



HAL
open science

Répercussion du développement morphocinétique précoce sur la transition materno-embryonnaire : comparaison des transcriptomes d'embryons bovins au stade 16-32 cellules

Audrey Teste

► **To cite this version:**

Audrey Teste. Répercussion du développement morphocinétique précoce sur la transition materno-embryonnaire : comparaison des transcriptomes d'embryons bovins au stade 16-32 cellules. Médecine vétérinaire et santé animale. 2020. dumas-04782500

HAL Id: dumas-04782500

<https://dumas.ccsd.cnrs.fr/dumas-04782500v1>

Submitted on 14 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Année 2020

**RÉPERCUSSION DU DÉVELOPPEMENT
MORPHOCINÉTIQUE PRÉCOCE SUR LA TRANSITION
MATERNO-EMBRYONNAIRE : COMPARAISON DES
TRANSCRIPTOMES D'EMBRYONS BOVINS AU STADE
16-32 CELLULES**

THÈSE

pour obtenir le grade de

DOCTEUR VÉTÉRINAIRE

présentée et soutenue publiquement devant

LA FACULTÉ DE MÉDECINE DE CRÉTEIL

le 29 septembre 2020

par

Audrey, Colette, Géraldine TESTE

sous la direction de

Alline DE PAULA REIS

Président du jury :

Pr Alexandre DE LA TAILLE

Professeur à la Faculté de Médecine de CRÉTEIL

1^{er} Assesseur :

Dr Alline DE PAULA REIS

Maître de Conférences à l'EnvA

2nd Assesseur :

Dr Lucie CHEVALLIER

Maître de Conférences à l'EnvA

Liste des membres du corps enseignant



Directeur : Pr Christophe Degueurce
 Directeur des formations : Pr Henry Chateau
 Directrice de la scolarité et de la vie étudiante : Dr Catherine Colmin
 Directeurs honoraires : MM. les Professeurs C. Pilet, B. Toma, A.-L. Parodi, R. Moraillon, J.-P. Cotard, J.-P. Mialot & M. Gogny

Département d'Elevage et de Pathologie des Équidés et des Carnivores (DEPEC) Chef du département : Pr Grandjean Dominique - Adjoint : Pr Blot Stéphane

<p>Unité pédagogique d'anesthésie, réanimation, urgences, soins intensifs</p> <ul style="list-style-type: none"> - Dr Fernandez Parra Rocio, Maître de conférences associée - Pr Verwaerde Patrick* <p>Unité pédagogique de clinique équine</p> <ul style="list-style-type: none"> - Pr Audigé Fabrice - Dr Bertoni Léila, Maître de conférences - Dr Bourzac Céline, Chargée d'enseignement contractuelle - Dr Coudry Virginie, Praticien hospitalier - Pr Denoix Jean-Marie - Dr Giraudet Aude, Praticien hospitalier - Dr Jacquet Sandrine, Praticien hospitalier - Dr Mespouhès-Rivière Céline, Praticien hospitalier* - Dr Moiroud Claire, Praticien hospitalier - Dr Tanqueret Ludovic, Chargé d'enseignement contractuel <p>Unité pédagogique de médecine et imagerie médicale</p> <ul style="list-style-type: none"> - Dr Benchekroun Ghita, Maître de conférences - Pr Blot Stéphane* - Dr Canonne-Guibert Morgane, Maître de conférences - Dr Freiche-Legros Valérie, Praticien hospitalier - Dr Maurey-Guénec Christelle, Maître de conférences 	<p>Unité pédagogique de médecine de l'élevage et du sport</p> <ul style="list-style-type: none"> - Dr Cabrera Gonzales Joaquin, Chargé d'enseignement contractuel - Dr Fontbonne Alain, Maître de conférences - Pr Grandjean Dominique* - Dr Hoummady Sara, Chargée d'enseignement contractuelle - Dr Maenhoudt Cindy, Praticien hospitalier - Dr Nudelmann Nicolas, Maître de conférences - Dr Ribeiro dos Santos Natalia, Praticien hospitalier <p>Unité pédagogique de pathologie chirurgicale</p> <ul style="list-style-type: none"> - Dr Decambon Adeline, Maître de conférences - Pr Fayolle Pascal - Dr Manassero Mathieu, Maître de conférences - Pr Viateau-Duval Véronique* <p>Discipline : cardiologie</p> <ul style="list-style-type: none"> - Pr Chetboul Valérie - Dr Saponaro Vittorio, Praticien hospitalier <p>Discipline : ophtalmologie</p> <ul style="list-style-type: none"> - Dr Chahory Sabine, Maître de conférences <p>Discipline : nouveaux animaux de compagnie</p> <ul style="list-style-type: none"> - Dr Pignon Charly, Praticien hospitalier - Dr Volait Laetitia, Praticien hospitalier
--	--

Département des Productions Animales et de Santé Publique (DPASP) Chef du département : Pr Millemann Yves - Adjoint : Pr Dufour Barbara

<p>Unité pédagogique d'hygiène, qualité et sécurité des aliments</p> <ul style="list-style-type: none"> - Dr Bolnot François, Maître de conférences - Pr Carlier Vincent - Dr Gauthier Michel, Maître de conférences associée - Dr Mimitet Narjes, Chargée d'enseignement contractuelle <p>Unité pédagogique de maladies réglementées, zoonoses et épidémiologie</p> <ul style="list-style-type: none"> - Dr Crozet Guillaume, Chargé d'enseignement contractuel - Pr Dufour Barbara* - Pr Haddad/Hoang-Xuan Nadia - Dr Rivière Julie, Maître de conférences <p>Unité pédagogique de pathologie des animaux de production</p> <ul style="list-style-type: none"> - Pr Adjou Karim - Dr Belbis Guillaume, Maître de conférences* - Dr Delsart Maxime, Maître de conférences associé - Pr Millemann Yves - Dr Plassard Vincent, Praticien hospitalier - Dr Ravary-Plumioën Béangère, Maître de conférences 	<p>Unité pédagogique de reproduction animale</p> <ul style="list-style-type: none"> - Dr Constant Fabienne, Maître de conférences* - Dr Denis Marine, Chargée d'enseignement contractuelle - Dr Desbois Christophe, Maître de conférences (rattaché au DEPEC) - Dr Mauffré Vincent, Maître de conférences <p>Unité pédagogique de zootechnie, économie rurale</p> <ul style="list-style-type: none"> - Dr Arné Pascal, Maître de conférences - Dr Barassin Isabelle, Maître de conférences - Pr Bossé Philippe* - Dr De Paula Reis Aline, Maître de conférences - Pr Grimard-Ballif Bénédicte - Pr Ponter Andrew <p>Rattachée DPASP</p> <ul style="list-style-type: none"> - Dr Wolgust Valérie, Praticien hospitalier
---	--

Département des Sciences Biologiques et Pharmaceutiques (DSBP) Chef du département : Pr Desquilbet Loïc - Adjoint : Pr Pilot-Storck Fanny

<p>Unité pédagogique d'anatomie des animaux domestiques</p> <ul style="list-style-type: none"> - Dr Boissady Emilie, Chargée d'enseignement contractuelle - Pr Chateau Henry - Pr Crevier-Denoix Nathalie - Pr Robert Céline* <p>Unité pédagogique de bactériologie, immunologie, virologie</p> <ul style="list-style-type: none"> - Pr Boulouis Henri-Jean - Pr Eloit Marc - Dr Lagrée Anne-Claire, Maître de conférences - Pr Le Poder Sophie - Dr Le Roux Delphine, Maître de conférences* <p>Unité pédagogique de biochimie, biologie clinique</p> <ul style="list-style-type: none"> - Pr Bellier Sylvain* - Dr Deshuillers Pierre, Maître de conférences - Dr Lagrange Isabelle, Praticien hospitalier <p>Unité pédagogique d'histologie, anatomie pathologique</p> <ul style="list-style-type: none"> - Dr Cordonnier-Lefort Nathalie, Maître de conférences - Pr Fontaine Jean-Jacques - Dr Lalot Eve, Maître de conférences - Dr Reyes-Gomez Edouard, Maître de conférences* <p>Unité pédagogique de management, communication, outils scientifiques</p> <ul style="list-style-type: none"> - Mme Conan Muriel, Professeur certifié (Anglais) - Pr Desquilbet Loïc, (Biostatistique, Epidémiologie) - Dr Legrand Chantal, Maître de conférences associée - Dr Maignac Geneviève, Maître de conférences* - Dr Rose Hélène, Maître de conférences associée 	<p>Unité de parasitologie, maladies parasitaires, dermatologie</p> <ul style="list-style-type: none"> - Dr Blaga Radu, Maître de conférences (rattaché au DPASP) - Dr Briand Amaury, Assistant d'Enseignement et de Recherche Contractuel (rattaché au DEPEC) - Dr Cochet-Faivre Noëlle, Praticien hospitalier (rattaché au DEPEC) - Pr Guillot Jacques* - Dr Polack Bruno, Maître de conférences - Dr Risco-Castillo Veronica, Maître de conférences <p>Unité pédagogique de pharmacie et toxicologie</p> <ul style="list-style-type: none"> - Dr Kohlhauer Matthias, Maître de conférences - Dr Perrot Sébastien, Maître de conférences* - Pr Tissier Renaud <p>Unité pédagogique de physiologie, éthologie, génétique</p> <ul style="list-style-type: none"> - Dr Chevallier Lucie, Maître de conférences (Génétique) - Dr Crépeaux Guillemette, Maître de conférences (Physiologie, Pharmacologie) - Pr Gilbert Caroline (Ethologie) - Pr Pilot-Storck Fanny (Physiologie, Pharmacologie) - Pr Tiret Laurent (Physiologie, Pharmacologie)* - Dr Titeux Emmanuelle (Ethologie), Praticien hospitalier <p>Discipline : éducation physique et sportive</p> <ul style="list-style-type: none"> - M. Philips Pascal, Professeur certifié
---	---

* responsable d'unité pédagogique

Professeurs émérites : Pr Combrisson Hélène, Pr Enriquez Brigitte, Pr Panthier Jean-Jacques, Pr Paragon Bernard.

Remerciements

Au Président du Jury de cette thèse, Professeur à la Faculté de Médecine de Créteil,

Qui nous a fait l'honneur d'accepter la présidence du jury de cette thèse,
Hommage respectueux.

À Mme Alline DE PAULA REIS, Maître de conférences à l'EnvA,

Pour avoir placé votre confiance en moi pour l'écriture de cette thèse dont le sujet vous tenait tant à cœur et pour m'avoir épaulé, aiguillé, rassuré tant de fois au cours de ce travail difficile, pour n'avoir jamais baissé les bras et pour votre bonne humeur perpétuelle,
Mes plus sincères remerciements.

À Mme Lucie CHEVALLIER, Maître de conférences à l'EnvA,

Pour avoir acceptée d'être l'assesseure de cette thèse, pour votre relecture consciencieuse et pour vos remarques qui m'ont beaucoup touchée et rassurée sur mon travail,
Sincères remerciements.

À Véronique DURANTHON, Directrice de recherche, Catherine ARCHILLA, Assistante ingénieur, Luc JOUNEAU, Ingénieur de recherche et Sophie CALDERARI, Chargée de recherche à l'INRAE,

Pour votre disponibilité, ainsi que pour vos conseils et votre aide précieuse dans l'analyse des résultats de cette thèse,
Sincères remerciements.

À mes parents,

Pour m'avoir emmenée là où j'en suis aujourd'hui, pour m'avoir donné tant d'amour et pour ce lien indéfectible qui nous unit.

À ma famille,

Pour vos encouragements, votre écoute, votre soutien et votre amour.

À Aurélien,

Pour avoir accepté de m'aimer inconditionnellement.

À mes amis,

Qui m'ont apporté tant de bonheur au quotidien et m'ont permis de me donner à fond, pour avoir teinté ma vie de rires, d'aventures et de souvenirs incroyables.

Table des matières

Table des matières.....	1
Liste des figures.....	5
Liste des tableaux.....	7
Liste des abréviations.....	9
Introduction.....	19
Première partie : Bibliographie.....	21
I. Le développement embryonnaire précoce chez les bovins.....	21
1. L'ovocyte avant la fécondation.....	21
a. La maturation nucléaire de l'ovocyte.....	21
b. La maturation cytoplasmique de l'ovocyte.....	22
c. La maturation moléculaire de l'ovocyte.....	22
➤ Les ARN.....	23
➤ Les protéines.....	23
d. La régulation des ARN maternels.....	23
➤ La polyadénylation et le masquage des ARNm.....	24
➤ Le démasquage : initiation de la traduction dépendante du CPE.....	24
➤ La régulation temporelle du démasquage.....	26
➤ La déadénylation des ARNm maternels pendant la maturation de l'ovocyte.....	26
➤ La déadénylation des ARNm maternels pendant le développement embryonnaire.....	26
2. Le clivage embryonnaire : du zygote au blastocyste.....	27
3. La transition materno-embryonnaire (MET).....	29
II. La transcription chez les eucaryotes.....	30
1. Le mécanisme général de la transcription.....	30
2. Illustration avec l'exemple de la transcription par l'ARN polymérase II chez les eucaryotes.....	31
3. La maturation des ARN.....	33
a. L'addition de la coiffe à l'extrémité 5'.....	33
b. La polyadénylation en 3'.....	33
c. L'épissage alternatif.....	34
III. Régulation de la transcription.....	36
1. La modulation de l'accessibilité à l'ADN.....	36
2. Les séquences cis-régulatrices de l'ADN.....	37
3. Le phénomène d'interférence d'ARN.....	38
➤ Zoom sur la régulation par les ARNmi.....	40
IV. L'activation du génome embryonnaire.....	42
1. Le moment de l'activation : activations mineure et majeure.....	42
2. Le contrôle de l'activation du génome embryonnaire.....	44
a. L'utilisation des protéines et des ARNm maternels stockés par l'ovocyte.....	44
b. Les modifications de la chromatine.....	45
➤ La déméthylation de l'ADN.....	45
➤ La répression de la transcription par la méthylation de l'ADN et son lien avec la modification des histones.....	47
c. Le rôle de la réplication de l'ADN.....	48
3. La dégradation des transcrits maternels.....	49

a.	La voie de l'interférence.....	49
➤	Chez le xénope, la souris	50
➤	Chez les bovins	51
b.	Les protéines de liaison à l'ARN	52
c.	L'autophagie	53
➤	Chez la souris.....	53
➤	Chez le porc.....	54
4.	Les gènes et les voies fonctionnelles activés lors de l'EGA	54
a.	Comment différencier les transcrits maternels des transcrits <i>de novo</i> au cours de l'EGA ?	55
b.	Etude de l'abondance des transcrits produits <i>de novo</i> par l'embryon et identification des gènes associés	56
c.	La fonction des gènes transcrits <i>de novo</i>	58
➤	Au stade quatre cellules	58
➤	Au stade huit cellules	59
➤	Au stade seize cellules.....	63
➤	Au stade blastocyste	64
d.	Les voies enrichies par les gènes transcrits <i>de novo</i>.....	64
➤	Au stade quatre cellules	65
➤	Au stade huit cellules	65
➤	Au stade seize cellules.....	66
➤	Au stade blastocyste	66
	Deuxième partie : Étude expérimentale	69
I.	Introduction.....	69
II.	Matériel et méthode	70
1.	<i>Production des embryons.....</i>	70
2.	<i>Maturation in vitro (MIV).....</i>	70
3.	<i>Fécondation in vitro (FIV) et culture embryonnaire.....</i>	70
4.	<i>Images en Temps réel (Time Lapse)</i>	71
5.	<i>Classification et formation des lots pour l'analyse transcriptomique</i>	71
6.	<i>Extraction de l'ARN et mesure de la pureté des échantillons.....</i>	72
7.	<i>Quantification de l'abondance des ARN.....</i>	72
8.	<i>Analyse des données</i>	72
9.	<i>Contrôle qualité.....</i>	72
10.	<i>Analyse des résultats.....</i>	73
a.	Analyse descriptive	73
➤	Clustering hiérarchique	74
➤	Analyse en composantes principales (ACP).....	74
b.	Analyse différentielle.....	74
➤	Abondance des transcrits des gènes étudiés	74
➤	Analyse de la comparaison quantitative entre catégories	75
➤	Analyse des profils des gènes différentiellement exprimés dans la comparaison mettant en évidence le plus grand nombre de différences transcriptomiques	76
➤	Etude d'enrichissement des ensembles de gènes différentiellement exprimés dans la comparaison mettant en évidence le plus grand nombre de différences transcriptomiques.....	76
III.	Résultats.....	77
1.	<i>Analyse descriptive.....</i>	77
a.	Clustering hiérarchique de corrélation	78
b.	Analyse en composantes principales (ACP)	79
2.	<i>Analyse différentielle.....</i>	80
a.	Abondance des transcrits des gènes étudiés	80
b.	Comparaison quantitative entre catégories	81

c. Diagrammes de Venn	82
3. Focus sur la comparaison A vs D	87
a. Gènes connus significativement surexprimés en A par rapport à la catégorie D	87
b. Gènes connus significativement surexprimés en D par rapport à la catégorie A	89
c. Profil des gènes différentiellement exprimés dans la comparaison A vs D	96
d. Voies fonctionnelles significativement enrichies par les gènes différentiellement exprimés dans la comparaison A vs D	98
IV. Discussion	109
1. Les gènes significativement surexprimés dans les embryons des catégories A ou D	110
a. Un retard d'utilisation des transcrits maternels par les embryons de la catégorie D ?	110
b. Un retard de dégradation des transcrits maternels chez les embryons de la catégorie D ?	111
➤ Implication de la voie de l'interférence ?	111
➤ Pourquoi la voie de l'interférence serait-elle dysfonctionnelle dans les embryons de la catégorie D ?	112
➤ D'autres mécanismes possiblement impliqués ?	113
c. Un retard d'activation transcriptionnelle chez les embryons de la catégorie D ?	113
➤ Un défaut de déméthylation ?	113
➤ Un problème dans la réplication de l'ADN ?	114
2. Les voies fonctionnelles significativement enrichies par les gènes surexprimés dans les embryons avec un développement morphocinétique arythmique (catégorie D)	115
Conclusion	121
Publications	123
Liste des références bibliographiques	125
Annexe 1 : Liste des gènes surexprimés en A par rapport aux catégories B, C et D	137
Annexe 2 : Liste des gènes surexprimés en B par rapport aux catégories A, C et D	141
Annexe 3 : Liste des gènes surexprimés en C par rapport aux catégories A, B et D	145
Annexe 4 : Liste des gènes surexprimés en D par rapport aux catégories A, B et C	149

Liste des figures

Figure 1 : Le développement embryonnaire précoce chez les bovins, de la fécondation au stade blastocyste.....	28
Figure 2 : Représentation schématique des motifs de liaison à l'ADN des facteurs de transcription (Richard <i>et al.</i> , 2012).....	31
Figure 3 : L'initiation de la transcription par l'ARN polymérase II (Lévine, 2011).....	32
Figure 4 : Les étapes de la polyadénylation des ARNm (Richard <i>et al.</i> , 2012).....	34
Figure 5 : Le mécanisme de l'épissage par le spliceosome (Richard <i>et al.</i> , 2012).....	35
Figure 6 : Le principe de l'épissage alternatif (Jaspard, 2012).....	36
Figure 7 : Le mode d'action des activateurs et des répresseurs de la transcription via leurs interactions avec les séquences amplificatrices (<i>enhancers</i>) ou silencieuses (<i>silencers</i>) (Richard <i>et al.</i> , 2012).....	38
Figure 8 : Les voies de l'interférence d'ARN, rôle des ARNm (à gauche) et des ARNs (à droite) (Richard <i>et al.</i> , 2012).....	39
Figure 9 : De l'ARNm à la protéine : les étapes de la transcription à la traduction et les régulations associées (en rouge).....	41
Figure 10 : Profil de méthylation du génome paternel (courbe bleue) et du génome maternel (courbe rose) au cours du développement précoce embryonnaire chez la souris (a) et chez le bovin (b). Les courbes noires représentent la méthylation <i>de novo</i> dans les lignées cellulaires embryonnaires (EM) et extra-embryonnaires (EX) (Dean <i>et al.</i> , 2003).....	47
Figure 11 : Voies fonctionnelles enrichies par les gènes activés <i>de novo</i> par l'embryon à différents stades de développement (créée à partir des résultats de Misirlioglu <i>et al.</i> , 2006 ; Vigneault <i>et al.</i> , 2009 ; Graf <i>et al.</i> , 2014b).....	67
Figure 12 : Lectures des <i>spikes</i> en fonction de leur concentration dans les 16 échantillons analysés.....	73
Figure 13 : <i>Clustering</i> hiérarchique de corrélation des quatre répétitions de chaque catégorie (A, B, C et D).....	78
Figure 14 : Analyse en Composantes Principales des quatre répétitions de chaque catégorie (A, B, C et D).....	79
Figure 15 : Gènes significativement surexprimés dans la catégorie A par rapport aux catégories B, C et D (FC > 2 et p-adj < 0,05).....	83

Figure 16 : Gènes significativement surexprimés dans la catégorie B par rapport aux catégories A, C et D (FC > 2 et p-adj < 0,05)	84
Figure 17 : Gènes significativement surexprimés dans la catégorie C par rapport aux catégories A, B et D (FC > 2 et p-adj < 0,05)	85
Figure 18 : Gènes significativement surexprimés dans la catégorie D par rapport aux catégories A, B et C (FC > 2 et p-adj < 0,05)	86
Figure 19 : Pourcentage de gènes différentiellement exprimés dans la comparaison A vs D en fonction de leur profil dans l’embryon bovin produit <i>in vivo</i> (déterminé à partir des données obtenues dans EmbryoExpress).	97
Figure 20 : Visualisation des voies significativement surexprimées en D par rapport à A au sein des grands ensembles de voies régissant les processus biologiques, les fonctions moléculaires et les composants cellulaires.	102
Figure 21 : Voies fonctionnelles enrichies par les gènes surexprimées en D par rapport à A (Groupe I : GO:0000977 (1) et GO:0000976 (2)).	104
Figure 22 : Voies fonctionnelles enrichies par les gènes surexprimées en D par rapport à A (Groupe II : GO:0001228 (1), GO:0045944 (2), GO:0045892 (3), GO:0045893 (4), GO:0006357 (5) et GO:1903508 (6)).	105
Figure 23 : Voies fonctionnelles enrichies par les gènes surexprimées en D par rapport à A (Groupe III : GO:0071900).	107
Figure 24 : Voies enrichies par les gènes surexprimées en D par rapport à A (Groupe IV : GO:0050777).	108
Figure 25 : Schéma bilan de la discussion.	119

Liste des tableaux

Tableau 1 : Nombre de gènes identifiés comme transcrits <i>de novo</i> par l'embryon à différents stade de développement en fonction des stratégies utilisées.	58
Tableau 2 : Nombre d'ARN absents (0 transcrits lus), faiblement présents (entre 0 et 10 transcrits lus), moyennement présents (entre 10 et 100 transcrits lus) ou fortement présents (>100 transcrits lus), dans les quatre répétitions de chaque catégorie (A, B, C et D)	80
Tableau 3 : Nombre de gènes inclus dans les comparaisons entre les différentes catégories d'embryons (A, B, C, D) (Total) et nombre de gènes différentiellement exprimés en fonction des seuils de <i>Fold Change</i> (<0,5 ou >2) et de la valeur p ajustée (<0,05) choisis dans la présente étude	82
Tableau 4 : Gènes connus surexprimés en A par rapport à D avec description de leur fonction et de leur <i>Fold-Change</i> (FC > 2 dans la comparaison A vs D ; p-adj < 0,05)	87
Tableau 5 : Gènes connus surexprimés en D par rapport à A et impliqués dans des voies fonctionnelles différentiellement enrichies avec description de leur fonction et de leur <i>Fold-Change</i> (FC < 0,5 dans la comparaison A vs D ; p-adj < 0,05)	89
Tableau 6 : Proportion de gènes différentiellement exprimés dans la comparaison A vs D en fonction de leur profil dans l'embryon bovin produit <i>in vivo</i>	97
Tableau 7 : Voies fonctionnelles significativement enrichies par les gènes surexprimés dans la catégorie D par rapport à la catégorie A.	99

Liste des abréviations

3-MA : 3-méthyladénine

ACP : Analyse en Composantes Principales

ADAMTSL2 : *A Disintegrin And Metalloproteinase with Thrombospondin Motifs-Like 2*

adj p-value : valeur p ajustée = p-adj

ADN : Acide Désoxyribonucléique

ADNc : Acide Désoxyribonucléique complémentaire

ADP : *Adenosine Diphosphate*

AMPc : Adénosine Monophosphate cyclique = PKA

ARE : *Adenylate-Uridylate-Rich Element*

ARE-BP : *Adenylate-Uridylate-Rich Element Binding Protein*

ARN : Acide Ribonucléique

ARNm : Acide Ribonucléique messenger

ARNmi : Acide Ribonucléique *micro*, micro acide ribonucléique = miR

ARNr : Acide Ribonucléique ribosomal

ARNsi : Acide Ribonucléique *small interfering*, petit acide ribonucléique interférent

ARNsn : Acide Ribonucléique *small nuclear*, petit acide ribonucléique nucléaire

ARNsno : Acide Ribonucléique *small nucleolar*, petit acide ribonucléique nucléolaire

ARNt : Acide Ribonucléique de transfert

ARNtot : Acide ribonucléique total

ATG5 : *Autophagy Related 5*

ATP : *Adenosine Triphosphate*

ATP5A1 : *Adenosine Triphosphate Synthase F1 Subunit Alpha*

ATP5B : *Adenosine Triphosphate Synthase F1 Subunit Beta*

ATP5F1 : *Adenosine Triphosphate Synthase Peripheral Stalk-Membrane Subunit B*

ATP5J2 : Adenosine Triphosphate Synthase Membrane Subunit F

AURKA : Aurora Kinase A

AURKB : Aurora Kinase B

AURKC : Aurora Kinase C

BECN1 : Beclin 1

BMP15 : Bone Morphogenetic Protein 15

BMPR1B : Bone Morphogenetic Protein Receptor Type 1B

BREED : Biologie de la Reproduction, Environnement, Épigénétique et Développement

BrUTP : 5-Bromouridine 5'-Triphosphate

BYSL : Bystin-Like

CASR : Calcium Sensing Receptor

CCDC172 : Coiled-Coil Domain Containing 172

CCL17 : Chemokine Ligand 17

CCR4/POP2/NOT : Carbon Catabolite Repressor 4/Pyrin Domain-Only Protein 2/Negative Regulator Of Transcription

Cdc2 : Cell Division Cycle Protein 2 = Cdk1

Cdc25 : Cell Division Cycle Protein 25

CDH1 : Cadherin 1

Cdk : Cyclin-Dependent Kinase

Cdk1 : Cyclin-Dependent Kinase 1 = Cdc2

Cdk2 : Cyclin-Dependent Kinase 2

CDKN1A : Cyclin Dependent Kinase Inhibitor 1A

Cdx2 : Caudal Type Homeobox 2

CHD2 : Chromodomain Helicase Deoxyribonucleic Acid Binding Protein 2

C-mos : mos Proto-oncogene, Serine/Thréonine Kinase

COC : Complexe Ovocyte Cumulus

CPE : *Cytoplasmic Polyadenylation Element*

CPEB : *Cytoplasmic Polyadenylation Element Binding Protein*

CPEB1 : *Cytoplasmic Polyadenylation Element Binding Protein 1*

CPEB4 : *Cytoplasmic Polyadenylation Element Binding Protein 4*

CPSF : *Cleavage and Polyadenylation Specificity Factor*

CSTF : *Cleavage Stimulation Factor*

CTD : *Carboxy-Terminal Domain*

CUG-BP : *Cytosine-Uracile-Guanine Triplet Repeat Ribonucleic Acid-Binding Protein*

CUG-BP1 : *Cytosine-Uracile-Guanine Triplet Repeat Ribonucleic Acid-Binding Protein 1*

CUL4 : *Cullin-4*

CycA2 : *Cyclin A2*

DCAF : *Deoxyribonucleic Acid Damage-Binding Protein 1- Cullin-4 Associated Factor*

DCP2 : *messenger Ribonucleic Acid-Decapping Enzyme 2*

DDB1 : *Deoxyribonucleic Acid Damage-Binding Protein 1*

DDX5 : *Sequence D-E-A-D (Aspartate-Glutamate-Alanine-Aspartate)-box Helicase 5*

DEAD-box : *Sequence D-E-A-D (Aspartate-Glutamate-Alanine-Aspartate)-box*

DEAH-box : *Sequence D-E-A-H (Aspartate- Glutamate- Alanine- Histidine)-box*

DHX15 : *Sequence D-E-A-H (Aspartate- Glutamate- Alanine- Histidine)-box Helicase 15*

DHX38 : *Sequence D-E-A-H (Aspartate- Glutamate- Alanine- Histidine)-box Helicase 38*

DHX8 : *Sequence D-E-A-H (Aspartate- Glutamate- Alanine- Histidine)-box Helicase 8*

DNase : *Deoxyribonuclease*

DNMT : *Deoxyribonucleic Acid Methyltransferase*

DNMT1 : *Deoxyribonucleic Acid Methyltransferase 1*

DNMT3B : *Deoxyribonucleic Acid Methyltransferase 3 Beta*

DSC2 : *Desmocolin 2*

eCPE : *embryonic type Cytoplasmic Polyadenylation Element*

EDEN : *Embryonic Deadenylation Element*

EDEN-BP : *Embryonic Deadenylation Element Binding Protein*

EGA : *Embryonic Genome Activation*, activation du génome embryonnaire

EGF : *Epidermal Growth Factor*, facteur de croissance épidermique

EIF2C2 : *Protein Argonaute-2*

EIF3 : *Eukaryotic Initiation Factor 3*

eIF4E : *eukaryotic Translation Initiation Factor 4E*

eIF4G : *eukaryotic Translation Initiation Factor 4G*

ELAVL1 : *Embryonic Lethal Abnormal Vision-Like 1*

ELAVL2 : *Embryonic Lethal Abnormal Vision-Like 2*

ELK3 : *E26 Transforming-Specific Transcription Factor Elk3*

ELP2 : *Elongator Acetyltransferase Complex Subunit 2*

EM : Embryonnaire

ePAB : *embryonic Polyadenosine Binding Protein*

EPCAM : *Epithelial Cell Adhesion Molecule*

ESE : *Exonic Splicing Enhancer*

EX : Extra-embryonnaire

FAM186A : *Family With Sequence Similarity 186 Member A*

FC : *Fold-Change*

FIGLA : *Factor In the Germline Alpha*

FIV : *Fécondation In Vitro*

FOSL1 : *Fos-Related Antigen 1*

FSH : *Follicle-Stimulating Hormone*, hormone folliculo-stimulante

FURIN : *Furin Paired Basic Amino Acid Cleavage Enzyme*

FUZ : *Fuzzy Planar Cell Polarity Protein*

GABPA : *Guanine-Adenine-Binding Protein Alpha Chain*

GAP43 : *Growth Associated Protein 43*

GATA2 : *Guanine-Adenine-Thymine-Adenine Binding Protein 2*

GATA3 : *Trans-Acting T-Cell-Specific Transcription Factor*

GDF9 : *Growth Differentiation Factor 9*

GLD2 : *Germ Line Development 2*

GM-CSF : *Granulocyte-Macrophage Colony-Stimulating Factor*

GO : *Gene Ontology*

GTF : *General Transcription Factor, facteurs généraux de transcription*

GV : *Germinal Vesicle, vésicule germinale*

GVBD : *Germinal Vesicle Break Down*

H2A : *Histone 2A*

H2AFZ : *H2A Histone Family, Member Z*

H2B : *Histone 2B*

H3 : *Histone 3*

H4 : *Histone 4*

HAT : *Histone Acetyltransferase*

HDAC : *Histone Deacetylase*

HEX : *Polyadenylation Hexanucleotide*

HNRNPA2B1 : *Heterogeneous Nuclear Ribonucleoproteins A2/B1*

HSD3B7 : *Hydroxy-Delta-5-Steroid Dehydrogenase, 3 Beta- and Steroid Delta-Isomerase 7*

Hsp70 : *70 kilodalton Heat Shock Protein*

HSPA1A : *Heat Shock Protein Family A Member 1A*

HTR6 : *5-Hydroxytryptamine Receptor 6*

I2BC : *Institut de Biologie Intégrative de la Cellule*

IFIT1 : *Interferon Induced Protein With Tetratricopeptide Repeats 1*

IGF1R : *Insulin-Like Growth Factor 1 Receptor*

IL37 : *Interleukin 37*

INRAE : Institut National de Recherche pour l'Agriculture et l'Environnement

IP₃ : Inositol Triphosphate

JPH1 : *Junctophilin 1*

jpi : jour post-insémination

KCNE4 : *Potassium Voltage-Gated Channel Subfamily E Regulatory Subunit 4*

KCNH6 : *Potassium Voltage-Gated Channel Subfamily H Member 6*

KCTD21 : *Potassium Channel Tetramerization Domain Containing 21*

KLF17 : *Krüppel-Like Factor 17*

KLF4 : *Kruppel-Like Factor 4*

KO : *Knock-Out*

LC3 : *Microtubule Associated Protein 1 Light Chain 3 Alpha*

LED : *Light-Emitting Diode*

LH : *Luteinizing Hormone*, hormone lutéinisante

MAD1L1 : *Mitotic Arrest Deficient 1-Like 1*

MalAGE : Mathématiques et Informatique Appliquées du Génome à l'Environnement

MBD : *Methyl-Cytosine-Phosphate-Guanine Binding Domain Protein*

MBD1 : *Methyl-Cytosine-Phosphate-Guanine Binding Domain Protein 1*

MBD2 : *Methyl-Cytosine-Phosphate-Guanine Binding Domain Protein 2*

MBD3 : *Methyl-Cytosine-Phosphate-Guanine Binding Domain Protein 3*

MCI : Masse Cellulaire Interne

mCPE : *maturation type Cytoplasmic Polyadenylation Element*

MeCP2 : *Methyl-Cytosine-Phosphate-Guanine Binding Protein 2*

MET : *Maternal-Embryonic Transition*, transition materno-embryonnaire

MGI : *Mouse Genome Informatics*

MII : Métaphase II

miR : micro Acide Ribonucléique = ARNmi

MIV : Maturation *In Vitro*

MLL2 : *Histone-Lysine N-Methyltransferase*

MNF1 : *Mitochondrial Nucleoid Factor 1*

MPF : *Maturation Promoting Factor*

mRNP : *messenger Ribonucleoprotein*

MTA2 : *Metastasis Associated 1 Family Member 2*

MYB : *Myb Proto-oncogene, Transcription Factor*

Myc : *Myelocytomatosis*

Myt1 : *Myelin Transcription Factor 1*

NANOG : *Nanog Homeobox*

nc : non connu

NCAN : *Neurocan*

NFYA : *Nuclear Transcription Factor Y Subunit Alpha*

NOBOX : *Newborn Ovary Homeobox Protein*

NRXN1 : *Neurexin 1*

NuRD : *Nucleosome Remodeling Deacetylase*

Oct4 : *Octamer-Binding Transcription Factor 4*

OLR1 : *Oxidized Low Density Lipoprotein Receptor 1*

ORM1 : *Orosomuroid 1*

p53 : *Tumor Protein 53*

PABPC : *Polyadenosine Binding Protein Cytoplasmic*

p-adj : valeur p ajustée = adj p-value

PAG11 : *Pregnancy-Associated Glycoprotein 11*

PARN : *Polyadenosine-Specific Ribonucléase*

PBS : *Phosphate-Buffered Saline*, tampon phosphate salin

PDCD2 : *Programmed Cell Death Protein 2*

PFN2 : *Profilin 2*

PITPNM2 : *Phosphatidylinositol Transfer Protein Membrane Associated 2*

PKA : Protéine Kinase A = AMPc

Poly(A) : *Polyadenosine*

Poly(T) : *Polythymidine*

POU5F1 : *Octamer-Binding Protein Domain, Class 5, Transcription factor 1*

PPARGC1A : *Peroxisome Proliferator-Activated Receptor Gamma Coactivator 1 Alpha*

pRB : *Retinoblastoma Protein*

Pré-ARNm : Acide Ribonucléique Prémessager

PREB : *Prolactin Regulatory Element Binding*

RBP : *Ribonucleic Acid Binding Protein*

RIN : *Ribonucleic Acid Integrity Number*

RISC : *Ribonucleic Acid Induced Silencing Complex*

RNAseq : *Ribonucleic Acid Sequencing, séquençage de l'ARN*

RT-PCR : *Reverse Transcription-Polymerase Chain Reaction*

SALL1 : *Sal-Like 1*

SALL4 : *Sal-Like 4*

Samd4 : *Sterile Alpha Motif Domain-Containing Protein 4*

SFRS3 : *Splicing Factor Arginine/Serine-Rich 3*

SLAMF1 : *Signaling Lymphocytic Activation Molecule Family Member 1*

SLC12A5 : *Solute Carrier Family 12 Member 5*

SMAD5 : *Mothers Against Decapentaplegic Family Member 5*

SMARCA4 : *Switch/Sucrose Non-Fermentable-Related, Matrix Associated, Actin Dependent Regulator of Chromatin, Subfamily A, Member 4*

SMARCAL1 : *Switch/Sucrose Non-Fermentable -Related Matrix-Associated Actin-Dependent Regulator of Chromatin Subfamily A-Like Protein 1*

SMOC1 : *Secreted Protein Acidic and Rich in Cysteine Related Modular Calcium Binding 1*

SNP : *Single-Nucleotide Polymorphism*

snRNP : *small nuclear Ribonucleoprotein*

SNW1 : *SNW Domain-Containing Protein 1*

SOF : *Synthetic Oviduct Fluid*

SOX2 : *Sex Determining Region Y-box 2*

SOX5 : *Sex Determining Region Y-Box Transcription Factor 5*

SP1 : *Specificity Protein 1*

SR : *Serine and Arginine-Rich*

SRE : *Smaug Response Element*

SWI/SNF : *Switch/Sucrose Non-Fermentable*

TBP : *TATA Binding Protein*

TBX3 : *T-Box 3*

TEAD4 : *Transcriptional Enhancer Factor Domain Family Member 4*

TFCP2L1 : *CP2 Family Transcription Factor*

TFIIA : *Transcription Factor IIA*

TFIIB : *Transcription Factor IIB*

TFIID : *Transcription Factor IID*

TFIIE : *Transcription Factor IIE*

TFIIF : *Transcription Factor IIF*

TFIIH : *Transcription Factor IIH*

TNF α : *Tumor Necrosis Factor Alpha*

TP3 : *Putative Spermatidal Transition Protein 3*

TPM : *Transcripts Per Million*

TPST2 : *Tyrosylprotein Sulfotransferase 2*

TPT1 : *Tumor Protein Translationally-Controlled 1*

TRIB1 : *Tribbles Pseudokinase 1*

TRIM71 : *Tripartite Motif Containing 71*

TSG101 : *Tumor Susceptibility Gene 101*

UBE2D1 : *Ubiquitin Conjugating Enzyme E2 D1*

UCP3 : *Uncoupling Protein 3*

UMR : *Unité Mixte de Recherche*

UTR : *Untranslated Transcribed Region*

VEGF : *Vascular Endothelial Growth Factor A*

WISP1 : *Wnt Family Member 1 Inducible Signaling Pathway Protein 1*

Xrn1 : *5'-3' Exoribonuclease 1*

Xrn2 : *5'-3' Exoribonuclease 2*

YES1 : *Yamaguchi Sarcoma Viral Oncogene Homolog 1*

ZACN : *Zinc Activated Ion Channel*

ZFP36L2 : *Zinc Finger Protein 36-Like 2*

ZFP42 : *Zinc Finger Protein 42*

ZNF41 : *Zinc Finger Protein 41*

ZNF770 : *Zinc Finger Protein 770*

ZP1 : *Zona Pellucida Sperm-Binding Protein 1*

ZP2 : *Zona Pellucida Sperm-Binding Protein 2*

ZP3 : *Zona Pellucida Sperm-Binding Protein 3*

Introduction

La transcriptomique est l'étude du transcriptome, c'est-à-dire de l'ensemble des ARN issus de la transcription du génome d'un organisme.

Jusqu'à récemment, le transcriptome était souvent étudié grâce à des techniques de *microarray*. Cependant, ce genre de technique possède des limites telles que la dépendance à l'égard des connaissances existantes sur la séquence génomique de l'espèce étudiée, ou la difficulté de comparaison de niveaux d'expression entre différents objets d'étude.

Depuis quelques années, le développement de nouvelles méthodes de séquençage d'ADN à haut débit a permis de trouver une nouvelle technique permettant de cartographier et quantifier les transcriptomes : le *RNA-seq* (ou séquençage de l'ARN) (Wang *et al.*, 2009).

Sa capacité à quantifier à un instant t , l'abondance des transcrits présents dans les cellules, et donc à fournir une information sur l'expression génique aux différents stades du développement, en fait la technique de choix pour les études sur le développement embryonnaire.

Le *RNA-seq* est donc particulièrement indiqué pour l'étude de la transition materno-embryonnaire, moment où le contrôle du développement embryonnaire passe du transcriptome maternel au transcriptome embryonnaire.

En effet, l'embryon se développe, jusqu'à son activation génomique, grâce aux transcrits stockés par l'ovocyte lors de l'ovogénèse. Puis progressivement, il met en place sa propre machinerie transcriptionnelle, lui permettant de poursuivre son développement grâce à ses propres transcrits. Chez les bovins, cette activation du génome embryonnaire se déroule entre les stades huit et seize cellules.

Cependant, les mécanismes régissant cette étape sont encore en cours d'étude chez les bovins.

Par ailleurs, actuellement, les études portant sur le sujet considèrent des embryons de même stade embryonnaire comme similaires d'un point de vue de leur activation génomique.

Pourtant, l'observation du déroulement du développement précoce de l'embryon bovin produit *in vitro* a permis d'identifier de nombreux scénarios morphocinétiques de développement.

Parmi cette multitude de trajectoires que peuvent prendre les embryons post-fécondation, Reis *et al.* (2018) ont défini quatre grandes catégories de profil morphocinétique pour lesquelles les embryons ont le potentiel de se développer jusqu'au stade blastocyste.

Ces quatre catégories varient selon leur similitude avec le développement embryonnaire décrit dans la littérature.

Ainsi, la première catégorie représente les embryons avec une trajectoire morphocinétique très proche du développement normal décrit dans la littérature, appelés blastocystes à éclosion précoce (catégorie A). Les deux catégories suivantes représentent les embryons s'éloignant de cet « idéal » selon un gradient décroissant, appelés blastocystes éclos (catégorie B) et blastocystes retardés (catégorie C). La dernière et quatrième catégorie représente les embryons se détachant le plus du modèle de la littérature et sont appelés blastocystes arythmiques (catégorie D).

De plus, cette divergence de trajectoire étant très précoce dans le développement embryonnaire, il est actuellement possible de prédire dans quelle catégorie de blastocyste va se retrouver l'embryon (A, B, C ou D) dès le quatrième jour post-insémination, soit, avant l'activation du génome embryonnaire.

L'objectif général de cette thèse est donc de déterminer si le profil de développement morphocinétique précoce des embryons bovins appartenant à ces quatre catégories engendre des différences transcriptomiques révélatrices d'activations du génome embryonnaire divergentes. Un second objectif de cette thèse sera de proposer des pistes de réflexion pour des études à venir afin d'expliquer les principales différences transcriptomiques observées.

Pour ce faire, la première partie de cette thèse sera consacrée à une revue de la littérature. Nous aborderons tout d'abord l'ovogénèse et notamment les phénomènes de maturation de l'ovocyte qui entraînent le stockage de transcrits maternels permettant le développement précoce de l'embryon suite à la fécondation et ce, jusqu'à l'activation du génome embryonnaire.

Puis, un rappel des étapes du développement embryonnaire de la fécondation jusqu'au stade blastocyste sera fait.

Par la suite, nous détaillerons les mécanismes régissant la transcription ainsi que sa régulation.

Enfin, nous ferons un état des lieux des connaissances sur l'activation du génome embryonnaire bovin, son contrôle et ses mécanismes sous-jacents, ainsi que les gènes et les voies fonctionnelles actuellement identifiées comme significativement enrichies chez des embryons bovins produits *in vitro* au cours des principaux stades de développement : quatre cellules, huit cellules, seize cellules et blastocyste.

Dans la deuxième partie de cette thèse, nous décrirons notre étude expérimentale.

Celle-ci est dédiée à la comparaison des transcriptomes d'embryons bovins produits *in vitro* et appartenant à l'une des quatre catégories morphocinétiques précédemment décrites (A, B, C ou D). Les embryons inclus dans l'étude venaient d'accomplir le 4^e cycle embryonnaire depuis 0 à 12 heures et étaient tous au stade 16-32 cellules. Ils avaient donc tous théoriquement réalisé leur activation du génome embryonnaire.

Une analyse transcriptomique via le séquençage de l'ARN (*RNA-seq*) a été réalisée afin de quantifier l'abondance des transcrits présents dans chaque groupe d'embryons, permettant ainsi d'analyser les différences de niveau d'expression des gènes entre les catégories.

Nous réaliserons donc d'abord une comparaison statistique du transcriptome des quatre catégories morphocinétiques d'embryons puis nous étudierons de manière plus précise la comparaison des deux catégories présentant le plus de différences transcriptomiques.

Pour cette dernière comparaison nous nous intéresserons au profil des gènes différentiellement exprimés dans chacune des deux catégories ainsi qu'aux voies fonctionnelles enrichies par ces gènes.

Enfin, nous discuterons des questions que soulèvent ces résultats.

Le travail de cette thèse a été financé par le CRB-Anim, une infrastructure nationale financée par le programme Investissements d'Avenir de l'Agence Nationale de la Recherche (ANR).

Première partie : Bibliographie

I. Le développement embryonnaire précoce chez les bovins

1. L'ovocyte avant la fécondation

Avant d'être mature prêt à être fécondé, l'ovocyte passe par différentes phases de développement participant au processus que l'on appelle l'ovogénèse.

Comme expliqué par Picton (2001) et Van den Hurk et Zhao (2005), dans un premier temps, pour donner des ovocytes, les ovogonies doivent entrer en méiose.

Celle-ci n'est cependant pas complète et s'arrête au stade diplotène de la prophase I pour y rester en état de latence. Ce blocage sera levé par les signaux cellulaires induits par l'ovulation et la méiose reprendra.

A ce moment de l'ovogénèse, le noyau de l'ovocyte est appelé vésicule germinale (GV) (Dekel, 1995).

On dit que l'ovocyte est au stade GV et va alors subir trois types de maturation : nucléaire, cytoplasmique et moléculaire (Sirard *et al.*, 2006).

a. La maturation nucléaire de l'ovocyte

La maturation nucléaire de l'ovocyte correspond en fait à la reprise de la méiose.

Ce processus va permettre la diversité génétique ainsi que la réduction du bagage chromosomique de la cellule à 1N, c'est-à-dire à une copie de chaque chromosome (Van den Hurk et Zhao, 2005).

D'après Mermillod et Marchal (1999), le premier signe visible de la reprise du processus méiotique est le plissement de la membrane de la vésicule germinale. Les pores de cette enveloppe nucléaire disparaissent puis elle se rompt (Germinal Vesicle Breakdown ou GVBD) et disparaît rapidement (quatre à six heures chez les bovins).

La condensation des chromosomes a lieu en parallèle, les microtubules s'organisent en fuseau mitotique et la plaque métaphasique se forme.

Dekel (1995) et Van den Hurk et Zhao (2005) indiquent que l'ovulation (plus précisément le pic de LH) permet le déblocage de la méiose, et que l'expulsion du premier globule polaire signe la fin de cette première division méiotique.

Jones (2004) a précisé que la reprise de la méiose repose sur l'activation du MPF (*Maturation Promoting Factor*). Ce dernier est un hétérodimère composé d'une sous-unité catalytique, la kinase p34^{cdc2}, aussi appelée Cdk1 (*Cyclin-Dependent Kinase 1*) et d'une sous-unité régulatrice, la cycline B1.

Chez les mammifères, l'activation du MPF dépend de l'activité des kinases Wee1 et Myt1 et des phosphatases Cdc25. Les kinases Wee1 et Myt1 entraînent la phosphorylation de Cdk1 ce qui cause l'inactivation du MPF et l'arrêt de la méiose. Au contraire, la phosphatase Cdc25 entraîne l'activation du MPF par la déphosphorylation de Cdk1 : la méiose reprend.

Par conséquent, l'activation du MPF s'effectue lorsque les niveaux de Wee1/Myt1 sont bas et que les niveaux de Cdc25 sont élevés.

Une deuxième méiose est initiée mais là, de nouveau, elle se bloque, cette fois-ci au stade de métaphase II (MII). C'est la fécondation par le spermatozoïde qui lèvera ce nouveau blocage. Celle-ci aura pour conséquence l'expulsion du deuxième globule polaire, laissant l'ovocyte avec une seule copie de ses chromosomes, et donc prêt à fusionner avec le bagage génétique du spermatozoïde.

Cette fusion de deux cellules haploïdes (c'est-à-dire 1N) formera ainsi une cellule diploïde (2N), prête pour affronter la suite de son développement : le clivage embryonnaire (Saint-Dizier et Chastant-Maillard, 2014).

b. La maturation cytoplasmique de l'ovocyte

L'ovocyte subit également une maturation d'un point de vue cytoplasmique.

Hyttel *et al.* (1997) ont fait état d'une augmentation des réserves lipidiques. Celle-ci est d'une grande importance car les lipides seront une source d'énergie utilisable par l'embryon lors de son développement précoce. On observe également une diminution de taille de l'appareil de Golgi.

Les granules corticaux (vésicules cytoplasmiques de l'ovocyte) migrent vers la membrane plasmique et vont jouer un rôle dans l'inhibition de la polyspermie.

En effet, l'entrée du spermatozoïde dans l'ovocyte provoque une libération périodique du calcium stocké dans le réticulum endoplasmique via un messenger secondaire intracellulaire, l'inositol triphosphate (IP₃). Ces oscillations répétées de la concentration calcique intracellulaire provoquent une vague d'exocytose relâchant le contenu des granules corticaux vers l'extérieur de l'ovocyte.

Cela a pour effet de modifier biochimiquement des protéines de la zone pellucide (matrice extracellulaire glycoprotéique sulfatée entourant l'ovocyte) (Abbott et Ducibella, 2001).

Selon Sun (2003), il s'agit en fait de modification des sites de liaison au spermatozoïde sur la protéine ZP3 (*Zona Pellucida Sperm-Binding Protein 3*) par le N-acétylglucosaminidase et le durcissement de la zone pellucide par le clivage de la protéine ZP2 (*Zona Pellucida Sperm-Binding Protein 2*) grâce à des protéinases.

La zone pellucide devient ainsi imperméable à l'entrée d'autres spermatozoïdes, évitant la polyspermie qui conduirait à la création d'un zygote polyploïde qui ne se développerait pas ou posséderait des malformations importantes.

c. La maturation moléculaire de l'ovocyte

Dans une cellule, l'information génétique portée par l'ADN est transcrite dans le noyau sous forme d'ARN lors du phénomène appelé la transcription. Pour former les protéines nécessaires au métabolisme cellulaire, les ARN messagers (ARNm) sont traduits dans le cytoplasme par les ribosomes. Les détails de la transcription et de sa régulation seront traités plus en détails dans les parties II et III de cette thèse.

La maturation moléculaire est définie comme étant l'accumulation de molécules spécifiques (ARN et protéines) au sein de l'ovocyte en préparation aux événements suivant la fécondation (Sirard *et al.*, 2006).

➤ Les ARN

D'après Memili *et al.* (1998) et Tomek *et al.* (2002), les ovocytes au stade GV ont une transcription active (ARN ribosomiaux et/ou messagers). Cette transcription devient ensuite plus faible aux alentours de la GVBD, pour s'arrêter complètement lors du blocage de l'ovocyte en MII. L'ovocyte fabrique donc ses propres ARNm, appelés ARNm maternels.

Certains d'entre eux, comme les ARNm encodant les protéines de la zone pellucide (ZP1, ZP2 et ZP3), sont transcrits et traduits pendant la croissance de l'ovocyte.

D'autres ARNm maternels sont transcrits et accumulés pendant la croissance de l'ovocyte pour n'être traduits que plus tard lors de la maturation, la fécondation et/ou le développement embryonnaire.

En plus de la capacité de produire des protéines pour un usage immédiat, l'accumulation d'un grand pool d'ARNm permet à la cellule de croître et de se diviser sans activité génique, jusqu'à ce que le pool soit épuisé. La disponibilité des informations sous forme de molécules d'ARNm stables est donc particulièrement importante pour la suite du développement embryonnaire car pendant cette période, l'activité transcriptionnelle est faible, voire absente (Sommerville, 1999).

➤ Les protéines

Pendant l'ovogenèse, l'ovocyte synthétise et accumule des protéines jouant un rôle immédiat et/ou futur.

Par exemple, pendant l'ovogenèse les gènes spécifiques à ZP1, ZP2 et ZP3 sont transcrits et, suite à leur traduction, les protéines sont sécrétées et assemblées pour former la zone pellucide, qui jouera un rôle prépondérant dans le processus de fécondation (El-Mestrah *et al.*, 2002).

De même, Simon *et al.* (1989) ont montré l'importance de cette synthèse protéique pour la transition du stade GV au stade métaphase II de l'ovocyte. En effet, l'ajout de cycloheximide, un inhibiteur de la traduction, à des ovocytes de bovins, inhibe la GVBD, bloquant ainsi le déroulement normal de l'ovogénèse.

Tomek *et al.* (2002) ont par ailleurs montré une forte augmentation de la traduction, notamment via l'augmentation de la polyadénylation, au moment de la GVBD, laissant entendre que cette étape serait donc protéines-dépendante.

d. La régulation des ARN maternels

Les ARN maternels peuvent être stockés afin d'être utilisés plus tard au cours du développement de l'ovocyte ou de l'embryon.

Pour pouvoir être emmagasinés, les ARNm doivent remplir deux conditions : avoir une queue poly(A) courte et une séquence CPE (*Cytoplasmic Polyadenylation Element*).

➤ La polyadénylation et le masquage des ARNm

La polyadénylation correspond à l'ajout de résidus adényliques sur l'extrémité 3' des ARNm eucaryotes, afin de constituer la queue poly(A) qui servira notamment à traduire correctement cet ARN (voir section II.3.b pour plus de détails sur le mécanisme de la polyadénylation).

Dans le cas des ARNm maternels de l'ovocyte, la polyadénylation permet la traduction de certains ARNm spécifiques en fonction de la longueur de leur queue poly(A). L'activation de la traduction est associée aux ARNm ayant de longue queue poly(A) (80 à 150 nucléotides) alors que la répression de la traduction est observée chez les ARNm ayant une petite queue poly(A) (environ 20 nucléotides) (Richter, 1999).

Bien que les pré-ARNm synthétisés dans le noyau ovocytaire reçoivent tous une queue poly(A), cette dernière est soit conservée pour une traduction immédiate de l'ARNm, soit raccourcie, via la déadénylation cytoplasmique. Les ARNm déadénylés peuvent alors être dégradés ou stockés pour un usage ultérieur.

S'ils sont stockés, ils seront éventuellement réadénylés, via la polyadénylation cytoplasmique, afin de permettre leur traduction à un moment plus opportun (Wahle, 1995).

De plus, d'après Richter (2008), il existe certaines séquences spécifiques pour la traduction de ces ARNm, qui se situent dans leur 3'UTR (*Untranslated Transcribed Region*, région transcrite non traduite) :

- La séquence CPE (élément de polyadénylation cytoplasmique) qui est liée par la protéine CPEB (protéine de liaison du CPE).
- La séquence HEX (hexanucléotide de polyadénylation)

Lorsqu'un ARN possède une courte queue poly(A) et une séquence CPE dans son 3'UTR, il peut être « masqué ».

En effet, les ARNm maternels peuvent être accumulés et stockés dans un état de traduction inactive grâce à un empaquetage en particules ribonucléoprotéiques (mRNP, *messenger Ribonucleoprotein Particle*). Cet empaquetage est possible grâce à des protéines de masquage qui s'associent à l'intérieur des ARNm indépendamment de la séquence (Sommerville, 1999).

L'organisation des mRNP est bien décrite chez le xénope et notamment par Richter (2008). Cette espèce possède deux séquences CPE et donc deux protéines CPEB qui les lient. La protéine Maskin vient interagir avec ces deux CPEB et les trois protéines ainsi associées interagissent avec le facteur eIF4E, qui est un des facteurs permettant l'initiation de la traduction. Les CPEB, Maskin et eIF4E forment un complexe qui circularise l'ARNm et prévient ainsi l'interaction de eIF4E avec le deuxième facteur d'initiation de la traduction eIF4G. La traduction est inhibée.

➤ Le démasquage : initiation de la traduction dépendante du CPE

L'initiation de la traduction des ARNm masqués/stockés est associée au phénomène de polyadénylation cytoplasmique.

Là encore ce phénomène a été bien étudié chez le xénope.

Selon Richter (2008), sous l'effet de la stimulation par la progestérone, la kinase Aurora A, phosphoryle les CPEB liés aux séquences CPE. Cela entraîne un recrutement du CPSF (*Cleavage and*

Polyadenylation Specificity Factor) sur la séquence HEX, provoquant la polyadénylation par la poly(A) polymérase cytoplasmique, GLD2 (*Germ Line Development 2*).

La queue poly(A) ainsi allongée fixe alors la protéine ePAB (*embryonic Poly(A) Binding protein*) qui, en interagissant avec eIF4G, déplace la protéine Maskin, libérant ainsi eIF4E. La traduction peut ainsi être initiée.

Cependant, il existerait en fait deux types d'initiation de la traduction : une rapide et une retardée. Un modèle d'initiation de la traduction chez le xénope faisant état de cette différence a été proposé par Piqué *et al.* (2008).

Cela dépendrait en fait de la distance entre le CPE et le HEX. Si le CPE ne chevauche pas la séquence HEX, la phosphorylation des CPEB suffirait à briser l'interaction entre Maskin et eIF4E et à induire la polyadénylation. L'initiation de la traduction est rapide.

Si au contraire, un des deux CPE chevauche la séquence HEX, une compétition se crée entre CPEB et CPSF, empêchant ce dernier de se fixer sur HEX.

Pour débloquer cette situation, la Cdk1 doit intervenir. Elle induit elle-même la phosphorylation des CPEB et provoque la destruction de celui gênant l'interaction CPSF-HEX. CPSF peut alors de nouveau se fixer sur sa séquence et la polyadénylation peut avoir lieu. L'initiation de la traduction est donc retardée.

Cette régulation est en revanche moins bien décrite chez les mammifères.

Reyes et Ross (2016) ont exploré les différences entre le xénope et les mammifères concernant ce mécanisme de régulation de la traduction dans les ovocytes en maturation. Il semblerait que les facteurs clés de la régulation (comme CPEB et CPSF par exemple) soient conservés dans les trois espèces étudiées de mammifères (souris, humains et bovins) et qu'ils jouent un rôle similaire.

Il existe cependant quatre CPEB différents chez les mammifères. Si chez la souris il semble que cela soit CPEB1 qui joue le plus fort rôle dans la polyadénylation, chez les bovins c'est CPEB4 qui est le plus exprimé.

Uzbekova *et al.* (2008) ont également trouvé qu'il existait trois types de kinase Aurora chez les bovins (AURKA, AURKB et AURKC).

La kinase Aurora A jouerait le même rôle chez cette espèce que chez le xénope en phosphorylant les CPEB afin d'initier la polyadénylation.

La déplétion d'AURKA par la metformine dans les ovocytes bovins a par ailleurs entraîné l'arrêt de la progression méiotique, en empêchant l'extrusion du premier globule polaire.

De plus, selon Reyes et Ross (2016), chez les bovins, ePAB n'est pas exprimé mais une autre protéine semble jouer son rôle : PABPC1 (*Poly(A) Binding Protein Cytoplasmic 1*).

Une autre différence entre les mammifères et le xénope concerne la protéine Maskin. En effet, il semblerait qu'une protéine alternative ou un système modifié la remplace, et potentiellement, que la régulation de la traduction soit indépendante de celle-ci chez les bovins.

C'est une autre protéine, contenant un domaine en spirale, appelée Maskin-like, qui tiendrait ce rôle. Elle ne lierait néanmoins pas CPEB, mais interagirait directement avec eIF4E, inhibant la traduction dans les ARNm masqués de la même façon que chez le xénope.

➤ La régulation temporelle du démasquage

Le moment du démasquage des ARNm maternels est lui-même régulé. Il peut intervenir à n'importe quel moment du développement en fonction des besoins de l'ovocyte et de l'embryon par la suite.

Plusieurs facteurs déterminent ce moment.

Tout d'abord, il existe en réalité deux types de séquence CPE en fonction du moment de la polyadénylation cytoplasmique et donc de l'activation de la traduction : le type maturation (mCPE) généralement associé à la polyadénylation pendant la maturation de l'ovocyte ; et le type embryonnaire (eCPE) généralement associé à la polyadénylation pendant le développement embryonnaire (Richter, 1999).

De plus, d'autres facteurs joueraient un rôle comme le nombre de copies de séquences CPE ou la distance entre le CPE et le HEX.

D'après Reyes et Ross (2016), un transcrite contenant deux CPE ou plus serait plus traduit qu'un ARN avec un seul CPE chez les mammifères.

Piqué *et al.* (2008) ont montré quant à eux l'importance de la distance entre la séquence CPE et la séquence HEX.

En effet, chez le xénope, la traduction serait plus importante lorsque cette distance atteint 25 nucléotides (distance optimale) alors qu'elle serait faible lorsque cette distance est supérieure à environ 120 nucléotides.

➤ La déadénylation des ARNm maternels pendant la maturation de l'ovocyte

Les ARNm maternels qui ne possèdent pas de séquence CPE, qu'ils aient une queue poly(A) courte ou longue, sont déadénylés par défaut pendant la maturation de l'ovocyte, entraînant un arrêt de leur traduction. Cette déadénylation est prise en charge par la déadénylase PARN (Poly(A)-specific Ribonucléase) qui serait en fait relâchée lors de la dissolution de la vésicule germinale (Richter, 1999).

Cependant, d'après Lequarre *et al.* (2004), chez les bovins, une déadénylation pendant la maturation de l'ovocyte ne serait pas toujours synonyme de dégradation.

De même, si les séquences CPE et HEX sont trop éloignés, la déadénylation est enclenchée. En effet, chez la souris ils doivent être distants de moins de 100-120 nucléotides afin de subir une polyadénylation cytoplasmique (Richter, 1999).

➤ La déadénylation des ARNm maternels pendant le développement embryonnaire

Les ARNm possédant une séquence CPE sont donc polyadénylés pendant la maturation de l'ovocyte. Cependant, ils peuvent par la suite du développement embryonnaire (après la fécondation) être déadénylés afin de stopper leur traduction (Richter, 1999).

Cette déadénylation dépend de protéines de liaison à l'ARN et de la voie de l'interférence et est impliquée dans l'activation du génome embryonnaire.

Ces mécanismes seront traités plus en détails dans la partie IV.3 de cette thèse.

2. Le clivage embryonnaire : du zygote au blastocyste

Comme mentionné dans la partie précédente, la première étape du développement embryonnaire est la fécondation de l'ovocyte par un spermatozoïde. Elle consiste en la fusion des gamètes mâle et femelle pour former une cellule diploïde et se déroule chez les mammifères dans l'oviducte (plus précisément à la jonction entre l'ampoule et l'isthme) (Saint-Dizier et Chastant-Maillard, 2014).

L'ovocyte ovulé est alors bloqué en métaphase II et c'est l'entrée du spermatozoïde dans celui-ci qui déclenchera la reprise de la méiose, ce qui aura pour résultat d'enclencher la première division cellulaire une trentaine d'heures après la fécondation.

Par la suite, les mitoses continuent, résultant en un deuxième cycle de divisions 10-12 heures après la première. Les cellules de l'embryon sont également appelées blastomères (Hanzen *et al.*, 1999 ; Saint-Dizier et Chastant-Maillard, 2014).

A ce stade, l'embryon est toujours entouré par la zone pellucide et celle-ci fera notamment le lien pour les échanges entre les cellules de l'embryon et celles de la muqueuse de l'oviducte (nutriments, signaux...) (Besenfelder *et al.*, 2020).

Hanzen *et al.* (1999) précisent que lors des premières divisions, l'embryon conserve la même taille, seul le nombre de cellules augmente. Par ailleurs, jusqu'au stade huit cellules, les cycles embryonnaires sont plutôt synchrones et les cellules ont une taille similaire et sont pluripotentes, c'est-à-dire non différenciées.

Mais à partir du stade huit-seize cellules (soit, le 4^e cycle embryonnaire), les divisions deviennent largement asynchrones générant *in fine* deux types de populations cellulaires : des cellules de petite taille (divisions plus rapides), et d'autres plus grandes (divisions plus lentes). Les plus petites cellules formeront la masse cellulaire interne (MCI), responsable de la formation du fœtus, alors que les plus grandes constituent le trophoctoderme, impliqué dans l'attachement de l'embryon à la paroi utérine et dans la formation des tissus extra-embryonnaires (Saint-Dizier et Chastant-Maillard, 2014).

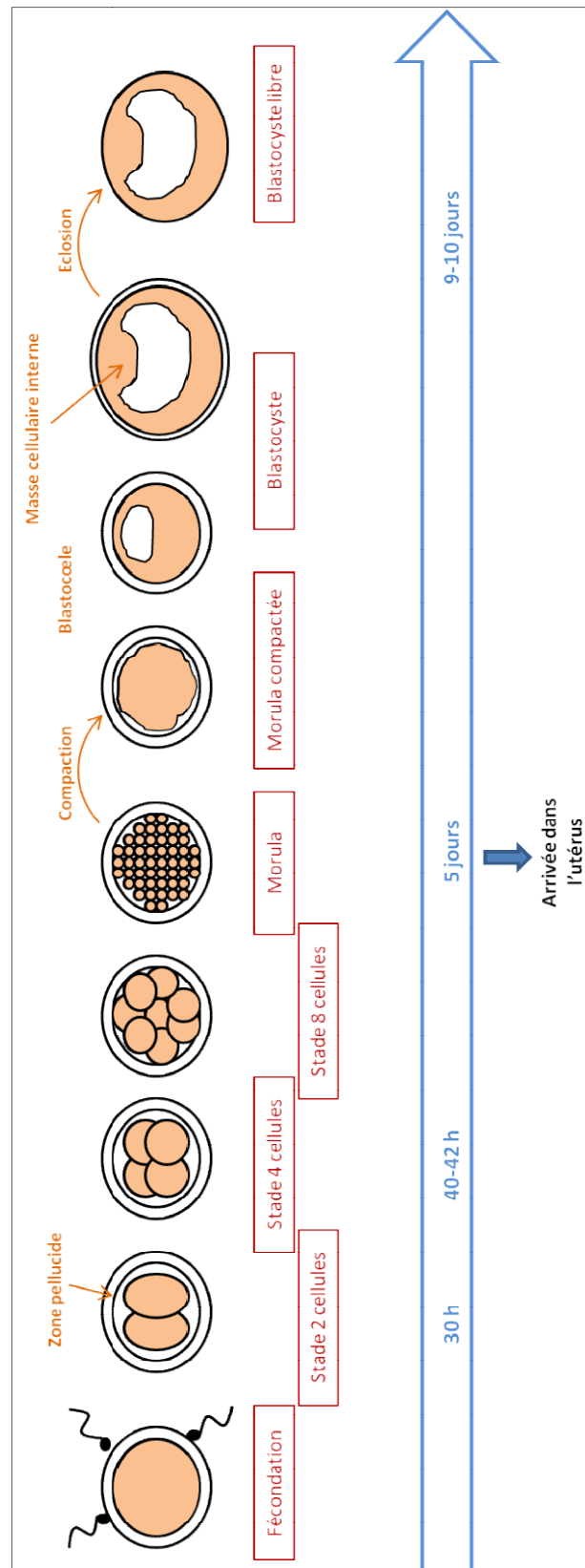
D'après Hanzen *et al.* (1999) et Saint-Dizier et Chastant-Maillard (2014), l'embryon arrive généralement dans l'utérus de la vache au 5^e jour de gestation, soit au stade seize cellules ou plus. Au stade 32-64 cellules, l'embryon prend le nom de morula et va subir une compaction. Cette dernière permet la mise en contact des cellules entre elles via des « gap-junctions ». Elle va également générer un appel d'eau dans la morula qui va ainsi former le blastocœle, on parle alors pour l'embryon de blastocyste.

L'expansion du blastocœle entraîne une augmentation de taille du blastocyste. De ce fait, la zone pellucide, sous pression, s'amincit et finit par se rompre, libérant le blastocyste : c'est l'éclosion. Ce phénomène arriverait environ entre le 9^e et le 10^e jour de gestation (Hanzen *et al.*, 1999).

Une représentation schématique de ces premières étapes du développement embryonnaire jusqu'au stade blastocyste est présentée en Figure 1.

Par la suite, le trophoctoderme débute ensuite une phase d'élongation de manière à atteindre la muqueuse utérine pour débiter l'implantation. D'après Saint-Dizier et Chastant-Maillard (2014), ces premiers contacts se dérouleraient au 20^e jour de gestation. Ces phases ne seront pas détaillées dans cette thèse.

Figure 1 : Le développement embryonnaire précoce chez les bovins, de la fécondation au stade blastocyste.



3. La transition materno-embryonnaire (MET)

La transition materno-embryonnaire (MET) représente la période au cours de laquelle le développement embryonnaire précoce, initialement régi par le matériel génétique maternel (ARN et protéines) stocké dans l'ovocyte au cours de sa maturation, passe progressivement sous le contrôle du génome embryonnaire.

Cette transition suppose le remplacement des transcrits maternels par les transcrits embryonnaires.

Pour cela, les transcrits et les protéines maternels vont diminuer en abondance via leur dégradation ou leur utilisation et l'embryon devra enclencher sa propre machinerie transcriptionnelle afin de produire ses propres ARN.

L'activation du génome embryonnaire (EGA) désigne ce phénomène selon lequel l'embryon commence à transcrire ces propres gènes.

L'EGA majeure, chez les bovins, se déroule entre le stade huit et seize cellules selon Frei *et al.* (1989) et Barnes et First (1991).

Avant celle-ci, l'embryon utilise donc le matériel maternel qui était stocké dans l'ovocyte : il est dans un état de silence transcriptionnel.

Selon Sirard (2010), celui-ci durerait environ trois jours chez les bovins.

D'après Schier (2007), il existe trois hypothèses causales expliquant ce silence.

La première concerne la méthylation de l'ADN qui réprime la transcription (voir IV.2.c).

La deuxième hypothèse serait que des facteurs de la machinerie transcriptionnelle soient inhibés ou détruits, comme la TBP (TATA Binding Protein) par exemple.

Et la troisième repose sur des modifications de l'enchaînement des phases du cycle cellulaire lors des premières divisions de l'embryon. En effet, à ce moment-là, les divisions sont très rapides grâce à un cycle cellulaire modifié. Celui-ci est en fait un enchaînement de phases S et M sans phases G1 et G2.

Or c'est durant ces phases G que la transcription se déroule habituellement, ces cycles rapides entraîneraient donc l'inhibition de la transcription.

Ce silence transcriptionnel met en lumière toute l'importance des stocks maternels qui vont combler cette lacune jusqu'à l'EGA.

En effet, Memili et First (1998) ont montré que bloquer la transcription par l'ARN polymérase II des gènes de profil embryonnaire avec de l' α -amanitine pendant les quatre premiers cycles cellulaires, n'affecte pas le taux d'embryons qui se développent jusqu'au stade huit-seize cellules.

Les réserves maternelles permettent donc à elles seules le développement jusqu'à ce stade.

Toutefois, aucun blastocyste n'est obtenu ce qui démontre que les transcrits embryonnaires sont tout de même nécessaires pour la suite du développement.

Cependant, et comme nous le verrons plus précisément dans la partie IV.1, il existe également une activation mineure du génome embryonnaire puisque certains auteurs ont retrouvé des traces de transcription embryonnaire dès le stade une cellule (Memili et First, 1999).

Dans ce concept de silence transcriptionnel, les auteurs ne prennent pas en compte cette activation mineure, du fait de sa moindre importance, puisqu'elle n'est pas une étape limitante du développement embryonnaire, les embryons dont la transcription a été bloquée atteignant tout de même le stade huit cellules.

II. La transcription chez les eucaryotes

Chez les eucaryotes, l'information génétique portée par l'ADN s'exprime sous forme de protéines ou d'ARN fonctionnels (ARN de transfert, ARN ribosomiaux, ARN interférents).

Les ARN sont produits à partir de l'ADN via un phénomène appelé transcription. Celle-ci a lieu dans le noyau des cellules grâce à l'action d'ARN polymérases.

Les ARN sont alors exportés du noyau via les pores nucléaires vers le cytoplasme. Les ARN dits messagers y sont alors transformés en protéines par les ribosomes via le phénomène appelé traduction.

Comme vu précédemment, la transcription est un mécanisme clé dans la maturation ovocytaire et le développement embryonnaire.

Elle est d'abord nécessaire à la constitution du stock d'ARNm maternels dans l'ovocyte afin de permettre à l'embryon de se développer jusqu'à la transition materno-embryonnaire, puis à l'embryon afin d'assurer la production de ses propres transcrits par la suite.

Comprendre comment elle se déroule et surtout comment elle est régulée permet d'éclairer les événements qui sont en jeu pour permettre l'activation du génome embryonnaire.

1. Le mécanisme général de la transcription

Richard *et al.* (2012) décrivent la transcription comme la synthèse d'un brin d'ARN à partir d'un brin d'ADN matrice par une ARN polymérase ADN-dépendante, par ajout de nucléotide complémentaire.

Elle se décompose en 3 étapes : l'initiation, l'élongation et la terminaison et aboutit, selon Leroux et Tosser-Klopp (2000), à la formation de différents types d'ARN :

- ARN messager (ARNm), traduit ensuite en protéines
- ARN de transfert (ARNt), transporteur des acides aminés jusqu'au ribosome afin de traduire l'ARNm en protéine
- ARN ribosomal (ARNr), possédant un rôle structural au sein des ribosomes
- Petit ARN interférent et micro-ARN (ARNsi et ARNmi), jouant un rôle dans l'expression des gènes par leur activité le plus souvent catalytique
- Petit ARN nucléaire (ARNsn), participant à l'épissage alternatif

D'après Leroux et Tosser-Klopp (2000), le type d'ARN produit est fonction du type d'ARN polymérase qui a réalisé la transcription. En effet, chez les eucaryotes, il existe 3 principales catégories d'ARN polymérases :

- L'ARN polymérase I, produisant les ARNr 5,8S ; 18S et 28S
- L'ARN polymérase II, produisant les ARNm et les ARNsn
- L'ARN polymérase III, produisant les ARNt, l'ARNr 5S et quelques ARNsn

L'initiation de la transcription fait intervenir des régions spécifiques de l'ADN.

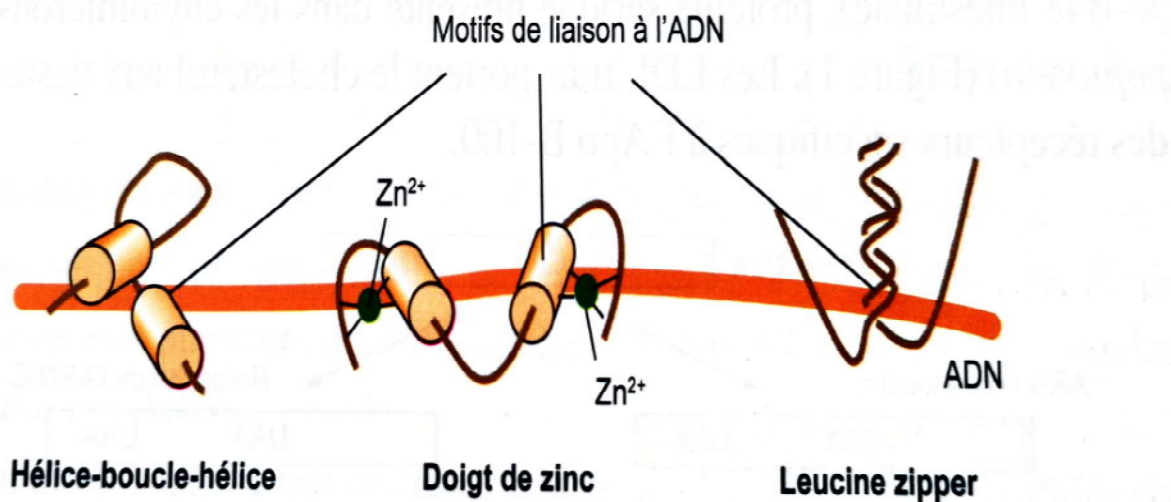
D'après Kornberg (2007), les ARN polymérases sont incapables de se fixer sur l'ADN à transcrire sans facteurs de transcription.

Selon Venters et Pugh (2009), les facteurs généraux de transcription (GTF) reconnaissent une séquence promotrice sur l'ADN et y recrutent l'ARN polymérase formant ainsi le complexe d'initiation.

Certains de ces facteurs généraux possèdent des activités hélicase et protéine kinase permettant de dérouler l'ADN au voisinage du site d'initiation et de libérer l'ARN polymérase une fois celle-ci recrutée pour que l'élongation puisse commencer.

Pabo et Sauer (1992) ont d'ailleurs étudié leurs domaines de fixation à l'ADN. Ils sont très structurés et contiennent au moins une hélice α selon plusieurs motifs possibles : par exemple doigts de zinc, hélice-boucle-hélice ou leucine zipper (Figure 2).

Figure 2 : Représentation schématique des motifs de liaison à l'ADN des facteurs de transcription (Richard *et al.*, 2012).



L'élongation correspond à la synthèse d'un ARN complémentaire au brin d'ADN matrice par l'ARN polymérase, par ajout de nucléotide complémentaire. La seule différence réside dans le fait qu'un brin d'ARN ne contient pas de thymines, qui sont remplacées par des uraciles.

Enfin, on appelle terminaison le moment où l'ARN polymérase cesse la synthèse de l'ARN et qu'elle et l'ARN naissant se détachent de la matrice d'ADN.

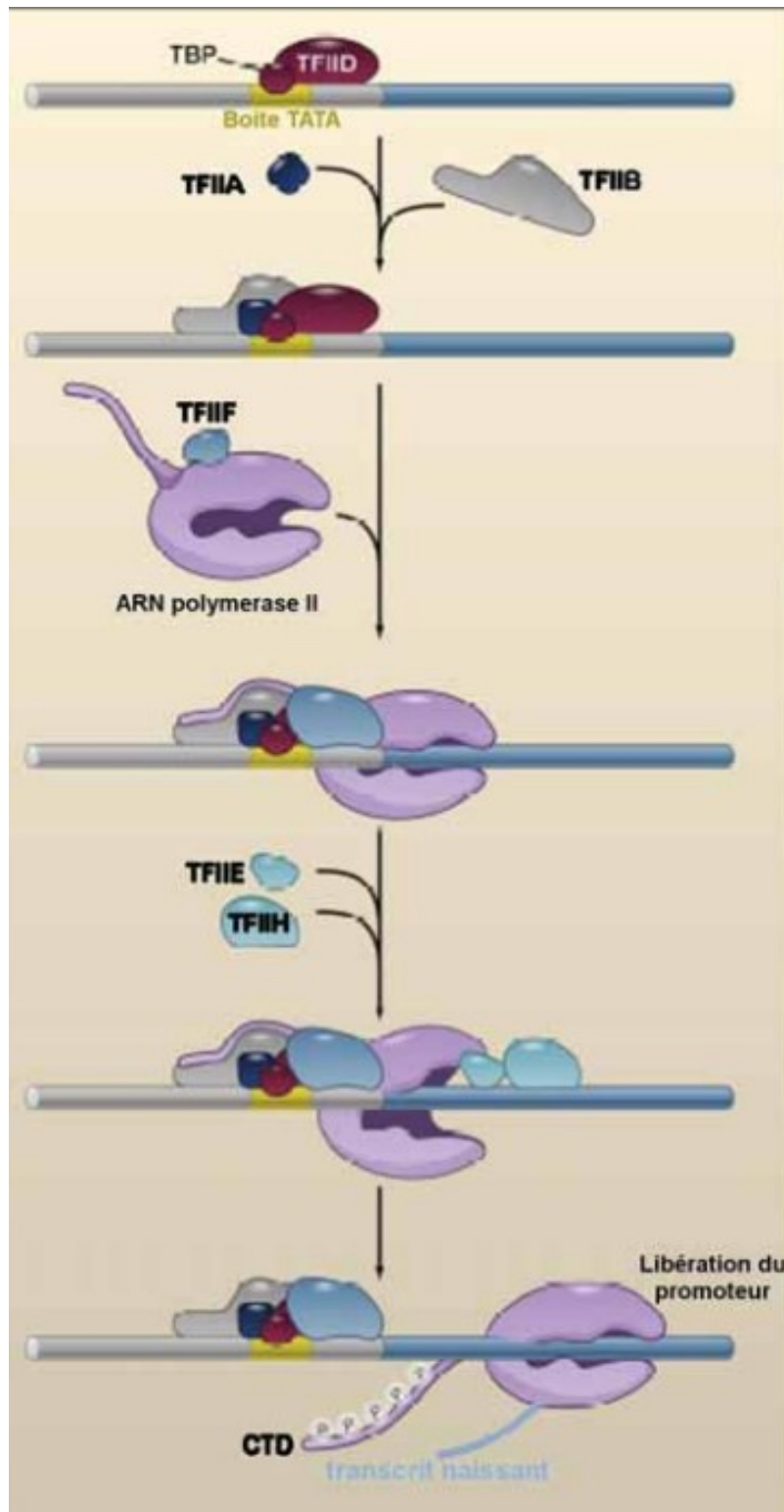
2. Illustration avec l'exemple de la transcription par l'ARN polymérase II chez les eucaryotes

D'après Venters et Pugh (2009), l'ARN polymérase II possède six facteurs de transcription généraux : TFIIA, TFIIB, TFIID, TFIIE, TFIIIF et TFIIH. Le facteur de transcription TBP (*TATA Binding Protein*) qui appartient au complexe TFIID reconnaît le promoteur contenant une boîte TATA, situé environ 25 nucléotides en amont du site d'initiation. TFIIA, TFIIB et TFIIIF se fixent alors à leur tour sur le promoteur et recrutent l'ARN polymérase II. Celle-ci recrute alors TFIIE qui recrute lui-même TFIIH.

Le complexe d'initiation ainsi formé reste cependant instable. En effet, le promoteur est en conformation fermée (structure en double hélice). TFIIH vient alors ouvrir l'ADN double brin grâce

à son activité hélicase ATP-dépendante. Cette dénaturation locale de l'ADN fait passer le promoteur en conformation ouverte, l'ARN polymérase peut accéder au brin matrice et commencer l'élongation (Figure 3).

Figure 3 : L'initiation de la transcription par l'ARN polymérase II (Lévine, 2011).



La terminaison de la transcription par l'ARN polymérase II est couplée fonctionnellement avec un évènement de maturation : la polyadénylation (voir II.3.b) (Logan *et al.*, 1987).

D'après Kuehner *et al.* (2011), suite à cet évènement, l'ARN nouvellement formé est clivé et un ARN résiduel auquel peut parfois être encore accroché le complexe de transcription (et donc l'ARN polymérase II) est ainsi créé.

Cet ARN résiduel possède une extrémité 5' libre, sans coiffe (protection ajoutée lors de la maturation de l'ARNm, voir II.3.a), ce qui permet à l'exoribonucléase 5'-3' 2 (Xrn2) de s'y fixer pour le détruire. Lorsque Xrn2 atteint l'ARN polymérase II en élongation, cela entraîne la terminaison en libérant cette dernière du brin d'ARN résiduel.

3. La maturation des ARN

Les ARN issus de la transcription subissent différents types de modifications post-transcriptionnelles permettant leur maturation avant leur traduction.

Les plus importantes modifications sont celles que subissent les ARN messagers. En effet, les produits de la transcription sont en fait des ARN pré-messagers et subissent trois phénomènes de maturation différents.

a. L'addition de la coiffe à l'extrémité 5'

Selon Shuman (2000), l'extrémité 5' du transcrit primaire est modifiée par ajout d'une 7-méthyl guanosine selon une réaction catalysée par un polypeptide bi-fonctionnel (possédant les activités triphosphatase et guanylyltransférase) et une méthyltransférase.

En effet, ces enzymes se lient toutes les deux spécifiquement et directement au domaine carboxy-terminal (CTD) de l'ARN polymérase II, lorsque celui-ci est phosphorylé au cours de l'élongation, générant l'addition de la coiffe.

La Figure 4 schématise le mécanisme de celle-ci.

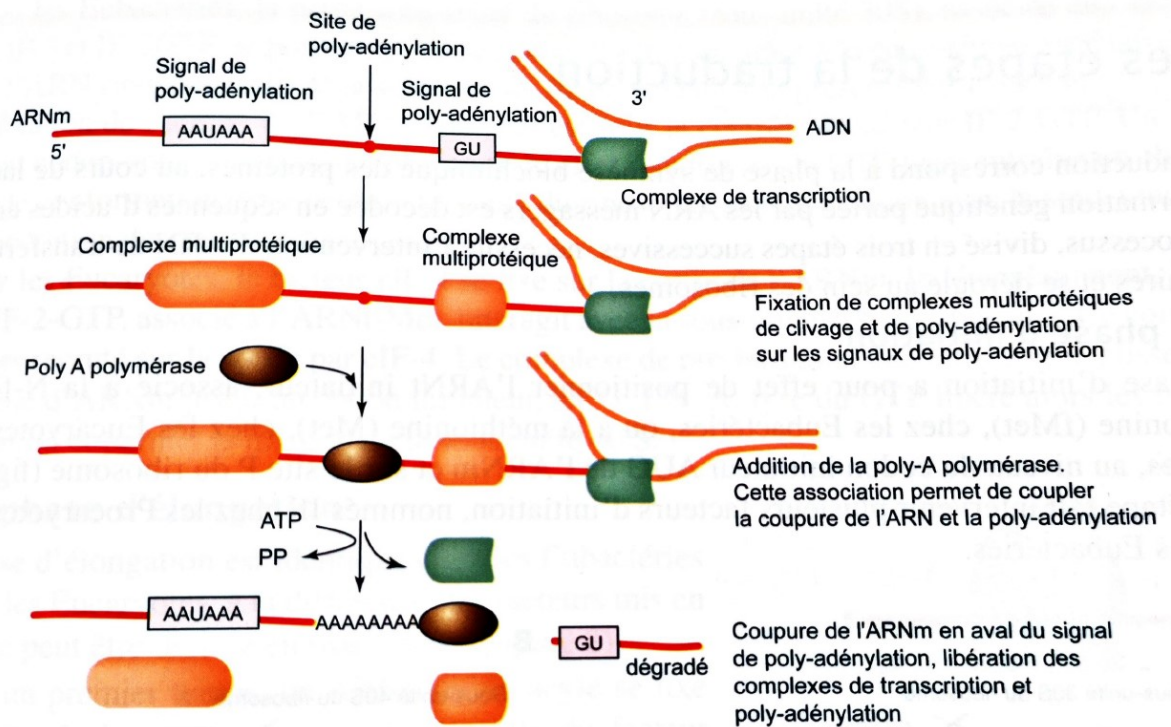
Cette coiffe définit ainsi l'extrémité 5' du premier intron, permettant un épissage adéquat (troisième type de maturation, voir II.3.c), participe au transport de l'ARNm dans le cytoplasme en le protégeant des nucléases et intervient dans l'initiation de la traduction (Richard *et al.*, 2012).

b. La polyadénylation en 3'

Hormis les ARNm des histones, l'extrémité 3' de tous les ARNm eucaryotes comporte une série de 50 à 200 résidus adényliques constituant la queue poly(A). Celle-ci permet la terminaison de la transcription mais participe également au transport des ARNm dans le cytoplasme, les protègent des dégradations et contribue à l'initiation de la traduction (Richard *et al.*, 2012).

D'après Kuehner *et al.* (2011), cette polyadénylation est initiée par la fixation de deux complexes protéiques, le facteur de spécificité de clivage et polyadénylation (CPSF) et le facteur stimulateur du clivage (CSTF), sur les signaux de polyadénylation (5'AAUAAA3' et une séquence riche en GU) encadrant le site de poly-adénylation. Cela déclenche le clivage de l'ARNm naissant au niveau de ce site de poly-adénylation et le recrutement d'une poly(A) polymérase qui forme alors la queue poly(A) (Figure 4).

Figure 4 : Les étapes de la polyadénylation des ARNm (Richard *et al.*, 2012).



c. L'épissage alternatif

Dye *et al.* (2006) ont indiqué que les gènes sont composés de séquences dites codantes, les exons, séparées entre elles par des séquences non codantes, les introns.

L'épissage consiste en l'excision des séquences non codantes, afin de rapprocher les séquences codantes entre elles.

Pour ce faire, les protéines SR (*Serine and arginine-rich proteins*) se lient sur des séquences spécifiques situées au sein des exons (appelées ESE) et jouent un rôle clé dans la reconnaissance des signaux d'épissage et dans l'assemblage de la machinerie d'épissage : le spliceosome.

Les spliceosomes sont des complexes comprenant de petits ARN nucléaires (ARNsn) U1, U2, U4, U5 et U6 assemblés avec des protéines spécifiques pour former des petites particules ribonucléoprotéiques nucléaires (snRNP). Les snRNP reconnaissent des séquences d'ARN spécifiques situées aux frontières exons/introns.

Le spliceosome va alors rapprocher une adénosine de l'intron et son extrémité 5' de manière à former une sorte de lasso, libérant ensuite l'intron en réunissant l'extrémité 3' de l'exon précédent avec l'extrémité 5' de l'exon suivant (Figure 5).

Cet épissage est dit alternatif car selon les introns éliminés et les exons gardés, il est possible d'obtenir plusieurs combinaisons différentes d'ARNm à partir d'un même gène (Leroux et Tossier-Klopp, 2000) (Figure 6).

Figure 5 : Le mécanisme de l'épissage par le spliceosome (Richard et *al.*, 2012).

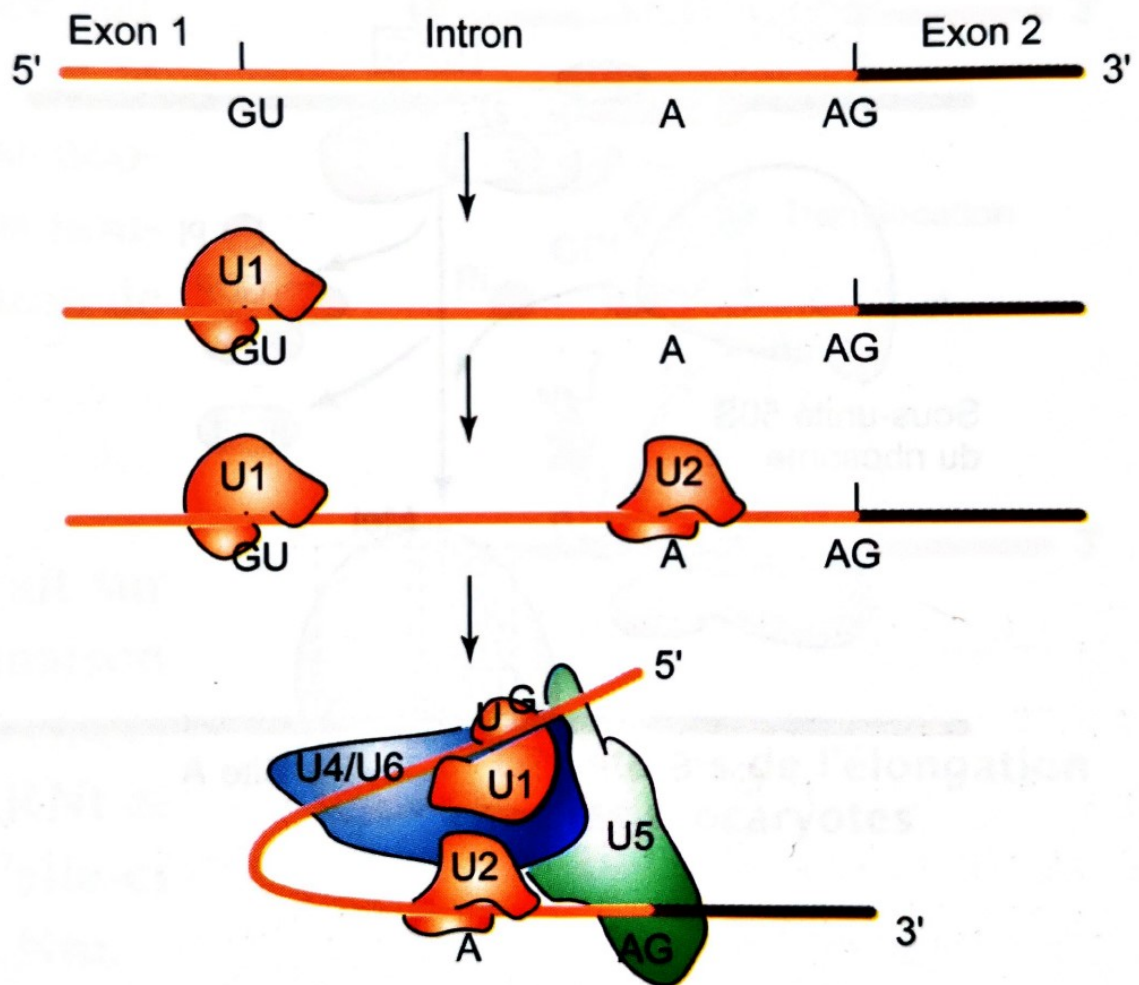
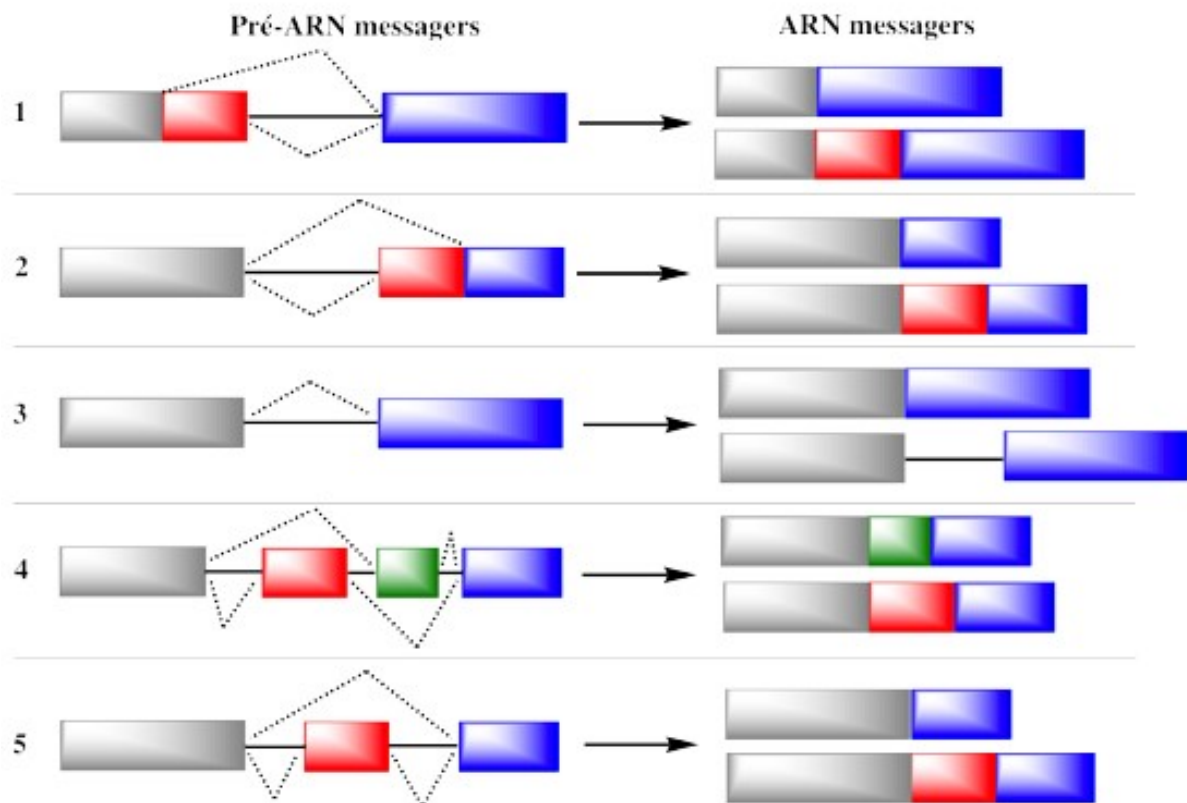


Figure 6 : Le principe de l'épissage alternatif (Jaspard, 2012).



III. Régulation de la transcription

Au vu de son importance et de sa complexité, la transcription est un mécanisme hautement régulé dans la cellule.

Cette régulation se fait à plusieurs niveaux.

Elle peut en effet concerner l'initiation de la transcription avec une modulation de l'accessibilité à l'ADN ou des séquences cis-régulatrices mais peut également impliquer des mécanismes post-transcriptionnels comme la voie de l'interférence.

1. La modulation de l'accessibilité à l'ADN

La mise en place des facteurs de transcription puis de toute la machinerie transcriptionnelle implique que l'ADN, empaqueté dans la chromatine, soit accessible.

Or cet ADN ainsi structuré se présente, dans le noyau de la cellule, enroulé autour de noyaux protéiques d'histones (octamère, formé de deux molécules d'histones H2A, deux H2B, deux H3 et deux H4). Cette association chromatine-histone est appelée nucléosome (Wolffe, 2001 ; Trotter et Archer, 2007).

Les modifications du degré de condensation de la chromatine interviennent donc dans les processus de régulation transcriptionnelle.

Tout d'abord, comme expliqué par Berger (2001), les histones peuvent subir des modifications covalentes du type acétylation, méthylation ou phosphorylation qui déstabilisent physiquement la chromatine en perturbant les interactions entre nucléosomes ou entre les histones et l'ADN.

Par exemple, l'acétylation des résidus lysine des queues des histones au niveau des régions promotrices de gènes spécifiques a pour effet de réduire les charges positives des histones, et donc d'affaiblir leur interaction avec l'ADN chargé négativement, le rendant plus accessible aux facteurs de transcription et donc, par la suite, à l'ARN polymérase.

De la même façon, Narlikar *et al.* (2002) et Venters et Pugh (2009) ont fait état de complexes protéiques de remodelage ATP-dépendants capables de modifier la structure et la répartition des nucléosomes sur l'ADN. Ils permettent leur glissement sur un même brin ou leur transfert sur un autre brin. Ils peuvent également altérer la structure des nucléosomes en déstabilisant les interactions ADN-histones.

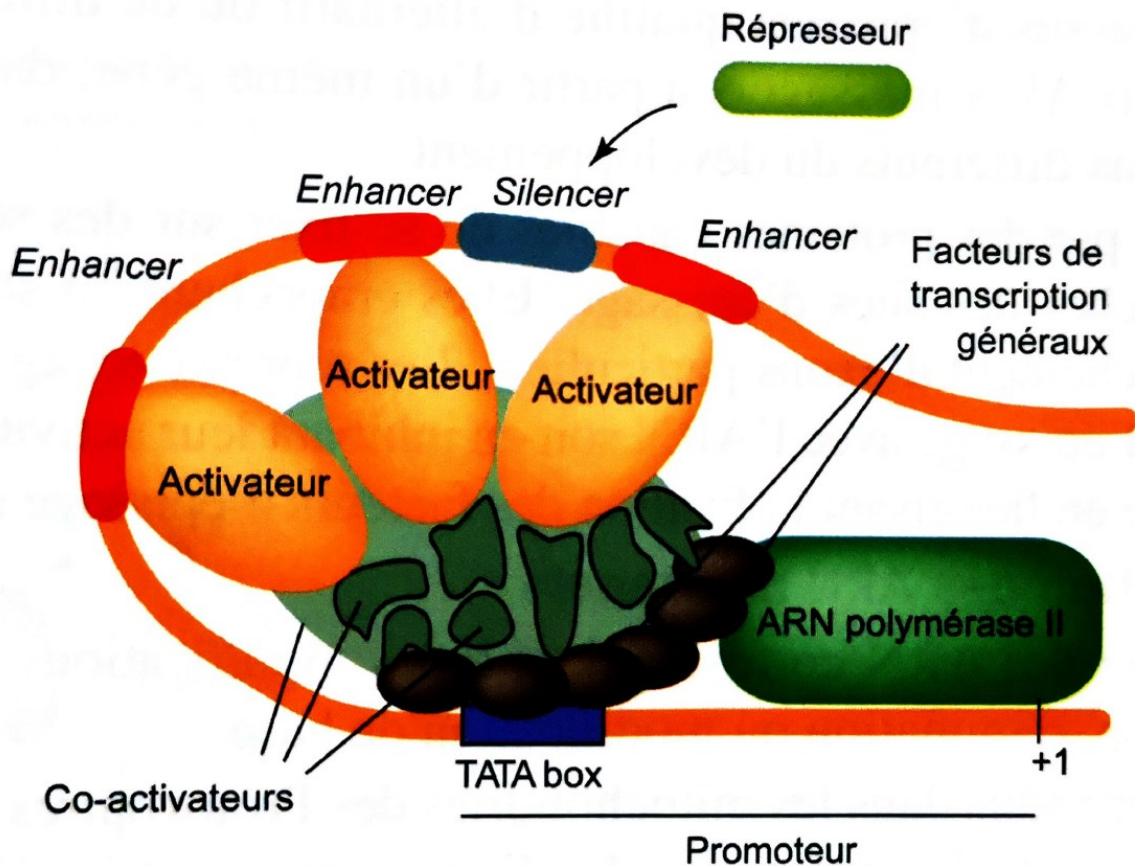
SWI/SNF (*Switch/Sucrose Non-Fermentable*) est un exemple de complexe de remodelage de la chromatine ATP-dépendant qui augmente l'accessibilité des facteurs de transcription aux séquences régulatrices en désorganisant ainsi les nucléosomes.

2. Les séquences cis-régulatrices de l'ADN

Les facteurs de transcription se fixent sur les promoteurs de l'ADN pour recruter l'ARN polymérase. Mais ils peuvent être modulés par des activateurs ou des répresseurs se fixant sur des régions encore plus en amont sur le brin d'ADN : les amplificateurs (ou *enhancers*) et les silencieux (ou *silencers*). Ces régions régulatrices peuvent aussi bien se trouver à proximité du promoteur, comme à plusieurs milliers de paires de bases de celui-ci. L'ADN forme alors une boucle, permettant à ses activateurs ou répresseurs d'interagir, soit directement, soit via des co-activateurs, avec le complexe d'initiation formé par les facteurs de transcription généraux et l'ARN polymérase (Ladunga, 2010) (Figure 7).

Parmi le nombre de co-activateurs connus pour cibler directement l'ARN polymérase II, le complexe Médiateur a émergé comme étant le plus crucial. Malik et Roeder (2005) ont décrit ce complexe protéique, constitué d'environ 30 sous-unités, comme jouant le rôle de pont entre les activateurs fixés sur les *enhancers* et les facteurs de transcription généraux, fixés sur l'ADN et l'ARN polymérase II. Cette interaction promouvrait la formation du complexe d'initiation de la transcription.

Figure 7 : Le mode d'action des activateurs et des répresseurs de la transcription via leurs interactions avec les séquences amplificatrices (*enhancers*) ou silencieuses (*silencers*) (Richard et al., 2012).



3. Le phénomène d'interférence d'ARN

Il existe une régulation des ARN de la cellule et donc de la modulation de l'expression des gènes via le phénomène d'interférence d'ARN.

Selon Khraiweh et al (2012), ce mécanisme repose sur la reconnaissance d'ARN double brin par un complexe nucléasique, la protéine Dicer, dans le cytoplasme.

Les ARN double brin peuvent être de deux origines : exogène ou endogène.

Les ARN double brin exogènes sont principalement des ARN viraux.

Les ARN double brin endogènes sont quant à eux de deux catégories :

- Des ARN qui possèdent la capacité de former une structure tige-boucle
- Des ARN complémentaires issus d'ARN transcrits à partir de répétitions inversées ou de transposons

Dautry et Ribet (2004) et Cartew et Sontheimer (2009) ont expliqué que la protéine Dicer reconnaît ces ARN double brin et les clive en petits fragments d'ARN simple brin d'une vingtaine de nucléotides. Selon le type d'ARN double brin, le produit final est différent.

Les ARN double brin possédant une structure tige-boucle seront transformés en micro-ARN (ARNmi), alors que les ARN double brin complémentaires deviendront des *small interfering* ARN (ARNsi).

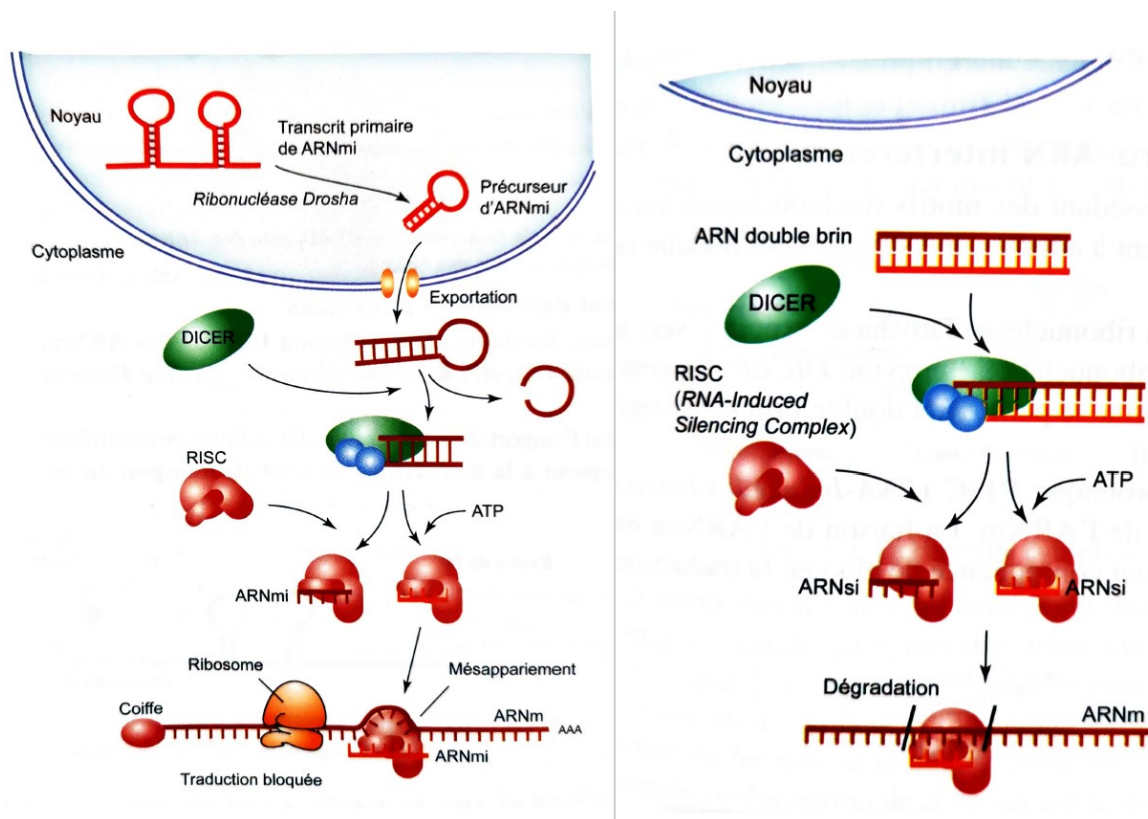
Ceux-ci sont alors pris en charge par un complexe protéique, le RISC (*RNA-Induced Silencing Complex*), qui les positionne sur l'ARNm dont ils sont les complémentaires.

Si l'appariement entre l'ARN interférent et l'ARN messager est parfait, le RISC induit une coupure endonucléolytique de l'ARNm qui conduira à sa destruction par des nucléases.

Si au contraire, l'appariement est imparfait, le RISC induit le blocage de la traduction de cet ARNm en protéines (dans la grande majorité des cas cela se fait via une dégradation exonucléolytique de celui-ci).

Généralement, le clivage endonucléolytique de l'ARNm est induit par les ARNsi et le blocage de la traduction l'est par les ARNmi mais ces phénomènes sont interchangeables entre les deux classes de petits ARN (Figure 8).

Figure 8 : Les voies de l'interférence d'ARN, rôle des ARNmi (à gauche) et des ARNsi (à droite) (Richard *et al.*, 2012).



La fonction de ce système serait, en plus de la modulation de l'expression génique, la protection de la cellule contre les infections virales, ainsi que le maintien de l'intégrité du génome en supprimant la mobilisation des transposons et l'accumulation d'ADN répétitif. En effet, la destruction des transposons rendrait les éléments répétitifs indisponibles pour des événements de recombinaison qui conduiraient à des translocations chromosomiques (Hannon, 2002).

La voie de l'interférence a également un grand rôle à jouer dans le développement embryonnaire précoce, car elle est soupçonnée d'être à l'origine de la dégradation du

transcriptome maternel lors de la MET. Ce phénomène sera étudié dans la partie IV.3.a de cette thèse.

➤ Zoom sur la régulation par les ARNmi

L'un des enjeux actuels est de savoir ce qu'il se passe une fois la protéine RISC et son ARNmi positionnés sur le brin d'ARNm. La traduction est bloquée, certes, mais de quelle façon ?

En effet, d'après Huntzinger et Izaurralde (2011), un consensus existe concernant la dégradation des ARNm ciblés (dégradation exonucléolytique) qui serait la voie principalement activée lors d'appariement imparfait entre l'ARNmi et l'ARNm.

Comme l'ont expliqué les auteurs, la protéine RISC et son ARNmi induisent une déadénylation de l'ARNm, provoquant le raccourcissement de la queue poly(A). L'ARN messenger ne peut ainsi plus être associé avec la PABPC (*Cytoplasmic Poly A Bonding Protein*), ce qui active le complexe de décoiffage DCP2.

L'extrémité 5' de l'ARNm est alors exposée à l'exonucléase cytoplasmique Xrn1 qui dégrade celui-ci.

Cependant, on ne peut exclure que certains ARNmi aient également une action au niveau de la traduction à proprement parler.

D'après Cartew et Sontheimer (2009), les ARNmi pourraient agir sur celle-ci par plusieurs voies : l'inhibition de l'initiation de la traduction, l'arrêt de l'élongation de la traduction, la terminaison prématurée de la traduction ou la dégradation de polypeptides naissants en cours de traduction.

Néanmoins, les mécanismes par lesquels l'ARNmi inhibe la traduction font l'objet d'un débat permanent. La question consistant à savoir si la répression se produit lors de l'initiation de la traduction ou après celle-ci n'a pas encore été résolue.

Certains auteurs comme Humphreys et al (2005) soutiennent la théorie du blocage de l'initiation en montrant que les ARNmi altèrent la fonction de la protéine eIF4E et ainsi bloquent la reconnaissance de la coiffe, première étape de la traduction par les ribosomes.

D'autres soutiennent que les ARNmi provoquent une dégradation rapide des polypeptides naissants (Nottrot et al., 2006) ou la dissociation prématurée des ribosomes (Petersen et al., 2006), favorisant l'hypothèse selon laquelle ce blocage arrive pendant l'élongation de la traduction.

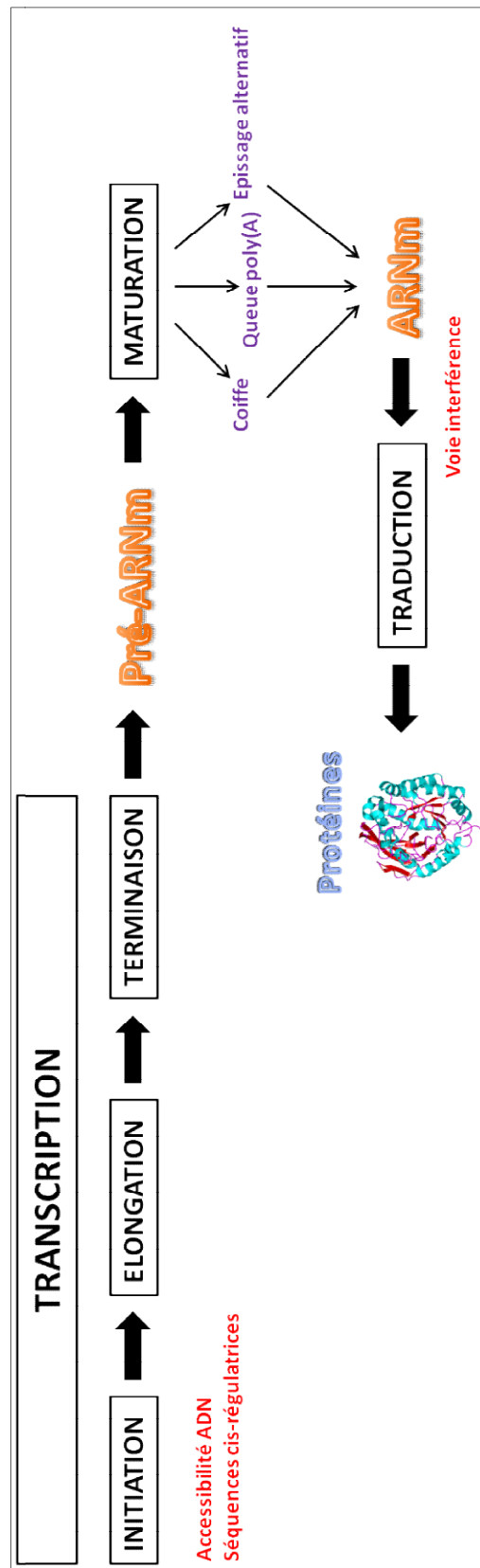
Enfin, un autre débat consiste à savoir si cette dégradation est concomitante du blocage de la traduction ou si elle est indépendante de cet événement.

Nous ne rentrerons pas dans les détails de cette question dans cette thèse.

Le mécanisme précis de régulation par les ARNmi semble donc loin d'être entièrement résolu et va certainement générer encore beaucoup d'interrogations (Cartew et Sontheimer, 2009).

Un schéma bilan des étapes de la transcription et de sa régulation est présenté en Figure 9.

Figure 9 : De l'ARNm à la protéine : les étapes de la transcription à la traduction et les régulations associées (en rouge).



IV. L'activation du génome embryonnaire

Lors de la transition materno-embryonnaire, le développement précoce, jusqu'alors sous contrôle du transcriptome maternel, passe sous le contrôle du transcriptome embryonnaire.

Le mécanisme selon lequel l'embryon commence à transcrire ses propres gènes est appelé l'activation du génome embryonnaire (EGA).

Cette activation implique une diminution de l'abondance des transcrits maternels par dégradation ou par leur utilisation afin de les remplacer par les transcrits de l'embryon.

Ces deux phénomènes sont inter reliés et s'inscrivent simultanément au sein de la transition materno-embryonnaire.

Cette partie sera traitée, autant que possible, de manière comparative entre différentes espèces : xénope, souris, porc et bovin.

1. Le moment de l'activation : activations mineure et majeure

L'activation majeure du génome embryonnaire est expérimentalement déterminée par le stade où, si on cultive l'embryon avec un inhibiteur de la transcription, il arrête de se développer. La transcription d'au moins un gène embryonnaire est donc nécessaire pour la poursuite immédiate du développement.

Le moment de l'activation majeure du génome embryonnaire diffère d'une espèce à l'autre.

Chez la souris elle se déroule au stade deux cellules (Aoki *et al.*, 1997), au stade quatre cellules chez le porc (Xu *et al.*, 2012) et au stade huit cellules chez l'homme (Gao *et al.*, 2018).

Chez les bovins, elle se produit entre le stade huit cellules et le stade seize cellules (Frei *et al.*, 1989 ; Barnes et First, 1991), soit entre le 3^e et le 4^e cycle embryonnaire.

La transcription est à ce stade régulée classiquement comme dans une cellule somatique normale, et la maturation des transcrits est complète.

Cependant, Viuff *et al.* (1998) ont découvert la présence d'ARN ribosomiaux dans des embryons bovins dès le stade quatre cellules grâce à une technique alliant l'hybridation *in situ* en fluorescence et une coloration argentique.

Un an plus tard, Memili et First (1999) ont étudié l'incorporation d'[³H]uridine et de [³⁵S]méthionine (pour suivre l'activité transcriptionnelle et traductionnelle respectivement) dans des embryons de bovins aux stades une et deux cellules traités ou non avec de l' α -amanitine (inhibiteur de l'ARN polymérase II et donc de la transcription). Les auteurs ont ainsi démontré la présence d'une transcription et d'une traduction d'origine embryonnaire dès le stade une cellule.

Cette transcription à des niveaux très inférieurs à celle détectée au stade huit cellules, serait donc le fruit d'une activation dite mineure, qui pourrait survenir chez les embryons bovins, au même titre que chez la souris (Aoki *et al.*, 1997), la drosophile (Edgar et Schubiger, 1986), le xénope (Kimelman *et al.*, 1987) ou l'homme (Xue *et al.*, 2013).

L'activation mineure du génome embryonnaire pourrait donc être définie comme le stade à partir duquel l'embryon est capable de transcrire, mais pour lequel l'inhibition de cette transcription n'entraîne pas d'arrêt immédiat du développement. Le nombre de gènes transcrits est moindre qu'au cours de la phase majeure et vraisemblablement le niveau de transcription de ces gènes est plus faible.

Au cours de celle-ci, les signaux de début et de fin de transcription sur le génome ne sont pas strictement respectés, on a donc transcription de régions du génome situées en périphérie des gènes qui ne seraient pas habituellement transcrites dans une cellule somatique. De plus la régulation post-transcriptionnelle des transcrits ne se fait pas toujours correctement, ainsi certains transcrits possèdent une maturation incomplète (épissage incomplet).

Sirard (2010) a tempéré tout de même cette notion d'activation mineure. En effet, bien que des preuves aient été fournies indiquant la présence de transcription avant le stade huit cellules chez le bovin, principalement par la comparaison d'embryons traités ou non avec de l' α -amanitine, elle n'a pas été clairement démontrée par des expériences utilisant des promoteurs spécifiques.

Selon l'auteur, cette apparition soudaine de « nouveaux » ARN avant le stade huit cellules pourrait être le fruit de la re-polyadénylation des ARN maternels jusque-là stockés sous forme de mRNP (donc avec une queue poly(A) courte). Lorsque la queue poly(A) est courte, elle nécessite des protocoles d'extraction qui n'utilisent pas de colonnes ou de filtres poly(T). Si les amorces utilisées pour la transcription inverse incluent une séquence poly(T), cela exclura une partie non définie de l'ARN stocké, et enfin, si une polymérase T7 est utilisée pour l'amplification de l'ARN dans les expériences de micro-puces, la présence d'une queue poly(A) fera probablement apparaître des ARN jusque-là cachés et qui seront donc faussement considérés comme nouvellement transcrits.

Pour remédier à cette problématique, Kanka *et al.* (2009) ont recherché des gènes transcrits *de novo* par l'embryon avant le stade huit cellules, chez les bovins, avec des amorces spécifiques lors de la transcription inverse afin d'éviter le biais d'amorçage spécifique de polyadénylation. Les auteurs ont trouvé un gène, impliqué dans l'épissage, possiblement transcrit durant cette activation mineure : SFRS3 (*Splicing Factor arginine/Serine-Rich 3*). L'abondance de l'ARNm de ce gène augmentait aux stades deux et quatre cellules ainsi qu'au stade tardif huit cellules. Ce gène était par ailleurs sensible au traitement par l' α -amanitine au stade quatre cellules et au stade tardif huit cellules.

Par ailleurs, Mondou *et al.* (2012) ont montré que le gène SFRS3 était également sensible à l' α -amanitine dans l'embryon à deux cellules. Ces résultats corroborent donc l'hypothèse de la présence d'une activité transcriptionnelle mineure dans l'embryon bovin précoce.

D'après Sirard (2010), et malgré les résultats de Kanka *et al.* (2009), la question d'une période de transcription mineure avant le stade huit cellules resterait tout de même à démontrer clairement dans les embryons bovins. Selon lui, cela pourrait être réalisé en utilisant des promoteurs spécifiques pour les gènes de profil embryonnaire, fusionnés avec des séquences rapportrices, et injectés rapidement après la fécondation.

2. Le contrôle de l'activation du génome embryonnaire

a. L'utilisation des protéines et des ARNm maternels stockés par l'ovocyte

Avant d'activer sa propre machinerie transcriptionnelle, l'embryon utilise les ressources maternelles pour se préparer à cette étape cruciale du développement.

Il peut notamment utiliser les protéines stockées lors de la maturation de l'ovocyte. Ces dernières sont nécessaires pour une EGA correcte.

La protéine kinase dépendant de l'AMPc (PKA) est un exemple illustrant cette nécessité des protéines maternelles pour l'EGA.

Poueymirou et Schultz (1989) ont démontré que l'inhibition de la PKA dans les zygotes de souris empêchait l'apparition de certaines protéines synthétisées spécifiquement au stade deux cellules, moment de l'EGA chez cette espèce.

Deux hypothèses pouvaient alors être formulées : soit la PKA joue un rôle dans l'activation de la traduction des ARNm maternels de ces protéines, soit elle joue directement un rôle dans l'activation de la transcription.

Or, le retrait des inhibiteurs de la PKA suivie de l'inhibition de la transcription par de l' α -amanitine, empêchait l'apparition de ces mêmes protéines lors de la MET.

Son action passe donc nécessairement par la transcription. En effet, si la PKA favorisait la traduction d'ARNm maternels stockés, la synthèse des protéines ne seraient pas inhibée par le blocage de la transcription.

Cette hypothèse a été corroborée par l'étude de Manejwala *et al.* (1991) dans laquelle les auteurs ont démontré que l'expression de l'ARNm codant pour la protéine maternelle Hsp70 au stade deux cellules n'est pas affectée par l'inhibition de la synthèse protéique dans l'embryon au stade une cellule mais est inhibée par l'inactivation de PKA.

La PKA a donc un effet au niveau de l'activation de la transcription et participe entre autres à la production d'autres protéines dans l'embryon au moment de son EGA.

Au même titre que les protéines, les ARNm maternels sont mobilisables par l'embryon au début de son développement via la polyadénylation cytoplasmique (voir I.1.d).

Les ARNm maternels sont également indispensables à l'EGA, comme le montre par exemple, le recrutement de l'ARNm de la cycline A2 pendant le développement embryonnaire précoce.

En effet, Hara *et al.* (2005) ont tout d'abord montré que cette cycline sous sa forme protéique est absente chez l'embryon au stade une cellule et apparaît seulement au cours du premier cycle cellulaire.

Son accumulation étant inhibée par la 3'-désoxyadénosine, qui inhibe la polyadénylation de l'ARNm, il semblerait donc qu'elle soit le produit de la traduction d'ARNm maternels.

La cycline A2 se lie avec Cdk2 (kinase 2 dépendante de la cycline) pour former le complexe CycA2-Cdk2, capable d'effectuer des phosphorylations.

Forjas de Borja *et al.* (2001) ont montré qu'un des facteurs phosphorylés par le complexe CycA2-Cdk2 est la protéine de spécificité 1 (SP1), un facteur de transcription, et que cette phosphorylation mène à l'augmentation de l'efficacité de cette dernière dans le processus de transcription. En effet, les cellules surexprimant la cycline A ont vu la capacité de liaison à l'ADN de SP1 augmenter.

Ainsi, la transcription des gènes dont les promoteurs sont stimulés par SP1 a été accrue.

Le complexe CycA2-Cdk2 peut également agir sur la protéine du rétinoblastome (pRB). Celle-ci lie les facteurs de transcription E2F réduisant ainsi leur activité transcriptionnelle. Or, E2F active l'expression de gènes pour permettre à la cellule d'entrer en phase S lorsqu'il est actif.

La phosphorylation de pRB par les Cdk (kinases dépendantes de cyclines) briserait le lien pRB-E2F et redonnerait sa pleine activité à E2F, activant ainsi le cycle cellulaire (Harbour et Dean, 2000).

En d'autres termes, la traduction de l'ARNm maternel de la cycline A2 lors du développement embryonnaire précoce engendre des réactions de type phosphorylations permettant *in fine* d'activer des facteurs de transcription ainsi que le cycle cellulaire. Cet ARNm maternel est donc nécessaire à la poursuite du développement de l'embryon et à l'activation de son génome.

Les protéines et les ARNm maternels sont donc utilisés par l'embryon et participent au bon déroulement de l'EGA, soit via leur action directe de régulation positive de la transcription, soit indirectement par la production ou l'activation d'acteurs nécessaires à l'EGA.

b. Les modifications de la chromatine

Pour activer le génome embryonnaire jusque-là inactif, une reprogrammation nucléaire est nécessaire.

Celle-ci dure environ trois jours chez les bovins, pendant lesquels un silence transcriptionnel est en place (Schier, 2007 ; Sirard, 2010).

Selon Graf *et al.* (2014b), les produits accumulés par l'ovocyte jouent un rôle prépondérant dans ces changements nucléaires en altérant la structure de la chromatine (modifications épigénétiques de l'ADN et modifications des histones par exemple).

Selon sa conformation, la chromatine permet aux facteurs de régulation d'atteindre ou non leur(s) cible(s) génomique(s). Un changement dans sa structure se traduit donc par la modification de l'expression des gènes.

De plus, comme mentionné dans la partie III.1, et par Dean *et al.* (2003) et Sirard (2010), toute modification des histones, qu'elle soit de type acétylation, méthylation, phosphorylation ou ubiquitination, altère la structure de la chromatine et donc change son statut transcriptionnel (actif ou inactif).

➤ La déméthylation de l'ADN

Une des modifications épigénétiques subies par l'ADN au cours de la MET est la déméthylation. La méthylation est généralement associée à la répression de la transcription.

D'après Dean *et al.* (2003), l'ADN des gamètes des mammifères est fortement méthylé sur les résidus cytosine au cours de la gamétogénèse et cette méthylation est maintenue jusqu'à la fécondation.

Suite à celle-ci, une vaste déméthylation des génomes se produit, débutant par une déméthylation active, très rapide, du pronucleus mâle, suivit d'une déméthylation progressive, passive, du génome maternel, au cours des premières divisions embryonnaires (Figure 9).

Selon les auteurs la déméthylation passive du génome maternel, c'est-à-dire dépendant de la réplication de l'ADN, serait la conséquence de l'exclusion de l'ADN méthyltransférase 1 (DNMT1) du noyau.

Cette enzyme rétablit normalement la méthylation de l'ADN pendant ou peu de temps après sa réplication. Son absence du noyau entraîne donc une baisse graduelle de la méthylation.

Chez les embryons bovin et murin, la déméthylation est progressive.

Celle-ci atteint son maximum au stade huit cellules chez les bovins (Dean *et al.*, 2003) (Figure 9).

Au vu du rôle de répresseur de la transcription de la méthylation, le fait que son minimum soit atteint au moment de l'EGA pourrait être un signe de son implication dans celle-ci.

On assiste par la suite à une méthylation *de novo* du génome de l'embryon dans sa progression vers le stade blastocyste.

Santos *et al.* (2002) et Dean *et al.* (2003) ont suggéré que celle-ci puisse être en fait dirigée.

En effet, certains gènes et séquences répétées se verraient ainsi re-méthylés différemment, attribuant plus ou moins de méthylation aux blastomères et conduisant ainsi à des différences entre les lignées cellulaires.

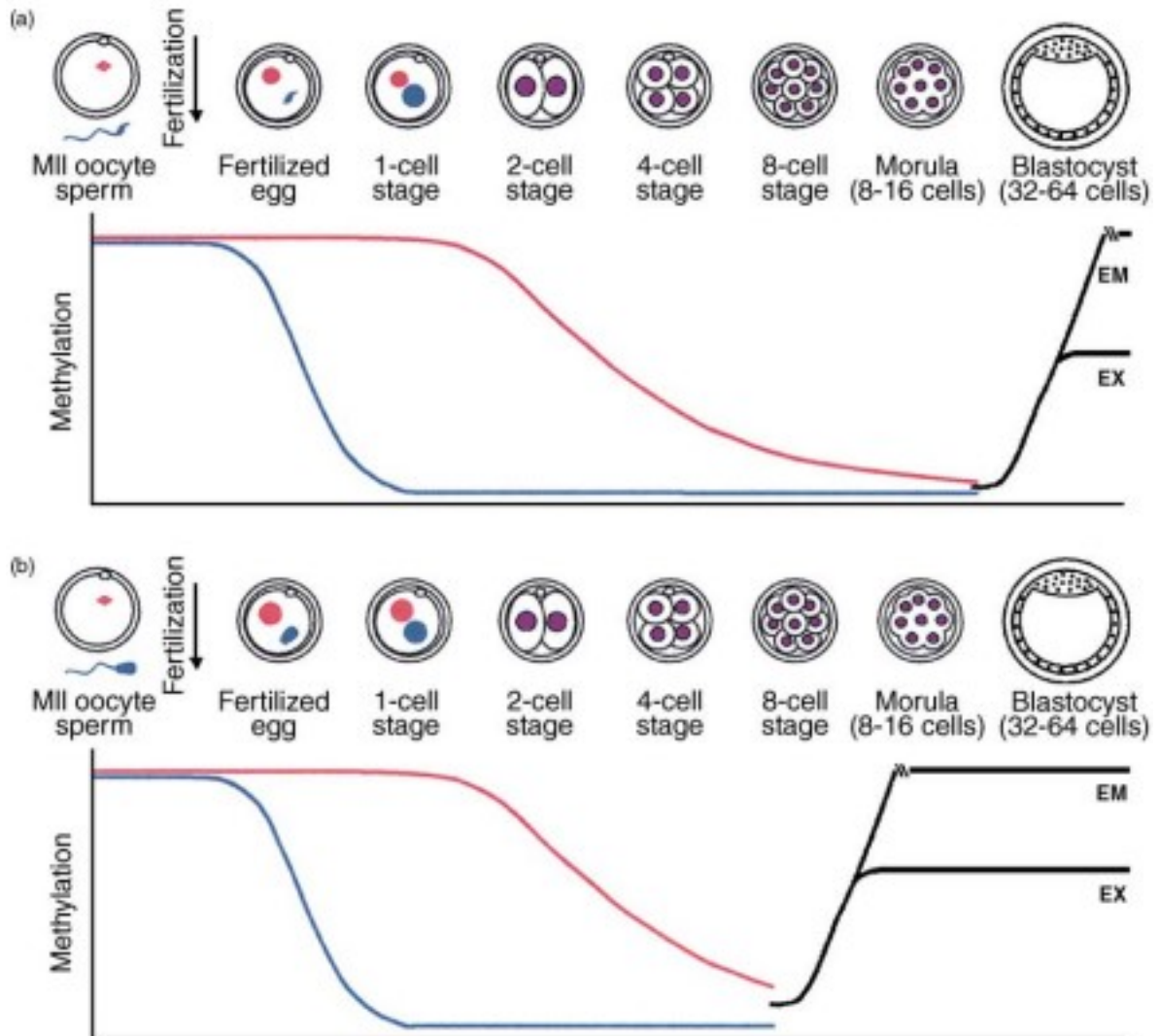
En effet, cela expliquerait que la méthylation *de novo* soit de niveau inégal entre les tissus embryonnaires (EM) et extra-embryonnaires (EX), de sorte que la MCI (d'où proviennent tous les tissus adultes) est fortement méthylée et que le trophoctoderme (donnant naissance à la plupart des tissus extra-embryonnaires comprenant le placenta) reste peu méthylé (Figure 9).

Cette re-méthylation est absolument nécessaire au développement correct de l'embryon.

Li *et al.* (1992) ont montré que la réduction de la méthylation de l'ADN, causée par une mutation du gène *DNMT1* murin, résultait en un phénotype létal récessif. Les embryons homozygotes retardaient leur développement et ne survivaient pas après la mi-gestation.

Cette étude a donc illustré l'importance de ce phénomène dans le développement précoce.

Figure 10 : Profil de méthylation du génome paternel (courbe bleue) et du génome maternel (courbe rose) au cours du développement précoce embryonnaire chez la souris (a) et chez le bovin (b). Les courbes noires représentent la méthylation *de novo* dans les lignées cellulaires embryonnaires (EM) et extra-embryonnaires (EX) (Dean *et al.*, 2003).



- La répression de la transcription par la méthylation de l'ADN et son lien avec la modification des histones

D'après Wade (2001), les protéines contenant des domaines de liaison aux cytosines méthylées (MBD) fournissent le lien critique entre la méthylation et l'établissement de cette répression transcriptionnelle. Quatre types de protéines seraient impliqués dans celle-ci : MeCP2, MBD1, MBD2 et MBD3.

Dans les cellules somatiques, une hypothèse expliquant ce mécanisme, émise par Bird et Wolffe (1999), implique le masquage du promoteur par l'encombrement causé par la liaison

des MBD à celui-ci, ou encore le recrutement des HDAC (histone désacétylases) par les MBD aux sites méthylés afin de provoquer une désacétylation des histones d'une région cible. Le degré d'acétylation des histones est un élément important de la régulation de la transcription, puisqu'une hyperacétylation des histones résulte en une chromatine plus permissive et donc en une levée de la répression de la transcription de la région touchée.

Alors que la répression de la transcription par les MBD a été démontrée dans les cellules somatiques, la présence de tous les sous-types de protéines MBD dans les embryons murins et bovins pré- et post-MET suggèrent un rôle analogue dans ceux-ci (Ruddock-D'Cruz *et al.*, 2008).

De même, des ADN méthyltransférases (DNMT), des histones acétyltransférases (HAT) et des histones désacétylases (HDAC) impliquées respectivement dans la méthylation de l'ADN, l'acétylation, et la désacétylation des histones sont présentes au cours du développement embryonnaire bovin. La présence de ces différentes classes de protéines suggère un très fort remodelage de la chromatine chez le jeune embryon bovin au cours de la MET (Ruddock-D'Cruz *et al.*, 2008).

c. Le rôle de la réplication de l'ADN

La réplication de l'ADN semble avoir un certain rôle dans la régulation de la transcription lors de l'EGA.

Aoki *et al.* (1997) ont montré que l'inhibition du premier cycle cellulaire par ajout d'aphidicoline (qui est un inhibiteur des polymérase répliquatives de l'ADN) dans des embryons de souris au stade une cellule, n'inhibe pas l'initiation de la transcription mais la réduit tout de même de 35% (mesurée grâce à l'incorporation de BrUTP, 5-Bromouridine 5'-Triphosphate, permettant de visualiser les nouveaux ARN transcrits dans des embryons traités avec l'aphidicoline en comparaison avec des embryons contrôles, non traités).

Les auteurs ont formulé l'hypothèse que le premier cycle cellulaire pourrait donc être activateur de la transcription grâce à la désorganisation que la réplication de l'ADN crée dans la chromatine, qui rendrait celle-ci plus accessible à la machinerie transcriptionnelle.

Cependant, Aoki *et al.* (1997) ont également montré qu'un état de répression de la transcription est décelable au cours du second cycle cellulaire. En effet, l'injection d'aphidicoline dans des embryons de souris au stade deux cellules augmente de 50% l'incorporation de BrUTP dans ces mêmes embryons, en comparaison aux embryons non traités.

Là encore une hypothèse est formulée : le second cycle cellulaire pourrait déplacer des complexes de transcription et/ou donner l'opportunité à des répresseurs d'atteindre leur(s) cible(s) suite à la désorganisation de la chromatine.

L'ajout de trapoxine (tétrapeptide cyclique induisant l'acétylation des histones) permet quant à elle de lever cette répression engendrée par le second cycle cellulaire, suggérant que l'hyperacétylation des histones serait promotrice de la transcription (Aoki *et al.*, 1997).

Latham et Shultz (2001) ont également proposé une interprétation des états alternativement permissif et répressif des cycles cellulaires quant à la transcription au moment de l'EGA.

La réplication de l'ADN permettrait un remodelage complet de la chromatine.

Les gènes exprimés seraient simplement ceux pour lesquels les facteurs de transcription nécessaires sont présents et pour lesquels le promoteur est accessible. Au contraire, d'autres gènes se retrouveraient différenciellement réprimés par ces modifications. Parmi ceux-ci, certains verront leur répression levée par l'action de promoteurs forts, c'est-à-dire régulés par des *enhancers* par exemple, ou par d'autres modifications telles que l'acétylation des histones.

D'ailleurs, certains auteurs, tels que Henery *et al.* (1995), ont remarqué que des embryons de souris au stade une cellule étaient capables de transcrire des gènes habituellement dépendants d'*enhancers*, sans ces derniers, tandis que leur présence était absolument nécessaire à des embryons de souris au stade deux cellules pour pouvoir transcrire ces mêmes gènes. Cela suggère que les *enhancers* soulagent une répression dans les embryons à deux cellules qui est absente dans les embryons à une cellule chez la souris.

Tous ces mécanismes agissant conjointement permettraient à l'embryon d'effectuer correctement son EGA en activant ou réprimant différenciellement les gènes spécifiques à son bon déroulement (Latham et Shultz, 2001).

3. La dégradation des transcrits maternels

Comme dit précédemment, l'activation du génome embryonnaire implique, en parallèle, une diminution d'abondance des transcrits d'origine maternelle afin de laisser la place aux transcrits d'origine embryonnaire.

Celle-ci peut être le résultat d'une dégradation ou d'une utilisation du matériel maternel.

Nous avons détaillé l'utilisation par l'embryon des protéines et des ARNm maternels dans la partie IV.2.a.

Cette partie-ci se focalisera sur la dégradation des transcrits d'origine maternelle, portée à la fois par des éléments provenant de la machinerie maternelle et par des éléments provenant de la machinerie embryonnaire, qui se mettent en place plus ou moins simultanément au cours de l'EGA (Barckmann et Simonelig, 2013).

a. La voie de l'interférence

Les micro-ARN (ARNmi) sont des régulateurs traductionnels qui en s'appariant à l'ARN messenger d'un gène cible sont capables d'inhiber son expression en réprimant sa traduction ou en le déadénylant (pour plus de détails se référer au III.3).

Ils font partie de la voie de l'interférence, qui est soupçonnée de jouer un rôle primordial dans la destruction du transcriptome maternel au cours de l'EGA.

Sirard (2010) a fait remarquer qu'il y a deux origines possibles pour ces ARNmi : comme bien d'autres ARN, ils sont stockés dans l'ovocyte et peuvent donc être d'origine maternelle, mais ils sont également les témoins d'une certaine capacité transcriptionnelle de l'embryon avant l'EGA, et peuvent donc être également d'origine embryonnaire.

Ceci pourrait par ailleurs expliquer l'activation mineure observée chez de nombreuses espèces dont les bovins par Memili et First (1999).

➤ Chez le xénope, la souris

Tout d'abord chez le xénope, Lund *et al.* (2009) ont découvert un ARNmi, le miR-427, produit en masse juste avant l'EGA dans cette espèce. Cet ARNmi agirait sur les 60 premiers nucléotides du 3'UTR des gènes codant pour les cyclines A1 et B2 (qui sont impliquées dans la régulation des différentes phases du cycle cellulaire, incluant le contournement de la phase G1) entraînant leur déadénylation.

Or, comme évoqué dans la partie I.3., une des hypothèses du silence transcriptionnel lors de la MET consiste à dire que les cycles cellulaires rapides, sans phases G, entraîneraient une inhibition de la transcription puisque c'est au cours de ces phases là qu'elle se déroule.

La neutralisation des cyclines A1 et B2 entraînerait donc un retour à un cycle cellulaire complet, incluant la phase G1, et donc, à un retour de la transcription, embryonnaire dans notre cas.

Le lien entre la voie de l'interférence, la dégradation du transcriptome maternel et le développement embryonnaire normal est ainsi clairement établi chez le xénope.

L'action des voies de l'interférence dans la dégradation des transcrits maternels a également été étudiée chez les mammifères.

Bernstein *et al.* (2003) ainsi que Murchison *et al.* (2007) ont montré que l'inactivation de la protéine Dicer dans des ovocytes au stade GV et des embryons de souris entraîne un phénotype létal.

En effet, les ovocytes se retrouvent bloqués à la première division méiotique (car Dicer jouerait un rôle dans l'intégrité du fuseau mitotique) et les embryons, eux, stoppent leur développement avant la gastrulation (car Dicer jouerait un rôle dans la différenciation cellulaire).

Cependant, Ma *et al.* (2010) ont montré que si les ovocytes de souris produisent des ARNmi en abondance, leurs ARNm cibles sont mal réprimés.

Les auteurs font donc remarquer qu'on ne peut pas attribuer le développement anormal d'un embryon de souris à l'inhibition des ARNmi en particulier car la protéine Dicer est commune aux deux voies de l'interférence. Les auteurs émettent donc l'hypothèse que la mortalité embryonnaire puisse être le résultat du dérèglement de la biogénèse des ARNs.

D'après Murchison *et al.* (2007), les transcrits régulés à la hausse dans des ovocytes dans lesquels la protéine Dicer a été mutée, sont fortement corrélés avec les ARNm qui sont dégradés pendant la maturation des ovocytes. Cela suggère donc que les petits ARN peuvent agir, directement ou indirectement, pour accélérer le renouvellement des transcrits au cours de ce processus de développement.

De plus, les ovocytes mutés pour Dicer contiennent des concentrations anormalement élevées de transcrits de transposon. La protéine Dicer pourrait donc réguler un nombre considérable de transcrits d'origine ovocytaire via des séquences de transposons intégrées dans leurs UTR.

Par ailleurs, les ARNm dont la proportion est en hausse dans les embryons mutants pour Dicer sont liés à plusieurs processus biologiques, parmi lesquels « les processus liés aux microtubules ». Ceci est particulièrement intéressant compte tenu de l'effet délétère que la perte de Dicer engendre sur le fuseau méiotique.

➤ Chez les bovins

Chez les bovins, le rôle de la voie de l'interférence dans la dégradation des transcrits maternels est moins bien étudiée.

Néanmoins, plusieurs auteurs ont mis en évidence une augmentation du taux d'expression d'ARNmi à des stades spécifiques du développement chez des ovocytes et des embryons bovins.

Tesfaye *et al.* (2009) ont montré une augmentation de l'expression de miR-125a, de miR-127 et de miR-145 dans des embryons bovins aux stades quatre et huit cellules.

Certains gènes cibles de miR-125a ont été identifiés, comme *TRIM71* (*Tripartite motif containing 71*), *EIF2C2* (*Protein argonaute-2*), *LOC407138* (une aminopeptidase), *HTR6* (*5-Hydroxytryptamine receptor 6*), *KCNH6* (*Potassium voltage-gated channel subfamily H member 6*), *MAD1L1* (*Mitotic arrest deficient 1 like 1*) et *ADAMTSL2* (*A disintegrin and metalloproteinase with thrombospondin motifs like 2*).

Les protéines codées par ces gènes sont connues pour être exprimées dans le processus d'ovogenèse des mammifères.

Puis, Tripurani *et al.* (2010) ont montré un taux d'expression stable de miR-10b et de miR-424 du stade GV jusqu'au stade seize cellules, puis une diminution aux stades morula et blastocyste, suggérant leur origine maternelle.

Les auteurs ont aussi observé la présence d'un micro-ARN ne correspondant à aucun ARNmi connu enregistré dans la base de données (mirbase.org), suggérant qu'il pourrait être spécifique aux embryons bovins.

Son abondance est relativement faible dans les ovocytes mais augmente au stade morula pour atteindre un pic au stade blastocyste.

L'étude de ses gènes cibles grâce à TargetScan a révélé un rôle dans les voies de signalisation et dans la différenciation cellulaire.

L'année suivante, Tripurani *et al.* (2011) ont analysé l'expression du miR-196a au cours du développement embryonnaire précoce. Cet ARNmi est exprimé à la fois dans les ovocytes et les embryons et a tendance à augmenter de manière constante du stade deux cellules au stade huit cellules.

De plus, les auteurs ont mis en évidence son rôle, chez la souris, de régulateur négatif du gène *NOBOX* (*Newborn ovary homeobox protein*) qui code pour un facteur de transcription essentiel à la folliculogénèse et qui semble également nécessaire pour le développement embryonnaire précoce bovin.

Les micro-ARN miR-21 et miR-130a ont été ciblé dans l'étude de Mondou *et al.* (2012).

Les auteurs ont montré une augmentation des taux d'expression de des formes matures de miR-21 et miR-130a, ainsi que la forme immature de miR130a, entre le stade une cellule et le stade huit cellules. Le précurseur de miR-21 augmente entre le stade MII de l'ovocyte et le stade deux cellules de l'embryon.

Une diminution de leur expression est observée si les embryons sont traités avec de l' α -amanitine au stade deux cellules, ce qui suggère qu'il existe également des ARNmi transcrits *de novo* par l'embryon.

Au vu du profil d'expression du précurseur de miR-21, les auteurs ont par ailleurs émis l'hypothèse que son augmentation pourrait être associée à l'activation mineure du génome embryonnaire

bovin détectée par Memili et First (1999), car celui-ci apparaît en forte quantité au même moment que celle-ci.

Mondou *et al.* (2012) ont par la suite recherché les cibles de ces ARNm dans le génome maternel. Les résultats suggèrent une implication dans des voies régulatrices du développement embryonnaire précoce.

Enfin, le miR-212 a été identifié chez le bovin par Tripurani *et al.* (2013). Son expression augmente entre le stade quatre cellules et le stade huit cellules. Il serait possiblement impliqué dans la régulation des transcrits maternels.

En effet, il serait un régulateur négatif du facteur de la lignée germinale alpha (FIGLA). FIGLA étant décrit comme un facteur de transcription spécifique à l'ovocyte, essentiel à l'expression de nombreux gènes nécessaires à la folliculogénèse, à la fécondation et à la survie embryonnaire précoce, ce micro-ARN pourrait donc jouer un rôle dans la dégradation des transcrits maternels.

b. Les protéines de liaison à l'ARN

Les protéines de liaison à l'ARN (RBP) ont également été incriminées dans la régulation de la dégradation des transcrits maternels lors de l'EGA chez certaines espèces.

Chez la drosophile, la protéine Smaug est un régulateur post-transcriptionnel, multifonctionnel qui agit comme un facteur de spécificité en liant les transcrits maternels qui contiennent des SRE (*Smaug Response Element*) (Tadros *et al.*, 2007). Son action passe par le recrutement du complexe CCR4/POP2/NOT-deadénylase entraînant l'élimination de la queue poly(A) des ARNm maternels et donc leur dégradation. Cette protéine serait requise pour la dégradation du deux tiers des ARNm maternels.

Deuxièmement, un enrichissement significatif des sites potentiels de liaison à la protéine Pumilio, un régulateur post-transcriptionnel impliqué dans la répression de la traduction et la déstabilisation des ARNm, a été trouvé dans les ARNm maternels éliminés au cours de la MET (De Renzis *et al.*, 2007).

Gerber *et al.* (2006) ont identifié deux transcrits en particulier (Bicoid et Hunchback) comme des cibles directes de dégradation de Pumilio dans des embryons de drosophile.

D'autres éléments fortement retrouvés dans les ARNm maternels sont les éléments riches en AU (ARE).

Selon Walser et Lipshitz (2011), par le recrutement de protéines de liaison aux éléments riches en AU (ARE-BP), les ARE peuvent stimuler le raccourcissement de la queue poly(A) et recruter l'exosome, une exoribonucléase.

À ce jour, le rôle potentiel des ARE-BP dans la déstabilisation des transcrits maternels chez la drosophile n'a pas été testé expérimentalement.

Cependant, chez le xénope, la voie médiée par les séquences ARE semble liée à EDEN-BP, qui déclenche la déadénylation des ARNm stockés.

En effet, chez cette espèce, la déadénylation des ARNm est entraînée par la séquence EDEN (*Embryonic Deadenylation Element*), qui est composée de séquences riches en GU (Guanosine et Uracile) localisées dans le 3'UTR et liées par la protéine de liaison à l'ARN EDEN-BP (*EDEN Binding Protein*). La déadénylase interagissant avec EDEN-BP n'a pas été identifiée, bien qu'un candidat potentiel soit la protéine PARN car celle-ci se lie à la protéine CUG-BP (*CUG triplet repeat RNA-*

binding protein), analogue d'EDEN-BP dans les cellules de mammifères (Barckmann et Simonelig, 2013).

D'après Walser et Lipshitz (2011), les homologues de plusieurs RBP impliqués dans l'élimination des transcrits chez la drosophile et le xénope sont également exprimés dans les embryons précoces de souris. Il s'agit notamment des homologues de Smaug (Samd4), EDEN-BP (CUG-BP1), ARE-BP (ELAVL1, ELAVL2 et ZFP36L2).

De plus, l'implication de la protéine ZFP36L2 (*Zinc finger protein 36 like 2*) dans la dégradation de l'ARNm a été démontrée dans les embryons de souris.

En effet, Lai *et al.* (2000) ont montré que la protéine ZFP36L2 joue un rôle physiologique important dans la déstabilisation de deux ARNm maternels contenant des ARE : celui codant pour le TNF α (aussi appelé cachexine, cytokine inflammatoire sécrétée par les macrophages) et celui codant pour le GM-CSF (facteur stimulant la production de granulocytes et de macrophages). Son effet semble de plus impliquer la déadénylation de ces ARNm.

Les souris *knock-out* (KO) pour ZFP36L2 sont caractérisées par un état inflammatoire chronique qui semble être en grande partie dû à un excès de TNF α circulant.

Cet ensemble de connaissances suggère l'importance des protéines de liaison à l'ARN dans le développement embryonnaire précoce, via leur action de dégradation du transcriptome maternel.

c. L'autophagie

De même que la voie de l'interférence ou les protéines de liaison à l'ARN, le processus d'autophagie semble jouer un rôle dans la dégradation des transcrits maternels lors de l'EGA chez la souris et le porc.

L'autophagie est un processus physiologique, intracellulaire, de dégradation d'éléments cytoplasmiques tels que des organites ou des protéines endommagés, par emprisonnement dans un autophagosome et fusion avec un lysosome.

➤ Chez la souris

Tsukamoto *et al.* (2008) ont étudié le rôle du phénomène d'autophagie dans la dégradation des transcrits maternels chez la souris.

Les auteurs ont montré que l'autophagie était régulée à la hausse, un court instant, juste après la fécondation, et qu'elle pourrait être induite par les oscillations calciques intracellulaires observées à ce moment-là du développement.

De plus, les souris KO pour le facteur Atg5, facteur nécessaire à la formation de l'autophagosome, subissent un arrêt de leur développement entre le stade quatre et huit cellules.

Un mécanisme possible, sous-jacent aux défauts de développement des embryons déficients en facteur Atg5, et donc en autophagie, pourrait être une altération du recyclage des protéines.

En effet, jusqu'au stade deux cellules, aucune différence de niveau de synthèse protéique n'a été décelée entre les embryons témoins et les embryons mutés.

Cependant, aux stades quatre et huit cellules, l'incorporation de [³⁵S]méthionine dans les protéines des embryons mutés a été réduite d'environ 70% par rapport à celle des embryons de type sauvage.

L'autophagie serait donc liée à la synthèse protéique de telle sorte qu'un faible niveau d'autophagie entraîne une diminution de la synthèse protéique. Étant donné l'importance de celle-

ci au cours du développement embryonnaire, ce déficit en protéines pourrait être un facteur impliqué dans l'arrêt du développement observé.

➤ Chez le porc

Xu et al. (2012) font état de l'autophagie comme un mécanisme permettant la dégradation des transcrits maternels au cours de l'EGA, chez le porc.

Dans leur étude, le phénomène d'autophagie a été étudié grâce au suivi des transcrits de gènes impliqués dans ce processus : *ATG5* (*Autophagy related 5*), *BECN1* (*Beclin 1*) et *LC3* (*Microtubule associated protein 1 light chain 3 alpha*), ainsi que son effet sur les transcrits de quatre gènes de profil maternel : *GDF9* (*Growth differentiation factor 9*), *BMP15* (*Bone morphogenetic protein 15*), *C-Mos* (*Mos proto-oncogene, serine/threonine kinase*) et *Cyclin B*.

Des embryons de porc traités avec de la 3-méthyladénine (3-MA, un inhibiteur de l'autophagie) ont montré une augmentation du taux de transcrits des gènes *GDF9*, *C-Mos* et *Cyclin B* aux stades deux et quatre cellules par rapport aux embryons non traités.

Au contraire, des embryons de porc traités avec de la rapamycine (un inducteur de l'autophagie) possèdent des taux plus faibles de transcrits des gènes *BMP15* et *Cyclin B* comparés aux embryons non traités.

Le traitement avec la 3-MA induit également une diminution du taux de développement des embryons jusqu'au stade blastocyste, alors que la rapamycine n'a aucun effet sur celui-ci.

Cependant, les traitements ont entraîné tous les deux une augmentation du taux d'apoptose et donc une diminution du nombre total de cellules dans les blastocystes.

Les auteurs ont donc démontré que l'autophagie est un phénomène impliqué dans la dégradation des transcrits maternels dans les embryons porcins. Leurs résultats suggèrent que ce phénomène est nécessaire au bon développement embryonnaire.

Ce phénomène n'a pas encore été étudié chez les bovins dans le cadre de la dégradation des transcrits maternels.

4. Les gènes et les voies fonctionnelles activés lors de l'EGA

De manière plus ou moins concomitante à la dégradation des transcrits maternels, l'embryon enclenche sa transcription et génère ses propres transcrits.

L'étude de cette activation passe tout d'abord par la différenciation des transcrits des gènes au profil dit embryonnaire, c'est-à-dire transcrits *de novo* au cours de l'EGA, et des transcrits des gènes au profil dit maternel, c'est-à-dire transcrits par l'ovocyte avant la fécondation et dont les produits sont censés être dégradés par l'embryon au cours de l'EGA.

Les transcrits repérés peuvent ensuite être associés à des gènes dont la fonction ainsi que l'implication dans des voies fonctionnelles peuvent être étudiées.

Les connaissances dans ce domaine sont plus ou moins récentes et, surtout, assez parcellaires.

Dans cette partie, une révision de littérature a été proposée pour mettre en commun les principales connaissances dans ce domaine chez le bovin.

a. Comment différencier les transcrits maternels des transcrits *de novo* au cours de l'EGA ?

Afin d'être capable d'étudier les gènes impliqués dans l'EGA, il est nécessaire de mettre au point des stratégies permettant de mettre en évidence la transcription embryonnaire. Il existe actuellement plusieurs approches pour aborder la question.

Les premières transcriptions d'origine purement embryonnaire ont été étudiées par RT-PCR et grâce aux micro-puces à ADNc par Misirlioglu *et al.* (2006) et Vigneault *et al.* (2009) en utilisant de l' α -amanitine.

En effet, les auteurs ont comparé des embryons traités et non traités afin d'identifier les transcrits qui étaient présents lorsque la transcription embryonnaire était fonctionnelle et qui ne l'étaient plus lorsque celle-ci était inhibée.

Ainsi les transcrits qui étaient significativement surexprimés chez les embryons non traités par rapport aux embryons traités étaient considérés comme transcrits *de novo*.

Une autre manière d'étudier les transcriptions embryonnaires implique le séquençage de l'ARN et consiste en la combinaison de plusieurs stratégies différentes comme l'ont fait Graf *et al.* (2014a).

La première consistait en une comparaison statistique entre différents stades pour identification des gènes exprimés dans les stades post-fécondation alors qu'ils ne l'étaient pas dans les stades ovocytaires.

Les gènes ont été considérés comme exprimés pour la première fois dans les embryons lorsque moins de cinq lectures de leurs transcrits avaient été trouvées aux deux stades ovocytaires (GV et MII) et qu'au moins vingt lectures à l'un des stades après la fécondation avaient été détectées.

De plus, ce nombre de lectures dans un des stades embryonnaires devait être significativement plus important que dans le stade précédent.

La deuxième stratégie consistait à détecter, grâce au suivi des SNP (*Single-Nucleotide Polymorphism* ou polymorphisme d'un seul nucléotide) spécifiques à la race du taureau, les gènes hérités du père, nouvellement transcrits par l'embryon.

En effet, si les transcrits détectés appartiennent à des gènes paternels, ils sont forcément le fruit d'une transcription embryonnaire vu que le spermatozoïde ne stocke pas de transcrits préalablement à la fécondation comme le fait l'ovocyte.

Pour ce faire, les auteurs ont utilisé deux races différentes et relativement éloignées (une de type *Bos taurus* et l'autre de type *Bos indicus*) pour produire les embryons.

Un grand nombre de SNP informatifs pour l'identification des transcrits provenant du génome paternel étaient attendus dans les embryons hybrides.

Enfin, la troisième stratégie impliquait la détection des séquences introniques afin de repérer les produits d'une transcription récente et donc d'une transcription *de novo*, par l'embryon.

Ces différentes stratégies possèdent néanmoins des limites.

En effet, un bon nombre de gènes ont un profil dit « mixte », c'est-à-dire materno-embryonnaire.

Il s'agit de gènes dont les transcrits sont stockés par l'ovocyte au cours de sa maturation (et dont l'abondance décroît au début du développement embryonnaire du fait de leur dégradation ou de

leur utilisation) mais qui sont ensuite re-transcrits par l'embryon à la suite de l'activation de sa machinerie transcriptionnelle.

Il existe ainsi une grande pluralité de profils et le choix d'une stratégie unique ne permet jamais une parfaite exhaustivité quant aux gènes mis en évidence.

Cela se voit notamment avec la troisième stratégie de Graf *et al.* (2014a) qui permet de mettre en évidence des milliers de gènes alors que leurs deux premières stratégies permettaient la mise en évidence de centaines de gènes seulement.

En combinant plusieurs stratégies de détection des transcrits embryonnaires, il est donc possible de pallier en partie à cette difficulté.

b. Etude de l'abondance des transcrits produits *de novo* par l'embryon et identification des gènes associés

Misirlioglu *et al.* (2006) ont comparé un groupe d'embryons traités à l' α -amanitine à des groupes témoins d'ovocytes (au stade MII) et d'embryons (au stade huit cellules) non traités.

La comparaison des embryons au stade huit cellules, sans aucun traitement, avec les ovocytes au stade MII, a permis de mettre en évidence 2 505 transcrits au total.

Parmi ceux-ci, 1 413 étaient significativement plus abondants dans les embryons en comparaison avec les ovocytes. Ces transcrits correspondaient à 258 gènes connus.

La comparaison des embryons au stade huit cellules non traités avec les embryons au stade huit cellules traités avec l' α -amanitine a permis la mise en évidence de 1 811 transcrits, dont 1 490 transcrits significativement plus abondants dans les embryons non traités par rapport aux embryons inhibés. 233 gènes de profil embryonnaire ont ainsi été mis en évidence.

Vigneault *et al.* (2009) ont également utilisé l' α -amanitine, cette fois-ci pour comparer l'abondance des transcrits d'embryons de bovins au stade huit cellules traités ou non traités.

Les auteurs ont ainsi trouvé 310 transcrits avec une abondance significativement plus élevée dans les embryons non traités par rapport aux traités, correspondant à 208 gènes différents connus.

Graf *et al.* (2014a), grâce à leurs trois stratégies combinées précédemment décrites, ont pu établir une cinétique transcriptionnelle aux différents stades de développement.

Pour cela, des pools d'ovocytes de Simmental allemande ont été prélevés au hasard, élevés *in vitro* et fertilisés avec le sperme d'un taureau unique de la race Brahmane (une race bovine issue de la sous-espèce *Bos indicus* : le Zébu).

Les ovocytes et embryons ont été arrêtés à différents stades de développement : GV, MII, quatre cellules, huit cellules, seize cellules et blastocyste.

Les auteurs ont détecté relativement peu de différences dans le nombre total de transcrits entre les ovocytes au stade GV et les ovocytes au stade MII.

Il en était de même entre les ovocytes au stade GV et les embryons au stade quatre cellules.

Le nombre de transcrits a été en revanche multiplié par dix entre les stades quatre et huit cellules.

Cette différence est encore plus élevée si l'on compare le stade quatre cellules aux stades supérieurs à huit cellules.

De plus, la proportion de transcrits qui présentaient une diminution de leur abondance était initialement faible (17% des transcrits voyaient leur abondance diminuer dans les embryons à quatre cellules par rapport aux ovocytes MII et 24% dans les embryons à huit cellules par rapport

aux embryons à quatre cellules). Mais celle-ci augmentait jusqu'à 55% dans les embryons à seize cellules si on les comparait aux embryons à huit cellules.

Enfin, les ovocytes GV et MII stockaient environ 13 000 transcrits de gènes, et seul un petit nombre de ceux-ci ont été re-transcrits par l'embryon après la fécondation (220 au total, soit 1,7% environ). 129 d'entre eux, soit 59%, l'ont été au stade huit cellules.

La très forte augmentation de l'abondance des transcrits à partir du stade huit cellules témoigne d'un réel enclenchement de la transcription à partir de ce stade. En parallèle, certains transcrits décroissent à partir du stade quatre cellules. Cette décroissance est à son maximum au stade seize cellules et serait donc, au moins en partie, le témoin de la dégradation des transcriptions maternelles.

De plus, les transcrits présents à chacun des stades (GV, MII, quatre cellules, huit cellules, seize cellules et blastocyste) ont été séquencés et comparés au génome de référence bovin pour identification.

Au total, 12,4 à 13,7x10³ transcrits de gènes différents ont été détectés.

Leur première approche qui consistait à repérer les gènes significativement surexprimés dans les stades embryonnaires par rapport aux stades ovocytaires leur a permis d'identifier huit gènes nouvellement transcrits au stade quatre cellules, 129 gènes au stade huit cellules, 36 gènes au stade seize cellules et 47 gènes au stade blastocyste.

En utilisant la première expression de l'allèle paternel comme marqueur (2^e stratégie), les auteurs ont trouvé 16 gènes activés au stade quatre cellules, 395 gènes au stade huit cellules, 314 gènes au stade seize cellules et 212 gènes au stade blastocyste.

Et grâce à leur troisième stratégie de détection de séquences introniques, ils ont mis en évidence 390 gènes s'activant au stade quatre cellules, 3 965 gènes au stade huit cellules, 628 gènes au stade 16 cellules et 1 865 gènes au stade blastocyste.

L'ensemble de ces données est résumé dans le Tableau 1.

Tableau 1 : Nombre de gènes identifiés comme transcrits *de novo* par l'embryon à différents stade de développement en fonction des stratégies utilisées.

Etudes	Stratégies	Nombre de gènes mis en évidence comme transcrits <i>de novo</i> par l'embryon			
		Stade 4 cellules	Stade 8 cellules	Stade 16 cellules	Stade blastocyste
Misirlioglu et al. (2006)	Comparaison ovocytes au stade MII – embryons au stade 8 cellules	/	258	/	/
	Comparaison embryons au stade 8 cellules non traités – embryons au stade 8 cellules traités avec α -amanitine	/	233	/	/
Vigneault et al. (2009)	Comparaison embryons au stade 8 cellules non traités – embryons au stade 8 cellules traités avec α -amanitine	/	208	/	/
Graf et al. (2014a)	Comparaison statistique des gènes surexprimés dans les stades embryonnaires par rapport aux stades ovocytaires	8	29	36	47
	Etude des SNP	16	395	314	212
	Détection des séquences introniques	390	3 965	628	1 865

c. La fonction des gènes transcrits *de novo*

Une fois les gènes transcrits *de novo* par l'embryon identifiés, la question de leur fonction se pose. Dans leurs études respectives, Misirlioglu *et al.* (2006), Vigneault *et al.* (2009) et Graf *et al.* (2014b) ont détaillé la fonction de quelques-uns des gènes ainsi mis en évidence.

➤ Au stade quatre cellules

Graf *et al.* (2014b), avec la combinaison de trois stratégies de détection, ont réussi à mettre en évidence 414 gènes qui s'activaient au stade quatre cellules.

Parmi ces gènes, un certain nombre d'entre eux est concerné par la maintenance de la pluripotence et l'auto-renouvellement des cellules souches (*HNRNPA2B1*, *KLF17*, *H2AFZ*, *EPCAM*, *MLL2*, *YES1*).

Le gène *HNRNPA2B1* (*Heterogeneous nuclear ribonucleoproteins A2/B1*) code pour la ribonucléoprotéine nucléaire hétérogène A2/B1. Celle-ci est connue pour interagir avec le facteur de transcription essentiel au maintien de la pluripotence : *SOX2* (*Sex determining region Y-box 2*) (Fang *et al.*, 2011).

Le gène *KLF17* (*Krüppel-like factor 17*) pourrait, d'après Van Vliet *et al.* (2006), être un activateur de la transcription. Il a aussi des fonctions liées à la pluripotence dans d'autres espèces comme l'homme ou le poisson-zèbre.

H2AFZ (*H2A histone family, member Z*) affecte la structure de la chromatine et promeut l'expression de gènes impliqués dans la pluripotence. Chez les embryons murins, Hu *et al.* (2013) ont montré qu'il était nécessaire à l'auto-renouvellement des cellules souches et à la différenciation.

L'inactivation du gène *EPCAM* (*Epithelial cell adhesion molecule*) chez la souris montre qu'il est lui aussi nécessaire à l'auto-renouvellement des cellules souches (González *et al.*, 2009).

MLL2 (*Histone-lysine N-methyltransferase*) régule quant à lui la transcription par modification des histones et est nécessaire pour le développement pré-implantatoire puisque les souris KO pour ce gène présentaient un arrêt de celui-ci (Glaser *et al.*, 2006).

Enfin, Tamm *et al.* (2011) ont montré que le gène *YES1* (*Yamaguchi sarcoma viral oncogene homolog 1*) joue un rôle dans l'auto-renouvellement des cellules souches et dans la maintenance de la pluripotence par régulation des gènes *POU5F1* (*POU domain, class 5, transcription factor 1*) et *NANOG* (*Nanog Homeobox*).

➤ Au stade huit cellules

Vigneault *et al.* (2009) ont remarqué six gènes d'intérêt qui avaient un profil d'expression un peu particulier : *TP3*, *HNRNPA2B1*, *BECN1*, et *ZNF41*.

En effet, les auteurs ont également réalisé une expérience de RT-PCR sur des embryons aux stades deux cellules, entre six et huit cellules (stade huit cellules précoce), entre huit et douze cellules (stade huit cellules tardif) et blastocyste.

Ils ont ainsi pu remarquer que ces gènes étaient exprimés de manière soudaine et transitoire juste avant l'EGA (entre le stade six cellules et le stade huit cellules précoce), suggérant leur rôle dans celle-ci.

Lorsque l'on s'intéresse à leur fonction, on observe que :

- *ZNF41* (*Zinc finger protein 41*) a un rôle putatif dans la transcription en tant que répresseur (Franzè *et al.*, 1991).
- D'après Landsberg *et al.* (2006), *HNRNPA2B1* (*Heterogeneous nuclear ribonucleoproteins A2/B1*) code pour une protéine de liaison à l'ADN impliquée dans le trafic et le traitement de l'ARNm.

Ce gène a d'ailleurs été trouvé comme activé au stade quatre cellules par la suite par Graf *et al.* (2014b) et est également impliqué dans le maintien de la pluripotence selon Fang *et al.* (2011).

Les protéines codées par ces deux gènes au profil d'expression transitoire pourraient donc être des précurseurs pour l'activation majeure qui a lieu au stade huit-seize cellules.

- *TP3* (*Putative spermatidal transition protein 3*) fait partie des protéines nucléaires de transition, connues pour remplacer les histones somatiques avant leur substitution par des protamines lors de la condensation de la chromatine dans la spermatogenèse (Meistrich *et al.*, 2003). Les auteurs n'ont pas pu expliquer sa présence et ont pointé le besoin d'études complémentaires pour mieux comprendre sa fonction chez l'embryon.

- Selon Ferraro et Cecconi (2007), le gène *BECN1* (*Beclin 1*) n'a pas de rôle dans la transcription mais est impliqué dans le processus d'autophagie.

Vigneault *et al.* (2009) ont attribué la surexpression de *BECN1* au milieu de culture qui pourrait ne pas avoir fourni un environnement optimal pour le développement embryonnaire mais n'ont pas exclu une autre explication à la présence de ses transcrits.

Au vu de l'importance du processus de l'autophagie dans la dégradation des ARNm maternels démontrée chez la souris et le porc (Tsukamoto *et al.*, 2008 ; Xu *et al.*, 2012) sa présence pourrait être le témoin d'une implication similaire chez les bovins.

Deuxièmement, les auteurs ont mis en évidence des gènes impliqués dans le maintien de la pluripotence à ce stade du développement.

En effet, les transcrits de *ZFP42* (*Zinc finger protein 42*) étaient absents des embryons bovins pré-MET, tandis que les embryons post-MET en présentaient une expression accrue.

Or son expression est étroitement régulée par les gènes *NANOG*, *SOX2* et *POU5F1*, qui sont directement impliqués dans le maintien de la pluripotence dans les cellules souches et les embryons.

En effet, *POU5F1* code pour la protéine Oct4 (*Octamer-binding transcription factor 4*) dont le niveau d'expression dans les cellules souches embryonnaires influence le devenir de celles-ci (maintien ou non de la pluripotence) (Niwa *et al.*, 2000).

Dans ces cellules, un niveau faible d'Oct4 induit une différenciation en lignée trophoblastique et un niveau élevé d'Oct4 induit une différenciation en endoderme primitif ou en mésoderme. Un niveau intermédiaire permet de maintenir la pluripotence.

Les cellules de la masse cellulaire interne doivent, pendant le développement précoce, garder leur pluripotence afin de permettre une constitution correcte de tous les types de tissus du futur embryon, le niveau d'expression d'Oct4 doit donc être finement régulé.

Chez la souris, un lien entre *ZFP42* et *POU5F1* a été démontré par Zhang *et al.* (2006). Selon ces auteurs, le gène *ZFP42* régule négativement *POU5F1* et donc le niveau d'expression d'Oct4. En contrepartie, Oct4, à de fortes concentrations, pourrait inhiber *ZFP42*. Ainsi, des cellules souches embryonnaires de souris KO pour *ZFP42* voient leur taux d'Oct4 fortement augmenter et se différencient en lignées endodermique et/ou mésodermique. *A contrario*, des cellules souches embryonnaires de souris dont l'expression de *ZFP42* est augmentée, se différencient en lignée trophoblastique (car le taux d'Oct4 chute en conséquence de cette surexpression).

Trois autres gènes ont également été identifiés par Vigneault *et al.* (2009) comme étant surexprimés dans leurs embryons au stade huit cellules non traités comparés aux embryons traités et jouant un rôle dans la pluripotence de l'embryon précoce :

- *TPT1* (*Tumor protein translationally-controlled 1*) ;
- *GABPA* (*GA-binding protein alpha chain*) ;
- *KLF4* (*Kruppel-like factor 4*).

En effet, ces gènes sont connus pour être des régulateurs de l'expression de *POU5F1* dans les cellules souches embryonnaires (Nakatake *et al.*, 2006 ; Kinoshita *et al.*, 2007 ; Koziol *et al.*, 2007).

TPT1 et *KLF4* suivent le modèle d'expression des gènes *SOX2*, *POU5F1* et *NANOG* en s'exprimant dès avant l'EGA chez les embryons bovins, témoignant de leur potentielle indépendance vis-à-vis de celle-ci.

La situation est différente pour *GABPA* qui, tout comme *ZFP42*, s'exprime, lui, au moment de l'EGA.

Le maintien de la pluripotence, grâce à la bonne régulation de l'expression de *POU5F1*, semble donc fondamental pour le développement embryonnaire.

En effet, Berstein *et al.* (2003) ont montré que l'inhibition de la voie de l'interférence entraînait un phénotype létal sur les embryons. Les auteurs ont suggéré que cette létalité serait engendrée par un lien potentiel entre la voie de l'interférence et Oct4. En effet, des embryons de souris KO pour la protéine Dicer présentaient de faibles taux d'Oct4 et donc une différenciation de leurs cellules souches en lignée trophoblastique.

Les auteurs n'ont cependant pas pu établir un lien direct entre Dicer et Oct4.

Misirlioglu *et al.* (2006) ont eux aussi pu identifier des gènes d'intérêt surexprimés dans les embryons au stade huit cellules comparés aux ovocytes au stade MII, et donc considérés comme transcrits *de novo*.

Parmi eux, on retrouvait les gènes *NFYA* (*Nuclear transcription factor Y subunit alpha*), *H2AFZ* (*H2A histone family, member Z*), *SMARCAL1* (*SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A-like protein 1*) et *DSC2* (*Desmocolin 2*).

Ils étaient respectivement 141, 99, sept et 319 fois plus transcrits dans les embryons que dans les ovocytes.

D'après Mantovani *et al.* (1992), le gène *NFYA* est un acteur de la synthèse d'ARNm.

H2AFZ est un régulateur de l'expression génique et de l'hétérochromatine silencieuse selon Hatch et Bonner (1990). Graf *et al.* (2014b) ont par la suite également trouvé *H2AFZ* parmi les gènes activés dans l'embryon (transcrits *de novo*) et celui-ci s'activait au stade quatre cellules.

Le gène *SMARCAL1* code pour une enzyme du complexe SWI/SNF jouant un rôle dans le remodelage de la chromatine ATP-dépendant en augmentant l'accessibilité des facteurs de transcription aux séquences régulatrices par désorganisation des nucléosomes (Narlikar *et al.*, 2002 ; Venters et Pugh, 2009). D'après De la Serna *et al.* (2001), ces changements conduisent à une structure de chromatine unique capable de maintenir la totipotence pendant l'embryogenèse et de conduire à une différenciation au cours du développement post-implantatoire.

Enfin, selon Ghassemifar *et al.* (2003), *DSC2* joue un rôle dans l'adhésion cellulaire.

Parmi les gènes détectés comme activés au stade huit cellules par Graf *et al.* (2014b), on peut distinguer plusieurs grands groupes fonctionnels.

Le premier groupe comprenait des gènes impliqués dans la pluripotence de l'embryon.

Par exemple, on retrouvait *SALL4* (*Sal-like 4*), un régulateur essentiel de *POU5F1* (Zhang *et al.*, 2006), ou le proto-oncogène *Myc* qui, dans les embryons murins, affine la population de cellules épiblastiques (les cellules avec un faible taux de *Myc* seraient alors éliminées) (Claveria *et al.*, 2013).

Il en était de même pour le gène *TBX3* (*T-box 3*) pour lequel Lu *et al.* (2011) ont montré sa nécessité au maintien de l'auto-renouvellement des cellules souches ainsi que leur différenciation par la suite en endoderme extra-embryonnaire chez la souris.

La protéine homéobox Nanog, codée par le gène *NANOG*, forme, d'après Tan *et al.* (2013), avec Oct4 et Sall4 une triade de protéines partageant une relation fonctionnelle étroite pendant le développement embryonnaire préimplantatoire.

En effet, ces protéines interagissent entre elles et co-occupent les promoteurs de nombreux gènes cibles.

Elles réguleraient à la baisse des gènes impliqués dans l'établissement et le maintien de la structure de la chromatine et favoriseraient l'expression de gènes impliqués dans la différenciation, la réponse au stress et les processus de développement. Bon nombre de leurs cibles sont également impliquées dans le cycle cellulaire.

Des souris KO pour l'une ou l'autre d'entre elles ne mènent pas à terme leur développement embryonnaire ce qui suggère leur rôle fondamental à cette période.

De plus, certains ARNm feraient partie intégrante du réseau de régulation de la triade Oct4-Sall4-Nanog. Celles-ci seraient en fait des régulateurs positifs de l'expression des ARNm.

Graf *et al.* (2014b) ont également mis en évidence le gène *KLF4* (*Krüppel-like factor 4*), qui prévient la différenciation cellulaire et régule l'expression du gène *NANOG* chez la souris (Zhang *et al.*, 2010). Dans cette espèce, il n'est pas indispensable au développement embryonnaire mais, selon Katz *et al.* (2002), les souris KO pour ce gène meurent peu après la naissance car possèdent une peau défectueuse qui génère une perte de fluides létale.

Vigneault *et al.* (2009) avaient d'ailleurs également identifié ce gène comme étant surexprimé dans des embryons bovins au stade huit cellules par rapport à des embryons dont la transcription avait été inhibée.

Un deuxième groupe de gènes rassemble des gènes impliqués dans les mécanismes de transcription et de traduction.

Graf *et al.* (2014b) ont en effet trouvé des transcrits de gènes codant pour les sous-unités du facteur 3 du complexe d'initiation de la traduction des eucaryotes EIF3.

Hinnebusch (2006) a expliqué qu'EIF3 est composé de 13 sous-unités organisées en modules : A (EIF3A, EIF3B, EIF3G, EIF3I), B (EIF3F, EIF3H, EIF3M), et C (EIF3C, EIF3D, EIF3E, EIF3K, EIF3L). C'est un facteur d'initiation, il est donc nécessaire au bon déroulement de la traduction des ARNm en protéines en recrutant certaines sous-unités ribosomales ou d'autres facteurs d'initiation par exemple. Il est également impliqué dans le désassemblage et le recyclage du complexe ribosomique lors de la terminaison de la traduction.

On retrouvait également les gènes de la famille DEAD- et DEAH-box (*DDX5*, *DHX8*, *DHX15*, *DHX38*). Ce sont, selon Abdelhaleem *et al.* (2003), des hélicases d'ARN impliquées dans le métabolisme de l'ARN tel que l'initiation de la transcription, la biogénèse des ribosomes et l'épissage des ARN pré-messagers. Elles auraient également un rôle dans la différenciation.

Un autre groupe de gènes retrouvés comme s'activant à ce stade comprenait des gènes plutôt tournés vers le métabolisme cellulaire.

Par exemple, la famille de facteurs associés à *DDB1* (*DNA damage-binding protein 1*) et *CUL4* (*Cullin-4*) (DCAFs).

L'ubiquitine ligase *CUL4-DDB1* régule la prolifération cellulaire, la réparation de l'ADN et le maintien de l'intégrité du génome par ubiquitination de régulateurs clés (Lee et Zhou, 2007).

De même étaient présents les gènes encodant des sous-unités de l'ATP synthase mitochondriale (*ATP5A1*, *ATP5B*, *ATP5F1*, *ATP5J2*).

Selon Jonckheere *et al.* (2012), l'ATP-synthase produit de l'ATP à partir d'ADP en présence d'un gradient de protons à travers la membrane mitochondriale, généré par le transport d'électrons dans la chaîne respiratoire.

L'ATP a de multiples fonctions, mais est entre autre impliqué dans la biosynthèse de nucléotides et le métabolisme (c'est notamment la forme d'énergie de la cellule).

Le dernier groupe de gènes activés au stade huit cellules était fonctionnellement relié à la méthylation de l'ADN et à la structure de la chromatine.

Parmi eux, on retrouvait *DNMT3B* (*DNA methyltransferase 3 beta*) et *MTA2* (*Metastasis associated 1 family member 2*).

Les souris KO pour le gène *DNMT3B* présentent des défauts de croissance et des anomalies du tube neural rostral (Okano *et al.*, 1999).

Ma *et al.* (2010) ont montré que *MTA2* est un composant du complexe de remodelage des nucléosomes et de la désacétylation des histones (NuRD) et est donc essentiel pour le maintien de l'expression mono-allélique des gènes soumis à empreinte (c'est-à-dire un gène dont la copie héritée de la mère et la copie héritée du père ne sont pas exprimées de la même manière, le plus généralement une des deux copies est totalement éteinte alors que l'autre est active).

Pour finir, le gène *SMARCA4* (*SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 4*), qui est un facteur de remodelage de la chromatine impliqué dans le maintien de la pluripotence et de l'auto-renouvellement des cellules souches trophoblastiques (Kidder et Palmer, 2010), faisait partie de ce groupe.

➤ Au stade seize cellules

De même qu'aux stades quatre et huit cellules, l'étude de la fonction des gènes trouvés comme activés au stade seize cellules par Graf *et al.* (2014b) révélait leur implication dans des processus de différenciation.

Si dans les stades précédents ces processus étaient clairement tournés vers un maintien de la pluripotence des cellules embryonnaires, au stade seize cellules, quelques gènes avec une fonction dans l'activation de la différenciation apparaissent.

On retrouvait le gène *ELP2* (*Elongator acetyltransferase complex subunit 2*), qui semble impliqué dans l'activation du promoteur du gène *POU5F1* (Barnea et Bergman, 2000), ainsi que le gène *SALL1* (*Sal-like 1*) qui interagit avec les gènes *NANOG* et *SOX2* pour jouer un rôle dans la pluripotence des cellules souches (Karantzali *et al.*, 2011).

Il en allait de même pour *PDCD2* (*Programmed cell death protein 2*). Selon Mu *et al.* (2010), les souris KO pour ce gène présentent une interruption de la croissance de la masse cellulaire interne, ce qui suggère son rôle dans l'auto-renouvellement des cellules pluripotentes embryonnaires.

TEAD4 (*TEA domain family member 4*) était également retrouvé et est nécessaire pour le développement et la différenciation du trophoctoderme. Nishioka *et al.* (2008) ont montré que les souris KO pour ce gène ne sont pas capables de former le blastocœle et meurent au moment de la pré-implantation. Il semblerait qu'il régule à la baisse le gène *POU5F1* (ce qui engendrerait la différenciation du trophoctoderme) et qu'il régule à la hausse la protéine homéobox *Cdx2* dans les cellules du tissu extra-embryonnaire.

Enfin, les auteurs ont mis en évidence le gène *BYSL* (*Bystin-like*) qui est un facteur important de la biogénèse de la sous-unité 40S du ribosome et est requis pour la formation du blastocyste. Adachi *et al.* (2007) ont montré que des souris dont les transcrits du gène *BYSL* sont dégradés par des ARNs ont une différenciation anormale du trophoctoderme.

D'autres gènes identifiés dans cette étude possèdent des fonctions impliquées dans la suite du développement de l'embryon.

Le gène *TSG101* (*Tumor susceptibility gene 101*) est requis pour la prolifération cellulaire avant la gastrulation ainsi que pour la régulation de la voie p53 (facteur de transcription régulant de multiples fonctions cellulaires importantes comme la régulation du cycle cellulaire, l'autophagie ou l'apoptose) (Ruland *et al.*, 2001).

L'étude du gène *FUZ* (*Fuzzy planar cell polarity protein*) chez la souris révèle son importance pour le transport membranaire et pour la voie de signalisation de la polarité planaire cellulaire (Gray *et al.*, 2009).

Le gène *FOSL1* (*Fos-related antigen 1*) est requis pour le développement de la lignée trophoblastique et est également un régulateur clé de la placentation du trophoblaste (Kent *et al.*, 2011).

Les auteurs ont également retrouvé un gène impliqué dans le remodelage de la chromatine : *CHD2* (*Chromodomain helicase DNA binding protein 2*) (Marfella *et al.*, 2006).

➤ Au stade blastocyste

Graf *et al.* (2014b) ont étudiés la fonction de certains gènes trouvés comme s'activant au stade blastocyste.

Là encore, des gènes impliqués dans les processus de différenciation ont été mis en évidence.

Leur fonction était majoritairement tournée vers une différenciation des cellules souches embryonnaires pour former les différentes lignées tissulaires.

Ces résultats s'inscrivaient dans la continuité des fonctions des gènes trouvés comme activés au stade seize cellules qui commençaient à être impliquées dans l'activation de cette différenciation.

Ainsi les auteurs ont repéré le gène *FURIN* (*Furin paired basic amino acid cleavage enzyme*), important dans le développement du placenta et dans la fusion du trophoblaste par traitement du récepteur au facteur de croissance insulino-like 1 (*IGF1R*) et du facteur de croissance de l'endothélium vasculaire (*VEGF*) (Zhou *et al.*, 2013).

D'ailleurs, *IGF1R* s'activait également à ce stade. Bedzhov *et al.* (2012) indiquent que ce gène régule la prolifération, la différenciation et la survie de beaucoup de types cellulaires différents. Sa mutation dans des embryons de souris provoque des anomalies de formation du trophoblaste. Son activation requiert la cadhérine E (*CDH1*) pour le maintien de la lignée trophoblastique. Cela concorde avec les observations de Graf *et al.* (2014b) qui ont retrouvé ce gène comme s'activant en parallèle d'*IGF1R* au stade blastocyste.

Le gène *SNW1* (*SNW domain-containing protein 1*) joue un rôle dans la différenciation de la crête neurale chez les vertébrés (Wu *et al.*, 2011).

De même, Graf *et al.* (2014b) ont retrouvé la protéine homéobox *Cdx2* (*CDX2*).

Chez les souris, *CDX2* réprime le gène *POU5F1* et induit donc la différenciation du trophoblaste (Niwa *et al.*, 2005).

Chez les bovins cependant, Berg *et al.* (2011) ont montré que *CDX2* n'influe pas sur les taux d'*Oct4*, suggérant un rôle différent chez cette espèce.

Enfin, *GATA3* (*Trans-acting T-cell-specific transcription factor*) agirait sur la différenciation du trophoblaste et régulerait même la transcription de *Cdx2* (Ralston *et al.*, 2010).

Le gène *CCL17* (*Chemokine ligand 17*) qui est impliqué dans l'invasion trophoblastique chez les humains (Li *et al.*, 2014) et le gène *TFCP2L1* (*CP2 family transcription factor*), impliqué dans le maintien et l'auto-renouvellement des cellules souches de façon très similaire à *NANOG*, ont également été retrouvés activés à ce stade (Ye *et al.*, 2013).

d. Les voies enrichies par les gènes transcrits *de novo*

Les gènes identifiés comme transcrits *de novo* par l'embryon au cours de l'EGA peuvent également être étudiés comme un ensemble.

En effet, certains de ces gènes agissent en synergie pour faire fonctionner une voie fonctionnelle dans l'embryon. On dit qu'ils enrichissent cette voie.

Pour déterminer si parmi leurs ensembles de gènes, certains enrichissaient significativement des voies fonctionnelles, les auteurs des trois études précédemment citées ont effectué une analyse d'ontologie génique (GO).

➤ Au stade quatre cellules

Les 414 gènes identifiés par Graf *et al.* (2014b) comme s'activant au stade quatre cellules ont subi une analyse d'ontologie génique (GO) afin de faire ressortir les voies fonctionnelles dans lesquelles ils sont impliqués.

Trois voies fonctionnelles ont été ainsi mises en évidence :

- Traitement de l'ARN
- Traduction
- Transport

Ces gènes étant transcrits avant l'EGA, ils pourraient donc être la conséquence de l'activation mineure décelée par Memili et First (1999) et permettraient à l'embryon de se préparer à l'activation majeure au stade huit-seize cellules.

➤ Au stade huit cellules

L'analyse d'ontologie génique (GO) des 258 gènes mis en évidence par Misirlioglu *et al.* (2006) dans leur comparaison ovocytes au stade MII – embryons au stade huit cellules (non traités) a révélé cinq voies fonctionnelles enrichies au stade huit cellules :

- Régulation de la transcription
- Adhésion cellulaire
- Transduction de signaux
- Transport cellulaire
- Métabolisme
- Réponse immune

Dans la comparaison des embryons au stade huit cellules non traités avec les embryons au stade huit cellules traités avec l' α -amanitine de Misirlioglu *et al.* (2006), les transcrits de 233 gènes avaient été trouvés comme étant surexprimés dans les embryons non traités.

Ces gènes sont sensiblement les mêmes que ceux surexprimés dans les embryons huit cellules non traités par rapport aux ovocytes, tout comme les voies fonctionnelles qui sont ressorties.

En ce qui concerne les 208 gènes surexprimés dans des embryons au stade huit cellules non traités par rapport à des embryons au stade huit cellules dont la transcription a été inhibée par Vigneault *et al.* (2009), l'analyse ontologique a révélé que 38% de ces gènes étaient impliqués dans la régulation de la transcription et le traitement de l'ARN et 18% étaient des gènes impliqués dans la biosynthèse protéique.

Graf *et al.* (2014b) avait mis en évidence 4 489 gènes activés au stade huit cellules. Leur analyse ontologique a révélé l'enrichissement des voies fonctionnelles suivantes :

- Transcription ADN-dépendante
- Métabolisme des nucléotides de type purines
- Ubiquitination des protéines

- Initiation de la traduction
- Processus métaboliques de l'ARN

Ces fonctions témoignent de l'activation majeure du génome embryonnaire (initiation de la transcription et de la traduction) mais également de la continuité de la dégradation des transcrits et des protéines maternels.

➤ Au stade seize cellules

Au stade seize cellules, l'activation de 978 gènes a été mise en évidence par Graf *et al.* (2014b), classés dans huit voies fonctionnelles :

- Epissage de l'ARN
- Régulation de la glycolyse
- Processus de biosynthèse de l'ATP
- Régulation négative de la transcription
- Transcription de l'ARNr 5S par l'ARN polymérase III
- Régulation de la transcription par l'ARN polymérase II des catabolites du carbone
- Homéostasie des ions calciques cytosoliques
- Initiation et élongation de la transcription par l'ARN polymérase II

On observe cette fois une tendance des gènes à initier le métabolisme cellulaire tout en maintenant le niveau de transcription et de traduction enclenché au stade précédent.

➤ Au stade blastocyste

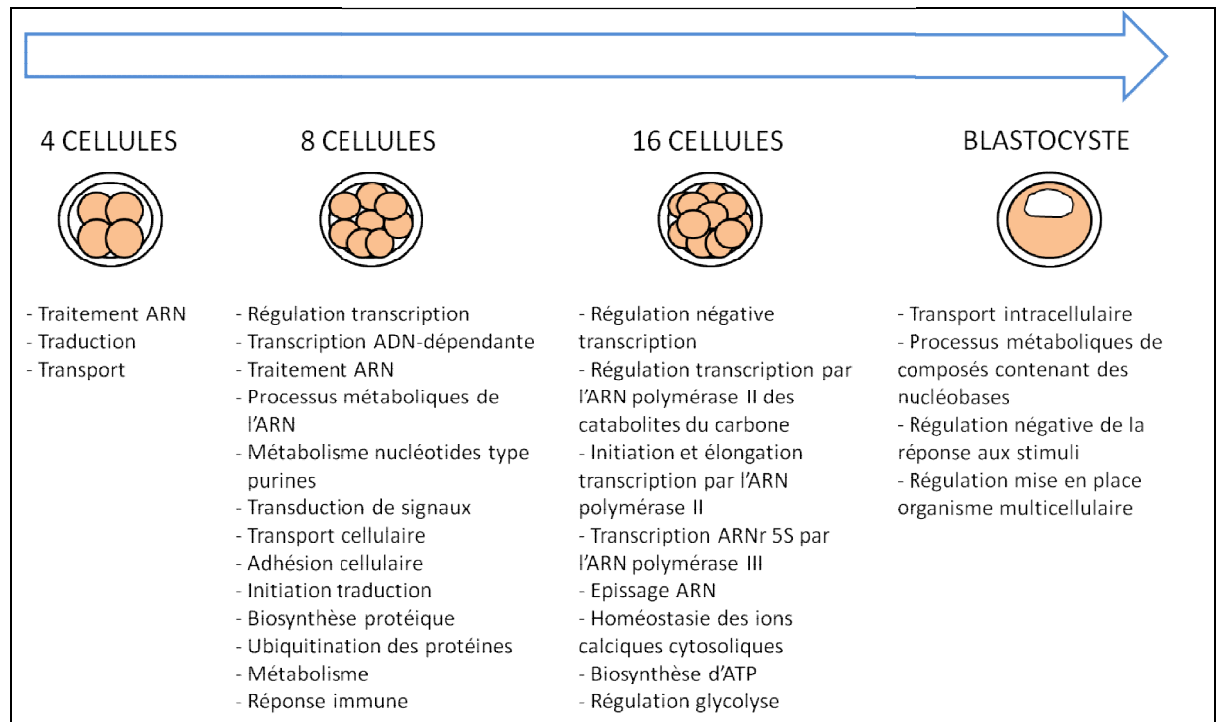
Au-delà du stade 16 cellules et jusqu'au stade blastocyste, 2124 autres gènes s'activaient dans l'étude de Graf *et al.* (2014b). Ils sont impliqués dans les voies fonctionnelles suivantes :

- Transport intracellulaire
- Régulation de la mise en place d'un organisme multicellulaire
- Régulation négative de la réponse aux stimuli
- Processus métabolique de composés contenant des nucléobases

Les premiers processus de différenciation cellulaire apparaissent.

Un schéma récapitulatif des différentes voies fonctionnelles enrichies à chaque stade du développement embryonnaire est présenté en Figure 10.

Figure 11 : Voies fonctionnelles enrichies par les gènes activés *de novo* par l'embryon à différents stades de développement (créée à partir des résultats de Misirlioglu *et al.*, 2006 ; Vigneault *et al.*, 2009 ; Graf *et al.*, 2014b).



En conclusion, ces différentes études quant aux gènes et voies enrichies par la transcription embryonnaire lors de la MET montrent qu'avant le stade huit cellules, l'embryon est soumis à une activation transcriptionnelle mineure qui va lui permettre d'engranger des transcrits utiles à l'activation majeure qui va suivre.

De même, il va devoir mettre en route certains des mécanismes permettant la dégradation du matériel maternel jusque-là stocké et utilisé.

Par ailleurs, avant le stade seize cellules, le transcriptome embryonnaire est également tourné vers un maintien de la pluripotence des cellules.

Suite à l'EGA majeure, aux stades seize cellules et blastocyste, les fonctions des transcrits embryonnaires semblent se diriger vers le maintien d'une transcription et d'une traduction efficace et autonome.

De plus, dès le stade seize cellules, on voit apparaître une différenciation des tissus pour permettre la création des deux types de lignées cellulaires qui forment le blastocyste (la masse cellulaire interne et le trophectoderme).

Les fonctions concernant le métabolisme apparaissent dès le stade huit cellules puis semblent s'intensifier et se développer plus précisément aux stades seize cellules et blastocyste.

Deuxième partie : Étude expérimentale

I. Introduction

Le rendement de la fécondation *in vitro* (FIV) reste globalement faible, avec un taux de production d'embryons entre 35 et 50% et un taux de gestation entre 20 et 55% après un transfert en frais selon les études (Ponsart *et al.*, 2004 ; Pontes *et al.*, 2009 ; Pontes *et al.*, 2011 ; Munoz *et al.*, 2014 ; Arreiseigor *et al.*, 2016).

Selon Le Brusq (2018), la méthode actuellement utilisée a été développée par Lindner et Wright (1983) sur l'embryon produit *in vivo*, cependant très différent des embryons produits *in vitro*.

De ce fait, et au vu des résultats variables après le transfert dans une receveuse, la méthode de classification des embryons est fréquemment remise en cause.

L'observation de la morphocinétique précoce d'embryons de bovins produits semi-individuellement *in vitro* a permis d'identifier une large variabilité de trajectoires de développement embryonnaire ouvrant une perspective de générer une nouvelle classification adaptée aux embryons produits *in vitro*.

Un travail récent développé en partenariat entre les laboratoires BREED et MaIAGE de l'INRAE, a permis de proposer une nouvelle classification des embryons bovins produits *in vitro* basée sur la morphocinétique précoce (Reis *et al.*, 2018). Celle-ci a pour but de déterminer, le plus précocement possible, les embryons qui mourront ou qui s'arrêteront au terme de leur développement *in vitro* d'une part, et les embryons qui continueront à se développer atteignant le stade blastocyste, stade au cours duquel un embryon est apte à être transféré dans une receveuse, d'autre part.

De plus, cette méthode de classification propose de répartir les blastocystes en quatre catégories selon leur trajectoire de développement morphocinétique précoce : catégorie A (blastocystes à éclosion précoce), et les catégories B (blastocyste éclos), C (blastocystes retardés) et D (blastocystes arythmiques). Ces catégories sont classées selon un gradient décroissant de similitude avec les standards de la littérature, jugés normaux. La catégorie D est donc l'extrême opposé de la catégorie A et présente des divisions anormales et/ou une cyclicité arythmique, mais atteint tout de même le stade blastocyste « transférable ».

Pour chaque embryon classé, un pourcentage d'appartenance aux différentes catégories lui est attribué.

Actuellement, cette prédiction est possible dès le 4^e jour post-insémination *in vitro* (jpi), soit au stade huit-neuf cellules.

On sait également que c'est pendant ces étapes de développement précoce que la transition materno-embryonnaire se déroule. En effet, le développement de l'embryon est tout d'abord sous le contrôle du transcriptome maternel avant de passer sous le contrôle de son propre génome (Denny et Tyler, 1964 ; Braude *et al.*, 1979 ; Telford *et al.*, 1990). Elle comprend deux étapes se chevauchant dans le temps : l'élimination des transcrits maternels et l'activation du génome embryonnaire (qui se déroule entre les stades huit et seize cellules chez les bovins).

À partir de l'observation de ces grandes différences de développement morphocinétique précoce, nous nous sommes donc posé la question de l'impact de celles-ci sur cette transition materno-embryonnaire.

Cependant, il est actuellement impossible de suivre en temps réel la transition materno-embryonnaire chez les embryons bovins de manière compatible avec sa survie.

Nous avons donc étudié cette problématique de manière indirecte en analysant le transcriptome d'embryons bovins prédits dans les quatre catégories morphocinétiques de blastocystes (A, B, C et D), le plus précocement possible après l'activation du génome embryonnaire (soit 0 à 12h après la fin du 4^e cycle embryonnaire, c'est-à-dire au stade 16-32 cellules).

L'objectif de cette analyse est de vérifier s'il existe des différences transcriptomiques résultant de ces développements morphocinétiques divergents, post-activation du génome embryonnaire. En cas d'observation d'une différence, un deuxième objectif de cette étude sera d'identifier les gènes et les voies différenciellement exprimées à ce stade.

II. Matériel et méthode

1. Production des embryons

Les ovocytes ont été récupérés post mortem sur des ovaires de bovins à l'abattoir. Ils ont été transportés dans des conditions contrôlées, à 38°C. Il s'est écoulé environ trois heures entre la collecte des ovocytes et leur mise en culture.

Les follicules ont été ponctionnés et le liquide folliculaire a été collecté dans un flacon stérile. Le contenu a ensuite été filtré sur un filtre en acier de 50µM et rincé avec du milieu EuroFlush (IMV®) sur une boîte de Pétri.

Les Complexes Ovocytes Cumulus (COC) ont été identifiés et rincés deux fois dans le milieu de rinçage Euroflush et une fois dans le milieu de maturation avant d'être placés en maturation *in vitro* (MIV).

La qualité des COC a été estimée selon des critères morphologiques conventionnels (Marquant-Le Guenne, 1998). Seuls les COC de qualité 1, 2 et 3 ont été sélectionnés pour la MIV.

2. Maturation in vitro (MIV)

La MIV a été réalisée pendant 22 h, à 38,5°C dans une atmosphère humidifiée à 5% de CO₂. Le milieu de maturation était composé de M199/NaHCO₃ additionné de 10% de sérum de veau fœtal, 10µg/mL de FSH (hormone folliculo-stimulante), 12µg/mL de LH (hormone lutéinisante), 1µg/mL de 17β-estradiol et 10ng/mL de facteur de croissance épidermique (EGF). Les COC ont été mis à maturer par lots de 50, ans 500µL de milieu de maturation.

3. Fécondation in vitro (FIV) et culture embryonnaire

La semence de quatre taureaux différents, de race Holstein, a été utilisée pour la fécondation.

Les spermatozoïdes mobiles ont été séparés du sperme décongelé par un gradient discontinu BoviPure® (40/80) selon les instructions du producteur (Nidacon International AB, Göthenborg, Suède).

La FIV a été réalisée comme indiqué par Parrish *et al.* (1986) dans un milieu Tyrode contenant 10µg/mL d'héparine, 20µM de pénicillamine, 10µM d'hypotaurine et 1µM d'épinéphrine. Les ovocytes ont été incubés avec 10⁶ spermatozoïdes/mL pendant 22 heures à 38,5°C dans 5% de CO₂ dans l'air humidifié.

Le moment dit de l'insémination correspond au moment où les ovocytes et les spermatozoïdes ont été placés ensemble dans la même boîte de culture pour la fécondation *in vitro*. Celui-ci est considéré comme le point de départ (T0) du développement précoce de l'embryon.

À 22h de mise en contact des gamètes, les cellules du cumulus ont été enlevées et les zygotes présumés ont été placés individuellement dans des micro-puits de boîtes de Pétri, adaptés à la production de prises de vues en temps réel (time lapse). Les micro-puits sont couverts par une goutte de 150µL de milieu SOF additionné de 1% de sérum de vache en œstrus, sous huile minérale, et placés dans l'incubateur à 38,5°C, 5% de CO₂, 5% d'O₂ et 90% de N₂ dans l'air humidifié (Holm *et al.*, 1998) jusqu'au jour 4,3 après insémination.

À la fin de la période de culture, les embryons ont été rincés trois fois individuellement dans du PBS, puis mis individuellement dans des tubes Eppendorfs pour une congélation à sec et stockés à -80°C. Un total de 295 embryons a été congelé.

4. Images en Temps réel (Time Lapse)

Les boîtes de Pétri adaptées au système PrimoVision® contenant les embryons ont été placées sur la micro-caméra dans l'incubateur. Une unité de contrôle externe équipée d'un logiciel en cours d'exécution a permis de focaliser l'image initiale et de régler la durée du laps de temps et la fréquence de capture d'image. Un faible courant électrique est envoyé au microscope pendant quelques secondes pour permettre la capture d'image. Afin d'éviter d'endommager l'embryon, le microscope est équipé d'une lumière LED verte homogène. La durée d'éclairage et l'intensité lumineuse sont minimisées sans affecter la résolution.

Pour un bon fonctionnement de PrimoVision®, la température et l'humidité relative doivent être comprises entre 0 et 40°C et 10 et 90% respectivement, ce qui est compatible avec les conditions de culture d'embryons. Dans cette étude, la fréquence de capture d'image a été fixée à 15 minutes. Le logiciel d'analyse Primo Vision® a compilé les images pour générer les vidéos.

5. Classification et formation des lots pour l'analyse transcriptomique

Les films des embryons congelés ont été observés et annotés selon la méthode décrite par Le Brusq (2018). Cinq embryons n'avaient pas complété le 4^e cycle embryonnaire et ont été éliminés de l'étude. Les embryons ayant complété le 4^e cycle embryonnaire ont été analysés par le prédicteur automatique développé au sein de l'équipe (Reis *et al.*, 2018).

Les embryons classés comme étant la catégorie «Ecllosion précoce» seront appelés catégorie A, «Eclos» : Catégorie B ; «Blastocyste retardé» : Catégorie C et «Arythmique» : Catégorie D.

Seize échantillons ont été préparés pour l'analyse transcriptomique : quatre échantillons pour chacune des quatre catégories (A, B, C, D). Chaque échantillon était composé de huit embryons. Ainsi, un total de 128 embryons a été inclus dans cette étude (quatre catégories x quatre échantillons x huit embryons). Afin de supprimer un éventuel effet séance, les embryons ont été distribués dans les échantillons de façon à ce que chaque séance de production soit représentée dans chacun d'eux. De la même façon, tous les taureaux étaient représentés dans tous les échantillons.

6. Extraction de l'ARN et mesure de la pureté des échantillons

L'ARN total des 16 échantillons de huit embryons a été extrait séparément en utilisant le kit PicoPure (Arcturus), selon les instructions du fournisseur. Une étape de DNase a été effectuée suivant les recommandations du kit. Les ARN totaux ont ensuite été analysés (quantitativement et qualitativement) à l'aide d'une puce pico sur le bioanalyzer 20100 (Agilent).

Enfin, la vérification de la qualité des ARN obtenus a été analysée par lecture du RIN (*RNA Integrity Number*) (Mueller *et al.*, 2004).

7. Quantification de l'abondance des ARN

Le séquençage de l'ARN (ARNseq) a été effectué à partir de 500pg d'ARN totaux (ARNtot) sur la plateforme de l'I2BC (Gif-sur-Yvette – France : <https://www.i2bc.paris-saclay.fr/spip.php?article399>).

8. Analyse des données

Les données brutes d'intensité ont été lues sur R (www.r-project.org). Le package Limma a été utilisé pour identifier les gènes différentiellement exprimés.

Compte tenu du grand nombre de comparaisons (environ 27 000), une correction est nécessaire afin de limiter les faux positifs (soit 5% du total). Pour ce faire, les valeurs p obtenues par cette analyse ont été ajustées pour les comparaisons multiples en utilisant la méthode de Benjamini et Hochberg appelée *false discovery rate* (taux de fausses découvertes).

Cela a permis de classer les valeurs p brutes par ordre croissant afin de leur donner un rang ($p\text{-value ajustée (adj } p\text{-value)} = p\text{-value brute} \times \text{nombre de test} / \text{rang}$).

9. Contrôle qualité

Des *spikes* ont été utilisés comme contrôle interne et permettent la validation du différentiel.

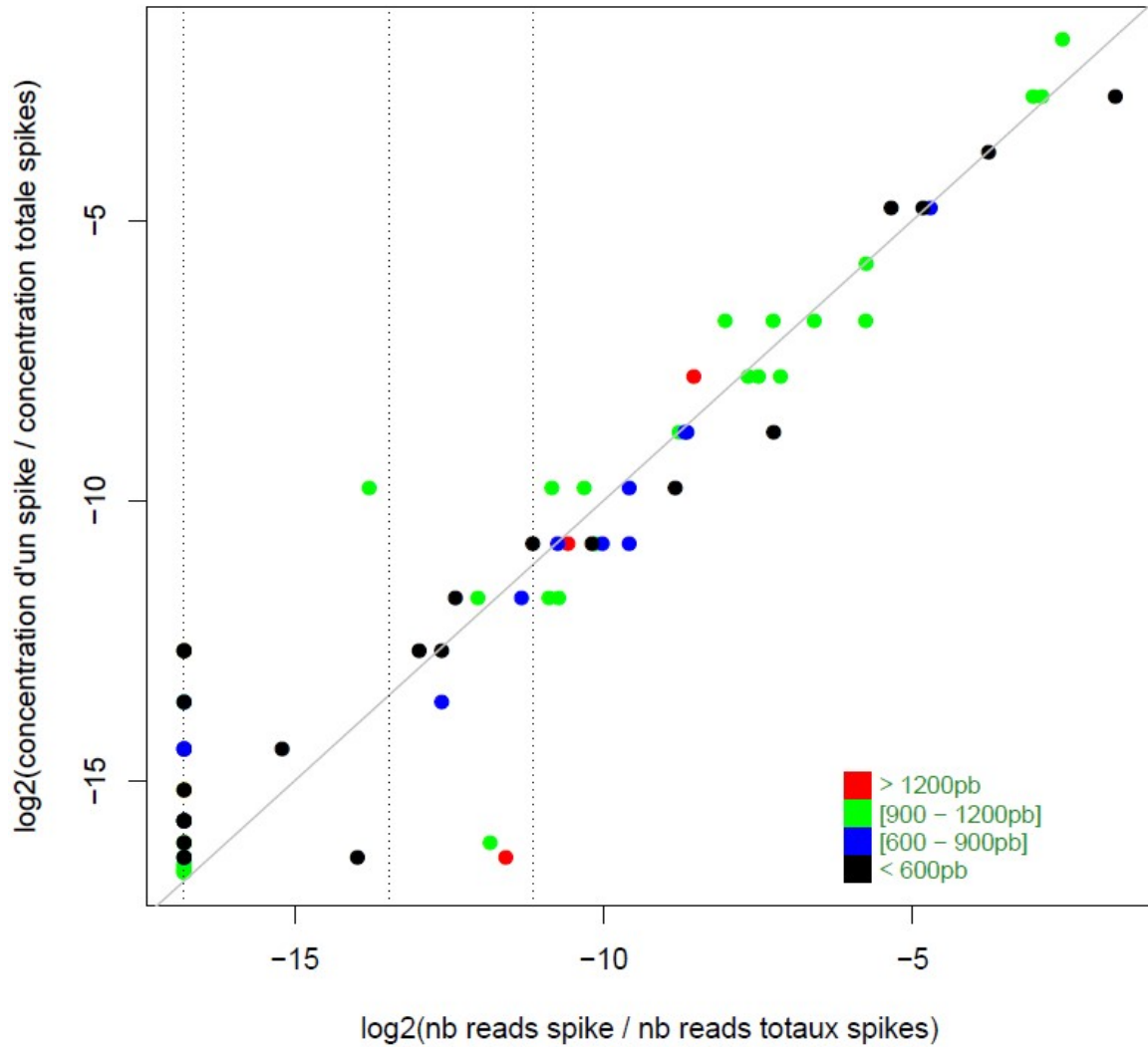
Les *spikes* sont un mélange d'ARN exogènes de taille variable, de séquence connue mais totalement artificielle, sans homologie avec la séquence du génome étudié. Ces ARN sont ajoutés sous forme de mix avec des concentrations variables de chaque ARN.

Dans notre étude, 113 253 *spikes* ont été utilisés, répartis dans nos 16 groupes d'embryon.

La Figure 12 montre que le nombre de lectures des *spikes* est proportionnelle à leur concentration, la fiabilité de nos comptes de lectures est donc correcte, autant pour les transcrits lus de nombreuses fois que pour ceux en moins grande abondance et aussi bien pour les *spikes* contenant

quelques centaines de paires de base que pour ceux de taille plus importante. D'où notre choix de ne pas restreindre notre analyse aux transcrits lus plus de 100 fois.

Figure 12 : Lectures des *spikes* en fonction de leur concentration dans les 16 échantillons analysés



10. Analyse des résultats

a. Analyse descriptive

Dans notre étude, les seize échantillons (quatre par catégorie) sont représentés et annotés A-R1, A-R2, A-R3, A-R4, A, B-R1, B-R2, B-R3, B-R4, C-R1, C-R2, C-R3, C-R4, D-R1, D-R2, D-R3 et D-R4.

➤ Clustering hiérarchique

Le *clustering* hiérarchique consiste à rassembler des individus selon un critère de ressemblance qui s'exprimera sous la forme d'une distance.

La distance existante entre chaque couple d'individus est alors représentée à l'image d'un arbre phylogénétique.

Deux échantillons d'embryons identiques auront une distance nulle et se retrouveront donc au même endroit, sur la même branche de l'arbre. Au contraire, plus les deux échantillons seront dissemblables, et plus la distance sera importante, ils seront donc représentés sur des branches différentes, aussi éloignées l'une de l'autre que le sont les deux échantillons.

Il existe deux façons possibles de calculer les distances rendant compte de la ressemblance entre deux éléments : les distances euclidiennes ou les distances basées sur la corrélation.

Si l'on veut identifier des échantillons avec les mêmes profils globaux, quelles que soient leurs amplitudes, l'utilisation de la distance basée sur la corrélation est le meilleur choix.

C'est particulièrement le cas dans l'analyse des données d'expression génique, où l'on pourrait vouloir considérer des ensembles d'embryons similaires lorsqu'ils expriment fortement les mêmes gènes, mais également s'ils expriment faiblement des gènes communs.

A contrario, la distance euclidienne considère des échantillons proches seulement s'ils expriment fortement les mêmes gènes.

C'est pour cette raison que nous choisirons d'exposer les résultats du clustering par corrélation (et non euclidien) dans la suite de ce travail.

➤ Analyse en composantes principales (ACP)

L'objectif de l'ACP est de décrire graphiquement un tableau de données d'individus avec des variables quantitatives de grande taille. Les échantillons sont représentés par un nuage de points dans un espace à n dimensions. L'ACP va projeter ce nuage de points sur un graphique à deux axes (c'est-à-dire à deux dimensions) en déterminant les axes qui expliquent le mieux la dispersion de ce nuage de points.

b. Analyse différentielle

➤ Abondance des transcrits des gènes étudiés

Les transcrits de chaque gène ont tout d'abord été dénombrés.

Chaque fois que l'ARN d'un gène est lu par la machine, il est comptabilisé (nombre de « *read* » ou lecture). Le nombre de lectures est proportionnel à l'abondance des ARN correspondants dans l'embryon.

Dans un premier temps, une analyse exploratoire a été proposée pour mieux connaître les résultats dans leur ensemble. Ainsi, un premier tableau a été construit afin de quantifier le nombre de transcrits absents (0 lecture), ceux qui ont un nombre de lectures strictement supérieur à zéro et allant jusqu'à dix lectures maximum, ceux qui ont un nombre de lectures strictement supérieur à dix et allant jusqu'à 100 lectures maximum et enfin ceux qui ont été lus strictement plus de 100 fois.

Chaque transcrit étant associé à un gène, on peut ainsi dire si un gène est pas du tout, faiblement, moyennement ou fortement exprimé dans les embryons des seize sous-groupes étudiés.

➤ Analyse de la comparaison quantitative entre catégories

Six comparaisons ont été effectuées à partir des données des puces (tel que décrit dans la partie II.8) : A vs B, A vs C, A vs D, B vs C, B vs D, C vs D. Celles-ci avaient pour but de quantifier puis d'identifier les gènes surexprimés ou sous-exprimés grâce à l'abondance de leurs transcrits dans les échantillons des catégories comparées.

Le nombre de gènes pour lesquels cette différence d'expression est statistiquement significative a été recensé.

Deux indicateurs statistiques ont été utilisés pour concrétiser cette notion d'expression différentielle : le *Fold-Change* (FC) et la valeur p ajustée.

Le *Fold-Change* est le rapport du niveau moyen d'expression d'un gène exprimé dans une condition par rapport à une autre. Comme le niveau d'expression entre les différents échantillons peut varier de 0 à plusieurs milliers, une transformation log a été appliquée pour rendre les résultats comparables (rapports symétriques par rapport à 0).

Dans une comparaison X vs Y, un gène avec un *Fold-Change* > 1 sera surexprimé dans la condition X par rapport à la condition Y. Au contraire si ce *Fold-Change* est < 1, le gène sera considéré comme sous-exprimé dans la condition X par rapport à la condition Y.

Les valeurs limites choisies dans cette étude sont 0,5 et 2, c'est-à-dire que pour un FC < 0,5, le gène est considéré comme sous-exprimé dans la première catégorie (deux fois moins exprimé), et pour un FC > 2, le gène est considéré comme surexprimé dans la première catégorie (deux fois plus exprimé).

Enfin, pour affiner ce travail, nous avons sélectionné les gènes dont la différence d'expression entre deux catégories était significative d'un point de vue statistique.

La p-value représente la probabilité d'obtenir au moins la même valeur du test statistique si l'hypothèse nulle H_0 était vraie, c'est-à-dire s'il n'existe pas de différence réelle entre les deux groupes comparés.

On compare la p-value à un seuil. Pour cette étude, nous avons choisi un intervalle de confiance de 95%. Ainsi :

- si la p-value est < 5%, H_0 est rejetée et la différence observée est significative (avec un risque d'erreur de 5%)
- si la p-value est > 5%, on ne rejette pas l'hypothèse H_0 et on ne peut rien conclure.

Nous avons utilisé la valeur p ajustée (p-adj), comme précédemment décrit dans la partie II.8. Les comparaisons résultant en une valeur p ajustée < 0,05 ont été considérées comme présentant une différence significative.

Ainsi, un recensement des gènes significativement sur ou sous-exprimés dans une catégorie par rapport à une autre a pu être effectué.

La suite de l'analyse des résultats s'est concentrée uniquement sur la comparaison qui mettait en évidence le plus grand nombre de différences transcriptomiques.

➤ Analyse des profils des gènes différentiellement exprimés dans la comparaison mettant en évidence le plus grand nombre de différences transcriptomiques

Le niveau d'expression attendu chez l'embryon *in vivo* pour les deux ensembles de gènes différentiellement exprimés dans la comparaison mettant en évidence le plus grand nombre de différences transcriptomiques a été étudié grâce à la base de données du logiciel propriétaire EmbryoExpress, disponible dans l'UMR BREED (<http://www.embryoexpress.org/welcome>).

Ces valeurs de référence (trois échantillons par stade) obtenues *in vivo* dans des ovocytes aux stades GV (Vésicule Germinale) et MII (Métaphase II) et dans des embryons aux stades quatre cellules, huit cellules, seize cellules et blastocyste nous ont permis d'établir notre propre classification de profil d'expression.

En effet, tous les gènes différentiellement exprimés dans la comparaison sélectionnée se sont vus attribuer un profil type :

- Profil maternel définit comme un gène transcrit au cours de l'ovogénèse et dont les produits (ARN et protéines) sont stockés dans l'ovocyte pour le développement précoce du futur embryon. Ce stock diminue ensuite, sans qu'une phase de transcription embryonnaire (ou *de novo*) ne soit identifiable lors de la transition materno-embryonnaire.
- Profil embryonnaire définit comme un gène transcrit pour la première fois par l'embryon, avec sa propre machinerie transcriptionnelle, à la suite de l'activation de son génome, et non présent dans l'ovocyte au moment de la fécondation.
- Profil materno-embryonnaire ou mixte définit comme un gène dont les transcrits sont stockés et présents dans l'ovocyte au moment de la fécondation puis diminuent lors de la transition materno-embryonnaire, mais que l'embryon va par la suite transcrire *de novo*, avec sa propre machinerie transcriptionnelle, à la suite de l'activation de son génome. Cette transcription embryonnaire est observable par une augmentation de l'abondance des transcrits à au moins un stade embryonnaire.

Pour ce faire, tout d'abord, une médiane du niveau d'expression des gènes a été calculée à chaque stade de développement disponible dans EmbryoExpress.

Si la médiane des lectures (ou *reads*) au stade MII était égale à zéro et que au moins une médiane des stades quatre cellules, huit cellules, seize cellules ou blastocyste était non nulle alors le gène était classé comme étant un gène au profil embryonnaire.

Pour les gènes dont les médianes aux stades GV et MII étaient non nulles : si la médiane du stade GV ou la médiane du stade MII était supérieure ou égale à la médiane du stade quatre cellules, elle-même supérieure ou égale à la médiane du stade huit cellules, elle-même supérieure ou égale à la médiane du stade seize cellules, elle-même supérieure ou égale à la médiane du stade blastocyste, alors le gène était considéré comme étant de profil maternel.

Les autres gènes étaient classés comme des gènes à profil materno-embryonnaire.

➤ Etude d'enrichissement des ensembles de gènes différentiellement exprimés dans la comparaison mettant en évidence le plus grand nombre de différences transcriptomiques

Les listes des gènes surexprimés dans l'une ou l'autre des deux catégories de la comparaison mettant en évidence le plus grand nombre de différences transcriptomiques ont été soumises à une étude d'enrichissement.

L'étude d'enrichissement est obtenue par une méthode statistique permettant d'inférer des connaissances sur la fonction d'ensembles de gènes issue d'études antérieures à un ensemble de gènes que l'on souhaite étudier.

En effet, un ensemble de gènes peut partager une fonction biologique commune. Les voies fonctionnelles dans lesquelles ils sont impliqués ont été découvertes grâce aux puces à ADN ou au séquençage de l'ARN et les gènes impliqués dans chaque voie sont ainsi annotés dans des bases de données (librairies).

Dans notre étude, celle-ci a été réalisée grâce à la plateforme Enrichr (Chen et al., 2013 ; Kuleshov et al., 2016 : <https://amp.pharm.mssm.edu/Enrichr/>) et a permis la mise en évidence des voies fonctionnelles significativement enrichies par les gènes différentiellement exprimés par les embryons des catégorie comparées.

En effet, chaque ensemble de gènes répertorié dans les bases de données disponibles sur Enrichr est associé à un terme fonctionnel ou à un terme d'enrichissement tel qu'une voie, une lignée cellulaire ou un phénotype.

Enrichr implémente alors quatre scores pour rendre compte des résultats d'enrichissement : valeur p, valeur q, rang (score Z) et score combiné.

La valeur p est une probabilité statistique. Elle permet de juger si les gènes que l'on souhaite étudier sont significativement impliqués dans les voies/fonctions identifiées par le logiciel.

La valeur q est une valeur p ajustée.

Le score de classement ou score z est calculé pour évaluer l'écart par rapport au classement attendu.

Le score combiné est une combinaison de la valeur p et du score z.

Dans le cas de notre étude, la sélection des voies fonctionnelles dans lesquelles les gènes différentiellement exprimés étaient impliqués a été réalisée en fonction de la valeur q. Lorsque la valeur q était inférieure à 0,05, alors nous considérons que la voie fonctionnelle était significativement enrichie dans la catégorie en question.

Cinq librairies ont été utilisées : le Kegg de souris, les processus biologiques, les fonctions moléculaires, les composants cellulaires et les phénotypes des mammifères.

Par la suite, une étude des connexions entre les différentes voies fonctionnelles de la base de données GO (*Gene Ontology*) ainsi mises en évidence a été effectuée grâce à la plateforme QuickGO (Binns et al., 2009 : <https://www.ebi.ac.uk/QuickGO/>).

Les voies fonctionnelles annotées dans la base de données MGI (Mouse Genome Informatics) ne sont pas répertoriées dans la plateforme QuickGo et n'ont donc pas été étudiées de cette façon.

III. Résultats

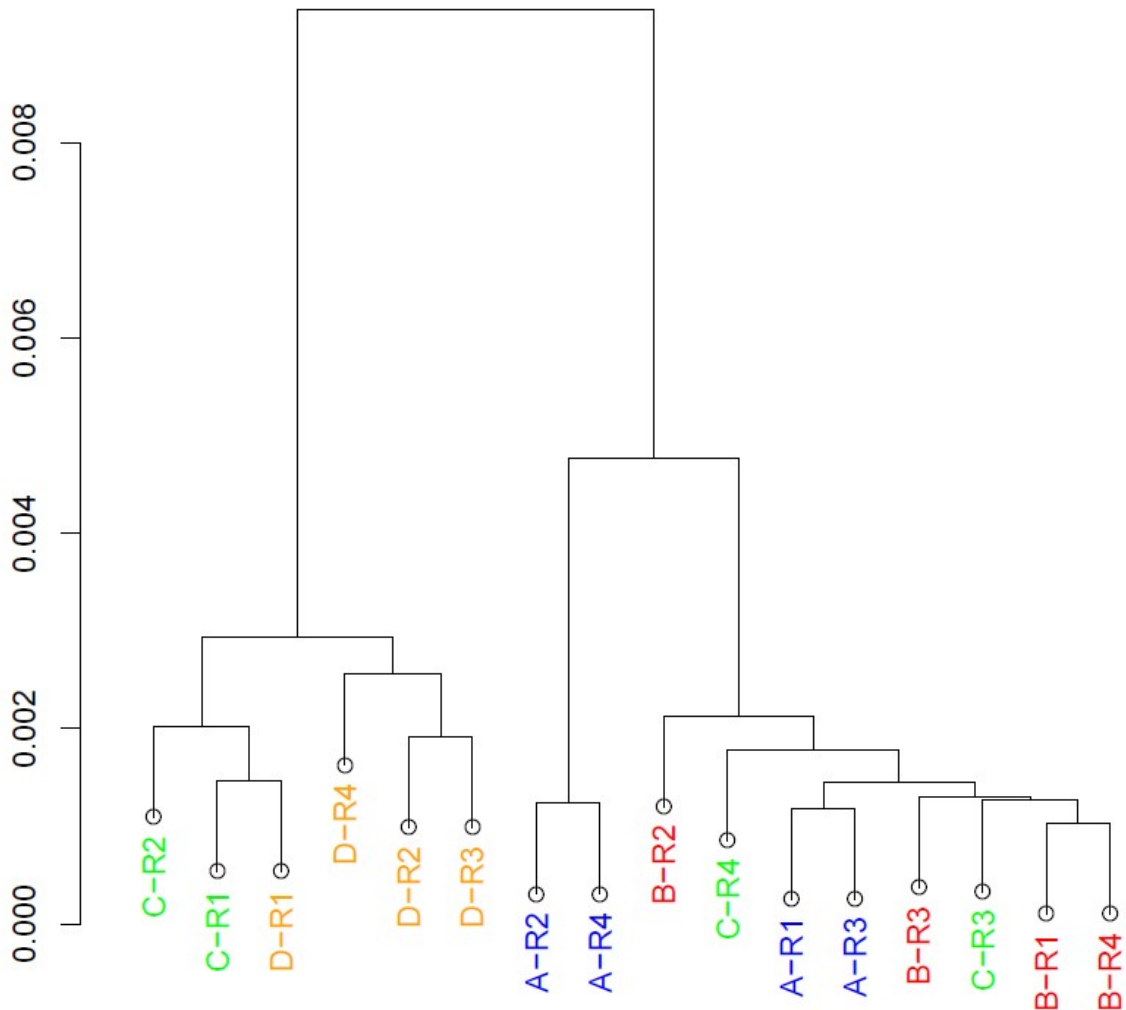
1. Analyse descriptive

Nous possédons quatre répétitions de huit embryons dans chacune des quatre catégories, soit quatre sous-groupes pour chaque profil : A-R1, A-R2, A-R3, A-R4 pour la catégorie A ; B-R1, B-R2, B-R3, B-R4 pour la catégorie B ; C-R1, C-R2, C-R3, C-R4 pour la catégorie C ; D-R1, D-R2, D-R3, D-R4 pour la catégorie D.

a. *Clustering* hiérarchique de corrélation

Les résultats du *clustering* hiérarchique de corrélation sont présentés en Figure 13.

Figure 13 : *Clustering* hiérarchique de corrélation des quatre répétitions de chaque catégorie (A, B, C et D)



On observe que l'ensemble des sous-groupes de la catégorie D possède une grande distance avec les catégories A et B et les sous-groupes C-R3 et C-R4. Les sous-groupes C-R1 et C-R2 se retrouvent proches des sous-groupes de la catégorie D.

Au contraire, les catégories A et B ont des distances relativement courtes entre elles. Deux sous-groupes de la catégorie C (C-R3 et C-R4) se retrouvent également à de faibles distances des catégories A et B.

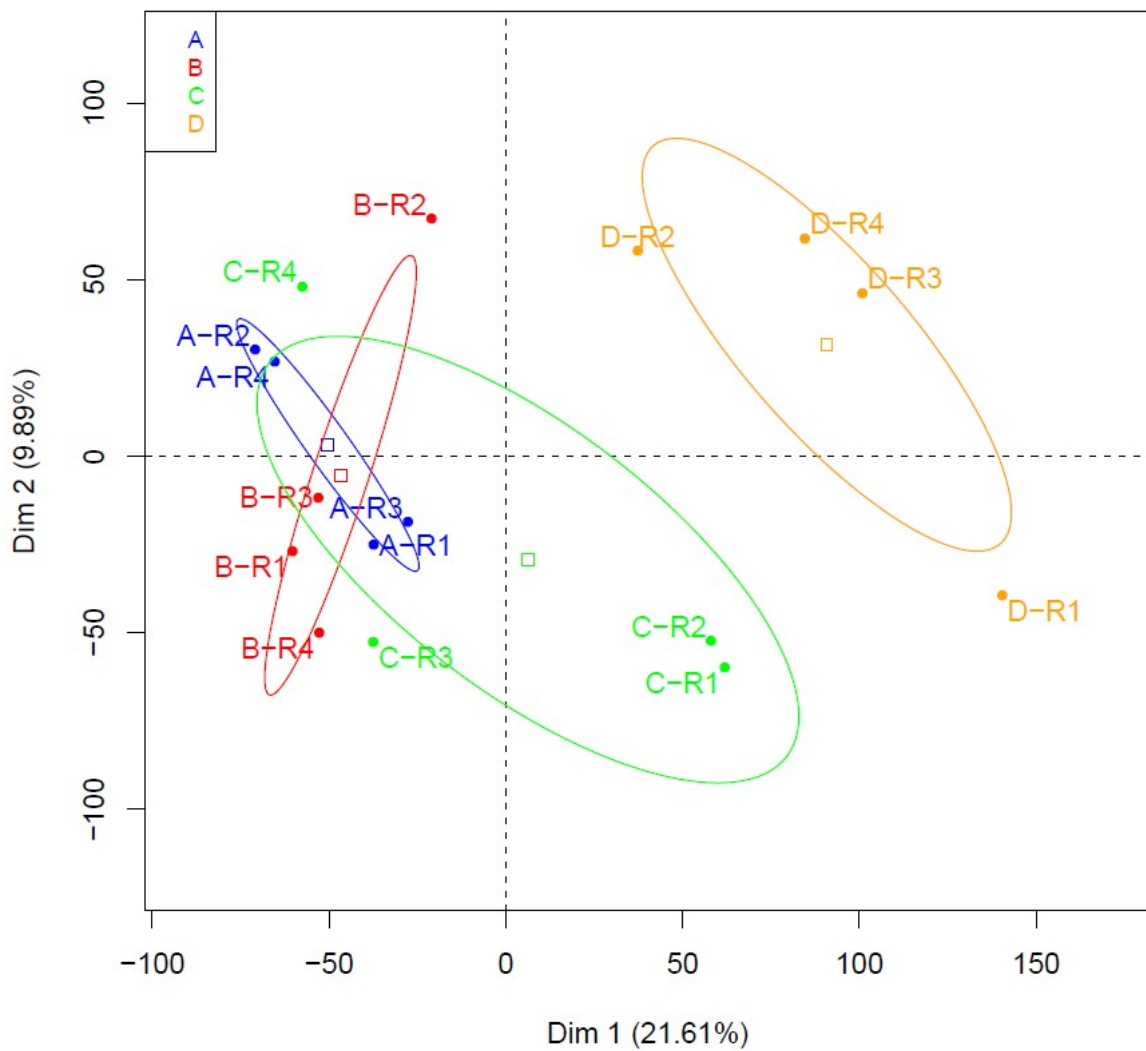
Il en ressort donc que les catégories A et B sont assez proches d'un point de vue de l'expression génique immédiatement après l'activation du génome embryonnaire et s'opposent à la catégorie D

qui, elle, est plus éloignée. La catégorie C semble être un intermédiaire entre le bloc A et B et la catégorie D.

b. Analyse en composantes principales (ACP)

Les résultats de l'analyse en composantes principales sont présentés en Figure 14.

Figure 14 : Analyse en Composantes Principales des quatre répétitions de chaque catégorie (A, B, C et D)



On observe que les nuages de points des catégories A et B sont plutôt proches, à gauche de l'axe vertical. La catégorie C est en partie proche de A et B, notamment les sous-groupes C-R3 et C-R4, qui se regroupent à gauche de l'axe vertical. Mais ses deux autres sous-ensembles C-R1 et C-R2 semblent se détacher à droite de l'axe vertical.

La catégorie D, elle, se détache complètement des trois précédentes en se plaçant totalement à droite de cet axe.

A nouveau, la catégorie D se différencie des autres catégories alors que les catégories A et B semblent être davantage proches.

2. Analyse différentielle

a. Abondance des transcrits des gènes étudiés

L'analyse du transcriptome des 16 sous-groupes comptabilisait un total de 27 325 transcrits, dont 20 230 étaient associés à des gènes connus.

Pour chacun des 16 sous-groupes, le nombre de lectures des transcrits des gènes a été rassemblé et étudié.

Les ARN ont alors été classés en quatre catégories selon s'ils sont absents (nombre de transcrits lus égal à zéro), faiblement présents (nombre de transcrits lus supérieur à zéro mais inférieur ou égal à 10), moyennement présents (transcrits lus plus de dix fois mais ne dépassant pas 100 lectures) ou fortement présents (transcrits lus plus de 100 fois) (Tableau 2).

Tableau 2 : Nombre d'ARN absents (0 transcrits lus), faiblement présents (entre 0 et 10 transcrits lus), moyennement présents (entre 10 et 100 transcrits lus) ou fortement présents (>100 transcrits lus), dans les quatre répétitions de chaque catégorie (A, B, C et D)

Sous-groupes	ARN absents (transcrits lus=0)	ARN faiblement présents (0<transcrits lus≤10)	ARN moyennement présents (10<transcrits lus≤100)	ARN fortement présents (transcrits lus>100)	Total
A-R1	8 293	2 756	4 923	11 353	27 325
A-R2	8 722	2 576	4 603	11 424	27 325
A-R3	8 384	3 043	4 609	11 289	27 325
A-R4	8 770	2 518	4 587	11 450	27 325
B-R1	8 620	2 948	4 572	11 185	27 325
B-R2	7 903	2 892	4 848	11 682	27 325
B-R3	8 289	2 919	4 810	11 307	27 325
B-R4	8 664	2 797	4 624	11 240	27 325
C-R1	7 745	3 263	4 871	11 446	27 325
C-R2	7 853	2 962	4 941	11 569	27 325
C-R3	8 782	2 748	4 616	11 179	27 325
C-R4	8 412	2 542	4 834	11 537	27 325
D-R1	7 188	3 640	4 784	11 713	27 325
D-R2	8 049	2 715	4 812	11 749	27 325
D-R3	7 349	3 134	4 890	11 952	27 325
D-R4	7 390	2 842	4 924	12 169	27 325

Par ailleurs, nous avons identifié 3 674 gènes pour lesquels aucun transcrit n'a été lu dans aucun des 16 échantillons de notre étude.

La question de la pertinence de la recherche de ces ARN chez l'embryon bovin à ce stade du développement s'est alors posée.

Afin d'y répondre, nous nous sommes référées à la base de données transcriptomiques de l'embryon bovin produit *in vivo* disponible dans EmbryoExpress.

Parmi ces 3 674 gènes, 14 avaient une médiane d'expression non nulle au stade seize cellules et au stade blastocyste sur l'ensemble des trois échantillons utilisés par EmbryoExpress : *FAM186A*, *HSD3B7*, *IL37*, *KCTD21*, *MNF1*, *PREB*, *RF01296*, *ZACN*, *ENSBTAG0000007333*, *ENSBTAG0000009273*, *ENSBTAG00000016173*, *ENSBTAG00000030714*, *ENSBTAG00000037632* et *ENSBTAG00000040098*.

Sept d'entre eux ont un profil embryonnaire selon la classification décrite en II.10.b.

Deux ont un profil maternel et cinq un profil materno-embryonnaire.

Deux gènes avaient une médiane d'expression supérieure à dix *reads* au stade seize cellules. Un de ces gènes n'est pas encore connu et a pour code Ensemble *ENSBTAG00000040098*. Il a un profil materno-embryonnaire. Ce gène semble être fortement transcrit lors de l'activation majeure puisqu'on observe une forte augmentation de l'abondance de ses transcrits entre les stades quatre et seize cellules. Puis, les transcrits se dégradent assez rapidement jusqu'au stade blastocyste.

De la même manière, le deuxième gène : *RF01296*, est très faiblement transcrit dans l'ovocyte (stade métaphase II) et voit l'abondance de ses transcrits augmenter fortement entre les stades huit et seize cellules. Mais ses transcrits restent à un niveau élevé par la suite du développement embryonnaire. Ce gène code pour l'ARNsno (ARN small nucleolar) U85.

Les ARNsno sont une classe de petits ARN qui jouent un rôle dans les modifications chimiques d'autres ARN, principalement les ARN ribosomiques, les ARN de transfert et les petits ARN nucléaires. Il en existe deux classes principales : les ARNsno de boîte C/D, qui sont associés à la méthylation, et les ARNsno de boîte H/ACA, qui sont associés à la pseudouridylation (la pseudouridine est un ribonucléotide résultant de la modification post-transcriptionnelle d'une uridine, que l'on retrouve dans les ARN non codants).

b. Comparaison quantitative entre catégories

Les différentes catégories (A, B, C, et D) ont été ensuite comparées deux à deux. On obtenait donc six comparaisons : A vs B, A vs C, A vs D, B vs C, B vs D, C vs D.

Les résultats de ces comparaisons quantitatives sont présentés dans le Tableau 3.

La colonne Total représente la totalité des gènes pour lesquels au moins dix lectures de leurs transcrits ont été comptabilisées dans au moins un des échantillons des quatre catégories comparées (A, B, C ou D).

Si dix lectures sont comptabilisées pour un gène donné qui n'apparaît que dans une des catégories, par exemple la catégorie D, il va tout de même se retrouver dans les résultats de toutes les analyses différentielles (A vs B ; A vs C, etc.). Cela explique que l'on ait le même nombre de gènes au total dans toutes les comparaisons, ainsi que la différence de 6 636 gènes entre le total de tous les sous-groupes (Tableau 2) et les gènes valorisés dans les comparaisons de catégories (Tableau 3). La colonne FC < 0,5 représente le nombre de gènes sous-exprimés dans la première catégorie par rapport à la seconde et la colonne FC < 0,5 et p-adj < 0,05 concerne les gènes significativement sous-exprimés dans la première catégorie de la comparaison. De la même manière, la colonne FC > 2 représente le nombre de gènes surexprimés dans la première catégorie par rapport à la seconde et la colonne FC > 2 et p-adj < 0,05 concerne les gènes significativement surexprimés dans la première catégorie de la comparaison.

Tableau 3 : Nombre de gènes inclus dans les comparaisons entre les différentes catégories d'embryons (A, B, C, D) (Total) et nombre de gènes différentiellement exprimés en fonction des seuils de *Fold Change* (<0,5 ou >2) et de la valeur *p* ajustée (<0,05) choisis dans la présente étude

Nombre de gènes	Total	FC < 0.5	FC < 0.5 et p-adj < 0.05	FC > 2	FC > 2 et p-adj < 0.05
A VS B	20 689 dont 4 391 non connus	1 754 dont 513 non connus	21 dont 4 non connus	1 656 dont 450 non connus	24 dont 2 non connus
A VS C	20 689 dont 4 391 non connus	2 033 dont 543 non connus	54 dont 6 non connus	1 576 dont 460 non connus	17 dont 7 non connus
A VS D	20 689 dont 4 391 non connus	3 507 dont 791 non connus	899 dont 114 non connus	1 113 dont 418 non connus	60 dont 24 non connus
B VS C	20 689 dont 4 391 non connus	1 777 dont 481 non connus	19 dont 1 non connu	1 437 dont 487 non connus	6
B VS D	20 689 dont 4391 non connus	3 461 dont 779 non connus	699 dont 78 non connus	1 074 dont 436 non connus	62 dont 29 non connus
C VS D	20 689 dont 4391 non connus	3 037 dont 727 non connus	258 dont 31 non connus	1 082 dont 371 non connus	40 dont 7 non connus

Un grand nombre de gènes présente des différences de niveau d'expression entre les différentes catégories (plus de 1 000 gènes dans toutes les comparaisons), mais seulement un petit nombre de gènes l'est de manière significative (moins de 1000 dans toutes les comparaisons).

Un grand nombre de gènes est significativement sous-exprimé dans les catégories A, B et C par rapport à la catégorie D. Il est possible de voir un gradient quantitatif, où, 899, 699 et 258 gènes sont sous-exprimés en A, B et C respectivement par rapport à D. Toujours quantitativement, les catégories B et C semblent être les plus proches, vu que seulement 25 gènes sont différentiellement exprimés entre elles (sous-exprimés en B (n=19) + surexprimés en B (n=6)). Les catégories A et D semblent être les plus éloignées : 959 gènes différentiellement exprimés (sous-exprimés en A (n=899) + surexprimés en A (n=60)).

c. Diagrammes de Venn

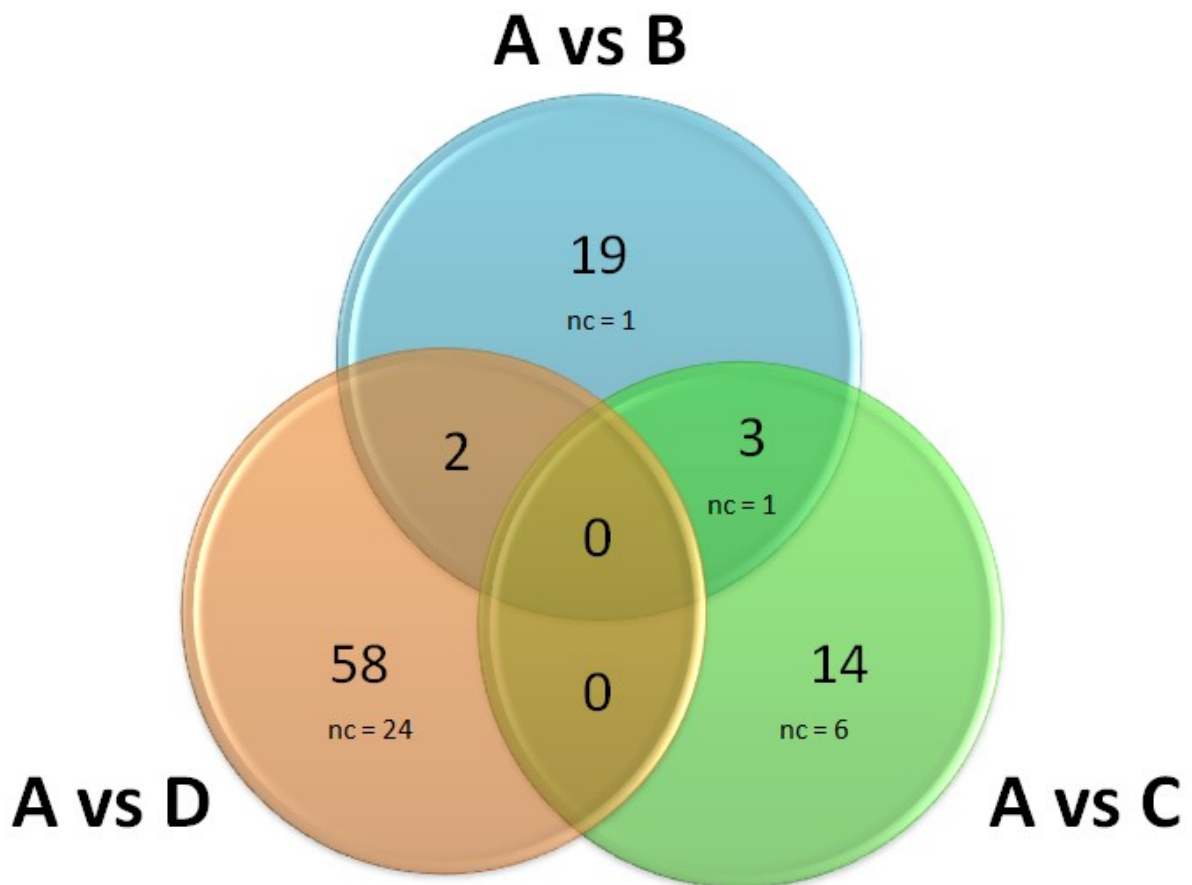
Pour aller plus loin, les résultats du Tableau 2 ont été représentés sous forme de quatre diagrammes de Venn, en prenant en compte uniquement les gènes significativement sous ou surexprimés dans chaque comparaison (FC > 2 ou FC < 0,5 et p-adj < 0,05).

Les gènes connus et non connus y sont représentés, les non connus sont annotés par le sigle « nc » sur les diagrammes.

La liste complète des gènes différentiellement exprimés entre les quatre catégories se trouve en Annexes 1, 2, 3 et 4.

Dans la Figure 15 sont représentés les gènes surexprimés dans la catégorie A par rapport aux trois autres catégories (A vs B ; A vs C ; A vs D).

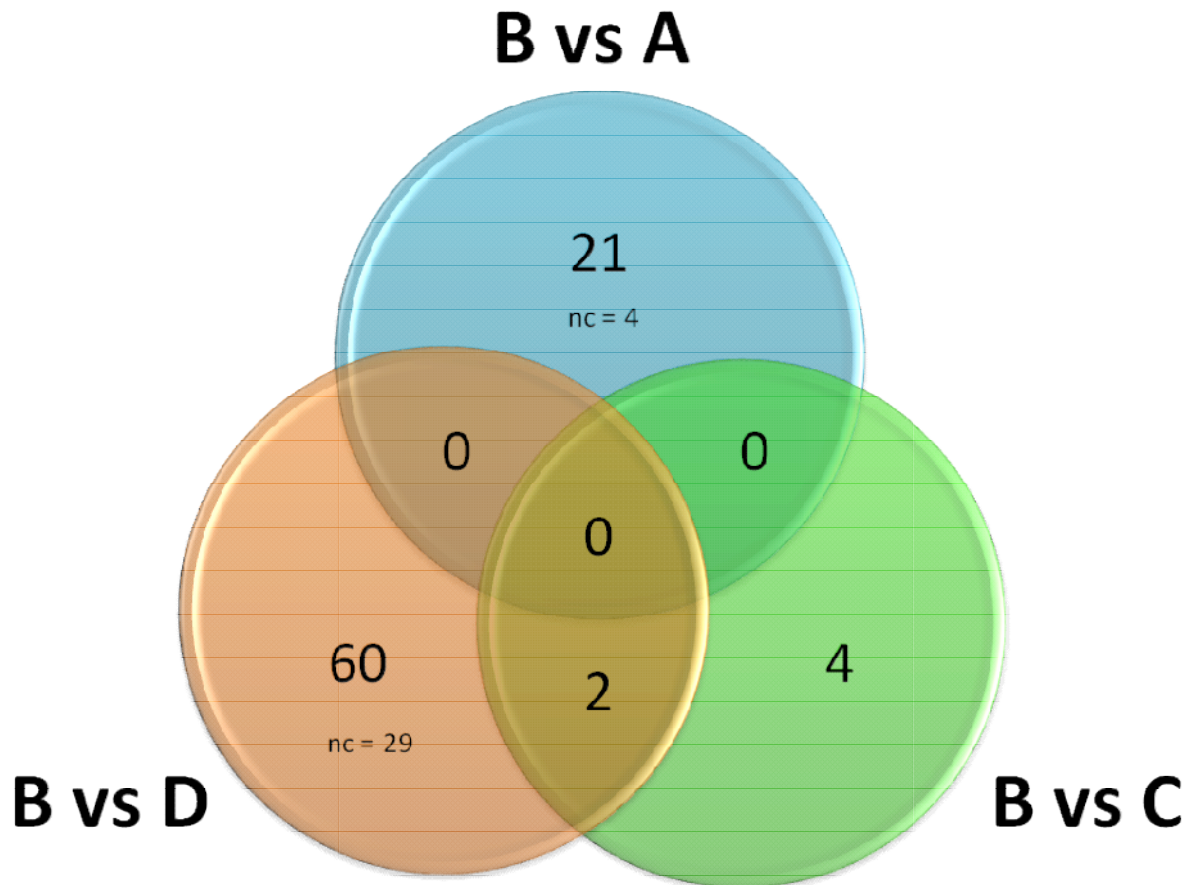
Figure 15 : Gènes significativement surexprimés dans la catégorie A par rapport aux catégories B, C et D (FC > 2 et p-adj < 0,05)



Aucun gène n'est surexprimé en A par rapport aux trois autres catégories (B, C et D). Deux gènes sont surexprimés en A par rapport à B et D (les gènes *KCNE4* et *TPST2*). Trois gènes sont surexprimés en A par rapport à B et C, dont un non connu (*OLR1*, *IFIT1* et *ENSBTAG00000003892*). Les 91 autres gènes surexprimés en A sont exclusifs à une des comparaisons : 58 sont surexprimés en A par rapport à D ; 19 sont surexprimés en A par rapport à B et 14 sont surexprimés en A par rapport à C.

La Figure 16 illustre les gènes surexprimés en B par rapport aux trois autres catégories (B vs A ; B vs C ; B vs D).

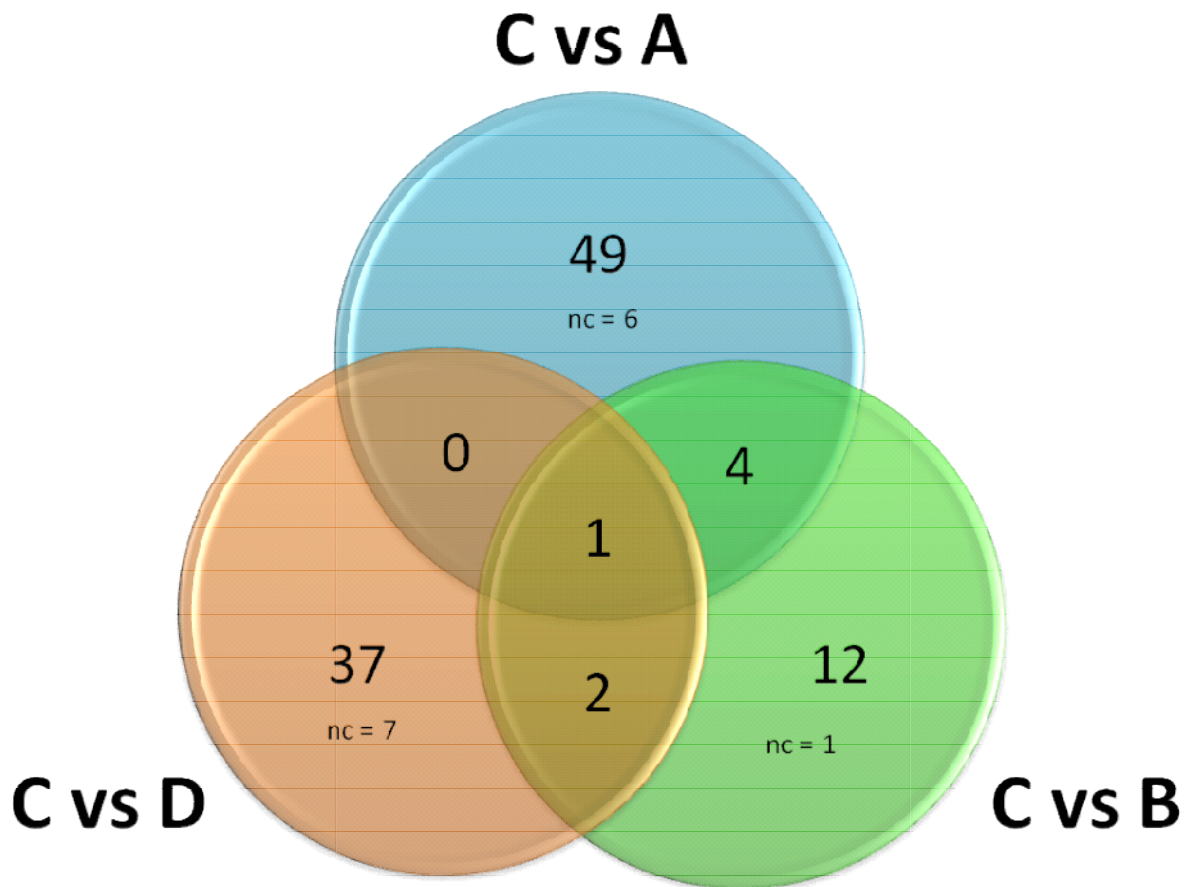
Figure 16 : Gènes significativement surexprimés dans la catégorie B par rapport aux catégories A, C et D (FC > 2 et p-adj < 0,05)



Aucun gène n'est surexprimé en B par rapport aux trois autres catégories (A, C et D). Seulement deux gènes sont surexprimés en B par rapport à C et D (*CCDC172* et *ZNF770*). Les 85 autres gènes surexprimés en B sont exclusifs à une des comparaisons : 60 sont surexprimés en B par rapport à D ; 21 sont surexprimés en B par rapport à A et 4 sont surexprimés en B par rapport à C.

Dans la Figure 17 sont représentés les gènes surexprimés dans la catégorie C par rapport aux trois autres catégories (comparaisons : C vs A ; C vs B ; C vs D).

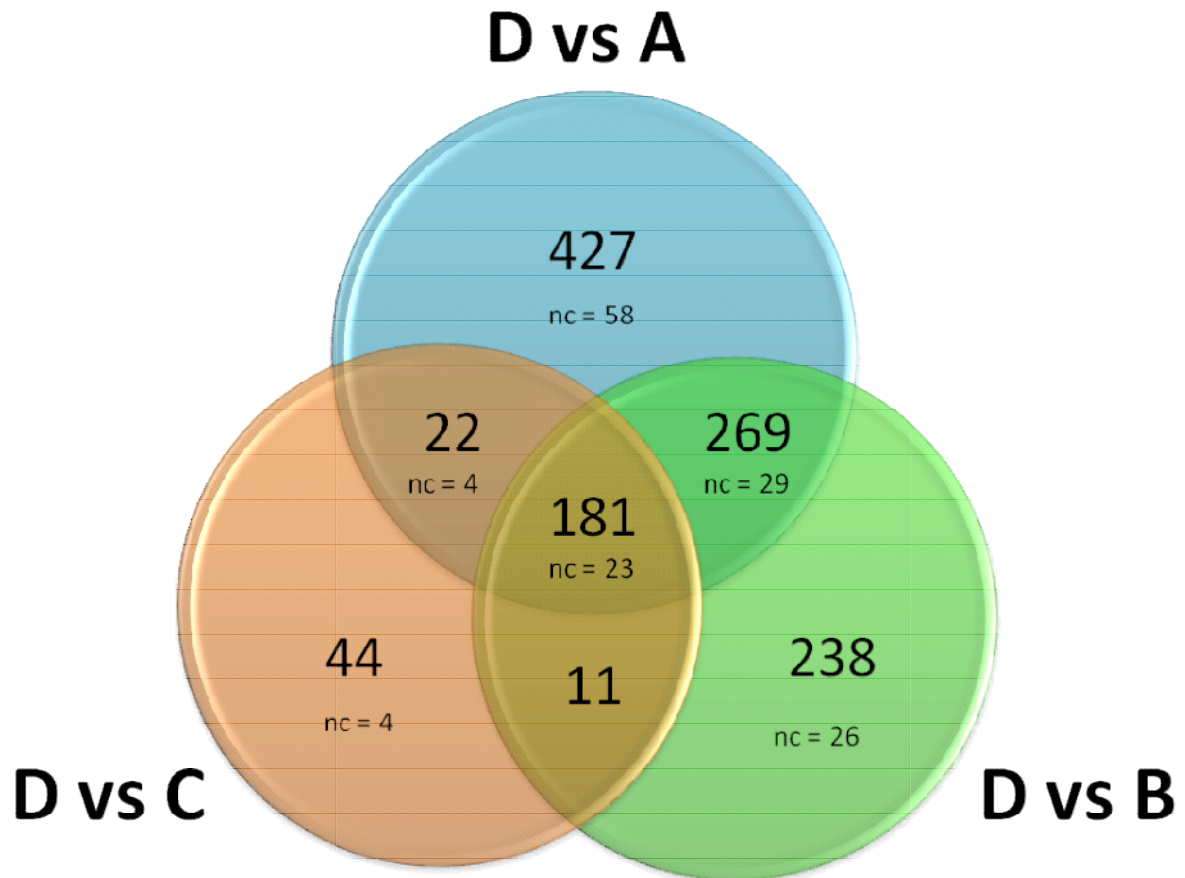
Figure 17 : Gènes significativement surexprimés dans la catégorie C par rapport aux catégories A, B et D (FC > 2 et p-adj < 0,05)



Un gène est communément surexprimé en C par rapport aux trois autres catégories (A, B et D) (*SLAMF1*). Deux gènes sont surexprimés en C par rapport à B et D (*KCNE4* et *RF00263*). Quatre gènes sont surexprimés en C par rapport à A et B (*WISP1*, *PAG11*, *HSPA1A* et *ORM1*). Les 98 autres gènes sont surexprimés en C de manière exclusive à une des comparaisons : 49 sont surexprimés en C par rapport à A ; 37 sont surexprimés en C par rapport à D et 12 sont surexprimés en C par rapport à B.

La Figure 18 illustre les gènes surexprimés en D par rapport aux trois autres catégories (D vs A ; D vs B ; D vs C).

Figure 18 : Gènes significativement surexprimés dans la catégorie D par rapport aux catégories A, B et C (FC > 2 et p-adj < 0,05)



La catégorie D présente un grand nombre de gènes surexprimés en comparaison aux autres catégories. 181 gènes sont communément surexprimés en D par rapport aux trois autres catégories (A, B et C). 269 gènes sont surexprimés en D par rapport à A et B. 22 gènes sont surexprimés en D par rapport à A et C. 11 gènes sont surexprimés en D par rapport à B et C. Les 709 autres gènes sont surexprimés en D de façon exclusive à une des comparaisons : 427 sont surexprimés en D par rapport à A ; 238 sont surexprimés en D par rapport à B et 44 sont surexprimés en D par rapport à C.

En résumé, la catégorie D est celle qui présente le plus grand nombre de gènes surexprimés par rapport aux trois autres catégories (A, B, C). De plus, la plus grande différence est observée entre les catégories A et D (qui sont également les catégories les plus divergentes en matière de trajectoire morphocinétique). Par la suite, nous nous focaliserons donc sur la comparaison A vs D afin de mieux étudier les différences existantes entre les deux catégories les plus divergentes au niveau du transcriptome, immédiatement après l'activation présumée du génome embryonnaire. Par ailleurs, nous incluons dans l'analyse uniquement les gènes actuellement connus.

3. Focus sur la comparaison A vs D

a. Gènes connus significativement surexprimés en A par rapport à la catégorie D

Une liste des 36 gènes connus sur les 60 surexprimés dans la catégorie A par rapport à la catégorie D mis en évidence dans la partie précédente a été dressée et est présentée dans le Tableau 4. La liste complète des gènes (connus et non connus) se trouve en Annexe 1.

Tableau 4 : Gènes connus surexprimés en A par rapport à D avec description de leur fonction et de leur *Fold-Change* (FC > 2 dans la comparaison A vs D ; p-adj < 0,05)

Nom du gène	Code Ensemble	Description fonction	<i>Fold-Change</i>
ADGRF5	ENSBTAG00000004347	Bos taurus adhesion G protein-coupled receptor F5 (ADGRF5), mRNA	3,1
AKR1C4	ENSBTAG00000022564	Bos taurus aldo-keto reductase family 1, member C4 (chlordecone reductase; 3-alpha hydroxysteroid dehydrogenase, type I; dihydrodiol dehydrogenase 4) (AKR1C4), mRNA	2,2
ANKFN1	ENSBTAG00000021292	Ankyrin repeat and fibronectin type III domain containing 1	2,1
BLNK	ENSBTAG00000021358	Bos taurus B cell linker (BLNK), mRNA	10,6
BRB	ENSBTAG00000004950	Bos taurus brain ribonuclease (BRB), mRNA	2,5
bta-mir-2887-2	ENSBTAG000000054661	Bta-mir-2887-1	2,3
C1QTNF3	ENSBTAG00000017071	Bos taurus C1q and TNF related 3 (C1QTNF3), mRNA	4,9
CR2	ENSBTAG00000038496	Bos taurus complement C3d receptor 2 (CR2), mRNA	73,8
CRISP2	ENSBTAG00000002166	Bos taurus cysteine rich secretory protein 2 (CRISP2), mRNA	2,1
CYP26A1	ENSBTAG00000021118	Bos taurus cytochrome P450, family 26, subfamily A, polypeptide 1 (CYP26A1), transcript variant 1, mRNA	2,7
EMX2	ENSBTAG00000003027	Bos taurus empty spiracles homeobox 2 (EMX2), mRNA	2,1
F2RL2	ENSBTAG00000000898	Bos taurus coagulation factor II thrombin receptor like 2 (F2RL2), mRNA	3,4
FBN2	ENSBTAG00000015307	Bos taurus fibrillin 2 (FBN2), mRNA	2,6
FCER1A	ENSBTAG00000012887	Bos taurus Fc fragment of IgE receptor Ia (FCER1A), mRNA	128,5

Nom du gène	Code Ensemble	Description fonction	Fold-Change
FKBP11	ENSBTAG00000015828	FK506 binding protein 11	2,3
FXVD4	ENSBTAG000000051295	FXVD domain containing ion transport regulator 4	2,1
HPD	ENSBTAG000000004175	4-hydroxyphenylpyruvate dioxygenase	2,9
KCNE4	ENSBTAG000000006432	Bos taurus potassium voltage-gated channel subfamily E regulatory subunit 4 (KCNE4), mRNA	2884306,1
KLRA1	ENSBTAG000000009252	Bos taurus killer cell lectin-like receptor subfamily A, member 1 (KLRA1), mRNA	2,0
KRTDAP	ENSBTAG000000005330	Bos taurus keratinocyte differentiation associated protein (KRTDAP), transcript variant 1, mRNA	4,1
LUZP2	ENSBTAG000000001043	Bos taurus leucine zipper protein 2 (LUZP2), mRNA	2,7
MCMDC2	ENSBTAG000000000620	Minichromosome maintenance domain containing 2	8,8
MMP9	ENSBTAG000000020676	Bos taurus matrix metalloproteinase 9 (MMP9), mRNA	2,2
MT1E	ENSBTAG000000038706	Metallothionein 1E	2,4
NADK2	ENSBTAG000000011334	Bos taurus NAD kinase 2, mitochondrial (NADK2), mRNA	2,1
OR5R1	ENSBTAG000000045739	Olfactory receptor 1038	83,2
PLA2G10	ENSBTAG000000021522	Phospholipase A2 group X	6,0
RF00090	ENSBTAG000000042256	Small nucleolar RNA, H/ACA box 74A	2,4
RF00425	ENSBTAG000000043258	Inconnue	2,5
RF00553	ENSBTAG000000044441	Small Cajal body-specific RNA 1	3,2
RF00586	ENSBTAG000000043222	Small nucleolar RNA, H/ACA box 12	3,2
RF01241	ENSBTAG000000044490	Inconnue	2,3
S100A10	ENSBTAG000000015147	Bos taurus S100 calcium binding protein A10 (S100A10), mRNA	2,0
SASH3	ENSBTAG000000002041	SAM and SH3 domain containing 3	2,2
TPST2	ENSBTAG000000010368	Bos taurus tyrosylprotein sulfotransferase 2 (TPST2), mRNA	2,3
TRIM63	ENSBTAG000000005085	Bos taurus tripartite motif containing 63 (TRIM63), mRNA	2,1

b. Gènes connus significativement surexprimés en D par rapport à la catégorie A

Une liste des gènes connus, surexprimés dans la catégorie D par rapport à la catégorie A mis en évidence dans la partie précédente a été dressée et est présentée dans le Tableau 5.

Pour un souci de lisibilité, le Tableau 5 contient uniquement les gènes identifiés comme participant à l'enrichissement différentiel d'au moins une voie fonctionnelle dans la catégorie D par rapport à la catégorie A et qui sera discuté par la suite. Ainsi, seuls 192 gènes apparaissent ci-dessous sur les 785 gènes connus et surexprimés dans la catégorie D par rapport à la catégorie A.

La liste complète des 899 gènes (connus et non connus) surexprimés dans la catégorie D se trouve en Annexe 4.

Tableau 5 : Gènes connus surexprimés en D par rapport à A et impliqués dans des voies fonctionnelles différentiellement enrichies avec description de leur fonction et de leur *Fold-Change* (FC < 0,5 dans la comparaison A vs D ; p-adj < 0,05)

Nom du gène	Code Ensemble	Description fonction	<i>Fold-Change</i>
ABL1	ENSBTAG00000017976	Bos taurus ABL proto-oncogene 1, non-receptor tyrosine kinase (ABL1), mRNA	0,30
ADARB1	ENSBTAG00000017486	Adenosine deaminase, RNA specific B1	0,35
ADGRA2	ENSBTAG00000008814	Adhesion G protein-coupled receptor A2	0,29
ADRB2	ENSBTAG00000002144	Bos taurus adrenoceptor beta 2 (ADRB2), mRNA	0,39
AHR	ENSBTAG00000007746	Bos taurus aryl hydrocarbon receptor (AHR), mRNA	0,24
ARC	ENSBTAG00000021639	Bos taurus activity regulated cytoskeleton associated protein (ARC), mRNA	0,01
ATXN1	ENSBTAG00000019675	Ataxin 1	0,15
BARHL2	ENSBTAG00000005938	BarH like homeobox 2	0,03
BARX2	ENSBTAG00000020643	BARX homeobox 2	0,14
BCL3	ENSBTAG00000009493	B cell CLL/lymphoma 3	0,26
BCL7A	ENSBTAG00000032517	BCL tumor suppressor 7A	0,06
BICC1	ENSBTAG00000010694	BicC family RNA binding protein 1	0,14
BMP15	ENSBTAG00000045782	Bos taurus bone morphogenetic protein 15 (BMP15), mRNA	0,07
BMP6	ENSBTAG00000019234	Bone morphogenetic protein 6	0,40
BMPR1B	ENSBTAG00000002081	Bos taurus bone morphogenetic protein receptor type 1B (BMPR1B), mRNA	0,27
BNC2	ENSBTAG00000003669	Basonuclin 2	0,27
BRD3	ENSBTAG00000010582	Bos taurus bromodomain containing 3 (BRD3), mRNA	0,37
CALCA	ENSBTAG00000010089	Calcitonin-related polypeptide alpha	0,01

Nom du gène	Code Ensemble	Description fonction	Fold-Change
CAND2	ENSBTAG00000015273	Cullin associated and neddylation dissociated 2 (putative)	0,18
CASR	ENSBTAG00000003865	Bos taurus calcium sensing receptor (CASR), mRNA	0,33
CCNQ	ENSBTAG00000011350	Cyclin Q	0,16
CDK5R2	ENSBTAG000000031052	Cyclin dependent kinase 5 regulatory subunit 2	0,15
CDKN1A	ENSBTAG00000008353	Bos taurus cyclin dependent kinase inhibitor 1A (CDKN1A), mRNA	0,28
CIART	ENSBTAG00000021487	Circadian associated repressor of transcription	0,18
CLIP3	ENSBTAG00000019674	CAP-Gly domain containing linker protein 3	0,22
CPS1	ENSBTAG00000016662	Bos taurus carbamoyl-phosphate synthase 1 (CPS1), mRNA	0,03
CREB5	ENSBTAG00000018909	Bos taurus cAMP responsive element binding protein 5 (CREB5), mRNA	0,32
CXCR4	ENSBTAG00000001060	Bos taurus C-X-C motif chemokine receptor 4 (CXCR4), mRNA	0,42
CYFIP2	ENSBTAG00000008916	Cytoplasmic FMR1 interacting protein 2	0,18
DAB1	ENSBTAG00000013802	DAB1, reelin adaptor protein	0,25
DAB2IP	ENSBTAG00000008232	DAB2 interacting protein	0,24
DACT1	ENSBTAG00000019421	Dishevelled binding antagonist of beta catenin 1	0,07
DBX1	ENSBTAG00000002733	Bos taurus developing brain homeobox 1 (DBX1), mRNA	0,12
DLX6	ENSBTAG00000021458	Distal-less homeobox 6	0,39
DMTF1	ENSBTAG00000020878	Cyclin D binding myb like transcription factor 1	0,43
DYRK1B	ENSBTAG00000012509	Dual specificity tyrosine phosphorylation regulated kinase 1B	0,26
EBF1	ENSBTAG00000037508	EBF transcription factor 1	0,10
EGR4	ENSBTAG00000024058	Early growth response 4	0,03
ELF3	ENSBTAG00000008756	E74 like ETS transcription factor 3	0,18
ELK1	ENSBTAG00000015817	Bos taurus ELK1, ETS transcription factor (ELK1), mRNA	0,49
ELK3	ENSBTAG00000001509	Bos taurus ELK3, ETS transcription factor (ELK3), mRNA	0,26
ELN	ENSBTAG00000019517	Bos taurus elastin (ELN), mRNA	0,30
ELOVL4	ENSBTAG00000015498	Bos taurus ELOVL fatty acid elongase 4 (ELOVL4), mRNA	0,27
ELP1	ENSBTAG00000004991	Elongator complex protein 1	0,49
EOMES	ENSBTAG00000016920	Eomesodermin	0,05

Nom du gène	Code Ensemble	Description fonction	Fold-Change
ESR2	ENSBTAG00000004498	Estrogen receptor 2	0,37
ETV1	ENSBTAG000000015981	Bos taurus ETS variant 1 (ETV1), mRNA	0,29
ETV5	ENSBTAG000000014915	Bos taurus ETS variant 5 (ETV5), mRNA	0,22
ETV6	ENSBTAG000000014605	ETS variant 6	0,31
F3	ENSBTAG000000007101	Bos taurus coagulation factor III, tissue factor (F3), mRNA	0,37
FANK1	ENSBTAG000000022917	Fibronectin type III and ankyrin repeat domains 1	0,29
FER	ENSBTAG000000003051	Bos taurus FER tyrosine kinase (FER), mRNA	0,26
FOSL1	ENSBTAG000000006194	Bos taurus FOS like 1, AP-1 transcription factor subunit (FOSL1), mRNA	0,29
FOSL2	ENSBTAG000000023929	Bos taurus FOS like 2, AP-1 transcription factor subunit (FOSL2), mRNA	0,33
FOXI3	ENSBTAG000000040093	Forkhead box I3	0,49
FOXJ3	ENSBTAG000000015887	Forkhead box J3	0,43
FOXN3	ENSBTAG000000046684	Bos taurus forkhead box N3 (FOXN3), mRNA	0,47
FOXO1	ENSBTAG000000044105	Forkhead box O1	0,31
FOXO3	ENSBTAG000000011234	Bos taurus forkhead box O3 (FOXO3), mRNA	0,27
FOXP3	ENSBTAG000000013279	Forkhead box P3	0,18
GAP43	ENSBTAG000000006451	Growth associated protein 43	0,01
GATA2	ENSBTAG000000019707	Bos taurus GATA binding protein 2 (GATA2), mRNA	0,34
GBX2	ENSBTAG000000008898	Bos taurus gastrulation brain homeobox 2 (GBX2), mRNA	0,40
GDF9	ENSBTAG000000009478	Bos taurus growth differentiation factor 9 (GDF9), mRNA	0,08
GPC3	ENSBTAG000000020406	Glypican 3	0,18
GRHL3	ENSBTAG000000019791	Grainyhead like transcription factor 3	0,26
GRIP1	ENSBTAG000000033726	Glutamate receptor interacting protein 1	0,16
HAVCR2	ENSBTAG000000000495	Bos taurus hepatitis A virus cellular receptor 2 (HAVCR2), mRNA	0,28
HCFC1	ENSBTAG000000011904	Host cell factor C1	0,48
HDAC10	ENSBTAG000000011000	Bos taurus histone deacetylase 10 (HDAC10), mRNA	0,32
HES7	ENSBTAG000000012436	Bos taurus hes family bHLH transcription factor 7 (HES7), mRNA	0,18

Nom du gène	Code Ensemble	Description fonction	Fold-Change
HEY2	ENSBTAG00000000919	Bos taurus hes related family bHLH transcription factor with YRPW motif 2 (HEY2), mRNA	0,50
HGF	ENSBTAG00000017664	Bos taurus hepatocyte growth factor (HGF), mRNA	0,01
HHAT	ENSBTAG00000048628	Hedgehog acyltransferase	0,28
HLF	ENSBTAG00000006618	Bos taurus HLF, PAR bZIP transcription factor (HLF), mRNA	0,17
HNF1B	ENSBTAG00000004710	HNF1 homeobox B	0,17
HOMEZ	ENSBTAG000000032686	Homeobox and leucine zipper encoding	0,21
HOXD13	ENSBTAG00000004313	Homeobox D13	0,29
HSPB11	ENSBTAG00000007112	Bos taurus heat shock protein family B (small) member 11 (HSPB11), mRNA	0,49
ID3	ENSBTAG000000030425	Bos taurus inhibitor of DNA binding 3, HLH protein (ID3), mRNA	0,12
IGF1R	ENSBTAG000000021527	Bos taurus insulin like growth factor 1 receptor (IGF1R), mRNA	0,36
IKZF2	ENSBTAG000000037804	IKAROS family zinc finger 2	0,20
INSM1	ENSBTAG000000045526	INSM transcriptional repressor 1	0,25
IRF2	ENSBTAG000000010002	Bos taurus interferon regulatory factor 2 (IRF2), mRNA	0,16
IRF2BPL	ENSBTAG000000025329	Interferon regulatory factor 2 binding protein like	0,19
IRF4	ENSBTAG000000002929	Interferon regulatory factor 4	0,12
IRF8	ENSBTAG000000017824	Interferon regulatory factor 8	0,27
JAZF1	ENSBTAG000000019024	Bos taurus JAZF zinc finger 1 (JAZF1), mRNA	0,17
JPH1	ENSBTAG000000008842	Bos taurus junctophilin 1 (JPH1), mRNA	0,35
KCTD13	ENSBTAG000000015982	Bos taurus potassium channel tetramerization domain containing 13 (KCTD13), mRNA	0,39
KIAA1958	ENSBTAG000000016259	KIAA1958	0,29
KLF11	ENSBTAG000000046218	Bos taurus Kruppel like factor 11 (KLF11), mRNA	0,49
KLF7	ENSBTAG000000044097	Kruppel like factor 7	0,38
LEF1	ENSBTAG000000006844	Lymphoid enhancer binding factor 1	0,12
LHX4	ENSBTAG000000003992	Bos taurus LIM homeobox 4 (LHX4), mRNA	0,04
LPL	ENSBTAG000000012855	Bos taurus lipoprotein lipase (LPL), mRNA	0,23
LRP6	ENSBTAG000000011761	LDL receptor related protein 6	0,42

Nom du gène	Code Ensemble	Description fonction	Fold-Change
MICAL2	ENSBTAG00000014855	Microtubule associated monooxygenase, calponin and LIM domain containing 2	0,48
MSX1	ENSBTAG00000010875	Bos taurus msh homeobox 1 (MSX1), mRNA	0,17
MXI1	ENSBTAG00000017263	Bos taurus MAX interactor 1, dimerization protein (MXI1), mRNA	0,32
MYB	ENSBTAG00000012074	Bos taurus MYB proto-oncogene, transcription factor (MYB), mRNA	0,42
MYF5	ENSBTAG00000026972	Bos taurus myogenic factor 5 (MYF5), mRNA	0,07
MYH10	ENSBTAG00000021151	Myosin heavy chain 10	0,49
NALCN	ENSBTAG00000037786	Sodium leak channel, non-selective	0,01
NBEA	ENSBTAG00000022991	Neurobeachin	0,22
NEDD4L	ENSBTAG00000005412	Neural precursor cell expressed, developmentally down-regulated 4-like, E3 ubiquitin protein ligase	0,31
NFATC1	ENSBTAG00000000656	Nuclear factor of activated T cells 1	0,35
NFIA	ENSBTAG00000000074	Bos taurus nuclear factor I A (NFIA), mRNA	0,10
NLRP14	ENSBTAG00000002289	NLR family pyrin domain containing 14	0,16
NLRP5	ENSBTAG00000013247	Bos taurus NLR family pyrin domain containing 5 (NLRP5), mRNA	0,29
NLRP8	ENSBTAG00000006965	NLR family pyrin domain containing 8	0,25
NLRP9	ENSBTAG00000004725	Bos taurus NLR family pyrin domain containing 9 (NLRP9), mRNA	0,20
NOBOX	ENSBTAG00000009832	NOBOX oogenesis homeobox	0,13
NOD1	ENSBTAG00000038235	Bos taurus nucleotide binding oligomerization domain containing 1 (NOD1), mRNA	0,27
NPAS3	ENSBTAG00000004462	Neuronal PAS domain protein 3	0,28
NR1H3	ENSBTAG00000010681	Bos taurus nuclear receptor subfamily 1 group H member 3 (NR1H3), mRNA	0,46
NR3C2	ENSBTAG00000027182	Bos taurus nuclear receptor subfamily 3 group C member 2 (NR3C2), mRNA	0,28
NRXN1	ENSBTAG00000024021	Neurexin 1	0,20
NRXN3	ENSBTAG00000050101	Neurexin 3	0,10
ONECUT2	ENSBTAG00000022819	One cut homeobox 2	0,16
PAX3	ENSBTAG00000009950	Bos taurus paired box 3 (PAX3), mRNA	0,24
PAX6	ENSBTAG00000004561	Bos taurus paired box 6 (PAX6), mRNA	0,36

Nom du gène	Code Ensemble	Description fonction	Fold-Change
PER1	ENSBTAG00000003889	Bos taurus period circadian regulator 1 (PER1), mRNA	0,21
PER2	ENSBTAG00000001879	Period circadian regulator 2	0,20
PFN2	ENSBTAG00000001150	Bos taurus profilin 2 (PFN2), mRNA	0,47
PHOX2A	ENSBTAG000000019168	Bos taurus paired like homeobox 2A (PHOX2A), mRNA	0,43
PODXL	ENSBTAG000000010452	Podocalyxin like	0,38
POU4F3	ENSBTAG000000018817	Bos taurus POU class 4 homeobox 3 (POU4F3), mRNA	0,01
PPARGC1A	ENSBTAG000000017024	Bos taurus PPARG coactivator 1 alpha (PPARGC1A), mRNA	0,20
PPEF2	ENSBTAG000000015449	Bos taurus protein phosphatase with EF-hand domain 2 (PPEF2), mRNA	0,07
PPP1R9A	ENSBTAG000000024426	Protein phosphatase 1 regulatory subunit 9A	0,27
PRAME	ENSBTAG000000033890	Preferentially expressed antigen in melanoma	0,34
PRDM1	ENSBTAG000000000816	Bos taurus PR/SET domain 1 (PRDM1), mRNA	0,42
PRKG1	ENSBTAG000000018404	Protein kinase cGMP-dependent 1	0,40
RALB	ENSBTAG000000032021	Bos taurus RAS like proto-oncogene B (RALB), mRNA	0,31
RARA	ENSBTAG000000012500	Bos taurus retinoic acid receptor alpha (RARA), mRNA	0,39
RASL11A	ENSBTAG000000005244	RAS like family 11 member A	0,40
REM2	ENSBTAG000000016381	RRAD and GEM like GTPase 2	0,24
RORA	ENSBTAG000000015904	RAR related orphan receptor A	0,43
SALL2	ENSBTAG000000013249	Spalt like transcription factor 2	0,38
SBNO2	ENSBTAG000000020780	Bos taurus strawberry notch homolog 2 (SBNO2), mRNA	0,38
SCN8A	ENSBTAG000000021339	Bos taurus sodium voltage-gated channel alpha subunit 8 (SCN8A), mRNA	0,28
SFN	ENSBTAG000000009223	Bos taurus stratifin (SFN), mRNA	0,07
SHANK2	ENSBTAG000000003171	SH3 and multiple ankyrin repeat domains 2	0,23
SHOX2	ENSBTAG000000005230	Bos taurus short stature homeobox 2 (SHOX2), mRNA	0,11
SIX1	ENSBTAG000000012139	SIX homeobox 1	0,08
SIX3	ENSBTAG000000027017	Bos taurus SIX homeobox 3 (SIX3), mRNA	0,06
SLC12A1	ENSBTAG000000010145	Solute carrier family 12 member 1	0,01
SLC12A5	ENSBTAG000000014179	Bos taurus solute carrier family 12 member 5 (SLC12A5), mRNA	0,18
SLC6A5	ENSBTAG000000014300	Solute carrier family 6 member 5	0,08

Nom du gène	Code Ensemble	Description fonction	Fold-Change
SLC7A1	ENSBTAG00000018577	Solute carrier family 7 member 1	0,33
SLCO2A1	ENSBTAG00000001324	Solute carrier organic anion transporter family member 2A1	0,03
SMAD5	ENSBTAG000000055062	Bos taurus SMAD family member 5 (SMAD5), mRNA	0,46
SMARCA2	ENSBTAG00000007494	Bos taurus SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2 (SMARCA2), mRNA	0,28
SMOC1	ENSBTAG000000030599	SPARC related modular calcium binding 1	0,44
SMURF1	ENSBTAG00000007118	SMAD specific E3 ubiquitin protein ligase 1	0,50
SNAP91	ENSBTAG000000009355	Synaptosome associated protein 91	0,40
SNCA	ENSBTAG000000024957	Bos taurus synuclein alpha (SNCA), mRNA	0,34
SOCS7	ENSBTAG000000008535	Suppressor of cytokine signaling 7	0,25
SOX5	ENSBTAG000000022360	Bos taurus SRY-box 5 (SOX5), mRNA	0,08
SPAG17	ENSBTAG000000001080	Bos taurus sperm associated antigen 17 (SPAG17), mRNA	0,19
SPI1	ENSBTAG000000021709	Spi-1 proto-oncogene	0,19
SPOP	ENSBTAG000000005824	Bos taurus speckle type BTB/POZ protein (SPOP), mRNA	0,43
STK11	ENSBTAG000000011639	Serine/threonine kinase 11	0,29
SULF2	ENSBTAG000000007490	Bos taurus sulfatase 2 (SULF2), mRNA	0,21
TACC3	ENSBTAG000000011044	Bos taurus transforming acidic coiled-coil containing protein 3 (TACC3), mRNA	0,28
TAF9B	ENSBTAG000000000895	Bos taurus TATA-box binding protein associated factor 9b (TAF9B), mRNA	0,36
TBX15	ENSBTAG000000007767	T-box 15	0,22
TBX18	ENSBTAG000000018161	T-box 18	0,28
TCF4	ENSBTAG000000016462	Bos taurus transcription factor 4 (TCF4), mRNA	0,14
TCF7	ENSBTAG000000001002	Bos taurus transcription factor 7 (TCF7), mRNA	0,20
TCL1A	ENSBTAG000000019580	T cell leukemia/lymphoma 1A	0,10
TENM2	ENSBTAG000000025071	Teneurin transmembrane protein 2	0,30
TET3	ENSBTAG000000022381	Tet methylcytosine dioxygenase 3	0,10
TFAP2B	ENSBTAG000000020426	Bos taurus transcription factor AP-2 beta (TFAP2B), mRNA	0,12
THRA	ENSBTAG000000010745	Bos taurus thyroid hormone receptor alpha (THRA), mRNA	0,09
TNFAIP3	ENSBTAG000000000436	TNF alpha induced protein 3	0,35

Nom du gène	Code Ensemble	Description fonction	Fold-Change
TOP2B	ENSBTAG00000004593	DNA topoisomerase II beta	0,28
TOX3	ENSBTAG00000000418	TOX high mobility group box family member 3	0,08
TRIB1	ENSBTAG00000023179	Bos taurus tribbles pseudokinase 1 (TRIB1), mRNA	0,40
TRIM44	ENSBTAG00000037389	Tripartite motif containing 44	0,28
TUBB3	ENSBTAG00000023730	Bos taurus tubulin beta 3 class III (TUBB3), mRNA	0,31
UBE2D1	ENSBTAG00000020796	Bos taurus ubiquitin conjugating enzyme E2 D1 (UBE2D1), mRNA	0,33
UHRF1	ENSBTAG00000002224	Bos taurus ubiquitin like with PHD and ring finger domains 1 (UHRF1), mRNA	0,27
UNC79	ENSBTAG00000008017	Unc-79 homolog, NALCN channel complex subunit	0,02
USP2	ENSBTAG00000009749	Bos taurus ubiquitin specific peptidase 2 (USP2), mRNA	0,25
VCAN	ENSBTAG00000014906	Bos taurus versican (VCAN), mRNA	0,39
ZBTB16	ENSBTAG00000011266	Bos taurus zinc finger and BTB domain containing 16 (ZBTB16), mRNA	0,09
ZHX2	ENSBTAG00000026309	Bos taurus zinc fingers and homeoboxes 2 (ZHX2), mRNA	0,07
ZNF398	ENSBTAG00000020445	Bos taurus zinc finger protein 398 (ZNF398), mRNA	0,20
ZNF462	ENSBTAG00000019187	Zinc finger protein 462	0,40
ZXDC	ENSBTAG00000009874	ZXD family zinc finger C	0,49

c. Profil des gènes différentiellement exprimés dans la comparaison A vs D

Cette étude a été réalisée sur des embryons au stade 16-32 cellules, soit, théoriquement, juste après l'activation de leur génome embryonnaire.

Les transcrits d'origine maternelle sont donc en diminution alors que les transcrits d'origine embryonnaire sont en augmentation.

Afin de connaître le profil des gènes différentiellement exprimés dans la comparaison A vs D, nous avons utilisé la base de données EmbryoExpress qui recense les données du transcriptome d'ovocytes et embryons bovins collectés *in vivo* aux stades suivants : ovocytes GV et MII, embryons aux stades quatre cellules, huit cellules, seize cellules et blastocyste.

Le profil *in vivo* de 40 des 60 gènes surexprimés en A par rapport à D était connu et a pu être établi à partir des données du logiciel EmbryoExpress.

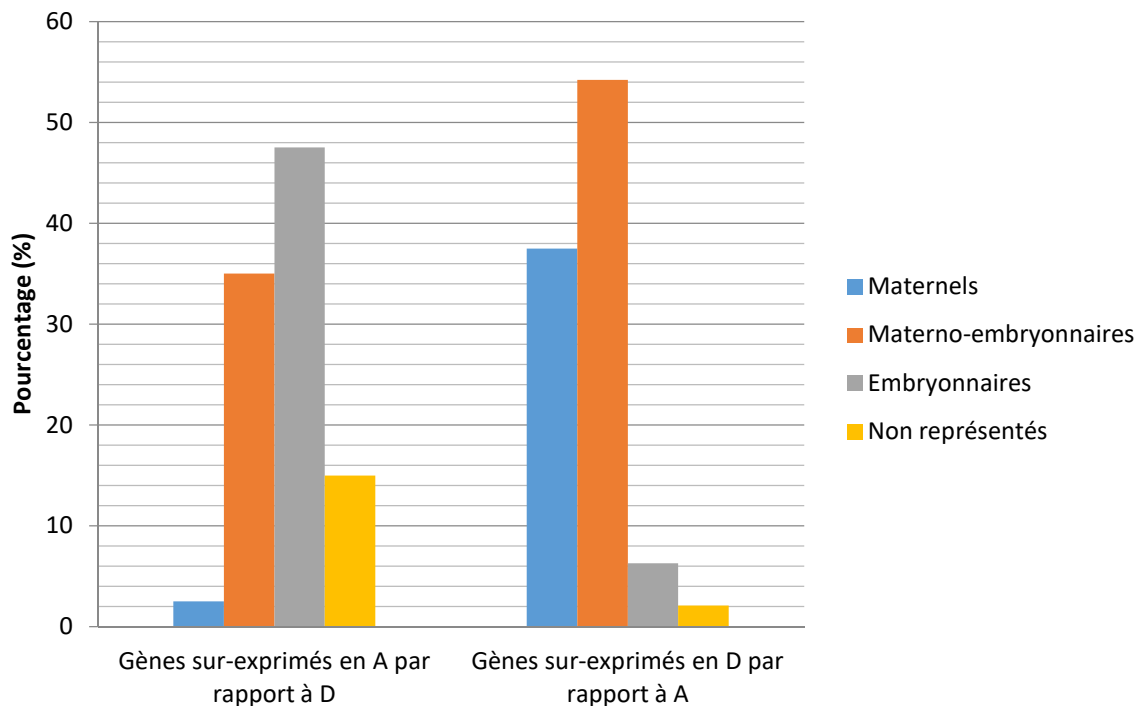
De même, sur un total de 899 gènes surexprimés en D par rapport à A, le profil *in vivo* de 816 d'entre eux était connu et a pu être étudié.

Les résultats de notre analyse sont présentés dans le Tableau 6 et la Figure 19.

Tableau 6 : Proportion de gènes différentiellement exprimés dans la comparaison A vs D en fonction de leur profil dans l'embryon bovin produit *in vivo*.

Comparaisons		Profils des gènes chez l'embryon <i>in vivo</i>				Total
		Maternels	Materno-embryonnaires	Embryonnaires	Non classés	
Gènes surexprimés en A par rapport à D	Nombre de gènes	1	14	19	6	40
	Pourcentage	2,5%	35,0%	47,5%	15,0%	100%
Gènes surexprimés en D par rapport à A	Nombre de gènes	306	442	51	17	816
	Pourcentage	37,5%	54,2%	6,3%	2,1%	100%

Figure 19 : Pourcentage de gènes différentiellement exprimés dans la comparaison A vs D en fonction de leur profil dans l'embryon bovin produit *in vivo* (déterminé à partir des données obtenues dans EmbryoExpress).



Un contraste est visible dans le profil des gènes surexprimés dans l'une et l'autre des deux catégories.

En effet, les gènes surexprimés dans la catégorie A ont un profil majoritairement embryonnaire et materno-embryonnaire, et seulement une infime partie (2,5%) a un profil maternel.

A contrario, parmi les gènes surexprimés dans la catégorie D, la majorité a un profil maternel et materno-embryonnaire et seule une minorité a un profil embryonnaire (6,3%).

Il existe donc un retard quant à la diminution en abondance des transcrits maternels chez les embryons de la catégorie D par rapport à la catégorie A.

Par ailleurs, il est important de souligner la présence de 23 gènes (six dans la catégorie A et 17 dans la catégorie D) présents dans les embryons de notre étude et non présents dans les embryons collectés *in vivo* de la base de données de référence.

d. Voies fonctionnelles significativement enrichies par les gènes différentiellement exprimés dans la comparaison A vs D

Une analyse d'enrichissement à partir des gènes différentiellement exprimés dans la comparaison A vs D a été réalisée sur la plateforme Enrichr.

Aucune voie fonctionnelle n'est significativement enrichie par les gènes surexprimés dans la catégorie A.

A l'inverse, on observe 13 voies significativement enrichies par les gènes surexprimés dans la catégorie D.

Ces voies fonctionnelles appartiennent à trois bibliothèques différentes : Processus biologiques (*Gene Ontology GO*), Fonctions moléculaires (*Gene Ontology GO*) et Phénotypes mammifères (*Mouse Genome Informatics MGI*).

Celles-ci sont présentées dans le Tableau 7.

Tableau 7 : Voies fonctionnelles significativement enrichies par les gènes surexprimés dans la catégorie D par rapport à la catégorie A.

Nom de la voie	Catégorie fonctionnelle	Valeur p ajustée	Nombre de gènes étudiés impliqués dans cette voie	Fonction
Positive regulation of transcription, DNA-templated (GO:0045893)	GO Processus biologiques	0,001416	79	Régulation positive de la machinerie de la polymerase II
Positive regulation of transcription from RNA polymerase II promoter (GO:0045944)	GO Processus biologiques	0,002199	63	Régulation positive de la machinerie de la polymerase II
Regulation of transcription from RNA polymerase II promoter (GO:0006357)	GO Processus biologiques	0,001891	95	Régulation positive de la machinerie de la polymerase II
Positive regulation of nucleic acid-templated transcription (GO:1903508)	GO Processus biologiques	0,02372	40	Régulation positive de la machinerie de la polymerase II
Negative regulation of transcription, DNA-templated (GO:0045892)	GO Processus biologiques	0,03229	56	Régulation négative de la machinerie de la polymerase II
Transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific binding (GO:0001228)	GO Fonctions moléculaires	0,003049	29	Régulation positive de la machinerie de la polymerase II
RNA polymerase II regulatory region sequence-specific DNA binding (GO:0000977)	GO Fonctions moléculaires	0,003261	39	Régulation de la machinerie de la polymerase II

Nom de la voie	Catégorie fonctionnelle	Valeur p ajustée	Nombre de gènes étudiés impliqués dans cette voie	Fonction
Transcription regulatory region DNA binding (GO:0044212)	GO Fonctions moléculaires	0,02966	31	Régulation de la transcription de l'ADN
Regulation of protein serine/threonine kinase activity (GO:0071900)	GO Processus biologiques	0,04591	12	Modulation du taux, de la fréquence ou de l'importance de l'activité de la kinase sérine/thréonine
Negative regulation of immune response (GO:0050777)	GO Processus biologiques	0,03664	9	Processus bloquant ou réduisant la fréquence, le taux ou l'ampleur de la réponse du système immunitaire à un stimulus immunogène
Neonatal lethality, complete penetrance (MP:0011087)	MGI Phénotypes mammifères	0,001108	46	Décès de tous les organismes d'un génotype donné dans une population au cours de la période néonatale
Postnatal lethality, complete penetrance (MP:0011085)	MGI Phénotypes mammifères	0,02053	37	Décès prématuré à tout moment entre la période néonatale et l'âge de sevrage de tous les organismes d'un génotype donné dans une population
Abnormal miniature excitatory postsynaptic currents (MP:0004753)	MGI Phénotypes mammifères	0,04862	14	Défaut dans la taille ou la durée des courants spontanés détectés dans les cellules post-synaptiques qui se produisent en l'absence d'impulsion excitatrice

On observe que les voies fonctionnelles enrichies dans la catégorie D par rapport à la catégorie A sont principalement annotées dans les bases de données « Processus biologiques » et « Fonctions moléculaires ».

La grande majorité d'entre elles sont des **voies fonctionnelles liées à la transcription** (huit voies sur les 14 mises en évidence), et parmi celles-ci, sept voies sont impliquées dans la machinerie de la polymérase II.

On retrouve également une voie impliquée dans **la régulation négative du système immunitaire** et enfin une voie impliquée dans **la régulation de l'activité de la protéine kinase sérine/thréonine**.

Trois voies sont, elles, retrouvées dans la base de données « Phénotypes mammifères». Deux d'entre elles sont impliquées dans la létalité précoce de l'organisme et une dans des anomalies de courant dans les cellules post-synaptiques.

Pour avoir un aperçu des liens que peuvent avoir ces voies fonctionnelles entre elles et de leur implication au sein de plus grands ensembles de voies aboutissant à des fonctions précises, celles-ci ont été recherchées dans la plateforme QuickGO.

Ce travail n'a pu être réalisé qu'avec les voies GO car il n'existe aucune plateforme similaire pour les voies MGI.

Une vue d'ensemble du résultat est présentée dans la Figure 20.

Les voies ont été entourées avec la même couleur que celles qui les représentent dans le Tableau 7.

QuickGo a attribué un code couleur aux flèches reliant les voies afin d'illustrer leur impact les unes sur les autres.

La liaison d'une voie X à une voie Y par une flèche noire signifie que la voie X est un sous-type de la voie Y. Par exemple, la voie « transcription » est un sous-type de la voie « processus métabolique des acides nucléiques ».

La liaison d'une voie X à une voie Y par une flèche bleue signifie que la voie X est une partie de la voie Y. Par exemple, la voie « transcription » fait partie de la voie « expression des gènes ».

La liaison d'une voie X à une voie Y par une flèche jaune signifie que la voie X régule la voie Y.

Si la flèche entre les deux voies est verte, cela signifie que la voie X régule positivement la voie Y.

Inversément, si cette flèche est rouge, la voie X régule négativement la voie Y.

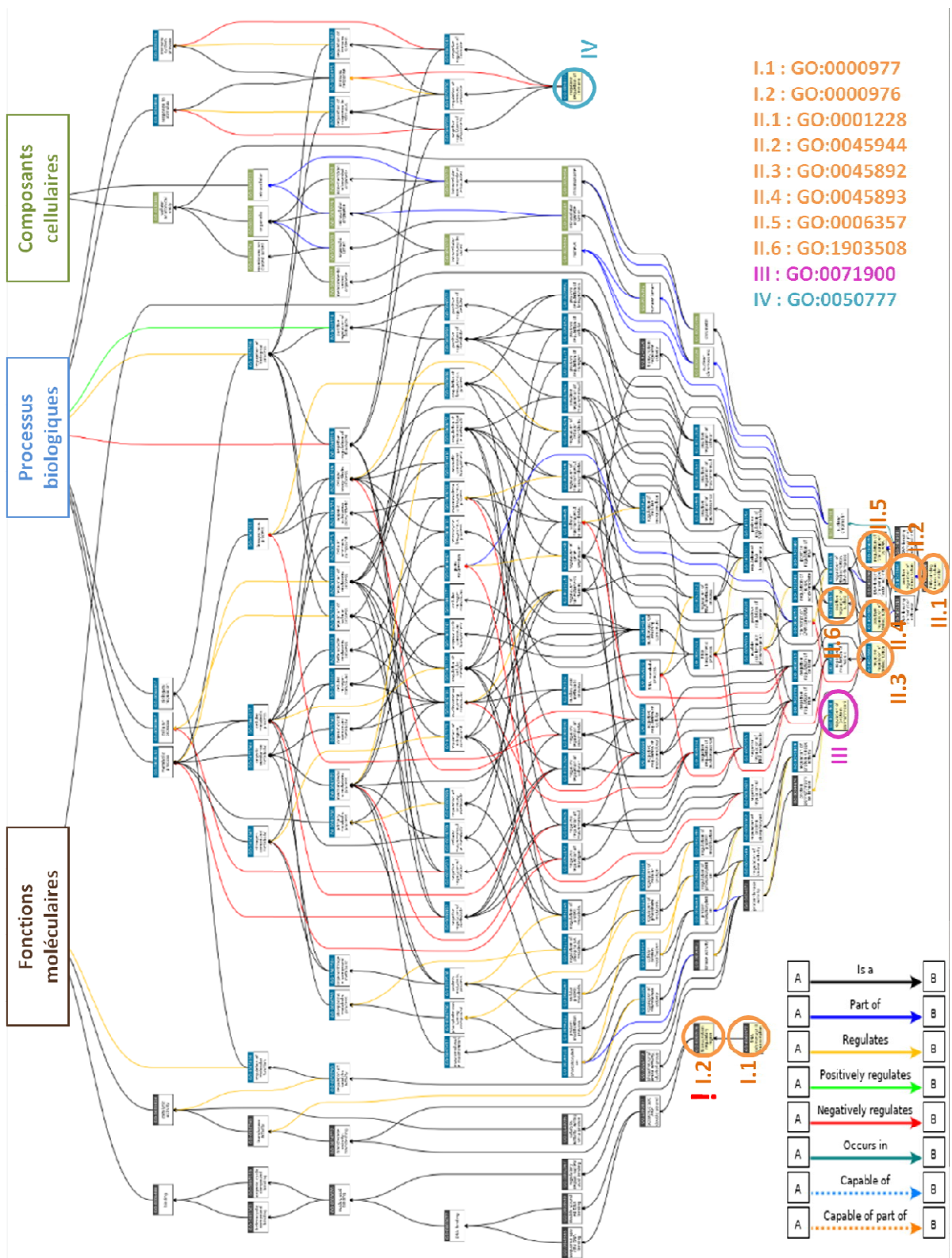
En ce qui concerne les trois dernières couleurs de la légende, représentant les liens « Occurs in », « Capable of » et « Capable of part of », la plateforme QuickGo n'a pas encore fourni les informations relatives à leur signification.

Par ailleurs, depuis une mise à jour de la plateforme QuickGo (réalisée le 27 avril 2020), la voie **GO:0044212** est fusionnée avec la voie **GO:0000976** (*Transcription regulatory region sequence-specific DNA binding*).

Elle apparaît donc sur les Figures 20 et 21 sous ce nouveau nom. Elle sera entourée de la même couleur que la voie d'origine et annotée avec un point d'exclamation rouge.

A ce jour, Enrichr n'a pas effectué cette mise à jour. Il sera donc nécessaire à l'avenir de vérifier si cette voie reste significativement enrichie par les gènes surexprimés en D par rapport à A dans notre étude.

Figure 20 : Visualisation des voies significativement surexprimées en D par rapport à A au sein des grands ensembles de voies régissant les processus biologiques, les fonctions moléculaires et les composants cellulaires.



La Figure 20 est visualisable plus en détails en double-cliquant sur l'icône PDF ci-dessous :



Figure 20

Cette vue d'ensemble permet de se rendre compte de quelle façon les voies fonctionnelles identifiées sont liées aux processus biologiques, aux fonctions moléculaires et aux composants cellulaires. On observe également que ces voies sont liées entre elles de manière plus ou moins directe.

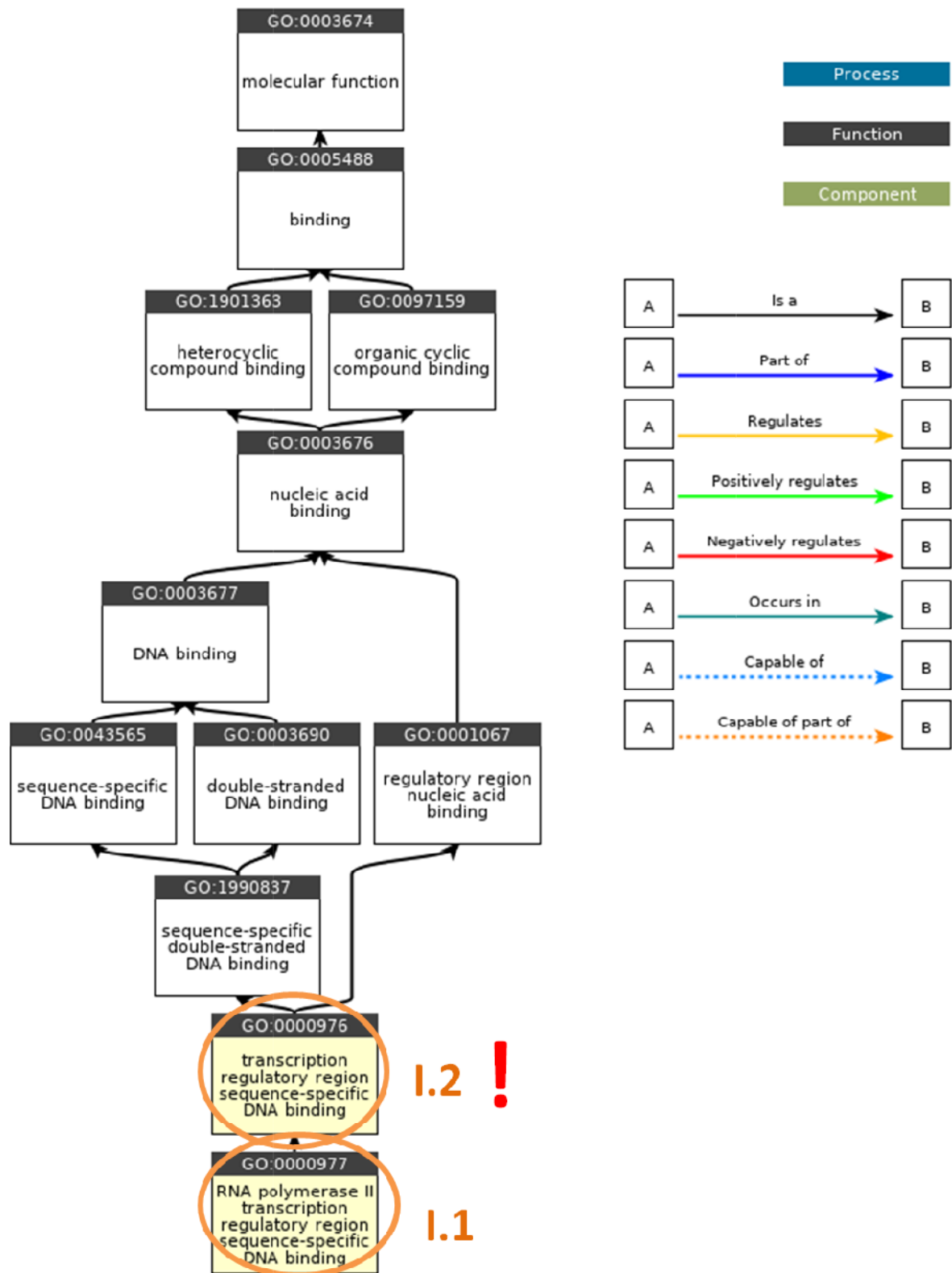
On distingue ainsi quatre grands groupes :

- **Groupe I**: RNA polymerase II regulatory region sequence-specific DNA binding ([GO:0000977](#)) et Transcription regulatory region DNA binding ([GO:0000976](#))
- **Groupe II**: Positive regulation of transcription, DNA-templated ([GO:0045893](#)), Positive regulation of transcription from RNA polymerase II promoter ([GO:0045944](#)), Regulation of transcription from RNA polymerase II promoter ([GO:0006357](#)), Positive regulation of nucleic acid-templated transcription ([GO:1903508](#)), Negative regulation of transcription, DNA-templated ([GO:0045892](#)) et Transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific binding ([GO:0001228](#))
- **Groupe III**: Regulation of protein serine/threonine kinase activity ([GO:0071900](#))
- **Groupe IV**: Negative regulation of immune response ([GO:0050777](#))

Pour avoir une vue plus précise et rapprochée, nous exploiterons par la suite le lien entre les voies à l'intérieur de ces 4 groupes.

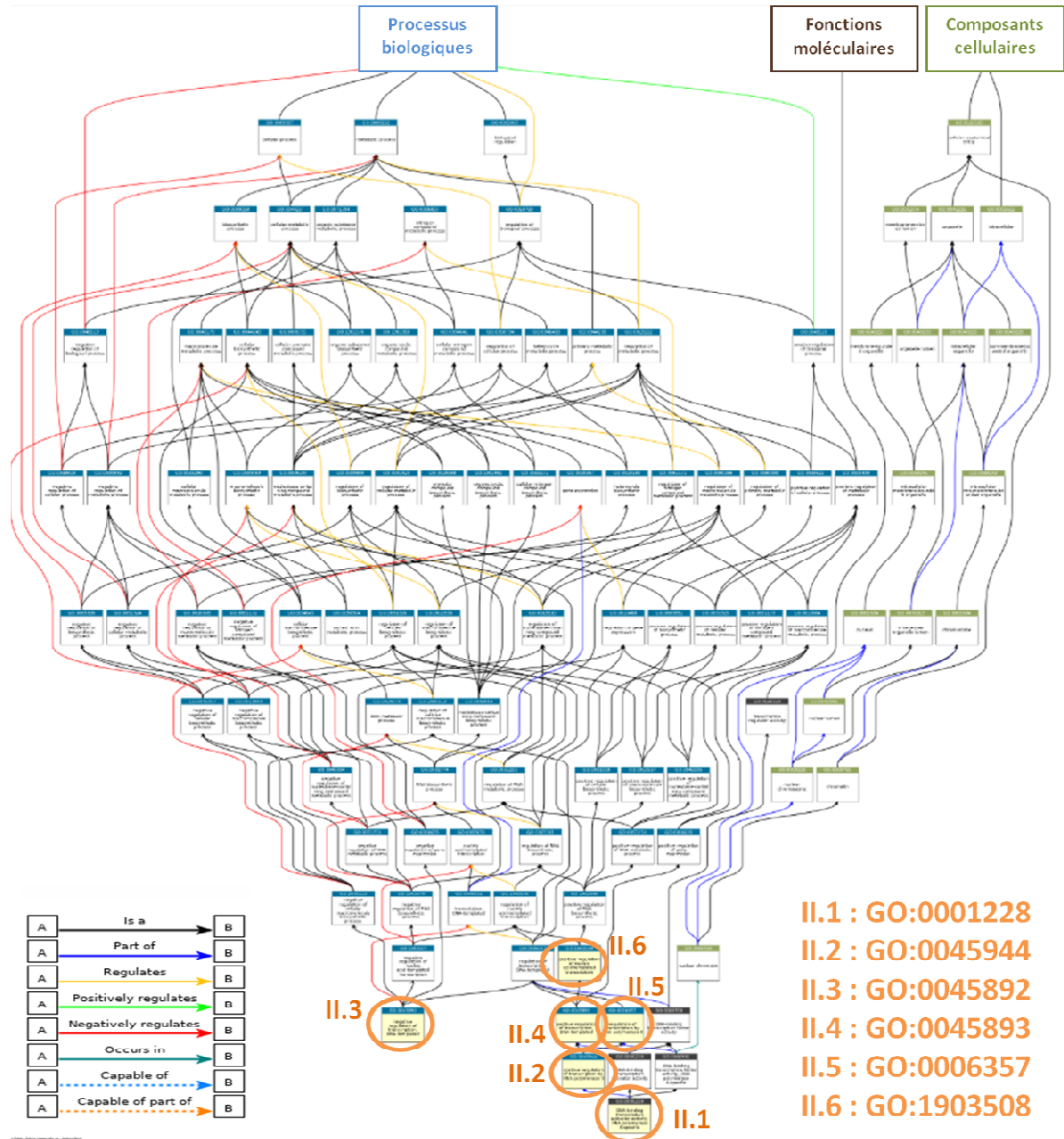
Les résultats obtenus sont illustrés en Figures 21, 22, 23 et 24.

Figure 21 : Voies fonctionnelles enrichies par les gènes surexprimés en D par rapport à A (Groupe I : GO:000977 (1) et GO:000976 (2)).



Le groupe I est essentiellement impliqué dans les fonctions moléculaires en relation avec la régulation de l'initiation de la transcription par liaison sur des séquences ou des motifs particuliers de l'ADN.

Figure 22 : Voies fonctionnelles enrichies par les gènes surexprimés en D par rapport à A (Groupe II : GO:0001228 (1), GO:0045944 (2), GO:0045892 (3), GO:0045893 (4), GO:0006357 (5) et GO:1903508 (6)).



La Figure 22 est visualisable plus en détails en double-cliquant sur l'icône PDF ci-dessous :



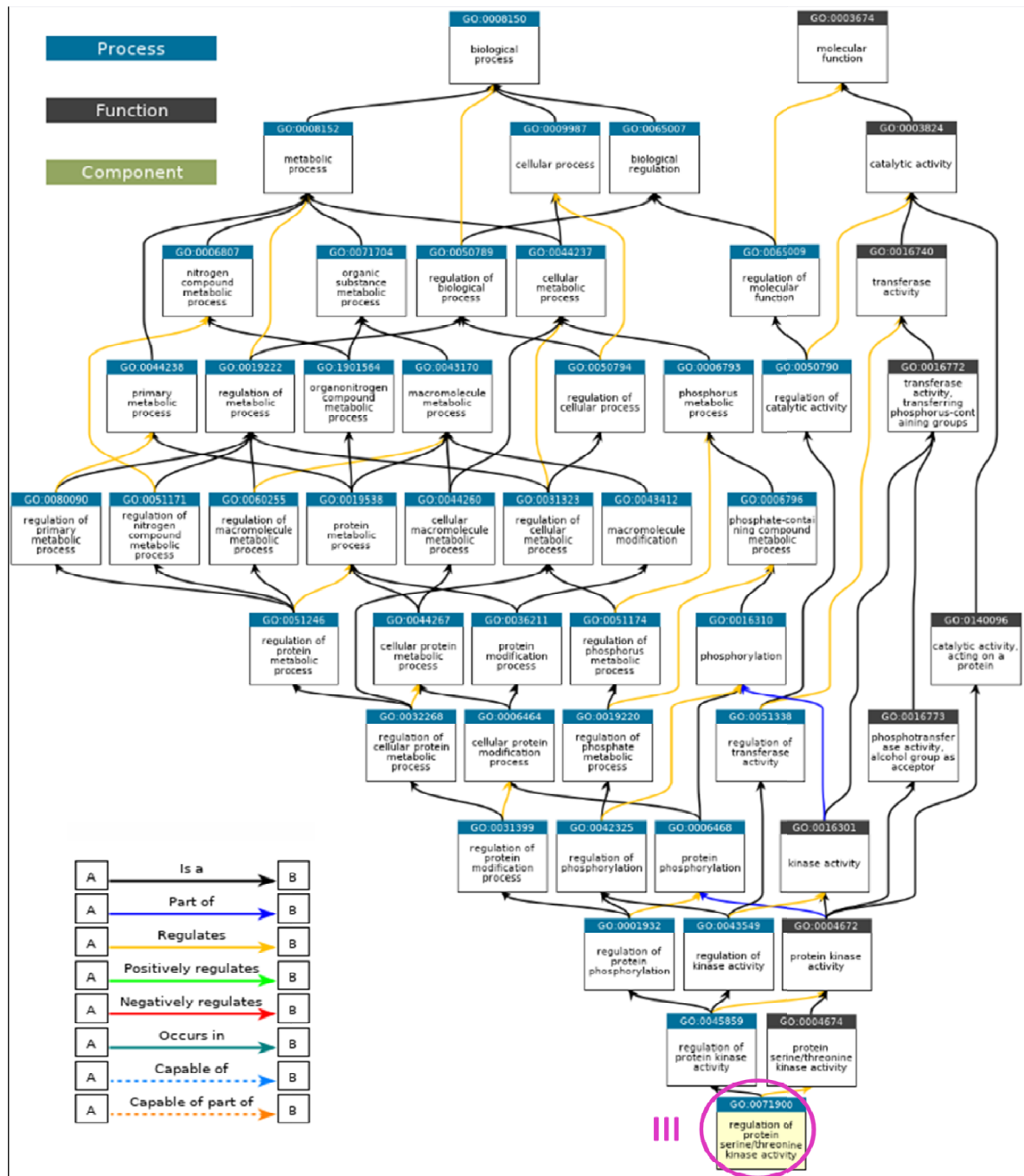
Figure 22

Le groupe II est majoritairement relié aux processus biologiques et notamment à leur régulation, ainsi qu'au métabolisme cellulaire.

Les voies GO:0045944 (II.2), GO:0045893 (II.4), GO:0006357 (II.5) et GO:1903508 (II.6) sont plutôt impliquées dans la régulation positive de processus biologiques, alors que la voie GO:0045892 (II.3) est, elle, plutôt liée à la régulation négative.

La voie GO:0001228 (II.1) est la seule à avoir également des liens avec les fonctions moléculaires (en lien avec la transcription) et les composants cellulaires (en lien avec la chromatine).

Figure 23 : Voies fonctionnelles enrichies par les gènes surexprimés en D par rapport à A (Groupe III : GO:007190).



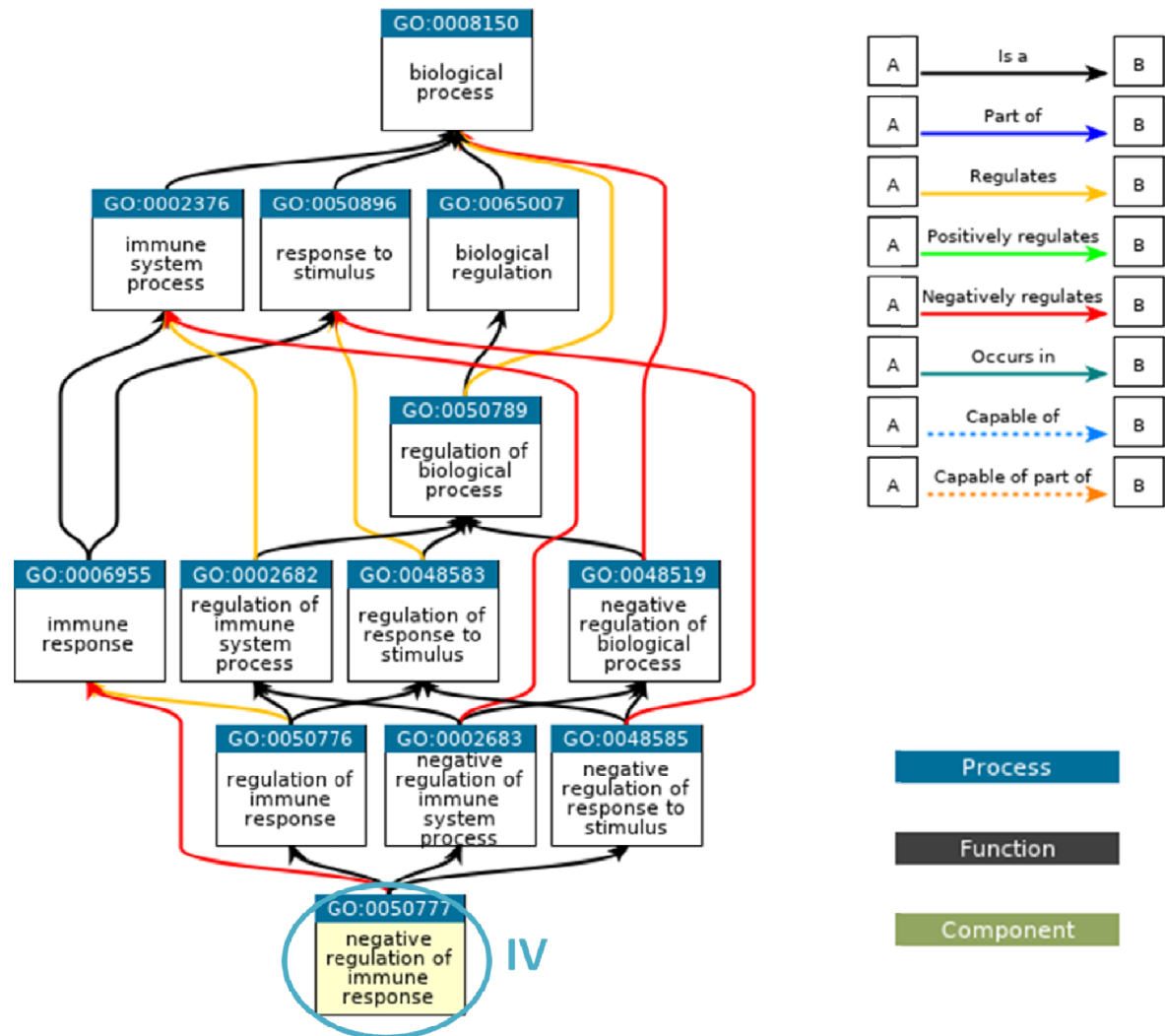
La Figure 23 est visualisable plus en détails en double-cliquant sur l'icône PDF ci-dessous :



Figure 23

Le groupe III est constitué d'une voie fonctionnelle ayant un lien avec les fonctions moléculaires et les processus biologiques. Cette voie est impliquée dans la régulation du métabolisme et des modifications des protéines de manière générale, mais elle possède également un lien avec un ensemble de voies liées plus spécifiquement au métabolisme enzymatique.

Figure 24 : Voies enrichies par les gènes surexprimés en D par rapport à A (Groupe IV : GO:0050777).



Le groupe IV est principalement impliqué dans les processus biologiques. Elle possède notamment un lien avec la régulation, majoritairement négative, de la réponse aux stimuli.

IV. Discussion

Dans notre étude, les transcriptomes d'embryons ayant complété le 4^e cycle embryonnaire depuis 0 à 12h (soit post-activation du génome embryonnaire) et avec des trajectoires morphocinétiques différentes ont été comparés.

Notre but était d'identifier si ces différences de développement morphocinétique pouvaient se refléter dans des transitions materno-embryonnaires divergentes entre les catégories.

Tout d'abord, des différences de profil transcriptomique ont pu être mises en évidence entre les quatre catégories morphocinétiques (A, B, C et D). Les transcriptomes des embryons présentant un développement précoce sans anomalie morphocinétique évidente (catégorie A) et celui des embryons présentant un développement précoce arythmique d'un point de vue morphocinétique (catégorie D) étaient ceux qui présentaient le plus de différences.

La comparaison entre ces deux catégories extrêmes a permis de mettre en évidence un très grand nombre de gènes surexprimés dans les embryons de la catégorie D par rapport à ceux de la catégorie A (au total 899 gènes, dont 785 connus), et un plus petit nombre de gènes surexprimés dans la catégorie A par rapport à la catégorie D (au total 60 gènes, dont 36 connus).

Les gènes surexprimés dans les embryons de la catégorie D étaient en grande partie des gènes à profil transcriptomique maternel (37,5%) et très peu de profil embryonnaire (6,3%).

Dans la catégorie A, au contraire, parmi les gènes surexprimés, très peu avaient un profil maternel (2,5%) alors qu'une forte proportion avaient un profil embryonnaire (47,5%).

De plus, si les 60 gènes surexprimés dans les embryons de la catégorie A par rapport à ceux de la catégorie D n'enrichissaient significativement aucune voie fonctionnelle, les 899 gènes surexprimés dans les embryons de la catégorie D enrichissaient plusieurs voies fonctionnelles (n=13).

La classification *Gene Ontology* a révélé que huit de ces voies étaient impliquées dans la régulation positive ou négative de la transcription, dans la régulation de la kinase sérine-thréonine et dans la régulation négative du système immunitaire.

Par ailleurs, deux voies enrichies étaient impliquées dans des phénotypes générant de la létalité néo- et post-natale et une voie concernait les anomalies de signaux dans les synapses. Comme nous n'avons aucune information sur le phénotype à la naissance de ces catégories, nous n'avons pas traité ces dernières dans la présente étude.

Ces résultats ont donné lieu à de nouvelles questions qui seront abordées à deux niveaux différents.

Premièrement, nous discuterons de ce que peut nous apporter l'étude des gènes différentiellement exprimés dans les embryons de ces deux catégories morphocinétiques.

Nous nous concentrerons notamment sur les potentielles explications du profil en grande partie maternel des gènes surexprimés dans les embryons au développement précoce arythmique (catégorie D).

Deuxièmement, nous nous focaliserons sur les voies fonctionnelles enrichies par les gènes surexprimés dans les embryons de la catégorie D, afin d'essayer de comprendre l'intérêt qu'elles peuvent avoir dans le développement embryonnaire et le lien potentiel avec des cycles cellulaires précoces arythmiques.

1. Les gènes significativement surexprimés dans les embryons des catégories A ou D

La plus forte proportion de transcrits de gènes à profil maternel surexprimés dans les embryons de la catégorie D par rapport à la catégorie A ainsi que la très faible présence de transcrits de gènes de profil embryonnaire dans la première suggèrent un retard de transition materno-embryonnaire chez ces embryons.

Celle-ci étant définie par l'élimination des transcrits maternels et l'activation du génome embryonnaire, ce retard pourrait être imputé à l'un ou l'autre de ces phénomènes.

Les embryons de la catégorie D pourraient donc faire face à :

- Un retard d'utilisation des transcrits maternels ;
- Un retard de dégradation des transcrits maternels ;
- Un retard d'activation transcriptionnelle.

Cependant, il ne faut pas perdre de vue que l'étude des profils des gènes a été faite sur la partie des gènes différentiellement exprimés dans l'une ou l'autre des catégories.

Or, les embryons de la catégorie D exprimaient au total un peu moins de 20 000 gènes parmi lesquels 899 étaient surexprimés en comparaison avec la catégorie A. Ce qui signifie qu'ils expriment de manière similaire à la catégorie A plusieurs milliers d'autres gènes.

De plus, les embryons inclus dans cette étude ont tous été prédits pour atteindre le stade blastocyste.

Donc, si défaut il y a chez les embryons de la catégorie D, celui-ci est partiel et n'affecte pas des gènes absolument essentiels à la survie des embryons jusqu'au stade blastocyste.

a. Un retard d'utilisation des transcrits maternels par les embryons de la catégorie D ?

Une des hypothèses pouvant expliquer le retard de MET chez les embryons de la catégorie D serait que ces embryons possèdent un retard d'utilisation des transcrits maternels.

Les transcrits maternels sont stockés dans une conformation bien précise, en particules ribonucléoprotéiques (mRNP). Cet assemblage en mRNP se fait grâce à des protéines de masquage qui inhibent la traduction en empêchant la liaison de ses facteurs initiateurs aux protéines fixées sur l'ARNm (Sommerville, 1999 ; Richter, 2008). Leur utilisation nécessite donc un démasquage afin d'initier la traduction, lui-même dépendant d'une polyadénylation cytoplasmique.

Il se pourrait donc que les embryons de la catégorie D aient commencé plus tard le démasquage de leur ARNm maternels que ceux de la catégorie A et qu'ils n'aient donc pas pu encore avoir accès à la totalité du stock maternel. L'abondance des transcrits maternels n'a donc pas autant diminué que dans la catégorie A.

D'autre part, un défaut fonctionnel d'une des protéines actrices du mécanisme de démasquage des transcrits maternels engendrerait également une indisponibilité de certains ARNm pour les embryons de la catégorie D, et expliquerait cette plus forte abondance de transcrits maternels chez eux.

Il faudrait cependant que ce défaut soit partiel car l'utilisation des ressources maternelles par l'embryon est déterminante pour un développement précoce correct (Harbour et Dean, 2000 ; Forjas de Borja *et al.*, 2001 ; Hara *et al.*, 2005). Or, tous les embryons inclus dans notre étude sont prédits pour devenir blastocyste.

Si un retard d'utilisation des transcrits maternels existe donc chez les embryons de la catégorie D, l'hypothèse selon laquelle ces embryons ont débuté tardivement le démasquage de leurs ARNm maternels semble plus plausible que celle d'un défaut fonctionnel dans la régulation de ce démasquage.

b. Un retard de dégradation des transcrits maternels chez les embryons de la catégorie D ?

Un retard de MET chez les embryons de la catégorie D pourrait également être dû à un retard de dégradation des transcrits maternels.

Plusieurs mécanismes peuvent impacter cette dégradation : la voie de l'interférence, les protéines de liaison à l'ARN et l'autophagie.

➤ Implication de la voie de l'interférence ?

L'un des mécanismes potentiellement impliqués dans le processus de dégradation des transcrits maternels est la voie de l'interférence.

Un dysfonctionnement de celle-ci engendrerait une élimination incomplète ou inefficace des transcrits maternels.

La question qui se pose alors est de savoir si l'on retrouve des transcrits des gènes cibles d'ARNmi (acteurs principaux de cette voie de l'interférence) parmi les transcrits en surabondance (associés aux gènes surexprimés) dans les embryons des catégories A et D, et ce, afin de déterminer si la voie de l'interférence aurait pu être plus ou moins efficace dans l'une des deux catégories.

Mondou *et al.* (2012) ont établi la liste des 946 gènes du génome bovin cibles des ARNmi miR-21 et miR-130a.

On retrouve 60 gènes de cette liste parmi les gènes surexprimés dans les embryons de la catégorie D par rapport à ceux de la catégorie A.

Parmi ces 60 gènes, à l'exception de deux dont les profils n'ont pas pu être établis, 21 ont un profil maternel et 37 un profil materno-embryonnaire.

En ce qui concerne les gènes surexprimés dans la catégorie A par rapport à la catégorie D, un seul, de profil materno-embryonnaire, s'avère être une cible des ARNmi miR-21 et miR-130a : *S100A10*.

D'après Stelzer *et al.* (2016), ce gène code pour une protéine de la famille S100 contenant deux motifs main EF (motif structurel constitué de deux hélices α enserrant un ion Ca^{2+}).

Les protéines de cette famille sont impliquées dans un grand nombre de processus cellulaires tels que la progression du cycle cellulaire et la différenciation.

Or, cette dernière fonction est un phénomène plutôt tardif du développement embryonnaire (stade blastocyste). La surexpression de ce gènes dans les embryons de la catégorie A est donc plus probablement due à sa transcription de novo et non pas à un manque de dégradation de ses transcrits maternels.

Un autre ARNmi d'intérêt vis-à-vis de nos résultats est le miR-196a. En effet, Tripurani *et al.* (2011) ont montré que le miR-196a était présent du stade deux cellules au stade huit cellules dans l'embryon bovin et qu'il agissait sur les transcrits du gène *NOBOX*.

Ce gène fait également partie des gènes surexprimés par les embryons de la catégorie D par rapport aux embryons de la catégorie A et a un profil maternel.

Selon Li *et al.* (2017), *NOBOX* joue un rôle dans l'arrêt de méiose en prophase I dans les ovocytes humains. Un dérèglement de ce gène ou de sa régulation entraîne donc des cycles cellulaires anormaux.

En résumé, 6,8% des gènes surexprimés dans les embryons de la catégorie D sont des gènes cibles des ARNmi miR-21, miR-130a ou miR-196a (61 gènes cibles d'ARNmi sur 899 gènes au total).

Quatre des gènes cibles identifiés ne sont pas exprimés aux stades seize cellules et blastocyste dans les embryons produits *in vivo* selon la base de référence EmbryoExpress : *GAP43*, *NCAN*, *PITPNM2* et *UCP3*.

En comparaison, seul un gène sur les 60 surexprimés dans les embryons de la catégorie A par rapport à la catégorie D (soit 1,7%) est un gène cible des ARNmi miR-21 et miR-130a. Cependant, ce gène étant materno-embryonnaire et présent chez des embryons produits *in vivo* aux stades seize cellules et blastocyste selon EmbryoExpress, sa surexpression pourrait être le fruit d'une transcription *de novo* par les embryons de la catégorie A.

Il y a donc une plus forte proportion de gènes cibles d'ARNmi (relativement, 4 fois plus) parmi les gènes surexprimés dans les embryons de la catégorie D que parmi ceux des embryons de la catégorie A.

Ces résultats confortent l'hypothèse selon laquelle les embryons avec un développement morphocinétique arythmique présenteraient un retard ou un manque de dégradation de leurs transcrits maternels, au moins en partie en lien avec la voie de l'interférence.

➤ Pourquoi la voie de l'interférence serait-elle dysfonctionnelle dans les embryons de la catégorie D ?

Dans l'hypothèse où le retard de MET serait lié, au moins en partie, à un manque de dégradation par la voie de l'interférence de certains transcrits maternels chez les embryons de la catégorie D, la question qui se pose alors est de savoir pourquoi celle-ci aurait moins bien fonctionné dans ces embryons là.

Une première hypothèse serait que ces embryons ne possèdent pas ou de manière trop peu importante les ARNmi nécessaires à la dégradation des transcrits maternels, de par leur activation transcriptionnelle retardée.

En effet, certains ARNmi, à l'instar des miR-21 et miR-130a, sont les produits d'une transcription embryonnaire (Mondou *et al.*, 2012), donc si les embryons n'ont pas encore complètement enclenché leur machinerie transcriptionnelle, il se peut qu'une partie des ARNmi normalement nécessaires à la dégradation des transcrits maternels à ce stade ne soient pas encore présents ou en quantité trop infime pour avoir une réelle action dans l'embryon.

Une deuxième possibilité serait que, malgré la présence de ces ARNmi, leurs cibles ne soient pas accessibles.

En effet, les transcrits maternels étant stockés en particules ribonucléoprotéiques grâce à des protéines de masquage qui inhibent la traduction en empêchant la liaison de ses facteurs initiateurs aux protéines fixées sur l'ARNm (Sommerville, 1999 ; Richter, 2008), on peut imaginer que dans cet état conformationnel, l'ARNm puisse être également difficile d'accès pour les ARNmi. Un retard démasquage des ARNm maternels pourrait expliquer une moindre efficacité de la voie de l'interférence.

Il apparaissait également comme une explication probable à l'hypothèse d'un retard d'utilisation des transcrits maternels dans les embryons de la catégorie D.

Les deux hypothèses peuvent donc être liées : un retard de démasquage génèrerait une indisponibilité de certains ARNm pour l'embryon ainsi qu'une action incomplète des acteurs de la voie de l'interférence, engendrant à la fois un manque d'utilisation et un manque de dégradation des transcrits maternels.

➤ D'autres mécanismes possiblement impliqués ?

L'éventuel retard de MET lié à un manque de dégradation des transcrits maternels chez les embryons de la catégorie D peut être la conséquence d'autres mécanismes que la voie de l'interférence.

En effet, nous avons également vu que les protéines de liaison à l'ARN sont des acteurs de cette étape de la MET (Lai *et al.*, 2000 ; Walser et Lipshitz, 2011).

De la même manière que pour les ARNm, les embryons au développement morphocinétique précoce arythmique peuvent manquer de ces protéines ou alors leurs ARNm peuvent être moins accessibles à celles-ci.

Enfin, l'autophagie a un rôle important dans la dégradation des transcrits maternels chez la souris et le porc (Tsukamoto *et al.*, 2008 ; Xu *et al.*, 2012), il n'est pas donc exclu qu'un dysfonctionnement de celle-ci puisse avoir des conséquences dans la MET chez les bovins également. Cependant, des études complémentaires sont nécessaires afin d'en apporter la preuve et, de ce fait, nous ne considérerons pas son implication dans le retard de MET observé chez les embryons de la catégorie D comme plausible.

c. Un retard d'activation transcriptionnelle chez les embryons de la catégorie D ?

Une troisième hypothèse quant au retard dans l'accomplissement de la MET chez les embryons de la catégorie D serait que ces derniers peinent à mettre en route leur machinerie transcriptionnelle. Un retard d'activation du génome embryonnaire seul ne permettrait pas d'expliquer la plus grande quantité de transcrits maternels chez les embryons de la catégorie D par rapport à la catégorie A. Néanmoins, ce phénomène pourrait se surajouter à un manque d'utilisation et/ou de dégradation des transcrits maternels par l'embryon et ainsi accentuer le retard de MET.

La méthylation et la réplication de l'ADN sont deux mécanismes potentiellement impliqués dans l'activation transcriptionnelle des embryons bovins, une défaillance partielle dans un de ces deux phénomènes pourrait alors engendrer un retard d'EGA.

➤ Un défaut de déméthylation ?

Des phénomènes de déméthylation et méthylation *de novo* se produisent tout au long du développement embryonnaire précoce.

La méthylation de l'ADN engendre une répression de la transcription.

Chez les bovins le génome est déméthylé depuis la fécondation jusqu'à un seuil maximum atteint vers le stade huit cellules, au moment de l'EGA (Dean *et al.*, 2003).

Un retard partiel de déméthylation dans les embryons de la catégorie D engendrerait un déficit en transcription puisque l'inhibition de celle-ci ne serait pas totalement levée, et entraînerait alors un retard dans l'EGA de ces embryons.

Un mécanisme sous-jacent à ce manque de déméthylation pourrait être lié à l'ADN méthyltransférase 1 (*DNMT1*).

En effet, d'après Dean *et al.* (2003), la déméthylation du génome est le résultat de l'expulsion du noyau de la *DNMT1* au cours des cycles réplicatifs de l'ADN.

Étant donné que les embryons de la catégorie D initient des mitoses anormales au début de leur développement, peut-être que la *DNMT1* n'est pas expulsée comme elle le devrait et engendre donc une reméthylation de l'ADN suite à sa réplication, maintenant ainsi l'inhibition de la transcription.

Une deuxième hypothèse proviendrait des conditions de culture *in vitro* des embryons de notre étude.

Salilew-Wondim *et al.* (2018) ont montré que des conditions de culture *in vitro* avant ou pendant l'EGA entraînent un schéma de méthylation de l'ADN aberrant chez les embryons bovins.

Les conditions de culture *in vitro* pourraient donc être à l'origine d'un défaut de déméthylation dans les embryons de la catégorie D, soit directement en empêchant la déméthylation de certaines régions génomiques, soit indirectement via l'hypométhylation de gènes dont la transcription serait propice à la reméthylation ou défavorable à la déméthylation de l'ADN.

Cette deuxième hypothèse a tout de même une faille : les embryons de la catégorie A ont été soumis aux mêmes conditions de culture que les embryons D, et, eux, ne semblent pas présenter de retard dans leur MET.

Ainsi, la théorie la plus plausible pour expliquer un défaut de déméthylation serait donc celle d'une origine intrinsèque à l'ovocyte et/ou à l'embryon plutôt qu'une origine extrinsèque liée à l'environnement de production *in vitro*.

Cependant, à l'état actuel des connaissances sur ces catégories morphocinétiques d'embryons, nous ne pouvons pas écarter la possibilité que les embryons de la catégorie A se soient mieux adaptés que ceux de la catégorie D à l'environnement de production auquel ils ont été soumis.

➤ Un problème dans la réplication de l'ADN ?

En plus de ce défaut de déméthylation que pourraient engendrer des cycles réplicatifs de l'ADN défectueux, ceux-ci pourraient également entraîner un remodelage imparfait de la chromatine.

En effet, Aoki *et al.* (1997) avaient démontré les états alternativement permissif et répressif des cycles cellulaires quant à la transcription et Latham et Shultz (2001) avaient suggéré que ce phénomène permettait un remodelage complet de la chromatine qui activait ou réprimait différenciellement les gènes spécifiques au bon déroulement de l'EGA.

Des cycles réplicatifs anormaux pourraient donc engendrer la répression ou la non activation ponctuelle de gènes nécessaires à l'activation transcriptionnelle des embryons et expliquer un retard dans la MET de ceux-ci.

EN RÉSUMÉ

La plus grande part de gènes à profil maternel ainsi que la faible part de gènes à profil embryonnaire surexprimés dans les embryons de la catégorie D par rapport aux embryons de la catégorie A suggèrent que les embryons avec un développement morphocinétique précoce arythmique sont en retard dans leur transition materno-embryonnaire par rapport aux embryons n'ayant pas présentés d'anomalie morphocinétique de division.

Ces embryons pourraient faire face à un retard d'utilisation et/ou de dégradation des transcrits maternels.

Un retard de démasquage des ARNm maternels expliquerait qu'une partie de ceux-ci n'aient pas pu être traduits et donc utilisés par les embryons, alors qu'une voie de l'interférence moins efficace pourrait expliquer un retard de dégradation des transcrits maternels.

En effet, nous avons identifié plusieurs gènes cible d'ARNmi surexprimés dans la catégorie D et en plus forte proportion que dans les gènes surexprimés par la catégorie A.

De plus, les deux hypothèses peuvent être liées, puisqu'un retard de démasquage des ARNm maternels pourrait entraîner non seulement un retard d'utilisation des transcrits par l'embryon mais également un retard de dégradation de ceux-ci en bloquant l'accès des ARNmi à leurs cibles.

Un manque de dégradation des transcrits maternels peut également être expliqué par un défaut fonctionnel des protéines de liaison à l'ARN dans les embryons de la catégorie D. Cependant, nous n'avons pas d'éléments nous permettant de vérifier cette hypothèse.

En parallèle d'un retard d'élimination des transcrits maternels, les embryons de la catégorie D peuvent faire face à un retard d'activation transcriptionnelle qui accentuerait leur retard de MET.

Un défaut de méthylation ou dans la réplication de l'ADN lors des premiers cycles cellulaires arythmiques de ces embryons auraient pu engendrer ce retard d'EGA. Mais actuellement nous n'avons pas d'éléments nous permettant d'aller plus loin dans la validation de ces hypothèses.

2. Les voies fonctionnelles significativement enrichies par les gènes surexprimés dans les embryons avec un développement morphocinétique arythmique (catégorie D)

Nous allons maintenant nous intéresser aux voies fonctionnelles enrichies par les gènes surexprimés dans les embryons de la catégorie D par rapport aux embryons de la catégorie A.

Nous chercherons d'abord à établir si l'enrichissement observé semble cohérent avec le stade embryonnaire auquel se trouvent nos embryons. Puis, nous proposerons des hypothèses pouvant expliquer nos différentes observations.

Pour rappel, les embryons de la catégorie D surexprimaient des gènes impliqués dans des voies fonctionnelles liées à la régulation positive ou négative de la transcription, à la régulation de la kinase sérine-thréonine, à la régulation négative du système immunitaire, à la létalité néo- et post-natale et aux anomalies de signaux dans les synapses.

La question qui se pose alors est pourquoi ces voies sont-elles plus enrichies dans la catégorie D que dans la catégorie A ?

D'après Graf *et al.* (2014b), au stade seize cellules, on observe chez des embryons bovins produits *in vitro* l'enrichissement des voies fonctionnelles en lien avec la régulation de la transcription et notamment la transcription par l'ARN polymérase II :

- Régulation négative de la transcription ;
- Régulation de la transcription par l'ARN polymérase II des catabolites du carbone ;
- Initiation et élongation de la transcription par l'ARN polymérase II.

L'enrichissement de ces voies fonctionnelles par les gènes surexprimés dans les embryons de la catégorie D semble donc cohérent avec leur stade de développement (16-32 cellules) par comparaison avec des embryons bovins, au même stade, produit *in vitro*, sans différenciation de profil morphocinétique.

A contrario, les embryons de la catégorie D enrichissent également une voie fonctionnelle liée à la régulation négative du système immunitaire (GO:0050777) dont un des sous-types est la voie fonctionnelle : « Régulation négative de la réponse aux stimuli » (voir III.3.d).

Or Graf *et al.* (2014b) ont retrouvé cette dernière activée au stade blastocyste chez l'embryon bovin produit *in vitro*.

Par rapport aux embryons de la catégorie A qui n'enrichissent aucune voie fonctionnelle, les embryons de la catégorie D enrichissent donc simultanément des voies fonctionnelles spécifiques au stade seize cellules et au stade blastocyste.

Une des premières hypothèses que génère l'observation d'une telle différence est que les deux catégories d'embryons sont en réalité dans des cycles cellulaires différents.

Celle-ci est très peu plausible car un protocole minutieux a été mis en place pour être sûr que l'étude porterait sur des groupes comparables.

En effet, une lecture minutieuse a été réalisée pour nous assurer que tous les embryons inclus dans l'étude avaient accompli le 4^e cycle embryonnaire, et ce, depuis zéro à douze heures maximum, ce qui ne correspond même pas à un cycle cellulaire complet.

On parle d'embryons au stade 16-32 cellules car dans la catégorie D, les embryons ont subi des anomalies morphocinétiques qui ont engendré des divisions dites directes à trois, voir quatre cellules au cours du premier cycle embryonnaire. Les embryons de cette catégorie peuvent donc posséder plus de cellules que ceux de la catégorie A mais les embryons des deux groupes ont strictement effectué le même nombre de cycles embryonnaires (n=4), à un demi-cycle de division près.

Une autre hypothèse permettrait d'expliquer la différence d'enrichissement entre les catégories A et D.

Il se pourrait que l'enrichissement des voies fonctionnelles dans la catégorie D soit en fait le résultat de la grande quantité de transcrits maternels encore présents dans ces embryons comparés à ceux de la catégorie A.

En effet, le démarrage arythmique du développement des embryons de la catégorie D aurait pu engendrer un retard d'élimination (soit par manque d'utilisation soit par manque de dégradation) des transcrits maternels. Ceux-ci auraient alors enrichi de manière artificielle des voies fonctionnelles « typiques » du stade de développement auquel ces embryons se trouvent (celles spécifiques du stade seize cellules) et « atypiques » (celles associées au stade blastocyste et celles non identifiées comme différentiellement enrichies à aucun stade du développement par les études de Misirlioglu *et al.* (2006), Vigneault *et al.* (2009) et Graf *et al.* (2014b)).

Cette hypothèse est confortée par le fait que, parmi les 61 gènes cibles d'ARNmi surexprimés dans les embryons de la catégorie D, 19 sont impliqués dans l'enrichissement de voies fonctionnelles chez ces embryons : *BMPR1B*, *CASR*, *CDKN1A*, *ELK3*, *FOSL1*, *GAP43*, *GATA2*, *JPH1*, *MYB*, *NOBOX*, *NRXN1*, *PFN2*, *PPARGC1A*, *SLC12A5*, *SMAD5*, *SMOC1*, *SOX5*, *TRIB1* et *UBE2D1*.

Les voies fonctionnelles principalement enrichies par ces gènes sont les voies concernant la régulation positive de la transcription (GO:0045893 ; GO:0045944 ; GO:0006357).

Par ailleurs, la voie fonctionnelle impliquée dans la régulation de la kinase sérine/thréonine est également une voie fonctionnelle enrichie par les gènes surexprimés dans les embryons de la catégorie D par rapport à la catégorie A.

Celle-ci ne fait pas partie des voies fonctionnelles enrichies par l'embryon bovin produit *in vitro* mises en évidence par Misirlioglu *et al.* (2006), Vigneault *et al.* (2009) et Graf *et al.* (2014b), pour aucun des stades étudiés par les auteurs.

Cette voie est liée à des processus métaboliques protéiques (et notamment enzymatiques) et de phosphorylations (voir partie III.3.d).

Son enrichissement dans les embryons de la catégorie D pourrait indiquer un besoin accru en protéines pour ces embryons-là, potentiellement lié à leur activité transcriptionnelle encore très importante.

L'enrichissement de cette voie fonctionnelle pourrait également être le résultat d'une activité accrue de kinases sérine/thréonine spécifiques. Or, il se trouve qu'une kinase sérine/thréonine très importante pour le développement embryonnaire précoce est la protéine Aurora A.

En effet, cette protéine phosphoryle les CPEB liés aux séquences CPE des ARNm, première étape du démasquage et donc initie la traduction des ARNm maternels stockés sous forme de mRNP dans l'embryon précoce (Richter, 2008 ; Uzbekova *et al.*, 2008). Il est donc possible que les embryons de la catégorie D soient en pleine phase de démasquage des ARNm maternels.

Si c'est le cas, cela corroborerait l'hypothèse émise dans la première partie de cette discussion, selon laquelle le manque d'utilisation et/ou de dégradation des transcrits maternels serait imputable à un retard de démasquage des ARNm maternels.

De plus, cela va également dans le sens d'un enrichissement différentiel de voies fonctionnelles dans les embryons de la catégorie D du fait de la forte abondance des transcrits maternels dans cette catégorie par rapport à A.

Par ailleurs, Uzbekova *et al.* (2008) ont souligné le rôle de la protéine Aurora A dans la progression du cycle cellulaire.

L'enrichissement d'une voie fonctionnelle régulant l'activité de protéines telles qu'Aurora A pourrait donc également être une réponse des embryons de la catégorie D suite à leurs divisions anormales au début du développement.

Un des scénarios qui semble donc le plus probable est que le retard de MET des embryons de la catégorie D a engendré une plus forte abondance de transcrits maternels dans ces embryons. Les gènes associés à ces transcrits ont, de ce fait, enrichi de manière artificielle des voies fonctionnelles dans la catégorie D par rapport à la catégorie A.

Quant à savoir si le retard de MET des embryons de la catégorie D est le résultat d'une adaptation de ceux-ci afin de pallier aux anomalies divisionnelles qu'ils ont supportées au début de leur développement ou non, cela ne pourra être résolu que dans le futur via de nouvelles études plus approfondies sur le sujet.

EN RÉSUMÉ

Les embryons de la catégorie D surexprimaient suffisamment de gènes pour enrichir treize voies fonctionnelles, contrairement aux embryons de la catégorie A qui n'enrichissaient aucune voie fonctionnelle.

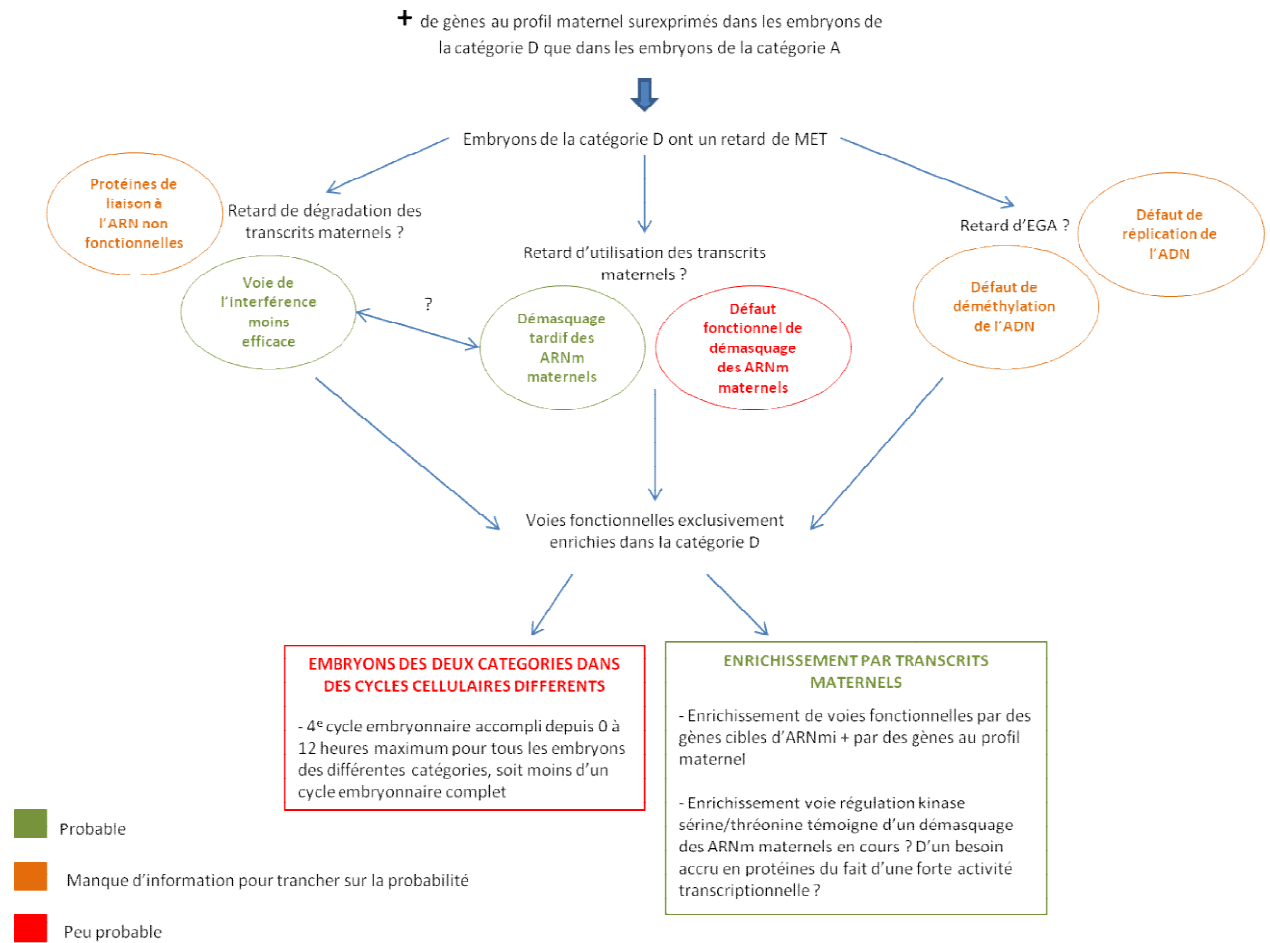
Du fait d'un protocole de sélection minutieux dans lequel seuls les embryons ayant complétés le 4^e cycle embryonnaire depuis 0 à 12 heures maximum, soit moins de temps qu'un cycle embryonnaire complet, ont été inclus, il est impossible que les embryons des deux catégories ne soient pas dans le même stade cellulaire.

Les transcrits maternels, présents en forte abondance dans les embryons avec un développement précoce arythmique, auraient pu engendrer l'enrichissement artificiel des voies fonctionnelles spécifiques et non spécifiques à leur stade de développement (16-32 cellules).

En effet, les divisions arythmiques qu'ont subi les embryons de la catégorie D auraient pu entraîner un réel retard de MET, engendrant un retard dans l'épuisement des stocks de transcrits maternels. Cependant, ce retard dans l'élimination des transcrits d'origine maternelle ne semble pas être létal pour ces embryons puisqu'il a été prédit qu'ils atteindraient le stade blastocyste. A l'heure actuelle, nous ne pouvons donc pas exclure que leur retard de transition materno-embryonnaire soit en fait le fruit d'une adaptation leur permettant d'acquérir une bonne compétence de développement et de survie malgré les anomalies morphocinétiques qu'ils ont subi au début de leur développement.

Un schéma récapitulatif des deux parties de la discussion de notre étude mettant en lumière les problématiques et les hypothèses qu'a généré ce travail est présenté dans la Figure 25.

Figure 25 : Schéma bilan de la discussion.



Notre étude montre donc qu'il existe de grandes différences transcriptomiques entre des embryons aux développements morphocinétiques différents. Or les études actuelles ne prennent pas en compte ces différences et travaillent sur des populations où les embryons sont tous considérés « similaires ». Le développement morphocinétique serait donc une information importante à prendre en compte, par la suite pour affiner les études et mieux comprendre l'activation du génome embryonnaire dans les différents individus ou sous-groupes d'individus à développement cellulaire similaire.

Conclusion

La transition materno-embryonnaire est un phénomène précoce du développement embryonnaire bovin désignant le transfert du contrôle du développement du transcriptome maternel au transcriptome embryonnaire.

Elle est caractérisée par deux événements : l'élimination des transcrits maternels et l'activation du génome embryonnaire.

La réussite de la transition materno-embryonnaire repose tout d'abord sur la maturation correcte de l'ovocyte lors de l'ovogénèse. Pendant cette étape du développement, certaines parties du génome maternel sont transcrites, produisant alors des ARN et des protéines qui vont être stockés dans l'ovocyte en attendant la fécondation.

Suite à cette dernière, l'embryon va pouvoir utiliser ces réserves pour les premières étapes de son développement, palliant ainsi à l'immatunité de son propre système de transcription et de traduction.

L'utilisation de ces réserves va faire diminuer l'abondance des transcrits maternels dans l'embryon, tout comme leur dégradation, dès le stade une cellule, *via* différents mécanismes encore peu élucidés de nos jours mais qui peuvent impliquer la voie de l'interférence, des protéines de liaison à l'ARN ou l'autophagie.

En parallèle à ce phénomène, entre les stades huit et seize cellules (3^e et 4^e cycles embryonnaires), le génome embryonnaire bovin s'active, mettant en route la machinerie transcriptionnelle.

Cette activation de la transcription permettra à l'embryon de poursuivre son développement au-delà de ces stades-là, en générant ses propres transcrits, afin de remplacer les réserves maternelles.

La méthylation et la réplication de l'ADN de l'embryon sont des phénomènes qui pourraient réguler l'activation du génome embryonnaire.

Cependant, d'après les résultats de notre étude, tous ces événements pourraient ne pas se dérouler de manière identique chez tous les embryons bovins.

En effet, nous avons identifié des différences significatives dans le transcriptome d'embryons bovins avec des développements morphocinétiques dissemblables, après l'activation de leur génome embryonnaire.

L'étude des gènes différentiellement exprimés entre nos quatre catégories révèle un grand nombre de différences, et notamment dans la comparaison opposant les deux catégories les plus éloignées (les embryons au profil morphocinétique sans anomalie apparente, la catégorie A, et les embryons au profil morphocinétique arythmique, la catégorie D).

Au total, 60 gènes étaient surexprimés dans les embryons sans anomalie de développement morphocinétique apparente en comparaison avec les embryons au développement morphocinétique arythmique, dont 2,5% avaient un profil maternel.

Inversement, 899 gènes étaient surexprimés dans les embryons au développement morphocinétique arythmique par rapport aux embryons sans anomalie visible de développement morphocinétique, dont 37,5% présentaient un profil maternel.

L'étude des voies fonctionnelles enrichies par les gènes surexprimés par l'une de ces deux catégories par rapport à l'autre dévoile que les embryons sans anomalie de développement morphocinétique apparente ne surexpriment aucune voie fonctionnelle alors que les embryons au

développement morphocinétique arythmique en surexprimant 13 par rapport aux premiers. Ces voies fonctionnelles sont liées à la régulation positive ou négative de la transcription, à la régulation de la kinase sérine-thréonine, à la régulation négative du système immunitaire, à des phénotypes générant de la létalité néo- et post-natale et à des anomalies de signaux dans les synapses.

L'enrichissement des voies fonctionnelles de manière exclusive à la catégorie d'embryons ayant un profil morphocinétique arythmique, ainsi que leur faible nombre de gènes surexprimés à profil embryonnaire, pourraient alors s'expliquer par la forte proportion de transcrits maternels encore présents chez ces embryons-là.

Ces résultats suggèrent donc qu'un retard de transition materno-embryonnaire existe chez des embryons avec un profil morphocinétique arythmique. Celui-ci pourrait être dû à une utilisation et/ou à une dégradation insuffisante des transcrits maternels.

Un retard de démasquage des ARNm maternels par exemple pourrait rendre inutilisable une partie de ceux-ci pour l'embryon ainsi qu'empêcher la voie de l'interférence de dégrader certains transcrits maternels cibles.

De plus, la faible proportion de gènes à profil embryonnaire dans ces embryons pourrait faire penser qu'ils font également face à un retard dans leur activation transcriptionnelle. Cela pourrait ainsi contribuer au retard de transition materno-embryonnaire déjà initié par le retard d'élimination des transcrits maternels.

Ce retard de transition materno-embryonnaire pourrait être une conséquence des anomalies morphocinétiques qu'ont subi les embryons de la catégorie D au début de leur développement.

Cependant, comme ces embryons vont tout de même se développer jusqu'au stade blastocyste, il n'est pas exclu que ce retard d'élimination des transcrits maternels soit en fait un phénomène adaptatif pour contrer les anomalies précoces.

Les perspectives concernant l'étude de l'activation du génome embryonnaire et notamment son lien avec le développement morphocinétique sont donc nombreuses et découlent du peu de choses connues à ce sujet chez les bovins.

L'une de ces perspectives serait, tout d'abord, de mieux comprendre l'origine du retard de transition materno-embryonnaire chez les embryons avec un profil morphocinétique arythmique.

Pour cela, il serait certainement intéressant d'explorer les mécanismes potentiellement impliqués dans ce retard chez ces embryons, et notamment ceux pouvant être liés à un manque de dégradation des transcrits maternels, en commençant par le rôle de la voie de l'interférence.

De plus, notre étude s'est principalement concentrée sur la comparaison opposant les deux catégories les plus éloignées d'un point de vue morphocinétique car c'est celle-ci qui faisait apparaître le plus grand nombre de différences transcriptomiques entre toutes les catégories, et qui donc, avait le plus de chance de nous apporter des résultats contrastés.

Cependant, il ne serait pas moins nécessaire d'aller plus loin et de s'intéresser aux autres comparaisons moins exploitées dans cette étude (A vs B ; A vs C ; B vs C ; B vs D ; C vs D).

De nombreuses autres études pourraient donc voir le jour dans un futur proche et permettront peut-être de répondre aux nombreuses questions qu'a soulevées notre travail.

Dans tous les cas, sachant que le profil morphocinétique des embryons bovins a des répercussions sur la transition materno-embryonnaire, dorénavant, ce paramètre devra être pris en compte pour tenter d'éliminer les biais qui lui sont potentiellement associés et pour affiner les connaissances autour de ce sujet.

Publications

51 Bovine embryos with distinct early morphokinetic pathways present different post-embryonic genome activation transcriptomic patterns and different cryotolerance

A. P. Reis^A, A. Jampy^D, A. Teste^{A,E}, B. M. Le Guienne^C, L. Laffont^A, S. Ruffini^A, E. Canon^A, C. Archilla^A, L. Jouneau^A, A. Trubuil^B, and V. Duranton^A

^AUMR BDR, INRA, ENVA, Université Paris Saclay, Jouy-en-Josas, France;

^BMaiAge, INRA, Jouy-en-Josas, France;

^CAlice, Recherche et Développement, Jouy-en-Josas, France;

^DAgroParisTech, Paris, France;

^EÉcole Nationale Vétérinaire d'Alfort, Maisons Alfort, France

We recently developed and validated a methodology based on early morphokinetics to predict 6 categories of bovine embryos, including 4 categories of blastocysts: early hatching blastocysts (EHB), hatching blastocysts (HB), low hatching blastocysts (LHB), and arrhythmic blastocysts (AB) (Reis *et al.* 2018 Anim. Reprod. 15, 601). We hypothesised that different early morphokinetic pathways could (1) be accompanied by different biological patterns (EGA, etc.) and (2) influence interaction with environmental constraints (cryopreservation, etc.). The objective of this study was (1) to investigate whether transcriptomic differences between the predicted EHB, HB, LHB, or AB could be observed immediately after the fourth embryonic cycle (just after EGA); (2) to assess cryotolerance of these morphokinetic categories. Time lapses were produced during 4 days post *in vitro* insemination (dpi) (Reis *et al.* 2018). Study 1: 128 embryos having finished the fourth embryonic cycle were individually dry frozen at 4.7 dpi, classified as EHB, HB, LHB, or AB, and pooled into four batches of eight embryos/category (total 16 samples). Total RNA was extracted, amplified using the SMARTseq V4 ultra low input kit (Clontech), libraries were prepared using the Nextera XT Illumina library preparation kit, and sequenced (paired-end 50-34 bp) on an Illumina NextSeqn 500 instrument. Identification of differentially expressed genes was achieved using Limma package (R). The *P*-values were adjusted using the Benjamini and Hochberg false discovery rate (Saenz-de-Juano *et al.* 2014). Fold change >2 or <0.5 and adjusted *P* < 0.05 were considered significant. Enrichment analysis was performed using the Web server enrichR. Study 2: 90 embryos showing signs of expansion (diameter >135 µm and -2 µm at zona pellucida) were slow frozen at 6.2 or 7 dpi, classified as EHB (*n* = 8), HB (*n* = 29), LHB (*n* = 32), or AB (*n* = 21), and further thawed and cultured *in vitro* for 72h. Cryotolerance was evaluated at 24 and 72 h post-thawing. Statistical analysis was performed using chi-square or Fisher test. Study 1: the AB category presented 258, 699, and 899 upregulated genes and 40, 62, and 60 downregulated in comparison to LHB, HB, and EHB, respectively. The functional comparison of the two extremes (EHB vs. AB), at 4.7 dpi, showed that AB embryos present intense transcription regulation activity. The monocarboxylic acid binding pathway was upregulated in EHB. Study 2: the post-thawing *in vitro* survival was similar between EHB, HB, LHB, and AB (50%, 34.5%, 31.3%, and 42.9%, respectively). The hatching/alive ratio in EHB category tended to be lower than AB (1/4 vs. 8/9; *P* = 0.052). Frozen/thawed EHB and HB presented reduced hatching rates, whereas LHB and AB significantly increased the hatching potential after thawing compared with their not-frozen counterparts of the learning database (25% vs. 100%; 50% vs. 100%; 70% vs. 11%; 88.9% vs. 18%, respectively). Morphokinetic prediction highlighted a new classification of bovine embryos with different biological features. This could contribute to improve practice on IVF embryo transfer taking into account these differences.

IETS 2020 abstracts. (2020) . *Reprod. Fertil. Dev.* 32(2), 125-244

Bovine embryos with distinct early morphokinetic pathways present different post EGA transcriptomic patterns and different cryotolerance.

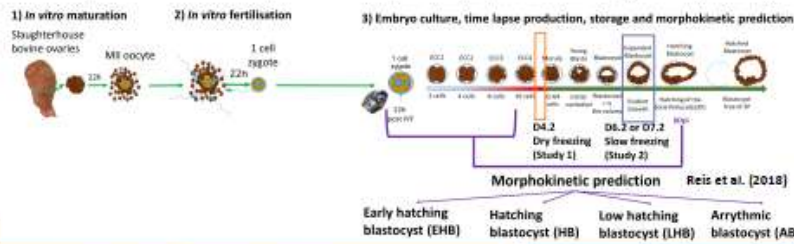
A. P. Reis^{1*}, A. Jampy¹, A. Teste¹, B. M. Le Guienne³, L. Laffont⁴, S. Ruffini¹, E. Canon¹, C. Archilla¹, L. Jouneau¹, A. Trubuil², V. Duranthon¹

Introduction :

We recently developed and validated a methodology based on the early morphokinetic to predict 6 categories of bovine embryos, including 4 categories of blastocysts: Early Hatching Blastocysts (EHB), Hatching Blastocysts (HB), Low Hatching Blastocysts (LHB) and Arrhythmic Blastocysts (AB) (Reis et al. 2018). The objectives of this study were 1) to investigate whether transcriptomic differences between the predicted EHB, HB, LHB or AB could be observed immediately after the completion of the 4th embryonic cycle – i.e. ECC4 completed, 16-32 cells (just after EGA); 2) to assess cryotolerance of these morphokinetic categories.

Material and methods:

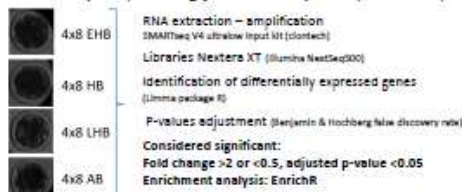
Figure 1: General description of material and methods: production of embryos, time-lapses, prediction and freezing time to explore transcriptomic profiles and cryotolerance of EHB, HB, LHB, AB.



Study 1

Figure 1.1: Material and methods (continuation – Study 1)

At 4.2dpi: dry freezing if ECC4 is completed (16-32 cells)



Results

Figure 2.1: Comparison of the post EGA (ECC4 completed) transcriptomic profiles of EHB (A), HB (B), LHB (C) and AB (D). Descriptive comparison of the 16 samples: Euclidean clustering (left) and PCA (center). Quantitative comparison of the number of genes over expressed on AB (D) compared to the three other categories: Venn diagram (right).

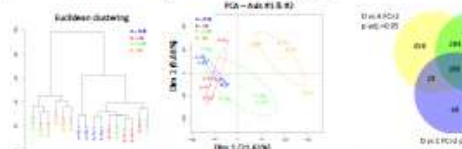


Table 1.1: Molecular function and biological process pathways differentially expressed between AB (D) and EHB (A).

Overexpressed on AB (D) compared to EHB (A)	Biological Process
	Transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific binding (GO:0002238)
	Positive regulation of transcription from RNA polymerase II promoter (GO:0045444)
	Regulation of transcription from RNA polymerase II promoter (GO:0006357)
	Positive regulation of transcription, DNA-templated (GO:0043889)
	Positive regulation of nucleic acid-templated transcription (GO:1909508)
	Negative regulation of transcription DNA-templated (GO:0045892)
	RNA polymerase II regulatory region sequence-specific DNA binding (GO:0000977)
	Transcription regulatory region DNA binding (GO:0044232)
	Regulation of protein serine/threonine kinase activity (GO:0071300)
	Acron guidance (GO:0007411)
	Negative regulation of immune response (GO:0050777)
Overexpressed on AB (A) compared to AB (D)	Monocarboxylic acid binding (GO:0032329)

At this stage (ECC4 completed, i.e. 16-32 cells) AB embryos present intense transcription regulation activity.

Study 2

Figure 1.2: Material and methods (continuation – Study 2)

At 6.2 or 7.2 dpi: slow freezing (Slow freezing (97%EG + 3% Sucrose) if inner diameter $\geq 135\mu\text{m}$ and ZP $\sim 2\mu\text{m}$)



Results

Table 1.2: In vitro cryotolerance of slow-frozen / thawed blastocysts of 4 morphokinetic categories (EHB, HB, LHB, AB).

Category	N (replicates)	Survival / thawed 24h n(%)	Survival / thawed 72h n(%)	Hatching / alive 72h n(%)
EHB	8 (2)	5 (62.5)	4 (50.0)	1 (25.0) ^a
HB	29 (8)	15 (51.7)	10 (34.5)	5 (50.0)
LHB	32 (8)	16 (50.0)	10 (31.3)	7 (70.0)
AB	21 (7)	12 (57.1)	9 (42.9)	8 (88.9) ^b

a vs b, p=0.052

The post-thawing in vitro survival was similar between EHB, HB, LHB and AB. The hatching/alive ratio in EHB category tended to be lower than AB (p=0.032). Frozen/thawed EHB and HB presented reduced hatching rates while LHB and AB significantly increased the hatching potential after thawing compared to their not frozen counterparts of the learning database (25% vs 100%; 50% vs 100%; 70% vs 11%; 88.9% vs 18%, respectively).

Conclusion:

Morphokinetic prediction highlighted a new classification of bovine embryos with different biological features. It could contribute to improve practice on IVF embryo transfer taking into account these differences.

¹Université Paris Saclay, INRAE, ENVA, BREID, 78350, Jouy-en-Josas, France

²INRAE, Unité MolAGE, Centre INRAE - Domaine de Vilvert - Bat 133 - 78350 Jouy-en-Josas, FRANCE

³Allice, Recherche et Développement, F-78350, Jouy-en-Josas, France

* co-senior author

*corresponding author



Domaine de Vilvert
78350 Jouy-en-Josas
alice.reis@inrae.fr

Liste des références bibliographiques

- ABBOTT A.L., DUCIBELLA T. (2001) Calcium and the control of mammalian cortical granule exocytosis. *Front. Biosci.* 6, D792-806
- ABDELHALEEM M., MALTAIS L., WAIN H. (2003) The human DDX and DHX gene families of putative RNA helicases. *Genomics* 81(6), 618-622
- ADACHI K., SOETA-SANEYOSHI C., SAGARA H., IWAKURA Y. (2007) Crucial role of Bysl in mammalian preimplantation development as an integral factor for 40S ribosome biogenesis. *Mol. Cell. Biol.* 27(6), 2202-2214
- AOKI F., WORRAD D.M., SCHULTZ R.M. (1997) Regulation of Transcriptional Activity during the First and Second Cell Cycles in the Preimplantation Mouse Embryo. *Dev Biol* 181(2), 296-307
- BARCKMANN B., SIMONELIG M. (2013) Control of maternal mRNA stability in germ cells and early embryos. *BBA-Gene Regul Mech* 1829(6-7), 714-724
- BARNEA E., BERGMAN Y. (2000) Synergy of SF1 and RAR in Activation of Oct-3/4 Promoter. *J. Biol. Chem.* 275(9), 6608-6619
- BARNES F.L., FIRST N.L. (1991) Embryonic transcription in in vitro cultured bovine embryos. *Mol. Reprod. Dev.* 29(2), 117-123
- BEDZHOV I., LISZEWSKA E., KANZLER B., STEMMLER M.P. (2012) Igf1r Signaling Is Indispensable for Preimplantation Development and Is Activated via a Novel Function of E-cadherin. *PLOS Genetics* 8(3), e1002609
- BERG D.K., SMITH C.S., PEARTON D.J., *et al.* (2011) Trophectoderm lineage determination in cattle. *Dev. Cell* 20(2), 244-255
- BERGER S.L. (2001) An embarrassment of niches: the many covalent modifications of histones in transcriptional regulation. *Oncogene* 20(24), 3007-3013
- BERNSTEIN E., KIM S.Y., CARMELL M.A., *et al.* (2003) Dicer is essential for mouse development. *Nat Genet* 35(3), 215-217
- BESENFELDER U., BREM G., HAVLICEK V. (2020) Review: Environmental impact on early embryonic development in the bovine species. *Animal* 14(S1), s103-s112
- BINNS D., DIMMER E., HUNTLEY R., *et al.* (2009) QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics* 25(22), 3045-3046
- BIRD A.P., WOLFFE A.P. (1999) Methylation-Induced Repression— Belts, Braces, and Chromatin. *Cell* 99(5), 451-454

- BRAUDE P., PELHAM H., FLACH G., LOBATTO R. (1979) Post-transcriptional control in the early mouse embryo. *Nature* 282(5734), 102-105
- C ARRESEIGOR, Y FILIPIAK, S GONZÁLEZ, *et al.* (2016) Pregnancy rates and embryo mortality from different cattle breeds in vitro produced embryos in a commercial program, In *18th International Congress on Animal Reproduction (ICAR)*, Tours, du 28 au 30 juin 2016
- CARTHEW R.W., SONTHEIMER E.J. (2009) Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136(4), 642-655
- CHEN E.Y., TAN C.M., KOU Y., *et al.* (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 14, 128
- CLAVERÍA C., GIOVINAZZO G., SIERRA R., TORRES M. (2013) Myc-driven endogenous cell competition in the early mammalian embryo. *Nature* 500(7460), 39-44
- DAUTRY F., RIBET C. (2004) L'interférence par l'ARN : vers une génomique fonctionnelle chez les mammifères ? *Med Sci (Paris)* 20(8-9), 815-819
- DE LA SERNA I.L., CARLSON K.A., IMBALZANO A.N. (2001) Mammalian SWI/SNF complexes promote MyoD-mediated muscle differentiation. *Nat. Genet.* 27(2), 187-190
- DE RENZIS S., ELEMENTO O., TAVAZOIE S., WIESCHAUS E.F. (2007) Unmasking activation of the zygotic genome using chromosomal deletions in the Drosophila embryo. *PLoS Biol.* 5(5), e117
- DEAN W., SANTOS F., REIK W. (2003) Epigenetic reprogramming in early mammalian development and following somatic nuclear transfer. *Semin Cell Dev Biol* 14(1), 93-100
- DEKEL N. (1995) Molecular control of meiosis. *Trends Endocrin Met* 6(5), 165-169
- DYE M.J., GROMAK N., PROUDFOOT N.J. (2006) Exon Tethering in Transcription by RNA Polymerase II. *Molecular Cell* 21(6), 849-859
- EDGAR B.A., SCHUBIGER G. (1986) Parameters controlling transcriptional activation during early Drosophila development. *Cell* 44(6), 871-877
- EL-MESTRAH M., CASTLE P.E., BOROSSA G., KAN F.W.K. (2002) Subcellular Distribution of ZP1, ZP2, and ZP3 Glycoproteins During Folliculogenesis and Demonstration of Their Topographical Disposition Within the Zona Matrix of Mouse Ovarian Oocytes1. *Biol Reprod* 66(4), 866-876
- FANG X., YOON J.-G., LI L., *et al.* (2011) Landscape of the SOX2 protein-protein interactome. *Proteomics* 11(5), 921-934
- FERRARO E., CECCONI F. (2007) Autophagic and apoptotic response to stress signals in mammalian cells. *Arch. Biochem. Biophys.* 462(2), 210-219
- FORJAS DE BORJA P.F. d. (2001) Cyclin A-CDK phosphorylates Sp1 and enhances Sp1-mediated transcription. *The EMBO Journal* 20(20), 5737-5747

- FRANZÈ A., ARCHIDIACONO N., ROCCHI M., MARINO M., GRIMALDI G. (1991) Isolation and expression analysis of a human zinc finger gene (ZNF41) located on the short arm of the X chromosome. *Genomics* 9(4), 728-736
- FREI R.E., SCHULTZ G.A., CHURCH R.B. (1989) Qualitative and quantitative changes in protein synthesis occur at the 8–16-cell stage of embryogenesis in the cow. *Reproduction* 86(2), 637-641
- GAO L., WU K., LIU Z., *et al.* (2018) Chromatin Accessibility Landscape in Human Early Embryos and Its Association with Evolution. *Cell* 173(1), 248-259.e15
- GERBER A.P., LUSCHNIG S., KRASNOW M.A., BROWN P.O., HERSCHLAG D. (2006) Genome-wide identification of mRNAs associated with the translational regulator PUMILIO in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 103(12), 4487-4492
- GHASSEMIFAR M.R., ECKERT J.J., HOUGHTON F.D., *et al.* (2003) Gene expression regulating epithelial intercellular junction biogenesis during human blastocyst development in vitro. *Mol. Hum. Reprod.* 9(5), 245-252
- GLASER S., SCHAFT J., LUBITZ S., *et al.* (2006) Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* 133(8), 1423-1432
- GONZÁLEZ B., DENZEL S., MACK B., CONRAD M., GIRES O. (2009) EpCAM is involved in maintenance of the murine embryonic stem cell phenotype. *Stem Cells* 27(8), 1782-1791
- GRAF A., KREBS S., ZAKHARTCHENKO V., *et al.* (2014) Fine mapping of genome activation in bovine embryos by RNA sequencing. *P Natl Acad Sci* 111(11), 4139-4144
- GRAF Alexander, KREBS S., HEININEN-BROWN M., *et al.* (2014) Genome activation in bovine embryos: Review of the literature and new insights from RNA sequencing experiments. *Anim Reprod Sci* 149(1-2), 46-58
- GRAY R.S., ABITUA P.B., WLODARCZYK B.J., *et al.* (2009) The planar cell polarity effector Fuz is essential for targeted membrane trafficking, ciliogenesis, and mouse embryonic development. *Nat Cell Biol* 11(10), 1225-1232
- HANNON G.J. (2002) RNA interference. *Nature* 418(6894), 244-251
- HANZEN C., DRION P., LOURTIE O., DEPIERREUX C., CHRISTIANS E. (1999) La mortalité embryonnaire. 1. Aspects cliniques et facteurs étiologiques dans l'espèce bovine. *Ann Med Vet* 143
- HARA K.T., ODA S., NAITO K., *et al.* (2005) Cyclin A2-CDK2 regulates embryonic gene activation in 1-cell mouse embryos. *Dev Biol* 286(1), 102-113
- HARBOUR J.W., DEAN (2000) The Rb/E2F pathway: expanding roles and emerging paradigms. *Gene Dev* 14(19), 2393-2409
- HATCH C.L., BONNER W.M. (1990) The human histone H2A.Z gene. Sequence and regulation. *J. Biol. Chem.* 265(25), 15211-15218

- HENERY C.C., MIRANDA M., WIEKOWSKI M., WILMUT I., DEPAMPHILIS M.L. (1995) Repression of gene expression at the beginning of mouse development. *Dev. Biol.* 169(2), 448-460
- HINNEBUSCH A.G. (2006) eIF3: a versatile scaffold for translation initiation complexes. *Trends Biochem. Sci.* 31(10), 553-562
- HOLM P., SHUKRI N.N., VAJTA G., *et al.* (1998) Developmental kinetics of the first cell cycles of bovine in vitro produced embryos in relation to their in vitro viability and sex. *Theriogenology* 50(8), 1285-1299
- HU G., CUI K., NORTHRUP D., *et al.* (2013) H2A.Z facilitates access of active and repressive complexes to chromatin in embryonic stem cell self-renewal and differentiation. *Cell Stem Cell* 12(2), 180-192
- HUMPHREYS D.T., WESTMAN B.J., MARTIN D.I.K., PREISS T. (2005) MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc Natl Acad Sci U S A* 102(47), 16961-16966
- HUNTZINGER E., IZAURRALDE E. (2011) Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet* 12(2), 99-110
- HYTTEL P., FAIR T., CALLESEN H., GREVE T. (1997) Oocyte growth, capacitation and final maturation in cattle. *Theriogenology* 47(1), 23-32
- JASPARD E. (2012) L'épissage des ARN messager pré-matures : le spliceosome. [<http://biochimej.univ-angers.fr/Page2/COURS/3CoursdeBiochSTRUCT/8Spliceosome/1Spliceosome.htm>] (consulté le 06/06/2020).
- JONCKHEERE A.I., SMEITINK J.A.M., RODENBURG R.J.T. (2012) Mitochondrial ATP synthase: architecture, function and pathology. *J. Inherit. Metab. Dis.* 35(2), 211-225
- JONES K.T. (2004) Turning it on and off: M-phase promoting factor during meiotic maturation and fertilization. *Mol Hum Reprod* 10(1), 1-5
- KAŇKA J., KEPKOVÁ K., NĚMCOVÁ L. (2009) Gene expression during minor genome activation in preimplantation bovine development. *Theriogenology* 72(4), 572-583
- KARANTZALI E., LEKAKIS V., IOANNOU M., *et al.* (2011) Sall1 regulates embryonic stem cell differentiation in association with nanog. *J. Biol. Chem.* 286(2), 1037-1045
- KATZ J.P., PERREAULT N., GOLDSTEIN B.G., *et al.* (2002) The zinc-finger transcription factor Klf4 is required for terminal differentiation of goblet cells in the colon. *Development* 129(11), 2619-2628
- KENT L.N., RUMI M.A.K., KUBOTA K., LEE D.-S., SOARES M.J. (2011) FOSL1 is integral to establishing the maternal-fetal interface. *Mol. Cell. Biol.* 31(23), 4801-4813
- KHRAIWESH B., ZHU J.-K., ZHU J. (2012) Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim Biophys Acta* 1819(2), 137-148

- KIDDER B.L., PALMER S. (2010) Examination of transcriptional networks reveals an important role for TCFAP2C, SMARCA4, and EOMES in trophoblast stem cell maintenance. *Genome Res.* 20(4), 458-472
- KIMELMAN D., KIRSCHNER M., SCHERSON T. (1987) The events of the midblastula transition in *Xenopus* are regulated by changes in the cell cycle. *Cell* 48(3), 399-407
- KINOSHITA K., URA H., AKAGI T., *et al.* (2007) GABPalpha regulates Oct-3/4 expression in mouse embryonic stem cells. *Biochem. Biophys. Res. Commun.* 353(3), 686-691
- KORNBERG R.D. (2007) The molecular basis of eukaryotic transcription. *Proc Natl Acad Sci USA* 104(32), 12955-12961
- KOZIOL M.J., GARRETT N., GURDON J.B. (2007) Tpt1 activates transcription of oct4 and nanog in transplanted somatic nuclei. *Curr. Biol.* 17(9), 801-807
- KUEHNER J.N., PEARSON E.L., MOORE C. (2011) Unravelling the means to an end: RNA polymerase II transcription termination. *Nat Rev Mol Cell Biol* 12(5), 283-294
- KULESHOV M.V., JONES M.R., ROUILLARD A.D., *et al.* (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 44(Web Server issue), W90-W97
- KUROSAKA S., ECKARDT S., MCLAUGHLIN K.J. (2004) Pluripotent lineage definition in bovine embryos by Oct4 transcript localization. *Biol. Reprod.* 71(5), 1578-1582
- LADUNGA I. (Éd.) (2010) Computational Biology of Transcription Factor Binding, *Methods in Molecular Biology*. Lincoln, Humana Press
- LAI W.S., CARBALLO E., THORN J.M., KENNINGTON E.A., BLACKSHEAR P.J. (2000) Interactions of CCCH Zinc Finger Proteins with mRNA binding of tristetraprolin-related zinc finger proteins to AU-rich elements and destabilization of mRNA. *J. Biol. Chem.* 275(23), 17827-17837
- LANDSBERG M.J., MORAN-JONES K., SMITH R. (2006) Molecular recognition of an RNA trafficking element by heterogeneous nuclear ribonucleoprotein A2. *Biochemistry* 45(12), 3943-3951
- LATHAM K.E., SCHULTZ R.M. (2001) Embryonic genome activation. *Front. Biosci.* 6, D748-759
- LE BRUSQ N. (2018) Élaboration d'un guide de lecture standardisée de paramètres morphocinétiques précoces pour la sélection d'embryons bovins produits in vitro en vue du transfert. Thèse Méd. Vét. Ecole nationale vétérinaire d'Alfort
- LEE J., ZHOU P. (2007) DCAFs, the missing link of the CUL4-DDB1 ubiquitin ligase. *Mol. Cell* 26(6), 775-780
- LEQUARRE A.S., TRAVERSO J.M., MARCHANDISE J., DONNAY I. (2004) Poly(A) RNA Is Reduced by Half During Bovine Oocyte Maturation but Increases when Meiotic Arrest Is Maintained with CDK Inhibitors1. *Biol Reprod* 71(2), 425-431

- LEROUX C., TOSSER-KLOPP G. (2000) La fonction du gène : les grandes étapes de l'utilisation de l'information génétique. *Productions Animales (Hors Série 2000)*, 21-28. (2000)
- LEVINE M. (2011) Paused RNA Polymerase II as a Developmental Checkpoint. *Cell* 145(4), 502-511
- LI E., BESTOR T.H., JAENISCH R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69(6), 915-926
- LI C.M., HOU L., ZHANG H., ZHANG W.Y. (2014) Induces Trophoblast Migration and Invasion by Regulating Matrix Metalloproteinase and Integrin Expression in Human First-Trimester Placenta. *Reprod Sci*, published online ahead of print, 2014 Jan 8. *Reprod Sci*. 2014;doi:10.1177/1933719113519170 (consulté le 06/06/2020)
- LI L., WANG B., ZHANG W., *et al.* (2017) A homozygous NOBOX truncating variant causes defective transcriptional activation and leads to primary ovarian insufficiency. *Hum. Reprod.* 32(1), 248-255
- LINDNER G.M., WRIGHT R.W. (1983) Bovine embryo morphology and evaluation. *Theriogenology* 20(4), 407-416
- LOGAN J., FALCK-PEDERSEN E., DARNELL J.E., SHENK T. (1987) A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. *Proc Natl Acad Sci U S A* 84(23), 8306-8310
- LU R., YANG A., JIN Y. (2011) Dual functions of T-box 3 (Tbx3) in the control of self-renewal and extraembryonic endoderm differentiation in mouse embryonic stem cells. *J. Biol. Chem.* 286(10), 8425-8436
- LUND E., LIU M., HARTLEY R.S., SHEETS M.D., DAHLBERG J.E. (2009) Deadenylation of maternal mRNAs mediated by miR-427 in *Xenopus laevis* embryos. *RNA* 15(12), 2351-2363
- MA J., FLEMR M., STEIN P., *et al.* (2010) microRNA activity is suppressed in mouse oocytes. *Curr Biol* 20(3), 265-270
- MA P., LIN S., BARTOLOMEI M.S., SCHULTZ R.M. (2010) Metastasis Tumor Antigen 2 (MTA2) Is Involved in Proper Imprinted Expression of H19 and Peg3 During Mouse Preimplantation Development. *Biol Reprod* 83(6), 1027-1035
- MALIK S., ROEDER R.G. (2005) Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem Sci* 30(5), 256-263
- MANEJWALA F.M., LOGAN C.Y., SCHULTZ R.M. (1991) Regulation of hsp70 mRNA levels during oocyte maturation and zygotic gene activation in the mouse. *Dev Biol* 144(2), 301-308
- MANTOVANI R., PESSARA U., TRONCHE F., *et al.* (1992) Monoclonal antibodies to NF-Y define its function in MHC class II and albumin gene transcription. *EMBO J.* 11(9), 3315-3322
- MARFELLA C.G.A., OHKAWA Y., COLES A.H., *et al.* (2006) Mutation of the SNF2 family member Chd2 affects mouse development and survival. *J. Cell. Physiol.* 209(1), 162-171

- MARQUANT-LE GUIENNE B. (1998) Petit atlas de l'ovocyte bovin. *Elevage Insémin*, 288, 24-30.
- MEISTRICH M.L., MOHAPATRA B., SHIRLEY C.R., ZHAO M. (2003) Roles of transition nuclear proteins in spermiogenesis. *Chromosoma* 111(8), 483-488
- MEMILI E., DOMINKO T., FIRST N.L. (1998) Onset of transcription in bovine oocytes and preimplantation embryos. *Mol. Reprod. Dev.* 51(1), 36-41
- MEMILI E., FIRST N.L. (1998) Developmental changes in RNA polymerase II in bovine oocytes, early embryos, and effect of alpha-amanitin on embryo development. *Mol. Reprod. Dev.* 51(4), 381-389
- MEMILI E., FIRST N.L. (1999) Control of Gene Expression at the Onset of Bovine Embryonic Development1. *Biol Reprod* 61(5), 1198-1207
- MERMILLOD P., MARCHAL R. (1999) La maturation de l'ovocyte de mammifères. *Med Sci (Paris)* 15(2), 148-156
- MISIRLIOGLU M., PAGE G.P., SAGIRKAYA H., *et al.* (2006) Dynamics of global transcriptome in bovine matured oocytes and preimplantation embryos. *P Natl Acad Sci* 103(50), 18905-18910
- MONDOU E., DUFORT I., GOHIN M., FOURNIER E., SIRARD M.-A. (2012) Analysis of microRNAs and their precursors in bovine early embryonic development. *MHR: Basic science of reproductive medicine* 18(9), 425-434
- MU W., MUNROE R.J., BARKER A.K., SCHIMENTI J.C. (2010) PDCD2 is essential for inner cell mass development and embryonic stem cell maintenance. *Dev. Biol.* 347(2), 279-288
- MUELLER O.M., SCHROEDER A., LIGHTFOOT S., SALOWSKY R., STOCKER S. (2004) RNA integrity number (RIN) - towards standardization of RNA quality assessment. *Cancer Res* 64(7 Supplement), 380-380
- MUÑOZ M., UYAR A., CORREIA E., *et al.* (2014) Prediction of pregnancy viability in bovine in vitro-produced embryos and recipient plasma with Fourier transform infrared spectroscopy. *J Dairy Sci* 97(9), 5497-5507
- MURCHISON E.P., STEIN P., XUAN Z., *et al.* (2007) Critical roles for Dicer in the female germline. *Genes Dev* 21(6), 682-693
- NAKATAKE Y., FUKUI N., IWAMATSU Y., *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
- NARLIKAR G.J., FAN H.-Y., KINGSTON R.E. (2002) Cooperation between Complexes that Regulate Chromatin Structure and Transcription. *Cell* 108(4), 475-487
- NISHIOKA N., YAMAMOTO S., KIYONARI H., *et al.* (2008) Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. *Mech. Dev.* 125(3-4), 270-283
- NIWA H., MIYAZAKI J., SMITH A.G. (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* 24(4), 372-376

- NIWA H., TOYOOKA Y., SHIMOSATO D., *et al.* (2005) Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* 123(5), 917-929
- NOTTROTT S., SIMARD M.J., RICHTER J.D. (2006) Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nat Struct Mol Biol* 13(12), 1108-1114
- OKANO M., BELL D.W., HABER D.A., LI E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99(3), 247-257
- PABO C.O., SAUER R.T. (1992) TRANSCRIPTION FACTORS: Structural Families and Principles of DNA Recognition. *Annu Rev Biochem* 61(1), 1053-1095
- PARRISH J.J., SUSKO-PARRISH J.L., LEIBFRIED-RUTLEDGE M.L., *et al.* (1986) Bovine in vitro fertilization with frozen-thawed semen. *Theriogenology* 25(4), 591-600
- PETERSEN C.P., BORDELEAU M.-E., PELLETIER J., SHARP P.A. (2006) Short RNAs Repress Translation after Initiation in Mammalian Cells. *Molecular Cell* 21(4), 533-542
- PICTON H.M. (2001) Activation of follicle development: the primordial follicle. *Theriogenology* 55(6), 1193-1210
- PIQUÉ M., LÓPEZ J.M., FOISSAC S., GUIGÓ R., MÉNDEZ R. (2008) A combinatorial code for CPE-mediated translational control. *Cell* 132(3), 434-448
- PONSART C., MARQUANT-LEGUIENNE B., HUMBLLOT P. (2004) Les biotechnologies de l'embryon bovin, évolution et perspectives., p.8
- PONTES J.H.F., NONATO-JUNIOR I., SANCHES B.V., *et al.* (2009) Comparison of embryo yield and pregnancy rate between in vivo and in vitro methods in the same Nelore (*Bos indicus*) donor cows. *Theriogenology* 71(4), 690-697
- PONTES J.H.F., MELO STERZA F.A., BASSO A.C., *et al.* (2011) Ovum pick up, in vitro embryo production, and pregnancy rates from a large-scale commercial program using Nelore cattle (*Bos indicus*) donors. *Theriogenology* 75(9), 1640-1646
- POUEYMIROU W.T., SCHULTZ R.M. (1989) Regulation of mouse preimplantation development: Inhibition of synthesis of proteins in the two-cell embryo that require transcription by inhibitors of cAMP-dependent protein kinase. *Dev Biol* 133(2), 588-599
- RALSTON A., COX B.J., NISHIOKA N., *et al.* (2010) Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. *Development* 137(3), 395-403
- REIS A.P., BROCARD G., BELGHITI M., *et al.* (2018) 40 Toward a standardised annotation of morphokinetic parameters for an automatic early prediction of the in vitro development potential of bovine embryos. *Reprod. Fertil. Dev.* 31(1), 146-146
- REYES J.M., ROSS P.J. (2016) Cytoplasmic polyadenylation in mammalian oocyte maturation: Oocyte cytoplasmic polyadenylation. *WIREs RNA* 7(1), 71-89

- RICHARD D., CHEVALET P., FOURNEL S., *et al.* (2012) *Biologie*, 2ème éd. Paris, Dunod
- RICHTER J.D. (1999) Cytoplasmic Polyadenylation in Development and Beyond. *Microbiol Mol Biol Rev* 63(2), 446-456
- RICHTER J.D. (2008) Breaking the Code of Polyadenylation-Induced Translation. *Cell* 132(3), 335-337
- RUDDOCK-D'CRUZ N.T., XUE J., WILSON K.J., *et al.* (2008) Dynamic changes in the localization of five members of the methyl binding domain (MBD) gene family during murine and bovine preimplantation embryo development. *Mol. Reprod. Dev.* 75(1), 48-59
- RULAND J., SIRARD C., ELIA A., *et al.* (2001) p53 accumulation, defective cell proliferation, and early embryonic lethality in mice lacking tsg101. *Proc. Natl. Acad. Sci. U.S.A.* 98(4), 1859-1864
- SAINT-DIZIER M., CHASTANT-MAILLARD S. (2014) *La reproduction animale et humaine, Synthèses*. Versailles, Editions Quae
- SALILEW-WONDIM D., SAEED-ZIDANE M., HOELKER M., *et al.* (2018) Genome-wide DNA methylation patterns of bovine blastocysts derived from in vivo embryos subjected to in vitro culture before, during or after embryonic genome activation. *BMC Genomics* 19
- SANTOS F., HENDRICH B., REIK W., DEAN W. (2002) Dynamic Reprogramming of DNA Methylation in the Early Mouse Embryo. *Dev Biol* 241(1), 172-182
- SCHIER A.F. (2007) The maternal-zygotic transition: death and birth of RNAs. *Science* 316(5823), 406-407
- SHUMAN S. (2000) Structure, mechanism, and evolution of the mRNA capping apparatus. *In Progress in Nucleic Acid Research and Molecular Biology*. Academic Press 66, 1-40
- SIMON M., JÍLEK F., FULKA J. (1989) Effect of cycloheximide upon maturation of bovine oocytes. *Reprod. Nutr. Dévelop.* 29(5), 533-540
- SIRARD M.-A., RICHARD F., BLONDIN P., ROBERT C. (2006) Contribution of the oocyte to embryo quality. *Theriogenology* 65(1), 126-136
- SIRARD M.A. (2010) Activation of the embryonic genome. *Soc Reprod Fertil Suppl* 67, 145-158
- SOMMERVILLE J. (1999) Activities of cold-shock domain proteins in translation control. *Bioessays* 21(4), 319-325
- STELZER G., ROSEN N., PLASCHKES I., *et al.* (2016) The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr Protoc Bioinformatics* 54, 1.30.1-1.30.33
- SUN Q.-Y. (2003) Cellular and molecular mechanisms leading to cortical reaction and polyspermy block in mammalian eggs. *Microsc. Res. Tech.* 61(4), 342-348

- TADROS W., GOLDMAN A.L., BABAK T., *et al.* (2007) SMAUG Is a Major Regulator of Maternal mRNA Destabilization in *Drosophila* and Its Translation Is Activated by the PAN GU Kinase. *Developmental Cell* 12(1), 143-155
- TAMM C., BÖWER N., ANNERÉN C. (2011) Regulation of mouse embryonic stem cell self-renewal by a Yes-YAP-TEAD2 signaling pathway downstream of LIF. *J. Cell. Sci.* 124(Pt 7), 1136-1144
- TAN M.H., AU K.F., LEONG D.E., *et al.* (2013) An Oct4-Sall4-Nanog network controls developmental progression in the pre-implantation mouse embryo. *Mol Syst Biol* 9, 632
- TELFORD N.A., WATSON A.J., SCHULTZ G.A. (1990) Transition from maternal to embryonic control in early mammalian development: A comparison of several species. *Mol. Reprod. Dev.* 26(1), 90-100
- TESFAYE D., WORKU D., RINGS F., *et al.* (2009) Identification and expression profiling of microRNAs during bovine oocyte maturation using heterologous approach. *Mol Reprod Dev* , 13
- TOMEK W., TORNER H., KANITZ W. (2002) Comparative Analysis of Protein Synthesis, Transcription and Cytoplasmic Polyadenylation of mRNA during Maturation of Bovine Oocytes in vitro. *Reprod Domest Anim* 37(2), 86-91
- TRIPURANI S.K., XIAO C., SALEM M., YAO J. (2010) Cloning and analysis of fetal ovary microRNAs in cattle. *Anim Reprod Sci*, 7
- TRIPURANI S.K., LEE K.-B., WEE G., SMITH G.W., YAO J. (2011) MicroRNA-196a regulates bovine newborn ovary homeobox gene (NOBOX) expression during early embryogenesis. *BMC Dev Biol* 11, 25
- TRIPURANI S.K., WEE G., LEE K.-B., *et al.* (2013) MicroRNA-212 Post-Transcriptionally Regulates Oocyte-Specific Basic-Helix-Loop-Helix Transcription Factor, Factor in the Germline Alpha (FIGLA), during Bovine Early Embryogenesis. *PLoS One* 8(9), e76114
- TROTTER K.W., ARCHER T.K. (2007) Nuclear receptors and chromatin remodeling machinery. *Mol Cell Endocrinol* 265-266, 162-167
- TSUKAMOTO S., KUMA A., MURAKAMI M., *et al.* (2008) Autophagy Is Essential for Preimplantation Development of Mouse Embryos. *Science* 321(5885), 117-120
- UZBEKOVA S., ARLOT-BONNEMAINS Y., DUPONT J., *et al.* (2008) Spatio-temporal expression patterns of aurora kinases a, B, and C and cytoplasmic polyadenylation-element-binding protein in bovine oocytes during meiotic maturation. *Biol. Reprod.* 78(2), 218-233
- VAN DEN HURK R., ZHAO J. (2005) Formation of mammalian oocytes and their growth, differentiation and maturation within ovarian follicles. *Theriogenology* 63(6), 1717-1751
- VAN VLIET J., CROFTS L.A., QUINLAN K.G.R., *et al.* (2006) Human KLF17 is a new member of the Sp/KLF family of transcription factors. *Genomics* 87(4), 474-482

- VENTERS B.J., PUGH B.F. (2009) How eukaryotic genes are transcribed. *Crit Rev Biochem Mol Biol* 44(2-3), 117-141
- VIGNEAULT C., GRAVEL C., VALLÉE M., MCGRAW S., SIRARD M.-A. (2009) Unveiling the bovine embryo transcriptome during the maternal-to-embryonic transition. *Reproduction* 137(2), 245-257
- VIUFF D., HYTTEL P., AVERY B., *et al.* (1998) Ribosomal Ribonucleic Acid Is Transcribed at the 4-Cell Stage in In Vitro-Produced Bovine Embryos¹. *Biol Reprod* 59(3), 626-631
- WADE P.A. (2001) Methyl CpG-binding proteins and transcriptional repression. *Bioessays* 23(12), 1131-1137
- WAHLE E. (1995) 3'-End cleavage and polyadenylation of mRNA precursors. *BBA-Gene Struct Expr* 1261(2), 183-194
- WALSER C.B., LIPSHITZ H.D. (2011) Transcript clearance during the maternal-to-zygotic transition. *Curr Opin Genet Dev* 21(4), 431-443
- WANG Z., GERSTEIN M., SNYDER M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10(1), 57-63
- WOLFFE A. (2001) Transcriptional regulation in the context of chromatin structure. *Essays Biochem* 37, 45-57
- WU M.Y., RAMEL M.-C., HOWELL M., HILL C.S. (2011) SNW1 Is a Critical Regulator of Spatial BMP Activity, Neural Plate Border Formation, and Neural Crest Specification in Vertebrate Embryos. *PLOS Biology* 9(2), e1000593
- XU Y.-N., SHEN X.-H., LEE S.-E., *et al.* (2012) Autophagy Influences Maternal mRNA Degradation and Apoptosis in Porcine Parthenotes Developing In Vitro. *J. Reprod. Dev.* 58(5), 576-584
- XUE Z., HUANG K., CAI C., *et al.* (2013) Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* 500(7464), 593-597
- YE S., LI P., TONG C., YING Q.-L. (2013) Embryonic stem cell self-renewal pathways converge on the transcription factor Tfcp2l1. *EMBO J* 32(19), 2548-2560
- ZHANG J.-Z., GAO W., YANG H.-B., *et al.* (2006) Screening for Genes Essential for Mouse Embryonic Stem Cell Self-Renewal Using a Subtractive RNA Interference Library. *Stem Cells* 24(12), 2661-2668
- ZHANG P., ANDRIANAKOS R., YANG Y., LIU C., LU W. (2010) Kruppel-like factor 4 (Klf4) prevents embryonic stem (ES) cell differentiation by regulating Nanog gene expression. *J. Biol. Chem.* 285(12), 9180-9189
- ZHOU Z., ZHANG Q., LU X., *et al.* (2013) The proprotein convertase furin is required for trophoblast syncytialization. *Cell Death Dis* 4, e593

Annexe 1 : Liste des gènes surexprimés en A par rapport aux catégories B, C et D

En rouge, les gènes étudiés plus en détails dans la comparaison A vs D.

Gènes uniquement surexprimés en A par rapport à la catégorie B :

ARHGD1B
BDH2
CCR1
DMRT1
DOK3
ENO2
FGF10
GEM
HDGFL3
NEBL
P4HA2
PCOLCE2
RETREG1
RGN
SLC25A24
SMIM3
TMEM17
TMEM207
ENSBTAG00000053807

Gènes uniquement surexprimés en A par rapport à la catégorie C :

ALOX15
PTGES
RNF125
SFRP2
TAGLN
TPM2
TTYH1
ZNF169
ERCC-00095
ERCC-00136
ERCC-00004
ERCC-00108
ERCC-00130
ENSBTAG00000052709

Gènes uniquement surexprimés en A par rapport à la catégorie D :

ADGRF5

AKR1C4
ANKFN1
BLNK
BRB
bta-mir-2887-2
C1QTNF3
CR2
CRISP2
CYP26A1
EMX2
F2RL2
FBN2
FCER1A
FKBP11
FXD4
HPD
KLRA1
KRTDAP
LUZP2
MCMDC2
MMP9
MT1E
NADK2
OR5R1
PLA2G10
RF00090
RF00425
RF00553
RF00586
RF01241
S100A10
SASH3
TRIM63
ENSBTAG00000050086
ENSBTAG00000020105
ENSBTAG00000004817
ENSBTAG00000032057
ENSBTAG00000020550
ENSBTAG00000025952
ENSBTAG00000033308
ERCC-00079
ENSBTAG00000032812
ERCC-00022
ERCC-00076
ENSBTAG00000021098
ENSBTAG00000050377
ENSBTAG00000055264
ENSBTAG00000033802
ENSBTAG00000001595

ENSBTAG00000054664
ENSBTAG00000006397
ERCC-00051
ERCC-00044
ERCC-00054
ENSBTAG00000050159
ENSBTAG00000032775
ENSBTAG00000050728

Gènes surexprimés en A par rapport aux catégories B et C :

IFIT1
OLR1
ENSBTAG00000003892

Gènes surexprimés en A par rapport aux catégories B et D :

KCNE4
TPST2

Annexe 2 : Liste des gènes surexprimés en B par rapport aux catégories A, C et D

Gènes uniquement surexprimés en B par rapport à la catégorie A :

ABCG5
AP3M2
CLOCK
DSPP
FMO1
KL
MYH2
NCAN
OR5A1
SAMD12
SLC22A7
SLC39A12
SLC6A5
SMCO1
SYCE3
TMEM196
ZDHHC23
ENSBTAG00000049405
ENSBTAG00000048611
ENSBTAG00000053488
ENSBTAG00000040202

Gènes uniquement surexprimés en B par rapport à la catégorie C :

RNF125
SFRP2
TAGLN
TFR2

Gènes uniquement surexprimés en B par rapport à la catégorie D :

bta-mir-2887-2
CR2
CYP26A1
DRD1
EMX2
F13A1
GPR171
HMGNS
KLRA1
KRT17
KRTDAP
LUZP2

MARVELD1
MCMDC2
MIRLET7I
MT1E
MT2A
PLCB1
PYY
REG3G
RF00425 (chromosome 16, ENSBTAG00000042363)
RF00425 (chromosome 29, ENSBTAG00000043258)
RF00553
RF00586
RF01164
SP140L
TERB1
TNC
TNFSF18
TTC30B
ZNF674
ENSBTAG00000052352
ERCC-00079
ENSBTAG00000052977
ENSBTAG00000021098
ERCC-00054
ERCC-00044
ENSBTAG00000052047
ERCC-00022
ERCC-00076
ERCC-00051
ENSBTAG00000052759
ENSBTAG00000035643
ENSBTAG00000050086
ERCC-00074
ERCC-00046
ENSBTAG00000054034
ENSBTAG00000004817
ENSBTAG00000025952
ERCC-00071
ERCC-00002
ENSBTAG00000038900
ENSBTAG00000049228
ENSBTAG00000050728
ENSBTAG00000054655
ENSBTAG00000051842
ENSBTAG00000048743
ERCC-00162
ENSBTAG00000005749
ENSBTAG00000038783

Gènes surexprimés en B par rapport aux catégories C et D :

CCDC172

ZNF770

Annexe 3 : Liste des gènes surexprimés en C par rapport aux catégories A, B et D

Gènes uniquement surexprimés en C par rapport à la catégorie A :

ABCA10
ABCG5
AP3M2
C7H19orf57
CORIN
CRAMP1
CYCT
DMTF1
DSPP
EIF4E1B
FAM117B
FAM163A
FBXL4
FMO1
FOXRED2
FRAT2
GIPC2
GRHL3
GRM2
HEPACAM
MSI1
MYF5
MYH2
NCAN
NMRK2
OR5A1
PER1
PEX5L
PHF7
PPP1R15A
PRKAB2
SAMD12
SEMA3F
SLC22A7
SLC39A12
SLC6A5
SMARCA2
SMCO1
TMEM196
UNC13A
ZDHHC23

ZFYVE1
ZNF792
ENSBTAG00000038891
ENSBTAG00000054666
ENSBTAG00000000623
ENSBTAG00000049405
ENSBTAG00000048611
ENSBTAG00000037710

Gènes uniquement surexprimés en C par rapport à la catégorie B :

B3GALT2
CCIN
CCR1
CYP2B6
DOK3
GEM
GLS2
NEBL
PTPRB
RETREG1
RGN
ENSBTAG00000053807

Gènes uniquement surexprimés en C par rapport à la catégorie D :

ADGRF5
AGPAT2
AKR1C4
ANKFN1
BLNK
CRISP2
F13A1
FBXO41
H2B
HNF4G
KLRA1
KRT10
MCMDC2
RF00045
RF00092
RF00265
RF00322
RF00334
RF00394
RF00406
RF00412
RF00413 (chromosome 26, ENSBTAG00000042224)
RF00413 (chromosome 26, ENSBTAG00000043641)
RF00421
RF00425

RF00443
RF00553
RF00561
RF00586
SPTA1
ENSBTAG00000052352
ENSBTAG00000021098
ENSBTAG00000020105
ENSBTAG00000046066
ENSBTAG00000025952
ENSBTAG00000049134
ENSBTAG00000030490

Gènes surexprimés en C par rapport aux catégories A et B :

HSPA1A
ORM1
PAG11
WISP1

Gènes surexprimés en C par rapport aux catégories B et D :

KCNE4
RF00263

Gène surexprimé en C par rapport aux trois autres catégories B, C et D :

SLAMF1

Annexe 4 : Liste des gènes surexprimés en D par rapport aux catégories A, B et C

En rouge, les gènes étudiés plus en détails dans la comparaison A vs D.

Gènes uniquement surexprimés en D par rapport à la catégorie A :

ABCC2
ABCG5
ACP7
ACTL8
ADARB1
ADCK1
ADD1
ADD2
ADGRA2
AGBL2
AMDHD1
ANKS4B
AP3M2
ARC
ARHGAP25
ARHGAP29
ARHGAP39
ARHGEF33
ARPP21
ASIC1
ASPHD2
ATP10B
BARHL2
BARX2
BICC1
BMP6
BMPR1B
BNC1
BRINP1
BSX
BTBD6
C10orf99
C1H3orf70
CA2
CACNA1E
CACNA2D2
CACNG7
CADM2
CALCA

CAND2
CARMIL2
CCDC142
CCDC85A
CCK
CCL25
CCNJL
CCR8
CDC25B
CDHR1
CDT1
CEP104
CEP164
CHAC1
CHL1
CMYA5
CNN2
CNTD2
CNTROB
COL4A6
CORO2B
CORO6
CPS1
CRAMP1
CREB5
CRYBG1
CSMD1
CYCT
DAB1
DAB2IP
DCLRE1A
DDB2
DGKA
DIRAS2
DMTF1
DOCK2
DOCK3
DOK5
DSPP
DYNC1I1
EEF1AKMT3
EFCAB8
EGR4
ELAVL3
ELF3
ELOVL4
ELP1
EML6
EOMES

EPB42
EPA6
EPS8L1
ERN1
ERRFI1
EXTL3
FAM124B
FAM149A
FAM163A
FAM193A
FAM19A2
FAM83A
FANK1
FAS
FBXO10
FBXO30
FCRLB
FER1L6
FFAR4
FGF14
FGF19
FIGNL2
FMN1
FMO1
FOSL2
FOXN3
FOXP3
FOXRED2
FRAT2
FYB2
GAS2
GAS7
GBX2
GJA8
GNGT1
GPC3
GPC4
GPR1
GPR143
GPR27
GPR84
GRAP2
GRHL3
GRIP1
GRM2
GRM7
GSE1
H4
HAAO

HAS2
HAVCR2
HCFC1
HDAC10
HEPACAM
HES2
HGF
HHAT
HLX
HOXC13
HOXD13
HPCAL1
HSPB11
HUNK
IBSP
IFT172
IGF2BP2
IGFBP1
IGSF21
IGSF9
IL10RA
IL7R
ILD1R
IQANK1
IRF2
IRX4
ITGAL
KCNA5
KCNH2
KCNK2
KCNN4
KIAA1211
KIAA1324L
KIF17
KIF3C
KL
KLF7
KRT23
KRT25
LCA5
LETM2
LHFPL5
LHX4
LIMCH1
LMO3
LPL
LRP6
LRRC8E
LRRN4

LTBP3
LYPD1
MAB21L1
MACO1
MAFA
MAGEB3
MAGI1
MARVELD2
MCF2
MCTP1
MDGA1
MFHAS1
MFN2
MFSD6
MICAL1
MICAL2
MPDZ
MPIG6B
MRVI1
MS4A1
MTUS2
MYB
MYH2
MYO1E
NAA80
NALCN
NBEA
NCAN
NCF2
NEK8
NFATC1
NFIA
NIPAL3
NLRP1
NLRP5
NLRP9
NMRK2
NOBOX
NOD1
NOXA1
NPAS3
NR1H3
ODF4
ONECUT2
OPLAH
OR5A1
OTOG
PABPC5
PADI3

PARP12
PARP4
PCDH8
PCSK1
PDE4A
PFKFB1
PHF2
PHKA2
PHLDA2
PITPNM2
PIWIL2
PKIG
PLCG2
PLPPR1
PNKD
PNLDC1
PNPLA6
POU4F3
PPEF2
PPP1R12B
PPP1R26
PPP1R9A
PRDM1
PRKCH
PRKG1
PROM2
PRSS27
PRSS33
PTPN14
PTPRN2
PTPRQ
RAI2
RALBP1
RALGDS
RALYL
RAP2C
RASGRP1
RASGRP4
RASL11A
RFFL
RHOA
RNF19B
RSPO4
RTKN2
SAA4
SCN4B
SDC3
SDK2
SEMA3D

SEMA3F
SF3A1
SFN
SH3YL1
SHANK2
SIL1
SLC12A5
SLC22A7
SLC24A2
SLC44A5
SLC6A5
SLC8A2
SLCO2A1
SLCO4C1
SMCO1
SMURF1
SNAP91
SNCAIP
SNTB1
SNX30
SNX33
SOX5
SPI1
SPNS3
SPRR4
SRGAP1
SSH2
SSTR2
SUSD3
SV2C
SYCE3
SYT10
SYT11
TAC3
TACR3
TAF9B
TBX18
TCAF2
TCN1
TDRD7
TENM2
TFAP2B
THPO
THRA
TIAM1
TMEM102
TMEM125
TMEM178A
TMEM185B

TMEM196
TMEM253
TMEM38A
TMEM45B
TMEM52
TNFAIP2
TNFAIP3
TPBGL
TSPAN19
TTC23L
UBASH3A
ULBP21
UNC79
USP13
USP20
VCAN
VCIPI1
WDR20
WEE2
WIZ
WNT3
ZCCHC3
ZCWPW1
ZDHHC23
ZFAT
ZFHX2
ZFYVE1
ZHX2
ZNF18
ZNF385C
ZNF462
ZNF529
ZNF532
ZNF804B
ZРАНB3
ZXDC
ENSBTAG00000000886
ENSBTAG00000009653
ENSBTAG00000012150
ENSBTAG00000012370
ENSBTAG00000012827
ENSBTAG00000014126
ENSBTAG00000014467
ENSBTAG00000015251
ENSBTAG00000015315
ENSBTAG00000018583
ENSBTAG00000019479
ENSBTAG00000019553
ENSBTAG00000024482

ENSBTAG00000030440
ENSBTAG00000030838
ENSBTAG00000030892
ENSBTAG00000035013
ENSBTAG00000037550
ENSBTAG00000038322
ENSBTAG00000038891
ENSBTAG00000040461
ENSBTAG00000044018
ENSBTAG00000046635
ENSBTAG00000046755
ENSBTAG00000047473
ENSBTAG00000048478
ENSBTAG00000048576
ENSBTAG00000048611
ENSBTAG00000048977
ENSBTAG00000049077
ENSBTAG00000049186
ENSBTAG00000049405
ENSBTAG00000049752
ENSBTAG00000050189
ENSBTAG00000050206
ENSBTAG00000050418
ENSBTAG00000050454
ENSBTAG00000050574
ENSBTAG00000051340
ENSBTAG00000051714
ENSBTAG00000051802
ENSBTAG00000051816
ENSBTAG00000051941
ENSBTAG00000052530
ENSBTAG00000052808
ENSBTAG00000052951
ENSBTAG00000053114
ENSBTAG00000053150
ENSBTAG00000053238
ENSBTAG00000053445
ENSBTAG00000053488
ENSBTAG00000053503
ENSBTAG00000053699
ENSBTAG00000053802
ENSBTAG00000053833
ENSBTAG00000054342
ENSBTAG00000054666
ENSBTAG00000055035

Gènes uniquement surexprimés en D par rapport à la catégorie B :

AARS
ABCA12

ABCA3
ADAMTS14
ADAMTS4
ADGRD1
ADRA1D
AGO1
AKAP2
ALOX12
ANP32E
ARHGAP9
ARHGDIB
ARHGEF6
ARMCX1
ATP10D
ATP6V0E2
ATP8B1
B3GALNT2
B3GALT2
BCAT1
BDH2
BEND4
BICD1
BMP2
BOLA
CA4
CACNG4
CCDC96
CCIN
CCR1
CD14
CD276
CDH18
CDH23
CDH3
CDK2AP2
CDKN1C
CGNL1
CHPT1
CIDEA
CLDN11
CMPK2
CNGB1
CNP
COL15A1
COL5A1
CPXM1
CRELD1
CRISPLD2
CTSZ

DMRT1
DOK3
DPP4
DTHD1
DUSP22
DYDC1
ELOA
EML1
ENO2
EPB41L3
EPHA2
ERICH4
ESAM
FAM107B
FAM110A
FAM209A
FAM92B
FBLN5
FGF10
FGF12
FGF5
FLT1
FNDC7
FSTL5
GABRA4
GALK2
GALNT9
GARNL3
GBA2
GEM
GFOD1
GJD2
GMPR
GSS
H1FX
HAS3
HECA
HHIPL2
HHLA2
HOGA1
HOXA7
IFFO1
INSIG1
ITGA11
JUN
KCNA3
KLF12
KLHL3
KMT5B

LBX1
LDB3
LMX1A
LRRC34
LRRC4
LRRC46
LSS
LTA
MAML1
MAPK4
MARC2
MBNL2
ME2
ME3
MEF2C
MEGF9
MFSD6L
MGAM
MSMP
MUC15
MYH15
MYO10
MYO1D
NAGLU
NEBL
NEXN
NFIB
NKX1-1
NLRC5
NNAT
NPB
NPNT
NPTX2
NR2E1
NUCB1
NWD2
OTX1
P4HA2
PACSIN2
PAM
PBLD
PCGF5
PCOLCE2
PFKL
PFKP
PGR
PIK3AP1
PKIA
PKP3

PLEKHB1
PNOC
POLM
PPFIBP1
PPP4R1
PRM2
PSMB8
PTGDR
PTRB
PTRM
RAB26
RAPGEFL1
RASGRF1
RASGRF2
RBM15
RETREG1
RF00026
RGN
RIMKLA
RNF43
ROM1
SCRN2
SFRP1
SFXN3
SIK3
SLC10A3
SLC10A4
SLC13A5
SLC16A13
SLC16A5
SLC1A2
SLC20A2
SLC25A29
SLC25A53
SLC38A8
SLC7A7
SLCO3A1
SOCS1
SPINDOC
SQLE
SUSD1
SVOP
SYNPO2L
SYT1
TAPBP
TAT
TINAG
TM7SF3
TRAF3IP3

TRIM32
TRIM9
TSPAN5
UNC80
WDR27
WWP2
XPO6
ZBTB37
ZBTB38
ZFHX3
ZFHX4
ZIM2
ZNF512B
ZNF879
ENSBTAG00000053807
ENSBTAG00000048460
ENSBTAG00000050057
ENSBTAG00000033656
ENSBTAG00000050737
ENSBTAG00000051440
ENSBTAG00000050871
ENSBTAG00000049205
ENSBTAG00000053613
ENSBTAG00000049646
ENSBTAG00000052058
ENSBTAG00000050721
ENSBTAG00000047294
ENSBTAG00000051436
ENSBTAG00000051258
ENSBTAG00000030416
ENSBTAG00000026836
ENSBTAG00000001704
ENSBTAG00000046425
ENSBTAG00000004608
ENSBTAG00000052895
ENSBTAG00000018549
ENSBTAG00000020516
ENSBTAG00000055266
ENSBTAG00000051328
ENSBTAG00000053420

Gènes uniquement surexprimés en D par rapport à la catégorie C :

ADAMTS18
ADHFE1
C13H20orf96
CCDC146
CCDC181
CDC42BPA
CFAP100

CLDN18
DUSP4
DYRK2
FABP7
FRS2
FSD1
GADD45B
GAS6
GREB1
HOXD11
JAKMIP3
JMJD8
KCNA2
LSM11
MAL2
OVCH2
PHF19
PHLPP1
PLSCR3
PNMA8A
PPP1R35
RNF125
SFRP2
SGMS2
SH3BP2
SIM2
SLC36A1
STK17B
TAGLN
TFR2
TMEM159
WIPI1
ZNF569
ENSBTAG00000052180
ENSBTAG00000047749
ENSBTAG00000024791
ENSBTAG00000054366

Gènes surexprimés en D par rapport aux catégories A et B :

ABCB10
ABLIM1
ACBD4
ACCSL
ADAMTSL1
ADRB2
AGPAT1
AHR
ALKBH5
ARFGEF3

ARHGAP24
ASTL
ATXN7L1
AZIN2
BAIAP2
BCL3
BCL6B
BMP15
BSPRY
BTBD2
C15H11orf52
C18H16orf46
C21H15orf39
C7H19orf57
CABLES2
CACHD1
CAMK2B
CCNO
CCNYL1
CDK5R2
CDKL1
CDKN1A
CDR2
CEACAM19
CECR2
CHAMP1
CLIP3
CNNM4
COCH
CORIN
CSGALNACT1
CUEDC1
CXCR4
CYP1A1
DACT1
DAZL
DDX25
DLG2
DNAJB4
DNAJC9
DPF1
DRC7
DUSP5
EIF4ENIF1
ELK1
ELK3
ELN
EMILIN2
EPB41

EPHA7
ESR2
ETV1
ETV5
ETV6
F3
FAM222B
FGF16
FKBP6
FOSL1
FOXI3
GCC1
GCNT3
GJA4
GJC1
GOLGA7B
GPR173
GTPBP1
HES6
HES7
HEY2
HOMEZ
HOOK3
HSPA1A
HSPB6
IGF2BP3
IKZF2
IL18RAP
ILDR2
INPP1
IPP
IQCA1L
IRF2BPL
ITSN2
JAML
KCNN3
KCTD13
KHDC3L
KIAA1147
KIAA1958
KLHL18
KLHL25
KLHL32
KPNA7
LACC1
LHFPL3
LMBRD1
LRIG1
LRR15

LRRC8B
MAP3K21
MEGF11
MGAT5B
MGST2
MOS
MPND
MRTFA
MS4A13
MSI1
MTMR12
MX11
MYH10
MYH7B
MYOZ1
NEDD4L
NEIL2
NHSL1
NME7
NOTO
NPTX1
NRARP
NRG3
NRXN1
NRXN3
OCLN
OLIG1
OOSP2
ORM1
OSBP2
OSBPL10
PAG11
PAK5
PER2
PEX5L
PFN2
PHF7
PIM3
PLEKHF2
PLEKHG4
PLEKHH1
PLXNA4
PODXL
PPARGC1A
PPIF
PPM1H
PPP1R15A
PRAG1
PRAME

PRELID3A
PTDSS1
QSOX1
RAB15
RAPGEF4
RARA
RASA1
RASD2
RASGEF1A
REM2
RGS16
RGS2 (chromosome 4, ENSBTAG00000020620)
RND1
RNF34
RNPEP
RORA
ROS1
RPAP1
RRAGD
RTP1
SAMD12
SBF2
SBNO2
SCN8A
SEC16B
SELP
SESN2
SEZ6
SH3BP5
SH3GLB2
SHTN1
SIX1
SIX3
SKAP2
SLC12A1
SLC26A8
SLC35E4
SLC39A12
SLC5A9
SLC7A1
SMARCA2
SMCO4
SMOC1
SNX25
SOX30
SPIRE1
STK11
STK33
STX6

SULF2
SYTL5
TCF4
TCF7
TDRD15
TECTB
TET3
TMEM145
TMEM163
TMOD2
TMTC2
TOX3
TRAT1
TRIM44
TRIM6
TRIM65
TRIM77
TRPM3
UCP3
UNC5D
USP44
VANGL1
WDR6
WDR93
WISP1
YBX2
YPEL4
ZBTB16
ZC2HC1A
ZMAT4
ZNF629
ZNF71
ZP2
ZP4
ENSBTAG00000050187
ENSBTAG00000025760
ENSBTAG00000053102
ENSBTAG00000046199
ENSBTAG00000054056
ENSBTAG00000046630
ENSBTAG00000023752
ENSBTAG00000019009
ENSBTAG00000049459
ENSBTAG00000009197
ENSBTAG00000052570
ENSBTAG00000049550
ENSBTAG00000051262
ENSBTAG00000009760
ENSBTAG00000033967

ENSBTAG00000026637
ENSBTAG00000002545
ENSBTAG00000054030
ENSBTAG00000048850
ENSBTAG00000048673
ENSBTAG00000054352
ENSBTAG00000053626
ENSBTAG00000054829
ENSBTAG00000006727
ENSBTAG00000052032
ENSBTAG00000049009
ENSBTAG00000038532
ENSBTAG00000052435
ENSBTAG00000053475

Gènes surexprimés en D par rapport aux catégories B et C :

ARID1B
CD302
CDK14
FRS3
FYN
GRM4
IFIT1
IRX6
LTBP1
MVB12B
OLR1

Gènes surexprimés en D par rapport aux catégories A et C :

BTK
FOXJ3
GAP43
HIST3H2A
KCTD12
KIAA0355
KLK7
LMO7
MAGIX
NUDT11
NUDT2
PALM3
RASD1
RHOBTB1
SALL2
SOCS7
TBX15
WISP3
ENSBTAG00000054434
ENSBTAG00000013920

ENSBTAG00000051378

ENSBTAG00000038815

Gènes surexprimés en D par rapport aux trois autres catégories A, B et C :

ABHD4

ABL1

ADAM19

ADAP1

ADAR

ARMC2

ASB11

ATXN1

BCAR4

BCL7A

BNC2

BPGM

BRD3

BTG4

CASR

CASTOR1

CCDC92

CCNI2

CCNQ

CD200

CDA

CDCA7L

CDO1

CIART

COPG2

CROT

CSNK1E

CTH

CYB5A

CYBRD1

CYFIP2

DBX1

DLX6

DND1

DUSP10

DYRK1B

EBF1

EIF4E1B

ELAVL2

ENSA

EPB41L4B

ESYT3

EXT1

FAHD1

FAM110B

FAM117B
FAM161A
FAM81A
FBRSL1
FBXL4
FBXO34
FER
FERMT2
FKBP5
FNBP1L
FOXO1
FOXO3
GATA2
GCA
GDF9
GGCT
GIPC2
GNG12
GNG4
GPR137B
H1FOO
HACD1
HIST2H2BF
HLF
HNF1B
ID3
IGF1R
IGFBP3
INKA2
INSM1
IRF4
IRF8
JAZF1
JPH1
KCTD9
KLF11
LEF1
LYSMD4
MPP7
MSX1
MYF5
MYLIP
NLRP14
NLRP8
NPM2
NR3C2
PADI6
PANX1
PAX3

PAX6
PDK1
PER1
PHOX2A
PHYHIPL
PRDM13
PROCA1
PRSS23
RAB30
RAB3C
RAI14
RALB
RBM38
RBPMS
RERE
RGS2 (chromosome 16, ENSBTAG00000034366)
RIMS2
RPGRIP1
RRAS2
SAXO1
SDR42E1
SELENBP1
SEMA3A
SHOX2
SLBP2
SLC7A3
SMAD5
SNCA
SPAG17
SPG21
SPOP
SPRED2
TACC3
TBC1D8
TCL1A
TDRP
TENM4
TENT5C
TMCC2
TMEM225B
TMEM233
TOP2B
TRIB1
TRIM59
TUB
TUBB3
UBE2D1
UHRF1
UNCX

USH2A
USP2
USP36
VGLL4
VPS26C
WASF3
WSB2
WTIP
ZAR1
ZAR1L
ZCCHC14
ZNF362
ZNF398
ZNF436
ZNRF2
ENSBTAG00000037710
ENSBTAG00000000623
ENSBTAG00000040202
ENSBTAG00000044125
ENSBTAG00000021747
ENSBTAG00000053893
ENSBTAG00000013274
ENSBTAG00000048544
ENSBTAG00000052473
ENSBTAG00000054813
ENSBTAG00000005429
ENSBTAG00000053831
ENSBTAG00000049111
ENSBTAG00000053278
ENSBTAG00000024204
ENSBTAG00000055205
ENSBTAG00000006961
ENSBTAG00000052673
ENSBTAG00000054147
ENSBTAG00000023666
ENSBTAG00000012389
ENSBTAG00000053612
ENSBTAG00000048635

RÉPERCUSSION DU DÉVELOPPEMENT MORPHOCINÉTIQUE PRÉCOCE SUR LA TRANSITION MATERNO-EMBRYONNAIRE : COMPARAISON DES TRANSCRIPTOMES D'EMBRYONS BOVINS AU STADE 16-32 CELLULES

AUTEUR : Audrey TESTE

RÉSUMÉ :

La transition materno-embryonnaire correspond au moment où le contrôle du développement embryonnaire passe du transcriptome maternel au transcriptome embryonnaire.

Elle implique l'élimination des transcrits maternels et l'activation du génome embryonnaire, qui se déroule entre les stades huit et seize cellules chez les bovins. Notre étude s'est intéressée à l'effet que pouvait avoir le développement morphocinétique précoce sur cette transition materno-embryonnaire chez les embryons bovins.

Une analyse transcriptomique a donc été réalisée sur des embryons bovins au stade 16-32 cellules (post-activation du génome embryonnaire) et prédits dans quatre catégories morphocinétiques de blastocystes : A, B, C et D, classées selon un gradient décroissant de similitude avec les standards de la littérature, jugés normaux. La catégorie D étant donc l'extrême opposé à la catégorie A, et présentant des divisions anormales et/ou une cyclicité arythmique, mais atteignant tout de même le stade blastocyste « transférable ».

La comparaison quantitative entre les différentes catégories a révélé que les catégories A et D étaient celles qui présentaient le plus de différences dans leur transcriptome. Soixante gènes étaient surexprimés dans la catégorie A par rapport à la catégorie D et inversement, 899 gènes étaient surexprimés dans la catégorie D par rapport à la catégorie A. 2,5% des gènes surexprimés dans les embryons de la catégorie A avaient un profil maternel et 47,5% avaient un profil embryonnaire. Dans les embryons de la catégorie D, 37,5% avaient un profil maternel contre 6,3% qui avaient un profil embryonnaire.

Aucune voie fonctionnelle n'était enrichie par les gènes surexprimés dans la catégorie A comparée à la catégorie D, alors que treize voies fonctionnelles ont été mises en évidence comme étant enrichies par les gènes surexprimés dans les embryons de la catégorie D. Ces voies étaient liées à la régulation positive ou négative de la transcription, à la régulation de la kinase sérine-thréonine, à la régulation négative du système immunitaire, à des phénotypes générant de la létalité néo- et post-natale et à des anomalies de signaux dans les synapses.

Cet enrichissement de voies fonctionnelles de manière exclusive aux embryons ayant présenté des anomalies morphocinétiques serait dû à la forte proportion de transcrits maternels dans ces embryons là. Ils possèdent donc un retard de transition materno-embryonnaire comparés aux embryons n'ayant pas présenté d'anomalie de développement morphocinétique visible, imputable à un retard de dégradation et/ou d'utilisation des transcrits maternels. De plus, ces embryons pourraient faire face à un retard d'activation de leur génome accentuant cette différence.

Dans tous les cas, cette étude est la première à montrer que le profil morphocinétique des embryons bovins a des répercussions sur leur transition materno-embryonnaire. Ce sera donc un facteur à prendre en compte lors de futures études sur cette étape du développement.

MOTS CLÉS : DÉVELOPPEMENT EMBRYONNAIRE / MORPHOCINÉTIQUE / TRANSITION MATERNO-EMBRYONNAIRE / ACTIVATION GÉNOME EMBRYONNAIRE / GÈNE / EMBRYON / TRANSCRIPTOME / DIVISION CELLULAIRE / BOVIN

JURY :

Président : Pr De La Taille Alexandre

1^{er} Assesseur : Dr De Paula Reis Alline

2nd Assesseur : Dr Chevallier Lucie

IMPACT OF EARLY MORPHOKINETIC DEVELOPMENT ON THE MATERNO-EMBRYONIC TRANSITION: COMPARISON OF TRANSCRIPTOMES OF BOVINE EMBRYOS AT STAGE 16-32 CELLS

AUTHOR: Audrey TESTE

SUMMARY:

The materno-embryonic transition corresponds to the moment when the control of embryonic development switches from the maternal transcriptome to the embryonic one.

It involves the delete of maternal transcripts and the activation of the embryonic genome, which takes place between stages eight and six cells in cattle. Our study looked at the effect that early morphokinetic development could have on this maternal-embryonic transition in bovine embryos.

A transcriptomic analysis was therefore carried out on bovine embryos at the stage 16-32 cell (post-activation of the embryonic genome) and predicted in four morphokinetic categories of blastocysts : A, B, C and D, classified according to a decreasing gradient of similarity with the standards of the literature, considered normal. Category D is therefore the extreme opposite to category A, and presenting abnormal divisions and/or an arrhythmic cyclicality, but still achieving the "transferable" blastocyst stage.

A quantitative comparison between the different categories revealed that categories A and D were the ones with the most differences in their transcriptome. Sixty genes were overexpressed in category A compared to category D and conversely, 899 genes were overexpressed in category D compared to category A. 2.5% of the genes overexpressed in embryos of category A had a maternal profile and 47.5% had an embryonic profile. In category D embryos, 37.5% had a maternal profile compared to 6.3% who had an embryonic profile.

No functional pathway was enriched by genes overexpressed in category A compared to category D, while thirteen functional pathways were highlighted as being enriched by genes overexpressed in category D embryos. These pathways were linked to the positive or negative regulation of transcription, to the regulation of serine-threonine kinase, to the negative regulation of the immune system, to phenotypes generating neonatal and postnatal lethality and to signal abnormalities in synapses.

This exclusive enrichment of functional pathways to embryos having presented morphokinetic anomalies would be due to the high proportion of maternal transcripts in these embryos. They therefore have a delay in materno-embryonic transition compared to embryos that have not shown any visible morphokinetic development anomaly, linked to a delay in degradation and/or use of maternal transcripts. In addition, these embryos could face a delay in activation of their genome enhancing this difference.

In all cases, this study is the first to show that the morphokinetic profile of bovine embryos has impacts on their materno-embryonic transition. This will therefore be a factor to take into account in future studies on this stage of development.

KEYWORDS: EMBRYO DEVELOPMENT / MORPHOKINETIC / MATERNO-EMBRYONIC TRANSITION / EMBRYONIC GENOME ACTIVATION / GENE / EMBRYO / TRANSCRIPTOME / CELLULAR DIVISION / BOVINE

JURY:

Chairperson: Pr De La Taille Alexandre

1st Assessor: Dr De Paula Reis Alline

2nd Assessor: Dr Chevallier Lucie