

Progrès dans le diagnostic moléculaire des tumeurs

L'analyse du méthylome de l'ADN

Dr méd. Jürgen Hench^a, Prof. Dr méd. Markus Tolnay^a, Prof. Dr méd. Stephan Frank^a

^a Abteilung Neuropathologie, Institut für Medizinische Genetik und Pathologie, Universitätsspital Basel



Vous trouverez l'éditorial relatif à cet article à la page 134 de ce numéro.

Depuis peu, l'analyse de la méthylation de l'ADN du génome entier permet de définir le type tumoral avec une bonne précision diagnostique. La technologie peut également être utilisée pour la détermination de marqueurs prédictifs et pronostiques majeurs, devenant ainsi une composante de plus en plus incontournable dans la prise en charge interdisciplinaire des patients cancéreux.

Introduction

Durant la transformation des cellules souches indifférenciées en cellules ou tissus différenciés, des activités génétiques spécifiques sont requises. Ces dernières sont régulées par des modifications chimiques des composants de l'ADN et des protéines histones. Le code génétique ne s'en trouve pas influencé, on parle de modifications épigénétiques (cf. glossaire dans le tableau 1).

Les profils épigénétiques résultant des modifications chimiques présentent une spécificité tissulaire et sont

en outre influencés par des mutations associées aux tumeurs. Ainsi, chaque profil épigénétique représente une signature tumorale spécifique, qui est un apport important pour le diagnostique et la médecine oncologique personnalisée. En particulier la détermination des *profils de méthylation de l'ADN* du génome entier (tab. 1) permet non seulement une classification diagnostique précise et spécifique dans une entité au sein d'un vaste spectre de tumeurs (carcinomes, sarcomes, tumeurs cérébrales, mélanomes, néoplasies lymphoprolifératives), mais elle peut aussi largement contribuer à l'optimisation du traitement via la détermina-

Tableau 1: Glossaire

Amplification génique prédictive: Les tumeurs avec des amplifications de gènes des récepteurs de facteurs de croissance présentent souvent une hyperactivité des voies de transduction du signal en aval; des points d'attaque thérapeutiques peuvent en découler (par ex. traitement par trastuzumab du carcinome mammaire avec amplification HER2).

CUP syndrome: Survenue de métastases sans tumeur primaire connue («cancer of unknown primary»). Parfois, les métastases n'arrivent pas à être attribuées à un tissu d'origine au moyen de méthodes histologiques, mais uniquement au moyen de méthodes moléculaires.

FISH: Hybridation in situ en fluorescence. La méthode est entre autres utilisée pour détecter des amplifications de gènes (par ex. du gène HER2).

«Machine learning»: Sert à générer des connaissances à partir de l'expérience dans les systèmes informatiques. Les algorithmes sous-jacents «apprennent» à partir d'exemples, ils les généralisent et ils sont ensuite capables de classer de nouvelles données inconnues sur cette base.

Méthylation de l'ADN: Concerne avant tout la cytosine, un composant de l'ADN fréquent dans les séquences promotrices. Les promoteurs hyperméthylés sont difficilement accessibles pour les facteurs de transcription et ils sont dès lors associés à une activité réduite des séquences génétiques en aval.

Modifications épigénétiques: Par le biais d'une modification chimique du composant de l'ADN cytosine (méthylation) ou des protéines histones (acétylation, méthylation), l'accessibilité des séquences d'ADN pour les facteurs de transcription et donc l'activité des gènes qui y sont codés sont régulées.

Puce à ADN: Systèmes d'analyse moléculaire miniaturisés pour la quantification simultanée de milliers de séquences d'acides nucléiques prédéfinies. En plus de la détection de changements du nombre de copies et de certaines mutations ponctuelles, ces systèmes permettent également de faire la distinction entre les séquences d'ADN méthylées et non méthylées.

TCGA («The Cancer Genome Atlas»): Grand projet américain destiné à collecter et cataloguer les mutations génétiques dans le cadre des cancers. Les données collectées, y compris les profils de méthylation de l'ADN des tumeurs, sont en partie accessibles au public.

Technologie des nanopores: Procédé de séquençage de l'acide nucléique; mesure les signaux électriques générés lors du passage de nucléotides à travers des nanopores biosynthétiques. Le principe ressemble à celui du compteur Coulter (compteur de cellules sanguines) utilisé en hématologie.

UMAP («uniform manifold approximation and projection»): Algorithme destiné à réduire la dimension des données multidimensionnelles, par ex. du profil de méthylation de l'ADN du génome entier. Sert entre autres à la détermination non supervisée de la similitude d'ensembles de données au sein d'un collectif de données. Méthode développée à partir de t-SNE («t-distributed stochastic neighbour embedding»).



Jürgen Hench

tion concomitante de marqueurs prédictifs majeurs. Par conséquent, bon nombre d'analyses immunohisto-chimiques et moléculaires complémentaires souvent chronophages et coûteuses qui étaient jusqu'alors indispensables deviennent superflues, si bien que l'analyse du méthylome de l'ADN représente de plus en plus une modalité diagnostique attrayante sur le plan économique. Aujourd'hui déjà, les algorithmes de «*machine learning*» (tab. 1) sur lesquels se fonde la classification des tumeurs basée sur le méthylome présentent un pouvoir discriminant nettement supérieur à celui du diagnostic immunohisto-chimique-histologique conventionnel. Dans cet article, nous souhaitons non seulement présenter le principe de base, mais aussi et surtout le potentiel de cette méthode.

Analyse du méthylome

L'analyse du méthylome repose sur des puces à ADN ou *DNA microarrays* («Infinium Methylation BeadChips», Illumina®) (tab. 1), qui déterminent la méthylation de l'ADN actuellement au niveau d'environ 850 000 sites du génome, de préférence au sein des régions pro-

motrices. Les données brutes des puces peuvent en outre être converties en un profil de nombre de copies génomiques. Ce profil (fig. 1) peut faire l'objet d'une interprétation diagnostique par le pathologiste, permettant ainsi par exemple la détection d'*amplifications géniques prédictives* (par ex. EGFR, CDK4/6 pour les gliomes; ERBB2/HER2 pour les carcinomes mammaires) (tab. 1).

Toutefois, une interprétation des valeurs de méthylation concernant l'activité d'un gène en particulier n'est pas possible. Il est certes généralement admis que les séquences promotrices hyperméthylées conduisent à une régulation négative du gène concerné, mais le profil de méthylation de l'ADN ne renseigne pas directement sur l'expression de gènes spécifiques, car outre la méthylation de l'ADN, de nombreux mécanismes épigénétiques et fonctionnels supplémentaires, y compris des mutations, influencent l'expression protéique.

Par contre, les résultats sur l'ensemble des données se prête parfaitement à des analyses comparatives avec l'utilisation de cas de référence par «*machine learning*». Les algorithmes sous-jacents s'avèrent essentiels aussi

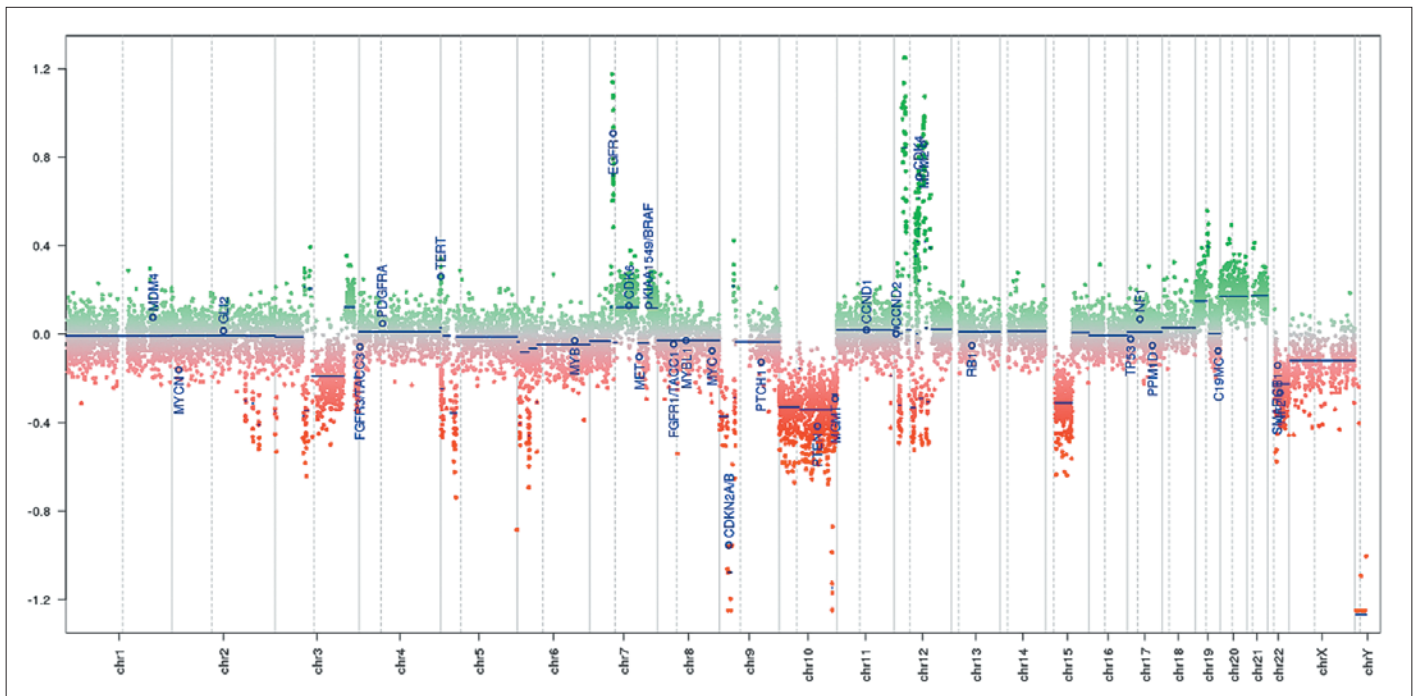


Figure 1: Profil de nombre de copies à l'échelle génomique.

Axe X: Sont représentés les chromosomes 1–22 ainsi que X et Y (de gauche à droite).

Axe Y: La ligne zéro correspond à un état diploïde (2n); la valeur –0,5 signale la perte d'une copie (1n), tandis que la valeur +0,5 signale le gain d'une copie (3n). Ces valeurs se rapportent à une teneur en cellules tumorales de 100% et ne sont dès lors pratiquement jamais atteintes dans les biopsies en raison de l'incorporation de cellules non néoplasiques.

Il s'agit ici du profil de nombre de copies d'un glioblastome avec IDH de type sauvage I (grade IV OMS) de sous-type moléculaire RTKII («receptor tyrosine kinase II»). On reconnaît les gains au niveau du chromosome 7 ainsi que la perte d'une copie du chromosome 10. En outre, une amplification prédictive du gène EGFR (sur le chromosome 7), des amplifications de CDK4 et MDM2 (chromosome 12) et une perte de CDKN2a/b (chromosome 9) sont perceptibles. Calcul et réalisation du graphique au moyen du logiciel libre conumee (<http://bioconductor.org/packages/conumee/>).

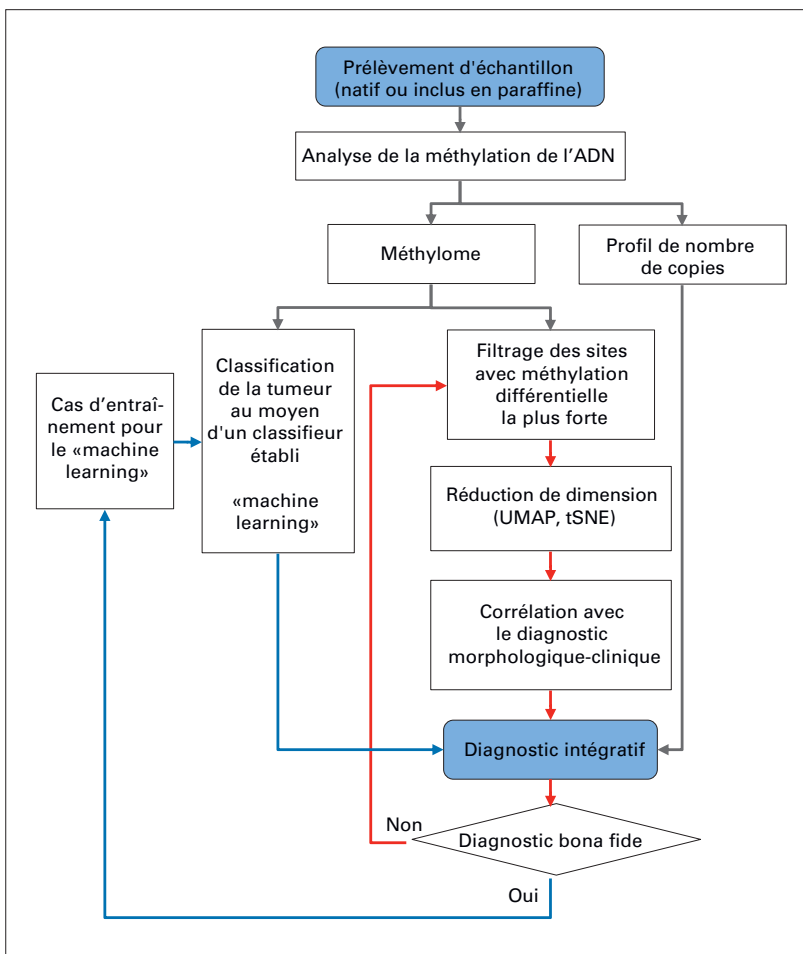


Figure 2: Déroulement du diagnostic basé sur le méthylome. A la différence du diagnostic pathologique classique, le diagnostic basé sur le méthylome ne se termine pas avec la pose du diagnostic, mais tous les cas restent dans le système apprenant. Le circuit marqué en rouge contient les nouvelles entités diagnostiques potentielles, tandis que le circuit marqué en bleu symbolise la contribution des échantillons annotés de façon détaillée. Le diagnostic intégratif transmis sous forme de rapport classique se base sur la synthèse des données cliniques, des résultats des analyses histologiques-immunohistochimiques conventionnelles, des examens moléculaires complémentaires et de l'analyse du méthylome de l'ADN. UMAP: «uniform manifold approximation and projection»; tSNE: «t-distributed stochastic neighbour embedding».

bien pour la constitution de collectifs de cas de référence appropriés que dans le contexte diagnostique (fig. 2).

Lorsqu'une classification se révèle impossible, les ensembles de données passent en continu à travers une sorte de boucle jusqu'à ce qu'ils puissent, en cas de détection d'une similitude avec d'autres cas alors souvent très rares, être attribués à une nouvelle entité diagnostique sur la base de leur profil de méthylation spécifique. Pour les patients et les cliniciens traitants, cela signifie qu'une tumeur rare, qui ne peut pas encore aujourd'hui être définitivement attribuée à un type connu, pourra à l'avenir être assignée à une entité moléculaire précisément définie. Cela ouvre une nouvelle dimension de la médecine personnalisée, car des in-

formations utiles peuvent potentiellement être extraites de chaque ensemble de données collecté en vue de perfectionner les algorithmes de classification par «machine learning». En nomenclature classique, cela signifie qu'une réévaluation a lieu en permanence dans les systèmes informatiques, si bien que tous les diagnostics sont continuellement tenus à jour dans le système.

Contrairement au diagnostic basé sur la morphologie, le profil de méthylation définit une entité connue de manière tranchée. Ceci n'est pas le cas d'un diagnostic basé sur l'histologie qui peut contenir un grand nombre d'entités moléculaires différentes. A l'opposé, des tumeurs ayant une origine très similaire peuvent présenter une histomorphologie qui peut varier [1].

La première utilisation pionnière des modifications épigénétiques dans le domaine du diagnostic pathologique général a été l'élaboration d'un algorithme basé sur les données relatives au méthylome du TCGA (tab. 1), qui était capable de faire la distinction entre les carcinomes, mélanomes, sarcomes, méningiomes et gliomes fréquents, ainsi que quelques néoplasies hémato-lymphoïdes [2]. Ainsi, cet algorithme servait avant tout à classifier la tumeur primaire dans le cadre du CUP syndrome (tab. 1). Depuis peu, l'analyse graphique UMAP (tab. 1, fig. 2) est disponible pour assigner les métastases aux tumeurs primaires correspondantes; grâce à elle, une banque de données portant sur le méthylome actualisée quotidiennement est interrogée lors de chaque pose de diagnostic. L'algorithme UMAP (<http://arxiv.org/abs/1802.03426>) nous permet de comparer simultanément les cas incertains à la fois avec des collectifs de cas diagnostiques de référence établis et avec des cas jusqu'alors non classifiables. Actuellement, nous réalisons une confrontation avec plus de 14500 ensembles de données méthylomiques se rapportant à diverses entités tumorales, ce qui permet aujourd'hui déjà de couvrir un spectre diagnostique extraordinairement vaste. Pour l'UMAP, y compris la préparation des données en amont, seuls des logiciels libres sont utilisés et, sur demande, nous mettons volontiers à disposition nos propres compléments (code source). La création d'une source librement accessible ou d'un serveur est en planification.

Les procédures diagnostiques basées sur le méthylome sont les plus avancées dans le domaine de la neuropathologie. Ainsi, dans notre institut, le «Brain Tumor Methylation Classifier», accessible au public, est déjà utilisé de manière routinière depuis 2017 (<http://www.moleculareuropathology.org> [MNP]) [1]. Nous utilisons cet outil pour le diagnostic de toutes les tumeurs du système nerveux central, y compris des méninges, de l'hypophyse et des gaines nerveuses périphériques,

en conjonction avec l'histomorphologie. La discrimination diagnostique exceptionnellement élevée de ce classificateur de tumeurs cérébrales permet d'atteindre une précision diagnostique jusqu'ici inégalée. Par ailleurs, le «Meningioma Methylation Classifier» [3] permet non seulement une classification diagnostique exacte, mais également une répartition en sous-types pertinents sur le plan pronostique et prédictif [4]; cette estimation du pronostic indépendante de l'histologie s'est révélée plus pertinente par rapport à la gradation basée sur la morphologie de l'Organisation mondiale de la santé (OMS) [3]. Depuis 2018, un classificateur des tumeurs des tissus mous est également disponible sur la plateforme MNP; grâce à cet outil, il est possible de classer de manière reproductible non seulement les sarcomes mais également les tumeurs du système nerveux périphérique (entre autres schwannomes, neurofibromes, tumeurs malignes des gaines nerveuses périphériques). Des classificateurs de méthylation dédiés ont également déjà été créés pour les problématiques auxquelles l'histologie permet à peine ou pas du tout d'apporter une réponse (adénocarcinomes des pommons et de l'intestin; carcinomes épidermoïdes selon leur origine) [5,6].

Perspectives

La précision de la classification basée sur le méthylome dépend principalement de la qualité des ensembles de données de référence, ce qui comprend également la tâche minutieuse mais actuellement non rémunérable financièrement de la curation de l'annotation par les experts. Il en résulte un champ de tension médico-scientifique. S'y ajoute que les coûts de l'analyse du méthylome des tumeurs sont aujourd'hui déjà significativement plus faibles que ceux des analyses combinées associant immunohistochimie, *FISH* (tab. 1) et séquençage génétique diagnostique. En conséquence, il est tout à fait possible que cela mène à des limitations motivées par des raisons commerciales de l'accès public aux systèmes de classification déjà établis [2]. Par ailleurs, la situation de monopole du fabricant des puces (Illumina®) ne doit pas être sous-estimée, même si elle a sans doute été la clé de l'établissement du diagnostic basé sur le méthylome. Dans ce contexte, le développement d'une méthode alternative faisant appel à des appareils basés sur les *nanopores* (Oxford Nanopore Technologies) peut être considéré comme une évolution positive (tab. 1) [7]. Les coûts d'achat de cette technologie (appareil et consommables), qui était initialement conçue pour la recherche sur le terrain, sont nettement inférieurs aux coûts de l'immunohistochimie, de l'analyse *FISH* et du

séquençage parallèle conventionnel, ce qui laisse également apparaître l'utilisation des nanopores par ordinateur portable comme un scénario réaliste dans les régions géographiques moins développées. Un autre avantage de la technologie des nanopores pourrait résider dans les courtes durées de traitement des échantillons: tandis que l'analyse du méthylome basée sur les puces à ADN prend 1–3 semaines en fonction du débit d'échantillons, la technologie de séquençage basé sur les nanopores permet d'obtenir une classification diagnostique des tumeurs en l'espace de 6–10 heures après le prélèvement de tissus. Des premiers tests internes avec de l'ADN isolé à partir de tissu natif se sont déroulés avec succès. Un protocole d'utilisation de tissus inclus en paraffine, qui permettrait également l'analyse d'échantillons tumoraux archivés, est actuellement testé par les auteurs.

En principe, l'analyse du méthylome peut à la fois être réalisée sur du tissu natif et sur du tissu inclus en paraffine archivé en raison de la stabilité élevée de la méthylation de l'ADN. Contrairement aux analyses mutationnelles, les ensembles de données «méthylomiques» livrent des informations intégratives sur les voies de signalisation altérées dans les cellules néoplasiques et réactives et fournissent ainsi aussi des éclaircissements sur le micromilieu tumoral, qui a pendant longtemps été négligé. Il en découle des domaines d'application potentiellement très pertinents sur le plan clinique allant au-delà de la classification diagnostique des tumeurs déjà évoquée, comme par ex. la prédiction de la réponse de la tumeur aux immunothérapies. Par ailleurs, mises à part les affections tumorales, les analyses du méthylome se concentrent actuellement de plus en plus sur les affections réactives et dégénératives. De premières modifications épigénétiques potentiellement exploitables dans le contexte diagnostique ont d'ailleurs déjà été mises en évidence.

Contrairement au diagnostic pathomorphologique classique, tous les ensembles de données méthylomiques générés à un moment donné peuvent toujours être comparés entre eux. Grâce à une curation continue des données assurée par des experts, un système «apprend» les diagnostics les plus rares afin d'ensuite pouvoir les poser lorsqu'il y est confronté. Le profil professionnel du pathologiste qui diagnostique les tumeurs est dès lors amené à évoluer, passant d'un diagnosticien de cas isolés à un curateur global avec une expertise extrêmement spécialisée. Ainsi, l'OMS a récemment défini pour la première fois des entités tumorales sur la base de leurs profils de méthylation avec une objectivité qui n'existait pas jusqu'alors et les a intégrées dans les classifications correspondantes. De plus, les patients dont le méthylome tumoral n'a pu

Correspondance:
Prof. Dr méd. Stephan Frank
Institut für Medizinische
Genetik und Pathologie
Universitätsspital Basel
Schönbeinstr. 40
CH-4031 Basel
stephan.frank[at]usb.ch

être attribué à aucune entité au moment de l'analyse peuvent être informés en cas de correspondance découverte ultérieurement. Le diagnostic basé sur le méthylome représente donc probablement la modalité la plus avancée de la «pathologie numérique», qui a notamment pour ambition de classer les maladies de la façon la plus complète possible, avec une comparabilité à l'échelle mondiale et avec une objectivité maxi-

male. Dans ce contexte, le diagnostic basé sur le méthylome représente aujourd'hui déjà un défi de taille et en même temps une grande opportunité, et ce non seulement pour nous, pathologistes, mais également pour toutes les disciplines spécialisées avec lesquelles nous interagissons.

Disclosure statement

Les auteurs n'ont pas déclaré d'obligations financières ou personnelles en lien avec le présent article.

L'essentiel pour la pratique

- L'analyse du méthylome de l'ADN permet un diagnostic tumoral objectif, robuste sur le plan technique et ainsi très reproductible. Les algorithmes de «machine learning» sous-jacents permettent une précision sans cesse croissante dans le diagnostic de quasiment toutes les tumeurs du système nerveux central, mais aussi de nombreux carcinomes, tumeurs des gaines nerveuses et des tissus mous (sarcomes), mélanomes et néoplasies hématolymphoïdes.
- Dans le cadre du «CUP syndrome», l'analyse du méthylome permet aussi souvent une assignation à la tumeur primaire.
- La méthode convient également pour la détermination de marqueurs pronostiques et prédictifs majeurs (HER2, EGFR, MLH1, MGMT, LOH 1p/19q, etc.).
- La technique fonctionne avec des échantillons inclus en paraffine, mais les tissus natifs, pour autant qu'ils soient disponibles, doivent néanmoins être privilégiés.
- Les coûts du diagnostic basé sur le méthylome sont comparativement faibles; ils correspondent à peu près à ceux d'une analyse FISH et sont donc plus bas que ceux des analyses de séquençage de nouvelle génération.

Références

- 1 Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, et al. DNA methylation-based classification of central nervous system tumours. *Nature*. 2018;555(7697):469–74.
- 2 Moran S, Martínez-Cardús A, Sayols S, Musulén E, Balañá C, Estival-Gonzalez A, et al. Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. *Lancet Oncol*. 2016;17(10):1386–95.
- 3 Sahn F, Schrimpf D, Stichel D, Jones DT, Hielscher T, Schefzyk S, et al. DNA methylation-based classification and grading system for meningioma: a multicentre, retrospective analysis. *Lancet Oncol*. 2017;18(5):682–94.
- 4 Paramasivam N, Hübschmann D, Toprak UH, Ishaque N, Neidert M, Schrimpf D, et al. Mutational patterns and regulatory networks in epigenetic subgroups of meningioma. *Acta Neuropathol (Berl)*. 2019;138(2):295–308.
- 5 Jurmeister P, Bockmayr M, Seegerer P, Bockmayr T, Treue D, Montavon G, et al. Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases. *Sci Transl Med*. 2019;11(509):eaaw8513.
- 6 Jurmeister P, Schöler A, Arnold A, Klauschen F, Lenze D, Hummel M, et al. DNA methylation profiling reliably distinguishes pulmonary enteric adenocarcinoma from metastatic colorectal cancer. *Mod Pathol*. 2019; 32(6):855–65.
- 7 Euskirchen P, Bielle F, Labreche K, Kloosterman WP, Rosenberg S, Daniau M, et al. Same-day genomic and epigenomic diagnosis of brain tumors using real-time nanopore sequencing. *Acta Neuropathol (Berl)*. 2017;134(5):691–703.