

Los NumtS como elementos móviles: revisión del campo emergente de la numtogenénesis desde la perspectiva del genoma humano

NUMTs as mobile elements: a revision of the emergent field of numtogenesis from the perspective of the human genome

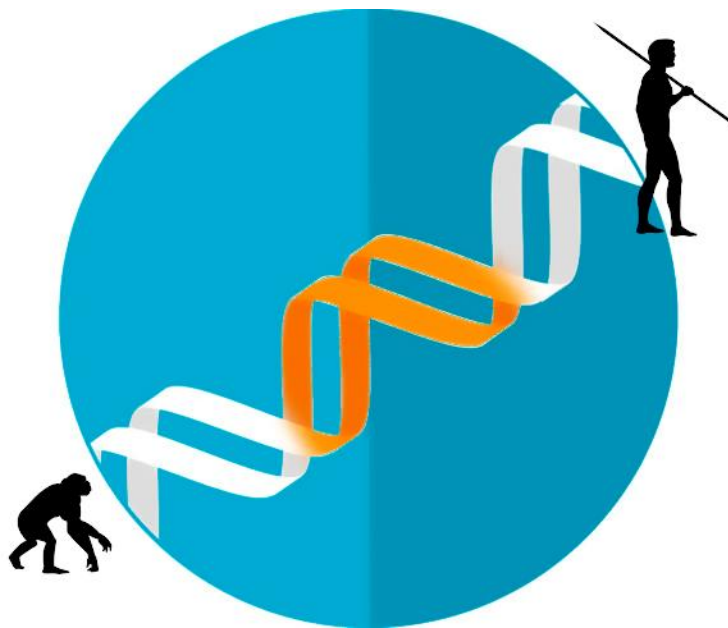
Os NUMTs como elementos móviles: revisión do campo emerxente da numtoxénesis desde a perspectiva do xenoma humano

Autor: David Blanco Remesar

Tutor: Horacio Naveira Fachal

Trabajo de Fin de Máster

A Coruña, Julio 2018



ÍNDICE	Pág.
RESUMEN.....	3
INTRODUCCIÓN.....	3
MATERIALES Y MÉTODOS.....	6
1.- Obtención de genomas.....	6
2.- Navegador genómico de la UCSC.....	6
3.- BLAT.....	7
4.- Ensembl	7
5.- Bioedit	7
6.- MEGA.....	8
RESULTADOS.....	8
1.- Detección de NumtS en todo el genoma humano.....	8
2.- Análisis del Numt en el cromosoma 7 humano.....	8
2.1.- Análisis del Numt en <i>R.roxellana</i>	10
2.2.- Análisis del Numt en <i>P.abelli</i>	11
3.- Análisis del Numt en el cromosoma 4 humano.....	11
4.- Análisis del Numt en el cromosoma 13 humano	12
4.1.- Análisis del Numt en <i>R.roxellana</i>	14
4.2.- Análisis del Numt en <i>N.larvatus</i>	14
5.- Estudio de SNPs presentes en NumtS	15
6.- Estudio filogenético de los NumtS del cromosoma 7 y cromosoma 13 con Bioedit/MEGA.....	16
DISCUSIÓN.....	16
ANEXO.....	19
REFERENCIAS.....	29

RESUMEN

Los NumtS (del inglés, Nuclear Mitochondrial DNA Segments) son fragmentos de ADN mitocondrial que se han transferido al genoma nuclear donde se insertan a través de recombinación no homologa o unión de extremos no homólogos (NHEJ, en inglés), ejerciendo un rol de protección en los cromosomas de delecciones nocivas durante la rotura de doble cadena (en inglés, DSB). No obstante, ciertos estudios sugieren que este proceso ocurre con más frecuencia durante el envejecimiento y puede ser el responsable de ciertas enfermedades, sobre todo las asociadas al cáncer. Los NumtS carecen de actividad transponible, y son considerados marcadores filogenéticos estables que permiten el estudio de la dinámica evolutiva de genomas entre especies. Mediante el uso del navegador genómico de la UCSC y la base de datos de Ensembl, así como el uso de herramientas y programas a mayores para el análisis comparativo de secuencias, en nuestro estudio hemos seleccionado varios NumtS presentes en el ser humano con el fin de dilucidar su origen y evolución en el genoma de los Primates.

INTRODUCCIÓN

Actualmente se sabe que las mitocondrias comparten un ancestro común con α -proteobacteria (Yang, Oyaizu, & Oyaizu, 1985). Estos orgánulos eucariotas presentan su propio ADN (mtADN), circular y de doble cadena, localizado en la matriz mitocondrial. El mtADN humano consta de 16569 pb y contiene 37 genes que codifican para 2 ARNs ribosómicos (rARN), 22 ARNs transferentes (tARN) y 13 polipéptidos en forma de mARN (Taanman, 1999) (Figura 1) También presenta una región de control no codificante (Wallace, 2010) conocida como D-loop, que contiene los promotores para la transcripción de las cadenas pesada (HSP, del inglés, High Strand Promoter) y ligera (LSP, del inglés, Light Strand Promoter) Esta región, además de ser el origen de replicación de la cadena pesada (Park & Larsson, 2011), está regulada por proteínas codificadas en el núcleo que se importan después de la traducción a las mitocondrias (Mercer *et al.*, 2012)

No obstante, el genoma mitocondrial es tan solo el remanente que ha quedado de un proceso evolutivo de transferencia horizontal continuada de genes desde la mitocondria al núcleo, que llevó a la reducción del tamaño de su genoma ancestral (Calabrese *et al.*, 2017; Hazkani-covo, 2007). Los fragmentos de genes parciales o completos de origen mitocondrial presentes en el ADN genómico son denominados NumtS (del inglés, Nuclear Mitochondrial Sequences), y son una gran herramienta para conocer la historia evolutiva de un taxón al carecer de actividad

transponible (Hazkani-covo, 2007) La presencia de un Numt en un locus específico puede ser usada como marcador genético, pues si está en más de un taxón indica que poseen un ancestro común. La presencia o ausencia de NumtS en loci específicos servirían para determinar el orden de ramificación filogenética de diferentes especies (Arctander, 1995)

Según McArthur *et al.* el proceso de migración de fragmentos del mtADN al ADN nuclear se produce durante la apoptosis. Las proteínas BAX y BAK, responsables de la formación de los poros en membrana mitocondrial exterior después de la pérdida de citocromo C, abren una vía de acceso de contacto del mtADN con el citosol, y su posterior migración al núcleo (McArthur *et al.*, 2018)

El mecanismo habitual de inserción de los NumtS parece ocurrir a través de recombinación no homologa o unión de extremos no homólogos (NHEJ, en inglés). Así los fragmentos de mtADN se unen con los extremos del ADN genómico durante la reparación de la rotura de doble cadena (DSB, en inglés) (Hazkani-covo, 2007) En contraste con la reparación por recombinación homóloga, el NHEJ clásico no busca secuencias verdaderamente homólogas como plantillas para reparación, sino que típicamente usa tramos cortos de identidad de secuencia (1-4 pb), o microhomología, para facilitar la unión final (Jensen-Seaman *et al.*, 2009). La inserción de NumtS también puede ocurrir sin microhomología, un proceso conocido como reparación de extremos romos. No obstante, NHEJ es intrínsecamente propenso a errores, y comúnmente implica pequeñas inserciones y deficiencias en el sitio de reparación (Hazkani-covo, Zeller, & Martin, 2010). Curiosamente, a lo largo de la historia evolutiva de los primates se conoce que más de la mitad de los eventos DSB que involucra NumtS muestran sólo pequeñas deficiencias o ninguna en absoluto. Esto es un indicativo de que los NumtS proporcionan a la maquinaria de unión final una herramienta para sellar los cortes sin la necesidad de procesar el ADN nuclear usando una nucleasa de manera adicional (Hazkani-covo & Covo, 2008; Hazkani-covo *et al.*, 2010), ejerciendo un rol de protección en los cromosomas de delecciones nocivas durante DSB. Aun así, existen estudios que sugieren que el proceso de numtogénesis (inserción de un Numt en el genoma nuclear) ocurre con más frecuencia durante el envejecimiento y puede ser el responsable de ciertas enfermedades, sobre todo las asociadas al cáncer (Ivessa, 2018).

La localización de los NumtS en seres humanos no es del todo fortuita sino que son varios los artículos que las relacionan con diversas características genómicas (Dayama, Emery, Kidd, & Mills, 2014). Existen evidencias de que los NumtS preferiblemente se insertan en regiones abiertas de cromatina, típicamente cerca de secuencias de oligómeros A + T (Dayama *et al.*,

2014; Tsuji *et al.*, 2012). Además aparentemente prefieren regiones abiertas de cromatina de gran curvatura y flexibilidad (Tsuji *et al.*, 2012). Tsuji *et al.* añaden que los NumtS tienden a insertarse cerca de retrotransposones, sin preferencia por la orientación del retrotransposón, y nunca tienden a hacerlo dentro de los mismos. Además, las endonucleasas codificadas por retrotransposones tales como LINE1 reconocen regiones ricas en A + T; y ha sido demostrada su capacidad de formación de numerosos DSB en células HeLa humanas, MCF7 y HIH 3T3 de ratón (Gasior *et al.*, 2006). Es por ello que se especula sobre la acción de la endonucleasa de L1 como mediadora de diversos eventos de inserción de NumtS en primates (Tsuji *et al.*, 2012).

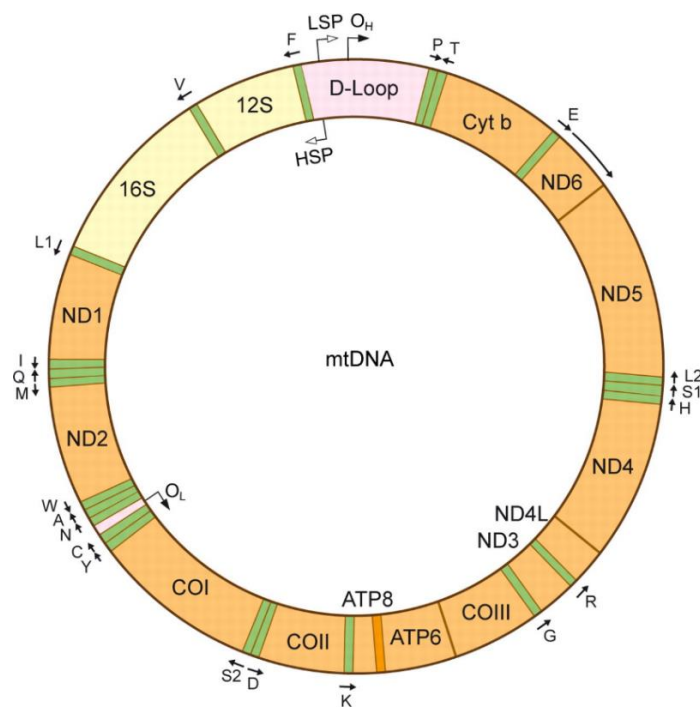


Figura 1. Representación esquemática de mtADN en humanos. La molécula de doble cadena de mtADN de ~16,5 Kb presenta una región no codificante denominada D-loop, que contiene los promotores para la transcripción de ambas cadenas (HSP y LSP) El origen de replicación de la cadena pesada (OH) se encuentra en esta región, mientras que el origen de replicación de la cadena ligera (OL) se encuentra en un clúster de genes tARN, señalizado en esta figura. Están representados en esta figura, a parte del D-loop, los genes de los 2 rARNs (12S y 16S), los 13 polipéptidos (ND1–6, ND4L, Cyt b, COI–III, ATP6, ATP8) y los 22 tARNs. El sentido de las flechitas al lado de cada una de las letras que representa a cada gen tARN de aminoácidos indica su localización, siendo el sentido horario en la cadena ligera, y el contrahorario en la cadena pesada, donde se sitúan, además, el resto de genes que no presentan notación. Imagen extraída de (Park & Larsson, 2011)

Para arrojar luz sobre el origen y la evolución de NumtS en primates hemos hecho un análisis comparativo de genomas recopilados de bases de datos como el navegador genómico UCSC

(<https://genome.ucsc.edu/>) o Ensembl (<http://www.ensembl.org/index.html>), centrándonos en varios NumtS de la vasta cantidad de ellos que existen en nuestro genoma, y que seguirán apareciendo a lo largo de la evolución de la especie humana (Dayama et al., 2014; Ricchetti, Tekaiia, & Dujon, 2004).

MATERIALES Y MÉTODOS

1.- Obtención de genomas

Se obtuvo el genoma de humano (hg38) de la página del navegador genómico de la Universidad de California Santa Cruz (<https://genome.ucsc.edu/>) Adicionalmente, descargamos los genomas mitocondriales de humano (hg38), langur chato dorado (JQ821835.1), orangután (NC_002083.1) y mono narigudo (KM889667.1) desde la base de datos genómico del NCBI (<https://www.ncbi.nlm.nih.gov/>)

2.- Navegador genómico UCSC

El navegador genómico UCSC fue creado en la etapa inicial del Proyecto del Genoma Humano y proporciona alineamientos de secuencias precisos entre secuencias consultadas (Lee, Oh, Kang, & Han, 2014) Así, los usuarios son capaces de obtener una gran variedad de información desde el estudio de conservación del genoma, hasta la visualización de diversos polimorfismos de un solo nucleótido (SNP), pasando por la distribución en el genoma de elementos transponibles. Así mismo, permite la alineación de diversos genomas; concretamente se utilizó en este estudio fundamentalmente el genoma de 30 primates.

A la hora de analizar los alineamientos entre especies de primates en el navegador, se distinguen visualmente anotaciones en forma de línea simple ‘-’, línea doble ‘=’, o regiones representadas por ‘N’. La línea simple informa de la ausencia de bases en esa especie. Esto puede ser explicado como una inserción o eliminación específica del linaje entre los bloques alineados en las especies alineadas. La línea doble informa sobre la presencia de bases en esta región, pero no alineables. Esto puede ser debido a una distancia evolutiva excesiva entre especies o ‘indels’ independientes en la región entre los bloques alineados en ambas especies. Finalmente, las ‘N’ son un reflejo de la incertidumbre en la relación entre el ADN de ambas especies, debido a la falta de secuencia en porciones relevantes de la especie que se está alineando (Genome Browser UCSC, 2018)

3.- BLAT

BLAT es un algoritmo de alineamiento de secuencias de ADN por pares diseñado para encontrar rápidamente secuencias con gran precisión (>95%) Los resultados de búsqueda BLAT en el navegador genómico UCSC muestran diversa información ordenada de la siguiente manera: puntuación (calculada según la longitud alineada y similitud de secuencia), inicio (posición de la primera coincidencia en la consulta), final (posición de la última coincidencia en la consulta), tamaño de la consulta (el tamaño de la secuencia de entrada), identidad (similitud de secuencia), coordenadas genómicas (posiciones genómicas de la secuencia combinada) y cadena (orientación de la secuencia coincidente en el genoma) (Genome Browser UCSC, 2018) (Figura 2C)

4.- Ensembl

Ensembl (<http://www.ensembl.org>) es un proyecto de bioinformática que organiza información biológica mediante una gran base de datos genómicos de diversas especies (Birney et al., 2004; Yates et al., 2016) Ensembl también ofrece recursos de genómica comparativa entre especies, en nuestro caso entre Primates, y que se basan en anotaciones genéticas para calcular evolución de genes, así como información de ortología de fragmentos de secuencias. Permite además construir alineamientos de ADN completos de secuencia de pares de genomas y secuencias múltiples. Finalmente, Ensembl integra diversas fuentes de datos dispares incluidos bases de datos de SNPs (dbSNP), mutaciones de genes humanos (HGMD), y de relaciones de variaciones humanos con fenotipos (ClinVar), y los presentan a través de una interfaz integrada y consistente (Yates et al., 2016)

5.- BioEdit

BioEdit (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>) consiste en un editor de secuencias destinado a proporcionar funciones básicas para la edición, alineación, manipulación y análisis de proteínas y secuencias de ácidos nucleicos (Hall T., 2001) BioEdit ofrece la herramienta ClustalW, un programa diseñado para construir múltiples alineamientos de secuencias biológicas. ClustalW alinea automáticamente muchas secuencias para que se puedan ver las identidades, similitudes y diferencias, calculando la mejor coincidencia para las mismas (Higgins, Thompson, & Gibson, 1996)

6.- MEGA

MEGA (del inglés, Molecular Evolutionary Genetics Analysis) es un software que permite inferir árboles filogenéticos a partir de construcciones de alineamientos de secuencias en diversos formatos (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013) Se infirieron árboles filogenéticos a través de las secuencias mitocondriales y nucleares de los distintos primates que podían contener NumtS ortólogos. Para este fin se usó la herramienta de MEGA, construyendo árboles máximo parsimoniosos sin muestreo “bootstrap” y mediante un método de búsqueda mini-mini heurístico.

RESULTADOS

1.- Detección de NumtS en todo el genoma humano

Mediante el uso de Ensembl, y en complementación con el navegador genómico de la UCSC, fue posible la detección de una amplia muestra de NumtS encontrados en el genoma humano (Anexo 1) De todos ellos, nuestro trabajo se centró en 3 de ellos (Anexo 1C)

2.- Análisis del Numt en el cromosoma 7 humano

En la posición 145997333-145997428 de la cadena (+) del cromosoma 7 (Chr7) se encuentra un Numt con un 99% de similitud con la secuencia mitocondrial de origen. La posición que ocupa este en el genoma mitocondrial (MT) es 1613-1708, llegando a abarcar parte del gen de tARN de la valina y del gen del ARN 16S. Esta secuencia de 96 nucleótidos muestra una discrepancia a nivel de la base Chr7: 145997339, la cual presenta una aparente transición de una citosina (C) de la secuencia original mitocondrial en posición MT:1619, a una timina (T) en esta posición del genoma nuclear (Figura 2).

Para acercarnos un poco a la antigüedad de este Numt se sometió a la secuencia mitocondrial original a un alineamiento con un clúster de secuencias de 12 especies de primates en Ensembl y analizando la base discrepante de este Numt entre ellos (Figura 3). Se apreció como la base de citosina en esta posición se encuentra en puntos anteriores en la ramificación del árbol filogenético hacia el género *Homo*, a pesar de la aparición de una timina en el género *Pan* (bonobo y chimpancé). Podemos observar también una timina en el género *Macaca*, a pesar de estar alejados filogenéticamente de los géneros anteriormente citados. Estos hechos evidenciarían la existencia de mutaciones paralelas con dos posibles orígenes (1) que el Numt

Seguidamente, con el fin de observar en perspectiva la evolución de este Numt en el orden de los Primates, se realizó un alineamiento en el navegador genómico UCSC de secuencias con 30 especies próximas filogenéticamente con el ser humano (Figura 4) Se observó cómo en este alineamiento hallamos otros NumtS en *Rhinopithecus roxellana* y en *Pongo abelli*.

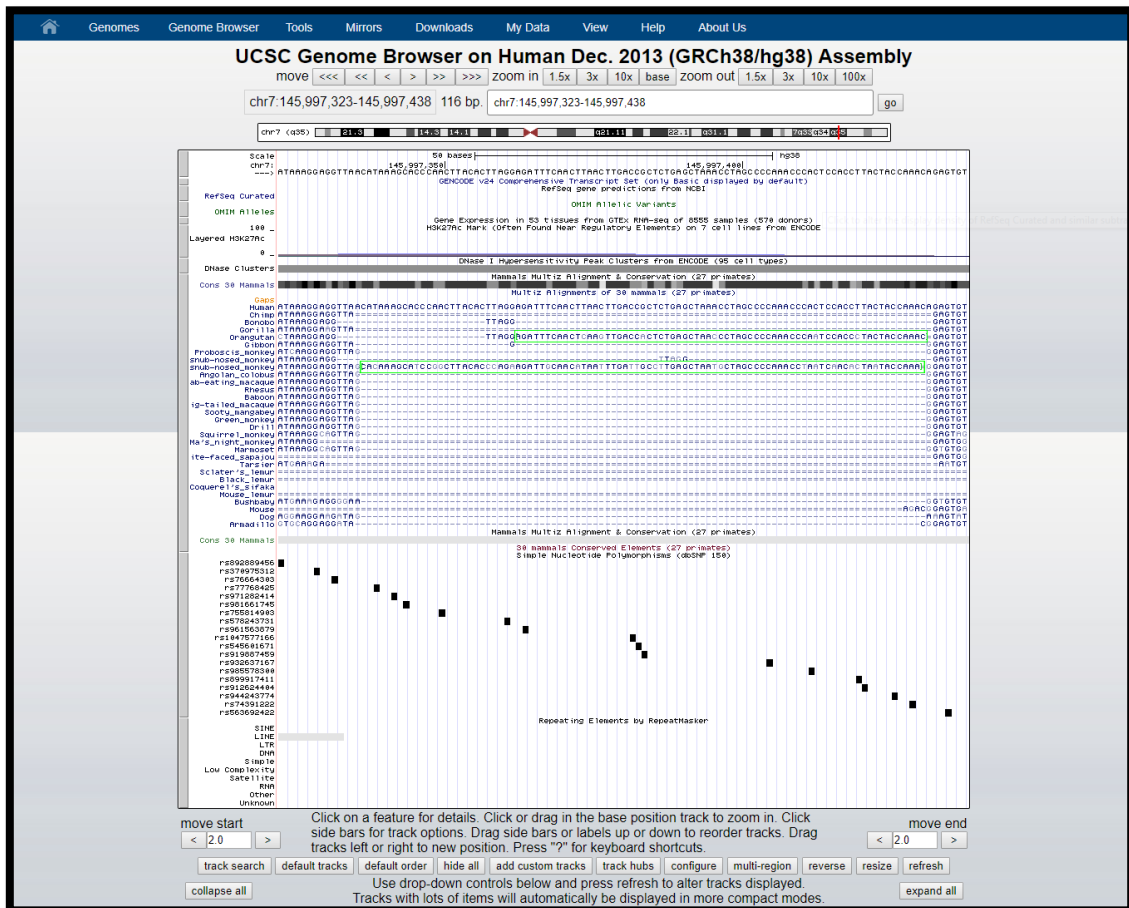


Figura 4. Alineamiento de secuencias del Numt en el cromosoma 7 (145997333-145997428) en humanos con otras secuencias en primates (parte superior). Las zonas remarcadas en verde indican la presencia de NumtS en esas dos especies: orangután (*P. abelli*) y el langur chato dorado (*R. roxellana*). En esta figura también se presenta en su parte inferior los distintos SNPs presentes en humanos en esta región del cromosoma 7

2.1.- Análisis del Numt en *R. roxellana*

Un estudio más a fondo de este Numt en esta especie usando el navegador genómico UCSC nos parte esta secuencia en dos bloques. El primer bloque de esta secuencia es CACAAAGCATCCGGCTTACACCCAGA. Se observó que la zona de origen exacta se encuentra en el cromosoma KN296453v1, en la posición 1886686-1886711 (Anexo 2A). El segundo bloque de este NUMT que divide el navegador genómico es AGATTGCAACATAATTTGATTGCCTTGAGCTAATGCTAGCCCCAAACCTAATCAACACTAATACC. Se observó que la zona de origen de esta secuencia pertenece al mismo

cromosoma que el primer bloque, en la posición 1886621-1886685 (Anexo 2B) Si se ordenan ambos bloques del cromosoma KN296453v1 se puede observar que difiere de la secuencia que aparece alineada en el navegador genómico. Esto es debido a que la secuencia que ahí se refleja es la inversa complementaria del Numt de esta especie. Seguidamente, se realizó un BLAT del genoma mitocondrial de *R.roxellana* con su genoma nuclear para así determinar su grado de similitud entre ambas secuencias, perteneciendo a una región con un 89,1% de similitud con la secuencia mitocondrial original. Se apreció también que el Numt en esta especie se insertó en la cadena complementaria del cromosoma KN296453v1 (Anexo 2C)

2.2- Análisis del Numt en *P.abelli*

Siguiendo un procedimiento similar con el Numt en *R.roxellana*, se realizó un estudio a fondo del Numt presente en *P.abelli*. En comparación con el caso anterior, esta secuencia (AGATTTCAACTCAAGTTGACCACTCTGAGCTAACCCCTAGCCCCAAACCCAATCCA CCCTACTACC) es presentada en un único bloque por el navegador, perteneciente al cromosoma 17_random de esta especie entre las posiciones 17151481-17151545 (Anexo 3A) A diferencia del Numt de *R.roxellana*, la secuencia que aparece alineada en el navegador es la misma que aparece aquí. Seguidamente, se realizó un BLAT del genoma mitocondrial de *P.abelli* con su genoma nuclear para así determinar su grado de similitud entre ambas secuencias, englobada en un hit con un 96,9% de similitud con la secuencia mitocondrial original (Anexo 3B). Se apreció también que el Numt en esta especie se insertó en la cadena principal de este cromosoma.

3.- Análisis del Numt en el cromosoma 4 humano

En la posición 55328160-55328290 de la cadena (+) del cromosoma 4 (Chr4) se encuentra un Numt con un 96.2% de similitud con la secuencia mitocondrial de origen. La posición que ocupa este en el genoma mitocondrial (MT) es 962-1092, llegando a abarcar parte del gen de ARN 12S. Esta secuencia de 131 nucleótidos, a diferencia del Numt del cromosoma 7, muestra varias discrepancias a nivel de diversas bases detalladas en la figura 5.

Para conocer la antigüedad de este Numt se sometió a la secuencia mitocondrial original a un alineamiento con un clúster de secuencias de 12 especies de primates en Ensembl y analizando la base discrepante de este Numt entre ellos (Anexo 4) Se apreció que, a excepción de la base MT:1011, el resto no presenta (MT:1085), presenta en un solo género (*Macaca*; MT: 1039) o presenta solo en humanos (MT: 1005; MT: 1018) distintas bases en la misma posición entre secuencias. Así mismo, la aparición de la base de citosina en MT:1011 parece datar de la época

la presencia de una adenina ancestral en lémures ratones que, una vez el árbol diverge en el infraorden Simiiformes aparece una timina en esa posición, la cual presenta todo este infraorden a excepción del ser humano y el cercopiteco verde, los cuales presentan una citosina. Esto puede ser debido a mutaciones paralelas convergentes en esta base en ambas especies. En cuanto al origen de este Numt en humanos provenir (1) de una mutación posterior a la inserción del Numt de la C a posteriormente una T o (2) un poco probable polimorfismo durante la especiación para ese rasgo T/C, cuando se produjo la inserción del Numt.

Para observar en perspectiva la evolución de este Numt en el orden de los Primates, se realizó un alineamiento en el navegador genómico UCSC de secuencias con 30 especies próximas filogenéticamente con el ser humano, al igual que los casos anteriores (Figura 7) No se encontró ningún alineamiento con los primates más próximos filogenéticamente, aunque sí se encontró una similitud de alineamiento con tres especies correspondientes a la subfamilia Colobinae (*R. roxellana*, *R. bieti* y *Nasalis larvatus*), las cuales presentan NumtS.

4.1.- Análisis del Numt en *R. roxellana*

Un estudio más a fondo de este Numt en esta especie usando el navegador genómico UCSC nos parte esta relativamente larga secuencia en varios bloques. Los bloques aparecen detalladamente en la Anexo 7. Toda esta secuencia alineada corresponde al cromosoma KN296453v1 de *R. roxellana*. La secuencia que aparece alineada en el navegador es la misma que aparece aquí. Seguidamente, se realizó un BLAT del genoma mitocondrial de *R. roxellana* con su genoma nuclear para así determinar su grado de similitud entre ambas secuencias, perteneciendo a una región con un 89,1% de similitud con la secuencia mitocondrial original. Nótese que este hit es el mismo obtenido al alinear secuencias en el caso del Numt del cromosoma 7 humano (Anexo 2C).

4.2.- Análisis del Numt en *N. larvatus*

Un estudio más a fondo de este Numt en esta especie usando el navegador genómico UCSC nos parte esta secuencia en los mismos bloques que en el caso de *R. roxellana*, llegando a tener secuencias muy similares. Los bloques aparecen detalladamente en la Anexo 8A. Toda esta secuencia alineada corresponde al cromosoma 19 de *N. larvatus*. La secuencia que aparece alineada en el navegador es la misma que aparece aquí. Seguidamente, se realizó un BLAT del genoma mitocondrial de *N. larvatus* con su genoma nuclear para así determinar su grado de similitud entre ambas secuencias, perteneciendo a una región con un 85,7% de similitud con la

secuencia mitocondrial original. Se apreció también que el Numt en esta especie se insertó en la cadena complementaria del cromosoma 19 (Anexo 8B).

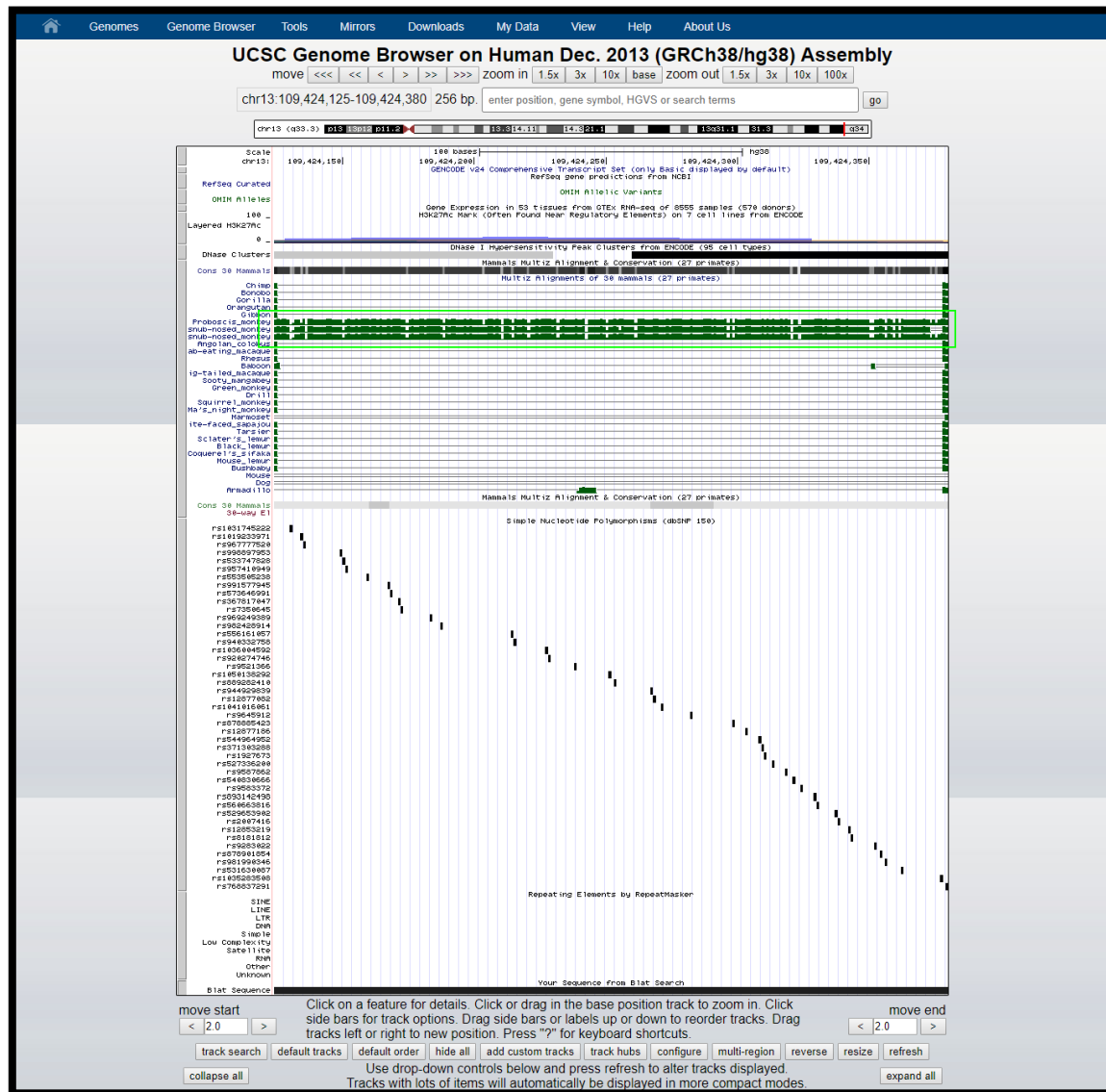


Figura 7. Alineamiento de secuencias del Numt en el cromosoma 13 en humanos con otras secuencias similares en primates (parte superior). Se puede observar cómo no se encuentra ningún alineamiento con nuestros parientes más próximos, aunque si se encuentra con tres especies de la subfamilia Colobinae (*R. roxellana*, *R. bieti* y *N. larvatus*). En esta figura también se presenta en su parte inferior los distintos SNPs presentes en humanos en esta región del cromosoma 13

5.- Estudio de SNPs presentes en NumtS

En cada uno de los NumtS de este trabajo se estudiaron varios SNPs (Anexo 9A) presentes en sus secuencias con el afán de arrojar luz sobre su posible función y si poseen algún significado clínico.

En general se encontraron que las variaciones de estos cuatro SNPs estudiados han sido reclutados en regiones regulatorias de algunas líneas celulares, llegando a funcionar como potenciadores o represores transcripcionales, tal y como apunta Richetti et al, sobre la integración de NumtS en el genoma, que se dirigen preferentemente a secuencias codificadoras o reguladoras (Ricchetti *et al.*, 2004) Además, no se constató ningún significado clínico de las variantes de SNPs. Es de especial interés, no obstante, destacar la frecuencia presente en el SNP rs8181812, debido a su extrema poca inusualidad, con un MAF (del inglés, Minor Allele Frequency) menor a 0,01% en la población mundial, siendo casos muy concretos los que presentan esta variación (Anexo 9B).

6.- Estudio filogenético de los NumtS del cromosoma 7 y cromosoma 13 con Bioedit/MEGA

Con el fin de estudiar estos NumtS en la historia evolutiva de los Primates, así como aproximar en el tiempo el momento de la inserción de estos en el genoma nuclear del ser humano, se utilizaron las secuencias de los NumtS del cromosoma 7 y cromosoma 13, junto con las secuencias de NumtS alineadas en otras especies de Primates con cada uno de estos, así como las secuencias mitocondriales de varios Primates. En Bioedit, y gracias a ClustalW, se alinearon las secuencias genómicas y mitocondriales, para posteriormente en MEGA, hacer uso de ellas para inferir árboles filogenéticos máximo parsimoniosos. Los resultados obtenidos para el Numt del cromosoma 7 y cromosoma 13 aparecen reflejados en el Anexo 10.

DISCUSIÓN

La integración de los genes mitocondriales en el genoma nuclear es un proceso fisiológicamente importante que contribuyó al origen y la evolución de la célula eucariota, y la transferencia de genes completos desde las mitocondrias al núcleo parece estar continuamente activa en algunas plantas (Knoop *et al.*, 1995). El análisis de NumtS, con ayuda de diversas herramientas con el navegador genómico UCSC o la base de datos de Ensembl, permitió una visualización más profunda mediante la genómica comparativa sobre la evolución y divergencia de los genomas entre especies.

La ausencia de estos NumtS en nuestros antepasados más próximos nos permite concluir que su inserción se ha producido recientemente en el ser humano. Contrariamente a otros estudios que apuntan sobre que la inserción de la mayoría de los NumtS en el ser humano se remonta a

un primate ancestral (Bensasson, Feldman & Petrov, 2004; Tourmen *et al.*, 2002), nuestros resultados no hacen sino alejarnos de esta idea. A raíz de la presencia de estos NumtS en otros Primates más alejados como *R.roxellana* o *N.larvatus* tenemos más similitud con ellos que con nuestros parientes más próximos. Aunque en principio es posible que una inserción se conserve en distintas ramas de una filogenia y se haya perdido en las demás, la gran similitud de secuencias entre Numt y genoma mitocondrial humano revela cuan reciente son estas inserciones. Se ha producido, concretamente en el Numt del cromosoma 7 y del cromosoma 13, la inserción de una misma secuencia mitocondrial en el mismo lugar del genoma nuclear varias veces a lo largo de la historia evolutiva de los primates, pero por procesos aparentemente independientes, debido a que no presentan un antecesor común que presente este Numt.

Los árboles filogenéticos del Numt del cromosoma 7 y del cromosoma 13 permitieron acercarnos al momento en el que se pudo producir la inserción de estos NumtS. En el caso del Numt del cromosoma 7 su secuencia mitocondrial de origen presenta un cambio nucleotídico compartido con el genoma mitocondrial del chimpancé y del bonobo. En la representación gráfica en el árbol filogenético máximo parsimonioso de este Numt, se observa como estos genomas mitocondriales de Primates están más próximos entre sí, que el del Numt del cromosoma 7 humano con su genoma mitocondrial. Con estos datos se podría concluir que la fuente de la inserción fue un genoma mitocondrial anterior a la aparición del que hoy es específico del hombre, y datar la inserción de este Numt durante la separación del linaje del chimpancé y el del ser humano. No obstante, en el caso del Numt del cromosoma 13, su secuencia mitocondrial de origen presenta una discrepancia nucleotídica que no aparece en el resto de Primates y, junto al árbol filogenético inferido, su proximidad entre Numt y genoma mitocondrial en el ser humano, sumado a su gran similitud de secuencias, reafirma aún más su posición como Numt adquirido muy recientemente.

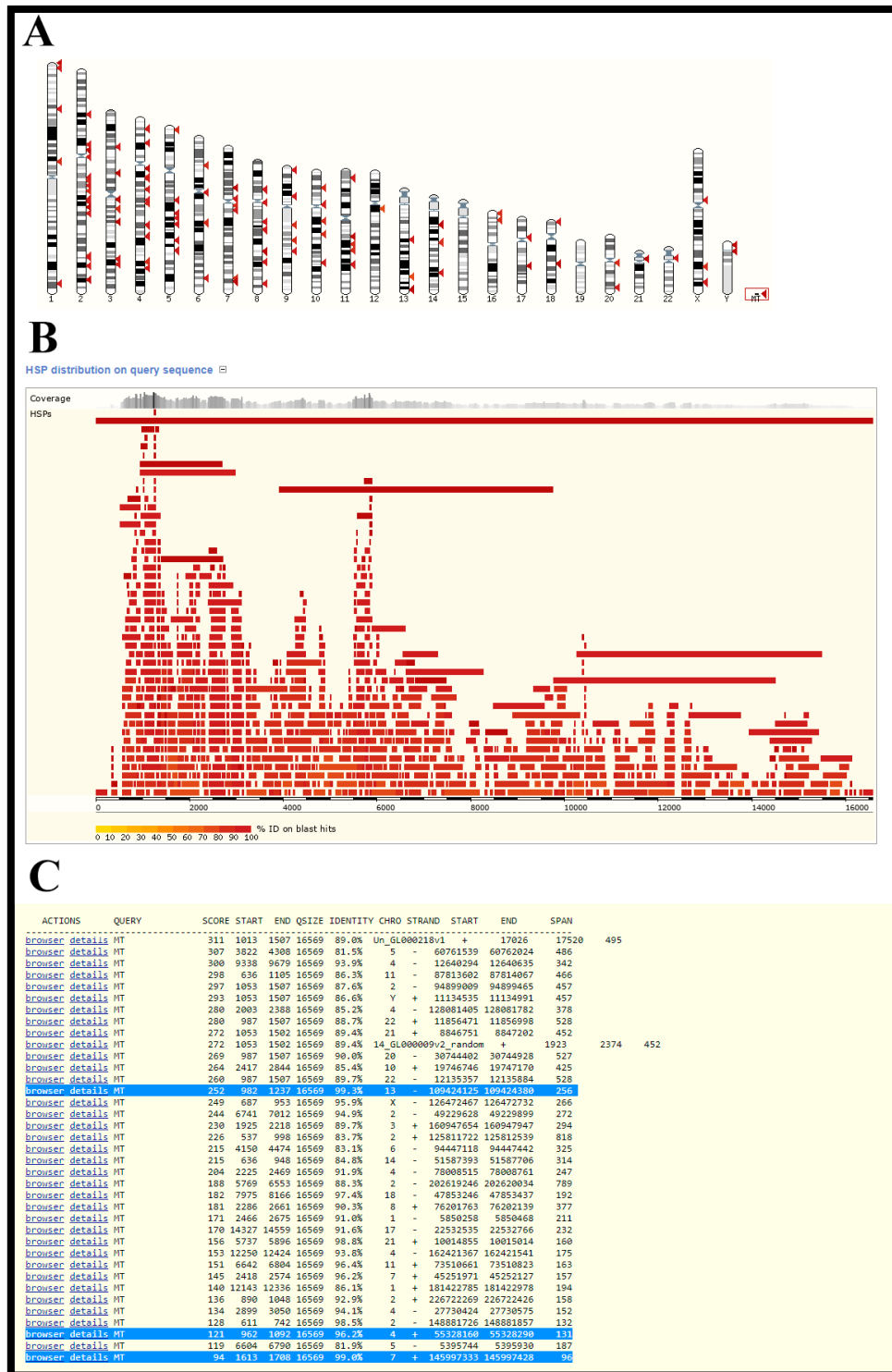
Es reseñable añadir que el Numt del cromosoma 4 y del cromosoma 13 provienen de la misma región mitocondrial. Aunque en nuestro estudio no lo hemos analizado, hay que destacar que solo un tercio de los Numt se integraron como nuevas secuencias provenientes del genoma mitocondrial, mientras que los dos tercios restantes se originaron como duplicaciones de Numt ya preexistentes en el genoma nuclear, siendo los más antiguos los que presentan mayor tasa de duplicación que los NumtS recientes (Hazkani-Covo, Sorek & Graur, 2003).

Con todo, los NumtS podrían llegar a tratarse como elementos transponibles (TE) siguiendo un “modelo de copia maestra” (Rouzic & Deceliere, 2005), pero con ligeras modificaciones, donde

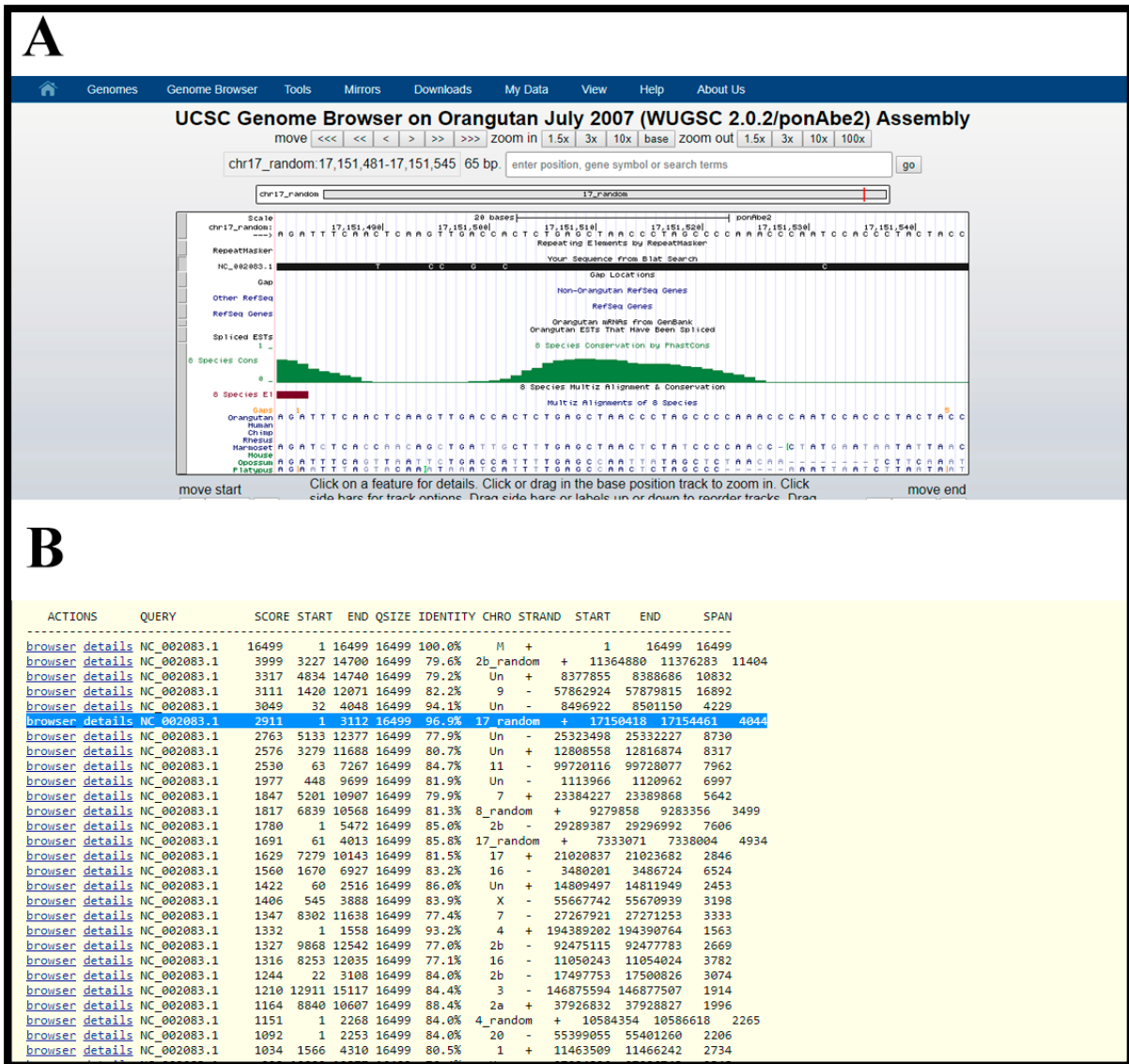
a diferencia de los TEs cuya copia residiría el genoma nuclear, esta copia en el caso de los NumtS residiría en el orgánulo mitocondrial. Esta reflexión da una explicación a la gran cantidad de hits que se encuentran en una misma región del genoma mitocondrial en el Anexo 1B.

A pesar de los resultados obtenidos en nuestro estudio que parece rebatir las conclusiones de una gran mayoría de estudios sobre la presencia de secuencias ortólogas en la evolución de los Primates (Tourmen *et al.*, 2002), es necesario que seamos conscientes de que es necesario el análisis de una cantidad considerable de NumtS para sacar unas conclusiones sólidas. No obstante, el proceso de inserción de secuencias mitocondriales en el genoma nuclear no es un proceso estático, sino un proceso continuo que requiere de un estudio constante tanto en la especie humana como en el resto de Primates y especies a lo largo de su historia evolutiva (Ricchetti *et al.*, 2004).

ANEXO



Anexo 1. Distribución en el genoma de los NumtS (A) y en la secuencia mitocondrial (x1000) (B). (A) Las flechas marcan la localización de los NumtS en cada uno de los cromosomas tras un BLAT del genoma mitocondrial humano contra hg38. Se puede observar como los cromosomas 15 y 19 carecen de secuencias de origen mitocondrial. (B) Las líneas rojas determinan el largo de la secuencia del Numt y su posición original en el genoma mitocondrial (C) NumtS seleccionados en nuestro estudio (marcados en azul) que se obtuvieron del navegador genómico UCSC. Corresponden a NumtS de secuencia relativamente corta y con >90% de similitud de secuencia presentes en los cromosomas 7, 4 y 13.



Anexo 3. Región del cromosoma 17_random de *P.abelli* donde se encuentra el Numt. (A) Se delimitó la región donde se encuentra el Numt (B) Tras realizar el BLAT entre el genoma mitocondrial y el genoma nuclear de esta especie, se obtuvieron varios hits de los cuales el encuadrado en azul pertenece a la región donde se encuentra el Numt, con un 96,9% de identidad y ubicado en la cadena principal del cromosoma 17_random.

Custom tracks

Export data

Share this page

Bookmark this page

Hidden and species

The following 1 species in the alignment are not shown - use "Configure this page" on the left to show them.

- Ancestral sequence

The following 3 species have no alignment in this region:

- Gibbon (*Nomascus leucogenys*)
- Marmoset (*Callithrix jacchus*)
- Olive baboon (*Papio anubis*)

LEGEND

Branch Length	Species	Expanded Alignments
— x1 branch length	SpeciesName high quality assembly	□ gap
--- x10 branch length	SpeciesName low quality assembly	■ aligned seq
--- x100 branch length		

Human > [chromosome:GRCh38.MT.962-1092.1](#)

Bonobo > [chromosome:panpan1.1.MT.382-512.1](#)

Chimpanzee > [chromosome:Pan_tro_3.0.MT.382-510.1](#)

Gorilla > [chromosome:gorGor4.MT.382-512.1](#)

Orangutan > [chromosome:PPYG2.MT.386-516.1](#)

Vervet-AGM > [chromosome:ChiSab1.1.MT.964-1093.1](#)

Crab-eating macaque > [chromosome:Macaca_fascicularis_5.0.MT.922-1051.1](#)

Macaque > [chromosome:Mmul_8.0.1.MT.915-1043.1](#)

Mouse Lemur > [chromosome:Mmur_3.0.MT.398-529.1](#)

Alignment

Currently showing full alignment. Please click the button below to show the alignment for first 120 columns.

[Hide full alignment](#)

```

Human      CCCC-ATAAAGCTAAAATTCACCTGAGTTGTAA---AAACTCCAGTGGAC---AAATTAAGTACTAGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGACCCAAACTGGGATT
Bonobo     CCCC-ATAAAGCTAAAATTCACCTGAGTTGTAA---AAACTCCAGTGGAC---AAATTAAGTACTAGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGACCCAAACTGGGATT
Chimpanzee --CC-ATAAAGCTAAAATTCACCTGAGTTGTAA---AAACTCCAGTGGAC---AAATTAAGTACTAGAAAGTGGCTTTAAATATCTGAAATACAAATAGCTAAGACCCAAACTGGGATT
Gorilla    CCCC-ATAAAGCTAAAATTCACCTGAGTTGTAA---AAACTCCAGTGGAC---AAATTAAGTACTAGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGACCCAAACTGGGATT
Orangutan  CCCC-ATAAAGCTAAAATTCACCTGAGTTGTAA---AAACTCCAGTGGAC---AAATTAAGTACTAGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGACCCAAACTGGGATT
Vervet-AGM GCTTAATAAAGCTAAAATTCACCTGAGTTGTAA---AAACTCCAGTGGAC---AAATTAAGTACTAGAAAGTGGCTTTAAATATCTGAAATACAAATAGCTAAGACCCAAACTGGGATT
Crab-eating macaque ACATA-ATAAAGCTAAAATTCACCTGAGTTGTAA---AAACTCCAGTGGAC---AAATTAAGTACTAGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGACCCAAACTGGGATT
Macaque    ATATA-ATAAAGCTAAAATTCACCTGAGTTGTAA---AAACTCCAGTGGAC---AAATTAAGTACTAGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGACCCAAACTGGGATT
Mouse Lemur TATAA-ATAAAGCTAAAATTCACCTGAGTTGTAA---AAACTCCAGTGGAC---AAATTAAGTACTAGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGACCCAAACTGGGATT

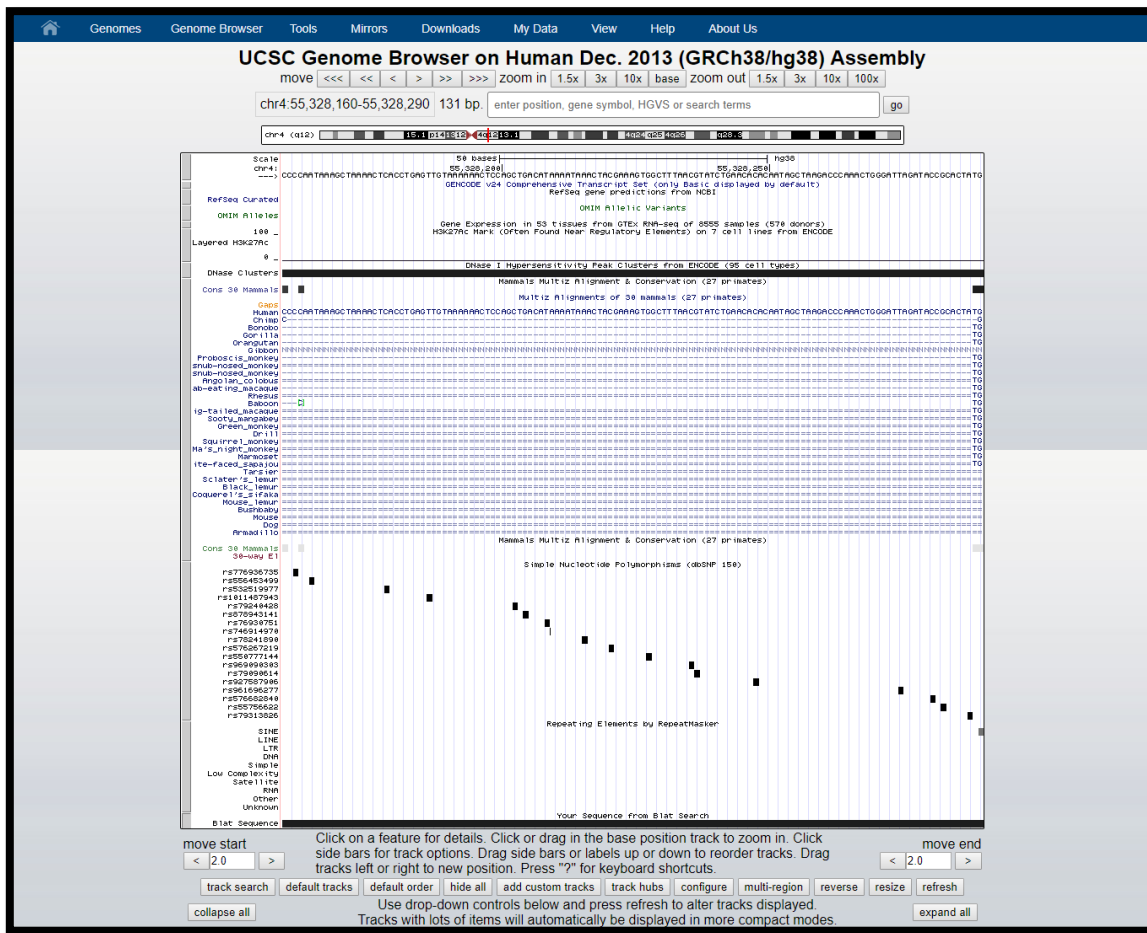
```

```

Human      AGATAC-C-CACTATG
Bonobo     AGATAC-C-CACTATG
Chimpanzee AGATAC-C-CACTATG
Gorilla    AGATAC-C-CACTATG
Orangutan  AGATAC-C-CACTATG
Vervet-AGM AGATAC-C-CACTATG
Crab-eating macaque AGACACCC-CACTATG
Macaque    AGACAC-C-CACTATG
Mouse Lemur AGATAC-C-CACTATG

```

Anexo 4. Alineamiento de la región mitocondrial humana MT: 962-1092 con otras 12 especies de primates. Se puede observar cómo no se encuentra emparejamiento en gibones, titis y papiones. Con el resto de primates emparejados sí que encuentra zonas de correlación, estableciendo un árbol filogenético en relación a la similitud de secuencias. La zona encuadrada en verde marca las bases discrepantes en este Numt.



Anexo 5. Alineamiento de secuencias del Numt en el cromosoma 4 en humanos con otras secuencias similares en primates (parte superior). Se puede observar cómo no se encuentra ningún alineamiento con nuestros parientes más próximos, aunque si con nuestros parientes más lejanos, aunque el alineamiento es muy deficiente. En esta figura también se presenta en su parte inferior los distintos SNPs presentes en humanos en esta región del cromosoma 4.

Ensembl BLAST/BLAT | BioMart | VEP | Tools | Downloads | Help & Docs | Blog

Human (GRCh38 p12) Search all species...

Location: MT:982-1,237

Location-based displays: Whole genome, Chromosome summary, Region overview, Region in detail, Comparative Genomics, Synteny, Alignments (image), **Alignments (text)**, Region Comparison, Genetic Variation, Resequencing, Linkage Data, Markers, Other genome browsers, UCSC, NCBI, Ensembl GRCh37

Configure this page | Custom tracks | Export data | Share this page | Bookmark this page

Chromosome MT: 982-1,237

Chr. MT

Alignments (text)

Alignment: 12 primates EPO Select another alignment

Location: MT:982-1237 Go Gene: Go

Hidden and species

The following 1 species in the alignment are not shown - use "Configure this page" on the left to show them.

- Ancestral sequence

The following 3 species have no alignment in this region:

- Gibbon (*Nomascus leucogenys*)
- Marmoset (*Callithrix jacchus*)
- Olive baboon (*Papio anubis*)

LEGEND

Branch Length: x1 branch length, x10 branch length, x100 branch length

Species: SpeciesName high quality assembly, SpeciesName low quality assembly

Expanded Alignments: gap, aligned seq

Human: [chromosome GRCh38:MT:982.1237.1](#)
 Bonobo: [chromosome panpan1.1.MT.402.657.1](#)
 Chimpanzee: [chromosome Pan_tro_3.0.MT.400.655.1](#)
 Gorilla: [chromosome gorGor4.MT.402.657.1](#)
 Orangutan: [chromosome PPyG2.MT.406.661.1](#)
 Vervet-AGM: [chromosome CHSsh1.1.MT.985.1236.1](#)
 Crab-eating macaque: [chromosome.Macaque.fascicularis.5.0.MT.942.1195.1](#)
 Macaque: [chromosome.Mmul_8.0.1.MT.935.1186.1](#)
 Mouse Lemur: [chromosome.Mmul_3.0.MT.418.677.1](#)

Alignment

Currently showing full alignment. Please click the button below to show the alignment for first 120 columns.

Hide full alignment

```

Human      ACCTGAGTTGTAA--AAAACCTCCAGTGGTCA-AAATTAACCTACGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGGCCAACTGGGATTAGATAC--CCACTATGCTTAG
Bonobo     ACCTGAGTTGTAA--AAAACCTCCAGTGGTCA-AAATTAACCTACGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGGCCAACTGGGATTAGATAC--CCACTATGCTTAG
Chimpanzee ACCTGAGTTGTAA--AAAACCTCCAGTGGTCA-AAATTAACCTACGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGGCCAACTGGGATTAGATAC--CCACTATGCTTAG
Gorilla     ACCTGAGTTGTAA--AAAACCTCCAGTGGTCA-AAATTAACCTACGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGGCCAACTGGGATTAGATAC--CCACTATGCTTAG
Orangutan  ACCTGAGTTGTAA--AAAACCTCCAGTGGTCA-AAATTAACCTACGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGGCCAACTGGGATTAGATAC--CCACTATGCTTAG
Vervet-AGM ACCTGAGTTGTAA--AAAACCTCCAGTGGTCA-AAATTAACCTACGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGGCCAACTGGGATTAGATAC--CCACTATGCTTAG
Crab-eating macaque ACCTGAGTTGTAA--AAAACCTCCAGTGGTCA-AAATTAACCTACGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGGCCAACTGGGATTAGATAC--CCACTATGCTTAG
Macaque    ACCTGAGTTGTAA--AAAACCTCCAGTGGTCA-AAATTAACCTACGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGGCCAACTGGGATTAGATAC--CCACTATGCTTAG
Mouse Lemur ACCTGAGTTGTAA--AAAACCTCCAGTGGTCA-AAATTAACCTACGAAAGTGGCTTTAAATATCTGAAACACAAATAGCTAAGGCCAACTGGGATTAGATAC--CCACTATGCTTAG

Human      CCTTAAACTTCAACAGTTAAA--TTACAAAACCTGCTGCCAGAACCTACGAGCCACAGCTTAAAACCTCAAAGGACCTGGCGGCTTCA--TA--TCCCTTAGAGGAGCCTGTCTT
Bonobo     CCTTAAACTTCAACAGTTAAA--TTACAAAACCTGCTGCCAGAACCTACGAGCCACAGCTTAAAACCTCAAAGGACCTGGCGGCTTCA--TA--TCCCTTAGAGGAGCCTGTCTT
Chimpanzee CCTTAAACTTCAACAGTTAAA--TTACAAAACCTGCTGCCAGAACCTACGAGCCACAGCTTAAAACCTCAAAGGACCTGGCGGCTTCA--TA--TCCCTTAGAGGAGCCTGTCTT
Gorilla     CCTTAAACTTCAACAGTTAAA--TTACAAAACCTGCTGCCAGAACCTACGAGCCACAGCTTAAAACCTCAAAGGACCTGGCGGCTTCA--TA--TCCCTTAGAGGAGCCTGTCTT
Orangutan  CCTTAAACTTCAACAGTTAAA--TTACAAAACCTGCTGCCAGAACCTACGAGCCACAGCTTAAAACCTCAAAGGACCTGGCGGCTTCA--TA--TCCCTTAGAGGAGCCTGTCTT
Vervet-AGM CCTTAAACTTCAACAGTTAAA--TTACAAAACCTGCTGCCAGAACCTACGAGCCACAGCTTAAAACCTCAAAGGACCTGGCGGCTTCA--TA--TCCCTTAGAGGAGCCTGTCTT
Crab-eating macaque CCTTAAACTTCAACAGTTAAA--TTACAAAACCTGCTGCCAGAACCTACGAGCCACAGCTTAAAACCTCAAAGGACCTGGCGGCTTCA--TA--TCCCTTAGAGGAGCCTGTCTT
Macaque    CCTTAAACTTCAACAGTTAAA--TTACAAAACCTGCTGCCAGAACCTACGAGCCACAGCTTAAAACCTCAAAGGACCTGGCGGCTTCA--TA--TCCCTTAGAGGAGCCTGTCTT
Mouse Lemur CCTTAAACTTCAACAGTTAAA--TTACAAAACCTGCTGCCAGAACCTACGAGCCACAGCTTAAAACCTCAAAGGACCTGGCGGCTTCA--TA--TCCCTTAGAGGAGCCTGTCTT

Human      GTAATCGATAAAACCCCGATCAACCTCA
Bonobo     GTAATCGATAAAACCCCGATCAACCTCA
Chimpanzee GTAATCGATAAAACCCCGATCAACCTCA
Gorilla     GTAATCGATAAAACCCCGATCAACCTCA
Orangutan  GTAATCGATAAAACCCCGATCAACCTCA
Vervet-AGM GTAATCGATAAAACCCCGATCAACCTCA
Crab-eating macaque GTAATCGATAAAACCCCGATCAACCTCA
Macaque    GTAATCGATAAAACCCCGATCAACCTCA
Mouse Lemur GTAATCGATAAAACCCCGATCAACCTCA
  
```

Ensembl release 92 - April 2018 © EMBL-EBI Permanent link - View in archive site

Anexo 6. Alineamiento de la región mitocondrial humana MT: 982-1237 con otras 12 especies de primates. Se puede observar cómo no se encuentra emparejamiento en gibones, titis y papiones. Con el resto de primates emparejados sí que encuentra zonas de correlación, estableciendo un árbol filogenético en relación a la similitud de secuencias. La zona encuadrada en verde marca las bases discrepantes en este Numt.

The image displays four sequential screenshots of the UCSC Genome Browser interface, each showing a different zoomed-in view of a genomic region on chromosome KN296453v1 of the Golden snub-nosed monkey (Oct. 2014, Rrox_v1/rhiRox1) assembly. Each screenshot includes a navigation bar at the top with options like Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, View, Help, and About Us. Below the navigation bar, the browser title and coordinates are shown, along with zoom controls (1.5x, 3x, 10x, 100x) and a search box. The main content area features several tracks: a Scale track showing the zoomed-in genomic coordinates, a RepeatMasker track showing masked repeats, a Gaps track showing gap locations, and an Other RefSeq track showing gene models. The screenshots show the browser zooming in from a 113 bp region to a 7 bp region, then to a 105 bp region, and finally to a 22 bp region.

Anexo 7. Región del cromosoma KN296453v1 de *R.roxellana* donde se encuentra el Numt (1887095-1887341). Representación de cada uno de los bloques en los que el navegador genómico UCSC divide esta secuencia de *R.roxellana* alineada con la secuencia del Numt del cromosoma 13 en el ser humano.

A

UCSC Genome Browser on Proboscis monkey Nov. 2014 (Charlie1.0/nasLar1) Assembly

chr19 18,531,381-18,531,493 113 bp

chr19 18,531,494-18,531,500 7 bp

chr19 18,531,501-18,531,605 105 bp

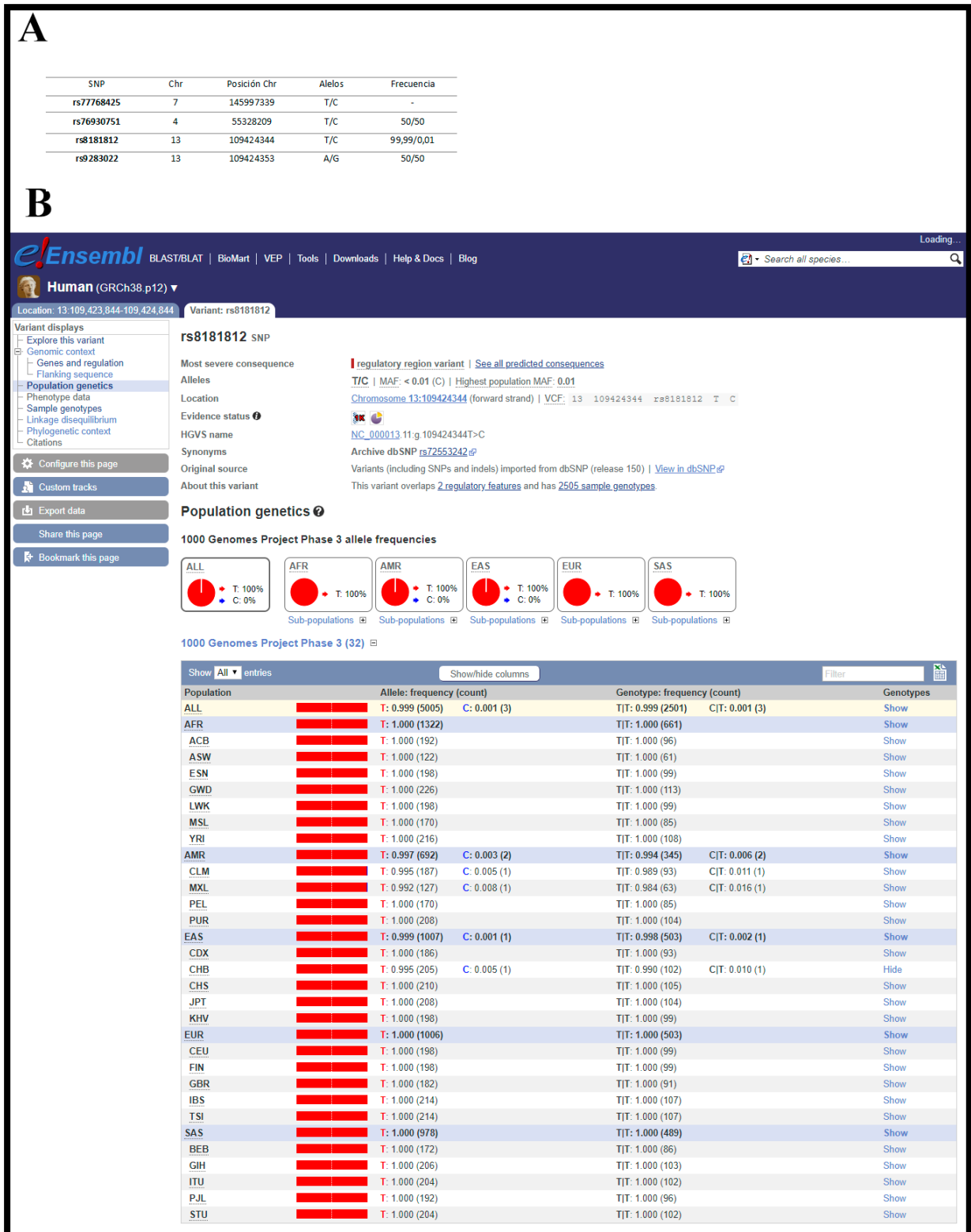
chr19 18,531,606-18,531,627 22 bp

chr19 18,531,628-18,531,632 5 bp

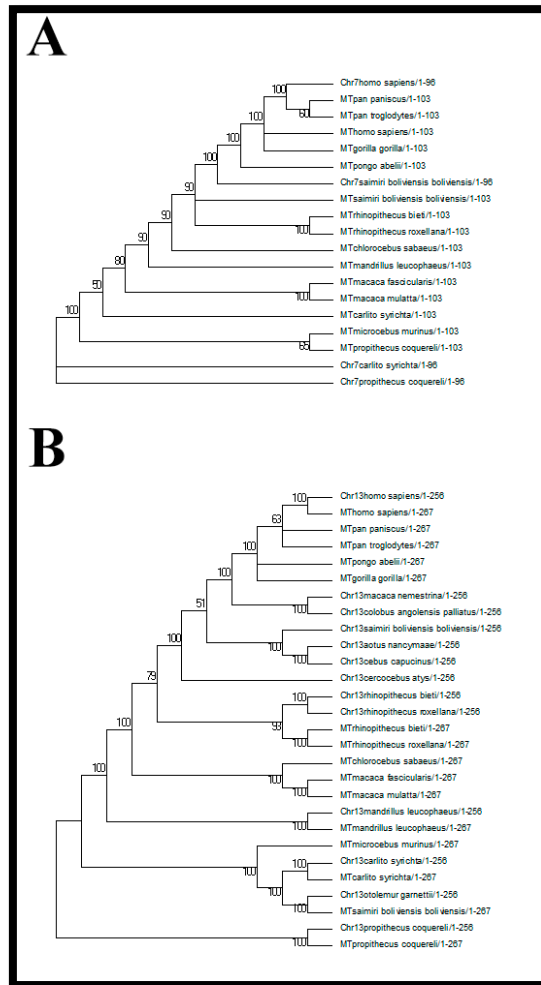
B

ACTIONS	QUERY	SCORE	START	END	QSIZE	IDENTITY	CHRO	STRAND	START	END	SPAN
browser details	KH889667.1	1614	6213	9372	16574	98.6%	1	+	103007864	103010209	2346
browser details	KH889667.1	1489	1074	3343	16574	97.9%	10	+	16240606	16242866	2259
browser details	KH889667.1	1412	13041	13456	16574	99.0%	6	-	103695922	103697337	1416
browser details	KH889667.1	1412	1	2522	16574	85.5%	3	+	165344782	165374407	29626
browser details	KH889667.1	1389	11252	13683	16574	82.5%	11	-	67454168	67459419	5252
browser details	KH889667.1	1362	9	3248	16574	83.5%	13	+	113331396	113334519	3214
browser details	KH889667.1	1358	19	5320	16574	83.0%	8	-	45677862	45682491	4630
browser details	KH889667.1	1315	5	2736	16574	84.1%	5	+	141507493	141510793	3301
browser details	KH889667.1	1305	154	2522	16574	85.1%	3	+	165338660	165338931	232
browser details	KH889667.1	1295	5	2736	16574	84.4%	10	+	6472738	6475429	2692
browser details	KH889667.1	1250	1459	4501	16574	82.6%	8	+	30773716	30776718	3003
browser details	KH889667.1	1222	14019	15274	16574	98.4%	Un	MH001000451	-	1	1253 1253
browser details	KH889667.1	1207	741	3797	16574	83.3%	X	-	109537060	109540103	3044
browser details	KH889667.1	1159	14229	16395	16574	84.2%	9	-	40468093	40470235	2143
browser details	KH889667.1	938	9	2181	16574	85.3%	1	+	207065693	207067650	2158
browser details	KH889667.1	895	1196	2414	16574	89.7%	6	-	70633359	70634569	1211
browser details	KH889667.1	882	8712	12995	16574	78.4%	3	+	63468544	63472627	4084
browser details	KH889667.1	865	1043	3118	16574	85.7%	8	+	91224312	91225996	1275
browser details	KH889667.1	864	472	11792	16574	81.4%	2	-	25887623	25896996	9374
browser details	KH889667.1	850	1	1321	16574	85.7%	19	-	18510717	18513026	1310
browser details	KH889667.1	787	1272	2522	16574	86.0%	3	+	165344035	165347266	1252
browser details	KH889667.1	781	3986	5303	16574	84.1%	5	+	43081375	43082690	1316
browser details	KH889667.1	776	443	2292	16574	81.5%	5	+	104571599	104573325	1727
browser details	KH889667.1	764	7313	8588	16574	83.0%	Un	MH001123921	-	14	1317 1304
browser details	KH889667.1	762	7034	8392	16574	85.5%	Un	MH001049836	-	257	1623 1367
browser details	KH889667.1	738	11347	12086	16574	99.0%	X	+	86299917	86300656	740
browser details	KH889667.1	726	9362	10780	16574	85.3%	Un	MH001301450	-	19	1360 1342
browser details	KH889667.1	700	740	2125	16574	87.8%	12	-	18359333	18362667	3335
browser details	KH889667.1	694	12906	16395	16574	79.1%	1	-	153681433	153682700	1368

Anexo 8. Región del cromosoma 19 de *N.larvatus* donde se encuentra el Numt (18531381-18531632). (A) Representación de cada uno de los bloques en los que el navegador genómico UCSC divide esta secuencia de *N.larvatus* alineada con la secuencia del Numt del cromosoma 13 en el ser humano. (B) Tras realizar el BLAT entre el genoma mitocondrial y el genoma nuclear de esta especie, se obtuvieron varios hits de los cuales el encuadrado en azul pertenece a la región donde se encuentra el Numt, con un 85,7% de identidad y ubicado en la cadena complementaria del cromosoma 19



Anexo 9. (A) Localización y frecuencia alélica de cada uno de los SNPs seleccionados y presentes en nuestros NumtS y (B) distribución de las variantes alélicas de rs8181812 en la población mundial. (A) El alelo representado a la izquierda consiste en el alelo de referencia, mientras que el de la derecha consiste en la variante. Ambos alelos están presentes en la cadena principal. La frecuencia de rs77768425 no está muy bien determinada. Datos obtenidos de la base de datos de SNPs de Ensembl (B) Distribución de los polimorfismos T/C en la población mundial, presentándose casos para el alelo C en tres individuos: uno asiático y dos americanos



Anexo 10. Árboles filogenéticos máximo parsimonios inferidos a partir de regiones alineadas en otros Primates con NumtS en el genoma nuclear humano, así como secuencias mitocondriales de Primates de la que provienen estos NumtS. (A) Árbol filogenético del Numt del cromosoma 7 (B) y del Numt del cromosoma 13. Las cifras sobre las ramas internas indican el % de árboles en los que aparece una rama como ésta, y dan apoyo a la topología obtenida. Para la creación de estos árboles filogenéticos se hizo uso de un método mini-mini heurístico.

REFERENCIAS

- Arctander, P. (1995). Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proceedings. Biological Sciences*, 262(1363), 13–19.
<https://doi.org/10.1098/rspb.1995.0170>
- Bensasson, D., Feldman, M. W., & Petrov, D. A. (2004). Rates of DNA Duplication and Mitochondrial DNA Insertion in the Human Genome. *Journal of Molecular Evolution*, 58(3), 361–361. <https://doi.org/10.1007/s00239-004-8301-1>
- Birney, E., Andrews, T. D., Bevan, P., Caccamo, M., Chen, Y., Clarke, L., ... Clamp, M. (2004). An Overview of Ensembl. *Genome Research*, 14(5), 925–928.
<https://doi.org/10.1101/gr.1860604>
- Calabrese, F. M., Balacco, D. L., Preste, R., Diroma, M. A., Forino, R., & Ventura, M. (2017). NumtS colonization in mammalian genomes. *Scientific Reports*, 1–10.
<https://doi.org/10.1038/s41598-017-16750-2>
- Dayama, G., Emery, S. B., Kidd, J. M., & Mills, R. E. (2014). The genomic landscape of polymorphic human nuclear mitochondrial insertions. *Nucleic Acids Research*, 42(20), 12640–12649. <https://doi.org/10.1093/nar/gku1038>
- Gasior, S. L., Wakeman, T. P., Xu, B., & Deininger, P. L. (2006). The Human LINE-1 Retrotransposon Creates DNA Double-strand Breaks. *Journal of Molecular Biology*, 357(5), 1383–1393. <https://doi.org/10.1016/j.jmb.2006.01.089>
- Hazkani-covo, E. (2007). Mitochondrial Insertions into Primate Nuclear Genomes Suggest the Use of numts as a Tool for Phylogeny, 1–5. <https://doi.org/10.1093/molbev/msp131>
- Hazkani-covo, E., & Covo, S. (2008). Numt-Mediated Double-Strand Break Repair Mitigates Deletions during Primate Genome Evolution, 4(10).
<https://doi.org/10.1371/journal.pgen.1000237>
- Hazkani-Covo, E., Sorek, R., & Graur, D. (2003). Evolutionary dynamics of large numts in the human genome: rarity of independent insertions and abundance of post-insertion duplications. *Journal of Molecular Evolution*, 56(2), 169–174.
<https://doi.org/10.1007/s00239-002-2390-5>
- Hazkani-covo, E., Zeller, R. M., & Martin, W. (2010). Molecular Poltergeists : Mitochondrial DNA Copies (numts) in Sequenced Nuclear Genomes, 6(2).

<https://doi.org/10.1371/journal.pgen.1000834>

Higgins, D. G., Thompson, J. D., & Gibson, T. J. (1996). Using CLUSTAL for multiple sequence alignments. *Methods in Enzymology*, 266, 383–402. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8743695>

Ivessa, A. S. (2018). Mitochondrial DNA escape through BAK / BAX macro-pores – implications in numtogenesis underlying aging and cancer, (2018). <https://doi.org/10.9777/rr.2018.10320.www.companyofscientists.com/index.php/rr>

Jensen-Seaman, M. I., Wildschutte, J. H., Soto-Calderón, I. D., & Anthony, N. M. (2009). A comparative approach shows differences in patterns of numt insertion during hominoid evolution. *Journal of Molecular Evolution*, 68(6), 688–699. <https://doi.org/10.1007/s00239-009-9243-4>

Knoop, V., Ehrhardt, T., Lättig, K., & Brennicke, A. (1995). The gene for ribosomal protein S10 is present in mitochondria of pea and potato but absent from those of Arabidopsis and Oenothera. *Current Genetics*, 27(6), 559–564. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7553942>

Lee, S., Oh, S., Kang, K., & Han, K. (2014). BLAT-Based Comparative Analysis for Transposable Elements: BLATCAT. *BioMed Research International*, 2014, 1–7. <https://doi.org/10.1155/2014/730814>

McArthur, K., Whitehead, L. W., Heddleston, J. M., Li, L., Padman, B. S., Oorschot, V., ... Kile, B. T. (2018). BAK/BAX macropores facilitate mitochondrial herniation and mtDNA efflux during apoptosis. *Science*, 359(6378), eaao6047. <https://doi.org/10.1126/science.aao6047>

Mercer, T. R., Neph, S., Dinger, M. E., Crawford, J., Smith, M. A., Shearwood, A. J., ... Mattick, J. S. (2012). The human mitochondrial transcriptome, 146(4), 645–658. <https://doi.org/10.1016/j.cell.2011.06.051>.The

Park, C. B., & Larsson, N.-G. (2011). Mitochondrial DNA mutations in disease and aging. *The Journal of Cell Biology*, 193(5), 809–818. <https://doi.org/10.1083/jcb.201010024>

Ricchetti, M., Tekaiia, F., & Dujon, B. (2004). Continued colonization of the human genome by mitochondrial DNA. *PLoS Biology*, 2(9). <https://doi.org/10.1371/journal.pbio.0020273>

- Rouzic, A. LE, & Deceliere, G. (2005). Models of the population genetics of transposable elements. *Genetical Research*, 85(3), 171. <https://doi.org/10.1017/S0016672305007585>
- Taanman, J.-W. (1999). The mitochondrial genome: structure, transcription, translation and replication. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1410(2), 103–123. [https://doi.org/10.1016/S0005-2728\(98\)00161-3](https://doi.org/10.1016/S0005-2728(98)00161-3)
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, 30(12), 2725–2729. <https://doi.org/10.1093/molbev/mst197>
- Tourmen, Y., Baris, O., Dessen, P., Jacques, C., Malthièry, Y., & Reynier, P. (2002). Structure and chromosomal distribution of human mitochondrial pseudogenes. *Genomics*, 80(1), 71–77. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12079285>
- Tsuji, J., Frith, M. C., Tomii, K., & Horton, P. (2012). Mammalian NUMT insertion is non-random. *Nature Reviews Genetics*, 13(12), 907–908. <https://doi.org/10.1093/nar/gks424>
- Wallace, D. C. (2010). Mitochondrial DNA mutations in disease and aging. *Environmental and Molecular Mutagenesis*, 41(5), n/a-n/a. <https://doi.org/10.1002/em.20586>
- Yang, D., Oyaizu, Y., & Oyaizu, H. (1985). Mitochondrial origins. *Journal of Molecular Evolution*, 21(4), 444–447.
- Yates, A., Akanni, W., Amode, M. R., Barrell, D., Billis, K., Carvalho-Silva, D., ... Flicek, P. (2016). Ensembl 2016. *Nucleic Acids Research*, 44(D1), D710–6. <https://doi.org/10.1093/nar/gkv1157>
- [on-line] <http://genome.ucsc.edu/cgi-bin/hgBlat>
- [on-line] Hall T. Manual of BioEdit version 7.0.0.