h de culture en conditions favorisant la différenciation. Nous avons émis l'hypothèse que ces gènes jouaient un rôle important dans le contrôle de la pluripotence. L'analyse des marqueurs de différenciation a révélé que ces deux gènes exercent des effets complémentaires dans le maintien des cellules à l'état pluripotent. En utilisant une lignée de cellules ES exprimant le gène marqueur *DsRed* sous le contrôle du promoteur *Sox17*, nous avons montré que l'inactivation de *Klf4* par interférence à l'ARN induit la différenciation en endoderme (extra-embryonnaire et définitif). Au contraire, en utilisant une lignée qui exprime le gène marqueur *GFP* sous le contrôle du promoteur *T/Brachyury*, nous avons montré que l'inactivation de *Klf5* induit la différenciation en mésoderme.

Afin d'élucider le mécanisme d'action de *Klf4* et *Klf5* dans ce processus, nous avons analysé le transcriptome des cellules ES exprimant les shRNA spécifiques de *Klf4* et *Klf5*. Nous avons ainsi démontré que l'atténuation de la transcription du gène *Klf4* active la

transcription des gènes *GATA-4*, *GATA-6*, *Sox7*, *Sox17*, *Dab2* et *amnionless* dans les cellules indifférenciées. Ces six gènes sont impliqués dans l’engagement de la différenciation dans le lignage endodermique (Fujikura et al., 2002; Futaki et al., 2004; Kalantry et al., 2001; Kanai-Azuma et al., 2002; Niimi et al., 2004; Qu et al., 2008; Seguin et al., 2008; Shimoda et al., 2007; Shimosato et al., 2007; Yang et al., 2002). De manière similaire, l’atténuation de la transcription du gène *Klf5* active la transcription du gène pan-mésodermique *T/Brachyury* dans les cellules indifférenciées. En conclusion, nos résultats indiquent que la voie LIF/STAT3 active deux voies parallèles, l’une réprimant la différenciation endodermique via le facteur *Klf4*, l’autre réprimant la différenciation mésodermique via le facteur *Klf5* (Article N°2).

2- Article n°1

Self-Renewal of Murine Embryonic Stem Cells Is Supported by the Serine/Threonine Kinases Pim1 and Pim3

Stem Cells, 2007, Vol.25, Issue 12, p.2296-3004

Self-Renewal of Murine Embryonic Stem Cells Is Supported by the Serine/Threonine Kinases Pim-1 and Pim-3

IRÈNE AKSOY,^{a,b,c} CALINE SAKABEDOYAN,^{a,b,c} PIERRE-YVES BOURILLOT,^{a,b,c} ANNA B. MALASHICHEVA,^{a,b,c}
JIMMY MANCIP,^{a,b,c} KENNETH KNOBLAUCH,^{a,b,c} MARIELLE AFANASSIEFF,^{a,b,c} PIERRE SAVATIER^{a,b,c}

^aInstitut National de la Santé et de la Recherche Médicale and ^bStem Cell and Brain Research Institute, Bron, France; ^cUniversité de Lyon, Lyon, France

Key Words. Embryonic stem cells • Leukemia inhibitory factor • STAT3 transcription factor • Pim kinases • Cell cycle

ABSTRACT

pim-1 and *pim-3* encode serine/threonine kinases involved in the regulation of cell proliferation and apoptosis in response to cytokine stimulation. We analyzed the regulation of *pim-1* and *pim-3* by the leukemia inhibitory factor (LIF)/gp130/signal transducer and activator of transcription-3 (STAT3) pathway and the role of Pim-1 and Pim-3 kinases in mouse embryonic stem (ES) cell self-renewal. Making use of ES cells expressing a granulocyte colony-stimulating factor:gp130 chimeric receptor and a hormone-dependent signal transducer and activator of transcription-3 estrogen receptor (STAT3-ER^{T2}), we showed that expression of *pim-1* and

pim-3 was upregulated by LIF/gp130-dependent signaling and the STAT3 transcription factor. ES cells overexpressing *pim-1* and *pim-3* had a greater capacity to self-renew and displayed a greater resistance to LIF starvation based on a clonal assay. In contrast, knockdown of *pim-1* and *pim-3* increased the rate of spontaneous differentiation in a self-renewal assay. Knockdown of *pim-1* and *pim-3* was also detrimental to the growth of undifferentiated ES cell colonies and increased the rate of apoptosis. These findings provide a novel role of Pim-1 and Pim-3 kinases in the control of self-renewal of ES cells. STEM CELLS 2007;25:2996–3004

Disclosure of potential conflicts of interest is found at the end of this article.

INTRODUCTION

The propagation of mouse embryonic stem (ES) cells is dependent on the presence of leukemia inhibitory factor (LIF), which engages a heterodimeric receptor complex consisting of LIF receptor (LIFR) and gp130. This complex activates the Janus-associated tyrosine kinases (JAKs), which phosphorylate the receptor chains. The phosphorylated tyrosines then act as docking sites for the signal transducer and activator of transcription-3 (STAT3) [1]. Phosphorylation of STAT3 by JAKs, followed by its dimerization and nuclear translocation, plays a crucial role in self-renewal of ES cells [2]. Hence, expression of a dominant-negative STAT3 mutant in ES cells forces differentiation [3], whereas expression of a hormone-dependent STAT3 mutant that can be activated directly by estradiol sustains self-renewal without the addition of LIF [4]. Activation of the LIFR:gp130 receptor also induces JAK-dependent tyrosine phosphorylation of the Sarc-homology domain phosphatase-2 (SHP-2) phosphatase. However, SHP-2 recruitment opposes STAT3-dependent self-renewal and facilitates differentiation by promoting the activation of the Ras/mitogen-activated protein kinase pathway [5].

A number of genes whose expression is activated by LIF have been identified in ES cells [4, 6–9], although only two, *c-myc* and *Krüppel-like factor (klf) 4*, have been shown to contribute to the inhibition of differentiation. *c-myc* is a target gene of the LIF/STAT3 pathway. Forced expression of a hormone-dependent Myc-ER temporarily inhibits the differentiation of ES cells induced by withdrawal of LIF [8]. Similarly,

forced expression of Klf-4 delays differentiation induced by LIF deprivation [9].

The *pim* genes encode the serine/threonine kinases Pim-1, Pim-2, and Pim-3, involved in the regulation of cell growth and apoptosis [10]. Pim-1 and Pim-2 both have the ability to cooperate with Myc in experimental T-cell lymphomagenesis [11, 12], a finding that established Pim-1 and Pim-2 as proto-oncogenes and important players in the process of malignant transformation. Pim kinases contribute to the regulation of cell proliferation in response to stimulation by cytokines and growth factors. The expression of the *pim-1* gene can be induced via the activation of STAT3 and STAT5, which bind directly to the *pim-1* promoter in eosinophils [13]. In lymphoid cells, *pim-1* is activated by interleukin-6 via STAT3 and regulates proliferation in cooperation with Myc [14]. All three Pim kinases also play an important role in the prevention of cell death [15, 16].

In this study, we examined the regulation of the *pim-1* and *pim-3* genes by the LIF/STAT3 pathway, as well as their role in the control of proliferation and differentiation, in ES cells. The data presented show that Pim-1 and Pim-3 play an important role in maintaining the ES cell identity by inhibiting differentiation and apoptosis.

MATERIALS AND METHODS

Plasmid Construction

The hormone-dependent STAT3-ER^{T2} transcription factor was generated by fusing the coding sequence of the mouse STAT3 tran-

Correspondence: Pierre Savatier, Ph.D., INSERM U846, 18 avenue Doyen Lépine, 69500 Bron, France. Telephone: 33-472-91-34-42; Fax: 33-472-91-34-61; e-mail: savatier@lyon.inserm.fr Received January 29, 2007; accepted for publication August 10, 2007; first published online in STEM CELLS EXPRESS August 23, 2007. ©AlphaMed Press 1066-5099/2007/\$30.00/0 doi: 10.1634/stemcells.2007-0066

STEM CELLS 2007;25:2996–3004 www.StemCells.com

scription factor to a 5'-Xhol/EcoRI-3' fragment containing the entire ER^{T2} domain [17]. The resulting cDNA was subcloned into the EcoRI site in *pPCAGIZ* [3] to generate *pPCAGIZ-STAT3-ER^{T2}*. Full-length cDNA encoding mouse Pim-1 was subcloned into the blunted BstXI site in *pPHCAG* [3] and into the blunted Xhol site in *pPHPGK* [18] to generate *pPHCAG-pim1* and *pPHPGK-pim1*, respectively. Full-length cDNA encoding mouse Pim-3 was subcloned into the blunted EcoRI site in *pPCAGIZ* [3] and into the Xhol site in *pPHPGK* to generate *pPCAGIZ-pim3* and *pPHPGK-pim3*, respectively.

The BLOCK-iT lentiviral RNAi Gateway vector kit (ref. K4943-00; Invitrogen, Carlsbad, CA, <http://www.invitrogen.com>) was used to generate lentiviral vectors expressing small hairpin (sh) RNA directed to green fluorescent protein (*GFP*), *pim-1*, and *pim-3*, according to the manufacturer's instructions. In the *pLenti6/BLOCK-iT* backbone vector, the blunted KpnI (blunt)/Xhol fragment containing the *PGK* promoter and the *neo*^r selectable gene. Oligonucleotides encoding shRNAs for *GFP*, *pim-1*, *pim-3*, and *stat3* were subcloned into the resulting *pLenti6/BLOCK-iT-PGKneo* vector to generate *pLenti6/BLOCK-iT-PGKneo-GFP*, *pLenti6/BLOCK-iT-PGKneo-pim1*, *pLenti6/BLOCK-iT-PGKneo-pim3*, and *pLenti6/BLOCK-iT-PGKneo-stat3*, respectively. The sequences of shRNA are as follows: *GFP*, 5'-GCCACAACTGC-TATATCATTCAAGAGAATGATAGACGTTGTGGC-3'; *pim-1*, 5'-GCAAGACCTCTTCGACTTTATTCAAGAGAATAAGTCG-AAGAGGTCTTG-3'; *pim-3*, 5'-CACGGTCTACACTGACTTGTGA-TGTTCAAGAGACATCAAAGTCAGGTAGACCGTG-3'; *stat3*, 5'-GGAGCTGTTCAAGAACTTATTCAAGAGATAAGTTCTGA-ACAGCTCC-3'.

ES Cell Culture, Electroporation, and Infection

All ES cell lines were routinely cultured in Glasgow's modified Eagle's medium supplemented with 10% fetal calf serum and 1,000 U/ml LIF as described previously [19]. To induce differentiation, cells were allowed to aggregate in hanging drops in ES cell medium without LIF [20]. After 2 days, embryoid bodies were collected and further grown in suspension for 1–4 days. For episomal supertransfection, embryonic day 14 T (E14/T) cells [18] were electroporated with 20 µg of supercoiled vectors at 200 V and 960 µF in a 0.4-cm cuvette. Cells were plated at 5 × 10⁴ cells per 10-cm dish and cultured in the presence of 100 µg/ml hygromycin B (Roche Diagnostics, Basel, Switzerland, <http://www.roche-applied-science.com>) or 1 µg/ml zeocine (Cayla-InvivoGen, Toulouse, France, <http://www.cayla.com>) for 7 days. Resistant colonies were pooled and further propagated in selection medium for 8–10 days prior to analysis.

Lentiviral vectors expressing shRNA were produced using the BLOCK-iT lentiviral RNAi expression system (ref. K4944-00; Invitrogen) according to the manufacturer's instructions. For lentiviral infection, CGR8 were plated at a density of 10⁴ cells in 24-well plates in 1 ml of medium composed of 100 µl of ES cell medium and 900 µl of culture supernatant from virus-producer cells. After 48 hours, ES cells were trypsinized, replated at 10⁴ cells per gelatin-coated 10-mm tissue culture dish, and further cultured in complete ES cell medium supplemented with 250 µg/ml G418 for 6 days.

Stimulation with LIF, Granulocyte Colony-Stimulating Factor, and Tamoxifen: Cell Lysates and Immunoblotting

ES cells were plated at a density of 2 × 10⁶ cells per 10-cm dish and cultured for 24 hours in medium lacking LIF. The following day, cells were stimulated with 10,000 U/ml LIF, 30 ng/ml granulocyte colony-stimulating factor (G-CSF) (Calbiochem, San Diego, <http://www.emdbiosciences.com>), or 100 nM 4'-hydroxytamoxifen (4'OHT) (Calbiochem) for the times indicated. Cells were then washed and scraped off with ice-cold phosphate-buffered saline (PBS), centrifuged, and frozen at -80°C. Cell pellets were lysed in 20 mM Hepes (pH 7.4), 100 mM NaCl, 50 mM NaF, 1% Triton X-100, 10% glycerol, 1 mM dithiothreitol, 1 mM sodium orthovanadate, 1 mM phenylmethylsulfonyl fluoride, and cocktail of protease inhibitor (Roche

www.StemCells.com

Diagnostics) for 1 hour at 4°C. Protein lysates were then cleared by centrifugation (14,000 rpm for 20 minutes). For immunoblotting, 30 µg of total proteins were resolved by SDS-polyacrylamide gel electrophoresis and electroblotted onto nitrocellulose. After overnight treatment with blocking buffer (50 mM Tris-HCl, pH 7.6, 150 mM NaCl, 5% dry milk), the membranes were probed with specific monoclonal or polyclonal antibodies (anti-Pim1 [sc-13513; Santa Cruz Biotechnology Inc., Santa Cruz, CA, <http://www.scbt.com>] and anti-Cdc25A [sc-7389; Santa Cruz Biotechnology]). Blots were incubated with horseradish peroxidase-coupled anti-mouse, anti-rabbit, or anti-goat IgG and developed using enhanced chemiluminescence reagents (Amersham Biosciences, Piscataway, NJ, <http://www.amersham.com>).

Semiquantitative and Real-Time Polymerase Chain Reaction

RNA was extracted using RNeasy kits with on-column DNase digestion, and reverse transcription was carried out with Omniscript, according to the manufacturer's recommendations (Qiagen, Hilden, Germany, <http://www.qiagen.com>). Polymerase chain reactions (PCRs) were performed with a PerkinElmer thermal cycler (PerkinElmer Life and Analytical Sciences, Boston, <http://www.perkinelmer.com>), operating on a regimen of 96°C for 5 seconds, 55°C–57°C (according to primers) for 15 seconds, and 72°C for 60 seconds, for 35 cycles, followed by 72°C for 10 minutes. Primers and annealing temperatures used to detect *Nanog*, *Oct-4*, *pim-1*, *pim-2*, *pim-3*, *cde25A*, *rex-1*, and *β-actin* by real-time PCR were as follows:

Nanog: cDNA length, 438 base pairs (bp); annealing temperature, 55°C; 5' primer (5' to 3'), TACCTCAGCCTCCAGCAGA; 3' primer (5' to 3'), CCTCTCAAGTCACTGGCAG

Oct-4: cDNA length, 974 bp; annealing temperature, 56°C; 5' primer (5' to 3'), ATGGATCCTCGAACCTGGC; 3' primer (5' to 3'), TCAGTTGAATGCATGCATGGGAG

pim-1: cDNA length, 231 bp; annealing temperature, 55°C; 5' primer (5' to 3'), GCCCTCCTTGAAAGAAATCC; 3' primer (5' to 3'), GGACCTGGAGTCTGGAATGA

pim-2: cDNA length, 212 bp; annealing temperature, 55°C; 5' primer (5' to 3'), ACATGGTCTGTGGGGACATT; 3' primer (5' to 3'), TCCTTGGAGGGAGTTGATGG

pim-3: cDNA length, 167 bp; annealing temperature, 57°C; 5' primer (5' to 3'), AGCAGTGAACCTCTGACCCCT; 3' primer (5' to 3'), TCAAGTATCCACCCAGGGCA

cde25A: cDNA length, 711 bp; annealing temperature, 52°C; 5' primer (5' to 3'), TCTGTCTAGATCTCCCTGG; 3' primer (5' to 3'), CAGATGCCATAATTCTGGAG

rex-1: cDNA length, 130 bp; annealing temperature, 55°C; 5' primer (5' to 3'), CGTGTAAACATACACCATCCG; 3' primer (5' to 3'), GAAATCCTCTCCAGAATAGG

β-actin: cDNA length, 200 bp; annealing temperature, 55°C; 5' primer (5' to 3'), TGAAACAACATACAATTCCATCATGAGTGTGA; 3' primer (5' to 3'), AGGAGCGATAATCTGATCTTCATGGTGCT.

Quantitative PCR was performed using the LightCycler 1.5 system and the LightCycler Fast Start DNA Master SYBR Green I kit (Roche Diagnostics) according to the manufacturer's instructions. Reactions were carried out in a total volume of 20 µl, comprising 0.4 µM each primer, 3.5 mM MgCl₂, 2 µl of SYBR Green, and 2 µl of diluted cDNA. Amplification and online monitoring were performed using the LightCycler 1.5 system (Roche Diagnostics). Following 40 amplification cycles, melt-curve analyses were performed to verify that only the desired PCR product had been amplified. PCR efficiency of both the target and reference genes was calculated from the derived slopes of standard curves by the LightCycler software (version 4.0; Roche Diagnostics). These PCR efficiency values were used to calculate the relative quantification values for calibrator-normalized target gene expression by the LightCycler relative quantification software (version 4.0). In all cases, target genes were normalized to *β-actin*. A *β-actin* cDNA was used as

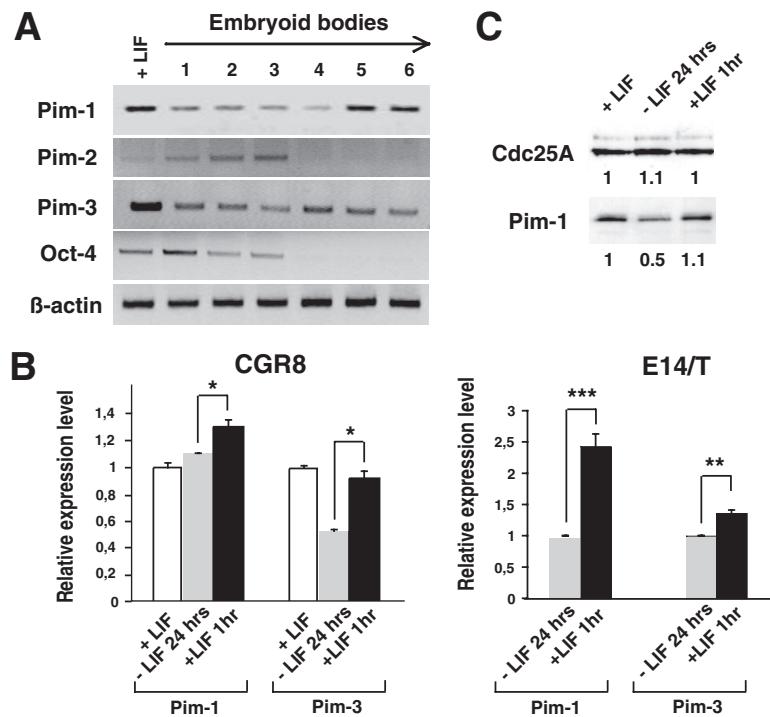


Figure 1. Regulation of *pim-1* and *pim-3* in ES cells and embryoid bodies. (A): Embryonic stem (ES) cells were induced to differentiate by withdrawal of LIF and formation of embryoid bodies for 1–6 days. Total RNA was subjected to reverse transcription polymerase chain reaction (RT-PCR) analysis to evaluate expression of *pim-1*, *pim-2*, and *pim-3*. Expression of *Oct-4* was used as control of differentiation and β -actin as control of the RT-PCR. (B): Influence of LIF on *pim-1* and *pim-3* expression in CGR8 ES cells (left panel) and E14/T ES cells (right panel). ES cells (+LIF) were plated at a density of 2×10^6 cells per 10-cm dish, LIF-starved for 24 hours (-LIF), and then restimulated with 10,000 U/ml LIF for 1 hour. Histograms represent the levels of *pim-1* and *pim-3* mRNAs measured by real time PCR in the -LIF 24 hours and +LIF 1-hour conditions normalized to the +LIF condition (means and SEs calculated from two [CGR8] or three [E14/T cell] replicates after normalization to β -actin; *, $p < .05$; **, $p < .01$; ***, $p < .001$, calculated with the Student *t* test). (C): Immunoblots of total cell lysates revealed by anti-Cdc25A and anti-Pim-1 antibodies. Experimental conditions were as described for (B). Numbers indicate levels of expression normalized to the +LIF condition. Abbreviations: E14, embryonic day 14 hr; LIF, leukemia inhibitory factor.

normalization control to ensure constant β -actin levels in all samples regardless of experimental conditions.

Abcys, Paris, France, <http://www.abcysonline.com>) was used to identify apoptotic cells. All reactions were performed according to the manufacturer's instructions.

Detection of Alkaline Phosphatase Activity

Dishes were fixed in methanol for 15 minutes and then stained for 15 minutes with a solution containing 1 mg/ml Fast Red TR salt (Sigma-Aldrich, St. Louis, <http://www.sigmaldrich.com>) dissolved in 0.1 M Tris, pH 9.2, containing 200 mg/ml naphtol AS-MX phosphate [21, 22].

Detection of Stem Cell-Specific Embryonic Antigen-1 Expression

ES cells were trypsinized, and the cell suspension incubated in culture medium for 6 hours on a rocker platform. Cells were subsequently Fc-blocked by treatment with 1 μ g of human IgG per 10^5 cells for 15 minutes at room temperature. Then, 10^5 cells in 25 μ l of PBS supplemented with 0.5% bovine serum albumin were incubated for 30–45 minutes at 4°C with 50 μ l (50 μ g/ml) of the phycoerythrin-conjugated mouse monoclonal anti-stem cell-specific embryonic antigen-1 (anti-SSEA1) (ref. FAB2155P; R&D Systems Inc., Minneapolis, <http://www.rndsystems.com>). Cells were washed twice with PBS. Fluorescence intensity was determined by flow cytometry on a FACScan instrument (FACSCanto II; BD Biosciences, San Diego, <http://www.bdbiosciences.com>), and data acquisition was performed using the FACSDiva software (BD Biosciences).

Terminal Deoxynucleotidyl Transferase dUTP Nick-End Labeling Assay

The ApopTag Red In Situ Apoptosis Detection Kit (ref. S7165;

RESULTS

Influence of LIF on the Transcriptional Regulation of *pim-1* and *pim-3*

Expression of *pim-1*, *pim-2*, and *pim-3* was examined by semi-quantitative RT-PCR in undifferentiated CGR8 cells and in 1–6-day-old embryoid bodies (Fig. 1A). ES cells expressed elevated levels of *pim-1* and *pim-3* transcripts, both of which declined abruptly during the first day of differentiation. By contrast, the *pim-2* transcript level increased during the first 2 days of differentiation and dropped from third day onwards. Thus, *pim-1* and *pim-3* exhibit a unique expression pattern characterized by a drop in transcript levels within 24 hours following withdrawal of LIF. Downregulation of *pim-1* and *pim-3* expression takes place prior to the loss of pluripotency marker *Oct-4* [23], occurring between day 3 and 4.

To examine the role of LIF in the regulation of *pim-1* and *pim-3* expression, CGR8 ES cells were LIF-deprived for 24 hours and then restimulated with 10,000 U/ml LIF for 1 hour. Withdrawal of LIF for 24 hours did not induce loss of ES cell pluripotency (supplemental online data 1). This stimulation protocol led to a strong activation of the LIF signaling pathway revealed by a 14-fold increase in the mRNA level of the LIF/STAT3 target gene *junB* [4, 6] (data not shown). *pim-1* and *pim-3* transcript levels increased 1.3 ± 0.1 -fold ($p < .06$) and 1.74 ± 0.13 -fold ($p < .05$) respectively, after LIF stimulation (Fig. 1B, left panel). Activation of *pim-1* and *pim-3* after stimulation by LIF was also observed in E14/T ES cells (*pim-1*: 2.4 ± 0.2 -fold, $p < .001$; *pim-3*: 1.35 ± 0.05 -fold, $p < .01$)

STEM CELLS

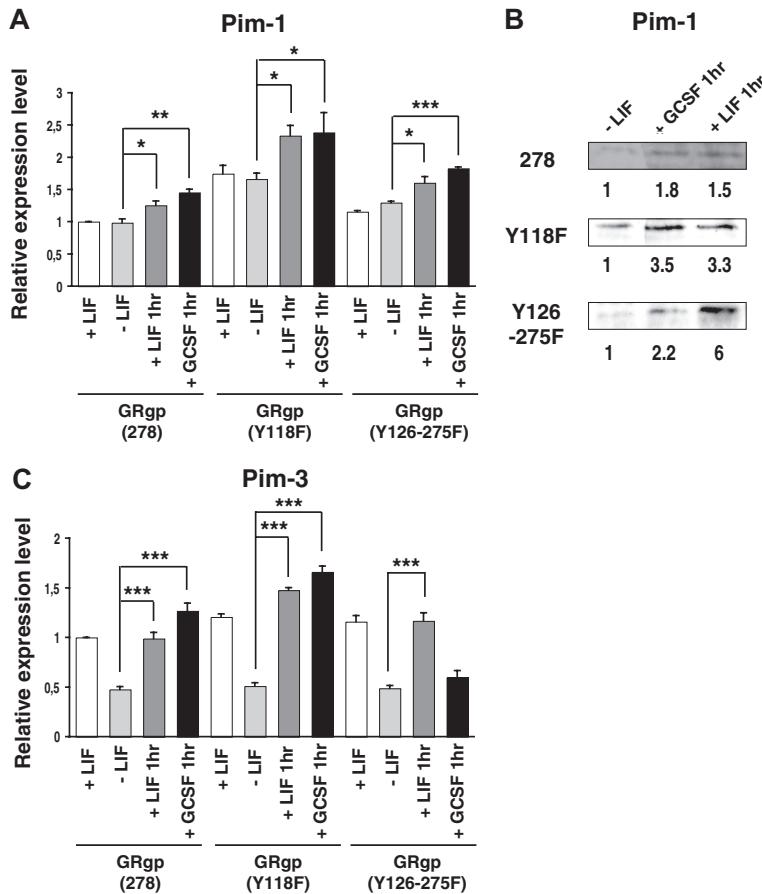


Figure 2. Regulation of *pim-1* (A, B) and *pim-3* (C) expression in embryonic stem (ES) cells expressing GRgp278, GRgpY126-275F, and GRgpY118F chimeric receptors. (A–C): ES cells (+LIF) were deprived of LIF for 24 hours (−LIF) and then restimulated either with 10,000 U/ml LIF (+LIF 1 hour) or with 30 ng/ml GCSF (+GCSF 1 hour). mRNA levels of *pim-1* and *pim-3* were analyzed by real-time polymerase chain reaction and normalized with β -actin cDNA. Histograms shown in (A) and (C) represent mRNA levels measured in the −LIF, +LIF 1-hour, and +GCSF 1-hour conditions and normalized to the +LIF condition (means and SEs calculated in three replicates; *, $p < .05$; **, $p < .01$; ***, $p < .001$; calculated with the Student *t* test). (B): Immunoblots of total cell lysates revealed by anti-Pim-1 antibody. Numbers indicate levels of expression normalized to the −LIF condition. Abbreviations: GCSF, granulocyte colony-stimulating factor; GRgp, G-CSF receptor fused to gp130; hr, hour; LIF, leukemia inhibitory factor.

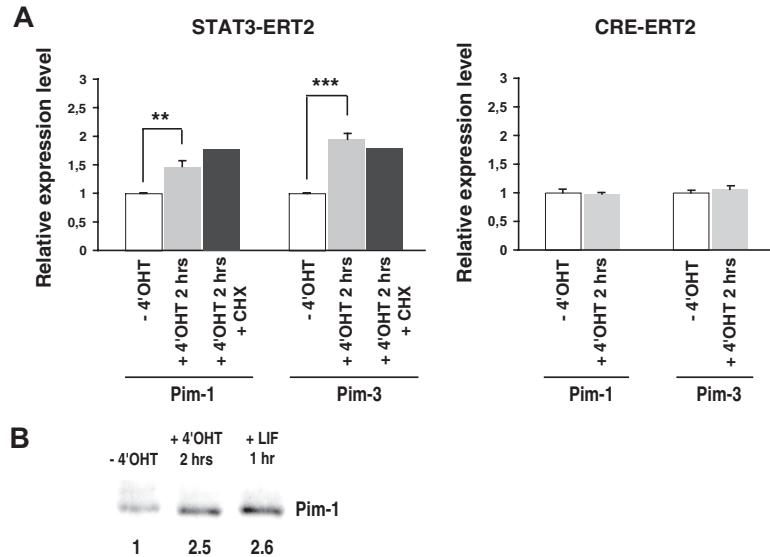
(Fig. 1B, right panel). The level of *pim-1* expression also increased at the protein level in response to LIF stimulation (2.1 ± 0.4 -fold, $p < .01$). In contrast, the levels of Cdc25A phosphatase, a substrate of Pim-1 [24], remained unchanged (Fig. 1C). Thus, the levels of *pim-1* and *pim-3* mRNA and the level of Pim-1 kinase were upregulated as a consequence of LIF receptor stimulation in ES cells.

Contribution of STAT3 to the Transcriptional Regulation of *pim-1* and *pim-3*

We made use of engineered ES cell lines stably expressing GRgp278, GRgp(Y118F), and GRgp(Y126-275F) receptors [3, 5] to study the role of STAT3 in the regulation of *pim-1* and *pim-3* expression. GRgp278 is a chimeric receptor consisting of the extracellular domain of the G-CSF receptor fused to the transmembrane and cytoplasmic region of gp130. Homodimerization of the GRgp278 receptor, mediated by G-CSF, induces recruitments of both STAT3 and SHP-2 and promotes self-renewal. GRgp(Y118F) is a mutated chimeric receptor in which phenylalanine has been substituted for tyrosine 118 in the intracytoplasmic domain, which prevents the recruitment of SHP-2 by the activated receptor. In the GRgp(Y126-275F) receptor, phenylalanines have been substituted for tyrosine 126 and 275, which strongly reduces recruitment of STAT3 by the activated receptor (supplemental online data 2). ES cells expressing wild-type, SHP-2 binding-deficient, and STAT3 binding-deficient chimeric receptors were LIF-deprived for 24 hours

and restimulated for 1 hour either with 10,000 U/ml LIF, to activate the endogenous gp130 receptor, or with 30 ng/ml G-CSF, to activate the chimeric receptor. Stimulation of GRgp278 (wild-type) with G-CSF resulted in increased levels of *pim-1* (1.6 ± 0.2 -fold, $p < .01$) and *pim-3* (2.8 ± 0.4 -fold, $p < .001$) mRNAs (Fig. 2A, 2C). Similarly, stimulation of the GRgp(Y118F) cells with G-CSF resulted in increased levels of *pim-1* (1.5 ± 0.25 -fold, $p < .05$) and *pim-3* (3.35 ± 0.45 -fold, $p < .001$) mRNAs, indicating that recruitment of SHP-2 to gp130 receptor is not required for the transcriptional activation of *pim-1* and *pim-3* genes. In the STAT3 binding-deficient cell line (GRgp(Y126-275F)), LIF and G-CSF were equally efficient at activating the expression of *pim-1*, indicating that recruitment of STAT3 to the activated gp130 receptor is not essential to the transcriptional activation of *pim-1*. Examination of protein levels confirmed the activation of Pim-1 expression in response to LIF and G-CSF stimulation in the cell lines expressing the wild-type, SHP-2-deficient, and STAT3 binding-deficient receptors (Fig. 2B). By contrast, G-CSF failed to activate the expression of *pim-3* in GRgp(Y126-275F), indicating that activation of STAT3 is required for the transcriptional activation of *pim-3* (Fig. 2C).

pim-1 is a target gene of STAT3 that contributes to interleukin-6-dependent regulation of proliferation in lymphoid cells [14]. This prompted us to examine further whether STAT3 activity could regulate *pim-1* expression in the absence of gp130 receptor activation. To this end, a conditional mutant of STAT3 was generated by fusing the STAT3 coding sequence to ER^{T2},



a mutated ligand-binding domain of the human estrogen receptor [17, 25]. The STAT3-ERT^{T2} fusion protein was overexpressed by means of supertransfection of the *pPCAGIZ-STAT3-ERT^{T2}* episomal expression vector into the E14/T-cell line in the presence of 4'OH [18]. *pPCAGIZ-STAT3-ERT^{T2}* transfectants formed a uniform population of undifferentiated ES cells that was indistinguishable from the sister population of *pPCAGIZ-STAT3-ERT^{T2}* transfectants propagated in the presence of LIF (data not shown). To assess the regulation of *pim-1* and *pim-3* by STAT3 activity, STAT3-ERT^{T2} transfectants were deprived of 4'OH for 24 hours and then restimulated with 100 nM 4'OH for 2 hours (Fig. 3A). This resulted in increased levels of *pim-1* (1.5 ± 0.1 -fold, $p < .01$) and *pim-3* (2.0 ± 0.1 -fold, $p < .001$) mRNAs. Stimulation by 4'OH in the presence of cycloheximide led to a 1.7-fold increase in mRNA levels, indicating that 4'OH-induced transcriptional activation of *pim-1* and *pim-3* was not dependent on de novo protein synthesis. The observed effect is STAT3-specific since E14/T ES cells expressing a hormone-dependent Cre-ERT^{T2} showed no increase in *pim-1* and *pim-3* expression following stimulation with 4'OH. The increase of *pim-1* mRNA level observed in response to 4'OH was paralleled by a 2.5-fold increase in protein level (Fig. 3B). Together, these data indicate that the expression of *pim-1* and *pim-3* genes is regulated by STAT3 activity in ES cells.

Contribution of Pim-1 and Pim-3 Kinases to the Inhibition of ES Cell Differentiation

The role of Pim kinases in the inhibition of ES cell differentiation was examined by RNA interference, making use of lentiviral vectors encoding small hairpin RNA (shRNA) for *pim-1* (*pLenti6/BLOCK-iT-PGKneo-pim1*), *pim-3* (*pLenti6/BLOCK-iT-PGKneo-pim3*), and *GFP* (*pLenti6/BLOCK-iT-PGKneo-GFP*) and the *neo*^r selectable gene. Forty-eight hours after infection, CGR8 ES cells were replated at clonal density and further cultured in selection medium for 7 days to kill noninfected cells (approximately 50% of the population). The percentages of undifferentiated (alkaline phosphatase⁺ [AP⁺]), mixed (AP⁺/AP⁻), and differentiated (AP⁻) G418-resistant colonies were calculated (Fig. 4A). CGR8 ES cells expressing the shGFP maintained undifferentiated AP⁺ colonies at between 71% and 76% in this assay. Expression of shPim-1 or shPim-3 reduced, albeit moderately, the proportion of undifferentiated

Figure 3. Regulation of *pim-1* and *pim-3* by hormone-dependent STAT3-ERT^{T2}. Embryonic day 14 T cells expressing STAT3-ERT^{T2} (left panel) or Cre-ERT^{T2} (right panel) were deprived of 4'OH for 24 hours and then restimulated with 100 nM 4'OH for 2 hours in the presence of 100 µg/µl CHX or not. (A): mRNA levels of *pim-1* and *pim-3* were analyzed by real-time polymerase chain reaction and normalized with *β-actin* cDNA. Histogram represents mRNA levels measured in the +4'OH 2-hour and +4'OH 2-hour + CHX conditions and normalized to the -4'OH condition (means and SEs calculated in three replicates; *, $p < .05$; **, $p < .01$; ***, $p < .001$; calculated with the Student *t* test). (B): Pim-1 protein level analyzed by immunoblotting with anti-Pim-1 antibody. Numbers indicate levels of expression in +4'OH 2-hours and +LIF 1-hour conditions normalized to the -4'OH condition. Abbreviations: CHX, cycloheximide; hr, hour; LIF, leukemia inhibitory factor; 4'OH, 4'-hydroxytamoxifen; STAT3, signal transducer and activator of transcription-3.

colonies compared with control (shPim-1, -20%, $p < .01$; shPim-3, -27.9%, $p < .001$). The decrease in percentage of undifferentiated AP⁺ colonies observed in ES cells expressing shPim-1 and shPim-3 was paralleled by an increase in the percentage of mixed AP⁺/AP⁻ colonies (shPim-1, +17.3%, $p < .01$; shPim-3, +22.3%, $p < .001$) as well as of AP⁻, flat, differentiated colonies (shPim-1, +2.6%, $p < .05$; shPim-3, +5.5%, $p < .01$). Expression of *rex-1*, a marker of pluripotency [26], was analyzed by quantitative RT-PCR and showed 25% and 49% of reduction in transcript levels in cells expressing shPim-1 and shPim-3, respectively, compared with cells expressing shGFP (Fig. 4B). Cells expressing shSTAT3, a shRNA directed to STAT3, showed a 90% reduction in *rex-1* transcript levels.

Gene expression analysis by real-time PCR showed that the levels of *pim-1* and *pim-3* transcripts were reduced only twofold in the infected cell population (data not shown), possibly explaining the relatively small differences observed in the proportions of undifferentiated versus mixed/differentiated colonies in the self-renewal assay. CGR8 cells expressing shGFP, shPim-1, and shPim-3 were subcloned, and the level of expression of *pim-1* and *pim-3* was analyzed to select clones showing the highest interference (Fig. 4C). Clones 1, 2, and 7 (shPim-1) and clones 1, 3, and 5 (shPim-3)—all six showing a residual expression of *pim-1* and *pim-3* between 20% and 38% with respect to levels measured in the shGFP control—were selected for further analysis. All shPim-1 and shPim-3 clones formed 3.3–16-fold fewer AP⁺ undifferentiated colonies and 3.2–4-fold more AP⁻ differentiated colonies than the shGFP clone in the self-renewal assay (Fig. 4D). The stronger the interference of *pim-1* and *pim-3* expression is, the lower the percentage of undifferentiated colonies. Undifferentiated, mixed, and differentiated colonies were pooled and subsequently analyzed for the expression of pluripotency markers. All shPim-1 and shPim-3 clones displayed reduced levels of transcripts encoding Oct-4 (50%–66% reduction) and Nanog (44%–73% reduction) (Fig. 4E). They also displayed lower percentages of cells expressing the SSEA-1 antigen, a marker of pluripotent stem cells (shPim-1, 32%–50%; shPim-3, 25%–40%) compared with control cells (shGFP, 80%, $p < .001$) (Fig. 4F). In routine culture, all shPim-1 and shPim-3 clones displayed higher rates of spontaneous differentiation compared with the shGFP control. Interestingly, interference in

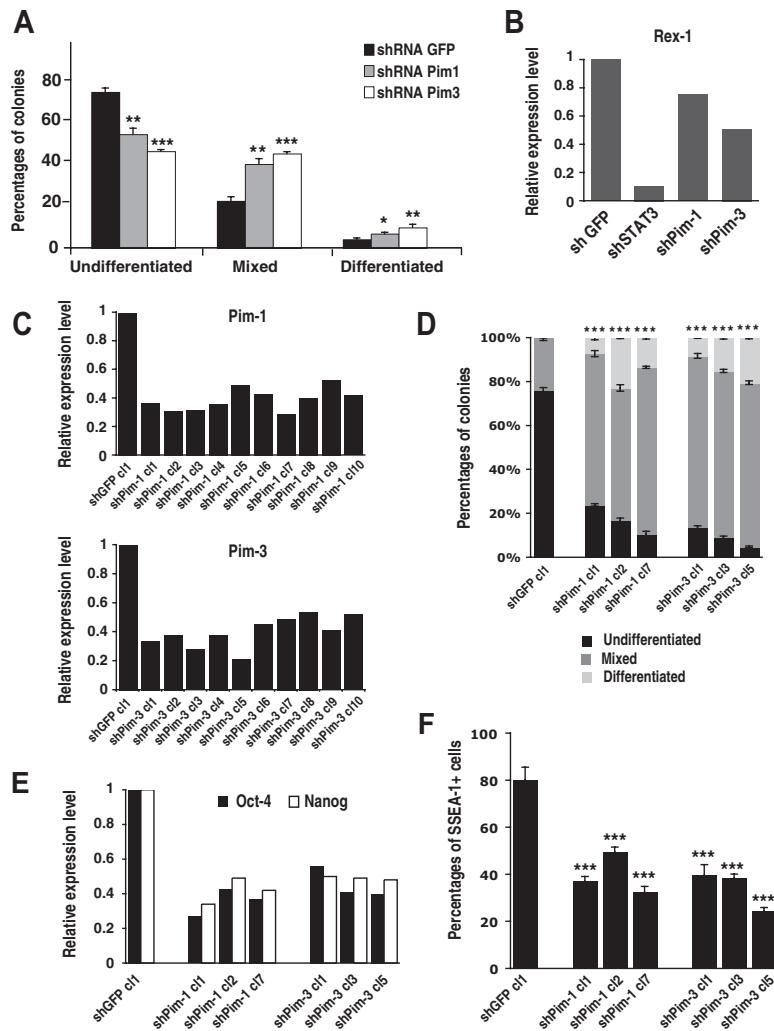


Figure 4. Knockdown of *pim-1* and *pim-3* expression by lentiviral vector-mediated RNA interference and evaluation of self-renewal in a clonal assay. (A): Percentages of undifferentiated, mixed, and differentiated colonies after staining to reveal alkaline phosphatase (AP) activity. (B): Level of *rex-1* mRNA in colonies harvested in (A), analyzed by real-time polymerase chain reaction (PCR) and normalized with *β-actin* cDNA. Histogram represents *rex-1* mRNA levels measured in colonies expressing shPim1, shPim3, and shSTAT3 after normalization to the *rex-1* mRNA level measured in embryonic stem (ES) cells expressing shGFP. (C): *pim-1* and *pim-3* expression in ES cell clones resulting from infection with the lentiviral vectors expressing shPim1 (top panel) and shPim3 (bottom panel). Levels of *pim-1* and *pim-3* mRNAs in shPim1 and shPim3 clones are normalized to levels measured in a shGFP clone. (D): Percentages of undifferentiated, mixed, and differentiated colonies after staining to reveal AP activity, obtained with shGFP clone 1; shPim1 clones 1, 2, and 7; and shPim3 clones 1, 3, and 5. (E): Level of *Oct-4* and *Nanog* mRNAs in colonies harvested in (D), analyzed by real-time PCR and normalized with *β-actin* cDNA. Histogram represents *Oct-4* and *Nanog* mRNA levels measured in colonies expressing shPim1 (clones 1, 2, and 7) and shPim3 (clones 1, 3, and 5), after normalization to the *Oct-4* and *Nanog* mRNA levels measured in colonies expressing shGFP (clone 1). (F): Percentage of SSEA-1⁺ cells in colonies harvested in (D) and analyzed by flow cytometry. (A, D, F): Means and SEs were calculated from three replicates. *, $p < .05$; **, $p < .01$; ***, $p < .001$; calculated with the Student *t* test (A, F) and the χ^2 test (D). Abbreviations: GFP, green fluorescent protein; sh, small hairpin; SSEA-1, stem cell-specific embryonic antigen-1; STAT3, signal transducer and activator of transcription-3.

the shPim1 and shPim3 clones was progressively lost during serial passagings, indicating that ES cells with low Pim-1 and Pim-3 content are counterselected (data not shown). These results demonstrate that expression of Pim-1 and Pim-3 kinases is required to inhibit differentiation of ES cells.

Overexpression of Pim-1 and Pim-3 Kinases Delays LIF-Induced Differentiation

Given that Pim-1 and Pim-3 kinases are necessary for self-renewal of ES cells, we asked whether overexpression of Pim-1 and Pim-3 would allow them to overcome spontaneous differentiation induced by LIF starvation. Pim-1 and Pim-3 kinases were overexpressed by means of supertransfection of *pPHPGK-pim1* and *pPHPGK-pim3* episomal expression vectors into the E14/T-cell line [18]. Engineered ES cells were plated at clonal density and further cultured for 7 days in the presence of LIF. The percentages of undifferentiated, mixed, and differentiated colonies were calculated (Fig. 5A). ES cells overexpressing Pim-1 and Pim-3 formed significantly more undifferentiated colonies (Pim-1, +5.3%, $p = .05$; Pim-3, +17.9%, $p < .001$) and significantly fewer mixed colonies (Pim-1, -5.6%, $p < .05$; Pim-3, -17.8%, $p < .001$) than control cells. These data indicate that overexpression of Pim-1 and Pim-3 reduces the rate of spontaneous differentiation in the presence of LIF.

In a second step, Pim-1 and Pim-3 transfectants were subjected to a clonal assay combined with a LIF rescue test to evaluate the recovery of ES cells following a short period of LIF starvation [27]. Cells were exposed to medium without LIF for 12, 24, 36, and 48 hours and were subsequently cultured for 5–7 days in medium supplemented with LIF. Then, the proportion of undifferentiated colonies was counted to assess the LIF rescue efficiency. In Pim-1 and Pim-3 transfectants, as in control cells, the proportion of undifferentiated colonies increasingly decreased with the duration of LIF deprivation, which indicates that overexpression of Pim-1 and Pim-3 is unable to fully sustain ES self-renewal in the absence of LIF (Fig. 5B). However, Pim-1 and Pim-3 transfectants showed a reduced sensitivity to LIF starvation. This was evidenced by a slower decrease in the proportion of undifferentiated colonies in ES cells overexpressing Pim-1 and Pim-3, compared with ES cells transfected with the empty vector (*pim-1* vs. control, $p < .001$; *pim-3* vs. control, $p < .05$). Together, these results are consistent with Pim-1 and Pim-3 overexpression partially

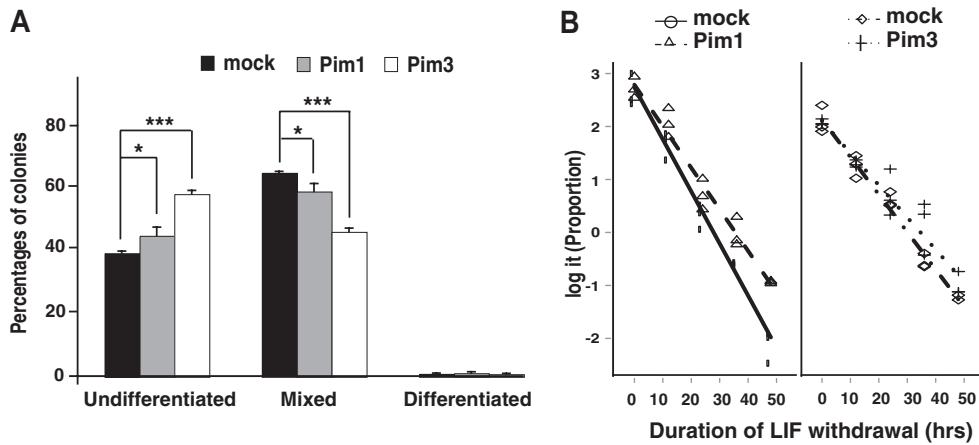


Figure 5. Overexpression of Pim-1 and Pim-3 by means of episomal supertransfection. (A): Embryonic day 14 (E14)/T (E14/T) embryonic stem (ES) cells were electroporated with *pPHPGK* (control), *pPHPGK-pim1*, and *pPHPGK-pim3* episomal vectors. ES cell transfectants were subsequently plated at 1×10^3 cells per 100-mm tissue culture dish and further cultured in complete ES cell medium for 7 days. Histograms represent the means and SEs, measured in three replicates, of the percentages of undifferentiated, mixed, and differentiated colonies. (B): E14/T ES cells were electroporated with *pPHCAG*, *pPHCAG-pim1*, *pPCAGIZ*, and *pPCAGIZ-pim3* episomal vectors and propagated in culture medium supplemented with hygromycin B (*pPHCAG* and *pPHCAG-pim1*) or with zeocin (*pPCAGIZ* and *pPCAGIZ-pim3*) for 15 days. The four ES cell transfectants were subsequently plated at 1×10^3 cells per 100-mm tissue culture dish. Cells were exposed to medium without LIF for 0, 12, 24, 36, or 48 hours and were subsequently cultured in complete ES cell medium with 1,000 U/ml LIF for 5–7 days (LIF rescue assay). Curves are based on the variations in the proportion of undifferentiated colonies as a function of the duration of LIF withdrawal in *pPHCAG* (control) and *pPHCAG-pim1* transfectants (left panel), or in *pPCAGIZ* (control) and *pPCAGIZ-pim3* transfectants (right panel) (hereafter termed the four conditions). The proportion of undifferentiated colonies as a function of the duration of LIF removal was modeled using logistic regression (generalized linear model with a binomial family). According to this model, the proportions, p , transformed as $\text{logit}(p) = \ln(p/[1 - p])$ are related linearly to the explanatory variables, gene and time. However, we noted a systematic oscillation in the residuals as a function of time for each condition on the transformed scale due to the upper and lower asymptotes differing from 1 and 0, respectively. This deviation from linearity was well described by a fixed sine wave across the four conditions, with frequency f estimated to be 0.025 cycles per hour (a period of approximately 40 hours) over the four conditions. This value was used in the subsequent analysis. The final model required four estimates of intercept and slope (a, b), one for each condition, and one estimate of the sine and cosine coefficients (c, d) for the four conditions [$\text{logit}(p) = a + b \times \text{time} + c \times \sin(2 \times \pi \times f \times \text{time}) + d \times \cos(2 \times \pi \times f \times \text{time})$]. After taking into account differences in intercept and the residual oscillatory component, the differences between slopes within each group were significant, indicating that overexpression of pim-1 and Pim-3 decreased the rate of decrease in the proportion of undifferentiated colonies as a function of the duration of LIF starvation (Pim-1 vs. control: diff = 0.02, SE (standard error) = 0.003, $z = 5.75$, $p < .001$; Pim-3 vs. control: diff = 0.009, SE = 0.004, $z = 2.08$, $p < .05$). The figure shows the data on the logit scale after normalization by the intercepts and removal of the oscillatory component with the best-fitting straight lines for each condition. Abbreviations: hrs, hours; LIF, leukemia inhibitory factor.

counterbalancing the effects of LIF deprivation on induction of differentiation.

Contribution of Pim-1 and Pim-3 Kinases to the Regulation of ES Cell Proliferation and Apoptosis

In the LIF-rescue experiment reported in Figure 5B, we observed that the colonies generated from Pim-1 transfectants were larger than the control colonies (data not shown), suggesting that Pim-1 is also involved in the regulation of ES cell growth. We therefore asked whether *pim-1* and *pim-3* knockdowns are detrimental to ES cell growth, making use of the ES cells expressing *pim-1*, *pim-3*, and GFP-specific shRNA. Because knockdown of *pim-1* and *pim-3* expression increases the rate of spontaneous differentiation and ES cell differentiation is associated with increased cell-cycle duration [19], we thought to analyze the effect of *pim-1* and *pim-3* knockdown on the size of residual undifferentiated AP⁺ colonies in the clonal assay. We observed that AP⁺ colonies expressing the *pim-1* and *pim-3* shRNA were, respectively, 31% and 41% smaller than AP⁺ colonies expressing the GFP shRNA ($p < .001$) (Fig. 6A). Terminal deoxynucleotidyl transferase dUTP nick-end labeling analysis of the undifferentiated colonies showed 28% (*pim-1*) and 35% (*pim-3*) increases in the number of positive cells ($p <$

.05), indicating that knockdown of *pim-1* and *pim-3* increases the rate of apoptosis (Fig. 6B, 6C). Together, these results are consistent with *pim-1* and *pim-3* contributing to the regulation of ES cell growth by inhibiting apoptosis.

DISCUSSION

First, we have shown that the expression of *pim-1* and *pim-3* can be activated by LIF in two different ES cell lines. We also demonstrated that *pim-1* and *pim-3* expression can be activated by hormone-dependent STAT3 in the presence of protein synthesis inhibitor. Together, these findings establish *pim-1* and *pim-3* as novel target genes of STAT3 in ES cells. The *pim-1* promoter is known to contain a DNA sequence that potentially binds STAT oligomers [14]. However, ES cells expressing the STAT3 binding-deficient G-CSF:gp130 chimeric receptor partially activate the expression of *pim-1* in response to G-CSF. This observation suggests that in the absence of STAT3 recruitment, another pathway can activate *pim-1*. Candidate pathways are the LIF/PI3K and LIF/Yes pathways, both of which contribute to inhibiting ES cell differentiation [22, 28]. Furthermore, STAT3 is not the only factor regulating *pim-1* and *pim-3* expression in ES cells. It was observed that deprivation of LIF for 24 hours reduces the

STEM CELLS

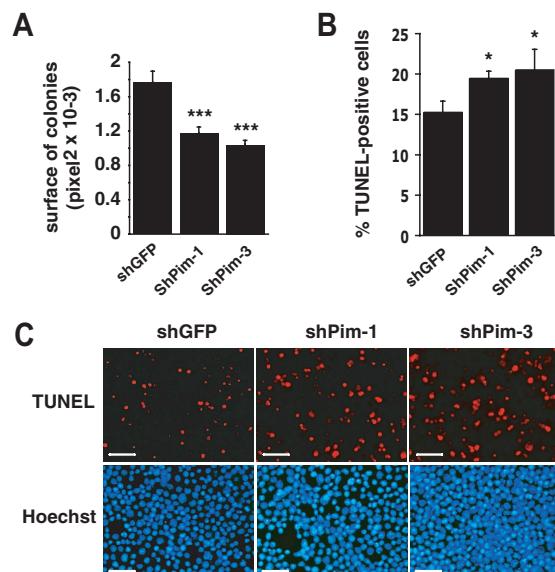


Figure 6. Knockdown of *pim-1* and *pim-3* expression by lentiviral vector-mediated RNA interference and evaluation of colony size (**A**) and apoptosis (**B, C**) in a clonal assay. The embryonic stem cell colonies expressing shPim-1, shPim-3, and shGFP were obtained as described in Figure 4A. (**A, B**): Histograms represent the means and SEs, measured in three replicates, of the surface of undifferentiated AP⁺ colonies (**A**) and the percentage of TUNEL-positive cells (**B**) (*, $p < .05$; ***, $p < .001$; calculated with the Student *t* test). (**C**): Photomicrographs of undifferentiated colonies expressing shGFP, shPim-1, and shPim-3, showing TUNEL-positive cells stained with Hoechst (representative fields). Scale bar = 5 μ m. Abbreviations: GFP, green fluorescent protein; sh, small hairpin; TUNEL, terminal deoxynucleotidyl transferase dUTP nick-end labeling.

pim-3 RNA level to 50% of its original level, and *pim-1* RNA level is not decreased by this treatment. This indicates that factors other than LIF and STAT3 contribute to the regulation of *pim-1* and *pim-3* expression. Loh et al. reported the mapping of Nanog-binding sites in ES cells [29]. Mining their expression profiling data revealed that Nanog-binding sites are present in the promoter of *pim-1*. Furthermore, we observed that *pim-1* and *pim-3* expression is downregulated upon Nanog knockdown in CGR8 ES cells (unpublished data). Hence, the residual expression of *pim-1* and *pim-3* observed after withdrawal of LIF might result from a Nanog-dependent transcriptional activation mechanism.

Second, we demonstrated that overexpression of Pim-1 and Pim-3 decreases the rate of spontaneous differentiation in a clonal assay and enhances the resistance of ES cells to LIF starvation in a LIF rescue assay, whereas the knockdown of Pim-1 and Pim-3 increases the rate of spontaneous differentiation. These observations indicate that Pim-1 and Pim-3 kinases contribute to maintain ES cell pluripotency. Overexpression of Pim-1 and Pim-3 delayed the appearance of mixed and differentiated colonies in the LIF rescue assay but failed to block differentiation in the absence of LIF. This is also true of *klf-4* and *c-myc*, both of which have been shown to contribute to the inhibition of ES cell differentiation [8, 9]. ES cells overexpressing Klf-4 demonstrated a trend toward reduced differentiation compared with wild-type cells but ultimately differentiated like their wild-type counterparts [9]. Similarly, ES cells overexpressing a hormone-dependent Myc-ER and propagated in the presence of tamoxifen had the ability to self-renew in the absence of LIF, but this property

was progressively lost with time in culture [8]. Since none of these four genes are able to mimic the effect of LIF on ES cell self-renewal when they are individually overexpressed, it is likely that the four of them have to act cooperatively to block ES cell differentiation and maintain pluripotency. A synergistic action of Myc and Pim-1 would be particularly worthy of examination. The *c-myc* and *pim-1* genes are activated at comparable levels in response to LIF or to 4'OH-T in the STAT3-ER^{T2} ES cells (data not shown). Myc is a transcriptional regulator of the *Cdc25A* gene [30], and Pim-1 phosphorylates, and thereby activates, the phosphatase Cdc25A, which is a positive G1-specific cell cycle regulator [24, 31]. Such a synergistic action of Myc and Pim-1 has been evidenced in lymphoid cells, where it regulates interleukin-6/STAT3-dependent proliferation [14].

Third, we showed that Pim-1 and Pim-3 regulate the growth rate of the population of undifferentiated ES cells. The method used is based on the calculation of the size of undifferentiated colonies obtained in self-renewal assays after knockdown of Pim-1 and Pim-3. This method allowed us to measure the growth rate of undifferentiated ES cells and, therefore, to distinguish the alterations in growth rates resulting from alterations in the proliferation rate of undifferentiated cells and those indirectly resulting from differentiation [19]. Inhibition of apoptosis is one mechanism by which Pim-1 and Pim-3 regulate the growth rate of undifferentiated ES cells. Pim kinases are known to play an important role in the prevention of cell death by inactivating the proapoptotic protein Bad [32, 33]. Antiapoptotic effects of Pim kinases have been demonstrated in several experimental systems [14, 16]. LIF has been shown to inhibit apoptosis in ES cells [6, 34]. Activation of Pim-1 and Pim-3 expression could thus be one mechanism by which LIF exerts its antiapoptotic effect.

To conclude, we provide evidence that the serine/threonine kinases Pim-1 and Pim-3 play an important role in the maintenance of the ES cell identity. This findings contribute to further deciphering the mechanisms by which the LIF/STAT3 pathway sustain self-renewal of mouse ES cells.

ACKNOWLEDGMENTS

We thank Dr. Austin Smith for the gift of *pPCAGIZ*, *pPHCAG*, and *pPHPGK* plasmids, as well as E14/T cells; Dr. Tom Burdon for the gift of ES cells expressing the G-CSF:gp130 chimeric receptors; and Dr. Anton Berns for the gift of mouse *pim-1* and *pim-3* cDNAs. This work was supported by research grants from Institut National de la Santé et de la Recherche Médicale (INSERM), Association Française contre la Myopathie and Vaincre La Mucoviscidose (contract no. 4CS18H), INSERM AVENIR 2002 program, and the European Union sixth Framework Programme (FunGenES, contract no. LSHG-CT-2003-503494), Région Rhône-Alpes (Thématisques prioritaires 2003–2005), and Fondation Bettencourt-Schueller. C.S. and A.M. were recipients of a fellowship from the Association pour la Recherche contre le Cancer and the Ligue Nationale contre le Cancer. I.A., C.S., and P.-Y.B. contributed equally to this work.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

The authors indicate no potential conflicts of interest.

REFERENCES

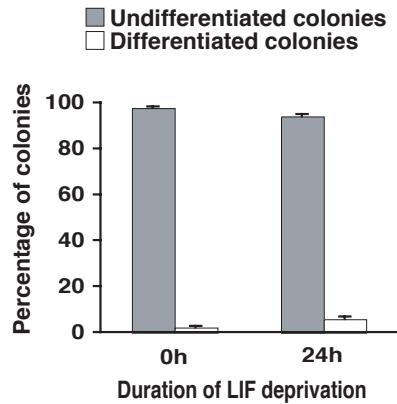
- 1 Burdon T, Chambers I, Stracey C et al. Signaling mechanisms regulating self-renewal and differentiation of pluripotent embryonic stem cells. *Cells Tissues Organs* 1999;165:131–143.
- 2 Boeuf H, Hauss C, Graeve FD et al. Leukemia inhibitory factor-dependent transcriptional activation in embryonic stem cells. *J Cell Biol* 1997;138:1207–1217.
- 3 Niwa H, Burdon T, Chambers I et al. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev* 1998;12:2048–2060.
- 4 Matsuda T, Nakamura T, Nakao K et al. STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *EMBO J* 1999;18:4261–4269.
- 5 Burdon T, Stracey C, Chambers I et al. Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. *Dev Biol* 1999;210:30–43.
- 6 Duval D, Reinhardt B, Kedinger C et al. Role of suppressors of cytokine signaling (Socs) in leukemia inhibitory factor (LIF)-dependent embryonic stem cell survival. *FASEB J* 2000;14:1577–1584.
- 7 Oka M, Tagoku K, Russell TL et al. CD9 is associated with leukemia inhibitory factor-mediated maintenance of embryonic stem cells. *Mol Biol Cell* 2002;13:1274–1281.
- 8 Cartwright P, McLean C, Sheppard A et al. LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development* 2005;132:885–896.
- 9 Li Y, McClintick J, Zhong L et al. Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. *Blood* 2005;105:635–637.
- 10 Bachmann M, Moroy T. The serine/threonine kinase Pim-1. *Int J Biochem Cell Biol* 2005;37:726–730.
- 11 Allen JD, Berns A. Complementation tagging of cooperating oncogenes in knockout mice. *Semin Cancer Biol* 1996;7:299–306.
- 12 Allen JD, Verhoeven E, Domen J et al. Pim-2 transgene induces lymphoid tumors, exhibiting potent synergy with c-myc. *Oncogene* 1997;15:1133–1141.
- 13 Stout BA, Bates ME, Liu LY et al. IL-5 and granulocyte-macrophage colony-stimulating factor activate STAT3 and STAT5 and promote Pim-1 and cyclin D3 protein expression in human eosinophils. *J Immunol* 2004;173:6409–6417.
- 14 Shirogane T, Fukada T, Muller JM et al. Synergistic roles for Pim-1 and c-Myc in STAT3-mediated cell cycle progression and antiapoptosis. *Immunity* 1999;11:709–719.
- 15 White E. The pims and outs of survival signaling: Role for the Pim-2 protein kinase in the suppression of apoptosis by cytokines. *Genes Dev* 2003;17:1813–1816.
- 16 Adam M, Pogacic V, Bendit M et al. Targeting PIM kinases impairs survival of hematopoietic cells transformed by kinase inhibitor-sensitive and kinase inhibitor-resistant forms of Fms-like tyrosine kinase 3 and BCR/ABL. *Cancer Res* 2006;66:3828–3835.
- 17 Indra AK, Warot X, Brocard J et al. Temporally-controlled site-specific mutagenesis in the basal layer of the epidermis: Comparison of the recombinase activity of the tamoxifen-inducible Cre-ER(T) and Cre-ER(T2) recombinases. *Nucleic Acids Res* 1999;27:4324–4327.
- 18 Aubert J, Dunstan H, Chambers I et al. Functional gene screening in embryonic stem cells implicates Wnt antagonism in neural differentiation. *Nat Biotechnol* 2002;20:1240–1245.
- 19 Savatier P, Lapillonne H, van Grunsven LA et al. Withdrawal of differentiation inhibitory activity/leukemia inhibitory factor upregulates D-type cyclins and cyclin-dependent kinase inhibitors in mouse embryonic stem cells. *Oncogene* 1996;12:309–322.
- 20 Wobus AM, Wallukat G, Hescheler J. Pluripotent mouse embryonic stem cells are able to differentiate into cardiomyocytes expressing chronotropic responses to adrenergic and cholinergic agents and Ca²⁺ channel blockers. *Differentiation* 1991;48:173–182.
- 21 Storm MP, Bone HK, Beck CG et al. Regulation of Nanog expression by phosphoinositide 3-kinase-dependent signalling in murine embryonic stem cells. *J Biol Chem* 2007;282:6265–6273.
- 22 Paling NR, Whealon H, Bone HK et al. Regulation of embryonic stem cell self-renewal by phosphoinositide 3-kinase-dependent signalling. *J Biol Chem* 2004;279:48063–48070.
- 23 Nichols J, Zevnik B, Anastassiadis K et al. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 1998;95:379–391.
- 24 Mochizuki T, Kitanaka C, Noguchi K et al. Physical and functional interactions between Pim-1 kinase and Cdc25A phosphatase. Implications for the Pim-1-mediated activation of the c-Myc signaling pathway. *J Biol Chem* 1999;274:18659–18666.
- 25 Vallier L, Mancip J, Markossian S et al. An efficient system for conditional gene expression in embryonic stem cells and in their in vitro and in vivo differentiated derivatives. *Proc Natl Acad Sci U S A* 2001;98:2467–2472.
- 26 Shi W, Wang H, Pan G et al. Regulation of the pluripotency marker rex-1 by Nanog and Sox2. *J Biol Chem* 2006;281:23319–23325.
- 27 Qu CK, Feng GS. Shp-2 has a positive regulatory role in ES cell differentiation and proliferation. *Oncogene* 1998;17:433–439.
- 28 Anneren C, Cowan CA, Melton DA. The Src family of tyrosine kinases is important for embryonic stem cell self-renewal. *J Biol Chem* 2004;279:31590–31598.
- 29 Loh YH, Wu Q, Chew JL et al. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* 2006;38:431–440.
- 30 Galaktionov K, Chen X, Beach D, Cdc25 cell-cycle phosphatase as a target of c-myc. *Nature* 1996;382:511–517.
- 31 Bartek J, Lukas J. Mammalian G1- and S-phase checkpoints in response to DNA damage. *Curr Opin Cell Biol* 2001;13:738–747.
- 32 Aho TL, Sandholm J, Peltola KJ et al. Pim-1 kinase promotes inactivation of the pro-apoptotic Bad protein by phosphorylating it on the Ser112 gatekeeper site. *FEBS Lett* 2004;571:43–49.
- 33 Li YY, Popivanova BK, Nagai Y et al. Pim-3, a proto-oncogene with serine/threonine kinase activity, is aberrantly expressed in human pancreatic cancer and phosphorylates bad to block bad-mediated apoptosis in human pancreatic cancer cell lines. *Cancer Res* 2006;66:6741–6747.
- 34 Duval D, Trouillas M, Thibault C et al. Apoptosis and differentiation commitment: Novel insights revealed by gene profiling studies in mouse embryonic stem cells. *Cell Death Differ* 2006;13:564–575.



See www.StemCells.com for supplemental material available online.

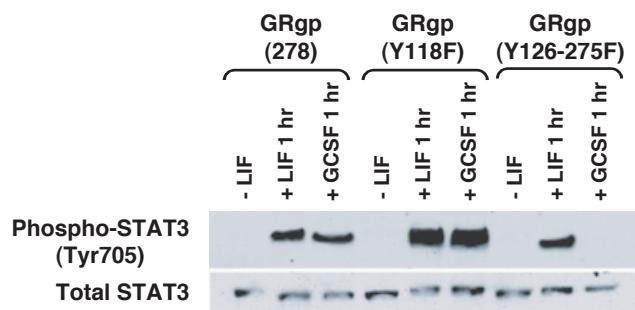
Supplementary data 1.

To rule out the possibility that LIF starvation for 24 hours would irreversibly commit ES cells to differentiation, self-renewal was quantified in a clonal assay. CGR8 mES cells were LIF-deprived for 24 hrs then replated at a density of 1000 cells per 10cm dish and further cultured in medium supplemented with LIF for five days. This resulted in a 4% decrease in the proportion of colonies retaining an morphologically undifferentiated phenotype (94.1 +/- 1% compared to 97.7 +/- 0.6% in non deprived cells; means and standard errors were measured in three replicates). Only 6% of the colonies originated from irreversibly committed cells and were fully differentiated. Thus, the vast majority of the CGR8 mES cells were still capable of self-renewal upon deprivation of LIF for 24 hrs.



Supplementary data 2

mES cells expressing wild-type, SHP2-binding deficient, and STAT3-binding deficient chimeric receptors were LIF-deprived for 24 hrs and restimulated for 1 hr, either with 10,000 U/ml of IF to activate the endogenous gp130 receptor or with 30ng/ml of G-CSF to activate the chimeric receptor. STAT3 phosphorylation was then analyzed by immunoblotting using an antibody to total STAT3 and phospho-Tyr705STAT3. In GRgp278 (wild-type) and GRgp(Y118F) (SHP2-binding deficient) transfectants, LIF and G-CSF were equally efficient at inducing STAT3 phosphorylation on Tyr705. By contrast, in GRgp(Y126-275F) (STAT3-binding deficient) transfectants, stimulation with G-CSF did not result in detectable STAT3 phosphorylation.



3- Article n°2

Novel STAT3 target genes, including, Klf4 and Klf5, exert distinct roles in the inhibition of mesoderm and endoderm differentiation in embryonic stem cells

Soumis à Plos Biology

**Novel STAT3 target genes, including Klf4 and Klf5, exert distinct roles in
the inhibition of mesoderm and endoderm differentiation in embryonic
stem cells**

*Irène Aksoy^{1,2,3}, *Pierre-Yves Bourillot^{1,2,3}, Valerie Schreiber^{1,2,3,4}, Herbert Schulz⁵, Oliver Hummel⁵, Tobias Nolden⁶, Michael X. Doss⁷, Norbert Hubner⁵, **Heinz Himmelbauer⁶, Jürgen Hescheler⁷, Agapios Sachinidis⁷, and Pierre Savatier^{1,2,3},

¹ Inserm, U846, 18 Avenue Doyen Lepine, 69500 Bron, France. ² Stem Cell and Brain Research Institute, 69500 Bron, France. ³ Université de Lyon, Université Lyon 1, 69003, Lyon, France. ⁴ PrimaStem, 18 Avenue Doyen Lepine, 69500 Bron, France. ⁵ Max Delbrück Center for Molecular Medicine, Robert-Rössle-Strasse 10, Berlin, Germany. ⁶ Max Planck Institute for Molecular Genetics, Ihnestrasse, 63-73, 14195 Berlin, Germany. ⁷ Center of Physiology and Pathophysiology, Institute of Neurophysiology, Robert-Koch-Strasse. 39, 50931 Cologne, Germany.

*Equal contributions

**Present address: Centre for Genomic Regulation (CRG), UPF, Barcelona, Spain

Keywords: embryonic stem cells, self-renewal, STAT3, Klf4, Klf5

Correspondence: Pierre Savatier, INSERM U846, 18 avenue Doyen Lépine, 69500 Bron, France. E-mail: pierre.savatier@inserm.fr; Tel: 33-472-91-34-42; Fax: 33-472-91-34-61.

Abstract

BACKGROUND: The Leukemia Inhibitory factor (LIF) activates the transcription factor STAT3, which results in the maintenance of mouse embryonic stem cells in the pluripotent state by inhibiting both mesodermal and endodermal differentiation. How the LIF/STAT3 pathway inhibits commitment to both mesoderm and endoderm lineages is presently unknown. **METHODOLOGY/PRINCIPAL FINDINGS:** Using a hormone-dependent STAT3 mutant and with microarray analysis, we identified 58 targets of STAT3 including 20 unknown genes. Functional analysis showed that 22 among the 23 STAT3 target genes analysed contribute to the maintenance of the undifferentiated state, as evidenced by an increase in the frequency of differentiated colonies in a self-renewal assay, and a concomitant elevation of early differentiation markers upon knockdown. STAT3 target genes fell into three distinct categories depending on their capacity to inhibit either mesoderm or endoderm differentiation, or both. Knockdown of *Krüppel like factor (Klf) 4* expression induced ectopic expression of the early endodermal regulators *GATA-4*, *GATA-6*, *Sox7*, and *Sox17* in undifferentiated ES cells and resulted in an increase in the proportion of cells that differentiate to both extraembryonic and definitive endoderm. By contrast, knockdown of *Klf5* expression induced ectopic expression of the panmesodermal regulator *T* and resulted in an increase in the proportion of cells that differentiate to mesoderm. **CONCLUSIONS/SIGNIFICANCE:** Our work provides for the first time evidence that the LIF/STAT3 pathway maintains the undifferentiated state of ES cells by activating the expression of two classes of target genes involved in the suppression of mesoderm and endoderm differentiation, respectively. It also highlights the central role played by *Klf4* and *Klf5* in the coordination of self-renewal and commitment to mesoderm and endoderm in mouse ES cells.

Introduction

Mouse ES cells can be easily propagated in defined medium lacking serum and feeder cells, under the sole action of recombinant Leukemia Inhibitory factor (LIF) and Bone Morphogenetic Protein (BMP) 4, which together cooperate to suppress differentiation (Chambers 2004). The LIF signaling pathway is involved in the inhibition of mesoderm and endoderm differentiation, while the BMP4 pathway is involved in the inhibition of the ectoderm pathway (Ying et al. 2003). LIF engages a heterodimeric receptor complex consisting of LIF receptor and gp130. This complex recruits and phosphorylates a number of factors including the signal transducer and activator of transcription (STAT)-3 (Boeuf et al. 1997), the phosphatidyl inositol 3 kinase (PI3K) (Paling et al. 2004), Yes, a member of the Src family of tyrosine kinases (Anneren et al. 2004), and the src-homology domain phosphatase 2 (SHP-2) (Burdon et al. 1999). Functional inactivation of STAT3, PI3K, and Yes, in the presence of LIF, induces spontaneous differentiation (Niwa et al. 1998; Anneren et al. 2004; Paling et al. 2004). In contrast, functional inactivation of SHP-2 facilitates self-renewal (Burdon et al. 1999), suggesting that LIF stimulates self-renewal by co-activating positive and negative signaling pathways in a coordinated manner. Expression of a hormone-dependent STAT3 mutant that can be activated directly by oestradiol sustains self-renewal without the addition of LIF (Matsuda et al. 1999). This suggests that the LIF/STAT3 pathway plays a pre-eminent role in the control of self-renewal and pluripotency of mouse ES cells.

Little is known about the effectors of STAT3-dependent self-renewal. The proto-oncogene *c-Myc* is a direct target of STAT3. The forced expression of a mutated form of cMyc temporarily inhibits the differentiation of mouse ES cells induced by withdrawal of LIF (Cartwright et al. 2005). We recently identified two novel STAT3 target genes, *Pim-1* and *Pim-3*, that encode members of the family of Pim serine/threonine kinases. *Pim-1* and *Pim-3* are activated by a ligand-dependent STAT3 mutant in the presence of protein synthesis inhibitor, indicating that both are direct targets of the LIF/STAT3 pathway. Their activities are necessary for ES cell self-renewal, as shown by the increased rates of spontaneous differentiation after knockdown (Aksoy et al. 2007). *Eed* was recently shown to maintain silencing of differentiation-associated genes upon direct activation by STAT3 in cooperation with the pluripotency gene *Oct-4* in undifferentiated ES cells (Ura et al. 2008). Contrasting with the four above mentioned genes, *Socs3* which encodes the Suppressor of cytokine signaling 3 (SOCS3) functions as a feedback regulator of gp130 signaling (Schmitz et al. 2000) that blocks self-renewal when overexpressed (Ying et al. 2003). *JunB* and *Grg5*, the

latter encoding the groucho-like protein Aes1, also have been shown to be regulated by LIF/STAT3 in ES cells (Matsuda et al. 1999; Sekkai et al. 2005), but evidence that they are effectors of self-renewal is lacking.

Klf4 and *Klf5* encode two closely related members of the Krüppel-like family of transcription factors. Together with *Klf2*, they have been shown to form a molecular circuitry that contributes to the maintenance of the undifferentiated state through activation of the pluripotency gene *Nanog* (Jiang et al. 2008). *Klf4* is also required for the establishment of pluripotency in somatic cells (Okita et al. 2007), delays differentiation in embryoid body culture (Li et al. 2005), and acts as a co-factor of the pluripotency-associated factors Oct4 and Sox2 to activate a particular set of target genes (Nakatake et al. 2006). *Klf5* was recently shown to contribute to the maintenance of the undifferentiate state of mouse ES cells (Parisi et al. 2008).

In order to investigate how STAT3 maintains mouse ES cells in the undifferentiated state, we characterized its transcriptome by microarray analysis. We identified 46 novel STAT3 target genes, of which 22 were found to contribute to the maintenance of the undifferentiated state. These 22 genes fall into three categories, depending on their capacity to inhibit either mesoderm or endoderm differentiation, or both. Two of them, *Klf4* and *Klf5*, were found to play crucial roles in suppressing differentiation towards endoderm and mesoderm lineages, respectively.

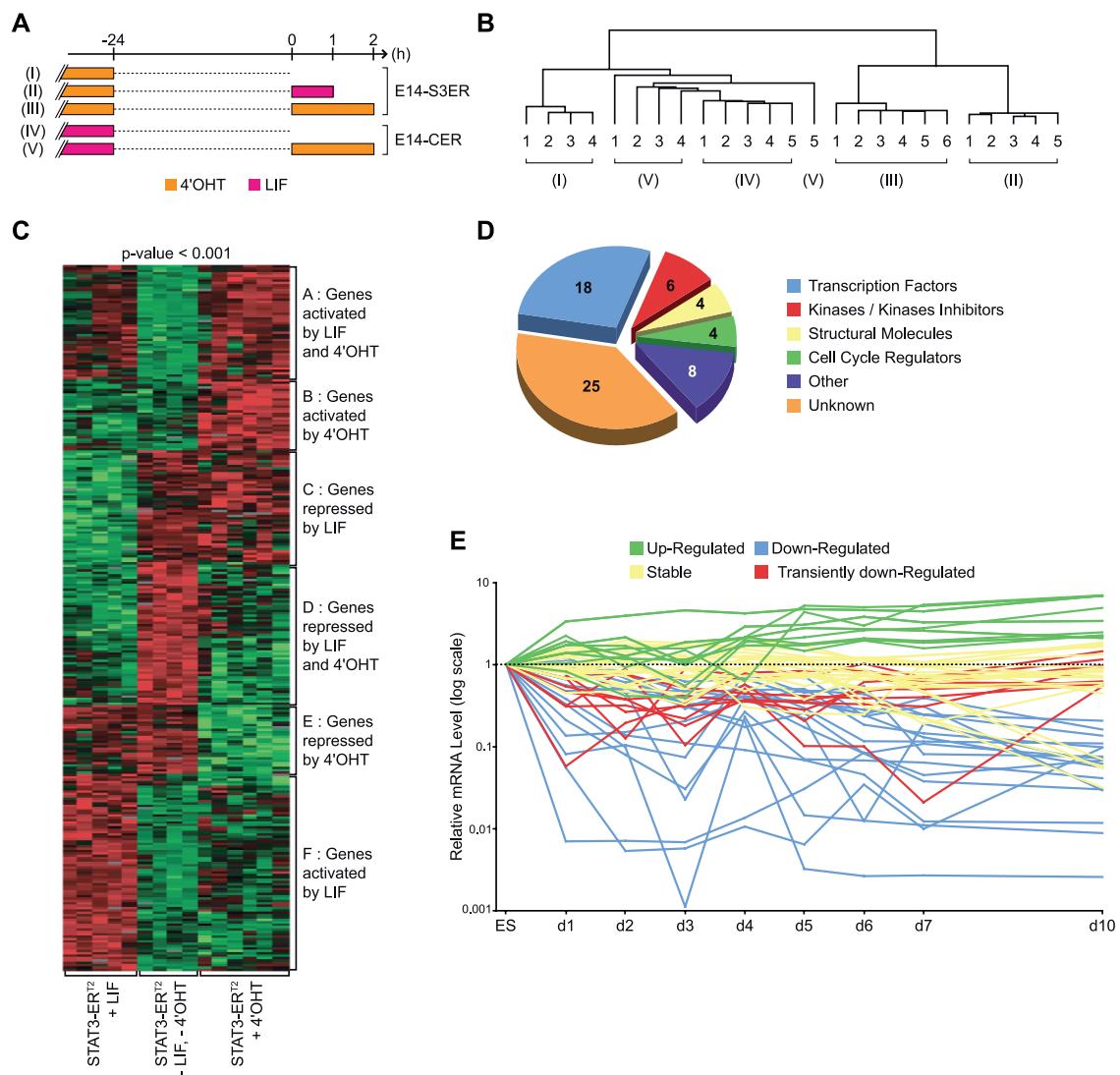


Figure 1

Aksoy, Bourillot *et al*,

Results

Identification of novel genes regulated by LIF/STAT3

E14/T-STAT3-ER^{T2} ES cells (Aksoy et al. 2007), referred to E14-S3ER, were used to identify the target genes of STAT3. E14-S3ER cells express the hormone-dependent STAT3-ER^{T2} from an episomal vector. They could self-renew with 10 nM of 4'hydroxy-tamoxifen (4'OHT) and differentiate after withdrawal of hormone (**Supplementary Figure 1**). By contrast, E14/T ES cells expressing the Cre-ERT2 recombinase (Vallier et al. 2001), referred to as E14-CER, were strictly dependent on LIF for propagation in the undifferentiated state. The gene expression profiles of E14-S3ER and E14-CER cells, before and after stimulation with LIF or 4'OHT, were compared using Mouse Genome 430 2.0 arrays. The following five conditions were examined (**Figure 1A**): (I) E14-S3ER cells deprived of LIF and 4'OHT for 24 hours (n = 4); (II) E14-S3ER cells treated as in (I) and re-stimulated with 10,000 U/ml LIF for 1 hour to reactivate endogenous STAT3 (n = 5); (III) E14-S3ER cells as in (I) and restimulated with 100 nM 4'OHT for 2 hours to activate STAT3-ER^{T2} (n = 6); (IV) E14-CER cells deprived of LIF for 24 hours (n = 5); (V) E14-CER cells as in (IV) and restimulated with 100 nM tamoxifen for 2 hours (n = 5). E14-CER cells were used to identify any 4'OHT-response genes, irrespective of STAT3 activity. Hierarchical clustering analysis of the relative expression of probe sets (ANOVA, $p < 0.05$) showed that E14-S3ER cells stimulated with LIF and 4'OHT appeared more similar to each other than to E14-S3ER and E14-CER cells after deprivation of LIF and 4'OHT (**Figure 1B**). This observation indicates that stimulation of STAT3 activity is the primary cause of the observed differences in gene expression signatures in these five ES cell populations. K-means clustering of probe sets differentially expressed in the E14-S3ER cells prior to, and after, stimulation with LIF and 4'OHT resulted in six independent clusters (**Figure 1C**). We reasoned that STAT3 target genes playing important roles in self-renewal should be up-regulated, by endogenous STAT3 - as a result of LIF stimulation - and by STAT3-ER^{T2} as a result of 4'OHT stimulation. Thus, all subsequent analyses were focused on cluster I, which contains the probe sets that were up-regulated by the two ligands. After elimination of redundant probe sets, elimination of probe sets activated by 4'OHT in the E14-CER control cells, and using filtering of 1.3-fold change in expression, 77 probe sets were selected. Sixty-five were confirmed by quantitative RT-PCR with a high degree of correlation ($r^2 = 0.95$), confirming the validity of the microarray data (**Figure 2, Supplementary Table I, Supplementary Figure 2**). Gene ontology analysis showed that the

Figure 2

Aksoy, Bourillot *et al.*,

annotated probe sets correspond to genes encoding transcription factors, kinases, kinase inhibitors, structural molecules and cell cycle regulators. Thirteen are RIKENs (**Figure 1D**, **Table I**).

The expression pattern of the 65 probe sets was analysed during differentiation induced by suspension culture (embryoid bodies) (**Figure 1E**, **Figure 2**). The 65 genes fell into four categories according to their expression pattern in undifferentiated *versus* differentiated cells: (i) 17 genes had their expression down-regulated as early as the first day of differentiation (blue); (ii) 8 genes had their expression down-regulated at the onset of differentiation and reactivated at later stages (red); (iii) 13 genes had their expression up-regulated during differentiation (green); (iv) 21 genes had their expression stable during differentiation (yellow).

To determine if the 65 probe sets were direct targets of STAT3, their expression was analysed in E14-S3ER cells following stimulation with 100 nM 4’OHT in the presence of cycloheximide. Seven STAT3 probe sets were not activated by 4’OHT, indicating that their transcriptional regulation by STAT3-ER^{T2} required *de novo* protein synthesis. In contrast, 58 probe sets were found to be resistant to cycloheximide treatment, strongly suggesting that they are direct transcriptional targets of STAT3-ER^{T2} (**Figure 2**).

Knockdown of STAT3 target genes is detrimental to self-renewal in LIF-supportive conditions

Twenty-three STAT3 target genes were selected for functional analysis. They encode transcription factors (*Klf4*, *Klf5*, *Smad7*, *Zfp36*, *Zfp36L1*, *Spry2*, *Sall4*, *Gbx2*), kinases (*pim-1*, *pim-3*, *sgk*), regulators of signaling (*Rgs16*, *Dact1*), a cyclin (*Cnnm1*), adhesion molecules (*ocln*, *Icam-1*), a stress-response gene (*Ier-3*), an extracellular matrix associated protein (*Cyr61*), a deadenylase (*Ccrn4l*), a sulfatase (*Sulf1*), and a protein of the cytoskeleton (*vim*). Two are Rikens (*5230400M03Rik* and *1700051K13Rik*) (**Figure 2**). To determine their role in the inhibition of ES cell differentiation, their expression was knocked-down by means of interfering lentiviral vectors and the engineered cells analyzed in a colony-forming assay. For all genes studied, a 35-87% reduction in transcript levels was observed in transient transfection experiments with the shRNA sequences selected (**Supplementary Table II**). Forty-eight hours after infection with interfering lentiviral vectors, CGR8 ES cells were dissociated, replated at clonal density, and further grown for 6 days in LIF-supportive conditions, in the presence of G418 to kill the non-infected cells. The resulting colonies were

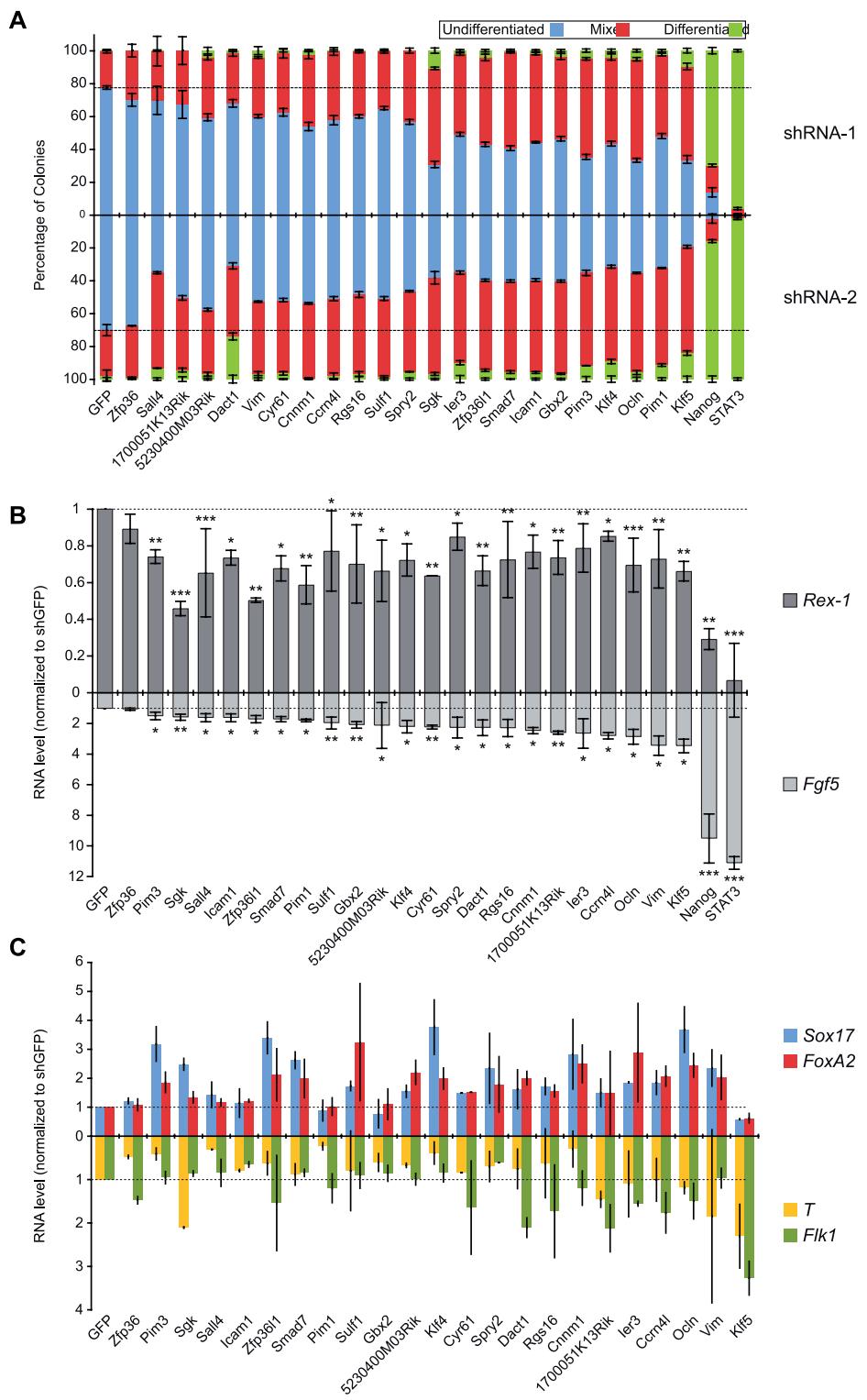


Figure 3

Aksoy, Bourillot *et al*,

stained for alkaline phosphatase (AP) activity (Paling et al. 2004), and the percentages of undifferentiated, mixed, and differentiated colonies calculated (**Figure 3A**, upper panel). Expression of control shRNA-1 (recognizing GFP) resulted in 77.5% of G418-resistant cells forming undifferentiated AP⁺ stem cell colonies, whereas knockdown of Nanog and STAT3 (positive controls) resulted in only 13.7% and 0.3% respectively. Knockdown of 22 out of the 23 target genes analysed resulted in a significant increase in the percentage of mixed/differentiated colonies at the expense of undifferentiated colonies, relative to the GFP control. The increase in the frequency of mixed/differentiated colonies ranged from 30% (*Sall4*) to 68% (*Sgk*). Comparable results were obtained for each gene with an independent shRNA sequence (shRNA-2) (**Figure 3A**, bottom panel). Interestingly, there was a significant correlation of the percentages of undifferentiated colonies following knockdown with shRNA-1 and shRNA-2 ($r^2 = 0.57$, $p < 0.01$). This correlation strongly suggests that the differences observed between the 22 genes in the frequency of undifferentiated colonies primarily resulted from the relative contribution of each gene to the suppression of differentiation, rather than from gene-to-gene variations in the level of interference. The effect of *Dact1*, *Sall4*, *E030031B1Rik*, and *1700051K13Rik* knockdown on the frequency of undifferentiated, mixed and differentiated colonies was confirmed using scrambled shRNA sequences (**Supplementary Figure 3**).

Undifferentiated, mixed and differentiated colonies obtained in colony-forming assays were pooled and subsequently analyzed by real-time PCR to quantify the expression of pluripotency and differentiation markers. Expression of the pluripotency marker *Rex-1* (*Zfp42*) was reduced relative to control after knockdown of all genes examined (1.1- to 2.2-fold) (**Figure 3B**, upper panel). Expression of the early ectodermal marker *Fgf5* was induced after knockdown of all genes examined (1.5 to 3.5-fold, **Figure 3B**, bottom panel). Expression of the endodermal markers *Sox17* and *FoxA2* was induced after knockdown of most genes except *Icam1*, *Pim-1*, *Gbx2*, *Zfp36*, and *Klf5* (**Figure 3C**, upper panel). Expression of the mesodermal markers *T* (Brachyury) and *Flk1* was induced after knockdown of *Klf5* and *1700051K13Rik* (**Figure 3C**, bottom panel). Together, these results show that the knockdown of 22 out of the 23 genes analysed induced differentiation in LIF-supportive conditions. We conclude that all 22 genes contribute to the maintenance of the undifferentiated state.

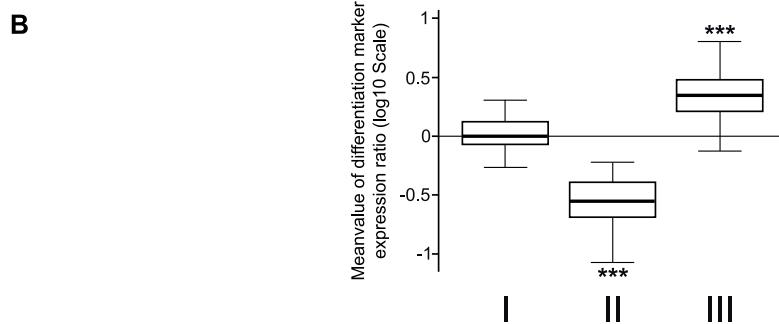
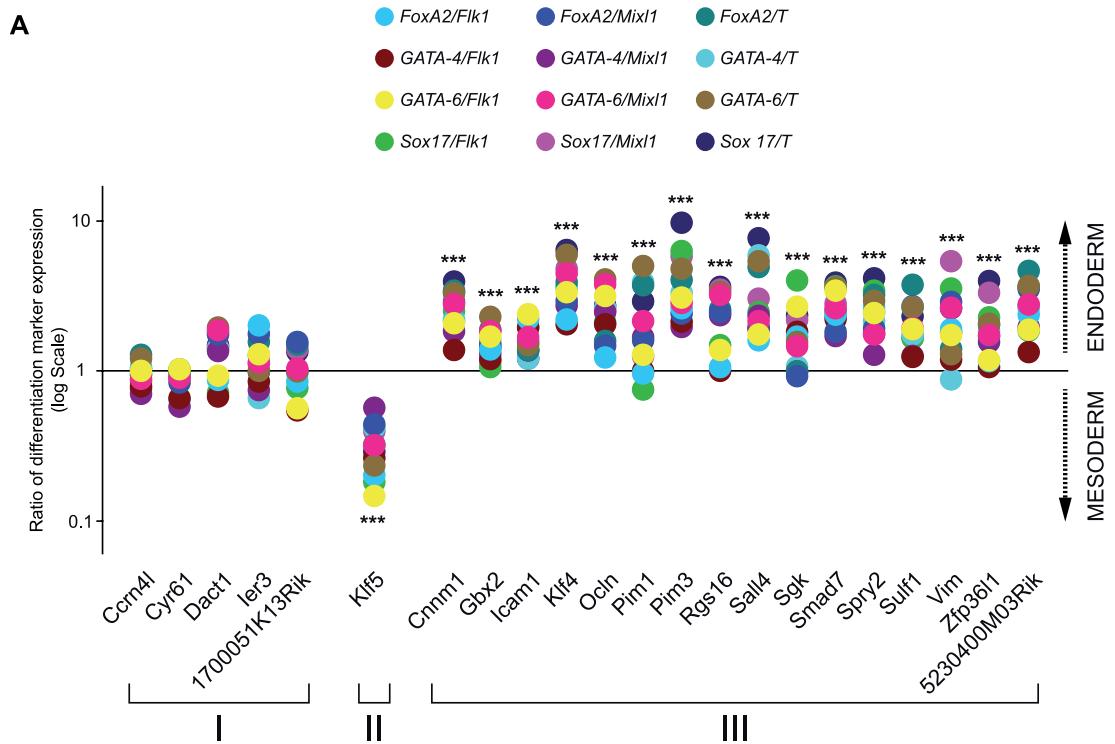


Figure 4

Aksoy, Bourillot *et al*,

Knockdown of STAT3 target genes leads to three distinct patterns of differentiation

Analysis of *Sox17*, *Foxa2*, *T*, and *Flk1* expression in ES cells infected with interfering vectors suggests that STAT3 target genes are not equally involved in the inhibition of mesoderm and endoderm differentiation. To address this issue further, expression of the mesoderm- (*Flk1*, *Mixl1*, *T*) and endoderm- (*FoxA2*, *GATA-4*, *GATA-6*, *Sox17*) associated genes was analysed by quantitative RT-PCR in pools of *neo^R* colonies resulting from the infection of ES cells with the interfering vectors (**Supplementary Table III**). To reveal a bias towards mesoderm or endoderm differentiation, endoderm/mesoderm ratios (R) were calculated for all 12 combinations of endoderm and mesoderm markers. In **Figure 4A**, each of the 22 genes is represented by a set of 12 R scores which together reflect the bias towards endoderm, or mesoderm, differentiation induced by gene knockdown. Of the 22 genes analysed, 16 showed a bias towards endoderm differentiation (*Cnnm1*, *Gbx2*, *Icam1*, *Klf4*, *Ocln*, *Pim1*, *Pim3*, *Rgs16*, *Sall4*, *Sgk*, *Smad7*, *Spry2*, *Sulf1*, *Vim*, *Zfp36L1*, and *5230400M03Rik*), one showed a bias towards mesoderm differentiation (*Klf5*), and five did not show any significant bias towards either lineages (*Ccrn4l*, *Cyr61*, *Dact1*, *Ier3*, *17700051K13Rik*). These results strongly suggest that some STAT3 target genes are involved in suppression of both endoderm and mesoderm differentiation, whereas others show lineage specificity.

Knockdown of Klf4 and Klf5 induces differentiation into endoderm and mesoderm, respectively

Knockdown of *Klf4* and *Klf5* showed contrasting effects on mesoderm and endoderm marker expression in the colony-forming assay (see **Figure 4**). To explore the role of *Klf4* and *Klf5* in the inhibition of mesoderm and endoderm differentiation, markers of both lineages were examined during differentiation induced by the formation of embryoid bodies and LIF withdrawal. ES cells were infected with lentiviral vectors expressing *shKlf4-1*, *shKlf4-2*, *shKlf5-1*, *shKlf5-2*, and a scrambled control, and subsequently cultured for 7 days at high density in medium supplemented with G418 to kill non-infected cells. These culture conditions were found to temporarily prevent spontaneous differentiation induced by either *Klf4* or *Klf5* knockdown (data not shown). Following *Klf4* knockdown with both shRNAs, expression of the endoderm-associated genes, *GATA-4*, *GATA-6*, *Sox17*, *Dab2* (Disabled-2), *FoxA2*, *Fxyd3* (FXYD protein 3), *Cxcr4*, and *Emp2* (Epithelial membrane protein 2) increased at all time points analysed (**Figure 5**). *Sox7* expression increased during the three first days of

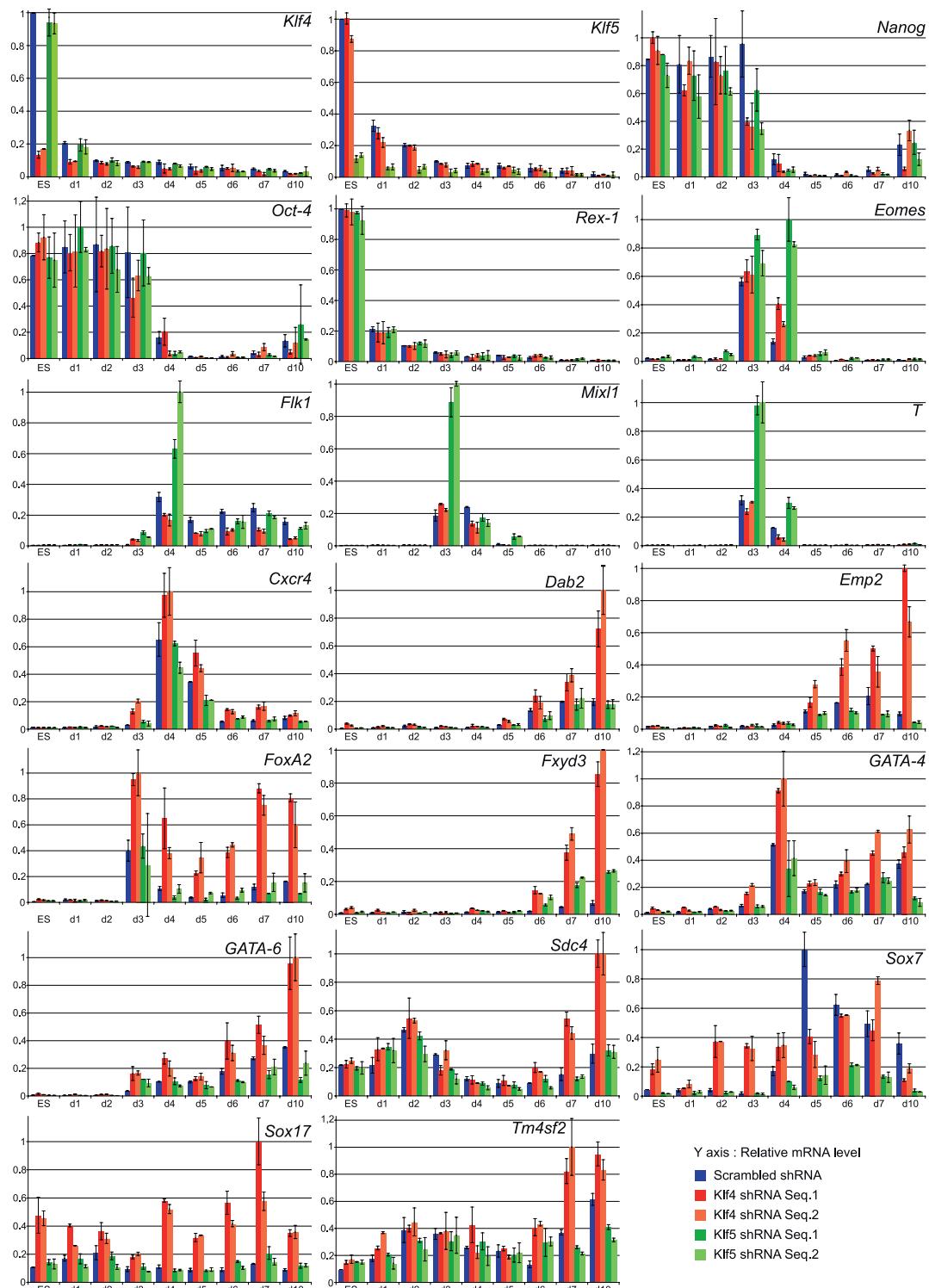


Figure 5

Aksoy, Bourillot et al,

differentiation whereas expression of *Sdc4* (Syndecan-4) and *Tm4sf2* (Transmembrane 4 superfamily 2) increased during the two last days. Thus, all endoderm-associated genes examined were up regulated upon *Klf4* knockdown. In contrast, expression of mesoderm-associated genes, *T*, *Mixl1*, *Flk1*, and *Eomes*, did not change relative to control. In contrast to *Klf4* knockdown, *Klf5* knockdown did not alter endoderm gene expression. However, *Klf5* knockdown did increase the expression of mesoderm genes, *T*, *Mixl1*, *Flk1*, and *Eomes*. After day 5, expression of *Flk1* was no longer increased in *Klf5* knockdown but was decreased in *Klf4* knockdown. This shift could reflect a depletion of mesoderm at the expense of endoderm following knockdown of *Klf5*. Together, these results demonstrate that, upon inactivation of LIF signaling and culture in suspension, knockdown of *Klf4* facilitates ES cell differentiation into cells expressing the markers of extra-embryonic and definitive endoderm, whereas knockdown of *Klf5* facilitates differentiation into cells expressing markers of mesoderm.

Knockdown of Klf4 and Klf5 alters the ratio of T⁺ versus Sox17⁺ cells in embryoid bodies.

Increase in endoderm and mesoderm markers induced by *Klf4* and *Klf5* knockdown could reflect an alteration in gene expression in the sub-population of cells committed to a mesoderm or an endoderm fate or, alternatively, an alteration in the percentage of cells which commit to either lineages. To distinguish between these two possibilities, two reporter ES cell lines (T-GFP-ES and Sox17-DsRed-ES) were engineered from CGR8 ES cells. T-GFP-ES cells express EGFP under the control of a 4417 bp promoter region upstream of the translation initiation site of the mesoderm-associated *T brachyury* gene (Doss *et al.* manuscript in preparation). Sox17-DsRed-ES cells express the DsRed fluorescent protein driven by the regulatory elements of the endoderm-associated *Sox17* gene, making use of an engineered 15.1 kbp BAC fragment which was integrated into the endogenous *Sox17* locus by homologous recombination (Nolden *et al.* manuscript in preparation). Both reporter cell lines were infected with interfering lentiviral vectors expressing *shKlf4-1*, *shKlf4-2*, *shKlf5-1*, *shKlf5-2*, and two scrambled controls, and subsequently cultured for 7 days at high density in medium supplemented with hygromycin to kill non-infected cells. Engineered ES cells were subsequently induced to differentiate in suspension in the absence of LIF. Knockdown of *Klf4* in the Sox17-DsRed cells resulted in an average 3.4-fold increase in DsRed⁺ cell number, (rang: 2.3-fold at day 5 to 5.5-fold at day 3) (**Figure 6A,B**). In contrast, knockdown of *Klf4* had only a minor effect on the number of GFP⁺ cells in T-GFP cells. In a mirror image, knockdown of *Klf5* in T-GFP cells resulted in an average 4.1-fold increase of GFP⁺ cell number (rang: 1.2-fold at day 5 to 10.1-fold at day 1). *Klf5* knockdown showed no

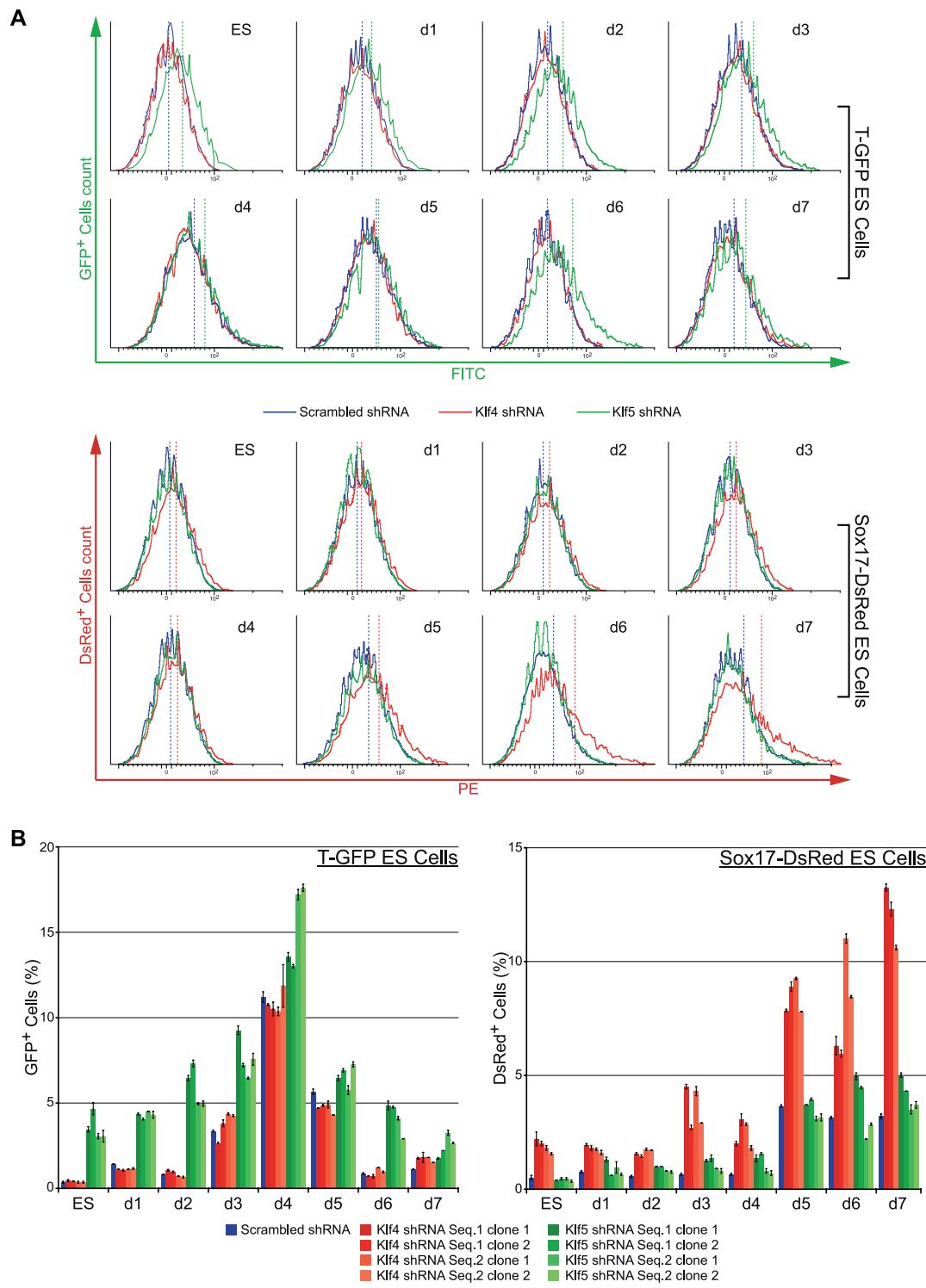


Figure 6

Aksoy, Bourillot *et al*,

reproducible effect on the number of DsRed⁺ cells in Sox17-DsRed cells. Together, these observations indicate that *Klf4* and *Klf5* knockdown increases the number of ES cells that commit to endoderm or mesoderm, respectively. In both cases, the increase in the number of fluorescent cells reached its maximum at early stages of differentiation (day 1 to day 3), thus two days before *T* activation and four days before *Sox17* activation in control cells. These observations suggest that *Klf4* and *Klf5* knockdown induces premature commitment to endoderm and mesoderm differentiation, a finding further supported by our previous observation that other differentiation markers exhibit premature expression. Collectively, these results indicate that *Klf4* and *Klf5* play a central role in the suppression of endoderm and mesoderm differentiation, respectively.

Knockdown of Klf4 and Klf5 activates the expression of mesoderm and endoderm differentiation-associated genes in pluripotent cells

In an effort to unravel the mechanisms by which *Klf4* and *Klf5* control endoderm and mesoderm commitment, we analysed the changes in gene expression profiles induced by either *Klf4* or *Klf5* knockdown in the undifferentiated ES cell. CGR8 ES cells were infected with interfering lentiviral vectors expressing *shKlf4-1*, *shKlf4-2*, *shKlf5-1*, *shKlf5-2*, and *shGFP*, and subsequently cultured for 10 days at high density in medium supplemented with G418 to kill the non-infected cells. Under such culture conditions, the resulting G418-resistant cell population showed a two-fold reduction in *Klf4* and *Klf5* transcript levels (**Figure 7A**), but exhibited no sign of differentiation, as evidenced by the morphology of the cells (**Figure 7B**), the expression of pluripotency markers, *Nanog*, *Oct-4* (*pou5f1*), and *Rex-1* (**Figure 7A**), the expression of the early ectoderm marker, *Fgf5* (**Figure 7A**), and the percentage of SSEA-1⁺ cells (**Figure 7C**), all of these markers were indistinguishable from control cells. Therefore, when cultured at high cell density, ES cells harboring a two-fold reduction in *Klf4* and *Klf5* levels are able to self-renew in the undifferentiated state within the time-frame of the experiment.

Transcriptomes of ES cells expressing *Klf4* and *Klf5* shRNAs were compared to that of control cells, using Mouse Genome 430 2.0 arrays. K-means clustering of probe sets differentially expressed between the five ES cell populations resulted in ten clusters corresponding to genes up-or down-regulated in one or several conditions (**Figure 7D**), a finding consistent with previous reports showing that *Klf4* and *Klf5* can function as both a transcriptional activator and inhibitor (Rowland and Peeper 2006; McConnell et al. 2007). After elimination of unidentified probe sets, and using filtering of 1.5-fold change in

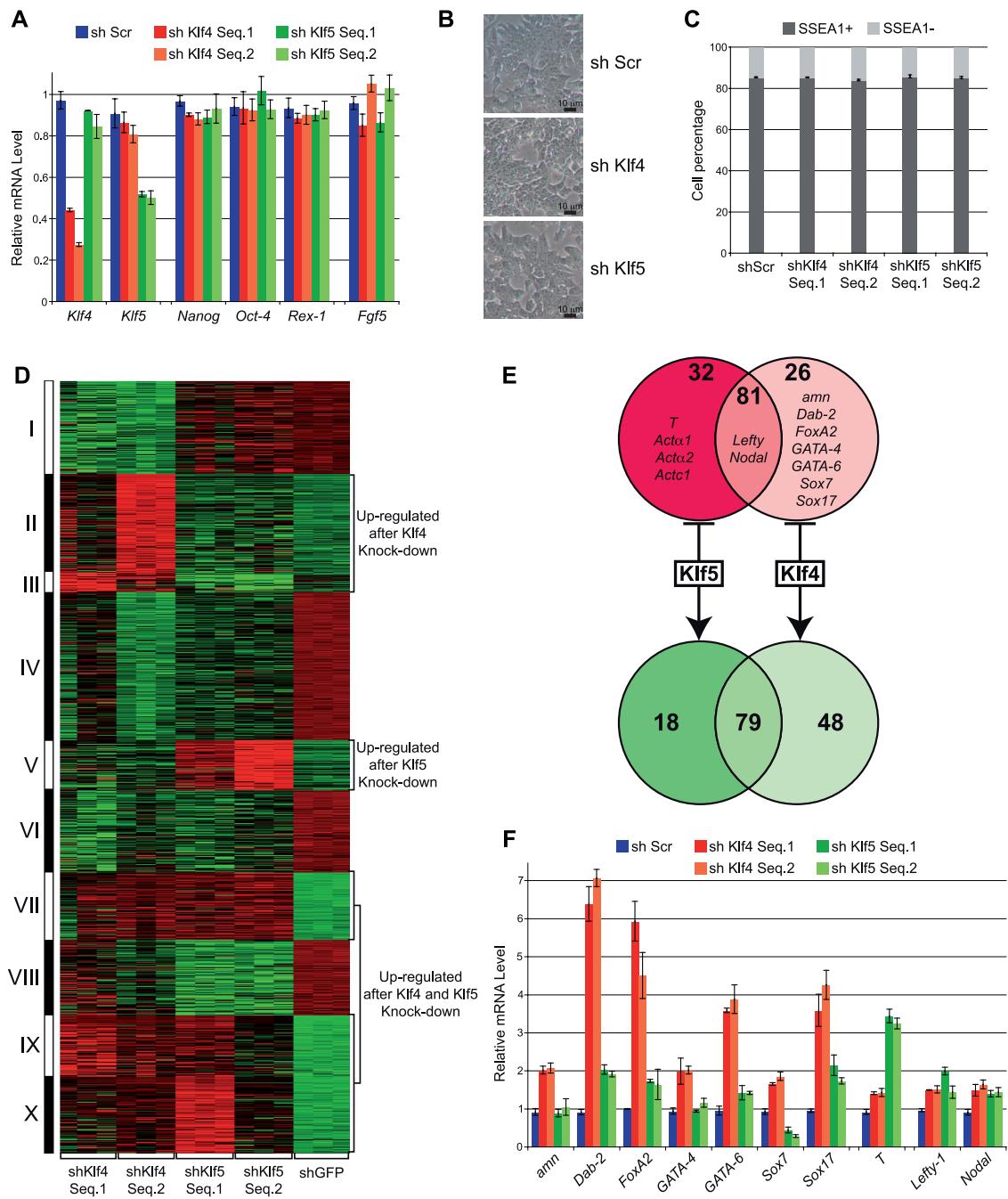


Figure 7

Aksoy, Bourillot *et al*,

expression ($p < 0.05$), 145 genes were found down-regulated (**Figure 7E, Supplementary Table IV**), of which 18 were down-regulated after *Klf4* knockdown, 48 after *Klf5* knockdown, and 79 after knockdown of either *Klf4* or *Klf5*. Using the same criteria of inclusion, 139 genes were found up-regulated after *Klf4/Klf5* knockdown. Eighty-one genes were up regulated after knockdown of both *Klf4* and *Klf5*. They include *Nodal* (1.6-fold) and *Lefty-1* (1.9-fold) (**Figure 7F**), which are expressed in ES cells and further up regulated upon differentiation induced by LIF withdrawal and formation of embryoid bodies (Dvash et al. 2004; Sekkai et al. 2005). Twenty-six genes were found up regulated in ES cells expressing both *Klf4* shRNAs, but not *Klf5* shRNAs (**Figure 7E**). They include the endoderm differentiation-associated genes *Sox7* (1.9-fold), *Sox17* (4.2-fold), *FoxA2* (5.2-fold), *GATA-4* (2.1-fold), *GATA-6* (4-fold), *Dab-2* (7.5-fold), and *ann* (2.1-fold) (**Figure 7F**). They encode transcription factors involved in the early commitment of embryonal carcinoma cells, embryonic stem cells, and primitive ectoderm of the early post-implantation embryo, to primitive and/or definitive endoderm (Kalantry et al. 2001; Fujikura et al. 2002; Kanai-Azuma et al. 2002; Yang et al. 2002; Futaki et al. 2004; Niimi et al. 2004; Yasunaga et al. 2005; Matsuura et al. 2006; Shimosato et al. 2007; Qu et al. 2008; Seguin et al. 2008). *Cubn* (cubulin), *nostrin* (nitric oxide synthase trafficker), *Klb* (klotho beta), *Cyp4f14* (polypeptide 4f14 of cytochrome P450), *neurog3* (Neurogenin-3), and *pthr1* (parathyroid hormone receptor 1), the expression of which is associated with hepatic, pancreatic, and thyroid endoderm differentiation, were also activated in ES cells expressing both *Klf4* shRNAs (**Supplementary Table IV**). Of note, the endoderm markers *Fxyd3*, *Cxcr4*, *Emp2*, *Sdc4*, *Tm4sf2*, and *Tmprss2* (transmembrane serine protease) were not activated. Importantly, no gene was found whose expression is implicated in the early stages of mesoderm differentiation was found. In a mirror image, 32 genes were up regulated in ES cells expressing both *Klf5* shRNAs, but not *Klf4* shRNAs (**Figure 7E**). They include *T*, involved in the early steps of mesoderm formation (Wilkinson et al. 1990) and activated 3.6-fold (**Figure 7F**), and genes whose expression is associated with smooth, skeletal, and cardiac muscle differentiation [*acta1* (actin α 1), *acta2* (actin α 2), and *actc1* (cardiac actin α)] . Other early mesoderm markers such as *Flk1*, *Mist1* and *Eomes*, were not activated in ES cells expressing *Klf5* shRNAs. No gene whose expression is associated with the early stages of endoderm differentiation was found amongst *Klf5*-response genes. Together, these results demonstrate that a two-fold attenuation of *Klf4* and *Klf5* expression leads, in undifferentiated ES cells, to a dramatic activation of genes involved in endoderm and mesoderm differentiation, respectively.

Discussion

STAT3 is an important regulator of mouse ES cell self-renewal (Niwa et al. 1998; Matsuda et al. 1999) which inhibits differentiation into both mesoderm and endoderm lineages (Ying et al. 2003). Its mechanism of action remains to be elucidated. To this aim, we identified 58 genes (35 known genes, 13 Rikens and 10 unidentified probe sets) whose expression is activated by a conditional ligand-dependent STAT3 mutant (in addition to activation by LIF) in the presence of a protein synthesis inhibitor. Of the 35 novel STAT3 target genes identified, 23 display STAT3 binding sites in their regulatory sequences (Chen et al. 2008) confirming the validity of our strategy to identify primary targets of the LIF/STAT3 pathway. Of note, 258 other genes are regulated by LIF – but not by ligand-dependent STAT3 – *i.e.* by a STAT3-independent mechanism. PI3K (Paling et al. 2004), cYes (Anneren et al. 2004) and SHP-2 (Burdon et al. 1999) are candidate-signaling molecules to functionally link the activated LIF receptor to these 258 genes.

We identified 22 STAT3 response genes whose knockdown of expression induces spontaneous differentiation measured in a colony-forming assay. While the effect of knockdown of individual genes on ES cell differentiation is sometimes rather modest, which may indicate a marginal contribution to the inhibition of differentiation, the sum of the individual effects results in the maintenance of the undifferentiated state. The 22 identified genes encode transcription factors, various types of proteins involved in protein and nucleic acid modification (serine/threonine kinases, sulfatase, deadenylase), adhesion molecules, extracellular matrix and cytoskeleton associated proteins, and a stress-response factor. Such a variety of molecules indicates that LIF/STAT3-dependent self-renewal results from the concomitant regulation of a wide-range of cellular activities, including the control of cell-cell and cell-matrix interactions. Eight genes appear to have a higher contribution to the inhibition of differentiation, evidenced by a stronger increase in the number of mixed/differentiated colonies and by a higher induction of differentiation marker expression following knockdown. They include *Pim-1* and *Pim-3* whose regulation by STAT3 and their role in self-renewal has been reported elsewhere (Aksoy et al. 2007), *Klf4*, *Klf5*, *Sgk*, *Smad7*, *Ocln* and *Icam-1*. Of note, these eight genes are highly expressed in undifferentiated ES cells and decline between day 1 and 4 of differentiation, suggesting that all eight contribute to ES cell identity. *Sgk* encodes the protein kinase B (PKB)-related protein serum and glucocorticoid-inducible kinase. SGK is a downstream effector of the PI3K/3-phosphoinositide-dependent protein kinases (PDK) signaling pathway (Kobayashi and Cohen 1999; Park et al. 1999),

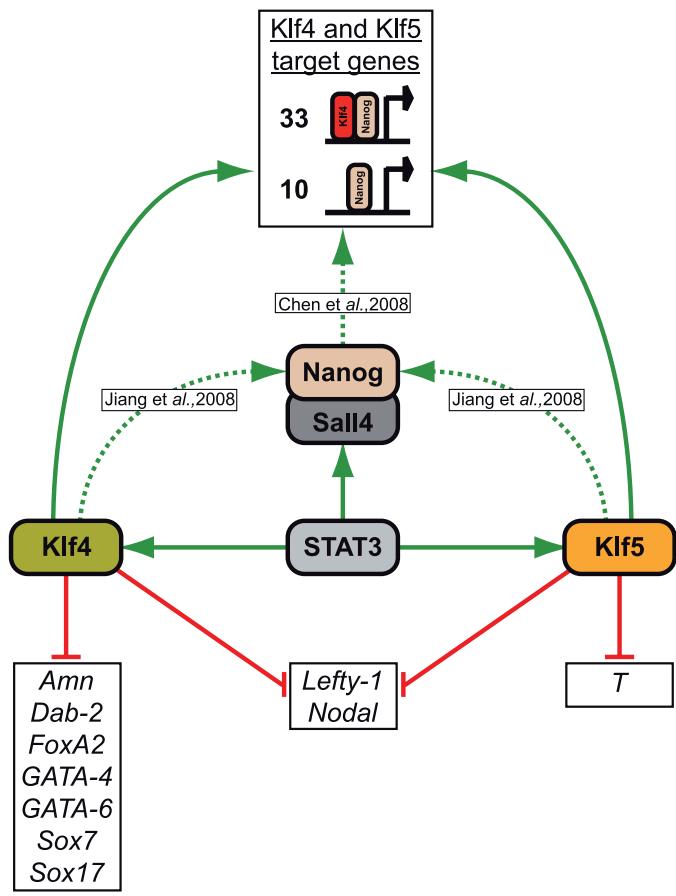


Figure 8

Aksoy, Bourillot *et al*,

which contributes to the maintenance of the undifferentiated state (Paling et al. 2004). The LIF/STAT3 and LIF/PI3K pathways may act synergistically to activate SGK and promote self-renewal. SGK - like PKB - inactivates glycogen synthase kinase (GSK)-3 *in vitro* (Kobayashi and Cohen 1999). Inhibition of GSK-3 activity is known to facilitate self-renewal of mouse (Ying et al. 2008) and human ES cells (Sato et al. 2004), providing a possible explanatory mechanism for SGK action.

Klf4 and *Klf5* encode two close members of the Krüppel-like family of transcription factors (Ghaleb et al. 2005). Knockdown of *Klf4* and *Klf5* induced by siRNA transfection does not result in observable differentiation (Jiang et al. 2008), a finding explained by the functional redundancy between *Klf2*, *Klf4* and *Klf5*. Using a sensitive method based on shRNA-mediated stable interference associated with a colony-forming assay, we were able to demonstrate that single knockdown of either *Klf4* or *Klf5* induces spontaneous differentiation of mouse ES cells in LIF-supportive conditions. Moreover, gene expression profiling revealed that Klf4 and Klf5 each regulates the expression of a unique set of genes in the undifferentiated ES cells. Together, these observations indicate that KLF4 and KLF5 support self-renewal by non-overlapping mechanisms and that both transcription factors are required for maintenance of the undifferentiated state of ES cells.

Recently *Klf4* and *Klf5*, together with *Klf2*, were shown to be part of a *Klf*-based molecular circuitry that regulates ES cell self-renewal through activation of *Nanog* expression (Jiang et al. 2008). Therefore, our finding that both *Klf4* and *Klf5* are regulated by STAT3 functionally links the LIF/STAT3 pathway to the regulation of *Nanog* expression and, indirectly, to the activity of the *Nanog/Oct-4/sox2* core regulatory network (Boiani and Scholer 2005). Interestingly, *Sall4* is also regulated by STAT3 (this work) and contributes to differentiation inhibition in agreement with a previous report (Zhang et al. 2006). SALL4 transcription factor forms complexes with Nanog and enhances its transcriptional activity (Wu et al. 2006). We therefore propose that up-regulation of Nanog transcriptional activity by KLF4 and KLF5 is reinforced by the activation of *Sall4*. In addition, of the 43 genes regulated by both KLF4 and KLF5 (see **Supplementary Table IV**), 33 display Nanog binding sites in their promoter (Chen et al. 2008), suggesting that KLF4, KLF5 and Nanog may act cooperatively, or redundantly, to activate the expression of common target genes (**Figure 8**).

The LIF/STAT3 signaling pathway is known to suppress differentiation to mesoderm and endoderm lineages (Niwa et al. 1998; Ying et al. 2003). Our lineage marker analysis demonstrates that STAT3 target genes are not equally involved in the inhibition of mesoderm and endoderm differentiation. Most of them show lineage specificity, suggesting that they

inhibit differentiation by preventing the activation of lineage-specific differentiation programs. Strikingly, *Klf4* and *Klf5* were found to exert opposite roles both in the inhibition of endoderm and mesoderm differentiation as well as in the repression of endoderm- and mesoderm-associated gene expression. We propose that *Klf4* inhibits ES cell differentiation into endoderm by repressing the expression of endoderm-specific regulators, including *GATA-4*, *GATA-6*, *Sox7*, and *Sox17*. Overexpressing *GATA-4* and *GATA-6* induces differentiation of ES cells to primitive endoderm (Fujikura et al. 2002; Shimosato et al. 2007). Forced expression of *Sox7* induces partial differentiation of F9 Embryonal Carcinoma (EC) cells to parietal endoderm (Futaki et al. 2004), whereas constitutive expression of *Sox7* and *Sox17* in human ES cells produces extraembryonic endoderm and definitive endoderm progenitors, respectively (Seguin et al. 2008). *Klf4* is likely to inhibit mouse ES cell differentiation into definitive endoderm, since the definitive endoderm-specific marker *Cxcr4* (McGrath et al. 1999; Yasunaga et al. 2005) is activated following *Klf4* knockdown. Similarly, we propose that *Klf5* inhibits ES cell differentiation into mesoderm, although the mechanism involved is not clear. The panmesodermal marker Brachyury (*T*) gene that regulates mesoderm formation in the early post-implantation embryo (Wilkinson et al. 1990) could be an effector of *Klf5*. Beside their distinct roles in differentiation inhibition, *Klf4* and *Klf5* also have redundant functions in the maintenance of the undifferentiated state of ES cells as shown, first by the capacity of each one of them, when overexpressed, to rescue pluripotency after *Klf2/Klf4/Klf5* triple knockdown-induced differentiation (Jiang et al. 2008), second by their common targets (this work). The negative regulation of the organizer gene *Nodal* by both *Klf4* and *Klf5* suggests that they may also share a common role in the initial steps of endomesoderm differentiation.

We provide evidence for the first time that two major regulators of ES cell pluripotency have distinct and contrasting roles in endoderm *versus* mesoderm differentiation inhibition. Our results suggest that the mechanisms regulating pluripotency and lineage choice are interdependant. Understanding these interdependancies will be instrumental to unraveling the mechanisms of lineage commitment in the pluripotent stem cell.

Materials and Methods

Plasmid construction

The hormone dependent STAT3-ER^{T2} transcription factor was generated by fusing the coding sequence of the mouse STAT3 transcription factor to a 5'-XhoI/EcoRI-3' fragment containing the entire ER^{T2} domain (Vallier et al. 2001). The resulting cDNA was subcloned into the EcoRI site in the episomal vector *pPCAGIZ* (Niwa et al. 1998) so to generate *pPCAGIZ-STAT3-ER^{T2}*. *Cre-ER^{T2}* was subcloned into the EcoRI site in *pPCAGIZ* to generate *pPCAGIZCre-ER^{T2}*. The BLOCK-iT lentiviral RNAi Gateway® vector kit (Invitrogen, Ref. K4943-00) was used to generate lentiviral vectors expressing small hairpin (sh)RNAs according to the manufacturer's instructions. In the *pLenti6/BLOCK-iT* backbone vector, the blunted KpnI (blunt)/XhoI fragment was replaced by a 1.3 Kb ClaI (blunt)/XhoI fragment containing the *PGK* promoter and the *neo^r* selectable gene in order to generate *pLenti6/BLOCK-iT-PGKneo^R*. Similarly, the blunted KpnI/XhoI fragment was replaced by a 1.7 Kbp EcoRV/EcoRV fragment, containing the *PGK* promoter and the *hygro^r* selectable gene, to generate *pLenti6/BLOCK-iT-PGKhygro^R*.

shRNA design

shRNA and siRNA sequences were designed using the siDESIGN®Center application of Dharmacon (<http://www.dharmacon.com>). For unassigned ESTs, the sequences of the Affymetrix probe sets were used to design the sequence of the shRNA. For each targeted gene, five independent shRNAs were cloned into pENTRY (Invitrogen, Ref. K4943-00), and the resulting pENTRY-shRNA vectors transfected into CGR8 ES cells so as to measure the interference by real-time PCR. For each targeted gene, the two shRNA sequences which show the highest interference in the transient transfection assay were cloned into *pLenti6/BLOCK-iT-PGKneo^r* or *pLenti6/BLOCK-iT-PGKhygro^r*. The sequences of selected shRNAs and the interference resulting from transfection of pENTRY-shRNA vectors in CGR8 cells are given in **Supplementary Table I**.

ES cell culture, electroporation and infection with lentiviral vectors

All ES cell lines were routinely cultured in Glasgow's Modified Eagle's Medium (GMEM) supplemented with 10% fetal calf serum (PerbioScience CRC0406) and 1000 U/ml of LIF. For plasmid transfection, 15 µg of supercoiled plasmid were mixed with 45 µl of

lipofectamine (Invitrogen Ref. 11668) in 3 ml of OPTIMEM (Invitrogen Ref. 31985). The resulting mixture was added to CGR8 ES cells at a density of 2×10^6 cells per 10 cm dish in 15 ml of medium. Cells were cultured for 48 hours prior to analysis. For episomal supertransfection, E14/T cells (Aubert et al. 2002) were electroporated with 20 μg of supercoiled vectors at 200 V and 960 μF in a 0.4 cm cuvette. Cells were plated at $5 \cdot 10^4$ cells per 10 cm dish, cultured in the presence of 100 $\mu\text{g}/\text{ml}$ of hygromycin B (Roche) for 7 days. Resistant colonies were pooled and further propagated in selection medium for one week. E14/T ES cells expressing STAT3-ER^{T2} were shifted to culture medium supplemented with 10 nM 4'-hydroxy-tamoxifen (4'OHT) lacking LIF (Calbiochem, Ref. 579002). Lentiviral vectors were produced using the BLOCK-iT lentiviral RNAi expression system (Invitrogen, Ref. K4944-00) according to the manufacturer's instructions. For infection, CGR8 were plated at a density of 10^4 cells in 24-well plates in 1 ml of medium composed of 100 μl of ES cell medium and 900 μl of culture supernatant from virus-producer cells. After 48 hours, ES cells were trypsinized, replated at 10^4 cells per gelatin-coated 10 mm tissue culture dish and further cultured in complete ES cell medium supplemented with 250 $\mu\text{g}/\text{ml}$ of G418 and 1000 U/ml LIF for 6 days. For colony-forming assay, dishes were fixed in methanol for 15 min, then stained for 15 min with a solution containing 1 mg/ml Fast Red TR saltTM (Sigma) dissolved in 0.1 M Tris, pH 9.2, containing 200 mg/ml naphtol AS-MX phosphate.

Semi-quantitative and real-time polymerase chain reaction (PCR)

RNA was extracted using RNAeasy kits with on-column DNase digestion and reverse transcription carried out with MuMLV-RT (Promega, Ref. M1701), according to the manufacturer's recommendations. Oligonucleotide sequences are given in **Supplementary Table V**. Quantitative PCR was performed using the LightCyclerTM 1.5 system and the LightCycler Fast Start DNA Master SYBR Green I kit (Roche Applied Science) according to the manufacturer's instructions. All normalizations were carried out with β -actin.

Differentiation via embryoid bodies and flow cytometry

To induce differentiation, cells were allowed to aggregate in hanging drops in ES cell medium without LIF (100 cells/drop). After 2 days, embryoid bodies were collected and further grown in suspension for 1 to 10 days in non-adherent Petri dishes. For analysis by flow cytometry, embryoid bodies were dissociated with 0.25% trypsin-EDTA. Cell were analysed using FACS Canto II (Beckton-Dickinson). Data were recorded and analysed using DiVa software.

Microarrays and bioinformatics

Total RNAs from ES cells was prepared with the Qiagen column kit (Qiagen) and treated with DNase (5 U/100 mg RNA, Sigma). Biotinylated cRNA was prepared according to the standard Affymetrix protocol (Expression Analysis Technical Manual, 1999; Affymetrix). In brief, double-stranded cDNA was synthesized from 10 µg total RNA using the SuperScript Choice System from InVitrogen and the Affymetrix T7- (dT)₂₄ primer that contains a T7 RNA polymerase promoter attached to a poly-dT sequence. The cDNA was transcribed into biotin-labeled cRNA using the Retic Lysate IVT™ kit (Ambion Inc., Woodward Austin, TX, USA). The cRNAs produced were purified using RNeasy kit (Qiagen) and fragmented to an average size of 30-50 bases according to Affymetrix recommendations. 15 µg of fragmented cRNA were hybridized for 16 hr at 45°C on the Mouse Genome 430 2.0 Array. Arrays were washed and stained in the Affymetrix Fluidics Station 450 and further scanned using the Affymetrix GeneChip® Scanner 3000. The image data were analyzed with the GeneChip® Operating Software (GCOS) using Affymetrix default analysis settings. Arrays - after passing the quality control - have been commonly normalized by the log scale robust multi-array analysis (RMA) (Irizarry et al. 2003). After outlier removal using the Nalimov test at $p < 0.001$ a parametric ANOVA (F-test) has been applied to obtain global expression differences between the different conditions. Hierarchical clustering of expression values of probe sets differentially expressed in the ANOVA was performed by using Cluster version 2.11 (Eisen et al. 1998) applying mean-centering and normalization of genes and arrays before average linkage clustering with uncentered correlation. Raw data are available in **Supplementary Table VI and VII.**

References

- Aksoy I, Sakabedoyan C, Bourillot PY, Malashicheva AB, Mancip J et al. (2007) Self-renewal of murine embryonic stem (ES) cells is supported by the serine/threonine kinases Pim-1 and Pim-3. *Stem Cells* 25: 2996-3004.
- Anneren C, Cowan CA, Melton DA (2004) The Src family of tyrosine kinases is important for embryonic stem cell self-renewal. *J Biol Chem* 279(30): 31590-31598.
- Aubert J, Dunstan H, Chambers I, Smith A (2002) Functional gene screening in embryonic stem cells implicates Wnt antagonism in neural differentiation. *Nat Biotechnol* 20(12): 1240-1245.
- Boeuf H, Hauss C, Graeve FD, Baran N, Kedinger C (1997) Leukemia inhibitory factor-dependent transcriptional activation in embryonic stem cells. *J Cell Biol* 138(6): 1207-1217.
- Boiani M, Scholer HR (2005) Regulatory networks in embryo-derived pluripotent stem cells. *Nat Rev Mol Cell Biol* 6(11): 872-884.
- Burdon T, Stracey C, Chambers I, Nichols J, Smith A (1999) Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. *Dev Biol* 210(1): 30-43.
- Cartwright P, McLean C, Sheppard A, Rivett D, Jones K et al. (2005) LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development* 132(5): 885-896.
- Chambers I (2004) The molecular basis of pluripotency in mouse embryonic stem cells. *Cloning Stem Cells* 6: 386-391.
- Chen X, Xu H, Yuan P, Fang F, Huss M et al. (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133(6): 1106-1117.
- Dvash T, Mayshar Y, Darr H, McElhaney M, Barker D et al. (2004) Temporal gene expression during differentiation of human embryonic stem cells and embryoid bodies. *Hum Reprod* 19(12): 2875-2883.
- Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95(25): 14863-14868.
- Fujikura J, Yamato E, Yonemura S, Hosoda K, Masui S et al. (2002) Differentiation of embryonic stem cells is induced by GATA factors. *Genes Dev* 16(7): 784-789.
- Futaki S, Hayashi Y, Emoto T, Weber CN, Sekiguchi K (2004) Sox7 plays crucial roles in parietal endoderm differentiation in F9 embryonal carcinoma cells through regulating Gata-4 and Gata-6 expression. *Mol Cell Biol* 24(23): 10492-10503.
- Ghaleb AM, Nandan MO, Chanchevalap S, Dalton WB, Hisamuddin IM et al. (2005) Kruppel-like factors 4 and 5: the yin and yang regulators of cellular proliferation. *Cell Res* 15(2): 92-96.
- Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* (Oxford, England) 4(2): 249-264.
- Jiang J, Chan YS, Loh YH, Cai J, Tong GQ et al. (2008) A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat Cell Biol* 10: 353-360.
- Kalantry S, Manning S, Haub O, Tomihara-Newberger C, Lee HG et al. (2001) The amnionless gene, essential for mouse gastrulation, encodes a visceral-endoderm-specific protein with an extracellular cysteine-rich domain. *Nat Genet* 27(4): 412-416.

- Kanai-Azuma M, Kanai Y, Gad JM, Tajima Y, Taya C et al. (2002) Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development* 129(10): 2367-2379.
- Kobayashi T, Cohen P (1999) Activation of serum- and glucocorticoid-regulated protein kinase by agonists that activate phosphatidylinositol 3-kinase is mediated by 3-phosphoinositide-dependent protein kinase-1 (PDK1) and PDK2. *Biochem J* 339 (Pt 2): 319-328.
- Li Y, McClintick J, Zhong L, Edenberg HJ, Yoder MC et al. (2005) Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. *Blood* 105(2): 635-637.
- Matsuda T, Nakamura T, Nakao K, Arai T, Katsuki M et al. (1999) STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *Embo J* 18(15): 4261-4269.
- Matsuura R, Kogo H, Ogaeri T, Miwa T, Kuwahara M et al. (2006) Crucial transcription factors in endoderm and embryonic gut development are expressed in gut-like structures from mouse ES cells. *Stem Cells* 24(3): 624-630.
- McConnell BB, Ghaleb AM, Nandan MO, Yang VW (2007) The diverse functions of Krüppel-like factors 4 and 5 in epithelial biology and pathobiology. *Bioessays* 29(6): 549-557.
- McGrath KE, Koniski AD, Maltby KM, McGann JK, Palis J (1999) Embryonic expression and function of the chemokine SDF-1 and its receptor, CXCR4. *Dev Biol* 213(2): 442-456.
- Nakatake Y, Fukui N, Iwamatsu Y, Masui S, Takahashi K et al. (2006) Klf4 cooperates with Oct3/4 and Sox2 to activate the Lefty1 core promoter in embryonic stem cells. *Mol Cell Biol* 26(20): 7772-7782.
- Niimi T, Hayashi Y, Futaki S, Sekiguchi K (2004) SOX7 and SOX17 regulate the parietal endoderm-specific enhancer activity of mouse laminin alphal gene. *J Biol Chem* 279(36): 38055-38061.
- Niwa H, Burdon T, Chambers I, Smith A (1998) Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev* 12(13): 2048-2060.
- Okita K, Ichisaka T, Yamanaka S (2007) Generation of germline-competent induced pluripotent stem cells. *Nature* 448: 313-319.
- Paling NR, Wheeldon H, Bone HK, Welham MJ (2004) Regulation of embryonic stem cell self-renewal by phosphoinositide 3-kinase-dependent signalling. *J Biol Chem* 279(46): 48063-48070.
- Parisi S, Passaro F, Aloia L, Manabe I, Nagai R et al. (2008) Klf5 is involved in self-renewal of mouse embryonic stem cells. *J Cell Sci* 121:2629-2634.
- Park J, Leong ML, Buse P, Maiyar AC, Firestone GL et al. (1999) Serum and glucocorticoid-inducible kinase (SGK) is a target of the PI 3-kinase-stimulated signaling pathway. *Embo J* 18(11): 3024-3033.
- Qu XB, Pan J, Zhang C, Huang SY (2008) Sox17 facilitates the differentiation of mouse embryonic stem cells into primitive and definitive endoderm in vitro. *Dev Growth Differ* (Epub ahead of print).
- Rowland BD, Peepo DS (2006) KLF4, p21 and context-dependent opposing forces in cancer. *Nat Rev Cancer* 6(1): 11-23.
- Sato N, Meijer L, Skaltsounis L, Greengard P, Brivanlou AH (2004) Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nat Med* 10(1): 55-63.
- Schmitz J, Weissenbach M, Haan S, Heinrich PC, Schaper F (2000) SOCS3 exerts its inhibitory function on interleukin-6 signal transduction through the SHP2 recruitment site of gp130. *J Biol Chem* 275(17): 12848-12856.

- Seguin CA, Draper JS, Nagy A, Rossant J (2008) Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. *Cell Stem Cell* 3(2): 182-195.
- Sekkai D, Gruel G, Herry M, Moucadel V, Constantinescu SN et al. (2005) Microarray analysis of LIF/Stat3 transcriptional targets in embryonic stem cells. *Stem Cells* 23(10): 1634-1642.
- Shimosato D, Shiki M, Niwa H (2007) Extra-embryonic endoderm cells derived from ES cells induced by GATA factors acquire the character of XEN cells. *BMC Dev Biol* 7: 80.
- Ura H, Usuda M, Kinoshita K, Sun C, Mori K et al. (2008) STAT3 and Oct-3/4 control histone modification through induction of Eed in embryonic stem cells. *J Biol Chem* 283(15):9713-23
- Vallier L, Mancip J, Markossian S, Lukaszewicz A, Dehay C et al. (2001) An efficient system for conditional gene expression in embryonic stem cells and in their in vitro and in vivo differentiated derivatives. *Proc Natl Acad Sci U S A* 98(5): 2467-2472.
- Wilkinson DG, Bhatt S, Herrmann BG (1990) Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* 343(6259): 657-659.
- Wu Q, Chen X, Zhang J, Loh YH, Low TY et al. (2006) Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. *J Biol Chem* 281: 24090-24094.
- Yang DH, Smith ER, Roland IH, Sheng Z, He J et al. (2002) Disabled-2 is essential for endodermal cell positioning and structure formation during mouse embryogenesis. *Dev Biol* 251(1): 27-44.
- Yasunaga M, Tada S, Torikai-Nishikawa S, Nakano Y, Okada M et al. (2005) Induction and monitoring of definitive and visceral endoderm differentiation of mouse ES cells. *Nat Biotechnol* 23(12): 1542-1550.
- Ying QL, Nichols J, Chambers I, Smith A (2003) BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* 115(3): 281-292.
- Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B et al. (2008) The ground state of embryonic stem cell self-renewal. *Nature* 453(7194): 519-523.
- Zhang J, Tam WL, Tong GQ, Wu Q, Chan HY et al. (2006) Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat Cell Biol* 8: 1114-1123.

Acknowledgements

We are grateful to Kenneth Knoblauch for help with the statistical analysis and Henry Kennedy for a critical reading of the manuscript.

Funding

This work was supported by research grants from the European Union 6th Framework Program (FunGenES, contract N°LSHG-CT-2003-503494), the Association Française contre la Myopathie, INSERM AVENIR program, Région Rhône-Alpes (Thématisques Prioritaires 2003-2005) and Fondation Bettencourt-Schueller. I.A. was a recipient of a fellowship from the Ministère de la Recherche et de l'Enseignement Supérieur and from the Fondation pour la Recherche Médicale.

Figure Legends

Figure 1: gene expression profiles of E14-S3ER and E14-CER ES cells and identification of STAT3 target genes.

(A) Schematic representation of the experimental strategy to identify STAT3 response genes. (B) Dendrogram of hierarchical clustering of probe sets differentially expressed between the five experimental conditions (ANOVA, $p < 0.05$). (C) K-means clustering of differentially expressed probe sets (ANOVA, $p < 0.001$) in E14-S3ER ES cells after withdrawal of both LIF and 4'OHT for 24 hours (E14-S3ER - LIF, - 4'OHT), followed by stimulation with LIF for 1 hour (E14-S3ER + LIF) or with 4'OHT for 2 hours (E14-S3ER + 4'OHT). Clusters C and F contain the probe sets that were up- and down-regulated upon stimulation with LIF only. Clusters B and E contain the probe sets that were up- and down-regulated following stimulation with 4'OHT only. Clusters A and D contain the probe sets that were up- and down-regulated upon stimulation with both LIF and 4'OHT (D) Chart pie representing the functional annotation of the selected probe sets according to DAVID 2007 bioinformatics “Gene Functional Classification” tool. (E) Expression of 59 unique probe sets measured by real-time PCR at different days after differentiation induced by the formation of embryoid bodies.

Figure 2: Genes regulated by STAT3 in E14-S3ER ES cells.

LIF, 4'OHT and 4'OHT + cyclo indicate fold-change in gene expression after stimulation with LIF, 4'OHT and 4'OHT + cycloheximide, respectively. Expression during differentiation is analysed by semi-quantitative RT-PCR during differentiation induced by formation of embryoid bodies (from day 1 to day 10). Gene symbols shown in gray boxes indicate correspond to genes which were functionally analysed.

Figure 3: Knockdown of STAT3-target genes with lentiviral vectors expressing shRNA.

(A) Histogram showing the percentages of undifferentiated (blue), mixed (red) and differentiated (green) colonies observed in clonal assays after knockdown of STAT3/Nanog response genes by means of lentiviral infection with the *pLenti6/BLOCK-iT-PGKneo^R* interfering vector expressing shRNA-1 (upper panel) and shRNA-2 (lower panel). CGR8 ES cells were infected with viral supernatants. Forty-eight hours after infection, cells were dissociated and replated at clonal density (100 cells/cm²). They were further grown for 6 days

in medium supplemented with LIF in the presence of G418 to kill non-infected cells. The resulting colonies were stained for alkaline phosphatase activity. Means and standard errors were calculated from three replicates using ANOVA test. Genes are shown according to the *P* value in the ANOVA test [from highest (left) to lowest (right)]. **(B)** Real-time PCR analysis of *Rex-1* (upper panel) and *Fgf5* (lower panel) expression in the colonies harvested in (A). Means and standard errors were calculated from three replicates using Student's t test (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). **(C)** Real-time PCR analysis of *Sox17* and *FoxA2* (upper panel) *T* and *Flk1* (lower panel) expression in the colonies harvested in (A). All normalizations was performed with β -actin.

Figure 4: expression of mesoderm and endoderm markers after knockdown of STAT3 target genes.

(A) Color dots represent ratios between endoderm (*FoxA2*, *GATA-4*, *GATA-6*, *Sox17* and) and mesoderm (*Flk1*, *Mixl1* and *T*) markers measured by real-time PCR. R scores are calculated for each of the 22 STAT3 target genes after knockdown. R scores are represented on a log scale. $R > 1$ indicates a bias towards endoderm differentiation, whereas $R < 1$ indicates a bias towards mesoderm differentiation. The 22 genes fall into three distinct categories. The difference of the log R scores from one in the three groups was tested with a one-way analysis of variance (cell means model) (***, $p < 0.001$). **(B)** The box-and-whiskers plots show the median values as solid lines with the interquartile range indicated by the box limits. The whiskers indicate approximately 1.5 times the distance away from the box. Group I shows no significant difference from 0 while the other two are highly significant ($p < 10^{-4}$).

Figure 5: Expression of markers associated with early differentiation into mesoderm and endoderm after knockdown of *Klf4* and *Klf5* expression.

ES cells were infected with *pLenti6/BLOCK-iT-PGKhygro^R*-derived lentiviral vectors expressing two different shRNAs for *Klf4* (*shKlf4-1* and *shKlf4-2*), two different shRNAs for *Klf5* (*shKlf5-1* and *shKlf5-2*), and a *Klf4/Klf5* scrambled shRNA as control (*shScrambled*). Histograms represent the expression levels of mesoderm (*Eomes*, *Flk1*, *Mixl1*, *T*), endoderm (*Cxcr4*, *Dab2*, *Emp2*, *FoxA2*, *Fxyd3*, *GATA-4*, *GATA-6*, *Sdc4*, *Sox7*, *Sox17*, and *Tm4sf2*) and ES (*Nanog*, *Oct-4*, *Rex-1*) markers, during differentiation induced by formation of embryoid bodies in the absence of LIF, and analysed by real-time PCR. d0 indicates undifferentiated ES cells, d1 to d10, days of differentiation in suspension.

Figure 6: Flow cytometry analysis of mesoderm and endoderm differentiation in embryoid bodies after knockdown of *Klf4* and *Klf5* expression.

T-GFP and Sox17-DsRed reporter ES cell lines were infected with *pLenti6/BLOCK-iT-PGKhygro*^R-lentiviral vectors expressing *shKlf4-1*, *shKlf4-2*, *shKlf5-1*, and *shKlf5-2*. After selection, hygromycin-resistant colonies were picked and analysed for the expression of *Klf4* and *Klf5* by real-time PCR. Normalization was performed with β -actin. With each one of the two reporter cell lines and each one of the four vectors, two independent clones showing strong interference were selected (data not shown). Clones expressing a *Klf4/Klf5* scrambled shRNA were engineered for control. Differentiation was induced by formation of embryoid bodies (day 1 to day 7) in hanging drops. d0 indicates undifferentiated ES cells. **(A)** Histograms representation of GFP- and DsRed-associated fluorescence in differentiating T-GFP-ES and Sox17-DsRed-ES clones, respectively. **(B)** Histograms showing the percentage of GFP⁺ and DsRed⁺ cells during differentiation of T-GFP-ES and Sox17-DsRed-ES clones, respectively. Means and standard deviations were calculated from three replicates using Student's t test.

Figure 7: Differential activation of mesoderm and endoderm markers in the pluripotent stem cell.

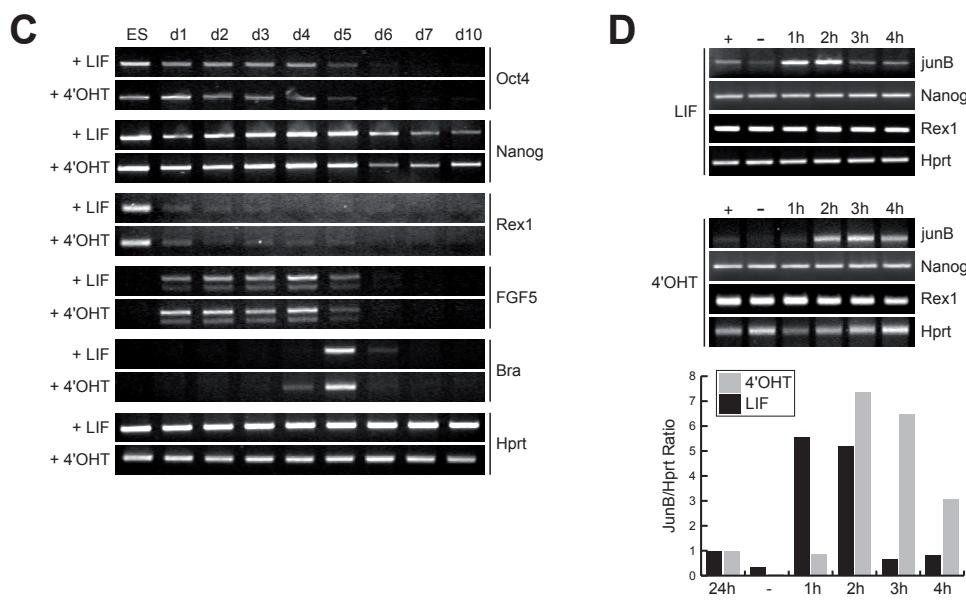
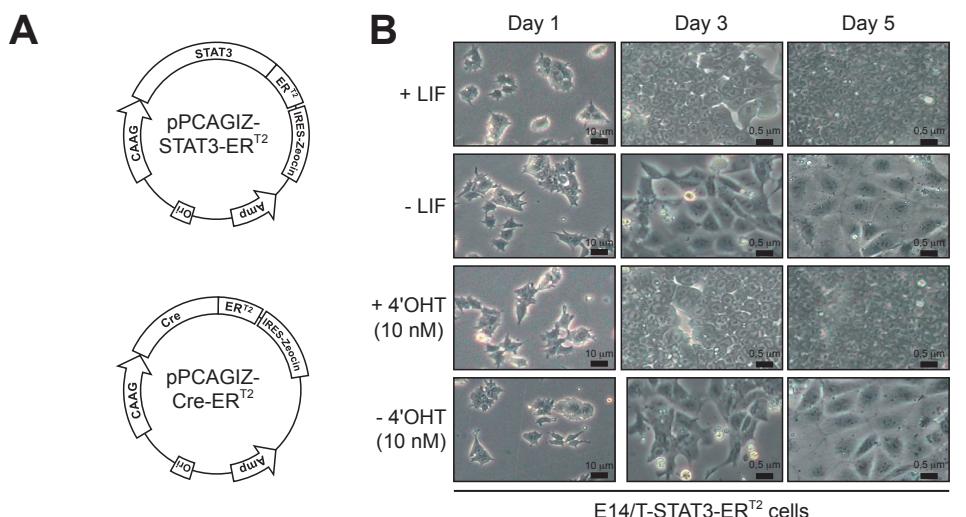
(A) Real-time PCR analysis of *Klf4*, *Klf5*, *Oct-4*, *Nanog*, *Rex-1*, and *Fgf5* expression in CGR8 ES cells stably expressing *Scrambled*, *Klf4* and *Klf5* shRNAs. Expression levels were normalized to control. Means and standard errors were calculated from two replicates using Student's t test. **(B)** Morphology of CGR8 ES cells expressing *Klf4*, *Klf5* and *Scrambled* shRNAs. Bar = 10 μ M. **(C)** Histogram showing the percentage of SSEA-1⁺ cells in the populations of ES cells expressing *Klf4*, *Klf5* and *Scrambled* shRNAs. **(D)** K-means clustering of differentially expressed probe sets (ANOVA, $p < 0.001$) in CGR8 ES cells expressing *Klf4*, *Klf5* and *GFP* shRNAs. Cluster V and VIII contain the probe sets that were up-regulated (cluster V) and down-regulated (cluster VIII) after *Klf4* knockdown. Cluster I, II and III contain the probe sets that were up-regulated (clusters II and III) and down-regulated (cluster I) after *Klf5* knockdown. Clusters VII, IX and X contain the probe sets that were up-regulated after both *Klf4* and *Klf5* knockdown, and cluster 4 and 6 those that were down-regulated after both *Klf4* and *Klf5* knockdown. Of note, each experimental condition resulted in more than one cluster. This results from differences in gene expression profiles after knockdown with two different shRNA. **(E)** Venn diagrams of the genes up- and down-

regulated by *Klf4* and *Klf5* in ES cells. (F) Real-time PCR analysis of *Sox7*, *Sox17*, *GATA-4*, *GATA-6*, *Dab-2*, *amn*, *Nodal*, *Lefty-1*, and *T* expression in CGR8 ES cells expressing shRNA of *Klf4*, *Klf5* and *Scrambled*, after normalization to β -actin. Means and standard errors were calculated from three replicates using Student's t test.

Figure 8: Model of regulation of pluripotency by *Klf4* and *Klf5*. Plain lines indicate positive (green) and negative (red) regulation of gene expression identified in this work. Dotted lines indicate the presence of Klf4 and Klf5 binding sites in the promoter of Nanog (Jiang et al. 2008), and of Nanog binding sites in the promoter of Klf4 and Klf5 target genes (Chen et al. 2008).

Supplementary Figure 1 : Episomal expression of ligand-dependent STAT3-ERT².

(A) *pPCAGIZ-STAT3-ERT²* episomal expression vector used to express STAT3-ERT² in E14/T ES cells. The plasmid carries a polyoma origin (*ori*) with the F101 mutation allowing episomal replication in ES cells. The cloned directionally within a transcription unit linked to the zeocine resistance gene (*zeo*) through an *IRES* and driven by the CAAG promoter. *pPCAGIZ-Cre-ERT²* was used as control (B) Morphology of E14/T ES cells expressing STAT3-ERT² after propagation for one month in medium supplemented with 1000 U/ml LIF (+ LIF), or with 10 nM 4'OHt (+ 4'OHt). - LIF and - 4'OHt indicate withdrawal of LIF and 4'OHt, respectively, to induce differentiation in monolayer. (C) Semi-quantitative RT-PCR analysis of *Oct4*, *Nanog*, and *Rex-1* (pluripotency markers), *Fgf5* and *T-Brachyury* (differentiation markers), and *hprt* (control) during differentiation of E14/T-STAT3-ERT² ES cells induced by formation of embryoiod bodies between day 1 (d1) and day 10 (d10). "+LIF" and "+4'OHt" indicate E14/T-STAT3-ERT² ES cells propagated for one month in the presence of 1000 U/ml LIF and 10 nM 4'OHt, respectively. (D) Semi-quantitative RT-PCR analysis of *Nanog*, *Rex-1* and *JunB* expression in E14/T-STAT3-ERT² cells following stimulation for 1-4 hours with 10,000 U/ml LIF to activate endogenous STAT3, or with 100 nM 4'OHt to activate the ligand-dependent STAT3. "+" indicate E14/T-STAT3-ERT² cells propagated for one month in medium supplemented either with 1000 U/ml LIF, or with 10 nM 4'OHt prior to experiment. "-" indicate deprivation of LIF or 4'OHt for 24 hours prior to restimulation. Histogram shows the quantification of *JunB* expression normalized to *hprt*.



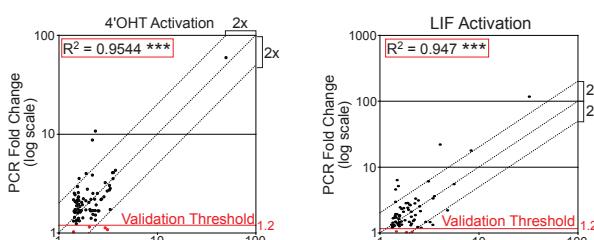
Aksoy, Bourillot et al,

Supplementary Figure 2 : Microarray data and PCR validation

p value and fold-change calculated from the microarray data and validated by real-time PCR for each gene selected.

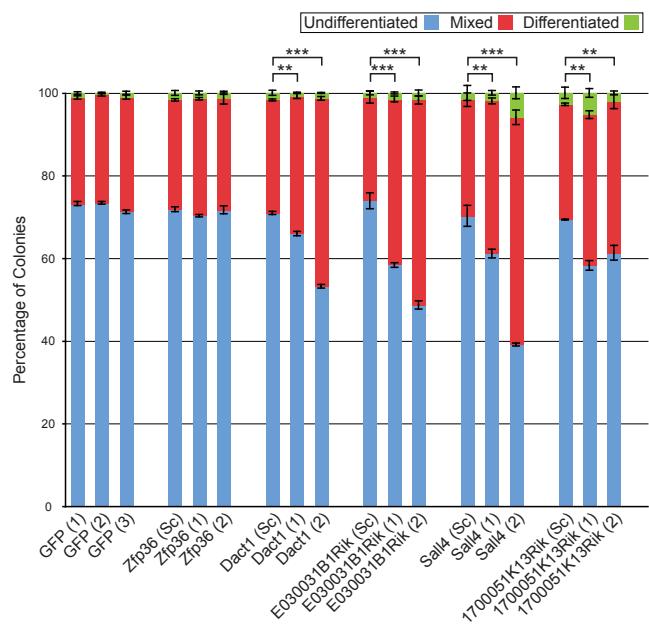
Probesets	Gene Symbol	STAT3-ER ^{+/+} /E14/T +/- 4'OH-T			STAT5-ER ^{+/+} /E14/T +/- LIF			Cre-ER ^{+/+} /E14/T +/- 4'OH-T			
		Fold Change	p-value	PCR	Fold Change	p-value	PCR	Fold Change	p-value		
1415899 at Jumb	-3.747086693	3.68266E-04	4.31	-0.453626284	1.26714E-02	22.05	1.172826857	1.01238E-01			
1400399 at Sfrp1	-1.940458221	1.00000E-05	1.00	-0.453626284	1.26714E-02	22.05	1.172826857	1.01238E-01			
1416047 at Sfrp2	-1.628845223	9.65503E-05	1.85	-1.510628771	1.48236E-03	1.78	1.002656534	9.1948E-02			
1416613 at Cyp11bt	-2.971532461	2.70854E-02	2.66	-2.349619122	6.49557E-03	5.52	-1.416668667	5.0941E-01			
1417394 at Kit	-3.306853387	6.20301E-04	1.75	-2.07556822	1.2811E-02	2.32	-1.70488473	9.76342E-01			
1417937 at Ddit1	-1.944332873	6.48622E-03	2.75	-1.377737853	3.20442E-04	1.32	1.068717088	8.8210E-01			
1420025 at Vim	-1.940458221	1.00000E-05	1.00	-0.453626284	1.26714E-02	22.05	1.172826857	1.01238E-01			
1420762 at Tnni3	-1.704516420	3.317882E-03	2.04	-1.507861383	7.57231E-03	1.27	1.026651118	5.59335E-02			
1421964 at Ier3	-1.873593098	7.94821E-03	1.71	-1.659962327	2.22824E-02	2.32	1.053653444	5.02123E-01			
1420337 at Gbx2	-3.25884974	3.24358E-03	2.75	-2.28059965	9.44598E-04	1.87	-1.01194415	9.4897E-01			
1420338 at Cnnm1	-1.568823709	7.80545E-04	2.58	-1.312339645	1.79181E-03	1.34	1.072163703	5.16332E-02			
1420365 at Sfrp2	-1.844945553	3.85025E-03	10.78	-1.490473784	3.8221E-02	6.35	1.0760398	4.83836E-01			
1423008 at Pim1	-1.444024947	4.09978E-04	1.35	-1.537891926	2.27056E-03	1.35	1.066724039	3.72181E-01			
1423100 at Fos	-2.95023014	9.50573E-06	3.23	-8.28715315	1.65684E-05	17.84	1.47425998	1.59833E-02			
1423388 at Smad2	-1.487998769	8.00112E-02	1.68	-1.753093204	1.69198E-03	2.1	-1.001848552	8.8210E-01			
1424047 at Smad1	-2.95023014	9.50573E-06	3.23	-8.28715315	1.65684E-05	17.84	1.47425998	1.59833E-02			
1424942 at Mmp1	-1.558877784	6.44178E-03	1.59	-1.413516085	5.58670E-02	3.33	-1.32471303	6.81778E-02			
1425503 at Gon2	-1.517998118	7.34241E-03	2.13	-1.408598059	1.77209E-02	1.41	-1.111221891	2.91195E-01			
1425837 at Comt4	-1.6016609	1.07978E-03	3.57	-1.94444119	2.94848E-02	5.16	-1.29831189	1.58405E-03			
1425922 at Nnyc1	-1.748700031	5.99558E-03	1.42	-1.409195624	3.21652E-03	2.98	-1.03967574	2.29599E-01			
1426025 at Rpl19	-2.302410804	9.29205E-03	1.32	-1.85044851	3.61522E-03	1.39	-1.05810852	3.09816E-01			
1426210 at Rpl19	-2.478317012	5.59296E-03	1.67	-2.61395772	3.43206E-02	1.2	-1.127075111	3.99442E-01			
1428780 at Gbbp1	-2.208277445	2.02048E-03	2.10	-2.072475792	3.26744E-03	1.99	-1.081563522	3.72172E-01			
1429438 at Bcor	-1.601409941	1.61563E-03	1.69	-1.335679527	1.45570E-02	1.2	-1.005324841	9.53856E-01			
1429600 at C530010507Rik	-2.01978022	5.53159E-02	2.04	-2.05010968	1.25172E-02	1.31	-1.18111588	4.75282E-01			
1429644 at Vim	-1.773624080	1.174454510	1.52	-1.886524080	3.61522E-03	1.39	-1.05810852	3.09816E-01			
1431094 at Slc4a1	-2.329383717	7.00110E-03	1.69	-2.294652038	1.58122E-03	3.01	-1.079027473	7.71047E-02			
1434025 at ---	-2.342099174	6.70396E-03	1.69	-1.97845442	1.20882E-02	2.41	-1.408617948	8.43447E-04			
1435284 at ---	-1.577972074	9.29205E-03	2.23	-2.249475074	1.57285E-02	2.35	-1.05810852	3.09816E-01			
1435654 at Sry2	-1.773624071	1.773624071	1.77	-1.886524071	3.61522E-03	1.39	-1.05810852	3.09816E-01			
1435954 at Sfrp2	-1.48304000M03Rk	1.849454543	4.78734E-03	2.10	-1.886524071	3.61522E-03	1.39	-1.05810852	3.09816E-01		
1437110 at Pim3	-1.844945553	3.19183E-03	1.41	-1.496736571	2.33904E-02	2.45	-1.05128666	3.16392E-01			
1437880 at Lbxw01	-2.492887745	3.78954E-03	1.58	-2.083485701	2.65660E-03	2.11	-1.12603288	2.43749E-03			
1438814 at ---	-1.492899051	3.39732E-03	1.32	-1.53565469	3.50482E-03	1.33	-1.037190143	6.08956E-01			
1439311 at E03001214Rik	-1.744045329	8.39732E-03	1.67	-1.598233449	1.89715E-02	2.03	-1.021229833	6.08919E-01			
1439465 at Tshz4	-1.503086152	7.87215E-04	3.03	-1.431932158	6.54753E-03	1.47	-1.011351749	8.84405E-01			
1440752 at Sdah4	-1.592959642	8.03275E-03	2.13	-1.497168552	1.77571E-02	1.5	-1.02140574	7.62111E-01			
1441231 at BC019575	-2.387362367	7.10046E-02	2.22	-2.90439604	1.67672E-02	1.48	-0.870710784	5.69203E-01			
1441718 at E032411E13Rik	-1.698184979	8.75976E-03	2.21	-1.815174423	5.65539E-03	2.85	-1.18860437	7.79346E-01			
1441843 at ---	-1.494545431	4.78734E-03	2.10	-1.886524071	3.61522E-03	1.39	-1.05810852	3.09816E-01			
1442009 at Vim	-2.732155004	7.43656E-03	2.08	-2.083485701	2.65660E-03	2.59	-1.05810852	3.09816E-01			
1443161 at E330040F24Rik	-2.339953337	2.11556E-03	871	-2.027277673	4.56553E-03	1.77	-0.717527773	2.78749E-01			
1443724 at ---	-1.409462164	3.39992E-03	1.32	-1.327283483	1.15779E-02	1.44	-1.04888794	4.68694E-01			
1444598 at 1700051K13Rik	-1.32867599	8.69852E-04	2.83	-1.398621298	1.58996E-03	2.21	-1.028921198	7.52194E-01			
1445245 at ---	-1.081857041	3.72151E-03	1.32	-1.497168552	1.77571E-02	1.5	-1.021229833	6.08919E-01			
1445625 at ---	-1.503086152	7.87215E-04	3.03	-1.431932158	6.54753E-03	1.47	-1.011351749	8.84405E-01			
1446550 at Gapt1	-1.61100738	9.42009E-03	1.33	-1.412652285	3.27035E-02	1.29	-1.163023945	1.46989E-01			
1447286 at ---	-1.457439756	3.71441E-03	1.86	-1.576913974	2.49159E-03	1.98	-1.094523856	2.07140E-01			
1448873 at Osnl	-2.153225806	3.43616E-03	1.36	-3.04310382	3.77930E-03	1.33	-0.872464394	5.18738E-01			
1449009 at Vim	-1.737155004	7.43656E-03	2.08	-2.083485701	2.65660E-03	2.59	-1.05810852	3.09816E-01			
1449244 at ---	-1.422271032	1.143411639	1.39	-1.449986031	6.55680E-04	1.93	-1.102861711	5.96234E-02			
14498075 at B930079_03Rik	-1.888022097	1.001902E-02	2.12	-1.490498777	5.50507E-04	1.72	-1.156641523	4.64889E-01			
1459961 at H3077C12-3N1A	-1.896375638	1.08749E-05	3.55	-1.762981029	1.41383E-04	3.61	-1.05896542	3.15441E-01			
1460700 at Stat3	-2.342254757	3.29125E-06	1.52	-1.937012683	1.30110E-02	2.25	-1.011252171	8.56463E-01			

Scatter plots showing comparison of gene activation after 4'OHT or LIF stimulation, between Affymetrix data and PCR quantification. Dotted lines indicate the diagonal and 2-fold changes between the two measurements.



Supplementary Figure 3 : Knockdown of *Dact1*, *Sall4*, *E030031B1Rik*, and *1700051K13Rik* increases the percentage of mixed + differentiated colonies in a colony-forming assay.

Histogram showing the percentages of undifferentiated (blue), mixed (red) and differentiated (green) colonies observed in a clonal assay after knockdown of STAT3/Nanog response genes. *pLenti6/BLOCK-iT-PGKneo^R* interfering vectors expressing two different shRNAs [(1) and (2)] and one scrambled shRNA (Sc) were used to knockdown *Zfp36*, *Dact1*, *E030031B1Rik*, *Sall4*, and *1700051K13Rik*. CGR8 ES cells were dissociated 48 hours after infection and replated at clonal density (100 cells/cm²). They were further grown for 6 days in medium supplemented with LIF in the presence of G418 to kill non-infected cells. The resulting colonies were stained for alkaline phosphatase activity. Knockdown of *Dact1*, *E030031B1Rik*, *Sall4*, and *1700051K13Rik* results in an increase in the proportion of undifferentiated colonies in the colony-forming assay. In contrast, knockdown of *Zfp36* does not increase the proportion of undifferentiated colonies. Means and standard errors were calculated from three replicates (ANOVA: *, p < 0.05; **, p < 0.01; ***, p < 0.001).



Aksoy, Bourillot *et al*,

Supplementary Table I : Sequence of shRNA and expression levels of STAT3 target genes after transfection of pENTRY-shRNA plasmids measured by real-time PCR.

Probe ID	Gene Symbol	shRNA Sequence	N°	Efficiency % WT expression RT-PCR
1416039_x_at	Cyr61	AAGTGCCTCCGGACAACCA	1	52
		CGACTGTACAGCCTATTCA	2	63
1416041_at	Sgk	AAGTATTCTATGCACTAAAGTT	1	49
		AAGTCCCTCAACAAATCAACC	2	60
1417394_at	Kif4	GGACCTAGACTTTATCCTTC	1	58
		GGTCATCAGTGTTAGCAA	2	55
1417937_at	Dact1	AAGTTCTGGAAAGAACATTT	1	58
		GTAGCACCGTATATCGTTACAAA	Sc	100
1419647_a_at	ler3	GCAGATGTAATCTAAATAC	2	60
		GAGACTTCTCCGTGCTAATTA	1	31
1420337_at	Gbx2	GGATTATGGCTGGATCTAA	2	61
		GTACCGCTTCAGCATAGA	1	35
1420354_at	Cnnm1	GAGGCCCAATAATACTTTG	1	35
		GGACATTCTGTTGCTAAAGA	2	63
1422528_a_at	Zfp36l1	CAGACTCTCCATCTCAGATGACT	1	53
		GCTGCCACTTCTCATAACG	2	51
1423006_at	Pim1	GCAGAACCTCTTGACTTTAT	1	53
		AAGGACGAGAACATCTAACGA	2	58
1423389_at	Smad7	GAGTCAGCTGGTACAGAAAGTGC	1	25
		GCTTCAGATTCCTAACTTCT	2	52
1424067_at	Icam1	GAGGAGGTGAATGTATAAGTTAT	1	57
		GCTGGCATTTCTCTAATGT	2	65
1425837_a_at	Ccrn4l	GAGCCCAGTCAGCTTAAGGAAC	1	61
		GCTTACAGCAGCACCAAT	2	53
1426037_a_at	Rgs 16	TTACCGAACAAACTTCCA	1	54
		GAGTCCTACCGAACAAACT	2	55
1434025_at	---	AAGCGATCTTAATTAGAACTT	1	56
		CACCTACATCAGAACATAACTC	2	20
1436584_at	Spry2	AAGGAAGAACAAACAGAAGTTGA	1	56
		AAGGACAAGTACCCAGCAGCAGCT	2	71
1437100_x_at	Pim3	AAGCATTGTAACGTGGCTAATT	1	45
		GACACGGTCTACACTGACTTT	2	55
1440752_at	Sal 4	AAGGTGATACTTACGTTGAGTC	1	62
		GTTGCTATTGAGAATCAGT	Sc	100
1441843_s_at	5230400M03Rik	GGTGATATCTACGTTGAGTC	2	50
		CATTCCAGGTAGAAATTAT	1	71
1444598_at	1700051K13Rik	TAGGCCAAATTCTGCTTGTGTT	2	60
		GAGCTACAGTTGAACCTTGATGT	1	58
1446525_at	---	GTCATGAGAACGTCTCGTATAT	Sc	100
		GGGCGATCCACTGAAGAATA	2	62
1448873_at	Ocln	CTACAGTATGTCCTCTAGA	1	70
		GTTGCTAGTCCACTAATA	Sc	100
1451021_a_at	Kif5	CAACGATGCGAACATAAT	2	62
		AAGAACCTCTATACAGATCT	1	32
1452519_a_at	Zfp36	GATTATAATGGTTGACAA	2	36
		AAACGGAACCTCTCCACAA	1	51
1456160_at	---	GACCGACCAAGCATACTAA	Sc	100
		ACCACTCTCTCGATACA	2	62
1456629_a_at	Vim	GAATGTCAGACTTATGT	1	60
		CAGCAGGATCTTTGAA	2	62
1457402_at	E030031B1Rik	AAGGAGGAATCCGGCAAGAAT	1	61
		GAAGAGTAAAGACCGGGAGAATCT	Sc	100
1457402_at	---	AAGCTGCCAACGGTATAAACAGT	2	49
		GFP	1	100
1457402_at	---	GCCACAACGTCTATCAT	2	100
		TGAACCCGATCGAGCTGAA	3	100
1457402_at	---	ACAGCCACAACGCTATAT	1	13
		STAT3	1	21
1457402_at	---	GGAGCTGTTCAGAAACTTA	1	21
		Nanog	1	21
1457402_at	---	ATGCTGCTCGCTCATAA	1	21
		Loop	1	21

Aksoy, Bourillot *et al*,

Supplementary Table II : quantification by real-time of mesoderm and endoderm markers in CGR8 ES cells expressing the interfering lentiviral vectors.

Probesets	Gene Symbol	Endoderm								Mesoderm							
		FoxA2		GATA-4		GATA-6		Sox17		Flk1		Mixl1		T			
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
1416039_x_at	Cyr61	1.560	0.071	1.080	0.113	1.675	0.148	1.680	0.283	1.640	1.075	1.865	1.421	1.683	1.198		
1416041_at	Sgk	1.449	0.243	1.560	0.128	2.284	0.960	3.420	1.643	0.855	0.064	1.567	1.333	1.462	1.101		
1417394_at	Klf4	1.864	0.329	1.729	0.280	2.821	1.661	3.051	1.392	0.845	0.205	0.635	0.163	0.475	0.241		
1417937_at	Dact1	1.823	0.340	1.413	0.743	1.953	0.862	1.497	0.517	2.100	0.226	1.047	0.869	1.008	0.555		
1419647_a_at	Ier3	3.113	1.266	1.317	0.382	2.003	0.446	1.960	0.191	1.550	0.057	1.763	1.012	1.996	1.652		
1420337_at	Gbx2	1.193	0.411	1.023	0.199	1.448	0.511	0.908	0.421	0.850	0.184	0.782	0.190	0.629	0.155		
1420354_at	Cnmm1	2.577	0.484	1.640	0.613	2.477	1.423	2.973	0.891	1.190	0.396	0.890	0.238	0.752	0.466		
1422528_a_at	Zfp36l1	1.778	0.861	1.632	0.299	1.815	0.763	3.479	0.419	1.535	1.096	1.046	0.571	0.879	0.490		
1423006_at	Pim1	1.147	0.313	1.198	0.215	1.534	0.480	0.895	0.255	1.195	0.332	0.717	0.221	0.307	0.153		
1423389_at	Smad7	1.991	0.468	1.866	0.750	2.866	1.824	3.071	0.814	0.840	0.085	1.093	0.182	0.794	0.225		
1424067_at	Icam1	1.432	0.381	1.259	0.139	1.537	0.584	1.253	0.402	0.645	0.064	0.919	0.402	1.040	0.442		
1425837_a_at	Ccnm4l	1.863	0.427	1.387	0.129	1.753	0.724	1.493	0.667	1.760	0.467	1.977	1.761	1.454	0.858		
1426037_a_at	Rgs 16	1.823	0.487	1.730	0.798	2.370	1.131	2.567	1.497	1.725	1.068	0.737	0.497	0.709	0.583		
1436584_at	Spry2	1.600	0.747	1.070	0.365	1.457	0.939	2.057	0.982	0.600	0.000	0.837	0.360	0.499	0.417		
1437100_x_at	Pim3	2.474	1.108	2.022	0.850	2.922	1.718	5.958	4.836	0.950	0.141	1.031	0.852	0.612	0.375		
1440752_at	Sall4	1.345	0.336	1.600	0.351	1.472	0.215	2.080	1.185	0.840	0.311	0.685	0.215	0.273	0.046		
1441843_s_at	5230400M03Rik	2.333	0.380	1.310	0.276	1.847	0.445	1.817	0.502	0.985	0.134	0.667	0.261	0.501	0.289		
1444598_at	1700051K13Rik	1.800	1.152	1.153	0.172	1.190	0.324	1.620	0.399	2.115	0.544	1.153	0.440	1.203	0.447		
1448873_at	Ocn	1.823	1.106	3.054	2.825	4.686	4.706	3.136	1.085	1.490	0.410	1.240	0.652	1.155	0.108		
1451021_a_at	Klf5	0.657	0.136	0.853	0.100	0.478	0.186	0.592	0.028	3.265	0.389	1.499	0.945	2.036	0.691		
1456292_a_at	Vim	1.820	0.642	1.140	0.269	1.670	0.426	3.410	1.898	0.960	0.226	0.633	0.454	1.300	1.692		
1457402_at	E030031B1Rik	2.397	2.044	1.117	0.176	1.703	0.530	1.480	0.440	0.900	0.297	0.887	0.196	0.641	0.699		

Aksoy, Bourillot *et al*,

Supplementary Table III: shKlf4 and shKlf5 Microarray Data.

Genes Up-regulated after Klf4 Knock-Down

Symbol	Title	shKLF4 (1) /shGFP	shKLF4(2) /shGFP
Srgn	serglycin	30.03748	17.786636
Dab2	disabled homolog 2 (Drosophila)	15.072913	9.436624
Sox17	SRY-box containing gene 17	13.896669	8.765994
Cubn	cubilin (intrinsic factor-cobalamin receptor)	12.158086	6.913168
Col4a2	procollagen, type IV, alpha 2	8.120743	5.835191
Col4a1	procollagen, type IV, alpha 1	7.283191	5.155701
Glipr1	GLI pathogenesis-related 1 (glioma)	6.982737	4.753714
Gata6	GATA binding protein 6	4.981869	3.579051
Amn	amnionless	4.553164	2.620456
Tcfec	transcription factor EC	4.402458	2.590347
Nostrin	nitric oxide synthase trafficker	3.87252	2.757978
Foxq1	Forkhead box Q1	3.545264	2.68902
Pdgfra	platelet derived growth factor receptor, alpha polypeptide	3.511711	2.858963
Klb	klotho beta	3.462642	2.88927
Bmper	BMP-binding endothelial regulator	2.911139	2.415208
Pthr1	parathyroid hormone receptor 1	2.458505	1.772142
Sox7	SRY-box containing gene 7	2.424708	1.584522
Stra8	stimulated by retinoic acid gene 8	2.386783	2.40465
Cdkn1c	cyclin-dependent kinase inhibitor 1C (P57)	2.370103	1.994821
Hs3st1	heparan sulfate (glucosamine) 3-O-sulfotransferase 1	2.303647	1.679485
Cyp4f14	cytochrome P450, family 4, subfamily f, polypeptide 14	2.212524	1.720016
Foxa2	forkhead box A2	1.910825	1.60336
Ceacam20	CEA-related cell adhesion molecule 20	1.830529	2.404308
Gata4	GATA binding protein 4	1.82836	1.657075
D4Wsu53e	DNA segment, Chr 4, Wayne State University 53, expressed	1.765902	1.627535
Enc1	ectodermal-neural cortex 1	1.670316	1.519116

Genes Up-regulated after Klf5 Knock-Down

Symbol	Title	shKLF5 (1) /shGFP	shKLF5(2) /shGFP
Thbs1	thrombospondin 1; similar to thrombospondin 1	28.464442	3.656341
Tagln	transgelin	16.380328	3.813579
Krt8	keratin 8	13.741447	3.311065
Krt18	keratin 18	11.827579	3.379037
Acta2	actin, alpha 2, smooth muscle, aorta	8.914854	1.903447
T	brachury	8.321972	2.539747
Pmp22	peripheral myelin protein	7.525864	2.273637
Acta1	actin, alpha 1, skeletal muscle	6.434925	1.88339
Cald1	caldesmon 1	5.75022	2.204188
Cxcl10	chemokine (C-X-C motif) ligand 10	4.976498	1.575941
Ghr	growth hormone receptor	4.954894	1.526022
Anxa1	annexin A1	4.35084	1.695184
Pitx2	paired-like homeobox transcription factor 2	4.188978	2.595477
Clu	clusterin	4.147603	3.434364
Actc1	actin, alpha, cardiac	4.147222	1.974137
Sfn	stratifin	3.477801	1.719875
Phlda1	pleckstrin homology-like domain, family A, member 1	3.461475	2.071678
Vgf	VGF nerve growth factor inducible	2.998383	1.595187
Sept4	septin 4	2.960566	1.606434
Axl	AXL receptor tyrosine kinase	2.937375	1.517
Ptg52	prostaglandin-endoperoxide synthase 2	2.752777	1.693648
Pou3f1	POU domain, class 3, transcription factor 1	2.585223	1.866069
Cited1	Cbp/p300-interacting transactivator with Glu/Asp-rich carboxy-terminal domain 1	2.56548	1.728614

Supplementary Table IV: Oligonucleotide sequences.

Probe ID	Gene	5' Primer	3' reverse
1455899_x_at	Socs3	TCAAGGGACAGCCTCTGA	TGTAAACCTCGAGGTGGAAG
1416039_x_at	Cyr61	CGGGCATTATTGCTCCAT	GAAGGAATGAGAGGCCTTT
1416041_at	Sgk	CAGACGGTTTGTGTGAGC	ACTGGGCATTGGTCATAA
1416613_at	Cyp1b1	CAATCCGTTTGACATCAGG	TCGAAGCAAACCAAGAATTG
1417394_at	Klf4	GAAGACCAGGATTCCTTGA	CCAAGCACCATCATTAGGC
1417937_at	Dact1	ACCTTGCACCTGTAGTGC	GTCAAGCCCAGTAAGGGTGA
1418025_at	Bhlhb2	AGCTTCTCAAGGCCAGCA	CACAAAGCAGACTGCATTGG
1418507_s_at	---	CAG GTG CAA CTC AAT CCT TG	TGT TGA TGT CCG TGG TCA GA
1418756_at	Trh	ATCCCTTAAGCAAGGGGT	GGAGGATGCGCTGAAGTTATAC
1419647_a_at	Ier3	GGCAACGCTTAACAGAAC	TGTTACAGCAGGGAGGCATA
1420337_at	Gbx2	GGCACCTCTAGATGTGAC	AAAACACTGCACCTGAGATCC
1420354_at	Cnnm1	TAGAGGAGCCGGTGTGAAAA	ATAGACTCGGAAGCCAAA
1422528_a_at	Zfp361	GTGCACCTTCTTCGGTCT	TTCAACCTTGTCCCTGGAG
1422914_at	Sp5	TCCTGCCAACAGACATCTTC	TACGACCCAGTTCTGTCC
1423006_at	Pim1	GCC CTC CTT TGA AGA AAA CC	GGA CCTGGA GTC TGG AAT GA
1423100_at	Fos	CTACTGTGTTCTGGCAATA	GTCAGAACATTCAAGACCACC
1423389_at	Smad7	ATCCCTGTGCTGAGCTCTG	TGCTCATTGAGCTAAGAACAGT
1424942_a_at	Myc	GACCGATCCCTGAATTGGA	CGAAGCTGTTGAGTTGTG
1425503_at	Gcnt2	GCCCCATATGCATACACACATAC	TCAGTCCAATGACAGCACAAGA
1425837_a_at	Ccrn4l	GGCACTTGTGCCTTCTGT	CCTCCAAGACGTGAGCTTC
1425922_a_at	Nmyc1	AATGGTGCTTAAGTCCAGC	CTTGCAGTCTAACTACTGGCC
1426037_a_at	Rgs 16	GAAACCAACTTGCCTGCCTA	GACAATCGGCACAAACACAAC
1424067_at	Icam1	GCTCAGATGCTGTGAGAGT	AAAGTAGGTGGGGAGGTGCT
1428055_at	Rian	CGAGATTTCGCCATGTAGGC	CAACTCATACGCCACCCCTT
1428765_at	Gtl2	TGGGCAGGTTTCTGTCTTC	CGACGGAACAAGAGTCCATT
1429438_at	Bcor	GCTCTGTGCTTGGCTAATT	CCACTGGATCACGTCCTTCAATA
1429600_at	C630010D07Rik	TTACTTTGTAGTGTGAAAGCC	ATCACAGTTACTTAATGACAGATT
1430448_at	6720418B01Rik	CTCTGAATTGCTGCCACAG	TCAAAGTGGCTTCTCCTC
1431094_at	1110006E14Rik	TTAAGGTCACTGGCTGAC	GAAACTGCAGACCCAAAAGC
1434025_at	---	CGTGTAAATAGGCGACAAATGG	GACAGCTGCTGCTGTATTTG
1435284_at	Rtn4	AACCCCTAGCACTGTGTT	CTGATGTCTGCACTTCTCA
1436584_at	Socs2	CACTCTAGGCAGAACAGTTGT	GTGACTGTGCCATGAAGCAT
1436766_at	A130042L18Rik	CCTACTAGAGACTAGCCTCAGAAAACC	CAACTAAACACCGTGGGGTAAG
1437100_x_at	Pim3	AGCAGTGACCTCTGACCCCT	TCAAGTATCCACCCAGGGCA
1437543_at	---	TGCAGCCCTGTAGTTGAT	CCCGGCTCAAGTTGTTCTT
1437880_at	Lbxcor1	TGGCTATGTCAATGGAGAC	GAGTCAGTTGTCGCCCTTT
1438814_at	---	CAGCAGTGTATGTGTGAA	TGGGCATCCTCTTAAGTAGA
1438906_at	---	CAGAGAACACAGCAGTTTCCAGTC	TAACCTGTCTATCCAGCAGTGAGG
1439311_at	B830012L14Rik	CATGTCTCAGCCCCCTAT	CCACCTGAAAACCAATCACA
1440001_at	---	CTGGTCTCCCATAGTTGTTCT	GATTGCACACACCTGTATGG
1440682_at	7030408E20Rik	CCTTGAACATCTCCAAAGC	TGCTAAGAACCCCTGCCAAT
1440752_at	Sal 4	AGTTGCGCTCGACCTCTCAA	CAGATAGCACCTAAAGACTCAACG
1441231_at	BCO19575	GCATTCTTAAAGAGCTCCGACA	TCTGCACTGCCAGCTTCAA
1441354_at	---	TAATGCAGCACAGGAGTGG	TATCAAGGGGACTGGAGAGATG
1441466_at	---	TCCGTGCTTTCTCTTGC	TGCGCAACTGAACATTACT
1441718_at	4432411E13Rik	GGAGGGAAAGCATCTCAT	GCCTGGTAACCTTCAAAT
1441769_at	---	CCGATAACCTTTCTGAAGC	CAAATGGTCTCTGAGCTGTC
1441843_s_at	5230400M03Rik	ATGACAGGCTAGGGAGTGG	GGGAAAGAACATGGGGTACT
1442484_at	---	CTGGCCCCATAATTCTACACC	AATGGTTCTGGTCCCCACTC
1443125_at	---	TGCAGGTATACTGACTGGCACT	TGAGCCCACAAATCAAGAGC
1443161_at	E330040F24Rik	CTGTGACACTGTAAGCAAATA	AGGTGCTGAAGGAGCAACAA
1443721_x_at	Stno	AGCAGGCTATACTGGCA	AGGTGACCTGATGCTGG
1443911_at	---	AAGCCTGAGTATTGGAAAACA	CTTATTCCTGAGCTGGACTGG
1444023_at	---	ATTCAATTGGTGACCGAAAGC	TTTCCTGGTTGGGCATA
1444598_at	1700051K13Rik	GGCGGATCCACTGAAGAATA	ACAGGGCAGCTGAGAAGAAA

**RÉSULTATS
COMPLÉMENTAIRES
ET
DISCUSSION**

Les cellules souches embryonnaires (cellules ES) de souris sont des cellules pluripotentes capable de s'autorenouveler de façon permanente sous l'action du LIF. Le facteur de transcription STAT3, recruté et activé au niveau du complexe ligand-récepteur LIF/LIFR-gp130, joue un rôle crucial dans le maintien de la pluripotence des cellules ES par l'inhibition de leur différenciation en endoderme et mésoderme (Matsuda et al., 1999 ; Niwa et al., 1998 ; Ying et al., 2003). En dépit du rôle central joué par la voie LIF/STAT3 dans le contrôle de l'autorenouvellement des cellules ES, ses gènes cibles dont l'activité serait responsable de l'inhibition de la différenciation sont pour la plupart inconnus. L'identification des gènes cibles du facteur de transcription STAT3 constitue donc une étape cruciale vers la compréhension des mécanismes de contrôle de l'autorenouvellement des cellules ES.

1 - Identification des gènes cibles de la voie LIF/gp130/STAT3.

1.1. Validation du modèle cellulaire.

Pour identifier les gènes cibles de STAT3 nous avons mis au point une lignée de cellules ES surexprimant le mutant conditionnel STAT3-ER^{T2}. Les cellules ES-STAT3-ER^{T2} sont capables de s'autorenouveler dans un milieu supplémenté avec du 4'OHT en l'absence de LIF. Ce résultat, déjà décrit par Matsuda et ses collègues, indique que l'activation du facteur STAT3 par le récepteur du LIF est suffisante pour inhiber la différenciation des cellules ES (Matsuda et al., 1999). L'inhibition de la différenciation induite par le 4'OHT est attestée (*i*) par l'aspect morphologique des colonies de cellules ES caractéristique des cellules maintenues en autorenouvellement sous l'action du LIF (Figure supplémentaire 1B de l'article n°2), et (*ii*) par la forte expression des gènes *Oct4*, *Nanog* et *Rex1* codant pour des facteurs de transcription spécifiques des cellules souches embryonnaires (Figure supplémentaire 1C de l'article n°2). Lorsque que le 4'OHT est supprimé du milieu de culture, les cellules ES se différencient spontanément. Cette différenciation est attestée (*i*) par l'aspect morphologique des cellules ES caractéristique des cellules différencierées (Figure supplémentaire 1B de l'article n°2) et (*ii*) par l'expression des gènes *Fgf-5* et *Brachyury*, des marqueurs de la différenciation des cellules souches embryonnaires (Figure supplémentaire 1C de l'article n°2). Ainsi, les cellules ES-STAT3-ER^{T2} maintenues en autorenouvellement sous l'action du 4'OHT (activation du facteur STAT3 conditionnel) et sous l'action du LIF (activation du facteur STAT3 endogène) sont indistinguables sur les quatre critères analysés.

Le domaine ER^{T2} que nous avons utilisé est différent de celui utilisé par Matsuda et ses collègues (Matsuda et al., 1999). Il contient, en effet, plusieurs mutations qui augmentent l'affinité du domaine ER pour le 4'OHT (Indra et al., 1999). Ainsi, la dissociation du domaine ER^{T2} de la protéine Hsp90 et l'activation consécutive du facteur STAT3-ER^{T2} sont obtenues avec une concentration de 10 nM alors qu'une concentration de 1000 nM est nécessaire à la dissociation du domaine ER utilisé par Matsuda et ses collègues.. Cette amélioration de l'efficacité du mutant conditionnel du facteur STAT3 est importante pour l'analyse du transcriptome. Elle permet de réduire la concentration de 4'OHT utilisée pour l'activation de STAT3 et de limiter ainsi les éventuels effets secondaires du 4'OHT sur l'activité génique des cellules ES.

Notre objectif était d'identifier les gènes cibles du facteur de transcription STAT3 dans les cellules ES surexprimant STAT3-ER^{T2} grâce à une analyse comparée des transcriptomes avant et après traitement par le 4'OHT. Il était donc important (*i*) de vérifier que l'expression d'un gène cible déjà connu, *junB*, est effectivement régulée par le 4'OHT et (*ii*) de déterminer les conditions de stimulation par le 4'OHT permettant d'optimiser le différentiel d'expression. Nous avons ainsi déterminé que des cellules ES-STAT3-ER^{T2} privées de LIF et de 4'OHT pendant 24 h puis restimulées par le 4'OHT pendant 2h réactivent l'expression du gène *junB*. Le différentiel d'expression entre les deux états est estimé à 10 fois, une valeur proche de celle calculée avec des cellules ES privées puis restimulées par le LIF (14 fois). Nous pouvons donc en conclure que les cellules ES-STAT3-ER^{T2} sont capables de réguler l'expression du gène *junB* via le facteur STAT3-ER^{T2} ou via le facteur STAT3 endogène avec des efficacités comparables.

En revanche, la cinétique d'activation du gène *junB* est différente selon que les cellules sont stimulées par le LIF ou par le 4'OHT :

- la stimulation par le LIF active le gène *junB* plus rapidement (dès 1 h) (Figure supplémentaire 1D de l'article n°2). Cette expression est transitoire et décroît très rapidement. Cette décroissance rapide résulte vraisemblablement d'un rétrocontrôle négatif exercé par le gène *Socs-3* (*suppressor of cytokine signaling-3*). Le gène *Socs-3* est lui-même un gène cible du facteur STAT3 et la protéine Socs-3 inhibe l'activité des kinases JAK nécessaires à la phosphorylation du facteur STAT3 endogène recruté par le récepteur (Duval et al., 2000) ;

- la stimulation par le 4'OHT active le gène *junB* seulement après 2h et cette expression est maintenue pendant au moins 3h (Figure supplémentaire 1D de l'article n°2). Une telle différence dans la durée d'expression pourrait s'expliquer par l'absence de

rétrocontrôle négatif de l'activité STAT3 dans le système expérimental STAT3-ER^{T2} dont le fonctionnement ne passe pas par les kinases JAK.

Nous avons utilisé cette lignée de cellules pour identifier les gènes cibles de STAT3 grâce à une analyse comparée des transcriptomes avant et après stimulation par le 4'OHT avec la technique des puces à ADN Affymétrix. Nous avons ainsi identifié 58 gènes cibles de STAT3. Parmi ces 58 gènes, 35 sont des gènes connus, 13 sont des RIKENs et 10 sont des gènes inconnus. Ces gènes sont activés à la fois par le facteur STAT3 endogène (après stimulation des cellules par le LIF) et par le mutant hormono-dépendant STAT3-ER^{T2} (après stimulation des cellules par le 4'OHT) en présence d'un inhibiteur de la synthèse protéique. Sur les 35 gènes cibles connus, mis en évidence dans notre étude, 23 comportent des sites de fixation de STAT3 dans leurs séquences régulatrices (Chen et al., 2008). De plus, les gènes *junB*, *c-fos*, *Socs-3*, *Klf4* et *c-myc* qui ont tous été identifiés comme des gènes cibles de STAT3 dans d'autres études (Matsuda et al., 1999) (Duval et al., 2000) (Li et al., 2005) (Cartwright et al., 2005) font partie de notre liste de gènes cibles. Ces observations confirment la validité de la stratégie expérimentale que nous avons utilisée pour identifier les gènes cibles directes de la voie LIF/STAT3.

1.2. Exhausitivité de l'analyse.

Nous pouvons nous demander si la liste des gènes cibles de STAT3 est vraiment exhaustive. Il semblerait que non. D'une part, les gènes *Dax1* et *Eed*, identifiés comme des gènes cibles du facteur de transcription STAT3 (Niakan et al., 2005) (Ura et al., 2008) ne figurent pas dans notre liste. D'autre part, nous avons réalisé les analyses de puces à ADN sur des cellules ES-STAT3-ER^{T2} préalablement adaptées à la culture en 4'OHT. Or, la transition de ces cellules d'un milieu de culture supplémenté en LIF à un milieu de culture privé de LIF et supplémenté en 4'OHT a été suivie d'une courte période (7 jours) pendant laquelle nous avons observé la différenciation et la mort d'une fraction de la population cellulaire. Il se peut que pendant cette « crise » certains gènes cibles de STAT3 aient été activés de manière spontanée afin de permettre la survie et le maintien en autorenouvellement des cellules. Ces gènes auraient alors échappé à l'analyse.

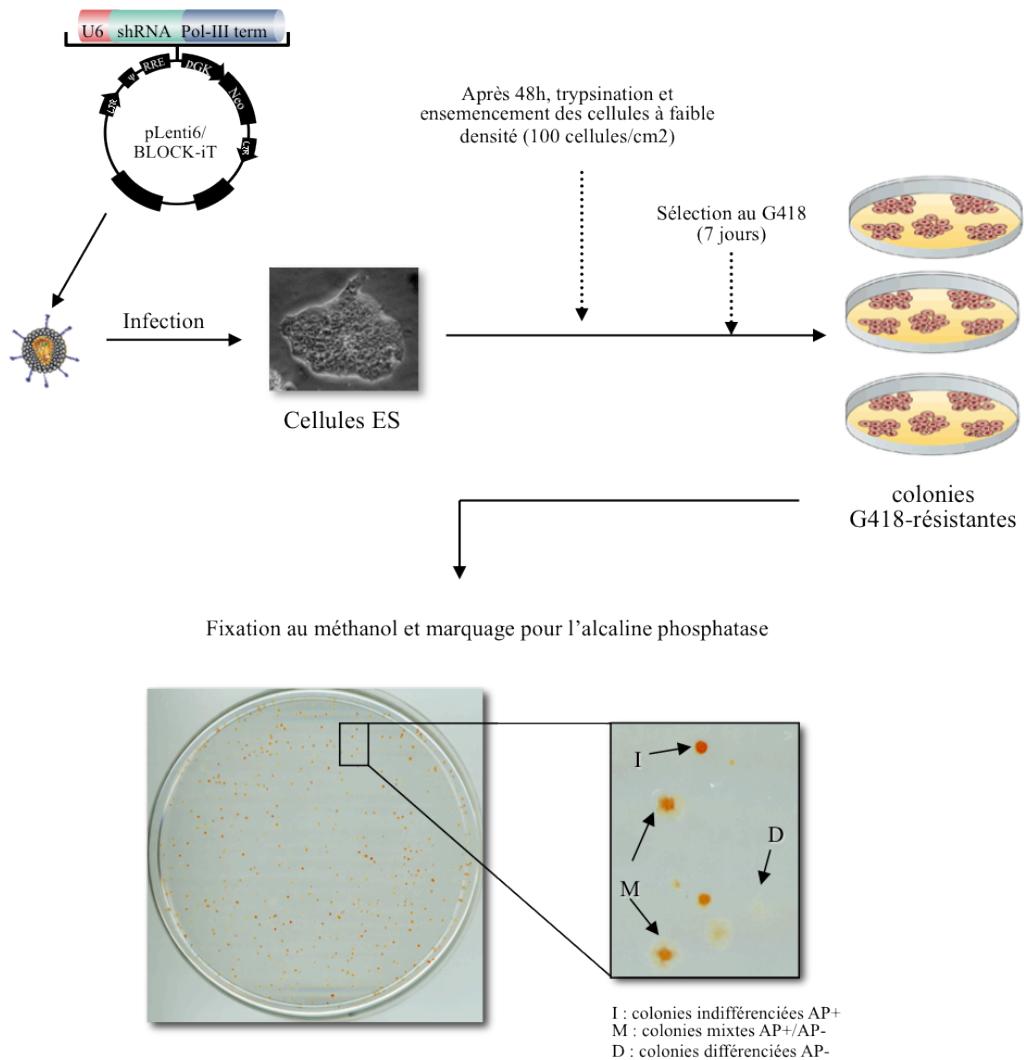


Figure 19 : Test d'autorenouvellement

Les cellules ES sont infectées avec deux vecteurs lentiviraux interférants (plenti6/BLOCK-iT commercialisé par Invitrogen) exprimant un gène de résistance à la néomycine et une séquence de shRNA spécifique de chaque gène. 48h après infection, les cellules sont trypsinées et ensemencées à faible densité à raison de 100 cellules / cm². Elles sont ensuite soumises à une sélection au G418 pendant 7 jours. Au bout des 7 jours, les colonies de cellules ES résistantes qui se sont formées sont fixées au méthanol et marquées pour l'acaline phosphatase (marqueur des cellules ES indifférenciées). Trois types de colonies sont obtenues : (i) I : les colonies indifférenciées composées de petites cellules très compactes ayant une forte activité acaline phosphatase AP⁺, (ii) M : les colonies mixtes composées de cellules indifférenciées au centre AP⁺ et de cellules différencierées à la périphérie AP⁻, et (iii) D : les colonies différencierées composées de cellules acaline phosphatase négatives AP⁻. Le pourcentage de colonies indifférenciées, mixtes et différencierées reflète l'état d'autorenouvellement de la population de cellules ES infectées.

1.3. Identification des gènes cibles du LIF, indépendants de STAT3.

Dans cette expérience d'analyse transcriptomique, la restimulation des cellules par le LIF induit l'activation de 258 gènes qui ne sont pas activés par le mutant hormono-dépendant STAT3-ER^{T2}. Nous avons vu que la stimulation du récepteur LIFR:gp130 active non seulement le facteur STAT3 mais aussi la voie de signalisation ERK via la phosphatase SHP-2 (Burdon et al., 1999), la voie PI3K (Paling et al., 2004) et la tyrosine kinase Yes (Anneren et al., 2004). La stimulation du récepteur LIFR:gp130 par le LIF induit donc vraisemblablement l'expression de nombreux gènes autres que les gènes cibles de STAT3. Nous nous sommes focalisés dans notre étude sur les gènes cibles de la voie LIF/gp130/STAT3.

2 - Les gènes cibles de STAT3 inhibent la différenciation des cellules ES.

2.1. Caractérisation préliminaire des gènes cibles grâce au test de formation de colonies (test d'autorenouvellement).

Parmi les 58 gènes cibles de STAT3 identifiés dans cette étude, nous en avons sélectionné 23 pour les analyses fonctionnelles. Cette sélection a été réalisée en premier lieu sur la base de leur profil d'expression au cours de la différenciation. Nous avons choisi des gènes dont le niveau d'expression diminue rapidement dès l'induction de la différenciation par formation de corps embryoides. Ces gènes sont en effet susceptibles de jouer un rôle clé dans le maintien des cellules ES à l'état indifférencié. Nous avons étudié le rôle de ces gènes grâce à une approche expérimentale originale associant l'infection des cellules par des vecteurs lentiviraux interférants et un test de formation de colonies (ou test d'autorenouvellement). Ce test repose sur l'analyse des capacités de croissance des cellules en conditions clonales. Les cellules ES cultivées à faible densité forment naturellement des colonies compactes exprimant la phosphatase alcaline, un marqueur des cellules indifférenciées (Figure 19). Une fraction (10-20%) forment des colonies dites mixtes, c'est-à-dire caractérisées par un centre indifférencié et une périphérie différenciée. Enfin, quelques cellules forment des colonies totalement différenciées n'exprimant pas la phosphatase alcaline. La proportion relative des trois types de colonies est un indicateur très sensible de l'état d'autorenouvellement d'une population de cellules ES.

L'inactivation partielle de l'expression des gènes cibles de STAT3 induit une augmentation de la proportion de colonies mixtes et différenciées dans le test

d'autorenouvellement. Notons que ce test est réalisé en présence de LIF afin de maintenir constante l'expression des autres gènes cibles. L'impact de l'inactivation de chacun des gènes cibles sur la différenciation est modéré. Nous supposons que c'est la somme des effets de tous les gènes cibles qui permet à STAT3 de maintenir les cellules ES en autorenouvellement. On peut légitimement spéculer que l'inactivation simultanée de plusieurs des gènes cibles de STAT3 aurait un effet beaucoup plus important sur la différenciation. Il serait intéressant de réaliser des inactivations multiples de gènes cibles appartenant à la même famille (*Pim1+Pim3*, *Klf4+Klf5*) pour réduire les redondances fonctionnelles entre gènes de la même famille.

Les 22 gènes cibles de STAT3 qui contribuent au maintien des cellules ES à l'état indifférencié codent pour des protéines ayant des rôles très variés. Ils codent pour des facteurs de transcription, des protéines de modification (sérine/thréonine kinases, sulfatase, déadénylase), des molécules d'adhésion, des protéines de la matrice extracellulaire et du cytosquelette ainsi qu'une protéine de régulation du stress cellulaire. Cette diversité suggère que le maintien de l'autorenouvellement des cellules ES de souris par la voie LIF/STAT3 passe par le contrôle de processus cellulaires très variés.

2.2. La kinase SGK (Serum and Glucocorticoid inducible Kinase).

Le gène *Sgk* code pour une sérine/thréonine kinase qui joue un rôle important dans la prolifération et la survie cellulaire (Brunet et al., 2001) (Buse et al., 1999). La kinase SGK est un effecteur de la PI3K (Kobayashi and Cohen, 1999) (Park et al., 1999). Or, nous avons vu que la voie PI3K, activé par la fixation du LIF sur son récepteur, joue aussi un rôle dans le maintien en autorenouvellement des cellules ES (Paling et al., 2004). Nous pouvons donc penser que la voie LIF/STAT3 et la voie LIF/PI3K fonctionnent en synergie pour activer la kinase SGK : (i) la voie LIF/STAT3 active la transcription du gène *sgk* ; (ii) la voie LIF/PI3K phosphoryle la kinase SGK (via la kinase PDK1) sur la thréonine 256 pour l'activer (Park et al., 1999).

Par ailleurs, la kinase SGK est capable d'inhiber la kinase Glycogen Synthase Kinase 3 (GSK-3) (Kobayashi and Cohen, 1999). L'inhibition de GSK-3 promeut l'autorenouvellement des cellules ES de souris (Ying et al., 2008) et d'Homme (Sato et al.,

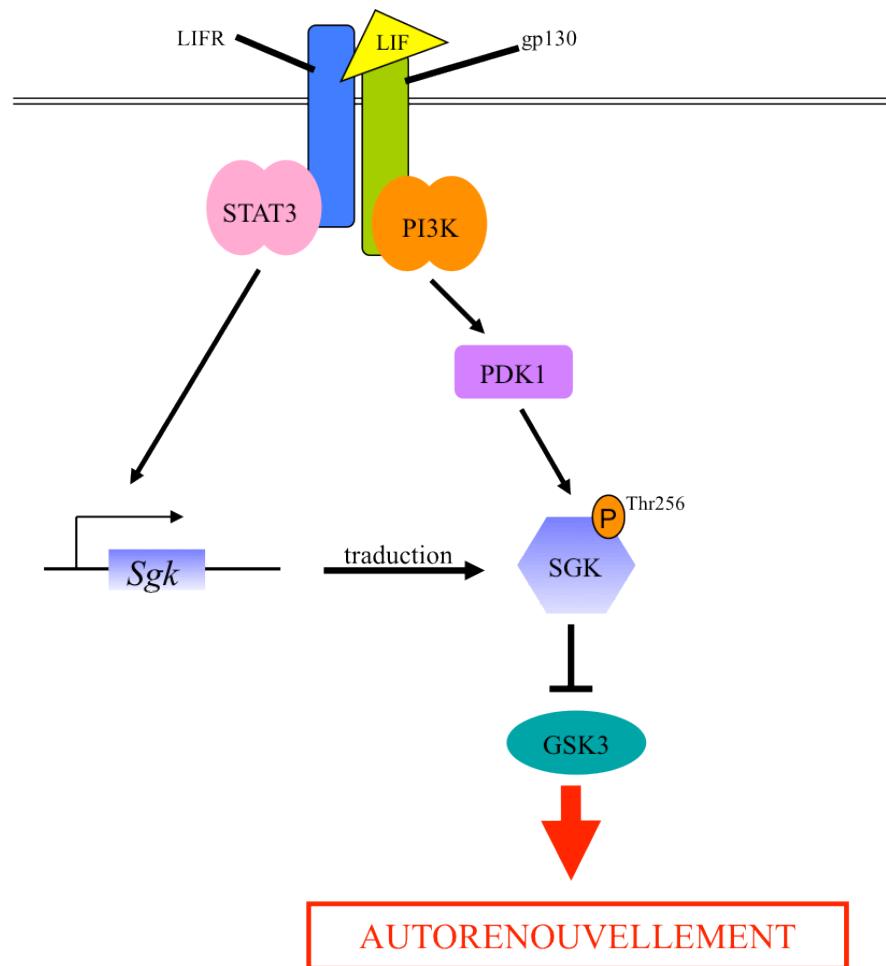


Figure 20 : Modèle de régulation de la kinase SGK par les voies LIF/STAT3 et LIF/PI3K

La voie LIF/STAT3 active l'expression du gène *Sgk*. Une fois traduit, la kinase SGK est phosphorylée par la voie LIF/PI3K (via la kinase PDK1) sur la thréonine 256 et est ainsi activée. Une fois activée, SGK inhibe la kinase GSK3 ce qui promeut l'autorenouvellement des cellules.

2004). Ainsi, l'activation transcriptionnelle de la kinase SGK par les voies LIF/STAT3 et LIF/PI3K pourrait promouvoir l'autorenouvellement des cellules ES de souris en inhibant la kinase GSK-3 (Figure 20).

2.3. Les sérine/thréonine kinases *Pim1* et *Pim3*.

2.3.1. Régulation de l'expression des kinases *Pim1* et *Pim3* par les facteurs STAT3 et Nanog.

Nous avons identifié les gènes *Pim1* et *Pim3* comme des gènes cibles de STAT3. Les régions régulatrices de *Pim1* et de *Pim3* comportent des sites de fixation de STAT3 (Shirogane et al., 1999) (Chen et al., 2008). Afin d'étudier plus précisément la régulation de *Pim1* et *Pim3* par STAT3, nous avons utilisé une lignée de cellules ES exprimant le récepteur chimérique GCSFR:gp130 muté sur les résidus tyrosine Y126-275F et ne pouvant ainsi plus fixer STAT3. Ce récepteur s'autodimérisé sous l'action du GCSF et active ainsi les autres voies de signalisation associées à gp130 dans les cellules ES. *Pim3* n'est pas activé par le GCSF indiquant que son expression est régulée exclusivement par STAT3. En revanche, *Pim1* reste activé en absence de recrutement de STAT3 indiquant que d'autres facteurs régulent son expression. Deux voies de signalisation impliquées dans le maintien en autorenouvellement des cellules ES peuvent activer l'expression de *Pim1* par le GCSF : la voie LIF/PI3 Kinase et la voie LIF/Yes (Paling et al., 2004) (Anneren et al., 2004). De plus, après privation des cellules de LIF pendant 24h le niveau d'expression de *Pim1* ne change pas et celui de *Pim3* ne diminue que de 50% ce qui suggère que d'autres facteurs, indépendants de la voie LIF, régulent leur expression. Le meilleur candidat semble être le facteur de transcription Nanog. En effet, des sites de fixation pour le facteur Nanog ont été identifiés dans les régions promotrices de *Pim1* et de *Pim3* (Loh et al. en 2005) (Chen et al. en 2008). Nous avons inactivé l'expression de Nanog par la technique d'interférence à l'ARN et montré que cela réduit les niveaux d'expression de *Pim1* et *Pim3*. Enfin, des sites de fixation pour un autre facteur majeur de la régulation de la pluripotence des cellules ES de souris, Oct4, ont été mis en évidence dans les séquences régulatrices de *Pim1* et de *Pim3* (Chen et al., 2008). Ainsi, l'activité résiduelle de l'expression de *Pim1* et *Pim3* après privation de LIF pourrait résulter de leur co-régulation par les deux facteurs de transcription Nanog et Oct4.

2.3.2. Les kinases *Pim1* et *Pim3* inhibent la différenciation des cellules ES.

L'inactivation de l'expression de *Pim1* et *Pim3* induit la différenciation des cellules ES alors que leur surexpression ralentit l'engagement des cellules dans la différenciation induite privation de LIF.

La différenciation induite par l'inhibition de l'expression de *Pim1* et *Pim3* se traduit par une augmentation du pourcentage de colonies mixtes et différenciées dans le test d'autorenouvellement ainsi que par la diminution de l'expression du marqueur de cellules ES indifférenciées *Rex1*. Un fort niveau d'interférence est primordial pour obtenir des effets plus drastiques. En effet, les résultats obtenus pour le test d'autorenouvellement à partir de clones sh*Pim1* et sh*Pim3* sont plus significatifs, de même que la diminution de l'expression dans ces cellules de *Nanog* et *Oct4*. Mais ce fort niveau d'interférence n'est cependant toujours pas suffisant pour obtenir des effets aussi drastiques que l'inhibition de *STAT3* ou de *Nanog*. La double inactivation *Pim1/Pim3* serait intéressante. En effet, elle permettrait d'annuler tout effet redondant dans la fonction de *Pim1* et de *Pim3* (Mikkers et al., 2004).

De même, dans les expériences de surexpression, même si *Pim1* et *Pim3* permettent aux cellules d'être plus résistantes à une privation de LIF, cela n'est pas suffisant pour permettre aux cellules de s'autorenouveler en l'absence complète de LIF. D'autres gènes cibles de *STAT3* tels que *Klf4* et *c-Myc*, jouent un rôle important dans l'inhibition de la différenciation des cellules ES de souris. La surexpression de *Klf4* retarde la différenciation induite par la formation de corps embryoïdes en l'absence de LIF, mais ces cellules se différencient malgré tout comme les cellules contrôles (Li et al., 2005). Quant à *c-Myc*, sa surexpression sous forme du mutant hormono-dépendant *Myc-ER* permet aux cellules de se maintenir temporairement en autorenouvellement en présence de 4'OHT et en absence de LIF (Cartwright et al., 2005). Ainsi, la surexpression individuelle de *Pim1*, *Pim3*, *Klf4* ou *c-Myc* ne permet pas de remplacer le LIF. Ici encore, la combinaison de plusieurs gènes est susceptible de donner de meilleurs résultats et de pallier ainsi à l'absence de LIF.

2.3.3. Les kinases *Pim1* et *Pim3* régulent la prolifération et l'apoptose dans les cellules ES.

Pim1 et *Pim3* semblent également avoir un rôle régulateur de la prolifération des cellules ES. Une mesure directe du temps de doublement de la population cellulaire est impossible car la présence des cellules différenciées, consécutive à l'inactivation de *Pim1* et *Pim3*, augmente le temps de doublement des cellules (Savatier et al., 1996). Nous avons donc mesuré la taille des colonies indifférenciées obtenues lors du test d'autorenouvellement dans les cellules où *Pim1* et *Pim3* ont été inhibés. Nous avons alors observé que l'inactivation de

Pim1 et Pim3 réduit la taille de ces colonies. Cette réduction de la taille des colonies peut résulter de deux phénomènes complémentaires : l'augmentation du taux d'apoptose et la réduction de la vitesse de prolifération.

Les kinases Pim1 et Pim3 sont connues pour leur rôle anti-apoptotique (Shirogane et al., 1999) (Adam et al., 2006). Elles exercent ce rôle en inactivant la protéine pro-apoptotique Bad (Aho et al., 2004) (Li et al., 2006). On sait par ailleurs que le LIF exerce une activité anti-apoptotique sur les cellules ES (Duval et al., 2000) (Duval et al., 2006). Nous avons montré que l'inactivation de Pim1 et Pim3 augmente le taux d'apoptose dans les cellules ES.

Nous n'avons pas étudié la vitesse de prolifération des cellules dans lesquelles l'expression de Pim1 et Pim3 a été atténuée. Nous pouvons néanmoins spéculer que l'inhibition de *Pim1* exerce un effet négatif sur le temps du cycle cellulaire. En effet, on sait que la kinase Pim1 est un régulateur positif de la phosphatase Cdc25A (Mochizuki et al., 1999), elle-même un régulateur positif des complexes cycline E :Cdk2 dont l'activité est nécessaire à la transition G1/S (Bartek and Lukas, 2001). De plus, Pim1 et Pim3 en phosphorylant les inhibiteurs de Kinases Dépendantes des Cyclines (CDK) p27^{kip1} et p21^{cip1}, favorisent la transition G1/S du cycle cellulaire en inhibant l'activité de ces inhibiteurs (Fujita et al., 2002) (Zhang et al., 2007). L'existence de ces régulations dans les cellules ES reste à démontrer.

2.4. *Klf4* et *Klf5*, en coopération avec *Sall4* : le lien entre la voie LIF/STAT3 et le réseau de régulation transcriptionnelle Nanog/Oct4/Sox2.

Nos résultats permettent d'établir pour la première fois un lien fonctionnel entre la voie LIF/STAT3 et le réseau de pluripotence Nanog/Oct4/Sox2.

En effet, nous avons montré que les gènes *Klf4* et *Klf5* sont des cibles de la voie LIF/STAT3. Par ailleurs, des expériences d'immunoprécipitation de chromatine indiquent que *Klf4* et *Klf5* sont de possibles régulateurs transcriptionnels de *Nanog*, formant avec ce dernier un réseau moléculaire régulant la pluripotence des cellules ES (Jiang et al., 2008). Parmi les gènes cibles de STAT3, nous avons également identifié celui codant pour le facteur de transcription *Sall4*. Le facteur SALL4 est un co-régulateur de la protéine NANOG avec lequel il forme des complexes promouvant ainsi son activation transcriptionnelle (Wu et al., 2006).

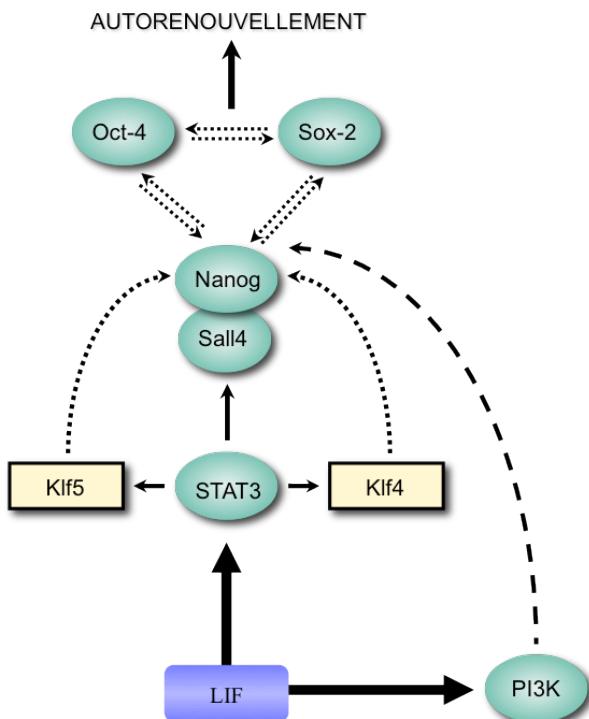


Figure 21 : Le LIF alimente le réseau de pluripotence Nanog/Oct4/Sox2

En activant STAT3 et la PI3-Kinase le LIF alimente le réseau de pluripotence Nanog/Oct4/Sox2.

L'activation de STAT3 par le LIF induit l'expression de *Klf4* et de *Klf5* qui vont à leur tour activer l'expression de *Nanog* (Jiang et al., 2008). *Klf4* et *Klf5* formeraient avec *Nanog* un réseau moléculaire régulant la pluripotence des cellules ES (Jiang et al., 2008). Un autre gène cible de STAT3, *Sall4* code pour un facteur de transcription décrit comme un co-régulateur de *Nanog* avec lequel il forme des complexes pour promouvoir son activité transcriptionnelle (Wu et al., 2006).

L'activation de la PI3-Kinase par le LIF contribue à l'activation transcriptionnelle de *Nanog* d'après les travaux réalisés par Storm et ses collègues (Storm et al., 2007).

En activant STAT3 et la PI3-Kinase le LIF alimenterait le réseau de pluripotence Nanog/Oct4/Sox2.

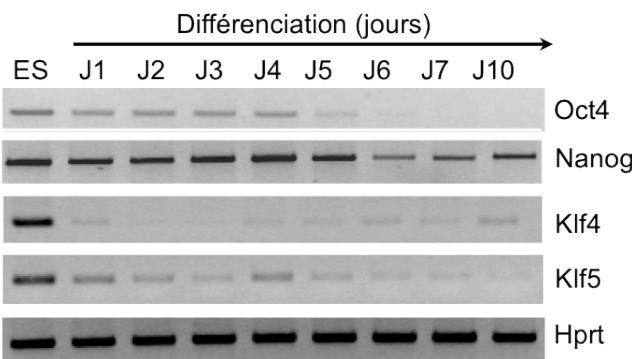


Figure 22 : Profil d'expression de *Oct4*, *Nanog*, *Klf4* et *Klf5* au cours de la différenciation

Analyse par RT-PCR semi-quantitative du niveau d'expression de *Oct4*, *Nanog*, *Klf4*, *Klf5* et *Hprt* (contrôle) dans les cellules ES induites à se différencier par formation de corps embryoides. Les analyses ont été réalisées sur les cellules indifférenciées (ES) et les cellules aux jours 1, 2, 3, 4, 5, 6, 7 et 10 de différenciation.

L'inactivation du gène *Sall4* induit la différenciation des cellules ES (Zhang et al., 2006). SALL4 constitue donc un autre lien entre la voie LIF/STAT3 et le réseau de pluripotence Nanog/Oct4/Sox2. Enfin, nous avons montré au laboratoire, en collaboration avec Mélanie Welham (Université de Bath, UK) (Storm et al., 2007), que la PI3K contribue à l'activation transcriptionnelle de Nanog. En conclusion, au moins trois mécanismes indépendants contribueraient à alimenter le réseau de pluripotence en réponse à la stimulation par le LIF (Figure 21).

Ce modèle de régulation prédit l'atténuation de l'expression de Nanog après privation de LIF et sa réactivation après restimulation. Il prédit également l'extinction de son expression parallèlement à *Klf4* et *Klf5* lors de la formation de corps embryoïdes. Aucune de ces deux prédictions ne sont vérifiées par l'expérience. En effet, le gène *Nanog* n'est pas apparu comme un gène cible de la voie LIF/STAT3 dans nos expériences. Par ailleurs, lors de la différenciation induite par la formation de corps embryoïdes, les gènes *Klf4* et *Klf5* s'éteignent dès le premier jour de différenciation alors que *Nanog* s'éteint au 4^{ème} jour, parallèlement à *Oct4* et *Sox2* (Figure 22). Ces observations remettent-elles en cause notre modèle ? Sans doute non. En effet, les gènes *Nanog*, *Oct4* et *Sox2* possèdent des sites de fixation pour les trois facteurs Nanog, Oct4 et Sox2 dans leurs séquences régulatrices, de sorte que ces trois gènes s'autorégulent et s'activent mutuellement. Dans ces conditions, on peut imaginer que le réseau de pluripotence constitué par ces trois gènes possède la capacité de s'auto-entretenir pendant 2 à 3 jours après la privation de LIF, expliquant ainsi le délai de 3 jours entre la privation de LIF et l'extinction de *Nanog*. Selon notre modèle, Le LIF alimenterait ce réseau de façon ponctuelle, maintenant ainsi le niveau d'expression de *Nanog* au-dessus d'un seuil critique.

3 - Les gènes cibles de STAT3 inhibent la différenciation des cellules ES de souris vers le lignage endodermique et/ou mésodermique.

3.1. Les gènes cibles de STAT3 constituent trois familles distinctes inhibant la différenciation vers le mésoderme et/ou vers l'endoderme

La voie de signalisation LIF/STAT3 permet le maintien en autorenouvellement des cellules ES en inhibant leur différenciation en mésoderme et en endoderme (Niwa et al., 1998) (Ying et al., 2003). L'analyse des marqueurs de différenciation exprimés dans les cellules ES après inactivation partielle de l'expression des gènes cibles a révélé des effets

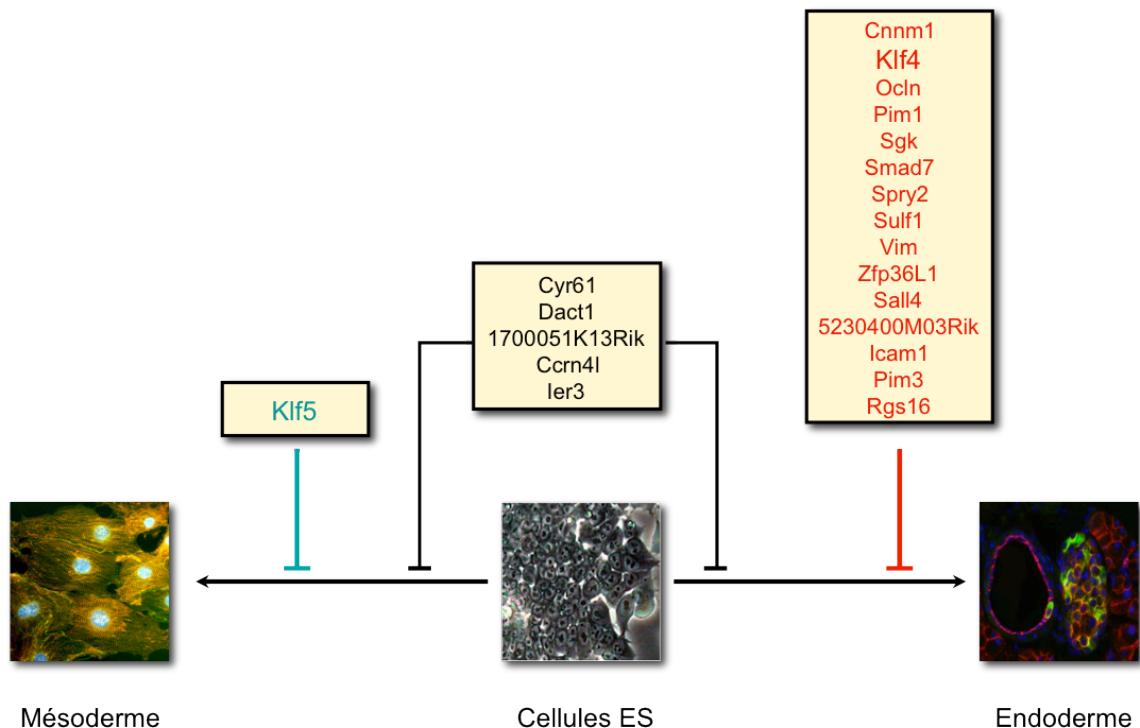


Figure 23 : Les trois familles de gènes cibles de STAT3

L'analyse des marqueurs de différenciation exprimés dans les cellules ES après inactivation partielle de l'expression des gènes cibles a permis d'identifier trois familles distinctes de gènes cibles de STAT3. La première famille est composé de 16 gènes (*Cnnm1*, *Klf4*, *Ocln*, *Pim1*, *Sgk*, *Smad7*, *Spry2*, *Sulf1*, *Vim*, *Zfp36L1*, *Sall4*, *5230400M03Rik*, *Icam1*, *Pim3* et *Rgs16*) qui sont impliqués dans l'inhibition de la différenciation vers l'endoderme, la deuxième famille ne compte qu'un seul membre, *Klf5*, impliqué dans l'inhibition de la différenciation vers le mésoderme, et la troisième famille composée de *Cyr61*, *Dact1*, *1700051K13Rik*, *Ccrn4l* et *Ier3* est constituée de gènes inhibant la différenciation en endoderme et en mésoderme.

inattendus. La plupart des gènes cibles de STAT3 sont impliqués dans l'inhibition de la différenciation vers l'endoderme, un seul est impliqué dans l'inhibition de la différenciation vers le mésoderme, quelques uns ne montrant aucune spécificité d'action vers l'un ou l'autre de ces deux lignages (Figure 23).

Un seul gène, *Klf5*, inhibe la différenciation vers le mésoderme de manière spécifique. Notons que parmi les 58 gènes cibles de STAT3 identifiés, nous en avons sélectionné 23 pour réaliser les analyses fonctionnelles par interférence à l'ARN. Il serait intéressant d'étudier de la même manière les 35 gènes restants afin d'identifier d'autres régulateurs de la différenciation mésodermique.

3.2. *Klf4* et *Klf5* exercent des effets opposés sur la différenciation en mésoderme et en endoderme.

Nous nous sommes intéressés aux gènes *Klf4* et *Klf5* qui, bien qu'appartenant à la même famille, semblent avoir des fonctions complémentaires dans le maintien de l'état indifférencié des cellules ES. Qu'il s'agisse de la différenciation spontanée induite uniquement par interférence à l'ARN en présence de LIF, ou de la différenciation induite par la formation de corps embryoïdes, les résultats obtenus démontrent que ces deux gènes exercent des effets radicalement différents. *Klf4* inhibe la différenciation en endoderme, alors que *Klf5* inhibe la différenciation en mésoderme. Cette conclusion est obtenue après analyse de l'expression de marqueurs de différenciation spécifiques de l'un ou de l'autre de ces deux lignages. L'utilisation de lignées exprimant un gène marqueur fluorescent, placé sous le contrôle des promoteurs des gènes *T* (mésoderme) et *Sox17* (endoderme), nous a permis de démontrer que l'inactivation partielle des gènes *Klf4* et *Klf5* modifie l'équilibre entre les deux lignages. Ainsi, l'inactivation de l'expression de *Klf4* augmente le nombre de cellules qui s'engagent vers l'endoderme, alors que l'inactivation partielle de *Klf5* augmente le nombre de cellules qui s'engagent vers le mésoderme. Il est intéressant de noter que l'augmentation du nombre de cellules se différenciant vers le mésoderme ne se fait pas aux détriment de l'endoderme. On peut émettre l'hypothèse selon laquelle cette augmentation se fait au détriment des cellules se différenciant en ectoderme.

L'ensemble de ces résultats suggère que *Klf4* et *Klf5* maintiennent les cellules ES en autorenouvellement en exerçant des rôles opposés : le premier en inhibant leur différenciation en endoderme et le second en inhibant leur différenciation en mésoderme. Des fonctions opposées pour *Klf4* et *Klf5* ont également été décrit dans d'autres mécanismes cellulaires

comme (i) la régulation transcriptionnelle : les deux facteurs se fixent sur les mêmes éléments de régulation mais Klf4 active l'expression de certains gènes alors que Klf5 inhibe l'expression de ces mêmes gènes (Dang et al., 2002) (Liu et al., 2003) (Piccinni et al., 2004) ; (ii) la prolifération : des études ont montré que le niveau d'expression de *Klf4* augmente lorsque les cellules rentrent en quiescence (induite par la privation de sérum). Au contraire, l'expression de *Klf5* augmente lorsque les cellules quiescentes sont stimulées par le sérum pour réentrer dans le cycle cellulaire (Shields et al., 1996) ; (iii) la transformation tumorale : *Klf4* a été décrit comme un suppresseur de tumeurs, notamment dans les cancers gastriques et colorectaux (Zhao et al., 2004), tandis que *Klf5* peut se comporter comme un oncogène dans ces mêmes cellules (Nandan et al., 2004). On peut s'interroger sur la signification de ces fonctions opposées dans la différenciation vers le mésoderme et l'endoderme.

3.3. *Klf4* et *Klf5* inhibent l'expression de régulateurs de la différenciation vers le mésoderme et l'endoderme.

Afin de comprendre comment les facteurs Klf4 et Klf5 inhibent la différenciation vers le mésoderme et l'endoderme, nous avons analysé le transcriptome des cellules ES dans lesquelles l'expression de *Klf4* et *Klf5* est diminuée de 50%. Nous avons ainsi montré que *Klf4* inhibe l'expression de plusieurs gènes contrôlant la différenciation vers l'endoderme, en particulier les gènes *GATA-4*, *GATA-6*, *Sox7*, *Sox17*, *FoxA2* et *Dab2*. De nombreuses données expérimentales indiquent que ces 6 gènes sont impliqués dans la différenciation des cellules souches pluripotentes en endoderme : (i) la surexpression de *GATA-4* et *GATA-6* induit la différenciation des cellules ES en endoderme extra-embryonnaire (Fujikura et al., 2002) (Shimosato et al., 2007). De plus, dans des souris invalidées pour le gène *GATA-6*, l'endoderme viscéral (extra-embryonnaire) est absent et les embryons meurent à E7.5 (Morrisey et al., 1998). Les cellules ES *GATA-4* -/- sont incapables de se différencier en endoderme viscéral (Soudais et al., 1995) ; (ii) la surexpression de *Sox7* induit la différenciation des cellules souches de carcinomes embryonnaires (EC) et des cellules ES humaines en endoderme extra-embryonnaire (Futaki et al., 2004) (Seguin et al., 2008) ; (iii) la surexpression de *Sox17* induit la différenciation des cellules ES de souris et d'Homme en endoderme définitif (endoderme intestinal) (Seguin et al., 2008) (Qu et al., 2008) ; (iv) le facteur de transcription *FoxA2* est nécessaire à la formation de l'endoderme extra-embryonnaire mais également à la formation d'une partie de l'endoderme définitif dans

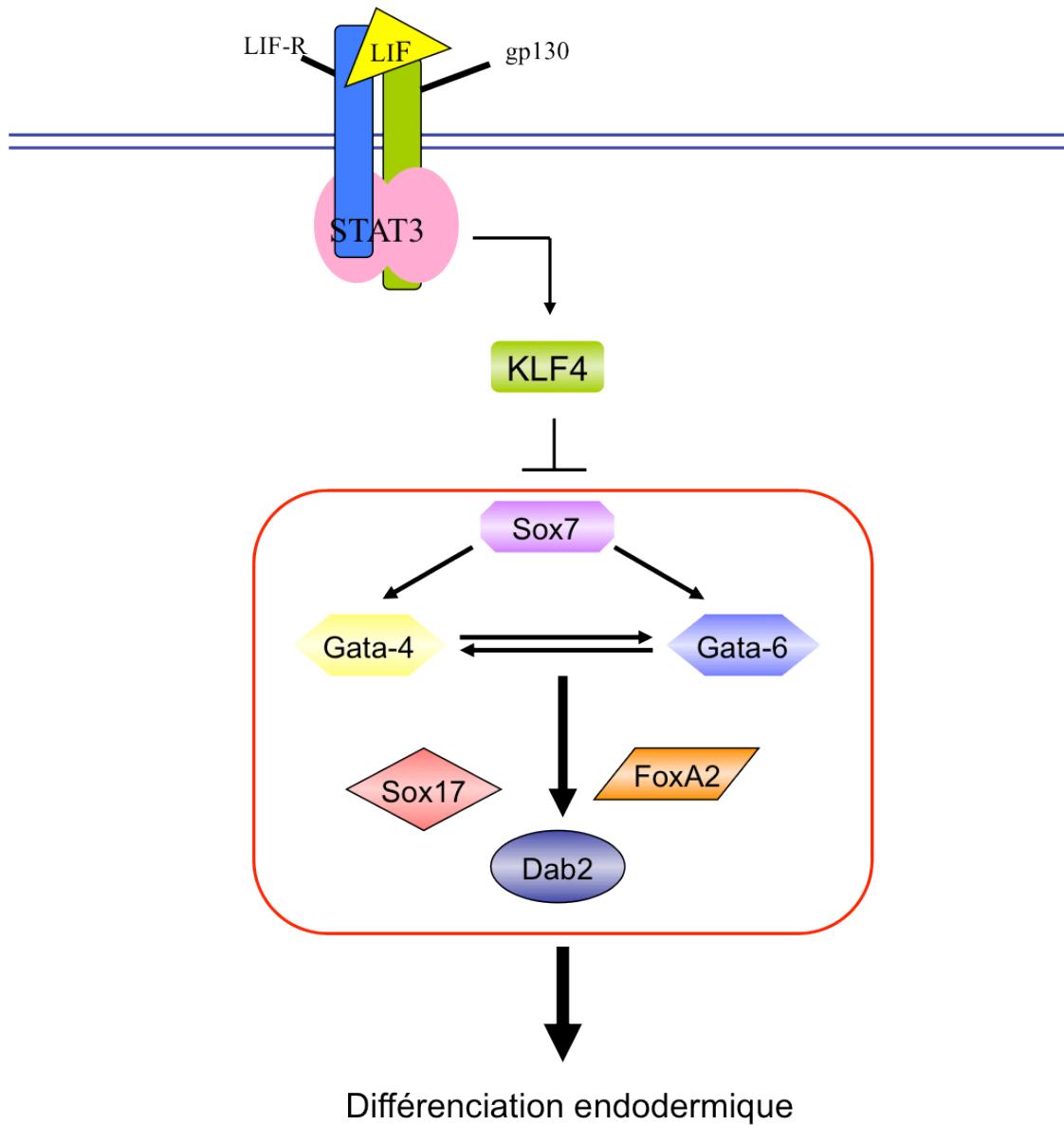


Figure 24 : Modèle d'action de Klf4 pour inhiber la différenciation des cellules ES en endoderme

Klf4 pourrait inhiber la différenciation des cellules ES en inactivant une cascade de régulation génique aboutissant à la différenciation des cellules pluripotentes en endoderme. Klf4 inhiberait ainsi l'expression de *Sox7* dans un premier temps ce qui conduirait à l'inhibition de *Gata-4* et *Gata-6* (gènes cibles de *Sox7*). Une fois inactivées, *Gata-4* et *Gata-6*, ne pourraient plus induire l'expression de *Sox17*, *Dab2* et *FoxA2* (gènes cibles de *Gata-6*). Ainsi, en inhibant cette cascade de régulation génique, Klf4 inhiberait la différenciation des cellules ES en endoderme.

l'embryon (Dufort et al., 1998). D'autres études ont montré que ce facteur est nécessaire à la formation de nombreux organes qui dérivent du feuillet endodermique comme le foie, le pancréas ou les poumons (Friedman and Kaestner, 2006). De plus, la surexpression de *FoxA2* dans des cellules ES de souris induites à se différencier par formation de corps embryoides active l'expression de marqueurs précoce de la différenciation endodermique (Levinson-Dushnik and Benvenisty, 1997) ; (v) le gène *Dab2* est essentiel à la formation de l'endoderme extra-embryonnaire viscéral *in vivo* (Yang et al., 2002) (Morris et al., 2002).

Ces six gènes forment une cascade de régulation. Dans les cellules souches de tératocarcinomes F9, *Sox7* active l'expression de *GATA-4* et *GATA-6*, et *GATA-4* et *GATA-6* activent l'expression de *Sox17*, *FoxA2* et *Dab2* (Futaki et al., 2004) (Morrisey et al., 2000). Ainsi, *Klf4* semble capable d'inhiber une cascade de régulation génique aboutissant à la différenciation des cellules pluripotentes en endoderme (Figure 24).

Les analyses de transcriptomes nous ont également permis d'identifier des gènes dont l'expression est inhibée à la fois par *Klf4* et par *Klf5*. L'un de ces gènes, *Nodal* joue un rôle essentiel dans la formation des feuillets embryonnaires chez la souris (Schier, 2003) (Zhou et al., 1993) (Conlon et al., 1994). De plus, la surexpression de *Nodal* induit la différenciation des cellules ES en mésoderme et en endoderme définitif (Takenaga et al., 2007). La régulation négative de *Nodal* par *Klf4* et *Klf5* suggère que ces deux facteurs pourraient, en plus de leur rôle spécifique, avoir une fonction commune dans les étapes précoce de la différenciation endomésodermique.

3.4. L'inactivation de *Klf4* induit-elle la différenciation en endoderme extra-embryonnaire ou en endoderme définitif ?

Les cellules ES sont capables de se différencier *in vitro* en endoderme extra-embryonnaire (endoderme viscéral et pariétal) et en endoderme embryonnaire, appelé également endoderme définitif, à l'origine des organes du système digestif et de la glande thyroïde. *Klf4* inhibe-t'il la différenciation dans les deux lignages endodermiques ? Nous n'avons pas pu répondre clairement à cette question car la plupart des marqueurs endodermiques analysés sont exprimés à la fois dans l'endoderme extra-embryonnaire et dans l'endoderme définitif. La distinction entre les deux types d'endoderme est donc difficile. *Cxcr4* est décrit comme un marqueur spécifique de l'endoderme définitif (McGrath et al., 1999) (Yasunaga et al., 2005). Son activation dans les cellules ES après inactivation partielle de l'expression du gène *Klf4* suggère donc que le facteur *Klf4* est capable d'inhiber la

différenciation en endoderme définitif. Pour confirmer cette hypothèse, il conviendrait d'étudier les effets de l'inactivation du gène *Klf4* dans la lignée de cellules ES *Gsc-GFP/Sox17-CD25* décrite par Yasunaga et ses collègues en 2005. Cette lignée comporte un « knock-in » de GFP dans le locus *goosecoid* et un « knock-in » de CD25 dans le locus *Sox17*. Les cellules de l'endoderme définitif sont *Gsc*⁺/*Sox17*⁺, alors que les cellules de l'endoderme viscéral sont *Gsc*⁻/*Sox17*⁺. Une analyse en cytométrie en flux après inactivation de *Klf4* permettrait de déterminer le rôle de *Klf4* dans la différenciation des deux catégories d'endoderme.

Aucun marqueur ne semble capable de discriminer positivement l'endoderme extra-embryonnaire, qu'il soit viscéral ou pariétal. Il est néanmoins vraisemblable que la différenciation en endoderme viscéral et pariétal est inhibée par *Klf4*. En effet, la surexpression de *GATA-4* et *GATA-6* induit la différenciation des cellules ES en endoderme extra-embryonnaire (Fujikura et al., 2002) (Shimosato et al., 2007). La surexpression de *Sox7* dans les cellules souches de carcinomes embryonnaires F9 induit leur différenciation en endoderme pariétal (Futaki et al., 2004). Enfin, l'expression constitutive de *Sox7* dans des cellules ES humaines induit la formation de progéniteurs de l'endoderme extra-embryonnaire (Seguin et al., 2008). Or, nous avons vu précédemment que ces trois gènes sont des cibles du facteur *Klf4* dans les cellules ES de souris.

CONCLUSION GÉNÉRALE

Les résultats obtenus dans cette étude nous permettent de mieux comprendre comment la voie de signalisation LIF/STAT3 contrôle l'autorenouvellement des cellules ES de souris et permet leur maintien à l'état indifférencié. C'est, à notre connaissance, la première fois que des gènes cibles d'une voie de signalisation sont montrés impliqués dans l'inhibition de la différenciation, soit vers le mésoderme (*Klf5*), soit vers l'endoderme (*Klf4*, *Pim1*, *Pim3*). Ces résultats soulèvent trois interrogations :

1/ les mécanismes identifiés dans les cellules ES de souris opèrent-ils également *in vivo* lors de la différenciation de l'épiblaste et la formation des feuillets embryonnaires ?

2/ les mécanismes identifiés dans les cellules ES de souris opèrent-ils également dans les cellules ES humaines ? Cette question nous semble particulièrement pertinente dans la mesure où les cellules ES de primates, humain et non humains, ne requièrent pas la voie LIF/STAT3 pour s'autorenouveler (Daheron et al., 2004) (Humphrey et al., 2004) (Sumi et al., 2004). En l'absence de cette voie fonctionnelle, quelle autre voie de signalisation active les gènes cibles de STAT3 en général, les gènes *Klf4* et *Klf5* en particulier ?

3/ les mécanismes que nous avons identifiés peuvent-ils servir à orienter la différenciation des cellules ES en cellules d'intérêt thérapeutique ? Il conviendra d'étudier la capacité des cellules dont l'expression de *Klf5* a été atténuée par interférence à l'ARN à se différencier en cellules dérivées du mésoderme (cardiomyocytes, cellules hématopoïétiques). De manière similaire, on pourra étudier l'effet de l'inactivation partielle de *Klf4* sur la capacité des cellules ES à se différencier en cellules dérivées de l'endoderme définitif (cellules endocrines β , hépatocytes).

Par ailleurs, notre étude suggère que la régulation de la pluripotence et la détermination du lignage sont des mécanismes étroitement liés. Une meilleure compréhension de ces mécanismes devrait permettre, à terme, de contrôler avec précision l'engagement des cellules souches pluripotentes (ES et iPS) dans l'ectoderme, le mésoderme, l'endoderme et le lignage germinal.

RÉFÉRENCES

BIBLIOGRAPHIQUES

- Adam, M., Pogacic, V., Bendit, M., Chappuis, R., Nawijn, M. C., Duyster, J., Fox, C. J., Thompson, C. B., Cools, J. and Schwaller, J.** (2006). Targeting PIM kinases impairs survival of hematopoietic cells transformed by kinase inhibitor-sensitive and kinase inhibitor-resistant forms of Fms-like tyrosine kinase 3 and BCR/ABL. *Cancer Res* **66**, 3828-35.
- Aho, T. L., Sandholm, J., Peltola, K. J., Mankonen, H. P., Lilly, M. and Koskinen, P. J.** (2004). Pim-1 kinase promotes inactivation of the pro-apoptotic Bad protein by phosphorylating it on the Ser112 gatekeeper site. *FEBS Lett* **571**, 43-9.
- Allen, J. D. and Berns, A.** (1996). Complementation tagging of cooperating oncogenes in knockout mice. *Semin Cancer Biol* **7**, 299-306.
- Allen, J. D., Verhoeven, E., Domen, J., van der Valk, M. and Berns, A.** (1997). Pim-2 transgene induces lymphoid tumors, exhibiting potent synergy with c-myc. *Oncogene* **15**, 1133-41.
- Amaravadi, R. and Thompson, C. B.** (2005). The survival kinases Akt and Pim as potential pharmacological targets. *J Clin Invest* **115**, 2618-24.
- Anderson, K. P., Kern, C. B., Crable, S. C. and Lingrel, J. B.** (1995). Isolation of a gene encoding a functional zinc finger protein homologous to erythroid Kruppel-like factor: identification of a new multigene family. *Mol Cell Biol* **15**, 5957-65.
- Anneren, C., Cowan, C. A. and Melton, D. A.** (2004). The Src family of tyrosine kinases is important for embryonic stem cell self-renewal. *J Biol Chem* **279**, 31590-8.
- Askew, D. S., Ashmun, R. A., Simmons, B. C. and Cleveland, J. L.** (1991). Constitutive c-myc expression in an IL-3-dependent myeloid cell line suppresses cell cycle arrest and accelerates apoptosis. *Oncogene* **6**, 1915-22.
- Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N. and Lovell-Badge, R.** (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* **17**, 126-40.
- Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H. F., John, R. M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M. et al.** (2006). Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* **8**, 532-8.
- Bachmann, M., Kosan, C., Xing, P. X., Montenarh, M., Hoffmann, I. and Moroy, T.** (2006). The oncogenic serine/threonine kinase Pim-1 directly phosphorylates and activates the G2/M specific phosphatase Cdc25C. *Int J Biochem Cell Biol* **38**, 430-43.
- Bachmann, M. and Moroy, T.** (2005). The serine/threonine kinase Pim-1. *Int J Biochem Cell Biol* **37**, 726-30.
- Bartek, J. and Lukas, J.** (2001). Mammalian G1- and S-phase checkpoints in response to DNA damage. *Curr Opin Cell Biol* **13**, 738-47.
- Baytel, D., Shalom, S., Madgar, I., Weissenberg, R. and Don, J.** (1998). The human Pim-2 proto-oncogene and its testicular expression. *Biochim Biophys Acta* **1442**, 274-85.
- Bazan, J. F.** (1990). Structural design and molecular evolution of a cytokine receptor superfamily. *Proc Natl Acad Sci U S A* **87**, 6934-8.
- Bissonnette, R. P., Echeverri, F., Mahboubi, A. and Green, D. R.** (1992). Apoptotic cell death induced by c-myc is inhibited by bcl-2. *Nature* **359**, 552-4.
- Blake, R. A., Broome, M. A., Liu, X., Wu, J., Gishizky, M., Sun, L. and Courtneidge, S. A.** (2000). SU6656, a selective src family kinase inhibitor, used to probe growth factor signaling. *Mol Cell Biol* **20**, 9018-27.
- Boeuf, H., Hauss, C., Graeve, F. D., Baran, N. and Kedinger, C.** (1997). Leukemia inhibitory factor-dependent transcriptional activation in embryonic stem cells. *J Cell Biol* **138**, 1207-17.
- Borg, K. E., Zhang, M., Hegge, D., Stephen, R. L., Buckley, D. J., Magnuson, N. S. and Buckley, A. R.** (1999). Prolactin regulation of pim-1 expression: positive and negative promoter elements. *Endocrinology* **140**, 5659-68.

- Borzillo, G. V., Endo, K. and Tsujimoto, Y.** (1992). Bcl-2 confers growth and survival advantage to interleukin 7-dependent early pre-B cells which become factor independent by a multistep process in culture. *Oncogene* **7**, 869-76.
- Bowman, T., Garcia, R., Turkson, J. and Jove, R.** (2000). STATs in oncogenesis. *Oncogene* **19**, 2474-88.
- Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G. et al.** (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-56.
- Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K. et al.** (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349-53.
- Bradley, A., Evans, M., Kaufman, M. H. and Robertson, E.** (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309**, 255-6.
- Brinster, R. L.** (1974). The effect of cells transferred into the mouse blastocyst on subsequent development. *J Exp Med* **140**, 1049-56.
- Brook, F. A. and Gardner, R. L.** (1997). The origin and efficient derivation of embryonic stem cells in the mouse. *Proc Natl Acad Sci U S A* **94**, 5709-12.
- Brown, M. T. and Cooper, J. A.** (1996). Regulation, substrates and functions of src. *Biochim Biophys Acta* **1287**, 121-49.
- Brunet, A., Bonni, A., Zigmond, M. J., Lin, M. Z., Juo, P., Hu, L. S., Anderson, M. J., Arden, K. C., Blenis, J. and Greenberg, M. E.** (1999). Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell* **96**, 857-68.
- Brunet, A., Park, J., Tran, H., Hu, L. S., Hemmings, B. A. and Greenberg, M. E.** (2001). Protein kinase SGK mediates survival signals by phosphorylating the forkhead transcription factor FKHLR1 (FOXO3a). *Mol Cell Biol* **21**, 952-65.
- Burdon, T., Chambers, I., Stracey, C., Niwa, H. and Smith, A.** (1999a). Signaling mechanisms regulating self-renewal and differentiation of pluripotent embryonic stem cells. *Cells Tissues Organs* **165**, 131-43.
- Burdon, T., Smith, A. and Savatier, P.** (2002). Signalling, cell cycle and pluripotency in embryonic stem cells. *Trends Cell Biol* **12**, 432-8.
- Burdon, T., Stracey, C., Chambers, I., Nichols, J. and Smith, A.** (1999b). Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. *Dev Biol* **210**, 30-43.
- Buse, P., Tran, S. H., Luther, E., Phu, P. T., Aponte, G. W. and Firestone, G. L.** (1999). Cell cycle and hormonal control of nuclear-cytoplasmic localization of the serum- and glucocorticoid-inducible protein kinase, Sgk, in mammary tumor cells. A novel convergence point of anti-proliferative and proliferative cell signaling pathways. *J Biol Chem* **274**, 7253-63.
- Cartwright, P., McLean, C., Sheppard, A., Rivett, D., Jones, K. and Dalton, S.** (2005). LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development* **132**, 885-96.
- Catena, R., Tiveron, C., Ronchi, A., Porta, S., Ferri, A., Tatangelo, L., Cavallaro, M., Favaro, R., Ottolenghi, S., Reinbold, R. et al.** (2004). Conserved POU binding DNA sites in the Sox2 upstream enhancer regulate gene expression in embryonic and neural stem cells. *J Biol Chem* **279**, 41846-57.
- Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S. and Smith, A.** (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643-55.

- Chanchevalap, S., Nandan, M. O., Merlin, D. and Yang, V. W.** (2004). All-trans retinoic acid inhibits proliferation of intestinal epithelial cells by inhibiting expression of the gene encoding Kruppel-like factor 5. *FEBS Lett* **578**, 99-105.
- Chen, X., Whitney, E. M., Gao, S. Y. and Yang, V. W.** (2003). Transcriptional profiling of Kruppel-like factor 4 reveals a function in cell cycle regulation and epithelial differentiation. *J Mol Biol* **326**, 665-77.
- Classon, M., Henriksson, M., Sumegi, J., Klein, G. and Hammarskjold, M. L.** (1987). Elevated c-myc expression facilitates the replication of SV40 DNA in human lymphoma cells. *Nature* **330**, 272-4.
- Conkright, M. D., Wani, M. A., Anderson, K. P. and Lingrel, J. B.** (1999). A gene encoding an intestinal-enriched member of the Kruppel-like factor family expressed in intestinal epithelial cells. *Nucleic Acids Res* **27**, 1263-70.
- Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B. and Robertson, E. J.** (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development* **120**, 1919-28.
- Conover, J. C., Ip, N. Y., Poueymirou, W. T., Bates, B., Goldfarb, M. P., DeChiara, T. M. and Yancopoulos, G. D.** (1993). Ciliary neurotrophic factor maintains the pluripotentiality of embryonic stem cells. *Development* **119**, 559-65.
- Copeland, N. G., Gilbert, D. J., Schindler, C., Zhong, Z., Wen, Z., Darnell, J. E., Jr., Mui, A. L., Miyajima, A., Quelle, F. W., Ihle, J. N. et al.** (1995). Distribution of the mammalian Stat gene family in mouse chromosomes. *Genomics* **29**, 225-8.
- Daheron, L., Opitz, S. L., Zaehres, H., Lensch, M. W., Andrews, P. W., Itsikovitz-Eldor, J. and Daley, G. Q.** (2004). LIF/STAT3 signaling fails to maintain self-renewal of human embryonic stem cells. *Stem Cells* **22**, 770-8.
- Dang, D. T., Pevsner, J. and Yang, V. W.** (2000). The biology of the mammalian Kruppel-like family of transcription factors. *Int J Biochem Cell Biol* **32**, 1103-21.
- Dang, D. T., Zhao, W., Mahatan, C. S., Geiman, D. E. and Yang, V. W.** (2002). Opposing effects of Kruppel-like factor 4 (gut-enriched Kruppel-like factor) and Kruppel-like factor 5 (intestinal-enriched Kruppel-like factor) on the promoter of the Kruppel-like factor 4 gene. *Nucleic Acids Res* **30**, 2736-41.
- Darnell, J. E., Jr.** (1997). STATs and gene regulation. *Science* **277**, 1630-5.
- Dautry, F., Weil, D., Yu, J. and Dautry-Varsat, A.** (1988). Regulation of pim and myb mRNA accumulation by interleukin 2 and interleukin 3 in murine hematopoietic cell lines. *J Biol Chem* **263**, 17615-20.
- Davis, S., Aldrich, T. H., Stahl, N., Pan, L., Taga, T., Kishimoto, T., Ip, N. Y. and Yancopoulos, G. D.** (1993). LIFR beta and gp130 as heterodimerizing signal transducers of the tripartite CNTF receptor. *Science* **260**, 1805-8.
- Diwan, S. B. and Stevens, L. C.** (1976). Development of teratomas from the ectoderm of mouse egg cylinders. *J Natl Cancer Inst* **57**, 937-42.
- Doetschman, T., Williams, P. and Maeda, N.** (1988). Establishment of hamster blastocyst-derived embryonic stem (ES) cells. *Dev Biol* **127**, 224-7.
- Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W. and Kemler, R.** (1985). The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morphol* **87**, 27-45.
- Dufort, D., Schwartz, L., Harpal, K. and Rossant, J.** (1998). The transcription factor HNF3beta is required in visceral endoderm for normal primitive streak morphogenesis. *Development* **125**, 3015-25.
- Durbin, J. E., Hackenmiller, R., Simon, M. C. and Levy, D. E.** (1996). Targeted disruption of the mouse Stat1 gene results in compromised innate immunity to viral disease. *Cell* **84**, 443-50.

- Duval, D., Reinhardt, B., Kedinger, C. and Boeuf, H.** (2000). Role of suppressors of cytokine signaling (Socs) in leukemia inhibitory factor (LIF) -dependent embryonic stem cell survival. *Faseb J* **14**, 1577-84.
- Duval, D., Trouillas, M., Thibault, C., Dembele, D., Diemunsch, F., Reinhardt, B., Mertz, A. L., Dierich, A. and Boeuf, H.** (2006). Apoptosis and differentiation commitment: novel insights revealed by gene profiling studies in mouse embryonic stem cells. *Cell Death Differ* **13**, 564-75.
- Eichmann, A., Yuan, L., Breant, C., Alitalo, K. and Koskinen, P. J.** (2000). Developmental expression of pim kinases suggests functions also outside of the hematopoietic system. *Oncogene* **19**, 1215-24.
- Ernst, M., Gearing, D. P. and Dunn, A. R.** (1994). Functional and biochemical association of Hck with the LIF/IL-6 receptor signal transducing subunit gp130 in embryonic stem cells. *Embo J* **13**, 1574-84.
- Ernst, M., Novak, U., Nicholson, S. E., Layton, J. E. and Dunn, A. R.** (1999). The carboxyl-terminal domains of gp130-related cytokine receptors are necessary for suppressing embryonic stem cell differentiation. Involvement of STAT3. *J Biol Chem* **274**, 9729-37.
- Ernst, M., Oates, A. and Dunn, A. R.** (1996). Gp130-mediated signal transduction in embryonic stem cells involves activation of Jak and Ras/mitogen-activated protein kinase pathways. *J Biol Chem* **271**, 30136-43.
- Evan, G. I., Wyllie, A. H., Gilbert, C. S., Littlewood, T. D., Land, H., Brooks, M., Waters, C. M., Penn, L. Z. and Hancock, D. C.** (1992). Induction of apoptosis in fibroblasts by c-myc protein. *Cell* **69**, 119-28.
- Evans, M. J. and Kaufman, M. H.** (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-6.
- Feldman, J. D., Vician, L., Crispino, M., Tocco, G., Baudry, M. and Herschman, H. R.** (1998). Seizure activity induces PIM-1 expression in brain. *J Neurosci Res* **53**, 502-9.
- Field, S. J., Johnson, R. S., Mortensen, R. M., Papaioannou, V. E., Spiegelman, B. M. and Greenberg, M. E.** (1992). Growth and differentiation of embryonic stem cells that lack an intact c-fos gene. *Proc Natl Acad Sci U S A* **89**, 9306-10.
- Forrai, A., Boyle, K., Hart, A. H., Hartley, L., Rakar, S., Willson, T. A., Simpson, K. M., Roberts, A. W., Alexander, W. S., Voss, A. K. et al.** (2006). Absence of suppressor of cytokine signalling 3 reduces self-renewal and promotes differentiation in murine embryonic stem cells. *Stem Cells* **24**, 604-14.
- Friedman, J. R. and Kaestner, K. H.** (2006). The Foxa family of transcription factors in development and metabolism. *Cell Mol Life Sci* **63**, 2317-28.
- Fuhrmann, G., Chung, A. C., Jackson, K. J., Hummelke, G., Baniahamad, A., Sutter, J., Sylvester, I., Scholer, H. R. and Cooney, A. J.** (2001). Mouse germline restriction of Oct4 expression by germ cell nuclear factor. *Dev Cell* **1**, 377-87.
- Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki Ji, J. and Niwa, H.** (2002). Differentiation of embryonic stem cells is induced by GATA factors. *Genes Dev* **16**, 784-9.
- Fujita, N., Sato, S., Katayama, K. and Tsuruo, T.** (2002). Akt-dependent phosphorylation of p27Kip1 promotes binding to 14-3-3 and cytoplasmic localization. *J Biol Chem* **277**, 28706-13.
- Fujita, N., Sato, S. and Tsuruo, T.** (2003). Phosphorylation of p27Kip1 at threonine 198 by p90 ribosomal protein S6 kinases promotes its binding to 14-3-3 and cytoplasmic localization. *J Biol Chem* **278**, 49254-60.
- Fukada, T., Hibi, M., Yamanaka, Y., Takahashi-Tezuka, M., Fujitani, Y., Yamaguchi, T., Nakajima, K. and Hirano, T.** (1996). Two signals are necessary for cell proliferation

induced by a cytokine receptor gp130: involvement of STAT3 in anti-apoptosis. *Immunity* **5**, 449-60.

Futaki, S., Hayashi, Y., Emoto, T., Weber, C. N. and Sekiguchi, K. (2004). Sox7 plays crucial roles in parietal endoderm differentiation in F9 embryonal carcinoma cells through regulating Gata-4 and Gata-6 expression. *Mol Cell Biol* **24**, 10492-503.

Garrett-Sinha, L. A., Eberspaecher, H., Seldin, M. F. and de Crombrugghe, B. (1996). A gene for a novel zinc-finger protein expressed in differentiated epithelial cells and transiently in certain mesenchymal cells. *J Biol Chem* **271**, 31384-90.

Gearing, D. P., Comeau, M. R., Friend, D. J., Gimpel, S. D., Thut, C. J., McGourty, J., Brasher, K. K., King, J. A., Gillis, S., Mosley, B. et al. (1992). The IL-6 signal transducer, gp130: an oncostatin M receptor and affinity converter for the LIF receptor. *Science* **255**, 1434-7.

Gearing, D. P., Thut, C. J., VandeBos, T., Gimpel, S. D., Delaney, P. B., King, J., Price, V., Cosman, D. and Beckmann, M. P. (1991). Leukemia inhibitory factor receptor is structurally related to the IL-6 signal transducer, gp130. *Embo J* **10**, 2839-48.

Gordon, J. I., Schmidt, G. H. and Roth, K. A. (1992). Studies of intestinal stem cells using normal, chimeric, and transgenic mice. *Faseb J* **6**, 3039-50.

Graves, K. H. and Moreadith, R. W. (1993). Derivation and characterization of putative pluripotential embryonic stem cells from preimplantation rabbit embryos. *Mol Reprod Dev* **36**, 424-33.

Guschin, D., Rogers, N., Briscoe, J., Witthuhn, B., Watling, D., Horn, F., Pellegrini, S., Yasukawa, K., Heinrich, P., Stark, G. R. et al. (1995). A major role for the protein tyrosine kinase JAK1 in the JAK/STAT signal transduction pathway in response to interleukin-6. *Embo J* **14**, 1421-9.

Hanna, L. A., Foreman, R. K., Tarasenko, I. A., Kessler, D. S. and Labosky, P. A. (2002). Requirement for Foxd3 in maintaining pluripotent cells of the early mouse embryo. *Genes Dev* **16**, 2650-61.

Hart, A. H., Hartley, L., Ibrahim, M. and Robb, L. (2004). Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human. *Dev Dyn* **230**, 187-98.

Heinrich, P. C., Behrmann, I., Muller-Newen, G., Schaper, F. and Graeve, L. (1998). Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway. *Biochem J* **334** (Pt 2), 297-314.

Hilberg, F. and Wagner, E. F. (1992). Embryonic stem (ES) cells lacking functional c-jun: consequences for growth and differentiation, AP-1 activity and tumorigenicity. *Oncogene* **7**, 2371-80.

Horvath, C. M., Wen, Z. and Darnell, J. E., Jr. (1995). A STAT protein domain that determines DNA sequence recognition suggests a novel DNA-binding domain. *Genes Dev* **9**, 984-94.

Humphrey, R. K., Beattie, G. M., Lopez, A. D., Bucay, N., King, C. C., Firpo, M. T., Rose-John, S. and Hayek, A. (2004). Maintenance of pluripotency in human embryonic stem cells is STAT3 independent. *Stem Cells* **22**, 522-30.

Iannaccone, P. M., Taborn, G. U., Garton, R. L., Caplice, M. D. and Brenin, D. R. (1994). Pluripotent embryonic stem cells from the rat are capable of producing chimeras. *Dev Biol* **163**, 288-92.

Iguchi-Ariga, S. M., Itani, T., Kiji, Y. and Ariga, H. (1987). Possible function of the c-myc product: promotion of cellular DNA replication. *Embo J* **6**, 2365-71.

Indra, A. K., Warot, X., Brocard, J., Bornert, J. M., Xiao, J. H., Chambon, P. and Metzger, D. (1999). Temporally-controlled site-specific mutagenesis in the basal layer of the

- epidermis: comparison of the recombinase activity of the tamoxifen-inducible Cre-ER(T) and Cre-ER(T2) recombinases. *Nucleic Acids Res* **27**, 4324-7.
- Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y. and Lemischka, I. R.** (2006). Dissecting self-renewal in stem cells with RNA interference. *Nature* **442**, 533-8.
- Jakob, H., Boon, T., Gaillard, J., Nicolas, J. and Jacob, F.** (1973). [Teratocarcinoma of the mouse: isolation, culture and properties of pluripotential cells]. *Ann Microbiol (Paris)* **124**, 269-82.
- Jaubert, J., Cheng, J. and Segre, J. A.** (2003). Ectopic expression of kruppel like factor 4 (Klf4) accelerates formation of the epidermal permeability barrier. *Development* **130**, 2767-77.
- Jiang, J., Chan, Y. S., Loh, Y. H., Cai, J., Tong, G. Q., Lim, C. A., Robson, P., Zhong, S. and Ng, H. H.** (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat Cell Biol* **10**, 353-60.
- Kaczynski, J., Cook, T. and Urrutia, R.** (2003). Sp1- and Kruppel-like transcription factors. *Genome Biol* **4**, 206.
- Kahan, B. W. and Ephrussi, B.** (1970). Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma. *J Natl Cancer Inst* **44**, 1015-36.
- Kefalas, P., Brown, T. R. and Brickell, P. M.** (1995). Signalling by the p60c-src family of protein-tyrosine kinases. *Int J Biochem Cell Biol* **27**, 551-63.
- Kim, J., Chu, J., Shen, X., Wang, J. and Orkin, S. H.** (2008). An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132**, 1049-61.
- Kobayashi, T. and Cohen, P.** (1999). Activation of serum- and glucocorticoid-regulated protein kinase by agonists that activate phosphatidylinositide 3-kinase is mediated by 3-phosphoinositide-dependent protein kinase-1 (PDK1) and PDK2. *Biochem J* **339 (Pt 2)**, 319-28.
- Kohlhase, J., Heinrich, M., Liebers, M., Frohlich Archangelo, L., Reardon, W. and Kispert, A.** (2002). Cloning and expression analysis of SALL4, the murine homologue of the gene mutated in Okihiro syndrome. *Cytogenet Genome Res* **98**, 274-7.
- Krumenacker, J. S., Buckley, D. J., Leff, M. A., McCormack, J. T., de Jong, G., Gout, P. W., Reed, J. C., Miyashita, T., Magnuson, N. S. and Buckley, A. R.** (1998). Prolactin-regulated apoptosis of Nb2 lymphoma cells: pim-1, bcl-2, and bax expression. *Endocrine* **9**, 163-70.
- Kuroda, T., Tada, M., Kubota, H., Kimura, H., Hatano, S. Y., Suemori, H., Nakatsuji, N. and Tada, T.** (2005). Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. *Mol Cell Biol* **25**, 2475-85.
- Lanahan, A., Williams, J. B., Sanders, L. K. and Nathans, D.** (1992). Growth factor-induced delayed early response genes. *Mol Cell Biol* **12**, 3919-29.
- Lee, J. H., Hart, S. R. and Skalnik, D. G.** (2004). Histone deacetylase activity is required for embryonic stem cell differentiation. *Genesis* **38**, 32-8.
- Lee, T. I., Jenner, R. G., Boyer, L. A., Guenther, M. G., Levine, S. S., Kumar, R. M., Chevalier, B., Johnstone, S. E., Cole, M. F., Isono, K. et al.** (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125**, 301-13.
- Levinson-Dushnik, M. and Benvenisty, N.** (1997). Involvement of hepatocyte nuclear factor 3 in endoderm differentiation of embryonic stem cells. *Mol Cell Biol* **17**, 3817-22.
- Li, M., Sendtner, M. and Smith, A.** (1995). Essential function of LIF receptor in motor neurons. *Nature* **378**, 724-7.
- Li, W., Nishimura, R., Kashishian, A., Batzer, A. G., Kim, W. J., Cooper, J. A. and Schlessinger, J.** (1994). A new function for a phosphotyrosine phosphatase: linking GRB2-Sos to a receptor tyrosine kinase. *Mol Cell Biol* **14**, 509-17.

- Li, Y., McClintick, J., Zhong, L., Edenberg, H. J., Yoder, M. C. and Chan, R. J.** (2005). Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. *Blood* **105**, 635-7.
- Li, Y. Y., Popivanova, B. K., Nagai, Y., Ishikura, H., Fujii, C. and Mukaida, N.** (2006). Pim-3, a proto-oncogene with serine/threonine kinase activity, is aberrantly expressed in human pancreatic cancer and phosphorylates bad to block bad-mediated apoptosis in human pancreatic cancer cell lines. *Cancer Res* **66**, 6741-7.
- Lilly, M. and Kraft, A.** (1997). Enforced expression of the Mr 33,000 Pim-1 kinase enhances factor-independent survival and inhibits apoptosis in murine myeloid cells. *Cancer Res* **57**, 5348-55.
- Lilly, M., Le, T., Holland, P. and Hendrickson, S. L.** (1992). Sustained expression of the pim-1 kinase is specifically induced in myeloid cells by cytokines whose receptors are structurally related. *Oncogene* **7**, 727-32.
- Lim, L. S., Loh, Y. H., Zhang, W., Li, Y., Chen, X., Wang, Y., Bakre, M., Ng, H. H. and Stanton, L. W.** (2007). Zic3 is required for maintenance of pluripotency in embryonic stem cells. *Mol Biol Cell* **18**, 1348-58.
- Liu, Y., Sinha, S. and Owens, G.** (2003). A transforming growth factor-beta control element required for SM alpha-actin expression in vivo also partially mediates GKLF-dependent transcriptional repression. *J Biol Chem* **278**, 48004-11.
- Lloyd, A. C.** (1998). Ras versus cyclin-dependent kinase inhibitors. *Curr Opin Genet Dev* **8**, 43-8.
- Loh, Y. H., Wu, Q., Chew, J. L., Vega, V. B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J. et al.** (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* **38**, 431-40.
- Losman, J., Chen, X. P., Jiang, H., Pan, P. Y., Kashiwada, M., Giallourakis, C., Cowan, S., Foltenyi, K. and Rothman, P.** (1999). IL-4 signaling is regulated through the recruitment of phosphatases, kinases, and SOCS proteins to the receptor complex. *Cold Spring Harb Symp Quant Biol* **64**, 405-16.
- Martin, G. R.** (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* **78**, 7634-8.
- Martin, G. R. and Evans, M. J.** (1975). Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro. *Proc Natl Acad Sci U S A* **72**, 1441-5.
- Masaki, H., Nishida, T., Kitajima, S., Asahina, K. and Teraoka, H.** (2007). Developmental pluripotency-associated 4 (DPPA4) localized in active chromatin inhibits mouse embryonic stem cell differentiation into a primitive ectoderm lineage. *J Biol Chem* **282**, 33034-42.
- Matoba, R., Niwa, H., Masui, S., Ohtsuka, S., Carter, M. G., Sharov, A. A. and Ko, M. S.** (2006). Dissecting Oct3/4-regulated gene networks in embryonic stem cells by expression profiling. *PLoS ONE* **1**, e26.
- Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T. and Yokota, T.** (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *Embo J* **18**, 4261-9.
- McGrath, K. E., Koniski, A. D., Maltby, K. M., McGann, J. K. and Palis, J.** (1999). Embryonic expression and function of the chemokine SDF-1 and its receptor, CXCR4. *Dev Biol* **213**, 442-56.
- Mikkers, H., Allen, J., Knipscheer, P., Romeijn, L., Hart, A., Vink, E. and Berns, A.** (2002). High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. *Nat Genet* **32**, 153-9.

- Mikkers, H., Nawijn, M., Allen, J., Brouwers, C., Verhoeven, E., Jonkers, J. and Berns, A.** (2004). Mice deficient for all PIM kinases display reduced body size and impaired responses to hematopoietic growth factors. *Mol Cell Biol* **24**, 6104-15.
- Miller, I. J. and Bieker, J. J.** (1993). A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. *Mol Cell Biol* **13**, 2776-86.
- Minami, M., Inoue, M., Wei, S., Takeda, K., Matsumoto, M., Kishimoto, T. and Akira, S.** (1996). STAT3 activation is a critical step in gp130-mediated terminal differentiation and growth arrest of a myeloid cell line. *Proc Natl Acad Sci U S A* **93**, 3963-6.
- Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M. and Yamanaka, S.** (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631-42.
- Miyazawa, K., Shinozaki, M., Hara, T., Furuya, T. and Miyazono, K.** (2002). Two major Smad pathways in TGF-beta superfamily signalling. *Genes Cells* **7**, 1191-204.
- Mochizuki, T., Kitanaka, C., Noguchi, K., Muramatsu, T., Asai, A. and Kuchino, Y.** (1999). Physical and functional interactions between Pim-1 kinase and Cdc25A phosphatase. Implications for the Pim-1-mediated activation of the c-Myc signalling pathway. *J Biol Chem* **274**, 18659-66.
- Morishita, D., Katayama, R., Sekimizu, K., Tsuruo, T. and Fujita, N.** (2008). Pim kinases promote cell cycle progression by phosphorylating and down-regulating p27Kip1 at the transcriptional and posttranscriptional levels. *Cancer Res* **68**, 5076-85.
- Morris, S. M., Tallquist, M. D., Rock, C. O. and Cooper, J. A.** (2002). Dual roles for the Dab2 adaptor protein in embryonic development and kidney transport. *Embo J* **21**, 1555-64.
- Morrisey, E. E., Musco, S., Chen, M. Y., Lu, M. M., Leiden, J. M. and Parmacek, M. S.** (2000). The gene encoding the mitogen-responsive phosphoprotein Dab2 is differentially regulated by GATA-6 and GATA-4 in the visceral endoderm. *J Biol Chem* **275**, 19949-54.
- Morrisey, E. E., Tang, Z., Sigrist, K., Lu, M. M., Jiang, F., Ip, H. S. and Parmacek, M. S.** (1998). GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes Dev* **12**, 3579-90.
- Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N. and Yamanaka, S.** (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol* **26**, 101-6.
- Nakano, T., Kodama, H. and Honjo, T.** (1994). Generation of lymphohematopoietic cells from embryonic stem cells in culture. *Science* **265**, 1098-101.
- Nandan, M. O., Yoon, H. S., Zhao, W., Ouko, L. A., Chanchevalap, S. and Yang, V. W.** (2004). Kruppel-like factor 5 mediates the transforming activity of oncogenic H-Ras. *Oncogene* **23**, 3404-13.
- Niakan, K. K. and McCabe, E. R.** (2005). DAX1 origin, function, and novel role. *Mol Genet Metab* **86**, 70-83.
- Nichols, J., Chambers, I. and Smith, A.** (1994). Derivation of germline competent embryonic stem cells with a combination of interleukin-6 and soluble interleukin-6 receptor. *Exp Cell Res* **215**, 237-9.
- Nichols, J., Chambers, I., Taga, T. and Smith, A.** (2001). Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines. *Development* **128**, 2333-9.
- Nichols, J., Davidson, D., Taga, T., Yoshida, K., Chambers, I. and Smith, A.** (1996). Complementary tissue-specific expression of LIF and LIF-receptor mRNAs in early mouse embryogenesis. *Mech Dev* **57**, 123-31.
- Nichols, J., Evans, E. P. and Smith, A. G.** (1990). Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity. *Development* **110**, 1341-8.

- Nishikawa, S. I., Nishikawa, S., Hirashima, M., Matsuyoshi, N. and Kodama, H.** (1998). Progressive lineage analysis by cell sorting and culture identifies FLK1+VE-cadherin+ cells at a diverging point of endothelial and hemopoietic lineages. *Development* **125**, 1747-57.
- Niwa, H., Burdon, T., Chambers, I. and Smith, A.** (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev* **12**, 2048-60.
- Niwa, H., Miyazaki, J. and Smith, A. G.** (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**, 372-6.
- No, D., Yao, T. P. and Evans, R. M.** (1996). Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc Natl Acad Sci U S A* **93**, 3346-51.
- Okumura-Nakanishi, S., Saito, M., Niwa, H. and Ishikawa, F.** (2005). Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J Biol Chem* **280**, 5307-17.
- Pages, G., Guerin, S., Grall, D., Bonino, F., Smith, A., Anjuere, F., Auberger, P. and Pouyssegur, J.** (1999). Defective thymocyte maturation in p44 MAP kinase (Erk 1) knockout mice. *Science* **286**, 1374-7.
- Pain, B., Clark, M. E., Shen, M., Nakazawa, H., Sakurai, M., Samarut, J. and Etches, R. J.** (1996). Long-term in vitro culture and characterisation of avian embryonic stem cells with multiple morphogenetic potentialities. *Development* **122**, 2339-48.
- Paling, N. R., Whealon, H., Bone, H. K. and Welham, M. J.** (2004). Regulation of embryonic stem cell self-renewal by phosphoinositide 3-kinase-dependent signaling. *J Biol Chem* **279**, 48063-70.
- Pan, G. J., Chang, Z. Y., Scholer, H. R. and Pei, D.** (2002). Stem cell pluripotency and transcription factor Oct4. *Cell Res* **12**, 321-9.
- Park, J., Leong, M. L., Buse, P., Maiyar, A. C., Firestone, G. L. and Hemmings, B. A.** (1999). Serum and glucocorticoid-inducible kinase (SGK) is a target of the PI 3-kinase-stimulated signaling pathway. *Embo J* **18**, 3024-33.
- Pasini, D., Bracken, A. P., Hansen, J. B., Capillo, M. and Helin, K.** (2007). The polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol Cell Biol* **27**, 3769-79.
- Pau, K. Y. and Wolf, D. P.** (2004). Derivation and characterization of monkey embryonic stem cells. *Reprod Biol Endocrinol* **2**, 41.
- Pennica, D., Shaw, K. J., Swanson, T. A., Moore, M. W., Shelton, D. L., Zioncheck, K. A., Rosenthal, A., Taga, T., Paoni, N. F. and Wood, W. I.** (1995). Cardiotrophin-1. Biological activities and binding to the leukemia inhibitory factor receptor/gp130 signaling complex. *J Biol Chem* **270**, 10915-22.
- Pesce, M. and Scholer, H. R.** (2001). Oct-4: gatekeeper in the beginnings of mammalian development. *Stem Cells* **19**, 271-8.
- Piccinni, S. A., Bolcato-Bellemin, A. L., Klein, A., Yang, V. W., Kedinger, M., Simon-Assmann, P. and Lefebvre, O.** (2004). Kruppel-like factors regulate the Lama1 gene encoding the laminin alpha1 chain. *J Biol Chem* **279**, 9103-14.
- Qu, C. K. and Feng, G. S.** (1998). Shp-2 has a positive regulatory role in ES cell differentiation and proliferation. *Oncogene* **17**, 433-9.
- Qu, X. B., Pan, J., Zhang, C. and Huang, S. Y.** (2008). Sox17 facilitates the differentiation of mouse embryonic stem cells into primitive and definitive endoderm in vitro. *Dev Growth Differ*.
- Resnitzky, D. and Kimchi, A.** (1991). Deregulated c-myc expression abrogates the interferon- and interleukin 6-mediated G0/G1 cell cycle arrest but not other inhibitory responses in M1 myeloblastic cells. *Cell Growth Differ* **2**, 33-41.
- Rodda, D. J., Chew, J. L., Lim, L. H., Loh, Y. H., Wang, B., Ng, H. H. and Robson, P.** (2005). Transcriptional regulation of nanog by OCT4 and SOX2. *J Biol Chem* **280**, 24731-7.

- Rodig, S. J., Meraz, M. A., White, J. M., Lampe, P. A., Riley, J. K., Arthur, C. D., King, K. L., Sheehan, K. C., Yin, L., Pennica, D. et al.** (1998). Disruption of the Jak1 gene demonstrates obligatory and nonredundant roles of the Jaks in cytokine-induced biologic responses. *Cell* **93**, 373-83.
- Rose, T. M., Weiford, D. M., Gunderson, N. L. and Bruce, A. G.** (1994). Oncostatin M (OSM) inhibits the differentiation of pluripotent embryonic stem cells in vitro. *Cytokine* **6**, 48-54.
- Rosner, M. H., Vigano, M. A., Ozato, K., Timmons, P. M., Poirier, F., Rigby, P. W. and Staudt, L. M.** (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* **345**, 686-92.
- Saba-El-Leil, M. K., Vella, F. D., Vernay, B., Voisin, L., Chen, L., Labrecque, N., Ang, S. L. and Meloche, S.** (2003). An essential function of the mitogen-activated protein kinase Erk2 in mouse trophoblast development. *EMBO Rep* **4**, 964-8.
- Sakai, I. and Kraft, A. S.** (1997). The kinase domain of Jak2 mediates induction of bcl-2 and delays cell death in hematopoietic cells. *J Biol Chem* **272**, 12350-8.
- Savatier, P., Lapillonne, H., van Grunsven, L. A., Rudkin, B. B. and Samarut, J.** (1996). Withdrawal of differentiation inhibitory activity/leukemia inhibitory factor up-regulates D-type cyclins and cyclin-dependent kinase inhibitors in mouse embryonic stem cells. *Oncogene* **12**, 309-22.
- Schier, A. F.** (2003). Nodal signaling in vertebrate development. *Annu Rev Cell Dev Biol* **19**, 589-621.
- Scholer, H. R., Dressler, G. R., Balling, R., Rohdewohld, H. and Gruss, P.** (1990). Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. *Embo J* **9**, 2185-95.
- Schorpp-Kistner, M., Wang, Z. Q., Angel, P. and Wagner, E. F.** (1999). JunB is essential for mammalian placentation. *Embo J* **18**, 934-48.
- Segre, J. A., Bauer, C. and Fuchs, E.** (1999). Klf4 is a transcription factor required for establishing the barrier function of the skin. *Nat Genet* **22**, 356-60.
- Seguin, C. A., Draper, J. S., Nagy, A. and Rossant, J.** (2008). Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. *Cell Stem Cell* **3**, 182-95.
- Selten, G., Cuypers, H. T. and Berns, A.** (1985). Proviral activation of the putative oncogene Pim-1 in MuLV induced T-cell lymphomas. *Embo J* **4**, 1793-8.
- Sharov, A. A., Masui, S., Sharova, L. V., Piao, Y., Aiba, K., Matoba, R., Xin, L., Niwa, H. and Ko, M. S.** (2008). Identification of Pou5f1, Sox2, and Nanog downstream target genes with statistical confidence by applying a novel algorithm to time course microarray and genome-wide chromatin immunoprecipitation data. *BMC Genomics* **9**, 269.
- Shi, Y., Glynn, J. M., Guilbert, L. J., Cotter, T. G., Bissonnette, R. P. and Green, D. R.** (1992). Role for c-myc in activation-induced apoptotic cell death in T cell hybridomas. *Science* **257**, 212-4.
- Shie, J. L., Chen, Z. Y., Fu, M., Pestell, R. G. and Tseng, C. C.** (2000). Gut-enriched Kruppel-like factor represses cyclin D1 promoter activity through Sp1 motif. *Nucleic Acids Res* **28**, 2969-76.
- Shields, J. M., Christy, R. J. and Yang, V. W.** (1996). Identification and characterization of a gene encoding a gut-enriched Kruppel-like factor expressed during growth arrest. *J Biol Chem* **271**, 20009-17.
- Shields, J. M. and Yang, V. W.** (1997). Two potent nuclear localization signals in the gut-enriched Kruppel-like factor define a subfamily of closely related Kruppel proteins. *J Biol Chem* **272**, 18504-7.

- Shimosato, D., Shiki, M. and Niwa, H.** (2007). Extra-embryonic endoderm cells derived from ES cells induced by GATA factors acquire the character of XEN cells. *BMC Dev Biol* **7**, 80.
- Shindo, T., Manabe, I., Fukushima, Y., Tobe, K., Aizawa, K., Miyamoto, S., Kawai-Kowase, K., Moriyama, N., Imai, Y., Kawakami, H. et al.** (2002). Kruppel-like zinc-finger transcription factor KLF5/BTEB2 is a target for angiotensin II signaling and an essential regulator of cardiovascular remodeling. *Nat Med* **8**, 856-63.
- Shirogane, T., Fukada, T., Muller, J. M., Shima, D. T., Hibi, M. and Hirano, T.** (1999). Synergistic roles for Pim-1 and c-Myc in STAT3-mediated cell cycle progression and antiapoptosis. *Immunity* **11**, 709-19.
- Singh, S. K., Kagalwala, M. N., Parker-Thornburg, J., Adams, H. and Majumder, S.** (2008). REST maintains self-renewal and pluripotency of embryonic stem cells. *Nature* **453**, 223-7.
- Smith, A. G.** (2001). Embryo-derived stem cells: of mice and men. *Annu Rev Cell Dev Biol* **17**, 435-62.
- Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M. and Rogers, D.** (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336**, 688-90.
- Solter, D., Skreb, N. and Damjanov, I.** (1970). Extrauterine growth of mouse egg-cylinders results in malignant teratoma. *Nature* **227**, 503-4.
- Soudais, C., Bielinska, M., Heikinheimo, M., MacArthur, C. A., Narita, N., Saffitz, J. E., Simon, M. C., Leiden, J. M. and Wilson, D. B.** (1995). Targeted mutagenesis of the transcription factor GATA-4 gene in mouse embryonic stem cells disrupts visceral endoderm differentiation in vitro. *Development* **121**, 3877-88.
- Stahl, N., Boulton, T. G., Farruggella, T., Ip, N. Y., Davis, S., Witthuhn, B. A., Quelle, F. W., Silvennoinen, O., Barbieri, G., Pellegrini, S. et al.** (1994). Association and activation of Jak-Tyk kinases by CNTF-LIF-OSM-IL-6 beta receptor components. *Science* **263**, 92-5.
- Stahl, N., Farruggella, T. J., Boulton, T. G., Zhong, Z., Darnell, J. E., Jr. and Yancopoulos, G. D.** (1995). Choice of STATs and other substrates specified by modular tyrosine-based motifs in cytokine receptors. *Science* **267**, 1349-53.
- Starr, R. and Hilton, D. J.** (1999). Negative regulation of the JAK/STAT pathway. *Bioessays* **21**, 47-52.
- Starr, R., Willson, T. A., Viney, E. M., Murray, L. J., Rayner, J. R., Jenkins, B. J., Gonda, T. J., Alexander, W. S., Metcalf, D., Nicola, N. A. et al.** (1997). A family of cytokine-inducible inhibitors of signalling. *Nature* **387**, 917-21.
- Stevens, L. C.** (1970). The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. *Dev Biol* **21**, 364-82.
- Stewart, C. L., Kaspar, P., Brunet, L. J., Bhatt, H., Gadi, I., Kontgen, F. and Abbondanzo, S. J.** (1992). Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. *Nature* **359**, 76-9.
- Storm, M. P., Bone, H. K., Beck, C. G., Bourillot, P. Y., Schreiber, V., Damiano, T., Nelson, A., Savatier, P. and Welham, M. J.** (2007). Regulation of Nanog expression by phosphoinositide 3-kinase-dependent signaling in murine embryonic stem cells. *J Biol Chem* **282**, 6265-73.
- Suemori, H., Tada, T., Torii, R., Hosoi, Y., Kobayashi, K., Imahie, H., Kondo, Y., Iritani, A. and Nakatsuji, N.** (2001). Establishment of embryonic stem cell lines from cynomolgus monkey blastocysts produced by IVF or ICSI. *Dev Dyn* **222**, 273-9.
- Sumi, T., Fujimoto, Y., Nakatsuji, N. and Suemori, H.** (2004). STAT3 is dispensable for maintenance of self-renewal in nonhuman primate embryonic stem cells. *Stem Cells* **22**, 861-72.

- Summy, J. M. and Gallick, G. E.** (2003). Src family kinases in tumor progression and metastasis. *Cancer Metastasis Rev* **22**, 337-58.
- Sun, R., Chen, X. and Yang, V. W.** (2001). Intestinal-enriched Kruppel-like factor (Kruppel-like factor 5) is a positive regulator of cellular proliferation. *J Biol Chem* **276**, 6897-900.
- Suske, G., Bruford, E. and Philipsen, S.** (2005). Mammalian SP/KLF transcription factors: bring in the family. *Genomics* **85**, 551-6.
- Takahashi, K. and Yamanaka, S.** (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-76.
- Takeda, K., Noguchi, K., Shi, W., Tanaka, T., Matsumoto, M., Yoshida, N., Kishimoto, T. and Akira, S.** (1997). Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. *Proc Natl Acad Sci U S A* **94**, 3801-4.
- Takenaga, M., Fukumoto, M. and Hori, Y.** (2007). Regulated Nodal signaling promotes differentiation of the definitive endoderm and mesoderm from ES cells. *J Cell Sci* **120**, 2078-90.
- Telerman, A., Amson, R., Zakut-Houri, R. and Givol, D.** (1988). Identification of the human pim-1 gene product as a 33-kilodalton cytoplasmic protein with tyrosine kinase activity. *Mol Cell Biol* **8**, 1498-503.
- Temple, R., Allen, E., Fordham, J., Phipps, S., Schneider, H. C., Lindauer, K., Hayes, I., Lockey, J., Pollock, K. and Jupp, R.** (2001). Microarray analysis of eosinophils reveals a number of candidate survival and apoptosis genes. *Am J Respir Cell Mol Biol* **25**, 425-33.
- Thomas, K. R. and Capecchi, M. R.** (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**, 503-12.
- Thompson, S., Clarke, A. R., Pow, A. M., Hooper, M. L. and Melton, D. W.** (1989). Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. *Cell* **56**, 313-21.
- Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S. and Jones, J. M.** (1998). Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145-7.
- Thomson, J. A., Kalishman, J., Golos, T. G., Durning, M., Harris, C. P. and Hearn, J. P.** (1996). Pluripotent cell lines derived from common marmoset (*Callithrix jacchus*) blastocysts. *Biol Reprod* **55**, 254-9.
- Ton-That, H., Kaestner, K. H., Shields, J. M., Mahatanankoon, C. S. and Yang, V. W.** (1997). Expression of the gut-enriched Kruppel-like factor gene during development and intestinal tumorigenesis. *FEBS Lett* **419**, 239-43.
- Torres, J. and Watt, F. M.** (2008). Nanog maintains pluripotency of mouse embryonic stem cells by inhibiting NFkappaB and cooperating with Stat3. *Nat Cell Biol* **10**, 194-201.
- Ura, H., Usuda, M., Kinoshita, K., Sun, C., Mori, K., Akagi, T., Matsuda, T., Koide, H. and Yokota, T.** (2008). STAT3 and Oct-3/4 control histone modification through induction of Eed in embryonic stem cells. *J Biol Chem* **283**, 9713-23.
- van der Lugt, N. M., Domen, J., Verhoeven, E., Linders, K., van der Gulden, H., Allen, J. and Berns, A.** (1995). Proviral tagging in E mu-myc transgenic mice lacking the Pim-1 proto-oncogene leads to compensatory activation of Pim-2. *Embo J* **14**, 2536-44.
- Vanhaezebroeck, B., Leevens, S. J., Ahmadi, K., Timms, J., Katso, R., Driscoll, P. C., Woscholski, R., Parker, P. J. and Waterfield, M. D.** (2001). Synthesis and function of 3-phosphorylated inositol lipids. *Annu Rev Biochem* **70**, 535-602.
- Vaux, D. L., Cory, S. and Adams, J. M.** (1988). Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells. *Nature* **335**, 440-2.

- Vlahos, C. J., Matter, W. F., Hui, K. Y. and Brown, R. F.** (1994). A specific inhibitor of phosphatidylinositol 3-kinase, 2-(4-morpholinyl)-8-phenyl-4H-1-benzopyran-4-one (LY294002). *J Biol Chem* **269**, 5241-8.
- Wagner, A. J., Small, M. B. and Hay, N.** (1993). Myc-mediated apoptosis is blocked by ectopic expression of Bcl-2. *Mol Cell Biol* **13**, 2432-40.
- Watanabe, S., Umebara, H., Murayama, K., Okabe, M., Kimura, T. and Nakano, T.** (2006). Activation of Akt signaling is sufficient to maintain pluripotency in mouse and primate embryonic stem cells. *Oncogene* **25**, 2697-707.
- Wen, J. P., Eder, C. and Lipps, H. J.** (1995). The processing of macronuclear-destined DNA sequences microinjected into the macronuclear anlagen of the hypotrichous ciliate *Stylonchia lemnæ*. *Nucleic Acids Res* **23**, 1704-9.
- Wiles, M. V. and Keller, G.** (1991). Multiple hematopoietic lineages develop from embryonic stem (ES) cells in culture. *Development* **111**, 259-67.
- Wilson, S. I., Rydstrom, A., Trimborn, T., Willert, K., Nusse, R., Jessell, T. M. and Edlund, T.** (2001). The status of Wnt signalling regulates neural and epidermal fates in the chick embryo. *Nature* **411**, 325-30.
- Wobus, A. M., Wallukat, G. and Hescheler, J.** (1991). Pluripotent mouse embryonic stem cells are able to differentiate into cardiomyocytes expressing chronotropic responses to adrenergic and cholinergic agents and Ca²⁺ channel blockers. *Differentiation* **48**, 173-82.
- Wozney, J. M. and Rosen, V.** (1998). Bone morphogenetic protein and bone morphogenetic protein gene family in bone formation and repair. *Clin Orthop Relat Res*, 26-37.
- Wu, Q., Chen, X., Zhang, J., Loh, Y. H., Low, T. Y., Zhang, W., Zhang, W., Sze, S. K., Lim, B. and Ng, H. H.** (2006). Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. *J Biol Chem* **281**, 24090-4.
- Yan, B., Zemskova, M., Holder, S., Chin, V., Kraft, A., Koskinen, P. J. and Lilly, M.** (2003). The PIM-2 kinase phosphorylates BAD on serine 112 and reverses BAD-induced cell death. *J Biol Chem* **278**, 45358-67.
- Yang, D. H., Smith, E. R., Roland, I. H., Sheng, Z., He, J., Martin, W. D., Hamilton, T. C., Lambeth, J. D. and Xu, X. X.** (2002). Disabled-2 is essential for endodermal cell positioning and structure formation during mouse embryogenesis. *Dev Biol* **251**, 27-44.
- Yang, Z. Z., Tschopp, O., Baudry, A., Dummler, B., Hynx, D. and Hemmings, B. A.** (2004). Physiological functions of protein kinase B/Akt. *Biochem Soc Trans* **32**, 350-4.
- Yasunaga, M., Tada, S., Torikai-Nishikawa, S., Nakano, Y., Okada, M., Jakt, L. M., Nishikawa, S., Chiba, T., Era, T. and Nishikawa, S.** (2005). Induction and monitoring of definitive and visceral endoderm differentiation of mouse ES cells. *Nat Biotechnol* **23**, 1542-50.
- Yin, T., Taga, T., Tsang, M. L., Yasukawa, K., Kishimoto, T. and Yang, Y. C.** (1993). Involvement of IL-6 signal transducer gp130 in IL-11-mediated signal transduction. *J Immunol* **151**, 2555-61.
- Ying, Q. L., Nichols, J., Chambers, I. and Smith, A.** (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **115**, 281-92.
- Yoon, H. S., Chen, X. and Yang, V. W.** (2003). Kruppel-like factor 4 mediates p53-dependent G1/S cell cycle arrest in response to DNA damage. *J Biol Chem* **278**, 2101-5.
- Yoon, H. S. and Yang, V. W.** (2004). Requirement of Kruppel-like factor 4 in preventing entry into mitosis following DNA damage. *J Biol Chem* **279**, 5035-41.
- Yoshida, K., Chambers, I., Nichols, J., Smith, A., Saito, M., Yasukawa, K., Shoyab, M., Taga, T. and Kishimoto, T.** (1994). Maintenance of the pluripotential phenotype of embryonic stem cells through direct activation of gp130 signalling pathways. *Mech Dev* **45**, 163-71.

- Yoshida, K., Taga, T., Saito, M., Suematsu, S., Kumanogoh, A., Tanaka, T., Fujiwara, H., Hirata, M., Yamagami, T., Nakahata, T. et al.** (1996). Targeted disruption of gp130, a common signal transducer for the interleukin 6 family of cytokines, leads to myocardial and hematological disorders. *Proc Natl Acad Sci U S A* **93**, 407-11.
- Yoshikawa, T., Piao, Y., Zhong, J., Matoba, R., Carter, M. G., Wang, Y., Goldberg, I. and Ko, M. S.** (2006). High-throughput screen for genes predominantly expressed in the ICM of mouse blastocysts by whole mount in situ hybridization. *Gene Expr Patterns* **6**, 213-24.
- Zhang, J., Tam, W. L., Tong, G. Q., Wu, Q., Chan, H. Y., Soh, B. S., Lou, Y., Yang, J., Ma, Y., Chai, L. et al.** (2006). Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat Cell Biol* **8**, 1114-23.
- Zhang, W., Geiman, D. E., Shields, J. M., Dang, D. T., Mahatan, C. S., Kaestner, K. H., Biggs, J. R., Kraft, A. S. and Yang, V. W.** (2000). The gut-enriched Kruppel-like factor (Kruppel-like factor 4) mediates the transactivating effect of p53 on the p21WAF1/Cip1 promoter. *J Biol Chem* **275**, 18391-8.
- Zhang, Y., Wang, Z. and Magnuson, N. S.** (2007). Pim-1 kinase-dependent phosphorylation of p21Cip1/WAF1 regulates its stability and cellular localization in H1299 cells. *Mol Cancer Res* **5**, 909-22.
- Zhao, W., Hisamuddin, I. M., Nandan, M. O., Babbin, B. A., Lamb, N. E. and Yang, V. W.** (2004). Identification of Kruppel-like factor 4 as a potential tumor suppressor gene in colorectal cancer. *Oncogene* **23**, 395-402.
- Zhong, Z., Wen, Z. and Darnell, J. E., Jr.** (1994). Stat3: a STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and interleukin-6. *Science* **264**, 95-8.
- Zhou, X., Sasaki, H., Lowe, L., Hogan, B. L. and Kuehn, M. R.** (1993). Nodal is a novel TGF-beta-like gene expressed in the mouse node during gastrulation. *Nature* **361**, 543-7.
- Zwaka, T. P. and Thomson, J. A.** (2005). Differentiation of human embryonic stem cells occurs through symmetric cell division. *Stem Cells* **23**, 146-9.

ANNEXE

Article n°3

*Genome analysis of self-renewal and differentiation in
mouse embryonic stem cells*

Soumis à Cell Stem Cell

Genomic analysis of Self-Renewal and Differentiation in mouse Embryonic Stem Cells

Herbert Schulz*, Raivo Kolde*, Priit Adler, Irène Aksoy, Konstantinos Anastassiadis, Michael Bader, Nathalie Billon, Hélène Boeuf, Pierre-Yves Bourillot, Frank Buchholz, Christian Dani, Michael Xavier Doss, Lesley Forrester, Ruairi Friel, Murielle Gitton, Domingos Henrique, Jürgen Hescheler, Heinz Himmelbauer, Norbert Hubner, Efi Karantzali, Androniki Kretsovali, Sandra Lubitz, Laurent Pradier, Jüri Reimand, Alexandra Rolletschek, Agapios Sachinidis, Pierre Savatier, Francis Stewart, Mike P. Storm, Marina Trouillas, Jaak Vilo, Melanie J. Welham, Johannes Winkler, Anna Wobus and Antonis K. Hatzopoulos†

From the “Functional genomics in embryonic stem cells” Consortium

* *The first two authors contributed equally to this work*

† Department of Medicine -Division of Cardiovascular Medicine and Department of Cell and Developmental Biology, Vanderbilt University, Nashville, Tennessee 37232, U.S.A.
& Institute of Clinical Molecular Biology and Tumor Genetics, Helmholtz Center Munich, Germany

Summary

Embryonic stem (ES) cells have practically unlimited self-renewal capacity and the potential to differentiate to a large variety of cell types. We have analyzed the transcriptome of three mouse ES cell lines in ten experimental settings using the Affymetrix 430 gene chips. Collective data analysis grouped transcripts in clusters containing genes with comparable expression patterns during differentiation. Comparison of gene expression profiles shows that individual ES cell lines share common differentiation programs, but also follow divergent expression patterns that correlate with their differentiation potential. The expression data have been organized in an interactive, open-access database that can be searched for the expression patterns of any transcripts during ES differentiation. The genomics data are also presented in animated graphs of KEGG pathways of signaling and metabolic cascades during ES cell differentiation. Finally, we have incorporated functional annotations for genes of interest to facilitate studies on the characteristics of embryonic stem cells.

Introduction

Stem cells hold great promise for tissue repair after injury¹. Studies in animal models and clinical trials indicate that stem cells may replace damaged cells improving organ recovery and function^{2,3}. For this reason, the genetic programs controlling self-renewal and differentiation of stem cells are of great interest for developing future tools to unlock the regenerative potential of stem cells. To this end, the mouse embryonic stem (ES) cells offer an excellent model system because they give rise to many different cell types in a reproducible manner, can be propagated practically indefinitely, have relatively stable karyotypes, and are easy to manipulate genetically⁴⁻⁷. Moreover, ES cell differentiation *in vitro* closely follows events that take place during early embryonic development including the formation of the three germ layers of ectoderm, mesoderm and endoderm, and the emergence of neuronal, endothelial, hematopoietic, cardiac, hepatic or pancreatic cells^{8,9}.

Functional studies have pointed out the critical roles of genes such as *Oct4* (*Pou5f1*)¹⁶⁻¹⁹, *Nanog*²⁰⁻²² and *Sox2*²³⁻²⁵ in the maintenance of ES pluripotency and suppression of differentiation pathways. Chromatin Immunoprecipitation and chip analysis revealed that both active and silenced genes are directly bound in ES cells by one or more of these three proteins (Boyer, Loh). The recent discoveries of new pluripotency factors including *Sall4*²⁶ and *Klf4*^{27,28} suggest that the expansion and fate of stem cells follows a complex course requiring the coordinated action of a significant number of yet to be characterized genes^{13,29,30}. Toward this goal, microarray studies have the potential to piece together groups of coregulated genes and discover novel components of regulatory pathways in ES cells.

In recent years, a number of genome-wide approaches have identified transcripts present in mouse and human ES cells or their differentiated derivatives using a variety of gene expression profiling methods¹⁰⁻¹³. This wealth of information also underscored a degree of variability among different datasets¹⁴ as well as the fact that high throughput approaches can generate “biological noise”¹⁵. The “Functional Genomics in Embryonic Stem Cells” consortium comprising 20 research groups (acronym FunGenES; <http://www.fungenes.org>) studied the transcriptome of ES cells during growth and differentiation towards ectodermal, endodermal and mesodermal lineages. Besides information gathered to answer specific experimental questions, the collective data offered the opportunity to search for coordinated

gene expression patterns in a systematic exploration of the mouse ES transcriptome under a battery of different experimental settings –minimizing laboratory-specific results.

Analysis of the genomics data grouped genes according to their expression pattern during differentiation, or their response to various experimental stimuli. Collective data comparison showed that individual ES cell lines have similar transcriptomes, yet, they can vary significantly in the gene expression profiles during differentiation. The expression data have been organized in an interactive, open-access database with a number of novel features to facilitate studies regarding the biology of embryonic stem cells.

Results

Coordinated analysis of the mouse ES cell transcriptome

The design of each independent experiment in the FunGenES consortium reflects the interests and hypotheses of individual consortium groups. However, from the outset of the collaboration there was an effort to streamline techniques and standardize experimental protocols. To this end, consortium members selected three ES cell lines (*CGR8*, *E14TG2a* and *RI*) for common use; ES cell clones were karyotyped and tested by alkaline phosphatase staining (to eliminate prematurely differentiated cells) before being distributed to most of the consortium groups and a number of laboratories shared serum batches and used a common LIF source. Finally, RNA samples were prepared following the same methods and subsequent microarray analyses were performed in a central facility using Affymetrix Mouse 430 v.2 arrays.

The configuration of each of the ten experiments and the RNA samples collected are summarized in **Table 1**. In brief, the studies consist of: six analyses on gene regulation during the ES cell expansion phase, focusing for example on LIF targets, Stat3 regulated genes, etc.; and four studies where ES cells were allowed to differentiate in monolayers or as embryoid bodies, in control culture media, or in the presence of various agents including retinoic acid, bFGF and Wnt pathway activators. Detailed descriptions of the individual experiments are included in the **Supplemental File 1**. The total number of tested conditions is 67, each performed in up to six, separate, biological replicates using a total of 258 Affymetrix arrays. The validation rate of the microarray expression profiling data is 90.7%, based on results obtained independently in eleven consortium laboratories. In brief, 330 of 364 genes, tested by quantitative or conventional PCR, gave comparable expression patterns as the genomics data obtained by microarray analysis.

Identification of gene sets that cluster across all experimental conditions

The synchronized genomic analysis presented the opportunity to compare the entire consortium data to search for coordinately expressed genes in responses to various stimuli or during the normal ES cell differentiation program. We hypothesized that transcripts which follow similar expression patterns in ES cells and their derivatives under a number of diverse experimental conditions are likely to be functionally interconnected or they might belong to related genetic pathways³⁴⁻³⁶.

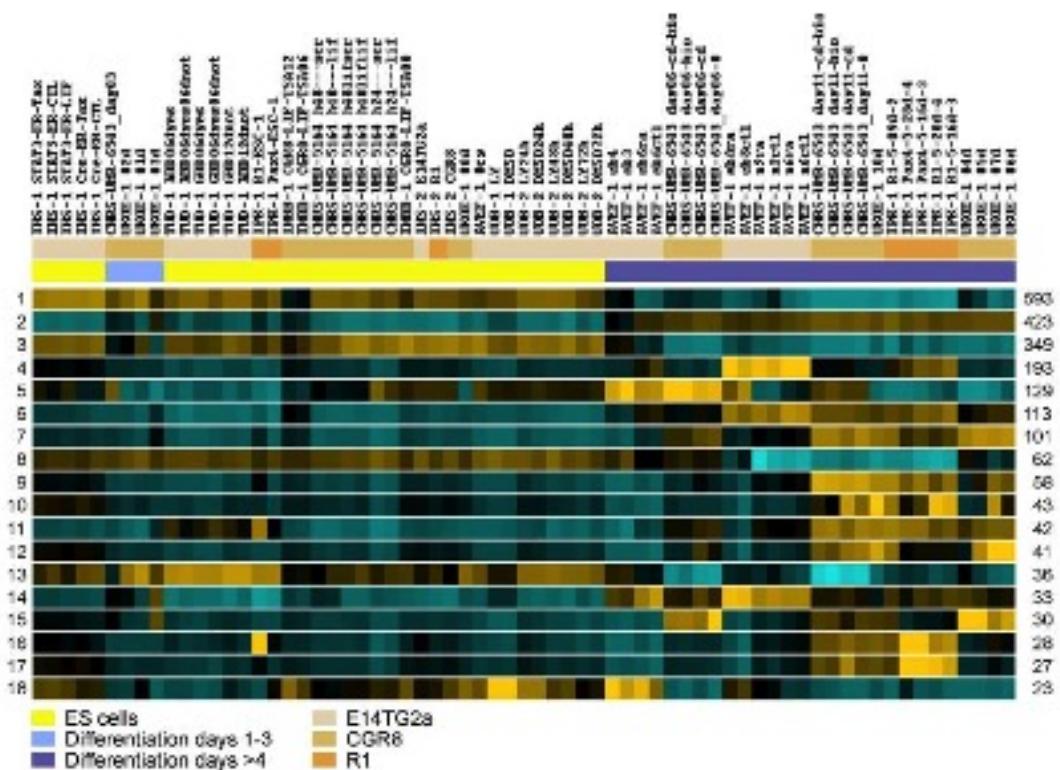


Figure 1

To explore this idea, we mined the genomics data of the consortium to identify gene sets the expression of which behaves the same way over the entire spectrum of experimental conditions. In order to facilitate interpretation of the bioinformatics outcome and enhance the biological significance of the computational data, we originally pre-selected probe sets coding for known genes. Specifically, we applied exclusion criteria to screen out transcripts without annotation and of unknown origin, as well as hypothetical transcripts or proteins. The selection reduced the number of transcripts from 45,101 to 32,020. We then removed redundant probe sets, and probe sets that showed no major differences in expression levels across all tests (standard deviation of the \log_2 signal over 67 conditions < 0.45). The selection criteria brought the number of transcripts used in the clustering analysis to 5,959.

Unsupervised hierarchical clustering of the 5,959 genes, using 100 random permutations, gave rise to 115 groups -containing a total number of 2,855 transcripts- with a probability of 95% or higher that clustering is not random. Cluster sizes range from 593 to 3 transcripts (**Supplemental File 2**). The heatmaps of the 18 largest clusters with more than 20 members are shown in **Figure 1**. The hierarchical clustering organized samples according to differentiation stages, with undifferentiated ES cells and early differentiation states at the left and progressively more differentiated states toward the right. The list of genes belonging to each cluster and the heatmaps of all 115 clusters are available in **Supplemental File 3** and in the accompanying database under the heading “Hierarchical Clusters” (<http://www.bioinf.ebc.ee/fungenes/Public/>).

Inspection of the data revealed many instances where clustered genes have been previously functionally associated. Specific examples that illustrate this point include: cluster 3, containing genes implicated in the maintenance of ES cell pluripotency such as *Nanog*, *Sox2*, *Pou5f1* (*Oct4*), *Klf2*, *Zpf42* (*Rex1*) and *Esrrb*; cluster 4, which consists mostly of genes associated with neuronal development and differentiation; cluster 8 that includes genes involved in RNA and DNA metabolism; cluster 11 with structural, enzymatic and regulatory genes of extracellular matrix deposition, e.g., *Collagen type I alpha 1*, *Collagen type III alpha 1*, *Collagen type V alpha 1*, *Mmp14*, *Timp2*, *TGFb2*, *Bmp1*, *Lysyl oxidase-like 1*, and *Integrin b1*; cluster 12, which is rich in genes involved in cardiogenesis like *Angiopoietin 1* and 2, *Mef2c*, *Myl4*, *cardiac Troponin T2*, *Tropomodulin 1*, *myosin binding protein C* and *Bves*; cluster 15 containing molecular regulators of early mesoderm development including *Bmp2*, *Bmp5*, *Msx1*, *Msx2*, *Cripto*, *Tbx20*, *Hey2*, *Smad6*, *Vegfr2 (Kdr)*, *Foxf1* and *Hand1*; cluster 20,

which comprises regulatory and structural genes linked to hemopoiesis such as *Gata1*, *Nfe2*, *Klf1*, *Tie1*, hemoglobins (*Hba-x*, *Hbb-b1*) and *Glycophorin A*; cluster 28 that consists almost entirely of genes encoding leucyl-, alanyl-, isoleucine-, methionine-, cysteinyl- and asparaginyl-tRNA synthases; cluster 30, the entire gene content of which appears during the gastrulation phase of embryogenesis and is involved in the formation of the three embryonic germ layers, e.g., *Goosecoid*, *Cerberus 1 homolog*, *Wnt3*, *Mesp1*, *Mixl1*, *mEomes* and *Even-skipped 1*; cluster 31 consisting of proteins involved in Ca^{2+} binding and homeostasis like *Gelsolin*, *Cystallin alphaB*, *Annexin A1* and *A3*, *S100 calcium binding protein*, and *thrombospondin*; cluster 35 that comprises *Hox* genes involved in the establishment of neural and somatic segmentation; cluster 44 with genes taking part in lipid biosynthesis; cluster 52 with intermediate filament encoding genes; cluster 54 with proteosome associated proteins; cluster 60 consisting of homeobox genes involved in lumbar spinal neuron development; and cluster 82, which is composed of transcription factors *Gata-4* and *Gata-6* together with their known target *Pdgrfa*³⁷.

It is likely that genes in individual clusters represent interconnected components of molecular and cellular pathways, or are targets of the same regulatory mechanisms, or both. Next to well-described genes, clusters often contain transcripts the function of which is either poorly understood or it has not been associated with the specific biological process. It is possible that a number of the latter transcripts are functionally related to previously characterized genes in the corresponding clusters providing novel unexplored links to known molecular pathways.

Identification of common and divergent components among expression data sets

To better illustrate the general structure of the entire matrix of 67 experiment conditions in relation to the expression pattern of the 5,959 clustered genes, we used the Principle Component Analysis (PCA) approach. PCA arranges expression data sets of individual samples in relation to each other and separates them in a multidimensional space where each dimension corresponds to a variance; every variance is based in the expression of particular gene clusters the nature of which may provide insights in the genetic pathways that are common or differ between time points, cell lines, or differentiation techniques.

The arrangement of the 67 experimental conditions or data sets in the first two dimensions representing the two most significant variances is shown in **Figure 2A**. As illustrated in the graph, samples that represent differentiated states separate along the x-axis, which represents

Principal Component 1 (PC1). Differentiating ES cells become increasingly divergent from the ES cell stage as they advance down the differentiation path. In contrast, gene expression data from undifferentiated ES cells, or ES cells exposed to different experimental stimuli congregate in one area suggesting that manipulation of ES cells under the particular experimental conditions (described in Table 1 and Supplemental File 1) does not lead to widespread induction of differentiation pathways. Of note, the most divergent ES samples were those obtained after treatment with the deacetylase inhibitor Trichostatin A, which causes genome-wide epigenetic modifications (Karantzali et al. in press).

Interestingly, both *E14TG2a* and *CGR8*, the two lines that data were available for most of the early and middle stages of differentiation (Days 0-11), show similar distancing from the ES cell state along the x-axis indicating that the expression of gene clusters supporting PC1 follow parallel changes in both lines. However, the opposite is true for PC2 in the y-axis where the differentiation stages of *E14TG2a* and *CGR8* lines greatly diverge suggesting that a number of gene clusters are specifically expressed in *E14TG2a* or *CGR8*. Interestingly, the *R1* line appears to be more similar to *CGR8* early on, whereas its gene expression profiles become closer to *E14TG2a* at later stages (**Figure 2A**).

To identify the gene clusters responsible for the arrangement in PC1 and PC2, we plotted the component loading of the 30 largest clusters. As shown in **Figure 2B**, the strongest contributors in PC1, are clusters 1, 3, 8, 13, 16, 24, 25 and 28, representing down regulated genes during ES cell differentiation, and clusters 2, 6, 9, 11, 14, 22, 23, and 26 encompassing genes that are upregulated during differentiation. As it shown from the heat maps in Figure 1 and in the database, both cluster types represent genes commonly regulated in both ES cell lines. Conversely, principal component 2 is mainly supported by gene clusters 4, 12, 15 and 20 -with cluster 4 predominantly expressed in *E14TG2a* cells during differentiation, whereas clusters 12, 15 and 20 are expressed in *CGR8* cells.

It is likely that gene clusters driving PC1 represent common pathways of differentiation between the two ES cell lines, i.e., a core differentiation program, whereas those that contribute to PC2 reflect the divergent differentiation potential of *CGR8* and *E14TG2a* cells. Consistent with this idea, clusters contributing to PC1 contain genes that take part in cell growth, replication, RNA, DNA, and protein synthesis, signaling cascades (e.g., various kinases), cell adhesion and extracellular matrix deposition (Table 2). In contrast, PC2 is

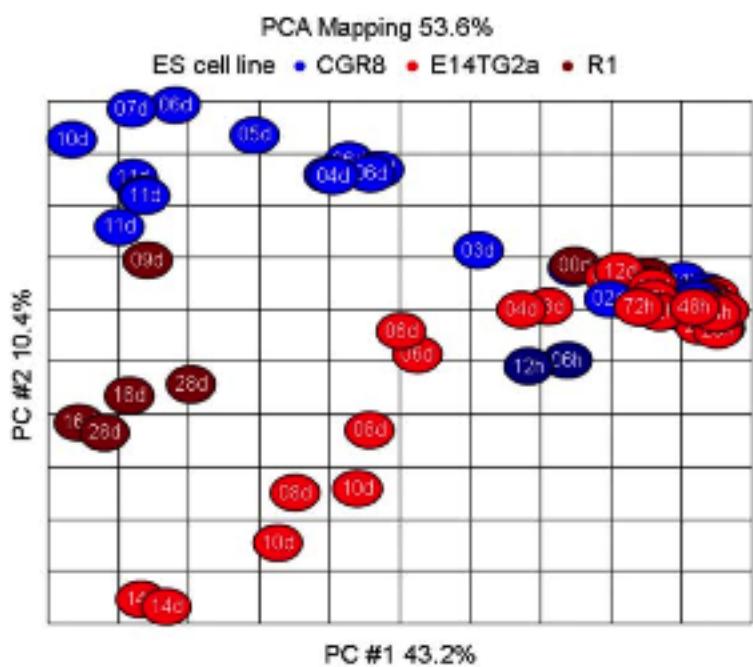


Figure 2A

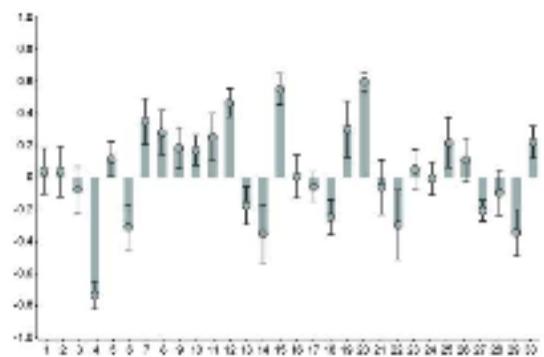
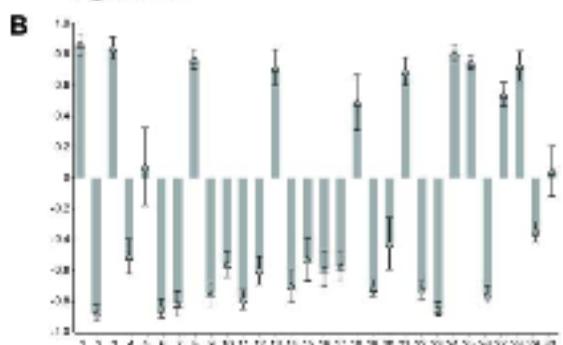


Figure 2B

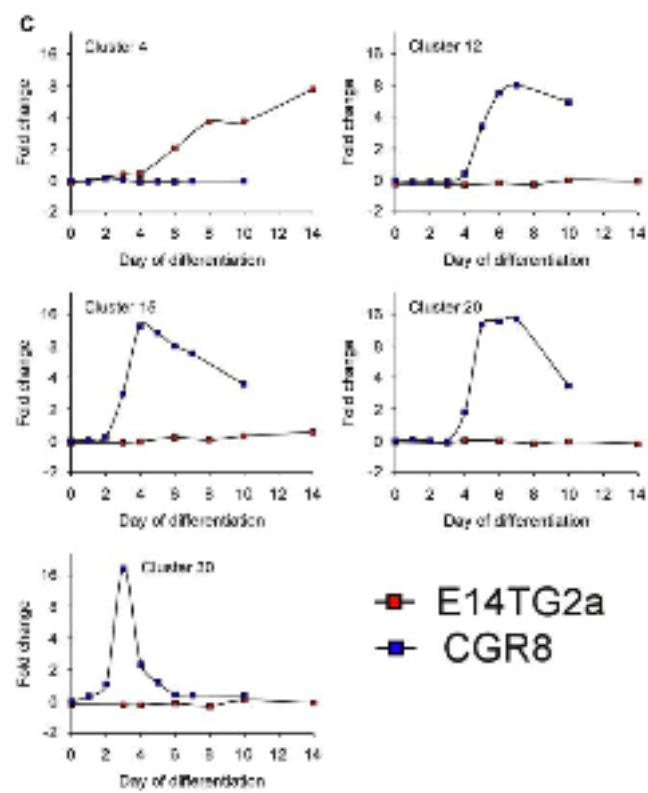


Figure 2C

driven by clusters containing neuronal-specific genes, which are expressed primarily in differentiating *E14TG2a* cells versus genes involved in cardiovascular and hematopoietic development, which predominate in *CGR8* cells. The contrasting expression patterns of clusters 4, 12, 15 and 20 in *CGR8* and *E14TG2a* cells are depicted in **Figure 2C**.

Of note, the earliest gene cluster that shows significant expression differences is cluster 30. As seen in **Figure 2C**, cluster 30 genes fail to express in *E14TG2a* cells, whereas there is a pronounced transient expression in *CGR8* cells at day 3 of differentiation. As mentioned above, genes in cluster 30 take part and regulate the process of gastrulation, a critical stage for the generation of mesendoderm during development. The failure to launch the “gastrulation gene expression program” in *E14TG2a* cells may explain the subsequent paucity in the expression of genes for mesodermal lineages like blood, heart and endothelium. Instead, the profiling data would suggest that *E14TG2a* are following a default pathway, which leads to differentiation of ectodermal lineages and expression of neuronal genes. At present we do not know whether this is an inherent difference between the two cell lines, or if it reflects the particular conditions employed during differentiation.

Changes in gene expression profiles during the early phases of ES differentiation

A selective illustration of gene expression changes during the early stages (days 0 to 5) of differentiation in *CGR8* cells, organized in three phases, is shown in **Figure 3**. The early phase, which begins immediately after LIF removal and the start of differentiation, is characterized by a swift and extensive loss of transcription factors, many of which have been implicated in regulation and maintenance of pluripotency. This group includes *Nanog*, *Oct4*, *Klf2*, *Sox2* and *Sall4* (**Figure 3**). Expression of *Klf2*, *Klf3*, *Klf4*, *Klf5*, *Gli1*, *Gli2* declines most rapidly while the expression of other genes drops gradually disappearing by day 4, e.g., *Klf9*, *Foxd3*, *Nanog*, *Oct4* and *Pcaf*. Concomitantly, this retreating group is replaced by a new set of transcriptional regulators including *Lef1*, *Tcf7*, *Myb*, *Sox3*, *Sall2* and *Hes6*.

A second phase centered on day 3 of differentiation is characterized by the robust and transient induction of genes that have been associated with the gastrulation stage of embryogenesis³¹, including *Goosecoid*, *Even-skipped 1*, *Cerberus 1 homolog*, *Wnt3*, *Mesp1*, *Mixl1* and *mEomes*. As the expression of gastrulation-associated genes subsides, there is an induction of a new set of transcription factors, including *Gata-1*, *Gata-4*, *Gata-5*, *Gata-6*, *Hand1*, *Hand2*, *Mef2a*, *FoxA1*, *FoxF1*, *Hoxa1*, *Hoxb2*, which are known to regulate early

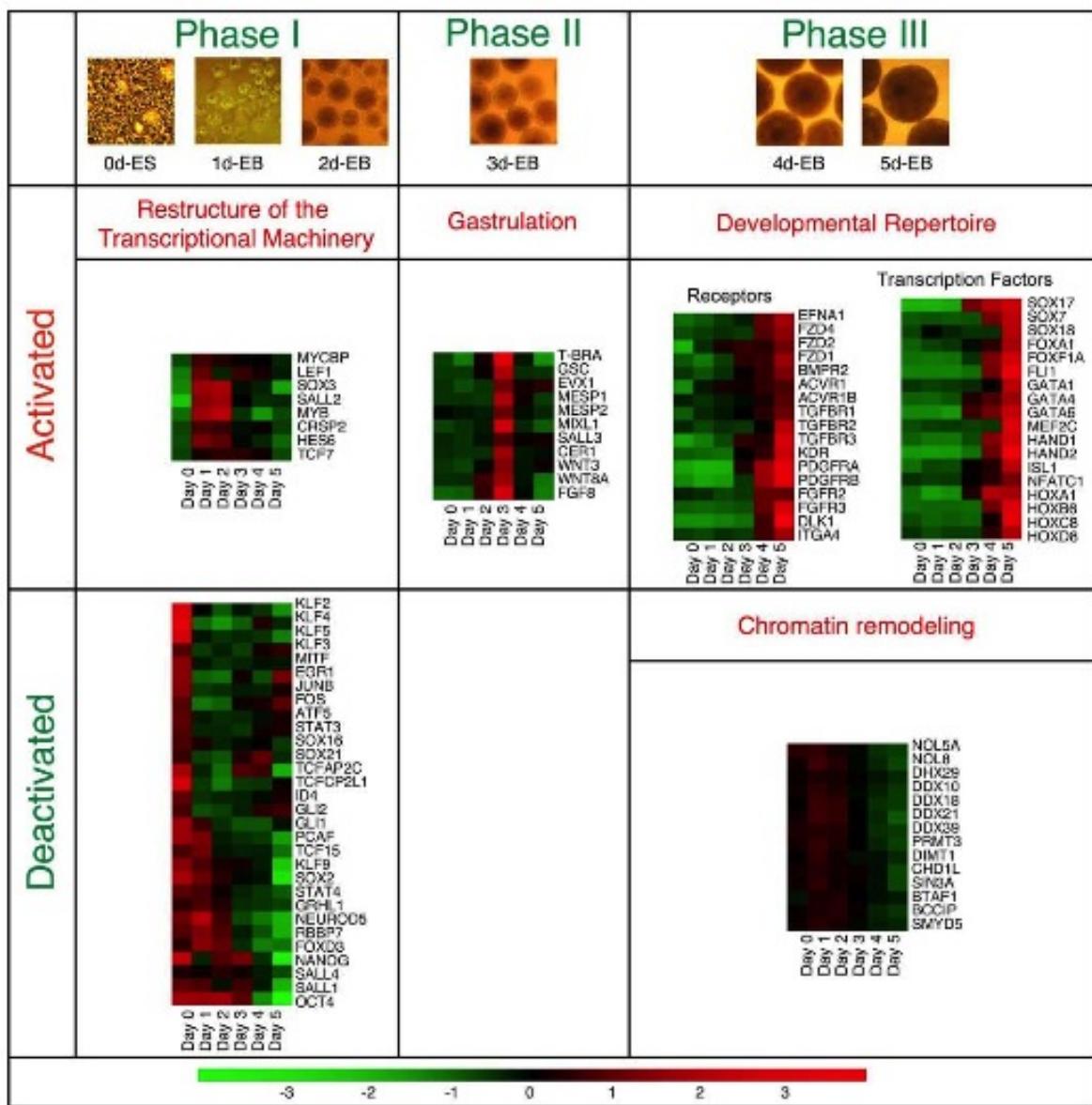


Figure 3

differentiation stages of hemopoiesis, cardiogenesis and vasculogenesis, as well as gut development and neurogenesis (**Figure 3**). This phase is also characterized by an enrichment in a wide variety of cell membrane receptors for the Wnt pathway (*Fzd1*, *Fzd2*, *Fzd4*), TGF β /BMP signaling (*Tgf β r1*, *Tgf β r2*, *Tgf β r3*, *Bmpr2*) and tyrosine kinase receptors (*Fgfr3*, *Pdgfra*, *Pdgfr β* , *Vegfr2*) endowing differentiating cells with a collection of receptors and regulatory molecules that take part in major signal transduction pathways.

Other characteristic events taking place at the molecular level in the third phase include down regulation of a large set of genes the products of which take part in chromatin remodeling, nucleolar structures and RNA processing^{32,33} and induction of genes encoding extracellular matrix proteins (collagens, laminins) and their cellular receptors (e.g., integrins; not shown).

Organizational design and special features of the FunGenES Database

To enhance the analytical power of the collected information, facilitate data mining, and provide public access of the consortium results and tools to the scientific community, we have organized the FunGenES genomics data in an open-access, interactive database adding several novel features (**Figure 4**). Besides the original 115 clusters described above, we performed a second round of k-means clustering analyses with the following modifications: (1) transcripts were clustered by time of expression; (2) we used the data from a subset of consortium samples representing untreated states without additional stimuli (26 probe sets listed in **Supplemental File 4**); and, (3) we included only transcripts with sufficient differential expression (standard deviation >0.45). The resulting “Time Series clustering” contains 8,211 genes, which have been organized in either 50 (Concise) or 200 (Analytical) clusters.

We have also carried out clustering of genes related to transcriptional activation to assist searches of interconnected circuits of gene expression regulators. To aid discovery of new molecules, we have analyzed ESTs separately to distinguish the ones that are expressed in ES cells or during the differentiation process. From approximately 12,000 ESTs, only 1,027 showed a specific expression pattern (8.6% of all ESTs present in the 430 2.0 microarray). This is in contrast to known genes where 21% had a particular pattern. This indicates that many ESTs in the microarrays might be cloning artifacts. However, the remaining 1,027 ESTs represent novel transcripts with potentially important functions in stem cell biology and

embryonic development. The 1,027 ESTs have been grouped in 50 clusters based on their timing of appearance; about half are expressed specifically in ES cells, the rest in differentiated counterparts. Transcription factor and EST clusters can be accessed in the “Specific Clusters” window of the FunGenES Database.

In the tools described above, the list of genes belonging to a cluster, together with the heatmaps of individual transcripts, appear by clicking on the corresponding cluster. Subsequent marking of any gene, or groups of genes, opens a link to the full gene name, its identification number and to g:Profiler that provides functional annotation to quickly assess the biological background of genes with interesting expression patterns³⁸. Terms of description include GO categories, KEGG and Reactome pathways, miRBase microRNA data, and TRANSFAC motifs. In addition to functional explanations, the g:Profiler provides convenient tools for dealing with different gene identifiers and finding orthologs from other organisms.

To better visualize co-regulated genes, the data have been organized as graphs representing the average levels of every gene that belongs to a particular expression pattern (“Expression waves”). The names of genes following this specific pattern are included below the graph. The corresponding graph representing genes expressed in the opposite manner is available on the same page for side-by-side comparisons. This way it is possible to search the database for groups of potentially interconnected genes as a starting point to decipher regulatory networks of transcription factors, signaling molecules and membrane receptors or for indications of genes that might be co-regulated by the same genetic pathways.

To maximize the analytical power of the database, we included an option to display the expression profile of any gene of interest, or group of genes, across the entire spectrum of the expression profiling data by using standard abbreviated gene names or any identifier supported by the Ensembl database [“Search your gene(s) of interest”]. This tool allows electronic analysis of the expression profile of any gene, or genes, in mouse ES cells and during the subsequent stages of differentiation and provides a fast assessment of expression profiling data obtained by RT-PCR or other techniques.

To examine the action of individual pathways *in toto* during the embryonic stem cell differentiation process, the database was given an additional feature to depict dynamic

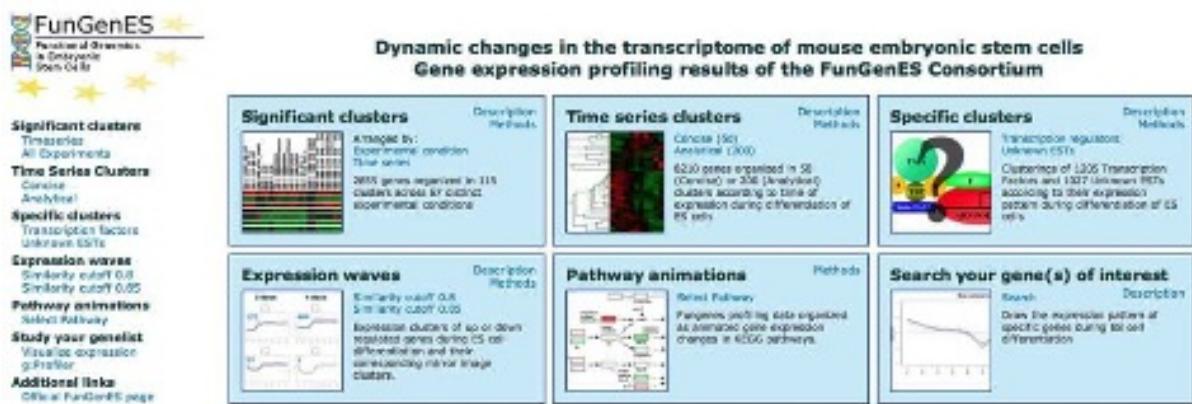


Figure 4

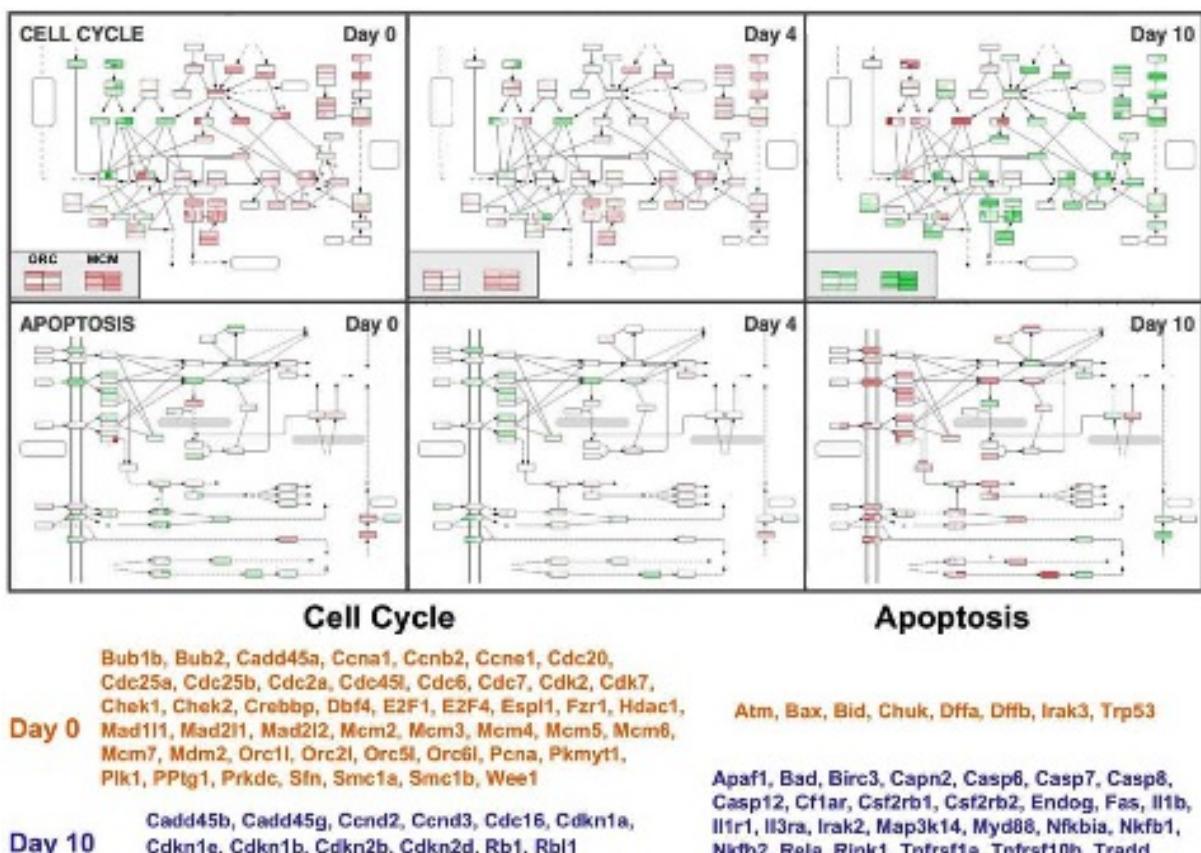


Figure 5

changes in specific genetic, signaling or metabolic pathways viewed in time-related animations based on the KEGG annotation³⁹ (“Pathway animations”). **Figure 5** depicts a stationary view of the KEGG pathways for Cell Cycle and Apoptosis at three time points; ES cells (day 0) have higher numbers of expressed genes involved in cell cycle (rectangles appearing in red color) compared to differentiated cells (day 10). Almost all of the genes expressed at day 0 have been silenced by day 10 (green) and replaced by a new set of genes.

The extensive changes in the expression profile from ES cells (Day 0) to differentiated cells at Day 10 is suggestive of a dramatic overhaul of the self-renewal machinery. The Cell Cycle animated pathway shows that genes encoding regulators of DNA replication are expressed at high levels in self-renewing ES cells and are progressively down regulated during differentiation. They include genes of the origin of replication complex (*orc*), the minichromosome maintenance (*mcm*), and the cell division cycle (*cdc*) families. Genes involved in DNA damage control and inhibition of DNA synthesis⁴⁰, in particular *Atm*, *Chk1* and *Chk2*, are also highly expressed in self-renewing ES cells, but decline during differentiation. These changes are indicative of the active replication machinery and the tightly controlled replication fidelity in proliferating ES cells⁴¹.

Self-renewing ES cells are also characterized by elevated levels of transcripts encoding the G1/S transition-promoting complex cyclin E1:Cdk2 and, by contrast, low levels of transcripts encoding D-type cyclins and Cdk4/6 inhibitors of the INK4 family (p15, p16, p18, p19; **Figure 5**). Differentiation is associated with a decrease in cyclin E1 and a concomitant elevation in D-type cyclins and cdk inhibitor transcripts⁴². These results are likely due to the progressive switch from a cyclin E-based autonomous cell cycle, which characterizes self-renewing ES cells, to the D-type cyclins/Retinoblastoma (Rb) protein-regulated cell cycle⁴³.

Conversely, few pro-apoptotic genes are expressed in ES cells (day 0; most boxes appear in green), but many are gradually induced during the differentiation process showing the exact opposite pattern from the genes involved in cell proliferation (**Figure 5**). As observed for cell-cycle genes, there is minimal overlap between apoptosis-associated genes expressed at day 0 and 10. This strikingly complementary pattern suggests a reciprocal interrelationship between the balance of pro- and anti-apoptotic genes in ES cells and their differentiated progeny.

Discussion

Genomic studies, which in principle group together co-regulated genes, have the potential to discover novel components of known regulatory pathways in ES cell biology that can be subsequently validated in functional studies. In recent years, genome-wide analysis combining Chromatin Immuoprecipitation (ChIP) and array hybridization (ChIP on chip) revealed that both active and silenced genes are directly bound in ES cells by one or more of the core pluripotency factors Oct4, Sox2 and Nanog (Boyer, Loh). Recently, functional analysis using loss-of-function and protein-protein interaction approaches and bioinformatics tools have began to piece together regulatory networks active in ES cells Ivanova 2006 , (REF 13, Zhou , Chipperfield. Melton and Wong).

It is also likely that the actual core factor set in ES cells is larger and more highly interconnected than previously suspected. Kim et al have performed a genome-wide analysis of target promoters for nine transcription factors, namely Oct4, Sox2, Klf4, c-Myc, Nanog, Dax1, Rex1, Zpf281, and Nac1. They found that target promoters bound by a single or few factors tend to be inactive or repressed, whereas promoters bound by more than four factors are active in the pluripotent state and become repressed upon differentiation. Interestingly, targets of Myc or Rex1 are implicated in protein metabolism, rather than in developmental processes, whereas targets of the other factors are enriched in genes for developmental processes. They established a hierarchy within the important pluripotency factors such that Klf4 serves as an upstream regulator of feed-forward circuits involving Oct4, Sox2, Nanog and Myc.

The increasing complexity of gene regulatory networks emerging from these studies, combined with the surging amount of genomics and proteomics work, underscore the need for innovative resources to mine available and future data by the scientific community. The FunGenES database provide a template for such resources with a number of innovative tools such as Animation Pathways, Expression Waves, Expression Time Series, Specific Clusters of gene classes such as ESTs or transcription factors and searches for the expression pattern of any gene or transcript during ES cell differentiation using standard names and IDs. Importantly, gene search results are linked to comprehensive annotation tools including the presence of common regulatory motifs in promoter areas and miRNA targeting information.

The analytical paradigms and expression pattern clusters presented here provide a scaffold for future comparative analysis of human ES cell lines. This information will be particularly important for future analyses of ES-like cells derived from somatic tissues that can be potentially used to derive pancreatic cells, cardiomyocytes or neurons for organ regeneration^{28,44-46}. We expect that not only regenerative medicine applications, but also basic science studies will benefit from the resources described here, especially when compared to expression profiling data obtained from loss- and gain-of-function approaches^{10,47}. Furthermore, the “new associations” of ESTs and genes to specific pathways provide a fresh collection of novel components that can be further explored in functional assays during embryonic development and in human diseases.

Experimental procedures

RNA isolation and microarray hybridization

Total RNA was isolated using the RNeasy Mini Kit from QIAGEN and treated with RNase-free DNase I (5 U / 100 μ g of nucleic acids, Sigma). Biotinylated cRNA was prepared according to the standard Affymetrix protocol (Genechip Expression Analysis Technical Manual, Affymetrix). In brief, double-stranded cDNA was synthesized from 10 μ g total RNA using the SuperScript Choice System (Invitrogen) and the Affymetrix T7-(dT)₂₄ primer. Following a phenol/chloroform extraction and ethanol precipitation, the cDNA was transcribed into biotin-labeled cRNA using the Retic Lysate IVT™ kit (Ambion). cRNA products were purified using the RNeasy kit (QIAGEN) and fragmented to an average size of 30-50 bases according to Affymetrix recommendations. 15 μ g of fragmented cRNA were used to hybridize to the Mouse Genome 430 2.0 Array for 16 hr at 45 °C. The arrays were washed and stained in the Affymetrix Fluidics Station 400 and scanned using the Hewlett-Packard GeneArray Scanner G2500A. The image data were analyzed with the GeneChip® Operating Software (GCOS) using Affymetrix default analysis settings. After a quality control test, arrays were normalized by the log scale robust multi-array analysis (RMA)⁴⁸.

We used 258 Affymetrix GeneChips to analyse the 67 individual experimental conditions outlined in **Table 1**. A detailed description of the individual experiments is provided in **Supplemental File 1** and the FunGenES Database.

Data preparation for unsupervised hierarchical clustering and principal component analysis

The first step in the data analysis was to average the biological replicates for each of the 67 experimental conditions. To identify genes which cluster together under the tested conditions, we excluded probe sets with a standard deviation in expression values of $> \log_2(0.45)$ from the vector mean. We then removed redundant gene/probe sets taking into account the ENTREZ, Unigene and RefSeq gene-id annotations. Among redundant probe sets, we selected the probe set with the highest average expression signal. We also took out probe sets of unknown origin, for example RIKEN sequences, or hypothetical transcripts/proteins. These criteria led to a dataset of 5,959 transcripts.

Unsupervised hierarchical clustering

Correlation of differentially expressed transcripts was detected by hierarchical clustering of expression values with the Cluster version 2.11 software³⁴ applying mean centering and normalization of genes and arrays before the computational clustering analysis. Average linkage hierarchical clustering of the data was carried out as described⁴⁹.

Permutations

The correlation significance of expression profiles between probe sets was assessed empirically by one hundred rounds of random permutations. For each round, the 67 values for each probe set were randomly redistributed and datasets clustered as described⁵⁰. The best similarity scores of each permuted and clustered dataset was collected to evaluate the 95th percentile of significant clusters in the original dataset. 5% of the permuted datasets gave rise to clusters containing no more than two genes at a similarity score > 0.85071 (**Supplemental File 5**). Clusters with 3 or more genes (115 clusters) were documented and selected for further analysis.

The genes in the 115 clusters were also organized according to average expression values at distinct differentiation time points without the influence of external stimuli. The selected 26 experimental conditions used for this analysis are depicted in **Supplemental File 4**.

Principal component analysis

Principal component analysis of 67 experimental conditions by 5959 transcripts has been performed using Partek Genomic Suite 6.3 beta, correlation as a dispersion matrix and normalized Eigenvector scaling.

Clustering of the data in time series

Besides the clustering described above that was based on the entire spectrum of experimental conditions, the original FunGenES expression data were also clustered according to timing of expression in a two-step strategy. First, probe sets were clustered with k-means into a small number of clusters using chord distance (Euclidean distance over vectors normalized to unit sphere). In a second step, the resulting clusters, represented by mean profiles, were clustered using average link hierarchical clustering with Pearson correlation distance measure, and visualized in heatmap representation³⁴. No filtering besides removing genes with low variation was applied to these datasets.

Expression waves

We developed a method to identify all genes that have characteristic expression patterns during ES cell differentiation. In brief, transcripts were included into a particular expression wave, if its correlation with the specific pattern was higher than a certain threshold and also highest among all other patterns. This analysis was done in two different stringent conditions with correlation thresholds of 0.8 and 0.85. The results are presented in a serie of graphs, with the list of genes that belong in the depicted pattern identified below. Each graph is juxtaposed to its “mirror image”, i.e., the graph representing genes that behave exactly the opposite way.

Pathway animations

To better visualize dynamic changes in entire genetic and signaling cellular circuits, we designed a set of animations based on pathways from the Kyoto Encyclopedia of Genes and Genomes³⁹ (KEGG). The animations use diagrams available at the KEGG webpage, which portray connections between pathway components. The expression levels of relevant genes are shown in the diagrams by the standard red (high) – green (low) color codes. In case a gene family represents a particular function, the corresponding box displays the expression patterns of all individual members of the family in adjacent stripes.

The “Search your gene(s) of interest” tool

1. This tool allows investigators to search and display the expression of any probe set across the entire spectrum of the FunGenES datasets. The program draws either clustered heatmaps or line graphs of expression. The linear graphs are drawn using statistics package R. The tool recognizes common gene identifiers; the conversion to Affymetrix probe set ids is done using Ensembl BioMart mappings. The heatmap representation is linked to the recently developed g:Profiler web toolkit for functional annotations of selected genes³⁸.

References

1. Rando, T. A. Stem cells, ageing and the quest for immortality. *Nature* **441**, 1080-1086 (2006).
2. Singec, I., Jandial, R., Crain, A., Nikkhah, G. & Snyder, E. Y. The leading edge of stem cell therapeutics. *Annu. Rev. Med.* **58**, 313-328 (2007).
3. Dimmeler, S., Zeiher, A. M. & Schneider, M. D. Unchain my heart: the scientific foundations of cardiac repair. *J. Clin. Invest.* **115**, 572-583 (2005).
4. Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156 (1981).
5. Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl Acad. Sci. USA* **78**, 7634–7638 (1981).
6. Smith, A. G. Embryo-derived stem cells: of mice and men. *Annu. Rev. Cell Dev. Biol.* **17**, 435-462 (2001).
7. Thomas, K. R. & Capecchi, M. R. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**, 503-512 (1987).
8. Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W. & Kemler, R. The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J. Embryol. Exp. Morphol.* **87**, 27-45 (1985).
9. Keller, G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev.* **19**, 1129-1155 (2005).
10. Ivanova, N. B. *et al.* A stem cell molecular signature. *Science* **298**, 601-604 (2002).
11. Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R. C. & Melton, D. A. "Stemness": transcriptional profiling of embryonic and adult stem cells. *Science* **298**, 597-600 (2002).
12. Sato, N. *et al.* Molecular signature of human embryonic stem cells and its comparison with the mouse. *Dev. Biol.* **260**, 404-413 (2003).
13. Ivanova, N. *et al.* Dissecting self-renewal in stem cells with RNA interference. *Nature* **442**, 533-538 (2006).
14. Assou, S. *et al.* A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. *Stem Cells* **25**, 961-973 (2007)
15. Tu, Y., Stolovitzky, G. & Klein, U. Quantitative noise analysis for gene expression microarray experiments. *Proc. Natl Acad. Sci. USA* **99**, 14031-14036 (2002).
16. Schöler, H. R., Hatzopoulos, A. K., Balling, R., Suzuki, N. & Gruss, P. A family of

- octamer-specific proteins present during mouse embryogenesis: evidence for germline specific expression of an Oct factor. *EMBO J.* **8**, 2543-2550 (1989).
- 17. Okamoto, K. *et al.* A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell* **60**, 461-472 (1990).
 - 18. Nichols, J. *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379-391 (1998).
 - 19. Niwa, H., Miyazaki, J. & Smith, A. G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature Genet.* **24**, 372-376 (2000).
 - 20. Mitsui, K. *et al.* The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631-642 (2003).
 - 21. Chambers, I. *et al.* Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643-655 (2003).
 - 22. Loh, Y. H. *et al.* The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature Genet.* **38**, 431-440 (2006).
 - 23. Avilion, A. A. *et al.* Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* **17**, 126–140 (2003).
 - 24. Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
 - 25. Boiani, M. & Schöler, H. R. Regulatory networks in embryo-derived pluripotent stem cells. *Nature Rev. Mol. Cell Biol.* **6**, 872-884 (2005).
 - 26. Zhang, J. *et al.* Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nature Cell Biol.* **8**, 1114-1123 (2006).
 - 27. Li, Y. *et al.* Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. *Blood* **105**, 635-637 (2005).
 - 28. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
 - 29. Wang, J. *et al.* A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364-368 (2006).
 - 30. Galan-Caridad, J. M. *et al.* Zfx controls the self-renewal of embryonic and hematopoietic stem cells. *Cell* **129**, 345-357 (2007).
 - 31. Tam, P. P. & Loebel, D. A. Gene function in mouse embryogenesis: get set for

- gastrulation. *Nat. Rev. Genet.* **8**, 368-381 (2007).
- 32. Tsai, R. Y. L. & McKay, R. D. G. A nucleolar mechanism controlling cell proliferation in stem cells and cancer cells. *Genes Dev.* **16**, 2991-3003 (2002).
 - 33. Xi, R. & Xie, T. Stem cell self-renewal controlled by chromatin remodeling factors. *Science* **310**, 1487-1489 (2005).
 - 34. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA* **95**, 14863- 14868 (1998).
 - 35. Brown, M. P. *et al.* Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proc. Natl Acad. Sci. USA* **97**, 262-267 (2000).
 - 36. Wu, L. F. *et al.* Large-scale prediction of *Saccharomyces cerevisiae* gene function using overlapping transcriptional clusters. *Nat. Genet.* **31**, 255-265 (2002).
 - 37. Wang, C. & Song, B. Cell-type-specific expression of the platelet-derived growth factor alpha receptor: a role for GATA-binding protein. *Mol. Cell. Biol.* **16**, 712- 723 (1996).
 - 38. Reimand, J., Kull, M., Peterson, H., Hansen, J. & Vilo, J. gProfiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucl. Acids Res.* **35**, W193-W200 (2007).
 - 39. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucl. Acids Res.* **28**, 27-30 (2000).
 - 40. Bartek, J. & Lukas, J. Pathways governing G1/S transition and their response to DNA damage. *FEBS Lett.* **490**, 117-122 (2001).
 - 41. Saretzki, G, Armstrong, L., Leake, A., Lako, M. & von Zglinicki, T. Stress defense in murine embryonic stem cells is superior to that of various differentiated murine cells. *Stem Cells* **22**, 962-971 (2004).
 - 42. Savatier, P., Lapillonne, H., van Grunsven, L. A., Rudkin, B. B. & Samarut, J. Withdrawal of differentiation inhibitory activity/leukemia inhibitory factor up-regulates D-type cyclins and cyclin-dependent kinase inhibitors in mouse embryonic stem cells. *Oncogene* **12**, 309-322 (1996).
 - 43. White, J. *et al.* Developmental activation of the Rb-E2F pathway and establishment of cell cycle-regulated cyclin-dependent kinase activity during embryonic stem cell differentiation. *Mol. Biol. Cell* **16**, 2018-2027 (2005).
 - 44. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313-317 (2007).

45. Wernig, M. *et al.* In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318-324 (2007).
46. Maherali, N. *et al.* Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. *Cell Stem Cell* **1**, 55-70 (2007).
47. Walker, E. *et al.* Prediction and Testing of Novel Transcriptional Networks Regulating Embryonic Stem Cell Self-Renewal and Commitment. *Cell Stem Cell* **1**, 71-86 (2007).
48. Irizarry, R. A. *et al.* Summaries of Affymetrix GeneChip probe level data. *Nucl. Acids Res.* **31**, DOI: 10.1093/nar/gng015 (2003).
49. Gasch, A. P. & Eisen, M. B. Exploring the conditional coregulation of yeast gene expression through fuzzy k-means clustering. *Genome Biol.* **3**, 1-22 (2002).
50. Yvert, G. *et al.* Trans-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcription factors. *Nature Genet.* **35**, 57-64 (2003).

Acknowledgements

The FunGenES Integrated Project was funded by the European Framework Programme 6.

Author Contributions

H.S. did the global data clustering analysis and compiled the graphs of super clusters. R.K. developed most of the public FunGenES Database features. N.H. directed the Affymetrix GeneChip processing facility. J.V. managed the bioinformatics team that built the public FunGenES Database. FunGenES principal investigators (M.B., H.B., F.B., C.D., L.F., R.F., A.K.H., D.H., H.H., J.H., A.K., L.P., A.S., P.S., F.S., M.J.W., A.W.) designed, took part and oversaw the individual experiments. Researchers (K.A., N.B., X.D., O.F., P.-Y.B., I.A., M.G., E.K., S.L., A.R., M.P.S., M.T.) performed ES cell cultures, RNA preparations and validated microarray data. P.A., J.R., J.W. contributed to the bioinformatics analysis. A.K.H. designed and supervised the global analysis of the FunGenES gene expression profiling data and wrote the manuscript draft. Results were discussed among all Consortium partners. Authors are included in alphabetical order except the first and the senior authors.

Correspondence and requests for materials should be addressed to antonis.hatzopoulos@vanderbilt.edu. Contact information for FunGenES Consortium PIs and participants is included in **Supplemental File 1**.

Table 1. Outline of the ten experiments of the FunGenES consortium

Sample ^a	Experiment ^b	ES clone ^c	Conditions ^d	Repeats ^e
INS-2	Comparison of ES clones	E14TG2a/CGR8/ R1	3	5
INS-1	Stat-3 targets in ES cells	E14TG2a	5	4-6
CNRS-UMR-5164	LIF targets in ES cells	CGR8	6	5
UOB-1/UOB-2	PI3-K targets in ES cells	E14TG2a	8	3-5
IMBB-1	TSA effects on ES cells	CGR8	3	3
TUD-1	Tag effects on ES cells	E14TG2a	6	3
UKOE-1	Standard ES differentiation	CGR8	9	3
AVEF-1	ES differentiation under neurogenic conditions	E14TG2a	11	4 ^f
CNRS-UMR-6543	ES differentiation under adipogenic conditions	CGR8	9	3
IPK-1	ES differentiation favoring the pancreatic lineage	R1	7	5

a: sample name abbreviation; **b:** short description of individual experiments; **c:** ES cell line used for the corresponding experiment; **d:** number of different conditions analysed; **e:** number of independent replicates; **f:** four replicates except sample AVEF-1 eb4 (two repeats).

Table 2. Gene content of clusters contributing to PC1 and PC2

Cluster	Gene number	Gene content	Down	Up
Principal Component 1				
1	593	Cellular machinery	✓	
2	423	Development		✓
3	349	Transcription	✓	
6	113	Cell adhesion		✓
8	62	Cell Cycle	✓	
9	58	Collagen - ECM		✓
11	42	Collagen - ECM		✓
13	36	Cell cycle house keeping	✓	
14	33	Transcription		✓
16	28	Signaling	✓	
22	14			✓
23	13			✓
24	13	Ribosomal	✓	
25	13	Replication/Cell cycle	✓	
26	13			✓

28 12 RNA/protein
machinery ✓

Principal Component 2

Cluster	Gene number	Gene content	E14TG2a	CGR8
4	193	Neuronal	✓	
7	101	Cardiac, Endothelial, Blood		✓
12	41	Cardiac, Blood		✓
15	30	Mesoderm		✓
20	17	Blood		✓

Figure Legends

Figure 1. Clustering analysis of the FunGenES Affymetrix microarray data. The average heatmaps of the eighteen largest clusters with at least 20 members are included. The cluster number is depicted on the left, the cluster size, i.e., the number of transcripts in the cluster, is shown on the right. Heatmap colors range from cyan (very low or no expression) over black (low/middle) to orange (high expression levels). The names of analyzed RNA samples are indicated on the top. The description of the acronyms and the experimental conditions for each sample are provided in **Table 1** and in **Supplemental File 1**, respectively. The names of the ES cell clones used in each experiment and the time of RNA isolation during the ES cell differentiation process are color-coded with bars above the heatmaps. Explanatory notes of color codes are indicated below.

Figure 2. Principal Component Analysis of the gene expression profiling datasets. **A:** Plot of the expression profiling data from the 67 experimental conditions in two dimensions according to Principal components PC #1 (in x-axis) and PC #2 (in y-axis). Samples from the different cell lines are color-coded as indicated above. The number inside each oval represents days of differentiation during EB formation, or exposure time in days or hours of ES cells to various stimuli. 43.2% of the dataset variance is due to differentiation (PC #1), another 10.4% is caused by ES cell lineage-specific expression patterns in differentiated EBs (PC #2). **B:** Component loading (y-axis), of the 30 largest clusters (x-axis) that were derived in the unsupervised hierarchical cluster analysis, to PC1 (top) and PC2 (bottom). The highest contributing gene clusters to PC #1 are 1, 3, 8, 13, 16, 24, 25 and 28 (down regulated genes during ES cell differentiation) and clusters 2, 6, 9, 11, 14, 22, 23, and 26 (upregulated). PC #2 is based on gene clusters 4 (specific to *E14TG2a* cells) and 7, 12, 15, and 20 (expressed predominantly in *CGR8* cells). **C:** Comparison of the average expression levels of clusters 4, 12, 15, 20 and 30 in *CGR8* (blue boxes) and *E14TG2a* ES cells (red boxes). The average expression differences of *E14TG2a* to *CGR8* expression values (y-axis, log₍₁₀₎ scale) are plotted against the differentiation time point (x-axis).

Figure 3. Gene expression changes during ES cell differentiation divided in three phases. Images of ES cells differentiating to embryoid bodies (EB) are shown along the top. Below, representative examples of up and down regulated genes in *CGR8* cells with the corresponding heatmaps have been assembled in three chronological phases. Phase I (left) involves a reorganization of the transcriptional machinery. The unique group of gastrulation-associated genes that appears transiently between days 2-4 of differentiation is illustrated in the middle panel as Phase II. At the end of this time period, differentiating cells become endowed with tissue-specific transcription factors and

a rich collection of membrane receptors (Phase III, right panel, top). At this time, expression of genes linked to chromatin remodeling and RNA processing is suppressed (Phase III, right panel, bottom). The heatmap scale at the bottom represents deviation from the mean expression of the corresponding gene on a log₂ scale, e.g., +3 represents 8-fold higher expression than the average level in the entire expression data.

Figure 4. Outline of the FunGenES database. The database interface is separated in six parts. The upper row windows provide entry to the different ways the FunGenES datasets have been organized: in clusters using the entire 67 experimental conditions (Hierarchical clusters, left); according to time of expression (Time series clusters) using a subset of experimental conditions (representing conditions without additional stimuli, middle); or depending on gene class, i.e., transcriptional regulators or ESTs (Specific clusters, right). The bottom row consists of three tools that offer: detailed organization of gene expression waves during the differentiation process (bottom left window); animation of KEGG pathways organized by successive days of differentiation; and, a search engine to obtain the expression of any microarray transcript (or groups of transcripts) during the ES cell differentiation process. In every entry point, buttons provide a “Description” of how the data have been organized and/or the “Methods” used to construct the various tools.

Figure 5. Snap shots from the animated Cell Cycle and Apoptosis KEGG pathways. The images depict day 0 (representing undifferentiated ES cells), and days 4 and 10 of differentiation. Each box represents a gene or a gene family that can be visualized by marking the box during animations. The different family members are represented with juxtaposed vertical rectangles within the same box. For genes having multiple probe sets in the Affymetrix microarray, the corresponding gene rectangle is split in vertical lines, each line depicting the expression level of an individual probe set. Red color marks expressed genes and green absence of detectable expression. Color intensity reflects expression levels. The table below contains the list of genes expressed at Day 0 and Day 10 from the Cell Cycle and Apoptosis pathways. ORC: origin of replication complex; MCM: mini-chromosome maintenance.

Supplemental File 1

Detailed overview of the microarray experimental designs & Contact Information

Project INS-2: comparison of three different ES cell lines

Cell lines: CGR8; E14TG2a; R1

Pierre-Yves Bourillot & Pierre Savatier

Contact: Pierre Savatier

INSERM U846

Stem Cell and Brain Research Institute

18 Avenue Doyen Lépine

69500 Bron, France

Tel: +33 4 72913442

Email: pierre.savatier@inserm.fr

Summary of the project

In order to identify differences in gene expression profiles between the three mouse ES cell lines used within the FunGenES consortium, RNA was prepared from undifferentiated CGR8, E14TG2a and R1 cells.

Abbreviation	Description	Arrays
CGR8	ES cell line CGR8	5
E14TG2a	ES cell line E14TG2a	5
R1	ES cell line R1	5

Project INS-1: Identification of Stat3 target genes

Cell lines: E14TG2a; E14(T)-Stat3-ER; E14(T)-Cre-ER

Irène Aksoy & Pierre Savatier

Contact: Pierre Savatier

INSERM U846

Stem Cell and Brain Research Institute

18 avenue Doyen Lépine

69500 Bron, France

Tel: +33 4 72913442

Email: pierre.savatier@inserm.fr

Summary of the project

The objective of this study is to identify the target genes of the STAT3 transcription factor in ES cells, making use of the hormone-dependent STAT3-ER^{T2}. STAT3-ER^{T2} is a fusion protein between STAT3 and the mutated ligand-binding domain of the human estrogen receptor. In the absence of ligand, STAT3-ER^{T2} is sequestered in the cytoplasm in an inactive state by association of the ligand-free ER^{T2} domain with the ubiquitous Hsp90 complex. In the presence of tamoxifen, STAT3-ER^{T2} is liberated and translocates into the nucleus. We made use of the pPCAGIZ vector and episomal supertransfection in E14/T cells to express STAT3-ER^{T2}.

After electroporation, cells were plated at 5x10⁵ cells per 10 cm culture dish in medium supplemented with 1,000 U/ml of LIF and cultured in the presence of 1 µg/ml of zeocine for 7 days. Resistant colonies were pooled and further propagated in selection medium for 10 days in the presence of 10 nM tamoxifen and in the absence of LIF.

STAT3-ER^{T2}-expressing ES cells do not require LIF for propagation in the undifferentiated state. They are strictly dependent upon tamoxifen for self-renewal. RNAs have been prepared from: (i) ES cells starved of LIF and tamoxifen for 24 hours; (ii) starved cells after re-stimulation with 1,000 U/ml of LIF for 1 hour; and, (iii) starved cells after re-stimulation with 100 nM of tamoxifen for 2 hours.

E14/T cells expressing Cre-ER^{T2}, a fusion protein between the Cre recombinase and the ER^{T2} domain have been also generated and used as control cells to identify -and eliminate off-target effects of STAT3-ER^{T2}. Cre-ER^{T2}-expressing ES cells require LIF for propagation in the undifferentiated state. RNA have been prepared from: (iv) ES cells starved of LIF and tamoxifen for 24 hours; and, (v) starved cells after re-stimulation with 100 µM tamoxifen for 2 hours.

Abbreviation	Description	Arrays
STAT3-ER-CTL	STAT3-ER not stimulated	4
STAT3-ER-Tax	STAT3-ER stimulated with Tax	6
STAT3-ER-LIF	STAT3-ER stimulated with LIF	5
Cre-ER-CTL	Cre-ER not stimulated	5
Cre-ER-Tax	Cre-ER stimulated with Tax	5

Project CNRS-UMR-5164: Identification of LIF target genes

Cell line: CGR8

Marina Trouillas and Hélène Boeuf

Contact: Hélène Boeuf

UMR 5164-CNRS- CIRID

Université Bordeaux 2

Bat. 1B, BP14

146 rue Léo Saignat

F-33076 Bordeaux Cedex, France

Tel: +33 5 57574633

Email: helene.boeuf@u-bordeaux2.fr

Summary of the project

The purpose of these experiments was to identify genes the expression of which is regulated by LIF in ES cells. To this end, CGR8 ES cells were grown in DMEM medium with 10% FCS with or without 20 ng/ml recombinant human LIF.

Five independent total RNA samples (Qiagen column) have been prepared from each cell growth condition. The proper LIF regulation has been confirmed by monitoring LIF-dependent activation of STAT3 using western blot with phospho-Tyr705 STAT3 on protein lysates prepared in parallel with each RNA preparation. The quality of RNA preparations was checked by RT-PCR with known LIF-dependent genes such as *Socs3*, *Klf4* and *JunB* (with *Hprt* serving as a non-regulated constant control).

Abbreviation	Description	Arrays
h24---ser	CGR8 cells grown in DMEM/10%FCS, without LIF for 24h have been re-induced for 25 minutes with DMEM/10%FCS	5
h24---lif	CGR8 cells grown in DMEM/10%FCS, without LIF for 24h have been re-induced for 25 minutes with DMEM/10%FCS/LIF	5
h48---ser	CGR8 cells grown in DMEM/10%FCS, without LIF for 48h have been re-induced for 25 minutes with DMEM/10%FCS.	5
h48---lif	CGR8 cells grown in DMEM/10%FCS, without LIF for 48h have been re-induced for 25 minutes with DMEM/10%FCS/LIF	5
h48lifser	CGR8 cells grown in DMEM/10%FCS/LIF for 48h have been re-induced for 25 minutes with DMEM/10%FCS	5
h48liflif	CGR8 cells grown in DMEM/10%FCS/LIF for 48h have been re-induced for 25 minutes with DMEM/10%FCS/LIF	5

Project UOB-1/UOB-2: Identification of PI3-K target genes

Cell line: E14TG2a

Mike P. Storm and Melanie Welham

Contact: Melanie Welham

Department of Pharmacy and Pharmacology

The University of Bath

Bath, BA2 7AY, UK

Tel: +44 1225 386428

Email: prsmjw@bath.ac.uk

Summary of the project

Aim: To identify genes regulated by PI3K signaling in ES cells.

E14tg2a cells cultivated in KO-DMEM supplemented with Knock-out serum replacement (Invitrogen) in the presence of LIF were either: (i) starved for 24h in media lacking LIF, pre-incubated for 30 minutes with 5 µM LY294002 in DMSO or DMSO alone and then treated for 2h with LIF, prior to preparation of RNA samples (UOB-1 samples); or, (ii) cultured in the continuous presence of LIF, plus DMSO alone or 5 µM LY294002 for 24 to 72h prior to RNA extraction (UOB-2 samples).

UOB-1:

Abbreviation	Description	Arrays
DMSO	DMSO + LIF 2h	5
LY	DMSO + LY294002 + LIF 2h	5

UOB-2:

Abbreviation	Description	Arrays
DMSO24h	24h DMSO	3
DMSO48h	48h DMSO	3
DMSO72h	72h DMSO	3
LY24h	24h DMSO+LY294002	3
LY48h	48h DMSO+LY294002	3
LY72h	72h DMSO+LY294002	3

Project IMBB-1: ES cells grown in the presence of LIF and the HDAC inhibitor TSA
Cell line: CGR8

Efi Karantzali and Androniki Kretsovali

Contact: Androniki Kretsovali

Institute of Molecular Biology & Biotechnology - FORTH

1527 Vassilika Vouton

GR-711 10 Heraklion, Crete, Greece

Tel: +30 2810 391191

Email: kretsova@imbb.forth.gr

Summary of the project

Aim: To investigate the effects of global histone acetylation (HAT/HDAC) in ES cell differentiation.

To assess the role of histone acetylases (HAT)/deacetylases (HDAC), CGR8 cells were cultivated in presence or absence of Trichostatin A (TSA, an HDAC inhibitor) for 0 (control), 6 and 12 hours, respectively.

Abbreviation	Description	Arrays
CGR8-LIF-TSA00	ESCs grown in presence of LIF and treated with TSA (HDAC inhibitor); 0 hours	3
CGR8-LIF-TSA06	ESCs grown in presence of LIF and treated with TSA (HDAC inhibitor); 6 hours	3
CGR8-LIF-TSA12	ESCs grown in presence of LIF and treated with TSA (HDAC inhibitor); 12 hours	3

Project TUD-1: TAg effects on DOX MIB and DOX DEX induced clones

Cell line: E14TG2a

Sandra Lubitz, Konstantinos Anastassiadis and Francis Stewart

Contact: Francis Stewart

Technische Universitaet Dresden

BioInnovationZentrum

Am Tatzberg 47

01307 Dresden, Germany

Tel: +49 351 46340129

Email: stewart@mpi-cbg.de

Summary of the project

The purpose of the experiment was to see if changes in the expression profile of the cells caused by transient expression of T-Antigen can be reversed. Two different E14TG2a lines (ABD and GBD*) were generated: both lines carry a construct containing the SV40 Large-T Antigen under the control of the tetracycline promoter. The ABD line also contains the reverse tet-repressor fused to VP16 and Androgen Binding Domain (ABD) under the control of the CAGGs promoter. The GBD*-clone carries a construct containing the reverse tet-repressor fused to VP16 and mutated Glucocorticoid Binding Domain (GBD*) under the control of the CAGGs promoter (GBD*-clone).

The term “induction” means that the cells were treated for 6 days with either Mibolerone and Doxycycline (for the ABD clone) or with Dexamethasone and Doxycycline (for the GBD* clone). This also signifies that the T-Antigen is expressed. “Deinduction” means that the cells were treated for 6 days with the above compounds and then cultured for an additional 6 days without the compounds. During the “deinduction” phase the expression of T-Antigen stops. Untreated cells means that the cells were cultured for the period of 12 days without any compounds, there is no expression of T-Antigen.

Abbreviation	Description	Arrays
ABD12dnot	clone 1 (ABD) 12 days untreated	3
ABD06dyes06dnot	clone 1 (ABD) 6 days induced and afterwards 6 days deinduced	3
ABD06dyes	clone 1 (ABD) 6 days induced	3
GBD12dnot	clone 2 (GBD) 12 days untreated	3
GBD06dyes06dnot	clone 2 (GBD) 6 days induced and afterwards 6 days deinduced	3
GBD06dyes	clone 2 (GBD) 6 days induced	3

Project UKOE-1: Differentiation of ES cells to Embryoid Bodies

Cell line: CGR8

Xavier Doss, Johannes Winkler, Jürgen Hescheler and Agapios Sachinidis

Contact: Agapios Sachinidis/Jürgen Hescheler

Institute of Neurophysiology

University of Cologne

Robert-Koch-Str. 39

50931 Cologne, Germany

Tel: +49-221-4787373

Email: a.sachinidis@uni-koeln.de; J.Hescheler@uni-koeln.de

Summary of the project

CGR8 cells were cultivated as embryoid bodies (EBs) without addition of selective reagents in hanging drop cultures (day 0-7) in a time course experiment. After 7 days, EBs were plated in gelatin-coated 6-well plates. RNA was prepared from ES cells (day 0), EBs (day 1 to 7) and plated EBs (day 10).

Abbreviation	Description	Arrays
00d	ES cell line CGR8 at day 0	3
01d	ES cell line CGR8 at day 1	3
02d	ES cell line CGR8 at day 2	3
03d	ES cell line CGR8 at day 3	3
04d	ES cell line CGR8 at day 4	3
05d	ES cell line CGR8 at day 5	3
06d	ES cell line CGR8 at day 6	3
07d	ES cell line CGR8 at day 7	3
10d	ES cell line CGR8 at day 10	3

Project AVEF-1: EB formation and differentiation into neuronal cells

Cell line: E14TG2a/Sox1TV2

Olivier Feraud, Murielle Gitton and Laurent Pradier

Contact: Laurent Pradier

Sanofi-Aventis

Centre de recherche de Paris

13, Quai Jules Guesde - BP 14

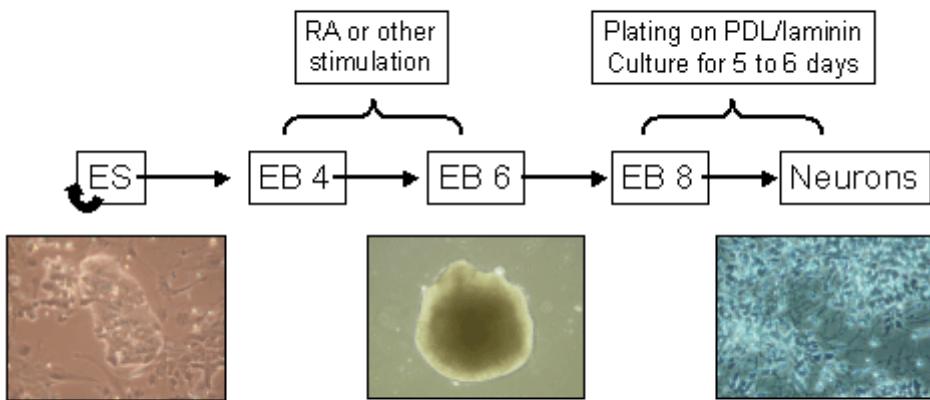
F-94403 Vitry sur Seine Cedex, France

Tel: +33 1 58933885

Email: Laurent.Pradier@sanofi-aventis.com

Summary of the project

EB protocol



ES cells have been allowed to form embryoid bodies in liquid culture condition onto non-adherent dishes. This step takes place during 8 days with a 48 hours retinoic acid stimulation between days 4 and 6. During this step samples have been prepared from ES cells, 3, 4, 6 and 8 days old embryoid bodies (ES, EB3-8) with (RA) or without (CTL) retinoic acid from day 6. At day 8 of differentiation EB have been dissociated and cells have been plated on poly-D-lysine/laminin coated dishes. During the first 48 hours of plating, bFGF has been added to the culture medium in order to increase the number of neuronal progenitor cells. Samples have been prepared at the end of this step (N2 for two days neurons). After removal of bFGF, cells are allowed to differentiate during 4 more days and the last samples have been prepared from this step (N6). RNA was prepared from four independent experiments and verified by RT-PCR of reference genes.

Abbreviation	Description	Arrays
0es	ES cells	4
eb3	3 days old embroyd bodies	4
eb4	4 days old embroyd bodies	2
eb6ctl	6 days old embroyd bodies; without retinoic acid stimulation	4
eb6ra	6 days old embroyd bodies; with retinoic acid stimulation	4
eb8ctl	8 days old embroyd bodies; without retinoic acid stimulation	4
eb8ra	8 days old embroyd bodies; with retinoic acid stimulation	4
n2ctl	(8 +) 2 days old neurons; without retinoic acid stimulation	4
n2ra	(8 +) 2 days old neurons; with retinoic acid stimulation	4
n6ctl	(8 +) 6 days old neurons; without retinoic acid stimulation	4
n6ra	(8 +) 6 days old neurons; with retinoic acid stimulation	4

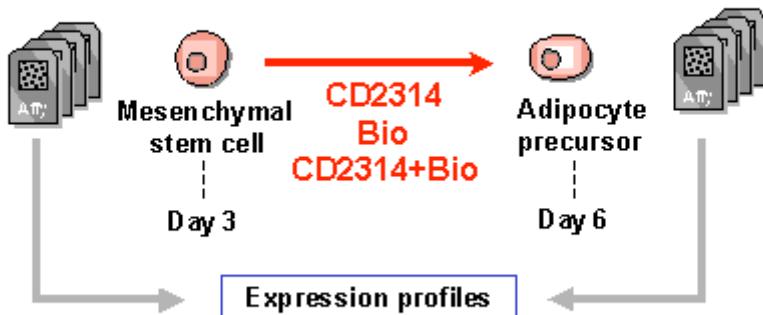
Project CNRS-UMR-6543: Differentiation to early mesoderm (day3) and further to adipocyte precursors (day 6) / Cell line: CGR8

Nathalie Billon and Christian Dani
Contact: Christian Dani
UMR 6543 CNRS / Centre de Biochimie
Parc Valrose
06108 Nice Cedex, France
Tel: +33 4 92076436
Email: dani@unice.fr

Summary of the project

Embryoid Bodies have been treated either with CD2314 (a Retinoic Acid Receptor β agonist), Bio (a canonical wnt pathway inducer) or both between day 3-6 of differentiation and RNA samples have been prepared at day 6 and at day 11.

Microarray experimental design



We have generated a high number of hanging drops, i.e., 62,100 drops, for each series of experiments and three independent series have been performed.

RNAs have been isolated and controlled for quality. Then, regulation of mesodermal differentiation for each condition and each series of experiments has been investigated by RT-PCR.

Abbreviation	Description	Arrays
day03	CGR8 cells at day 3; not stimulated	3
day06-0	CGR8 cells at day 6; not stimulated	3
day06-bio	CGR8 cells at day 6; stimulated with BIO	3
day06-cd	CGR8 cells at day 6; stimulated with CD2314	3
day06-cd-bio	CGR8 cells at day 6; stimulated with CD2314 and bio	3
day11-0	CGR8 cells at day 11; not stimulated	3
day11-bio	CGR8 cells at day 11; stimulated with BIO	3
day11-cd	CGR8 cells at day 11; stimulated with CD2314	3
day11-cd-bio	CGR8 cells at day 11; stimulated with CD2314and BIO	3

Project IPK-1: Differentiation of ESC to multilineage progenitors and further to committed pancreatic progenitors

Cell lines: R1; R1-PAX4

Alexandra Rolletschek and Anna Wobus

Contact: Anna Wobus

IPK Gatersleben

Correnstr.3

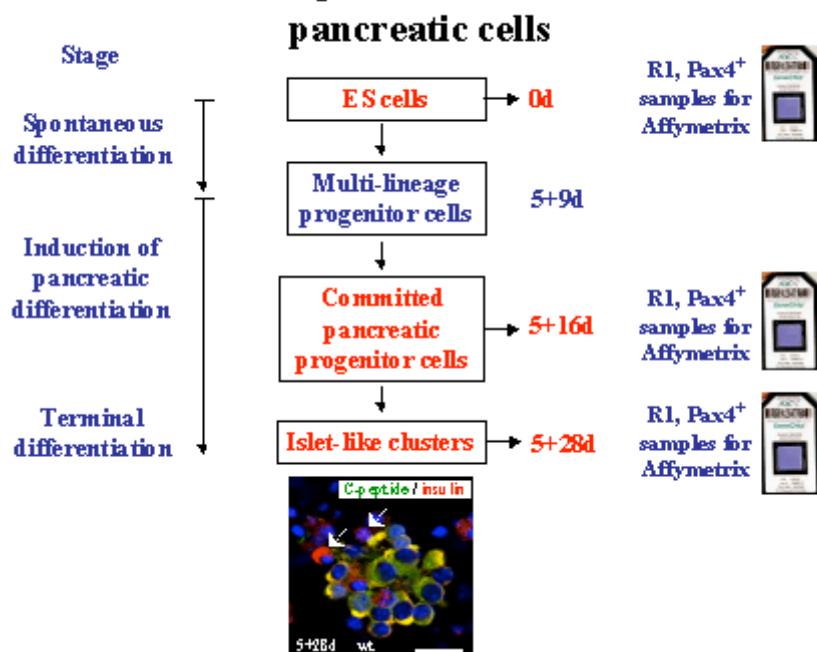
06466 Gatersleben, Germany

Tel: +49 394 825256

Email: wobusam@ipk-gatersleben.de

Summary of the project

Differentiation protocol for ES cell-derived



To characterize potential pancreatic progenitors during the process of induced pancreatic differentiation, wild-type R1 cells and R1 cells stably expressing the pancreatic control gene Pax4, RNA samples of undifferentiated cells, pancreatic precursor and differentiated cells were prepared.

Abbreviation	Description	Arrays
R1-ESC-1	R1; ES cells (untreated)	5
R1-5-09d-2	R1; 5 + 9 days	5
R1-5-16d-3	R1; 5 + 16 days	5
R1-5-28d-4	R1; 5 + 28 days	5
Pax4-ESC-1	R1-PAX; ES cells (untreated)	5
Pax4-5-16d-3	R1-PAX; 5 + 16 days	5
Pax4-5-28d-4	R1-PAX; 5 + 28 days	5

Affymetrix GeneChip Analysis Central Facility

Herbert Schulz and Norbert Hübner

Contact: Norbert Hübner

Max-Delbrück-Centre for Molecular Medicine (MDC) Berlin-Buch

Robert-Rössle-Str. 10

D-13092 Berlin, Germany

Tel: +49- 30 - 9406 2530

Email: nhuebner@mdc-berlin.de

Bioinformatics Team

Raivo Kolde, Priit Adler, Jüri Reimand and Jaak Vilo

Contact: Jaak Vilo

Institute of Computer Science, University of Tartu,

Liivi 2, 50409 Tartu, Estonia;

Estonian Biocentre,

Riia 23b, 51010 Tartu, Estonia; and

EGeen, Ülikooli 6a, 51003 Tartu, Estonia

Tel: +372 50 49 365

Email: Jaak.Vilo@ut.ee

FunGenES Principal Investigators

Michael Bader
Max-Delbrück-Centre for Molecular Medicine (MDC)
Robert-Rösslestr. 10
D-13092 Berlin-Buch, Germany
Tel: +49-30-9406-2193 / 2119
Email: mbader@mdc-berlin.de

Frank Buchholz
Max-Planck-Institute of Molecular Cell Biology and Genetics
Pfotenhauer Str. 108
01307 Dresden, Germany
Tel: +49-351-2102888
Email: buchholz@mpicbg.de

Lesley Forrester
Queens Medical Research Institute, E2.47
University of Edinburgh
47 Little France Crescent
Edinburgh EH16 4TJ, UK
Tel: +44 131 242 9163
Email: L.Forrester@ed.ac.uk

Ruairi Friel
Centre for Genome Research University of Edinburgh
Roger Land Building, Kings Building
West Mains Road
EH8 3JQ Edinburgh, UK
Tel: +44 131 650 5850
Email: ruairi.friel@stemcellsciences.com

Domingos Henrique
Instituto de Medicina Molecular
Faculdade de Medicina de Lisboa
Av. Prof. Egas Moniz
1649-028 Lisboa, Portugal
Tel: +351 21 7999516
Email: henrique@fm.ul.pt

Antonis K. Hatzopoulos
Department of Medicine and Department of Cell & Developmental Biology
Division of Cardiovascular Medicine
Vanderbilt University
2220 Pierce Ave – PRB 383
Nashville, TN 37232-6300, USA
Tel: +1-615-936 5529
Email: antonis.hatzopoulos@vanderbilt.edu

Heinz Himmelbauer
Max-Planck-Institute of Molecular Genetics
Ihnestr.73
D-14195 Berlin, Germany
Tel: +49-30-8413 1354
Email: himmelbauer@molgen.mpg.de