

L'arrivée des langues indo-européennes en Europe

Ruth Berger

La paléogénétique et la linguistique assistée par ordinateur livrent de nouveaux indices sur la diffusion des langues.

Avant que les langues indo-européennes ne conquièrent l'Europe, on y parlait notamment une langue apparentée au basque actuel, le vascon. Ainsi, en Europe occidentale et centrale, beaucoup de fleuves portent des noms dérivés de termes basques. Mais quand et comment les langues indo-européennes ont-elles supplanté celles de l'ancienne Europe ? Comme nous le verrons dans cet article, des études récentes de génétique et de linguistique ont apporté de nouveaux éléments de réponse à cette question.

Longtemps a prévalu l'hypothèse de l'archéologue américano-lituanienne Marija Gimbutas (1921-1994), selon laquelle des peuples de cavaliers nomades originaires des steppes entourant la mer Noire et la mer Caspienne auraient commencé, il y a 6000 ans au plus, à se répandre par vagues successives en Europe et en Asie. Quelques indices archéologiques en témoignent, mais ils n'ont pas apporté de réponse claire. En tout cas, ces peuples des steppes et leur culture patriarcale guerrière semblaient concorder avec certains éléments que l'on croyait connaître depuis le XIX^e siècle sur les anciens Indo-Européens et leur « patrie d'origine ».

L'archéologue britannique Colin Renfrew a élaboré à la fin des années 1980 une

hypothèse concurrente, selon laquelle les Indo-Européens seraient originaires d'Anatolie. Ils seraient arrivés bien plus tôt, il y a plus de 8000 ans, et ces colons auraient apporté l'agriculture à une Europe sauvage et faiblement peuplée.

La thèse de C. Renfrew est attrayante, parce qu'elle est simple : avec l'introduction de l'agriculture, l'Europe a connu sa plus grande révolution préhistorique et il tombe alors sous le sens de ramener la propagation des langues indo-européennes au même événement. Cela suppose toutefois que l'agriculture ait effectivement été introduite en Europe par l'immigration. C'est précisément ce dont doutent beaucoup d'archéologues. Selon eux, la population européenne ancestrale s'est convertie elle-même à un mode de vie agricole, contaminée en quelque sorte par le contact avec l'Anatolie, mais sans échanges importants de populations.

La génétique peut-elle nous aider à aller plus loin ? Depuis les années 1980, les

1. CES MASQUES SCULPTÉS de la ville de Myre, de l'ancienne Lycie (dans la région de l'actuelle Antalya, dans le Sud de la Turquie), semblent parler. Le lycien, supplanté par le grec au I^{er} siècle avant notre ère, est l'une des plus anciennes langues indo-européennes.



chercheurs tentent de déduire les anciens mouvements migratoires de l'étude des gènes d'individus vivant aujourd'hui. En Europe, ils ont trouvé une diminution progressive de la fréquence de certains variants de gènes du Sud-Est vers le Nord-Ouest. On a supposé que cela représentait effectivement une vague d'immigration, mais il était impossible de dire quand elle aurait eu lieu.

De telles déductions sont incertaines. Supposons que vous deviez déduire du paysage politique actuel quels étaient les groupes parlementaires composant le parlement en 1848. Il y a peu de chance que cela donne un résultat exact. Il en va de même pour les généalogies génétiques : il existe toujours plusieurs scénarios historiques aboutissant à la répartition actuelle, et de nombreux détails du passé resteront inconnus à jamais. Les progrès réalisés ces dernières années dans la lecture de l'ADN ancien représentent donc une aubaine.

Les hypothèses concurrentes relatives à la « révolution néolithique » en Europe, qu'il y ait eu immigration ou non, ne peu-

vent être testées directement que par des découvertes de fossiles.

Les chercheurs travaillant avec Barbara Bramanti, à l'Université de Mayence, ont isolé de l'ADN mitochondrial (voir l'encadré page 38) sur 26 squelettes des premiers agriculteurs de l'Est de l'Europe centrale (il existe deux types d'ADN, l'ADN du noyau cellulaire et l'ADN des mitochondries ; ce dernier est transmis par les mères). Les agriculteurs appartenaient à la civilisation dite de la céramique rubanée (ou culture rubanée), qui annonçait il y a quelque 7500 ans l'époque agricole dans nos régions. L'équipe de B. Bramanti a par ailleurs prélevé des échantillons d'ADN sur des restes de chasseurs-cueilleurs ayant vécu à la même époque.

Quand l'ADN témoigne

Résultat : les segments d'ADN analysés différaient considérablement entre les chasseurs et les agriculteurs. Ils paraissaient même génétiquement plus éloignés que les autochtones d'Australie le sont des

L'ESSENTIEL

✓ L'origine des langues indo-européennes reste une question très débattue, mais qui est aujourd'hui éclairée par des analyses d'ADN et des arbres de parenté linguistique élaborés par ordinateur.

✓ Deux hypothèses s'affrontent. L'une affirme que ces langues ont gagné l'Europe via des migrations de peuplades de guerriers nomades originaires des steppes entourant la mer Noire et la mer Caspienne, il y a moins de 6 000 ans.

✓ L'autre hypothèse invoque des migrations d'agriculteurs d'Anatolie, il y a plus de 8 000 ans.



L'AUTEUR



Ruth BERGER, biologiste et linguiste, a publié en 2008 un livre sur l'évolution des langues, *Warum der Mensch spricht* (Pourquoi l'homme parle).

Européens. Les premiers agriculteurs d'Europe centrale, appartenant à la civilisation de la céramique rubanée, étaient donc des immigrants. Cela concorde parfaitement avec la thèse de C. Renfrew, selon laquelle l'indo-européen serait arrivé en Europe par l'intermédiaire d'agriculteurs anatoliens.

Cependant, comme souvent, le diable se cache dans les détails. Certes, les chasseurs-cueilleurs ayant vécu à l'époque archaïque, sur laquelle a porté l'étude de B. Bramanti, contrastaient fortement avec les agriculteurs de la culture rubanée ; leur réservoir de gènes appauvri est typique de petits groupes isolés. Mais les variants d'ADN mitochondrial trouvés chez les chasseurs de l'âge de pierre (les variants U4 et U5) existent encore en Europe, avec une fréquence de 11 pour cent pour U5. De nos jours, le variant U4, rare en Allemagne, se rencontre avec la fréquence la plus élevée en Europe de l'Est, ainsi qu'au

Tadjikistan et dans les régions voisines ; il était considéré jusqu'à présent comme typiquement indo-européen.

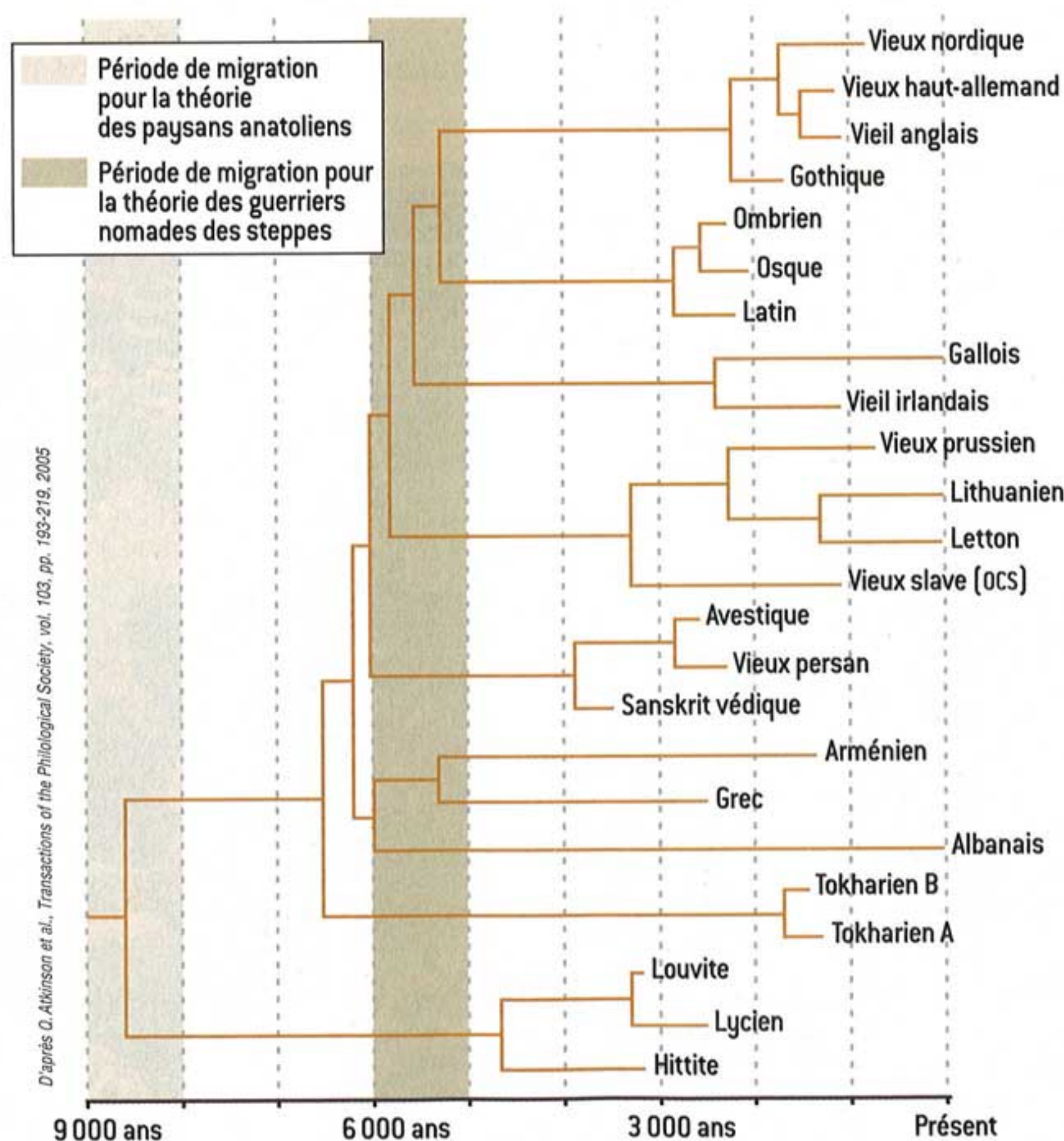
L'étude de B. Bramanti réservait une autre surprise : les agriculteurs de la culture rubanée présentaient une répartition des gènes différente de celle des Européens actuels ; la plupart d'entre eux portaient le variant N1a ou H de l'ADN mitochondrial. Le variant H est jusqu'à présent très répandu chez les Européens et même dominant, alors que N1a n'existe presque plus en Europe.

Quelles migrations ?

Ce résultat inattendu pourrait indiquer au moins deux grands événements migratoires au Néolithique de l'Europe centrale. D'abord seraient arrivés les peuples de la culture rubanée, qui auraient repoussé des chasseurs-cueilleurs porteurs des variants U4 et U5. Ils auraient ensuite eux-mêmes été repoussés si fortement par une vague de migration ultérieure, que leur type mitochondrial caractéristique N1a aurait pratiquement disparu. Cela correspond davantage à l'hypothèse de Gimbutas de l'immigration de peuples des steppes au début de l'âge du bronze qu'à la thèse de C. Renfrew. Les peuples de la culture rubanée pourraient avoir été les Basques primitifs et ce ne serait que lors de la migration ultérieure que les Indo-Européens auraient atteint l'Europe. Mais ce pourrait être l'inverse... ou ni l'un ni l'autre : en réalité, le tableau est encore plus complexe.

D'une part, la modification de la répartition des différents types d'ADN pourrait résulter en partie de la sélection et non de la migration. Des variants mitochondriaux auraient ainsi subsisté ou été éliminés selon l'endroit, sans qu'il y ait déplacement d'individus.

D'autre part, la civilisation de la céramique rubanée n'a pas été la seule agriculture primitive en Europe. Dans le Sud, le long des côtes méditerranéennes, il en existait une seconde, moins développée, dont les membres vivaient en partie dans des grottes. La céramique à empreintes de coquillages, caractéristique de la culture dite cardiale (*Cardium* est un genre de coquillage), est attestée au plus tôt en Thessalie dans l'actuelle Grèce ; mais elle s'étend le long de la côte Nord et Sud de la Méditerranée jusqu'à l'Atlantique. On a distingué un précurseur possible au Liban, tandis que les origines des peuples de la civilisation de la céramique rubanée se situent



2. CET ARBRE GÉNÉALOGIQUE des anciennes langues indo-européennes a été établi à l'aide d'une simulation informatique. L'ordinateur a traité des listes de vocabulaire de base de chaque langue à l'aide d'un algorithme statistique. La fiabilité de cet arbre reste à déterminer, tant que l'on ne disposera pas d'éléments d'information obtenus par d'autres moyens. L'analyse a été réalisée avec un autre choix de langues que celui des pages 36 et 37 et a donné des valeurs similaires, mais non identiques. (D'après Q. Atkinson et al., *Transactions of the Philological Society*, vol. 103, pp. 193-219, 2005.)

dans l'espace danubien occidental. La culture cardiale méditerranéenne, contrairement à celle de la céramique rubanée, est ainsi directement liée à l'Orient.

Les origines de certaines composantes du réservoir de gènes européens se trouvent-elles chez les représentants de la culture cardiale? Ces derniers, et non les peuples de la culture rubanée, seraient-ils les Indo-Européens recherchés? En 2005, Helen Chandler et ses collègues, de l'Université d'Oxford, ont mené une étude paléogénétique au Portugal. Là aussi, des chasseurs-cueilleurs locaux ont été comparés avec des agriculteurs du Néolithique et les chercheurs ont travaillé avec des marqueurs de l'ADN mitochondrial, le génome du noyau étant plus difficiles à isoler et séquencer à partir d'os anciens.

Des différences sont aussi apparues entre les chasseurs et les agriculteurs portugais, mais elles se sont révélées beaucoup moins marquées que celles notées chez les sujets examinés par l'équipe de B. Bramanti. Avant tout, les deux groupes portugais divergent de manière similaire de l'actuelle répartition moyenne européenne des différents types mitochondriaux: le variant J, considéré comme caractéristique du Proche-Orient, est totalement absent; en revanche, dans les deux groupes, les variants H et U sont respectivement très fortement et fortement représentés. La diversité est globalement réduite, comme on pouvait s'y attendre pour des populations vivant aux marges du continent à l'âge de pierre.

L'équipe de H. Chandler a comparé les anciens Portugais aux populations actuelles et en a conclu que les chasseurs tout comme les agriculteurs ressemblent, pour ce qui est de leur ADN mitochondrial, essentiellement aux Basques, aux Galiciens et aux Catalans. Du moins au Portugal, la première culture néolithique semble ainsi avoir été représentée par une population indigène vivant sur les côtes méditerranéennes occidentales. Dans cette région, on recherche sans doute les Indo-Européens en vain.

En résumé, la paléogénétique n'a pas apporté jusqu'ici d'éléments étayant la thèse de C. Renfrew, selon laquelle les premiers agriculteurs auraient immigré d'Anatolie. Concernant la civilisation de la céramique rubanée, il y a bien eu une migration, mais elle ne permet de remonter que jusqu'à l'espace danubien. De surcroît, comme la continuité génétique



3. SUR CETTE CARTE, LES COULEURS indiquent la fréquence de divers variants génétiques dans les populations. Les variants les plus fréquents en Orient se raréfient à mesure que l'on va vers le Nord-Ouest (et inversement), ce qui est compatible avec une colonisation en une ou plusieurs vagues à partir du Sud-Est.

directe entre les peuples de la culture rubanée et les Européens actuels fait défaut, il est probable qu'il y ait eu des migrations ultérieures vers la région considérée.

Si la génétique ne répond pas pour le moment à la question de l'origine des Indo-Européens, la linguistique nous aiderait-elle? En 2003 parut une étude, très débattue depuis lors, qui commençait par l'assertion qu'il était possible de trancher entre les thèses de Gimbutas et de C. Renfrew.

Des mots pour un arbre

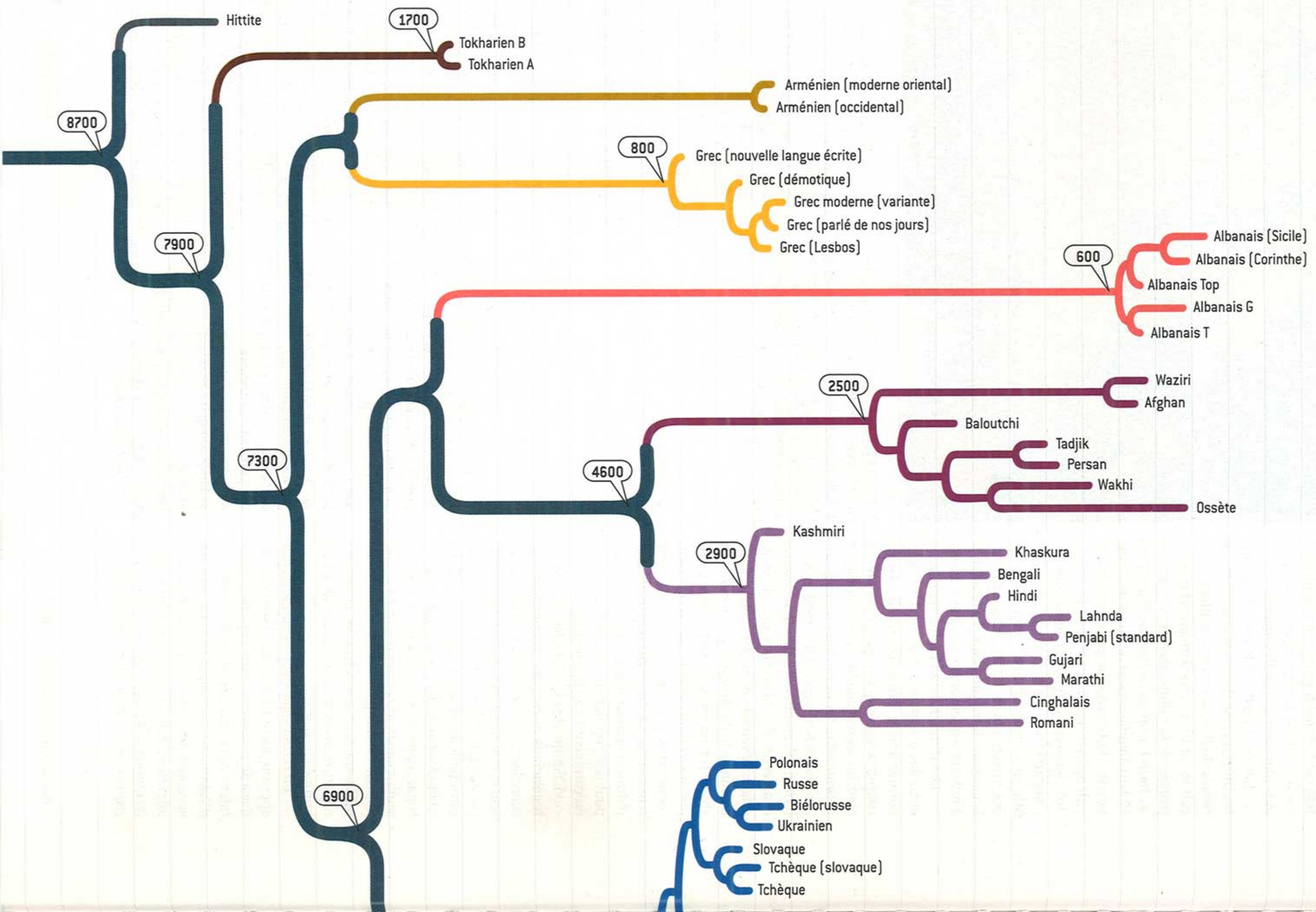
Quentin Atkinson et Russel Gray, deux psychologues de l'Université d'Auckland s'intéressant aux langues, déterminèrent à l'aide d'un programme informatique, qui sert normalement aux généticiens pour calculer les arbres généalogiques des humains, l'âge de la famille des langues indo-européennes. Ils ont fourni à l'ordinateur un vocabulaire de base de 200 mots et leurs équivalents dans les différentes langues indo-européennes, ainsi que des estimations sur le temps qu'il faut pour que dans une langue un mot soit remplacé par un autre d'origine différente. Le programme a alors calculé un arbre des langues et a déterminé l'âge de la racine et de l'ensemble des ramifications.

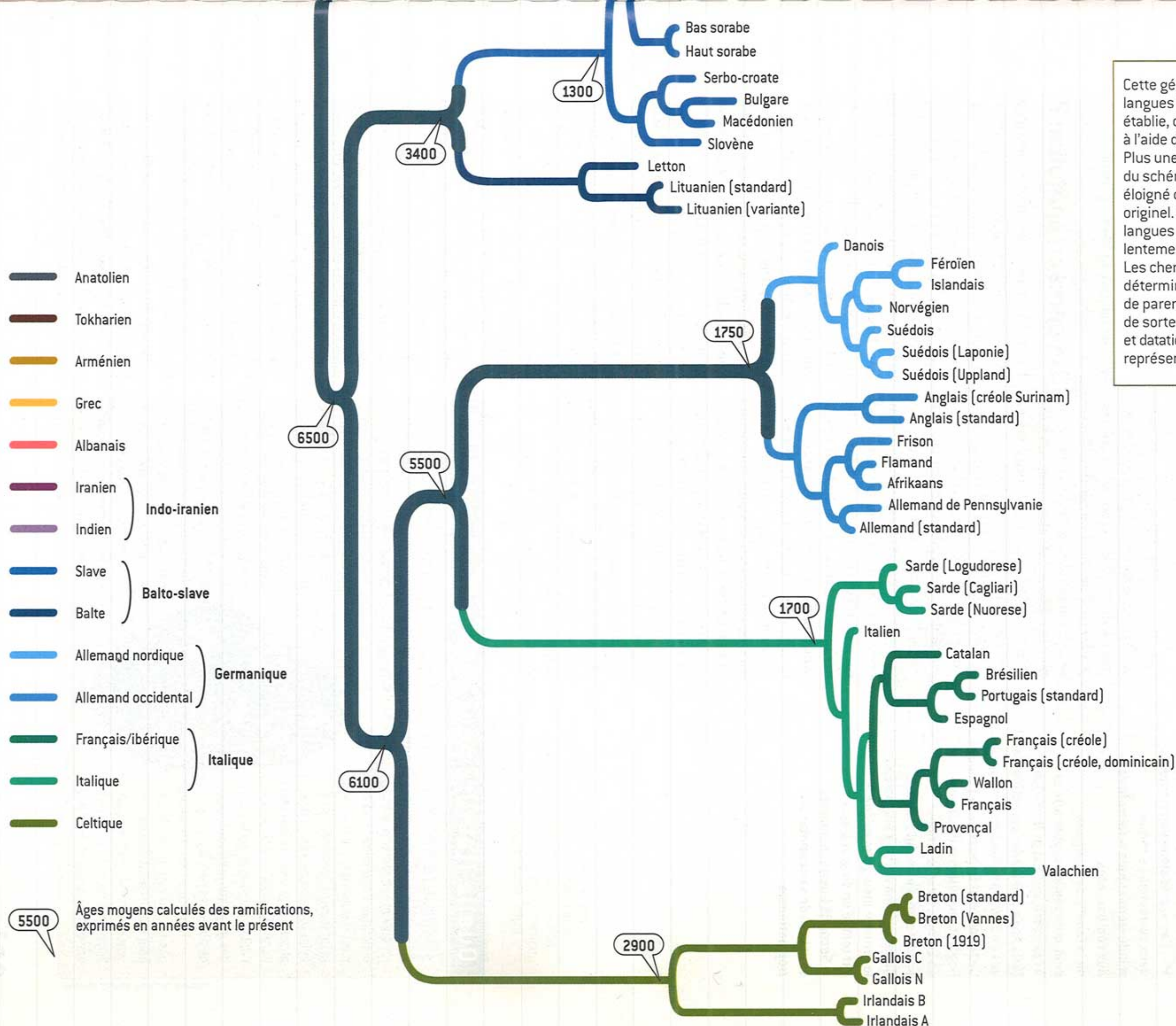
La méthode n'est pas aussi précise qu'il y paraît, les vitesses de modification du vocabulaire étant en fait variables: le contact avec les peuples voisins ou l'adoption d'une langue par un groupe parlant auparavant une autre langue peuvent accélérer

L'archéologie des langues

✓ **Les mots désignant le joug, la roue et l'axe sont très semblables dans les langues indo-européennes. On en a conclu que ces objets ont dû exister avant la ramification de l'indo-européen en langues filles. Ainsi, la famille de langues ne daterait pas de plus de 6 000 ans, car la roue n'existait pas auparavant. Les partisans de la thèse d'une ramification antérieure de l'indo-européen essaient au contraire d'expliquer les ressemblances par des emprunts et des apparitions simultanées survenues par hasard.**

L'arbre de parenté des langues indo-européennes





Cette généalogie de nombreuses langues indo-européennes a été établie, comme l'arbre de la figure 2, à l'aide de méthodes statistiques. Plus une langue est à droite du schéma, plus son vocabulaire s'est éloigné de celui de l'indo-européen originel. Il apparaît ainsi que certaines langues se sont transformées plus lentement que d'autres. Les chercheurs n'ont pas pu déterminer tous les anciens liens de parenté avec certitude, de sorte que d'autres ramifications et datations sont possibles. La variante représentée ici est la plus probable.

5500 Âges moyens calculés des ramifications, exprimés en années avant le présent

Indo-européen ou indo-germanique ?

✓ Ces expressions sont toutes deux correctes. Le terme « indo-germanique » fait allusion aux composantes de la famille des langues indo-européennes rencontrées à l'extrême Sud-Est et Nord-Ouest (Sri Lanka et Islande respectivement) de l'aire de répartition de ces langues. Ce terme s'est imposé dans l'espace linguistique allemand ; ailleurs en Europe, c'est le terme « indo-européen » qui prévaut. Dans certains cercles allemands, le terme « indo-germanique » est tombé en disgrâce après la Seconde Guerre mondiale, en raison de sa connotation trop nationale.

l'échange de mots. Les auteurs se sont efforcés de prendre en compte de tels impératifs. Et d'après Q. Atkinson et R. Gray, les faiblesses de la méthode sont largement compensées par l'utilisation d'un grand nombre de langues et de mots.

L'arbre établi par ordinateur (voir l'illustration pages 36 et 37) ne surprend pas et étaye des hypothèses qui ont été élaborées par des méthodes classiques. Par exemple, le statut de langue isolée de l'albanais se voit confirmé. De même, les grands groupes de langues actuelles, germaniques, romanes, slaves, etc., sont correctement représentés.

Mais la datation est-elle fiable ? Il ne faut pas se laisser leurrer par les valeurs plausibles pour la période historique (par exemple celle de la ramification des langues germaniques et romanes en leurs branches actuelles il y a au moins 1700 ans), car c'est précisément avec de telles dates déduites de connaissances historiques que la méthode a été étalonnée. Supposons ainsi que, pour les dates plus anciennes, elle fournisse au moins une approximation. Après tout, des analyses ultérieures ont conduit à des résultats similaires (voir la figure 2).

D'après la simulation, la première ramification de la langue originelle proto-indo-européenne s'est produite il y a déjà près de 8700 ans. L'hypothèse de Gimbutas serait donc réfutée. Les peuples de cavaliers des steppes de la Russie méridionale, que Gimbutas considérait comme étant les diffuseurs du proto-indo-européen, ont vécu plus de 2000 ans trop tard pour être

les Indo-Européens originels. Une ramification de l'indo-européen originel il y a environ 9000 ans est en revanche en accord avec la thèse de C. Renfrew. En Grèce orientale, les premières traces d'agriculture sont décelables à peine plus tard ; elles proviennent de l'Anatolie voisine et c'est précisément là-bas que l'arbre généalogique situe, avec le hittite, la plus ancienne branche des langues indo-européennes.

Un berceau anatolien ?

Il y a toutefois une difficulté. Le hittite n'est attesté qu'au II^e millénaire avant notre ère, des milliers d'années après la diffusion de l'agriculture. L'arbre ne révèle pas si le préhittite parlé il y a plus de 8700 ans doit aussi être situé en Anatolie, car il ne contient pas d'indications de lieu. Des localisations actuelles ou historiques peuvent tout au plus apporter de maigres indices. Les langues ne sont pas attachées à un lieu. Le celtique, aujourd'hui limité à la Bretagne et aux îles Britanniques, était autrefois parlé de l'Espagne jusqu'au Bosphore. Le turc, principale langue actuelle de l'Anatolie, n'y est arrivé qu'au haut Moyen Âge.

Les langues se déplacent ou disparaissent. L'histoire de l'Europe, de l'Est de l'Asie et des Indo-Européens s'étend sur neuf millénaires, au cours desquels des langues se sont déplacées, se sont modifiées ou ont disparu sans laisser de traces. C'est pourquoi nous ignorons quelle langue on parlait en Anatolie il y a plus de 8000 ans. D'après la thèse de C. Renfrew, il s'agissait du proto-indo-européen, qui serait ensuite arrivé avec l'agriculture dans les Balkans et serait devenu le précurseur du grec et de l'albanais. Le préhittite serait resté isolé en Anatolie, tandis que les autres langues indo-européennes auraient continué à diverger en Europe. Il aurait donc existé dans les Balkans et en Anatolie une continuité indo-européenne de plus de 8000 ans.

Mais il est aussi possible qu'il en fût autrement. L'écriture la plus ancienne d'Europe, déchiffrée partiellement, est le linéaire A sur l'île de Crète. Elle témoigne d'une langue qu'il a été impossible jusqu'ici de classer dans la famille indo-européenne. On débat aussi d'une origine non indo-européenne pour la langue des Pélasgiens, peuple préhistorique de la Grèce pratiquant l'agriculture, ainsi que celle des habitants de la haute Antiquité des îles de Limnos et de Chypre. Et l'Anatolie ? Lorsque les

Quelle fiabilité pour l'ADN mitochondrial ?

Les mitochondries (ci-contre) sont des organites cellulaires qui ont leur propre matériel génétique. L'ADN mitochondrial n'est transmis que par la mère, mute souvent, présente de multiples copies et peut être extrait plus souvent d'os anciens que l'ADN du noyau cellulaire. On l'utilise pour retrouver la trace des migrations préhistoriques de différents peuples.

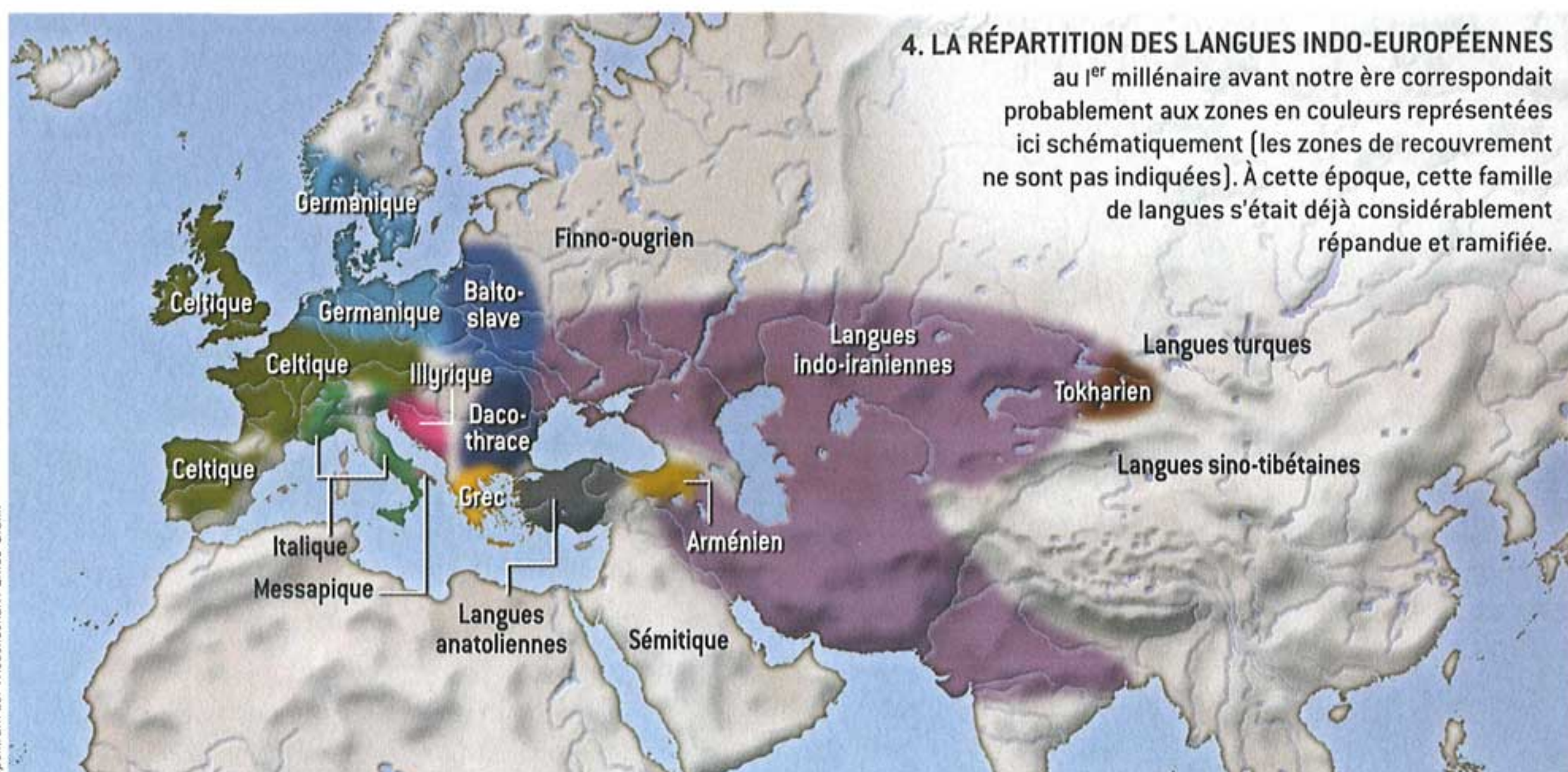
Il ne constitue toutefois pas un indicateur très fiable pour établir l'existence ou l'absence de parenté. D'une part, le génome nucléaire peut avoir une origine en partie différente

de celle des mitochondries. D'autre part, les différents variants de l'ADN mitochon-



drial ne sont pas seulement des marqueurs de parenté dépourvus de fonction.

Certains variants sont associés à des particularités du métabolisme ou à une vulnérabilité à certaines maladies. Dans la répartition actuelle des mitochondries, il est probable que la sélection soit intervenue tout autant que la parenté. Par exemple, le variant mitochondrial N1 pourrait avoir disparu par sélection chez les agriculteurs de la civilisation de la céramique rubanée, bien que ses porteurs soient peut-être nos ancêtres.



4. LA RÉPARTITION DES LANGUES INDO-EUROPÉENNES

au 1^{er} millénaire avant notre ère correspondait probablement aux zones en couleurs représentées ici schématiquement (les zones de recouvrement ne sont pas indiquées). À cette époque, cette famille de langues s'était déjà considérablement répandue et ramifiée.

Hittites commencèrent à y écrire, ils ont immortalisé, en même temps que leur propre langue indo-européenne, celle de l'État de leurs prédécesseurs Hatti, dont ils ont adopté la culture et la capitale. Cette langue n'était certainement pas indo-européenne. Il n'est donc nullement établi que les premiers agriculteurs d'Anatolie ou de Grèce aient été des Indo-Européens.

Un arbre très ramifié il y a près de 6 000 ans

Ainsi, l'arbre auquel sont parvenus R. Gray et Q. Atkinson, avec ses dates, est certes compatible avec la thèse selon laquelle l'indo-européen est arrivé en Europe avec l'agriculture, mais il ne peut la prouver.

L'arbre contient en revanche une autre date qui captive les chercheurs. Outre les langues disparues ou isolées comme le hittite ou le grec, il fixe aussi l'origine des grandes familles de langues indo-européennes : l'indo-iranien, le balte, le slave, le celtique, le romano-italique et le germanique. Toutes ces familles ont une origine commune, qui se situe environ 2000 ans après la divergence de l'hittite. La ramification en grandes familles existant aujourd'hui s'est alors produite dans une fenêtre temporelle de 1 400 ans seulement. Cela indique des migrations postérieures à la plus ancienne diffusion de l'agriculture. La première des grandes familles actuelles à diverger, il y a 6900 ans, fut l'indo-iranien. Ont ensuite suivi les familles des langues du Nord-Est de l'Eu-

rope, le balte et le slave. Il y a 6 100 ans, le celtique s'est séparé du tronc. Enfin, il y a 5500 ans, le reste s'est ramifié en germanique et en italique.

Selon C. Renfrew, cette énorme vague de migration, sur laquelle repose le succès des langues indo-européennes sur un vaste périmètre, est partie des Balkans. Il ne s'agit bien sûr que d'une option. Un coup d'œil sur la carte suggère une autre possibilité : la vague de ramifications et de migrations est peut-être partie des steppes du Nord de la mer Noire, c'est-à-dire de la région habitée par les peuples des steppes qui ont la faveur de Gimbutas. De là-bas, on arrivait facilement en Iran, et à la ramification de l'indo-iranien, aussi bien qu'en Europe de l'Est avec la ramification des langues baltes et slaves. On dispose même d'un élément qui étaye la présence très ancienne d'Indo-Européens dans la région de la mer Noire : les noms que portent les fleuves qui y coulent. Don, Donez, Dniepr, Dniestr, Danube, tous ces noms ont une racine indo-européenne, qui est le terme du vieux persan et du vieux celtique désignant le fleuve (*danu*).

Pour ce qui est de l'origine des Indo-Européens, les peuples des steppes sont donc toujours en lice. Il est possible qu'ils aient contribué à la répartition actuelle des langues indo-européennes. La vérité se situe sans doute quelque part entre les thèses de C. Renfrew et de Gimbutas. Une chose est sûre : dans ce domaine, les hypothèses les plus simples ne sont pas forcément les meilleures. ■

✓ BIBLIOGRAPHIE

B. Bramanti et al., Genetic discontinuity between local hunter-gatherers and Central Europe's first farmers, *Science*, vol. 326, pp. 137-140, 2009.

H. Chandler et al., Using ancient DNA to examine genetic continuity at the Mesolithic-Neolithic transition in Portugal, dans P. Arie et al. (éds.), *Actas del III Congreso del Neolítico en la Península Iberica*, Santander, pp. 781-786, 2005.

R. D. Gray et Q. D. Atkinson, Language-tree divergence times support the Anatolian theory of Indo-European origin, *Nature*, vol. 426, pp. 435-438, 2003.

E. Hamel et Th. Vennemann, Le vascon, première langue d'Europe, *Pour la Science*, n° 299, septembre 2002.

Collectif, *Les langues du monde*, Bibliothèque Pour la Science, 1999.

M. Ruhlen, *L'origine des langues*, Belin, 1997.