

produites par le métabolisme). À l'inverse d'une encyclopédie, ce modèle est converti en un logiciel qui peut agir sur les données pour faire des prédictions. Un robot scientifique peut alors combiner des modèles scientifiques abstraits avec la robotique de laboratoire pour tester et améliorer automatiquement les modèles.

Quand les chercheurs travaillent, ils formulent des hypothèses, puis les testent expérimentalement. De même, Adam émet des hypothèses nouvelles concernant la biologie de la levure, puis il en déduit les conséquences avec son modèle de métabolisme. Il réalise ensuite des expériences pour voir si les hypothèses concordent avec les observations.

D'abord, Adam a prédit quels gènes pouvaient coder les enzymes. La formulation de telles hypothèses met en jeu différentes capacités de raisonnement (voir l'encadré page 72). Pour ne retenir que les hypothèses les plus vraisemblables, Adam a utilisé ses connaissances. Par exemple, il connaissait une enzyme, la 2-aminoadipate transaminase, dont on ignorait le gène. Cette enzyme catalyse la réaction du 2-oxoadipate avec le L-glutamate pour donner le L-2-aminoadipate et le 2-oxoglutarate (la réaction inverse peut aussi avoir lieu). Ce composé participe à la multiplication des levures ; cette réaction est importante, car elle est une cible potentielle des médicaments antifongiques.

## Poser des hypothèses

Pour formuler une hypothèse, à savoir quel gène de la levure pourrait coder cette enzyme, Adam a interrogé sa base de données pour trouver des gènes d'autres organismes qui codent cette enzyme. Ainsi, chez *Rattus norvegicus* (le rat brun ou rat d'égout), un gène nommé *Aadat* code l'enzyme.

Adam a examiné la séquence d'acides aminés de l'enzyme codée par le gène *Aadat* et a cherché une séquence similaire codée par le génome de la levure. Adam sait que si des séquences d'acides aminés se ressemblent, il est raisonnable d'en déduire qu'elles sont « homologues », c'est-à-dire qu'elles partagent un ancêtre commun. Adam sait aussi que si des séquences sont homologues, alors la fonction de leur ancêtre commun a pu être conservée. Par conséquent, à partir de séquences d'acides aminés similaires, Adam peut conclure que les gènes

correspondants peuvent avoir la même fonction. Adam a trouvé trois gènes de levure ayant des séquences semblables à celle d'*Aadat* : *YER152c*, *YJL060w* et *YJL202w*. Il a supposé que chaque gène code l'enzyme 2-aminoadipate transaminase.

Pour tester ses hypothèses, Adam a utilisé les souches de levure de sa collection, conservées dans son congélateur, où chaque souche a un gène qui lui a été retiré. Le robot a examiné dans des conditions particulières la croissance de trois souches de levure, auxquelles il manquait

## LE PROCESSUS PERMETTANT À ADAM de faire des hypothèses et de confirmer expérimentalement les connaissances nouvelles ne nécessite pas d'intervention humaine.

respectivement les gènes *YER152c*, *YJL060w* et *YJL202w* ; il les a fait pousser avec des substances chimiques qui participent à la réaction catalysée par l'enzyme.

L'étape suivante serait de faire de multiples expériences avec chaque souche. Les financements de la recherche sont toujours limités, et les scientifiques font souvent la course pour être les premiers à résoudre un problème. Nous avons donc configuré Adam de façon à ce qu'il mette au point des expériences efficaces testant les hypothèses rapidement et à moindre coût. Ainsi, Adam suppose que chaque hypothèse a une certaine probabilité d'être vraie.

Cette supposition est controversée, et certains philosophes, tel l'Autrichien Karl Popper, ont rejeté l'idée que des hypothèses puissent avoir des probabilités associées. Toutefois, la plupart des chercheurs supposent tacitement que certaines hypothèses ont plus de chances d'être vérifiées que d'autres. Par exemple, ils se rangent souvent derrière l'idée du « rasoir d'Occam » : toutes choses égales par ailleurs, une hypothèse simple est plus probable qu'une hypothèse plus complexe. Adam prend également en compte le coût des expériences, que nous définissons comme le prix des produits chimiques utilisés. Une approche plus exacte considérerait aussi le « coût » du temps nécessaire.

Étant donné un ensemble d'hypothèses associées à des probabilités et un ensemble d'expériences ayant un coût, le but que nous avons fixé à Adam est de choisir une série d'expériences qui minimise le coût et élimine toutes les hypothèses

excepté une. Suivre cette approche de façon optimale est difficile à programmer, mais nos analyses ont montré que le choix réalisé par Adam sélectionne des expériences qui résolvent les problèmes plus rapidement et pour un meilleur coût que les autres choix possibles ; par exemple, il met en œuvre l'expérience la moins chère. Dans certains cas, Adam peut concevoir une seule expérience qui vérifie de nombreuses hypothèses. Les êtres humains ont du mal à en faire autant : ils ne considèrent en général qu'une hypothèse à la fois.

Quand le système d'intelligence artificielle d'Adam a déterminé les expériences les plus prometteuses, il utilise ses équipements robotiques pour les réaliser et observer les résultats. Adam n'est pas capable de voir directement les gènes

ou les enzymes ; il mesure quelle quantité de lumière traverse les cultures de levure. Avec ces données, Adam met en œuvre une chaîne de raisonnements compliquée qui lui permet de savoir si oui ou non les indices s'accordent avec les hypothèses sur les gènes et les enzymes. Ces chaînes de raisonnement existent souvent en science ; par exemple, les astronomes déduisent ce qui se passe dans des galaxies lointaines à partir du rayonnement qu'ils observent avec leurs instruments.

## Faire des découvertes

Adam a produit et confirmé expérimentalement 20 hypothèses suggérant quels gènes codent des enzymes de la levure. Comme toute découverte scientifique, celles d'Adam devaient être confirmées. Nous avons donc vérifié les conclusions d'Adam avec d'autres sources d'information qui ne lui étaient pas accessibles et de nouvelles expériences que nous avons réalisées nous-mêmes. Sept des conclusions d'Adam étaient connues, une semblait fausse, et 12 constituaient des nouveautés scientifiques.

En vérifiant ces résultats, nous avons confirmé que trois gènes (*YER152c*, *YJL060w* et *YJL202w*) codent l'enzyme 2-aminoadipate transaminase. Pourquoi nul n'avait découvert la fonction de ces gènes auparavant ? Parce que ces trois gènes codent la même enzyme, et celle-ci peut catalyser différentes réactions ; nous n'avions pas affaire à une simple correspondance entre un gène et une fonction enzymatique (le scénario classique). Les expériences et

## UN ROBOT QUI FAIT DE LA BIOLOGIE

Adam est un robot scientifique créé à l'Université d'Aberystwyth, au pays de Galles. Il étudie les gènes et leurs fonctions. Récemment, son système d'intelligence artificielle a formulé 20 nouvelles hypothèses portant sur les gènes de la levure susceptibles de coder des enzymes indispensables à sa multiplication. Puis Adam a réalisé des milliers d'expériences pour déterminer quelles hypothèses étaient justes ou fausses. Voici comment il a procédé.

### 1 Préparation des échantillons

Les bras du robot récupèrent des souches de levure congelées et les placent dans différents puits de plaques de culture qui contiennent un liquide nécessaire à leur développement – un milieu de culture (voir le cartouche en haut).

### 2 Croissance des levures

L'incubateur chauffe les plaques pendant 24 heures. Toutes les 40 minutes, un bras du robot installe chaque plaque de culture dans un lecteur optique qui contrôle la multiplication des levures (voir le cartouche en bas).

### 3 Séparation des cellules

Une centrifugeuse fait tourner chaque plaque pour séparer la levure de son milieu de culture.

Congélateur

Incubateur

Lecteur optique

Élimination des déchets

### 4 Élimination du milieu

Des lavages successifs éliminent la totalité du milieu de culture et les levures sont mises en suspension dans une solution saline.

### 5 Traitement des levures

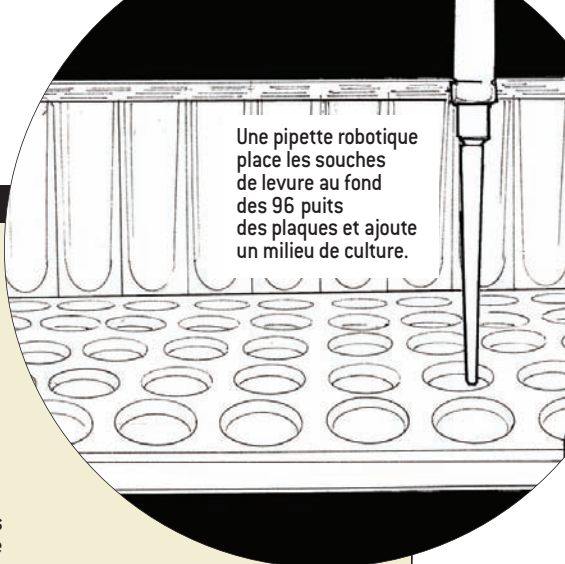
Pour tester différentes hypothèses, le robot ajoute aux levures diverses combinaisons de milieux de culture et de molécules qui pourraient modifier leur métabolisme.

### 6 Lecture des résultats

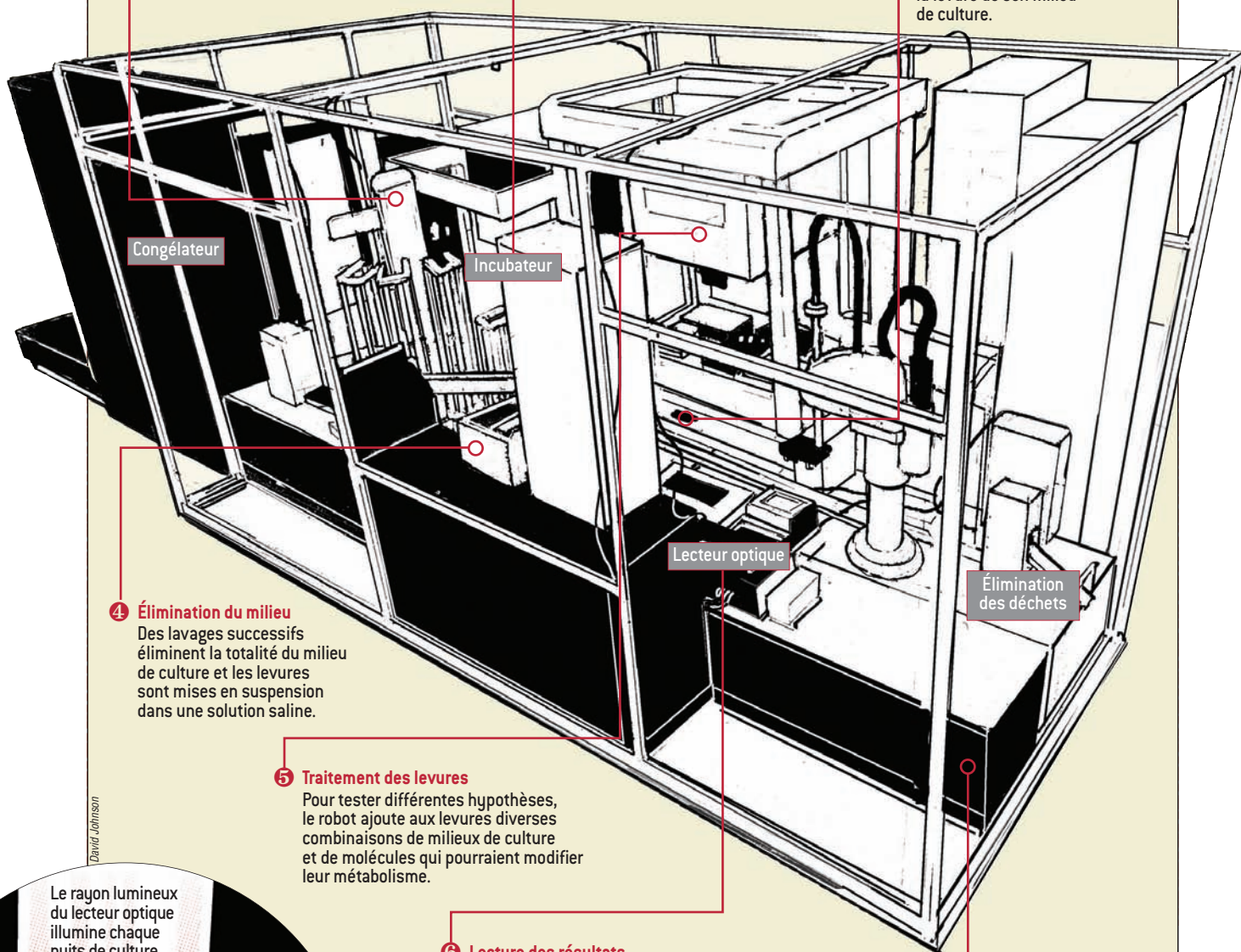
Les plaques restent dans l'incubateur plusieurs jours. Toutes les 20 minutes, un bras robotisé place les plaques dans le lecteur optique qui envoie les données de croissance des levures à un ordinateur.

### 7 Analyse des données

Les logiciels d'Adam analysent les résultats, ce qui peut prendre plusieurs heures.



Une pipette robotique place les souches de levure au fond des 96 puits des plaques et ajoute un milieu de culture.



David Johnson

Le rayon lumineux du lecteur optique illumine chaque puits de culture. L'intensité lumineuse qui traverse le puits mesure de combien les levures se sont multipliées.

