

# Commutation réversible entre les états épigénétiques dans les sous-organismes comportementaux des abeilles

[Brian R. Herb](#) , <sup>1,2,\*</sup> [Florian Wolschin](#) , <sup>3,4,†</sup> [Kasper D. Hansen](#) , <sup>1,6</sup> [Martin J. Aryee](#) , <sup>1,5</sup> [Ben](#)

[Langmead](#) , <sup>6</sup> [Rafael Irizarry](#) , <sup>6</sup> [Gro V. Amdam](#) , <sup>3,4,†</sup> et [Andrew P. Feinberg](#) <sup>1,2,6,7,†</sup>

[Auteur d'informations Droits d'auteur et licence Avertissement](#)

La version finale modifiée de cet article par l'éditeur est disponible à l' [adresse Nat Neurosci](#)

Voir d'autres articles dans PMC qui [citent](#) l'article publié.

## Données associées

[Matériel supplémentaire](#)

## Abstract

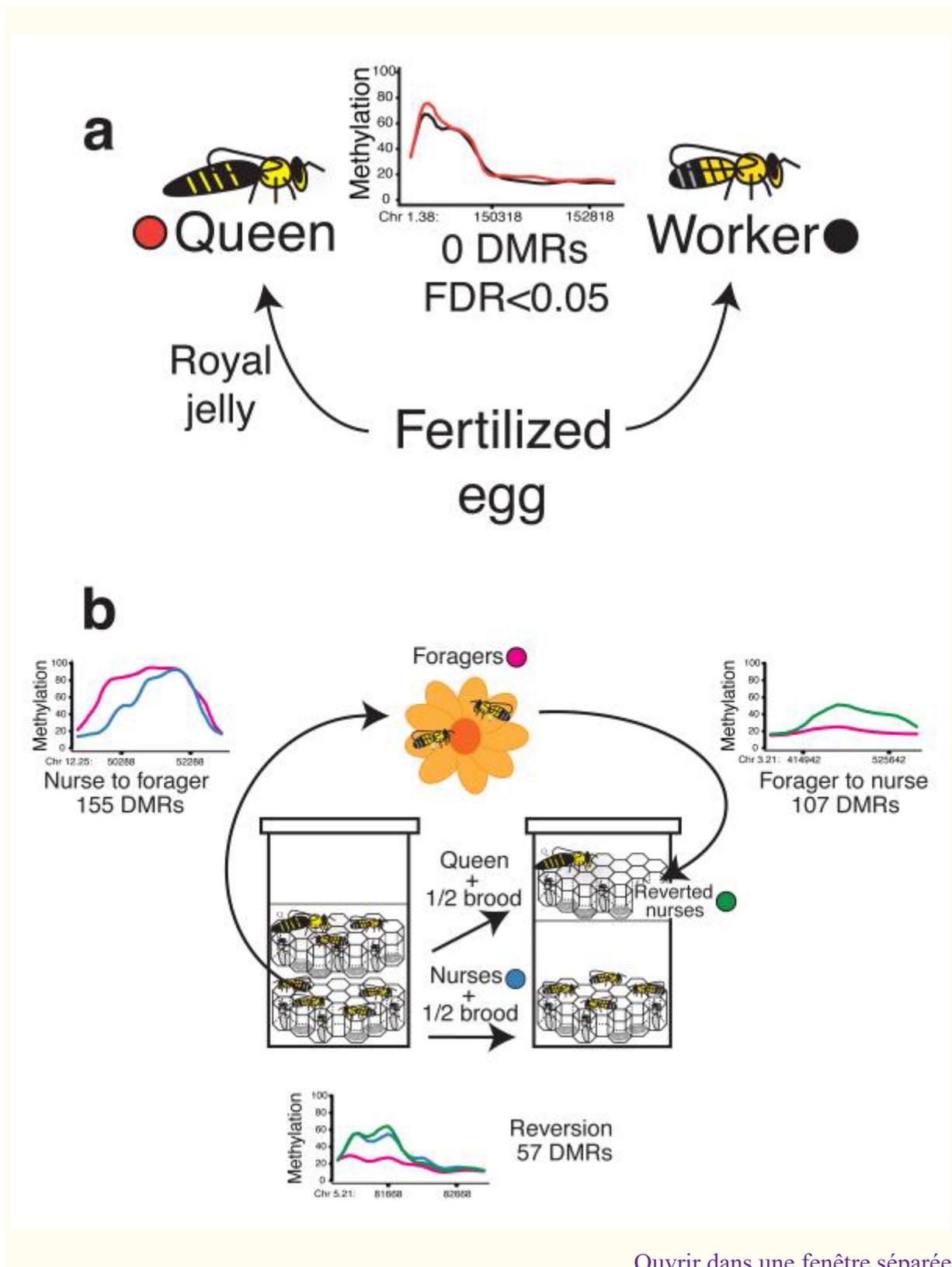
On pense que les modifications épigénétiques sont à la base de la différenciation spécifique à la lignée, car le modèle d'expression génique est modifié de manière stable, mais la séquence d'ADN reste la même. Récemment, l'épigénome d'une voie de différenciation spécifique a été cartographié, définissant des centaines de régions méthylées différenciellement (DMR) qui définissent l'engagement de la lignée chez les progéniteurs hématopoïétiques de souris. La méthylation de l'ADN semble essentielle dans ce système pour la spécificité de la lignée, car les cellules lymphoïdes présentent une méthylation globale de l'ADN supérieure à celle des cellules myéloïdes <sup>1</sup>. Cependant, les rôles de l'épigénome dans les changements globaux du remodelage de l'organisme ou du comportement n'ont pas encore été définis <sup>2</sup>.

L'abeille domestique *Apis mellifera* est un organisme modèle idéal pour de telles études <sup>3</sup>, car elle organise les structures sociales à partir de formes individuelles distinctes pouvant émerger d'un génome. Un embryon de femelle peut devenir une reine en suivant un régime de gelée royale et en consacrant sa vie à la ponte (lignée germinale) ou en devenant un «ouvrier» stérile (soma) <sup>4</sup>. Les travailleuses suivent un riche programme comportemental d'allaitement et passent ensuite à la recherche de nourriture, ce qui implique d'importants changements dans l'expression du gène dans le cerveau <sup>5</sup>. Contrairement aux reines, le comportement des ouvrières est remarquablement flexible: les ouvrières du même âge peuvent allaiter ou chercher du fourrage, et les butineuses peuvent reprendre des tâches d'allaitement <sup>6</sup>.

Afin d'étudier le rôle potentiel de la méthylation de l'ADN dans la définition des phénotypes de castes d'abeilles, nous avons comparé les méthylomes des reines soeurs aux ouvriers et infirmières soeurs des plantes fourragères par séquençage au bisulfite du génome entier (WGBS) et méthylation relative à haut débit à matrice complète (CHARM) ) analyse <sup>1</sup>. CHARM couvre 85% et le WGBS couvre 92% des CpG du génome de 270 Mo, les deux révélant une méthylation clairsemée dans tout le génome de l'abeille ( [Figure 1 supplémentaire](#) ).

Nous avons d'abord comparé cinq répliques biologiques de reines et de travailleuses, toutes deux recueillies dans les 4 heures suivant l'émergence du stade de puppe chez l'adulte ( [Fig. 1a](#) ). Le cerveau a été sélectionné en raison de son influence sur le comportement et, contrairement à l'ovaire, sa taille est semblable entre les reines et les ouvrières. L'analyse CHARM n'a révélé aucun DMR significatif par test FDR entre les reines et les ouvrières. Le

WGBS des mêmes échantillons n'a pas révélé de différence, en utilisant un seul test t CpG corrigé pour plusieurs tests. En outre, nous avons testé les différences les mieux classées selon CHARM, bien que statistiquement non significatives, par le pyroséquençage au bisulfite, une mesure indépendante de la méthylation de l'ADN au niveau de la base unique, et n'avons trouvé aucune différence spécifique à la caste ([Fig. Supplémentaire 2a à c](#)).

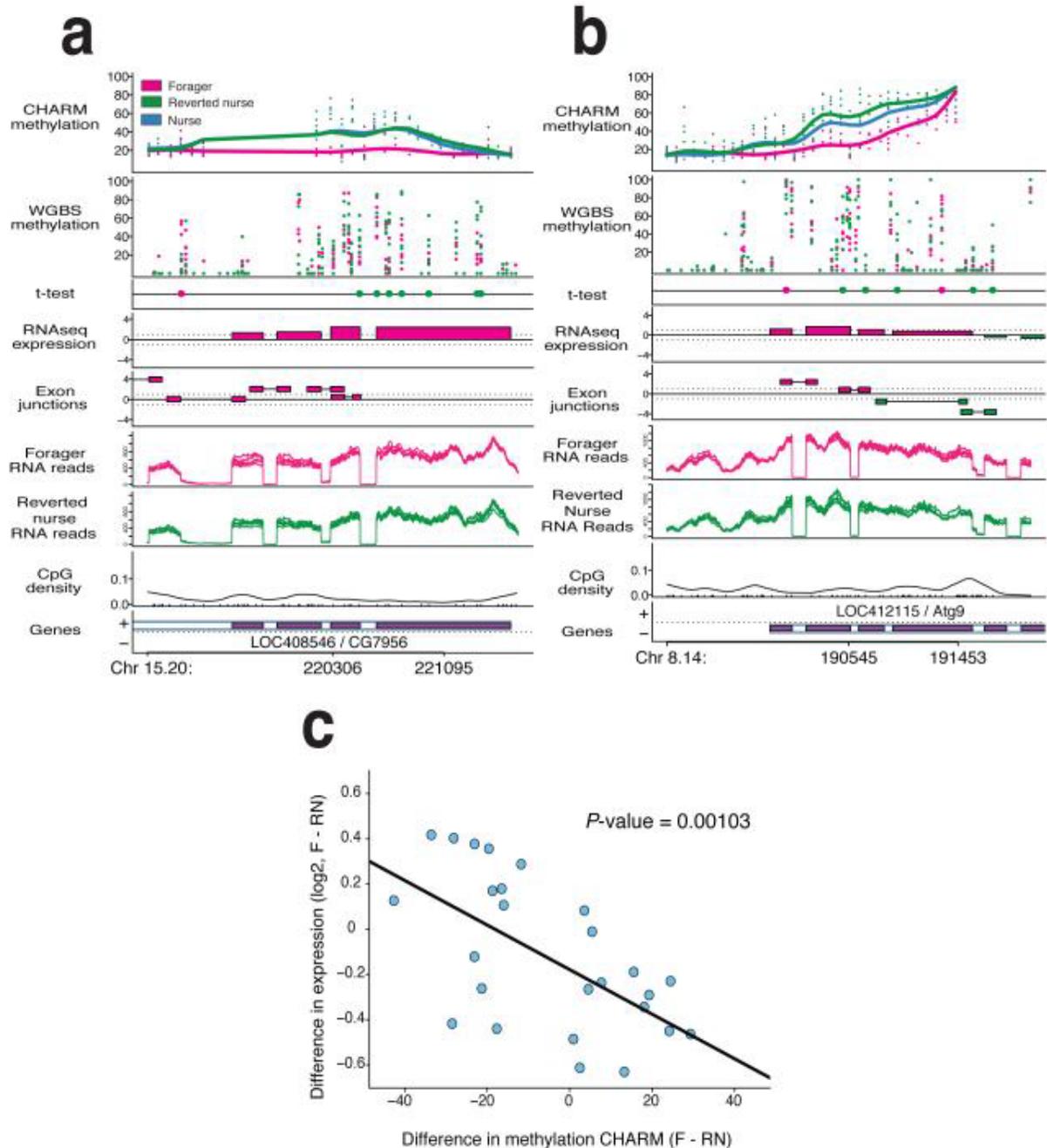


### Figure 1

#### **Modifications de la méthylation de l'ADN constatées entre les infirmières et les butineuses, mais pas entre les reines et les travailleuses**

**a** , Nous avons comparé des reines et des ouvrières émergentes utilisant CHARM ( $n = 5$  par phénotype) et n'avons trouvé aucune différence statistiquement significative. **b** , la méthylation de l'ADN change au cours de la transition entre infirmière et plante fourragère et revient pendant la transition de la fourrure à infirmière ( $n = 3$  par phénotype). Nous avons trouvé 155 rapports de détresse associés à la transition entre le système infirmier et le dispositif de repérage, 107 d'entre eux associés au changement de repaire et le système d'infirmière, et 57 rapports de rapport de médicament communs aux deux listes portant une signature spécifique.

Compte tenu de ces résultats négatifs, nous avons ensuite comparé les sous-groupes de travailleurs. Au début, la plupart des ouvriers sont des infirmières qui soignent les reines et les larves dans la ruche. Environ deux ou trois semaines plus tard, la majorité d'entre eux se tournent vers la recherche de nourriture et la collecte de pollen, de nectar et d'eau en dehors de <sup>5</sup>. À l'aide de CHARM, nous avons identifié 155 DMR distinguant les infirmières des cultures fourragères ( [Fig. 1b](#) , [Fig. 2a-b](#) , [Tableau 1](#) , [Tableau supplémentaire](#) , Validation du pyrosequencing au bisulfite, [Fig. 3a-e supplémentaire](#) ). Environ 70% des DMR chevauchent des exons (résumé dans le [tableau 1](#) et description complète de la localisation génomique des DMR indiquée dans le [tableau supplémentaire 1](#) ), similaire aux études précédentes <sup>7,8</sup>. Les gènes associés aux 155 DMR de nourrice à nourricière semblent être enrichis pour la régulation et le développement de gènes par le contrôle de la transcription et le remodelage de la chromatine. De nombreux auteurs de modifications d'histones, y compris LOC412350, une histone désacétylase similaire à Hdac3, JIL-1, une histone phosphotransférase et LOC411070, une histone H3 méthyltransférase <sup>9</sup> augmentant de méthylation au cours de la transition infirmière-infirmière. De plus, les gènes Iswi et spn-E de l'hélicase DEAD-box ont une capacité de remodelage de la chromatine et sont impliqués dans la morphogenèse <sup>10</sup>.



**Figure 2**

**La méthylation de l'ADN distingue les infirmières, les butineuses et les infirmières rétablies**

Deux exemples de DMR CHARM. Les panneaux du haut, **ab**, indiquent le pourcentage de méthylation pour les ensembles de données CHARM et WGBS, les points représentant des échantillons individuels et les lignes lissées représentant la moyenne du phénotype. Le panneau de test t affiche les CpG méthylés différenciellement à 1% supérieurs par le test t. La couleur du point indique quel phénotype présente une méthylation supérieure à cette CpG ( $n = 6$  par phénotype). Le panneau d'expression de RNAseq est une statistique t basée sur le nombre ou les lectures détectées dans les exons annotés, la couleur indiquant le phénotype exprimé le plus élevé. Le panneau Jonctions exon est une statistique t basée sur le nombre ou les lectures détectées couvrant les jonctions exon, tel que prédit par le programme TopHat, la couleur indiquant le phénotype exprimé le plus élevé. La commutation entre les jonctions exon d'infirmière et de forager plus exprimée est révélatrice d'autres événements d'épissage. Les panneaux de lectures d'ARN indiquent le nombre de lectures par

phénotype compilé par le programme TopHat ( $n = 6$  par phénotype). Les deux panneaux du bas montrent la densité de CpG et la position relative du gène. c .Parcelle d'expression relative des gènes comparant les cultures fourragères aux infirmières reconverties. 26 gènes associés aux DMR ont été testés pour les différences d'expression génique par PCR en temps réel ( $n = 12$  par phénotype). Le graphique illustre la différence entre l'expression moyenne de log2 et la différence moyenne de méthylation, déterminée par CHARM. L'analyse de corrélation donne une valeur de  $p$  de 0,001.

## Tableau 1

### Résumé des régions différemment méthylées (DMR)

| Type de DMR                         | Nombre de DMR | % de DMRs se chevauchant Exons | Exemples de catégories GO   | Exemple de gènes   |
|-------------------------------------|---------------|--------------------------------|---|--|
| Transition infirmière à butineuse   | 155           | 72,2%                          | Activité hélicase, remodelage de la chromatine, développement de neurones | alpha-Cat, Stat92E, Dhc64C, alpha-Spec, hts, spn-E         |
| Transition de nourrice à infirmière | 107           | 72,9%                          | Liaison nucléosidique, activité hélicase, organisation du cytosquelette   | Hel89B, Hsc70Cb, CG7177, Upf1, CG2017, eIF-4a, kis, Dhc64C |
| Chevauchement                       | 57            | 73,7%                          | Importation nucléaire, différenciation cellulaire, liaison à l'ATP        | Ranbp21, Fs (2), Ket, Mtor, eIF-4a, kis, poe, bur, BicD    |

Iswi en particulier joue un rôle dans la morphogenèse des dendrites <sup>11</sup> et peut contribuer aux changements constatés dans le cerveau nourricier avant la recherche de nourriture <sup>5</sup>. Afin de déterminer si les DMR que nous avons observés au cours de la transition infirmière-nourricière sont liés au phénotype, et pas simplement au résultat de la transition, nous avons renvoyé les fourragères aux infirmières en utilisant une stratégie de ruse de la ruche. Pour amorcer la réversion, les butineuses sont prêtes à retourner dans une ruche où seules les reines et les larves sont présentes ( [Fig. 1b](#) ). Les butineuses se sépareront ensuite en infirmières rétablies qui se chargent des tâches du soignant et des butineuses continues qui ne changent pas de comportement. <sup>6</sup>. La réversion sépare les changements dus au développement du système nerveux, à la maturation et à la recherche de nourriture, partagés entre infirmiers retournés et cultivateurs mais non infirmiers, et les changements fortement liés au comportement actuel partagés entre infirmiers infirmiers et infirmiers retournés mais pas chercheurs.

Avec CHARM, nous avons trouvé 107 DMR pour le précurseur de la transition infirmière inversée. Les gènes associés à ces CHARM DMR semblaient être enrichie en facteurs de transcription ainsi que des gènes d'hélicase à boîte DEAD, comme on le voit à l'infirmière de Forager DMR CHARM ( [Fig. 2a-b](#) , [Tableau complémentaire 2](#) , [Fig supplémentaire. 4](#) ). Sur ces 107 DMR CHARM, 57 chevauchaient avec des DMR CHARM associés à une transition infirmière-fourragère, une concordance remarquablement proche (valeur  $P < 2,2 \times 10^{-16}$  selon le test de Fisher, valeur  $P < 10^{-3}$  basée sur 1000 permutations, [Figure 1](#)). <sup>5</sup>). Ce sous-ensemble de gènes épigénétiquement réversibles a montré un enrichissement pour le développement, la liaison à l'ATP et la formation de pores nucléaires ( [tableau complémentaire 2](#) ). Ces gènes incluent l'orthologue de kismet, LOC726524, qui régule les gènes du développement tels que le hérisson, affecte l'apprentissage et la migration des axones chez *Drosophila* <sup>12,13</sup> et pourrait expliquer les différences observées dans l'apprentissage <sup>14</sup> entre infirmières et butineuses. De plus, les gènes LOC725306 <sup>15</sup> et LOC726524 de l' hélicase à boîte DEAD ont tous deux un rôle dans la transcription, alors que LOC411989 est impliqué dans la traduction <sup>16</sup>.

Nous voulions valider ce résultat de manière indépendante. Nous avons donc répliqué l'expérience de réversion et créé six nouveaux pools de six cerveaux pour les butineuses et les infirmières reconverties. Nous avons effectué WGBS sur ces 12 échantillons et nous avons constaté que les DMR à réversion 45/57 montrent le même sens de changement de la méthylation entre CHARM et WGBS. ( [Fig. 6 supplémentaire](#) ). Ce chevauchement des DMR entre les expériences répliquées est hautement significatif (  $P$ -value =  $3.3 \times 10^{-6}$  ). En outre, les 45 gènes corrélés au WGBS montrent un enrichissement pour la liaison à l'ATP et la formation de pores nucléaires ( [Tableau supplémentaire 3](#) ), conformément à notre analyse des DMR à inversion de 57 CHARM. Ces résultats fournissent la preuve d'un méthylome spécifique à l'infirmière qui doit être rétabli lors de la réversion.

Pour déterminer la signification de ces DMR réversibles, nous avons effectué le séquençage du transcriptome (RNAseq) sur 6 groupes chacun de butineuses et infirmiers reconvertis. Nous avons ensuite utilisé le programme TopHat pour analyser les données d'ARN-seq afin de prédire l'emplacement des exons annotés et non annotés et de déterminer la prévalence du saut d'exon. Cette analyse a identifié 22 des 45 DMR de réversion corrélés WGBS co-localisés avec des événements d'épissage alternatifs. Un exemple est présenté à la [figure 2b](#) dans le cinquième panneau et d'autres exemples sont illustrés aux [figures supplémentaires 7a à f](#). Ces données montrent une forte incidence d'événements d'épissage alternatifs dans les DMR et renforcent le rôle potentiel de la méthylation de l'ADN dans la régulation de l'épissage alternatif. <sup>8</sup>. Nous avons également trouvé une corrélation négative entre l'expression des gènes et les niveaux de méthylation de l'ADN entre les chercheurs de fourrage et les

infirmières reconverties pour 26 gènes par PCR en temps réel (valeur de  $p = 0,00103$ , [Fig. 2c](#), [Fig. 8a à f supplémentaire](#) ).

Nos données montrent un lien étroit entre la méthylation réversible de l'ADN et la transition et la réversion d'infirmières fourragères, mais aucune relation avec la ségrégation de la reine-ouvrière. Ces données contrastent avec une étude comparant des reines accouplées âgées de 2,5 semaines et des travailleurs aptes à se nourrir à l'alimentation <sup>8 ans</sup> <sup>8</sup>, ce qui s'explique probablement par la différence de chronologie des données, à savoir les reines et les travailleurs nouvellement apparus. Bien que la méthylation de l'ADN puisse jouer un rôle dans la distinction entre les reines et les ouvrières au cours du développement <sup>3</sup>, nos données montrent clairement que les méthylomes du cerveau de la reine et de l'ouvrier sont les mêmes au moment de l'émergence, malgré les différences de morphologie corporelle.

En résumé, nous avons trouvé des changements substantiels dans la méthylation de l'ADN qui accompagnent le changement de phénotype chez les subcastes d'abeilles domestiques. Les gènes associés à ces DMR peuvent potentiellement influencer les schémas généraux d'expression des gènes en modifiant la structure de la chromatine ou en régulant la machinerie de transcription. Les changements phénotypiques profonds entre infirmières et nourrisseuses peuvent être orchestrés par un sous-ensemble de gènes, eux-mêmes régulés par la méthylation de l'ADN. Les gènes régulateurs clés peuvent être soit différenciellement exprimés, soit épissés différenciellement, ce que nous corrélons aux modifications de la méthylation de l'ADN. Par exemple, l'homologue LOC411989 de eIF-4a, qui joue un rôle essentiel dans l'initiation de la traduction <sup>17</sup>, présente un épissage alternatif dans un exon qui code pour la liaison à l'ARN ([Fig. 7c supplémentaire](#)). Différentes isoformes de eIF-4a peuvent se lier à l'ARN avec une plus grande affinité, affectant ainsi globalement le taux ou la régulation de la traduction. Étant donné que cet exon épissé différenciellement se trouve dans un DMR, la méthylation pourrait être utilisée pour rappeler quelle isoforme exprimer chez les infirmières ou les butineuses. De manière remarquable, nous avons constaté que la méthylation de l'ADN est capable de revenir, parallèlement au retour expérimental des plantes fourragères aux infirmières, ce que nous avons démontré lors d'expériences répétées. Ceci suggère une signature de méthylation spécifique au sous-collage qui aide à la formation de phénotypes de sous-collage. Bien que les études chez les rongeurs aient révélé des modifications de la méthylation associées à l'apprentissage, ces modifications disparaissent après plusieurs heures et ne permettent pas d'établir un phénotype stable <sup>18, 19</sup>. De même, les soins peuvent induire des marques de méthylation de longue durée chez les rongeurs <sup>20</sup>. Notre étude est la première à montrer une méthylation réversible de l'ADN correspondant à un phénotype comportemental réversible chez toutes les espèces.

[Aller à:](#)

## METHODES

---

Les méthodes et les références associées sont disponibles dans la version en ligne du document.

[Aller à:](#)

## METHODES EN LIGNE

---

Préparation des abeilles

Pour chaque répliat, deux colonies ont été initialement préparées pour l'élevage des nourrices et des fourragères. Chacun d'entre eux était composé de 6 000 à 7 000 travailleurs nouvellement émergés (âgés de 0 à 24 h), nés de cinq reines sœurs de notre stock de recherche standard et d'une reine de type sauvage d'origine commerciale californienne. Tous les individus portaient une marque de peinture (Testors™) sur le thorax, puis après avoir établi un modèle de recherche de nourriture solide, les plantes butineuses étaient marquées de peinture sur l'abdomen afin de suivre leur parcours vital. La réversion comportementale a été réalisée essentiellement comme décrit avant <sup>21</sup>. Cette réversion a eu pour résultat deux colonies dérivées de la recherche de nourriture et deux colonies dérivées de la nourrice. Une jeune génération d'origine commerciale californienne a été fournie aux colonies afin de les inciter à inverser le comportement de la recherche de nourriture au profit du comportement d'allaitement. Pour le répliat utilisé dans les tests CHARM et le pyroséquençage de bisulfite, les abeilles de trois groupes ont été collectées 12 à 14 jours après la reconversion: 1. Aliments fourragers continus: abeilles butineuses qui ne sont pas revenues à un comportement d'allaitement; 2. Abeilles nourricières retournées: abeilles butineuses ayant un comportement semblable à celui de l'infirmière (tête dans les cellules du couvain, réponse lente à un défi de vol) dans la zone du couvain; 3. Infirmières continues: abeilles qui n'avaient jamais été observées en train de butiner et présentaient au moins un trait de comportement analogue à celui d'une infirmière au moment de la collecte. Pour le répliat utilisé dans les tests WGBS et RNAseq,

Les reines et les ouvrières étaient issues d'œufs produits par une reine d'abeilles mellifères inséminées par un seul drone appartenant à un stock de recherche standard avec un patrimoine génétique limité <sup>22</sup> et laissés à se développer en larves âgées de deux jours dans la ruche. Par la suite, les larves, à ce stade crucial du développement divergent des traits de reine et de travailleur, ont été autorisées à se développer en pupes ouvrières ou, pour le développement des reines, elles ont été greffées manuellement dans des cellules royales et élevées comme décrit précédemment <sup>23</sup>. Le jour précédant l'émergence (le jour du prélèvement de l'échantillon), les nymphes reine et ouvrière ont été transférées dans un incubateur à 33 ° C, humidité relative comprise entre 65 et 70%.

Pour toutes les abeilles, les individus des deux expériences ont été recueillis directement dans 2 ml d'éthanol refroidi à 80% et conservés à 4 ° C jusqu'à la dissection du cerveau central (<48 h). À la suite de dissections, les cerveaux ont été immédiatement transférés centraux en azote liquide <sub>2</sub> et stockés à -80 ° C jusqu'à utilisation ultérieure.

### Analyse de méthylation de l'ADN CHARM

L'ADNg a été isolé du cerveau à l'aide du kit Masterpure d'Epicentre. Les cerveaux issus de l'étude infirmière / fourrage ont été regroupés en groupes de sept pour chaque répliat biologique, trois réplats par phénotype. Les cerveaux de l'étude reine / travailleur ont été regroupés en groupes de huit pour chaque répliat biologique, cinq réplats par phénotype. La méthylation à l'échelle du génome a été évaluée par CHARM <sup>24, 25</sup> et réalisée comme décrit précédemment <sup>1</sup>. Un ensemble personnalisé de 2,1 millions d'éléments Nimblegen conçu pour le génome de l'abeille domestique, qui couvre environ 200 Mo de séquence génomique non répétée et comprend environ 8,7 millions de sites CpG couvrant environ 85% des quelque 10,2 millions de sites CpG du génome.

Les DMR de l'étude sur les nourrices et les fourragères ont été déterminés par un seuil d'écart d'au moins 10% de différence de méthylation <sup>24</sup> et les DMR avec une méthylation moyenne proche de la valeur initiale de 20% ont été éliminés. Les DMR issus de comparaisons par paires ont été combinés pour déterminer la méthylation relative des trois phénotypes. Les

différences de méthylation entre les infirmières continues et les infirmières inversées et les butineuses continues ont été déterminées, et ces différences ont été tracées. L'analyse de regroupement a identifié trois classes de DMR, décrites dans la [Figure supplémentaire 9](#). Analyse GO <sup>26</sup>de chaque classe a été réalisée en déterminant d'abord le gène d'Apis mellifera le plus proche de chaque DMR. Le gène orthologue de Drosophila melanogaster a été trouvé pour chaque gène d'Apis mellifera et cette nouvelle liste de gènes a été utilisée pour l'analyse GO à l'aide des outils Web disponibles dans la base de données bioinformatique de DAVID à l'[adresse david.abcc.ncifcrf.gov](http://adresse.david.abcc.ncifcrf.gov). Les DMR pour les études reine / travailleur ont été déterminés en calculant un score FDR pour chaque DMR potentiel.

### Pyrosequencing bisulfite

Environ 400 ng d'ADNg regroupé qui a également été utilisé pour l'analyse CHARM ont été convertis au bisulfite à l'aide du kit Zymo DNA-Methylation Gold. Nous avons utilisé la PCR nichée pour amplifier les régions d'intérêt au sein des DMR et quantifié le niveau de méthylation à l'aide du pyrosequencer Biotage PSQ HS96. Le pourcentage de méthylation pour chaque CpG dans notre région cible a été calculé à l'aide du logiciel de méthylation Q-CpG (Biotage). L'ADN témoin a été préparé en utilisant le kit Repli-G (Qiagen) de l'ADNg. L'ADN amplifié par Repli-G a servi de témoin méthylé à 0%. Un contrôle à 100% méthylé a été créé en traitant l'ADN amplifié de Repli-G avec la SSSI méthyltransférase (NEB), qui méthyle chaque site CpG. Des contrôles méthylés à 25%, 50% et 75% ont été créés en mélangeant des contrôles à 0% et 100%. Tous les témoins étaient traités au bisulfite avec le même kit Zymo que les échantillons de test. Séquences d'amorçage énumérées dans [Tableau supplémentaire 4](#).

### RT-PCR quantitative en temps réel

L'ARN a été extrait de cerveaux simples en lysant d'abord les cellules dans du tampon Chaos (thiocyanate de guanidinium 4,5 M, 2% de N-lauroylsarcosine, 50 mM d'EDTA, 25 mM de Tris-HCl, 0,1 M de b-mercaptoéthanol), puis purifié au phénol, chloroforme, puis purifié avec du Qiagen RNeasy. colonnes. L'ADNc a été synthétisé en utilisant le kit de transcription inverse Quantitect (Qiagen) et 1 ng d'ADNc a été utilisé pour chaque réaction PCR en temps réel. Fast Sybr green (Applied Biosystems) a été utilisé pour la réaction PCR en temps réel et quantifié par 7900HT (Applied Biosystems). Les séquences d'amorces sont énumérées dans le [tableau supplémentaire 5](#).

### Séquençage du bisulfite du génome entier

Des bibliothèques d'échantillons de reine / ouvrière et d'infirmière / de butineuse ont été créées à l'aide de kits de préparation de banque d'ADN TruSeq (Illumina) avec quelques modifications apportées au protocole standard. Des échantillons d'ADN génomique ont été préparés en homogénéisant des pools de cerveaux entiers. Pour les échantillons de reine / travailleur, 5 pools de 8 cerveaux par pool ont été préparés. Pour les échantillons infirmes / buveurs inversés, 6 pools de 6 cerveaux par pool ont été préparés. L'ADN génomique a été extrait de cerveaux homogénéisés en utilisant soit le kit Masterpure d'Epicentre (reine / travailleur), soit en lysant les cellules dans du tampon Chaos (thiocyanate de guanidinium 4,5 M, 2% de N-lauroylsarcosine, 50 mM d'EDTA, 25 mM de Tris-HCl, 0,1 M de ) et purifié avec des colonnes Qiagen DNeasy (infirmière retournée / fourragère). Pour tous les échantillons, l'ADN génomique a été cisailé à une taille moyenne de 350 pb à l'aide d'un sonificateur Covaris avec les paramètres suivants: Cycle de travail = 10%, Intensité = 5,0, Rafales par seconde = 200, Durée = 60 sec. Des extrémités franches ont été créées sur les fragments

d'ADN en utilisant un protocole unique pour éliminer l'introduction de cytosines non génomiques dans les fragments, qui seraient interprétées à tort comme des cytosines non méthylées au cours d'une analyse ultérieure. Pour ce faire, nous avons utilisé uniquement les nucléotides A, G et T avec un mélange des enzymes T4 DNA polymérase, Klenow DNA polymérase et T4 PNK (NEB) pour effectuer la réparation finale des fragments. Les adaptateurs Illumina ont ensuite été ligaturés aux fragments après l'addition d'un seul A, selon le protocole TruSeq. Les bibliothèques ont ensuite été sélectionnées en coupant un fragment de 400-500 pb dans un gel d'agrose (Agarose ultra basse gamme certifiée Bio-Rad, échelle d'ADN NEB-100 pb, coloration de gel d'acide nucléique Invitrogen-SYBR® Gold) et purifiées à l'aide de Qiagen MinElute. Kit d'extraction de gel. Les banques purifiées ont ensuite été converties au bisulfite et purifiées à l'aide de Zymo EZ DNA Methylation Gold. Les banques ont ensuite été amplifiées en utilisant un mélange de polymérases insensibles à l'Uracile; Denville Choice Taq et Agilent Turbo Pfu. Les échantillons de reine / travailleur ont été amplifiés pendant 12 cycles et retournés infirmière / fourragère pendant 15 cycles en utilisant les conditions de PCR TruSeq.

### Séquençage d'ARN

Les échantillons d'ARN des 6 groupes de 6 cerveaux d'infirmières et de butineuses inversées ont été dérivés du même lysat utilisé pour créer les bibliothèques WGBS d'infirmières et de butineuses inversées. Les bibliothèques RNAseq ont été créées à l'aide du kit de préparation d'échantillons d'ARN Illumina TruSeq sans modification du protocole standard. Ce kit s'enrichit en ARNm en utilisant des billes liées à un poly-T oligo pour se lier aux résidus poly-A de l'ARNm.

### Analyse des données du séquençage du bisulfite du génome entier chez la reine et le travailleur

Nous avons exécuté le pipeline d'alignement du bisulfite Bsmooth <sup>27</sup> (version 0.4.5-beta) sur les lectures de séquençage HiSeq 2000 appariées de 100 x 100 nt obtenues pour chaque groupe de serveurs Queen et Worker. Nous avons utilisé le pipeline d'alignement basé sur Bowtie 2 de Bsmooth, qui utilise une version de l'approche de conversion in-silico bisulfite non polarisée et efficace introduite par Lister *et al* <sup>28</sup>. Nous avons utilisé Bowtie 2 <sup>29</sup> version 2.0.0-beta5. Nous aligné sur un indice de référence composé du Baylor séquençage du génome humain Centre *A. mellifera* version 4.0 assemblage, la *A. mellifera* séquence mitochondriale, et le génome du phage lambda. [Le tableau supplémentaire 6](#) résume les résultats de l'alignement.

Nous avons ensuite utilisé Bsmooth pour extraire des mesures au niveau de la lecture. Une mesure au niveau de la lecture correspond à un cas où une lecture alignée chevauchait un CpG dans le génome de référence. La mesure enregistre la position génomique du CpG, l'allèle observé lors de la lecture, sa qualité de base, la qualité de la cartographie de l'alignement et d'autres mesures connexes.

À l'aide de Bsmooth, nous avons filtré les mesures au niveau de la lecture de trois manières. Premièrement, nous avons supprimé les mesures de niveau lecture des alignements avec une qualité de mappage inférieure à 10. Deuxièmement, nous avons supprimé les mesures de niveau lecture où l'allèle dans l'alignement n'était ni C ni T. Troisièmement, nous avons supprimé les mesures de niveau lecture des cycles de séquençage que nous avons jugés non fiable après inspection visuelle du graphe de «biais M». C'est-à-dire que nous avons tracé la fraction des mesures de niveau de lecture méthylées par rapport au cycle de

séquençage. Idéalement, ce graphique devrait être plat et horizontal, ce qui indique l'absence de relation forte entre le cycle de séquençage et la fraction des mesures de niveau de lecture méthylées. En pratique, nous avons trouvé des pics et des creux aux extrêmes des deux partenaires. Nous avons filtré les mesures des cycles affectés. Dans le cas de cet ensemble de données,

Après filtrage, nous avons utilisé Bsmooth pour trier les mesures au niveau de la lecture par coordonnée du génome et les compiler dans un tableau résumant les mesures de méthylation à chaque CpG du génome de référence. Pour estimer le taux de conversion du bisulfite, nous avons utilisé les mesures résumées de la méthylation sur le génome lambda, que nous supposons entièrement non méthylées. [Le tableau supplémentaire 7](#) récapitule les mesures de niveau de lecture obtenues, leur filtrage et les taux de conversion estimés du bisulfite. Après filtrage, nous avons uniquement inclus les preuves avec un score de qualité supérieur ou égal à 20 pour un CpG particulier.

Nous avons ensuite utilisé Bsmooth pour lisser les données et déterminer la corrélation en méthylation entre les données WGBS et les données CHARM. Comme environ 85% des CpG ayant plus de 25% de méthylation sont situés dans des gènes, nous avons comparé les niveaux de méthylation entre WGBS et CHARM en segmentant les gènes dans des fenêtres de 1000 pb et avons trouvé la valeur moyenne de méthylation lissée dans chaque fenêtre. Chaque fenêtre devait contenir au moins 4 sondes CHARM et 8 CpG. Cette analyse a été effectuée pour chaque échantillon de reine et d'ouvriers et les valeurs de corrélation vont de 0,691 à 0,807, moyenne = 0,755 ( [figures supplémentaires 1a à b](#) ). Les profils de méthylation ont été comparés pour déterminer la reproductibilité de la détection des régions de méthylation ( [Figure 1c-f supplémentaire](#) ).

Afin de déterminer s'il y avait une différence entre les reines et les travailleuses, nous avons effectué l'analyse suivante. Premièrement, nous avons uniquement analysé les CpG couvrant 10 échantillons de plus de 10 échantillons. Pour chacun de ces CpG, nous avons effectué un test t pour déterminer la différence de moyenne de méthylation entre les deux groupes. Le test t nous permet de mesurer la variabilité biologique. Les valeurs de p ont été corrigées pour de multiples tests en utilisant la procédure Benjamini-Horchberg et aucun CpG n'était significatif à un taux de découverte de 0,05 faux.

### Analyse des données du séquençage du bisulfite du génome entier par la nourrice et le retourné

Les mêmes méthodes ont été utilisées pour analyser les données de séquençage de Forager et Nurse inversées, ainsi que celles utilisées pour analyser les données de séquençage de Queen et Worker (voir ci-dessus). [Le tableau supplémentaire 8](#) récapitule les résultats de l'alignement, tandis que le [tableau supplémentaire 9](#) récapitule les mesures obtenues au niveau de la lecture, leur filtrage et les taux de conversion estimés en bisulfite.

Pour déterminer si les différences de méthylation entre les cultures fourragères et les infirmières reconverties découvertes à l'aide de CHARM existent également dans le jeu de données WGBS, des tests t ont été effectués sur des CpG individuels à moins de 500 pb des DMR de réversion CHARM. La différence moyenne entre les trois meilleurs CpG classés par importance dans le DMR a été calculée et comparée à la méthylation moyenne de CHARM dans le DMR. Ces résultats sont présentés sous forme de diagramme de dispersion dans la [Figure 6 supplémentaire](#) .

Afin d'évaluer l'importance du chevauchement dans la direction du changement entre CHARM et WGBS, nous avons utilisé le test suivant: en supposant qu'il n'y ait pas de corrélation entre CHARM et WGBS, il devrait exister une probabilité de 50% que la direction du changement soit la même. Par conséquent, nous calculons la probabilité d'observer une statistique plus extrême par  $P(X > 45) + P(X \leq 11)$ , X étant distribué binomialement avec 57 essais et 0,5 chance de succès.

### Analyse des données du séquençage de l'ARN infirmière et de la pépinière

Nous avons utilisé TopHat <sup>30</sup> v1.3.3 pour aligner les lectures de séquençage HiSeq 2000 appariées de 100 x 100 nt obtenues pour chaque pool d'infirmières et de groupes fourragers inversés. Nous nous sommes alignés sur un indice de référence constitué du centre de séquençage du génome humain Baylor, version 4.0 de l'assemblage de *A. mellifera* et de la séquence mitochondriale de *A. mellifera*. [Le tableau supplémentaire 10](#) résume les résultats de l'alignement.

Les alignements de TopHat ont été chevauchés avec des exons annotés obtenus à partir de NCBI pour former un tableau des comptes de chevauchement par gène et par échantillon. On a dit qu'un alignement chevauchait un exon s'il y avait une position de référence couverte à la fois par l'exon et par l'alignement. Un alignement épissé qui couvrait un exon sans que l'un des partenaires ne chevauche l'exon ne compte pas comme chevauchant l'exon.

Les jonctions et les nombres de jonctions émis par TopHat ont été combinés pour former un tableau de nombres par jonction et par échantillon. Deux jonctions non identiques mais dont les limites ne différaient pas de plus de 5 nt d'un côté ou des deux étaient considérées comme identiques. Cela était nécessaire car il existe une certaine variabilité quant à l'emplacement exact des limites de jonction dans TopHat. Le nombre de jonctions a été utilisé pour déterminer les jonctions exon exprimées de manière différentielle entre les cultures fourragères et les infirmières reconverties. Nous définissons les événements d'épissage alternatifs comme la présence d'au moins deux jonctions exon distinctes qui ont une expression opposée dans le même gène ou DMR. Pour déterminer la fréquence des événements d'épissage alternatifs dans les DMR, nous avons étendu le DMR de 500 pb de chaque côté et vérifié la présence de jonctions exon qui étaient davantage exprimées dans les cultures fourragères co-localisant avec des jonctions exon plus exprimées dans les infirmières infirmées. Des exemples sont présentés dans [Figure 2b](#) et [Figure supplémentaire 7](#).

Ref : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3518384/>