

**“ Les β -thalassémies, témoins des migrations préhistoriques
dans le bassin méditerranéen ”**

PERRIN Pascale

Action “ Origine de l’Homme, du Langage et des Langues ”

A. FICHE ADMINISTRATIVE

Titre du projet :

“ Les β -thalassémies, témoins des migrations préhistoriques dans le bassin méditerranéen ”

Mots-clés :

Archéologie/Anthropologie biologique (*Homo sapiens*)/ADN moderne/ ADN ancien/Génétique des populations/ Evolution/ Bases de données/Linguistique

Résumé du projet (10 lignes maximum) :

L’analyse de marqueurs génétiques peut contribuer de façon importante à éclairer l’histoire de nos origines. Le marqueur génétique utilisé ici est une mutation thalassémique, touchant le gène de β -globine (IVS-I-110). Cette mutation très fréquente dans le bassin méditerranéen y présente un gradient est-ouest. L’analyse des haplotypes en amont du gène de β -globine permet de retracer l’origine et la diffusion de cette mutation. Les données actuelles conduisent à proposer son origine dans la région anatolienne. L’identification de cette mutation sur des ADN anciens s’avère donc un excellent marqueur de populations à l’origine du processus de pré-néolithisation et de néolithisation de l’Europe. Afin de préciser l’origine géographique et chronologique de cette mutation, des échantillons seront prélevés sur les restes humains d’une série de sites de l’Anatolie et du levant nord. La confrontation des données archéo-anthropologiques, génétiques et linguistiques aboutiront à une meilleure connaissance des premières sociétés de chasseurs-cueilleurs du Paléolithique supérieur puis des populations du néolithique.

1. Responsable scientifique du projet

Nom **PERRIN**..... Prénom..**Pascale**.....

Grade.....**Maître de Conférences**.....

Discipline du responsable scientifique: ..**Biologie moléculaire, bioinformatique, évolution**

Établissement de rattachement**Université Claude Bernard – Lyon I** -.....

Adresse professionnelle : N°, rue , BP ..

Centre de Génétique Moléculaire et Cellulaire – UMR 5534

43, Boulevard du 11 Novembre 1918

Code postal **69622** Commune **Villeurbanne cedex**

Tél **04.72.43.29.48**

Fax **04.72.44.05.55**

E-Mail : **pperrin@cismsun.univ-lyon1.fr**

2. Laboratoire ou organisme de rattachement de l'équipe de recherche

Intitulé **Centre de Génétique Moléculaire et Cellulaire**

Type de formation (*cocher la case utile*)

- Unités CNRS : unité propre du CNRS unité associée ou mixte du CNRS

Préciser le code unité **UMR 5534**.

Préciser la délégation régionale :**Rhône-Alpes**.....

- Unités hors CNRS : unité universitaire (*Préciser l'université*)

.....
 Autre (*Préciser : EPST, Association, équipe étrangère...*)

Nom du directeur de l'organisme :**Dr Pierre COUBLE**.....

Adresse : N° , rue , BP **43, boulevard du 11 Novembre 1918**

Code postal **69622** Commune ...**Villeurbanne cedex**.....

Tél **04.72.43.13.26**

Fax **04.72.44.05.55**

3. Autre(s) laboratoires ou organisme(s) partenaires

Cette section est à reproduire autant de fois que nécessaire.

Intitulé **Maison de l'Orient méditerranéen Jean POUILLOUX – Université Lyon 2 -
CNRS – UPR 7537**

Discipline(s) couverte(s) par l'équipe : ...**Archéologie, paléoarchéologie, méthodes informatiques
appliquées à l'archéologie**

Nom du directeur de l'équipe **Dr Danièle STORDEUR**

Membre(s) partenaire(s) : et cf figure récapitulative des membres participants

Nom : **LE MORT**

Prénom : **Françoise**

Grade : ...**Chargée de recherche CNRS**.....E-Mail : **flm@mom.fr**

Adresse : N° , rue , BP...**7, rue Raulin.**

Code postal **69365** Commune **Lyon 07**.....

Tél **04.72.71.58.75**

Fax **04.72.72.08.65**

3. Autre(s) laboratoires ou organisme(s) partenaires

Cette section est à reproduire autant de fois que nécessaire.

Intitulé**Evolution Moléculaire et ADN fossile**

Discipline(s) couverte(s) par l'équipe : **Biologie moléculaire, ADN ancien**

Nom du directeur de l'équipe :..**Catherine HÄNNI** ...

Membre(s) partenaire(s) : et Cf figure récapitulative des membres participants

Nom : **HÄNNI**.....Prénom : **Catherine**

Grade : **Chargée de recherche CNRS**.....E-Mail : ...**hanni@cismsun.univ-lyon1.fr**

Adresse : N° , rue , BP **Université Claude Bernard – Lyon I**

Code postal **69622** Commune ..**Villeurbanne cedex**

Tél **04.72.43.29.50**

Fax **04.72.44.05.55**

3. Autre(s) laboratoires ou organisme(s) partenaires

Cette section est à reproduire autant de fois que nécessaire.

Intitulé **Département d'Archéologie, Faculté des Lettres et Sciences Sociales**

Discipline(s) couverte(s) par l'équipe : **Histoire, linguistique**

Membre(s) partenaire(s) :

Nom : ...**NACACHE**.....Prénom : ...**Albert**

Grade : **Professeur** E-Mail : **anaccash@nidal.com**.....

Adresse : N° , rue , BP.....**BP 113 – 6328 Beyrouth, Liban**

Tél **00.961.1.365.284**

Fax **00.961.1.363.306**

B. PROJET SCIENTIFIQUE

(10 pages maximum, bibliographie comprise à insérer ici)

Introduction

Parmi les grands tournants de l'histoire de l'homme moderne, l'émergence d'un langage complexe et la néolithisation constituent les deux éléments les plus déterminants. C'est, en effet, dans ces deux innovations fondamentales que s'enracine l'état présent de l'espèce humaine, non seulement dans le domaine de l'exploitation du milieu mais aussi dans ceux de la culture et des structures mentales. Or c'est au Proche Orient que les phénomènes de néolithisation ont commencé à se développer vers 11 000 B.P. (Cauvin 1997, Aurenche et Kozłowski 1999). L'analyse du spectre génétique des populations européennes révèle que le Proche-Orient a bien participé au peuplement de l'Europe suivant un processus dont les effets sur la composition génétique des populations d'à présent vont en s'atténuant à mesure que l'on s'éloigne du point de départ. C'est plutôt l'Anatolie qui est désignée comme zone de départ de cette diffusion linguistique ; la néolithisation de l'Europe, qui se serait effectuée à partir de l'Anatolie, est considérée comme ayant diffusé les langues proto-indo-européennes, à partir de 8 000 B.P. (Renfrew 1993).

Le complexe de gènes de globines fait partie des loci humains les plus étudiés, en raison du nombre de mutations affectant ses gènes. Certaines sont à l'origine des thalassémies, une des maladies génétiques les plus fréquentes dans le monde (Weatherall et Clegg 1996). En effet, selon l'Organisation Mondiale pour la Santé, plus de 150 millions de personnes sont considérés comme porteurs de -thalassémie, essentiellement dans les régions du monde où sévit la malaria: le bassin méditerranéen, le centre Est africain, l'Inde et l'Asie du Sud. Ces personnes, en général,

assymptomatiques, présentent une microcytose et une anémie légère ou modérée; les hétérozygotes sont protégés contre *Plasmodium falciparum*, l'agent infectieux de la malaria. Ainsi, dans certaines régions impaludées, une α -thalassémie peut présenter une forte prévalence, mais a contrario, dans d'autres régions tout aussi impaludées, elle sera présente à une faible fréquence. Ce remarquable degré d'hétérogénéité ne peut donc être la conséquence exclusive de la sélection par la malaria.

L'analyse moléculaire des chromosomes α -thalassémiques a démontré une grande hétérogénéité des mutations responsables de ces maladies (plus de 180). Parmi elles, deux mutations sont très fréquentes dans le bassin méditerranéen: **le stop codon 39 (C→T)** (introduction d'un codon stop au niveau du codon 39) **et l'IVS-I-110 (G→A)** (mutation G-A au nucléotide 110 de l'intron 1). En effet, la mutation stop codon 39 présente un gradient d'Ouest (96 % des α -thalassémies en Sardaigne (Rosatelli *et al.* 1992)) en Est (40 à 50% au Liban (Chehab *et al.* 1987) et en Turquie (Tadmouri *et al.* 1998)). La répartition de la mutation IVS-I-110 est différente: elle représente environ 80 % des α -thalassémies à Chypre (Baysal *et al.* 1992), 50 % au Liban (Chehab *et al.* 1987), 40 % en Turquie (Tadmouri *et al.* 1998), avec une diminution des fréquences vers les pays de l'ouest méditerranéen. Le fait que ces mutations soient prédominantes dans l'Est de l'Europe pour l'IVS-I-110 et dans l'Ouest de l'Europe pour le Stop Cd 39, suggère qu'elles sont les plus anciennes dans ces deux régions. **L'analyse des spectres de distribution des fréquences de ces mutations s'avère donc essentielle pour retracer l'origine et la diffusion des α -thalassémies dans le bassin méditerranéen.**

Région analysée

Depuis quelques années (Trabuchet *et al.* 1991 a et b, Bouhass, Perrin et Trabuchet 1994, Perrin *et al.* 1998, Fullerton *et al.* 1994, 2000), l'intérêt s'est porté sur le polymorphisme de séquence d'une région intergénique située en amont du gène de α -globine, d'environ 800 pb et comprenant 9 sites ponctuels polymorphes et un microsatellite composite (AT)_xTy (Cf. Figure suivante). **Nous avons mis au point différentes techniques pour l'analyse du polymorphisme et le cas échéant l'identification de la mutation β -thalassémique** : ARMS-PCR (PCR allèle-spécifique), qui permet chez un sujet hétérozygote d'établir indépendamment la structure de chacun de ses deux chromosomes (haplotype), double-ARMS et séquençage direct sur fragments PCR bicaténaires. Pour cela nous disposons d'une batterie d'amorces robustes (plus d'une dizaine) dont les principales sont mentionnées dans la figure 1.

Parmi les très nombreuses mutations responsables de thalassémies nous nous focaliserons essentiellement sur la mutation IVS-I-110. Elle est en effet, particulièrement intéressante compte tenu de sa fréquence plus élevée en méditerranée orientale et qui va décroissant vers les pays riverains de la Méditerranée occidentale (figure 2) :

RESULTATS

Cette mutation apparaît dans 7 contextes haplotypiques différents mais seul l'un de ces haplotypes lui est très fortement associé. C'est en Turquie que la diversité haplotypique est de loin la plus grande (figure 3). Cette diversité haplotypique pourrait alors s'expliquer par une occurrence multiple de la mutation sur des chromosomes différents. Cette hypothèse est cependant peu probable. Le site est, en effet, conservé dans les gènes de globine de différents primates (chimpanzé, gorille, bonobo) ce qui exclut la possibilité d'un point chaud de mutation. D'autre part, cette mutation est trouvée quasi exclusivement dans le bassin méditerranéen et non en Asie. L'origine unicentrique semble donc la plus plausible et la diversité observée serait engendrée par des événements de recombinaisons ou de conversions géniques aussi bien dans la partie 5' que dans la partie 3' du gène de α -globine. Pour que la mutation soit transférée d'un haplotype à un

autre, il faut que le chromosome porteur de la mutation IVS-I-110 recombine avec un chromosome de structure différente présent dans la population normale. Or, les données que nous avons obtenues sur les chromosomes ^A normaux sont cohérentes. **La mutation serait alors probablement apparue après la séparation des lignées caucasienne et indienne, il y a environ 10 à 30 000 ans B.P. (Cavalli-Sforza, Menozzi, Piazza, 1994) dans le corridor levantin ou le plateau anatolien. Sa diffusion serait intervenue dès les temps préhistoriques puis lors des événements de migrations historiques de façon plus récentes. La présence de cette mutation dans certaines régions peut être en effet la résultante d'événements de colonisation ou d'occupation.**

OBJECTIFS SCIENTIFIQUES

- **Le polymorphisme et les mutations du gène β -globine au niveau des ADN modernes et anciens**

Notre très bonne connaissance du contexte génétique des -thalassémies dans les populations méditerranéennes actuelles nous permet de disposer d'un corpus de données tout à fait exceptionnel puisqu'à ce jour nous disposons de l'haplotype de plus de 500 chromosomes ^A, -thalassémiques ou drépanocytaires. Dans ce contexte, l'exploitation et l'enrichissement de nos données préliminaires sur les ADN modernes puis le passage sur les ADN anciens s'avère essentiel. Cette approche s'inscrit directement dans cette nouvelle discipline qu'est la **paléopathologie moléculaire**. Pour la première fois, et ceci constitue une approche tout à fait originale, nous pourrions tester des scénarios construits sur les données actuelles, par des analyses au niveau des ADN extraits à partir d'ossements anciens. Afin de situer précisément l'origine géographique et chronologique de cette mutation, **des échantillons seront prélevés sur les restes humains d'une série de sites épipaléolithiques et néolithiques de l'Anatolie et du Levant nord**. Ceci sera facilité par le fait que des caractéristiques particulières du squelette accompagnent souvent les maladies de type thalassémie. Des modifications ostéologiques sévères sont, en effet, fréquemment trouvées sur les crânes ainsi que les os longs ou les dents des sujets malades (Ortner et Putschar 1985; Hershkovitz et al. 1997). Des modifications osseuses pouvant être en rapport avec des maladies de type thalassémie (hyperostose poreuse sévère notamment) ont été signalées dans de nombreux sites préhistoriques de la région considérée (Özbek 1989, 1998, 1999, inter alia). Ces seules constatations macroscopiques ne permettent cependant pas de poser un diagnostic sûr, car l'hyperostose poreuse semble pouvoir être liée aussi bien aux anémies ferriprives qu'aux anémies génétiques (Stuart Macadam 1987, 1991). Ce seul critère n'est donc pas suffisant pour diagnostiquer avec certitude une thalassémie mais il peut constituer un bon indicateur. La plupart des analyses sur l'ADN ont été menées jusqu'à présent sur l'ADN mitochondrial (Pääbo et al. 1988, Krings et al. 1997, Hänni et al. 1994 a,b) mais des travaux récents ont démontré la faisabilité sur l'ADN nucléaire (Greenwood et al. 1999) et notamment dans la cadre des pathologies bactériennes (Drancourt et al. 1998) ou génétiques (Filon et al. 1995). Nos amorces ont d'ailleurs été testées en vraie grandeur sur des échantillons d'ADN ancien vieux de 8000 à 12000 ans originaires du Maroc (Béraud-Colomb et al. 1995) et permettent d'analyser des fragments d'ADN courts (100 à 300 pb, tailles raisonnables sur des ADN dégradés), et chevauchants, ce qui permettra l'analyse d'une région assez informative.

- **Les sites archéologiques et l'échantillonnage**

Les fouilles effectuées dans la région Anatolie-Levant Nord, qui est fondamentale pour la compréhension de l'origine et des développements de la sédentarisation dans le bassin méditerranéen, ont amené la mise au jour de très nombreux restes humains néolithiques (environ 600 individus ont, par exemple, été découverts à Çayönü) et de quelques ossements épipaléolithiques. Les études anthropologiques sont cependant rares et une grande partie du matériel est restée inédite ou n'a fait l'objet que de brefs rapports préliminaires. Les populations épipaléolithiques et néolithiques de cette région sont très mal connues, tant du point de vue des pratiques funéraires que de celui de la morphologie, de la démographie ou de la pathologie. La sélection des sites s'est effectuée à l'issue d'un inventaire préalable des gisements ayant livré des restes humains. Les restes humains épipaléolithiques sont très rares. Les gisements de Sarkli Magara (Turquie) et Saaidé (Liban), où quelques ossements ont été trouvés, ont été retenus. Pour le Néolithique précéramique, les sites choisis se répartissent sur l'ensemble de la région (figure 4). La répartition chronologique et géographique des sites choisis permettra de cerner l'origine de la mutation et de suivre sa diffusion dans l'espace et dans le temps, ce qui devrait conduire à la mise

en évidence de mouvements de populations dans la région considérée. Cette approche paléogénétique devrait donc apporter des données permettant de confirmer ou d'infirmer les différentes hypothèses archéologiques concernant la néolithisation de l'Anatolie et sa diffusion vers les Balkans et l'ensemble du bassin méditerranéen et de tester l'hypothèse de l'origine de la mutation IVS-I-110. Les processus qui ont conduit à la néolithisation de l'ensemble de l'Anatolie n'apparaissent pas encore clairement mais ils pourraient être également liés à des mouvements de populations vers l'ouest, mouvements qui aboutiraient à l'introduction de l'agriculture et de l'élevage au cours du Néolithique céramique en Grèce et dans les Balkans .

Ces recherches contribueront également à une meilleure connaissance des populations néolithiques d'Anatolie et du Levant Nord, notamment de leur état sanitaire.

• **Données linguistiques :**

Les données disponibles concernant le Proche-orient ont l'avantage d'être abondantes et serrées donc plus faciles à connecter. Ainsi, il sera plus aisé d'élaborer des scénarios descriptifs de la préhistoire linguistique du Proche-orient, reflétant les processus de formation et de différenciation des communautés linguistiques Afro-asiatiques. En parallèle, les données archéologiques proche-orientales nombreuses, relatives à la période 12000-2000 av. J.C. (Naccache A. 1996, 1999) permettent de reconstituer une trajectoire sociale presque complète reliant les sociétés de chasseurs-cueilleurs du Paléolithique supérieur aux sociétés complexes de l'Age du bronze (urbaines, rurales, pastorales). **La convergence qui se fait jour entre les domaines archéologique, linguistique et génétique s'avère donc extrêmement prometteur quant à la richesse des informations que nous pourrions tirer sur les premières sociétés de chasseurs-cueilleurs du Paléolithique supérieur, puis des populations sédentarisées au pré-néolithique et au néolithique.**

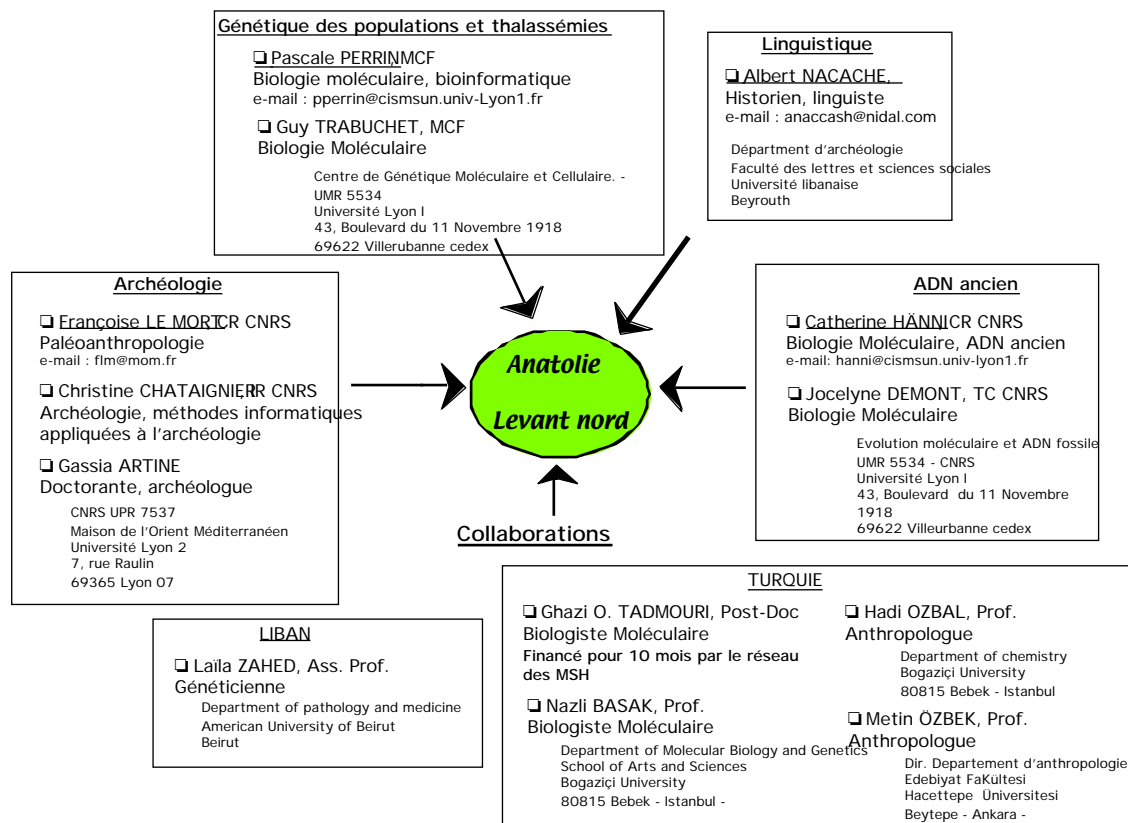
4. Création d'une base de données

Enfin, les données archéo-anthropologiques rassemblées à l'occasion de la sélection des sites et du prélèvement des échantillons serviront, tout comme les données génétiques, à la construction d'une banque de données sur les populations actuelles et anciennes de la région Anatolie - Levant Nord. Aux données archéologiques et génétiques, il sera particulièrement intéressant de confronter les données linguistiques connues dans la région du Proche-Orient.

Bibliographie :

- * Angel (1966) Science 153 : 760-763.
- * Aurenche O., Kozłowski S.K. (1999) La naissance du néolithique au proche-Orient. Paris Errance.
- * Béraud-Colomb et al. (1995) Am. J. Hum. Gen. 57: 1267-1274.
- * R. Bouhass, Perrin P. and Trabuchet G. (1994) Hemoglobin 18: 211-219.
- * Cauvin J. (1997) Naissance des divinités, naissance de l'agriculture. Paris CNRS Eds.
- * Cavalli-Sforza L.L., Menozzi P., Piazza A. (1996) The History and Geography of Human Genes. Princeton University press, New York USA.
- * Chehab F.F., Der Kaloustian V., Khouri F.P., Deeb S.S., and Kan Y.W. (1987). Blood. 69: 1141-1145.
- * Drancourt et al. (1998) PNAS 95: 12637-12640.
- * Filon et al. (1995) Nature Genetics 9,365-368.
- * Fullerton S.M., Harding R.M., Boyce A.J., Clegg J.B. (1994) Proc. Natl. Acad. Sci. USA 91 : 1805-1809.
- * Fullerton S.M., Bond, J., Schneider J.A., Hamilton B., Harding R.M., Boyce A.J., Clegg J.B. (2000) Mol. Biol. Evol. 17 : 179-188.
- * Greenwood A.D. et al. (1999) Mol. Biol. Evol. 16: 1466-1473
- * Hänni et al. (1994)a Genomics 22/ 487-489.
- * Hänni et al. (1994)b PNAS 91: 12336-12340.
- * Hershkovitz et al. (1997) Am. J. Phys. Anthr. 104 : 213-226.
- * Krings et al. (1998) Cell 90: 19-30.
- * Maggio A., Giambona A., Cai S.P., Wall J., Kan Y.W., and Chehab F.F. (1993). Blood. 81: 239-242
- * Naccache A. (1996) Proceed. Of the Seminar for Arabian studies 26 .
- * Naccache A. (1999) History and Theory 38 : 4.
- * Ortner and Putschar (1985) *Identification of Pathological Conditions in Human Skeletal Remains*. Smithsonian Contributions to Anthropology 28, Washington and London, Smithsonian Institution Press.
- * Özbek (1989) V. Arkeometri Sonuçları Toplantısı : 161-172.

- * Özbek (1998) In Arsebük, Mellink and Schirmer (eds) : *Light on Top of the Black Hill. Studies presented to Halet Çambel*, Istanbul, Ege Yayınları : 567-579.
- * Özbek M. (1999) Aeikli (Aksaray) Neolitik Caerkeinde Demir Eksikliinden Kaynaklanan Rahatsızlık. *Belleten C, LXII, 26* : 391-402.
- * Pääbo et al. (1988) *Nucl. Acids Res.* 16:9775-9787.
- * P. Perrin, Bouhass R., Mselli L., Garguier N., Nigon M., Bennani C., Labie D. and G. Trabuchet (1998) *Gene* 213 : 169-177.
- * Renfrew C. (1993) *The Roots of ethnicity, Archaeology, Genetics and the Origins of Europe*. Roma : Unione internazionale degli Istituti di Archeologia Storia e Storia dell'Arte in Roma.
- * Rosatelli M.C., Dozy A., Faà V., Meloni A., Sordu R., Saba L., Kan Y.W., and Cao A. (1992) *Am. J. Hum. Genet.* 50: 422-426.
- * Stuart-Macadam (1987) *Am. J. Phys. Anthr.* 74 : 521-526.
- * Stuart-Macadam (1991) In Ortner and Aufderheide (eds) : *Human Paleopathology. Current Syntheses and Future Options*, Washington and London, Smithsonian Institution Press : 36-39.
- * Tadmouri GO, Tüzmen S, Özçelik H, Ozer A, Baig SM, Senga EB, and Basak AN (1998) *American Journal of Hematology* 57 : 215-220.
- * Tadmouri GO, Garguier N, Demont J, Perrin P, and Basak AN, soumis.
- * Trabuchet G., Elion J., Baudot G., Pagnier J., Bouhass R., Nigon V.M., Labie D., Krishnamoorthy R. (1991 a). *Hum. Biol.* 63: 241-252.
- * Trabuchet G., Elion J., Dunda O., Lapoumeroulie C., Ducrocq R., Nadifi S., Zohoun I., Chaventre A., Carnevale P., Nagel R., Krishnamoorthy R., Labie D. (1991 b). *Human Genetics.* 87: 597-601.
- * Weatherall D.J. and Clegg J.B. (1996) *Nature Medicine.* 2: 847-849.



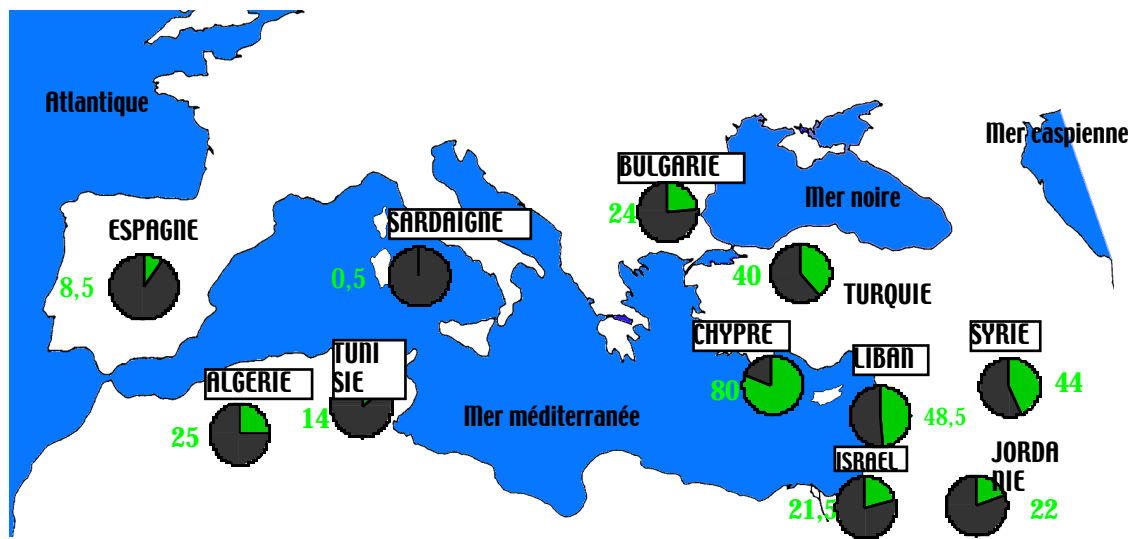


Figure 2 : Fréquence de la mutation IVS-I-110 dans le bassin méditerranéen.

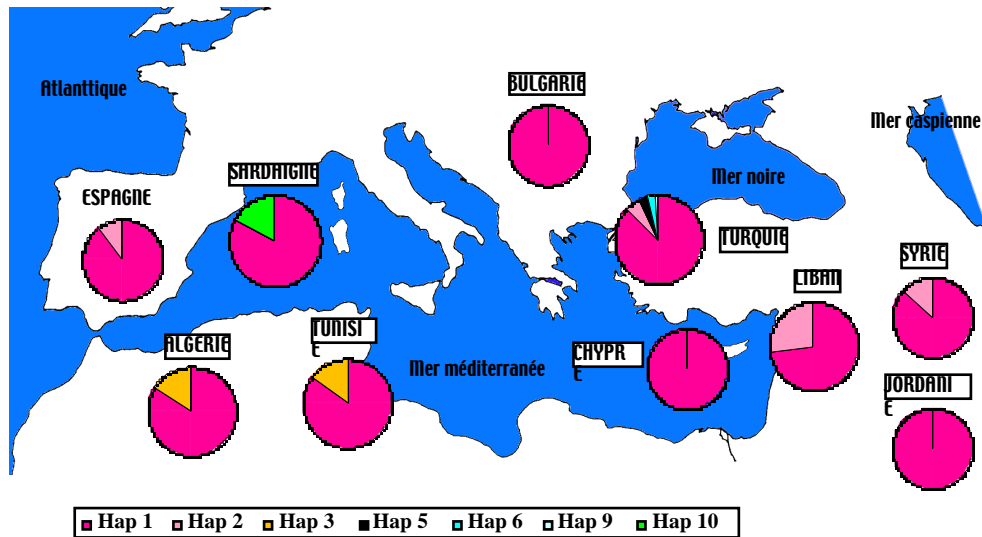


Figure 3 : Fréquence des haplotypes de séquences associés à la mutation IVS-I-110



Figure 4 : Sites du Néolithique précéramique.

