UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS. ***Fundada en 1551***
FACULTAD DE CIENCIAS BIOLÓGICAS
E.A.P. DE CIENCIAS BIOLÓGICAS

Detección de la variabilidad genética en la población de Puno utilizando marcadores STR DYS390, DYS391, DYS392 del cromosoma Y

TESIS: para optar el Título Profesional de: BIÓLOGO con mención en: BIOLOGÍA CELULAR Y GENÉTICA

PAUL WENCESLAO LÓPEZ GONZÁLEZ
ASESOR: Dra. BEATRÍZ RAQUEL LIZÁRRAGA DE OLARTE

LIMA -PERÚ 2004

. . 1

AGRADECIMIENTOS . 3

RESUMEN . 5

ABSTRACT . 7

INTRODUCCIÓN . 9

OBJETIVOS . . 11

OBJETIVO GENERAL: . 11

OBJETIVOS ESPECÍFICOS: . . 11

ANTECEDENTES . 13

MATERIALES Y MÉTODOS . 17

POBLACIÓN EN ESTUDIO . 17

TAMAÑO DE MUESTRA . 17

EXTRACCIÓN Y CUANTIFICACIÓN . . 18

AMPLIFICACIÓN . 18

VISUALIZACIÓN . 20

VALORACIÓN ESTADÍSTICA Y ANÁLISIS. . . 20

RESULTADOS . . 23

AMPLIFICACIÓN . 23

VALORACIÓN ESTADÍSTICA . . 24

DISCUSIÓN . 31

CONCLUSIONES . . 33

RECOMENDACIONES . 35

ILUSTRACIONES . 37

BIBLIOGRAFÍA . 49

ANEXOS . 53

DEDICATORIA *A mis padres Wenceslao y Lisbeth Por el apoyo y consejos a lo largo de mi formación personal y profesional. A la Dra. Beatriz Lizárraga Por las oportunidades y confianza otorgado en una formación profesional.*

"Programa Cybertesis PERÚ - Derechos son del Autor"

AGRADECIMIENTOS

A la Dra Beatriz Lizárraga Asesora de Investigación y Jefe del Laboratorio de Biología Molecular del Centro de Investigación de Bioquímica y Nutrición - Facultad de Medicina - UNMSM.

Al MSc. Ángel Medina Colque por lo aportado para dar inicio al presente trabajo de estudio con la población de Puno.

A los Blgs. Olimpio Ortega Dávila, Raul Y. Tito Tadeo, Gian Carlo Iannacone por el apoyo y asesoría técnica.

Al Laboratorio Genes en Medellín Colombia que a través del MSc. Juan José Builes permitió mejorar y resolver técnicamente lo planteado en el presente trabajo.

"Programa Cybertesis PERÚ - Derechos son del Autor"

RESUMEN

En el presente trabajo se estudió una muestra poblacional de ascendencia puneña, detectando su variabilidad genética a través de los marcadores microsatélites STR del Cromosoma Y: DYS390, DYS391 y DYS392 utilizando la técnica de amplificación Reacción en Cadena de la Polimerasa (PCR). Al comparar las frecuencias alélicas obtenidas con lo publicado a nivel mundial para la Población Amerindia, se encontraron diferencias significativas para el STR DYS392. Utilizando los datos publicados en la literatura se diseñaron dos cuadros de frecuencias haplotípicas poblacionales, uno a nivel de ocho poblaciones sudamericanas y otro de poblaciones a nivel mundial (tabla 16); mediante estos cuadros se ha encontrado que de las poblaciones sudamericanas la población de Puno estaría ubicada con las otras dos poblaciones peruanas, una de Arequipa y la segunda de Tayacaja trabajadas en la referencia, manteniéndose en nuestras poblaciones una distinción a nivel de Sudamérica. En el análisis con las diferentes poblaciones mundiales, la población puneña tendría mayor similitud genética con las poblaciones asiáticas, dándonos una muy buena correlación con la teoría acerca de la migración de los primeros hombres al continente americano así como de la conservación de caracteres genéticos de esta población.

"Programa Cybertesis PERÚ - Derechos son del Autor"

ABSTRACT

In this work one Puno population sample was studied by the PCR (Polymerase Chain Reaction) amplification technique to detect the genetic variability of Y Chromosome STRs Markers DYS390, DYS391 and DYS392. These were compared with alleles frequencies with closest populations reported (Amerindian Population), the results show significant differences for the DYS392 STR. In addition two tables of population haplotype frequencies were constructed to compare South American populations with our Puno population and the other to compare with world-wide populations. The Puno population fits with the other peruvian populations studied (Arequipa and Tayacaja) keeping its distinction in South America. When compared with world-wide populations, the Puno population is similar with the Asiatic populations showing good correlation with the theories of human migration for peopling of the Americas as well as the conservation of genetic characters in this population.

"Programa Cybertesis PERÚ - Derechos son del Autor"

INTRODUCCIÓN

Actualmente existe una gran necesidad de realizar estudios genético poblacionales para una mejor comprensión de lo ocurrido en el contexto bio-social de nuestros antepasados, los cuales a su vez sirven de apoyo elemental a los trabajos relacionados a las ciencias forenses.

El análisis de polimorfismos de DNA ha significado un cambio radical en la investigación de antropología, evolución, análisis de paternidad, así como de análisis forenses y diagnóstico de enfermedades. Antes del desarrollo de estas metodologías, se solucionaban la casi totalidad de los casos con los marcadores clásicos. El análisis del polimorfismo del DNA ha simplificado estas pruebas, las ha hecho más rápidas y baratas, permitiendo además una mejor solución a casos difíciles.

Los microsatélites STR (Short Tandem Repeat) son secuencias de nucleótidos que consisten de repeticiones en tandem, estas secuencias han sido encontradas frecuentemente por todo el genoma eucariótico y el número de unidades repetitivas son altamente polimórficos, es decir variables entre los individuos. Estas secuencias repetitivas pueden ser usadas como marcadores genéticos para estudios genético poblacionales, objetivo del presente trabajo, como para análisis de genética forense que usa por ejemplo frecuencias alélicas de marcadores polimórficos para valoración estadística de resultados genéticos (6, 7). Los microsatélites poseen una herencia mendeliana simple, lo que significa que un individuo hereda uno de los alelos de cada progenitor. En el caso de los STR de cromosoma Y, estos son heredados de padre a hijo (varón) mas no a la hija. Los marcadores utilizados en análisis patrilineal (10) en regiones

de no recombinación son transmitidos en bloque a los descendientes varones (9). Consecuentemente los haplotipos del cromosoma Y (9) pueden ser rápidamente diseñados por combinación de las variantes alélicas de múltiples marcadores y estos haplotipos son estables excepto cuando varían por una mutación que crea un nuevo polimorfismo. Los polimorfismos de esta manera proveen información acerca del parentesco del cromosoma Y, que nos guiaría hacia un antecesor común (17).

Con respecto a estos marcadores microsatélites STR, son pocos los trabajos publicados acerca de poblaciones peruanas, teniendo tan solo publicados a nivel mundial frecuencias de poblaciones amerindias (como datos más cercanos) que no necesariamente reflejan el contenido genético de muchas poblaciones como es el caso de las poblaciones peruanas. El presente trabajo busca acortar este vacío, además de aportar con nuevas metodologías para las diferentes aplicaciones que puedan derivarse.

A continuación se detallarán no solo los antecedentes que competen a dicho trabajo sino metodologías, resultados y sobre todo aportes a lo anteriormente expuesto.

"Programa Cybertesis PERÚ - Derechos son del Autor"

OBJETIVOS

OBJETIVO GENERAL:

- Establecer una base de datos de las frecuencias alélicas para los marcadores DYS390, DYS391 y DYS392 en la muestra poblacional en estudio.

OBJETIVOS ESPECÍFICOS:

- Estandarizar una metodología experimental apropiada para el análisis de estos microsatélites, para trabajos genético-poblacionales y ciencias forenses.
- Determinar las variaciones alélicas de los marcadores DYS390, DYS391 y DYS392 en la muestra poblacional en estudio, con respecto a los alelos publicados mundialmente.
- Determinar las frecuencias alélicas de los marcadores DYS390, DYS391 y DYS392 en las muestra poblacional en estudio.

- Comparar la población en estudio a nivel de frecuencias haplotípicas con otras poblaciones mundiales.

ANTECEDENTES

El DNA codificante, aunque es el más interesante desde el punto de vista biológico, posee en general poca variabilidad entre las personas, con la excepción de la región HLA. Desde el punto de vista de la identificación es mucho más interesante el DNA no codificante que no es transcrito a RNA y que no supone nada en términos de producción de proteínas, aunque tiene otras funciones biológicas importantes. Este tipo de DNA es, en general, muy variable entre las personas y por eso se dice "polimórfico". Representa, además, cuantitativamente, la mayor parte del genoma humano.

Aproximadamente la mitad del DNA no codificante es DNA repetitivo y aunque gran parte del mismo es extremadamente polimórfico, el DNA utilizado con diversos fines es el DNA repetido en tandem y dentro de él, el DNA minisatélite y microsatélite.

El DNA minisatélite y microsatélite consiste en repeticiones de secuencias de DNA de número variable, por lo que genéricamente se denominan VNTR ("variable number of tandem repeats"). Las repeticiones en el DNA microsatélite son de tamaño pequeño (de 1 a 6 pares de bases) por lo que se suelen denominar STRs ("short tandem repeats). Las repeticiones en un locus minisatélite tienen un tamaño medio de alrededor de 28 pares de bases", (se ha acuñado el término “variable number tandem repeat VNTR, para designar a los minisatélites si la secuencia repetitiva es de 9 a 64 nucleótidos) aunque esto se presta a confusión porque todos estos polimorfismos, finalmente, representan alguna forma de repetición en tandem (14), que son comúnmente usados en las técnicas de DNA Fingerprinting o Huella dactilar del DNA.

Poniendo un ejemplo, un STR puede tener una estructura como ACTT ACTT ACTT

ACTT ACTT ACTT ACTT ACTT...hasta un número n de repeticiones. Los individuos nos diferenciamos por el número de repeticiones de esa secuencia. Un individuo 8-12 para ese STR significa que tiene 8 veces la unidad de repetición (ACTT) en un lugar específico de un cromosoma (locus génico) y 12 veces en la misma posición de su cromosoma homologo.

Los minisatélites y microsatélites además de ser polimórficos (esto es, variables entre los individuos), poseen una herencia mendeliana simple. Esto significa que el individuo 8-12, utilizado de ejemplo, ha heredado uno de los alelos de su madre y otro de su padre biológico. En el caso de STR de cromosoma Y la segregación es de padre a hijo (varón) mas no a hija.

La PCR es una técnica de amplificación in vitro de pequeños segmentos de DNA con la que a partir de una sola copia, es decir a partir de una sola célula, se pueden hacer millones de copias, de modo que el producto amplificado puede ser fácilmente analizado, incluso sin recurrir al uso de sondas.

Los marcadores utilizados en análisis patrilineales (10) de regiones de no recombinación son transmitidos en bloque a los descendientes varones (9). Es decir, los haplotipos del cromosoma Y (9) pueden ser rápidamente construidos por combinación de las variantes alélicas de múltiples marcadores y estos haplotipos son estables excepto cuando varían por una mutación que crea un nuevo polimorfismo. Los polimorfismos de esta manera proveen información acerca del parentesco del cromosoma Y, que nos guiaría hacia un antecesor común (17).

Los STR sirven también para la identificación de individuos y para establecer relaciones de parentesco (figura 01). El poder de los análisis de DNA forenses está en los polimorfismos de los loci STR y el número de loci utilizados, la inclusión puede ser usualmente por estimación de genotipos STR o frecuencias alélicas en poblaciones relevantes de potencial inclusión, por lo tanto el análisis de STR de cromosoma Y está dentro de esta aplicación por ser de utilidad en identificación de individuos varones, en casos de violaciones y casos de paternidad con hijos varones que arrojan resultados insuficientes, ya sean estadísticos o por ausencia de familiares al momento del muestreo. Muchos trabajos forenses alrededor del mundo con DNA de cromosoma Y están centrados en 7 a 9 STR (figura 02), aunque existen reportes de haplotipos inusuales que fueron originados por mutación y locus duplicados (8).

Los polimorfismos únicos, son usados para Estudios Poblacionales cuando grupos de personas comparten el estado derivado de un marcador, las frecuencias de estos marcadores pueden ser usadas para calcular parentescos poblacionales utilizando múltiples marcadores (figura 02). Poblaciones con diferentes distribuciones de microsatélites son distintas, sin embargo poblaciones con distribuciones similares no son necesariamente similares. Esto es porque la evolución convergente puede guiar a diferentes cromosomas Y (definido por las mutaciones únicas) llevando la misma distribución de alelos (11), estas dificultades son ilustradas por algunos estudios (16, 17, 22, 23).

El STR DYS390 (Gen Bank G09611, **DYS390-27 (8-14-1-4), 227 bp.** que presenta 27 repeticiones) localizado en el cromosoma Y, marcador molecular en estudio, contiene

N repeticiones □TCTG□n □TCTA□m □TCTG□p □TCTA□q, donde N es igual a: n + m + p + q, entonces N puede ser 18, 19, 20,21, 22, 23, 24, 25, 26, 27. (Ver anexos tabla 1, 2 y cuadro1). Las frecuencias alélicas del DYS390 para las poblaciones a nivel mundial, han sido ya reportadas según anexo Tabla 05: DYS390 Global allele frequency distribution of DYS390 (4).

El STR DYS391 (Gen Bank G09613, **DYS391-11, 288 bp.** que presenta 11 repeticiones) localizado en el cromosoma Y, marcador molecular en estudio, contiene □TCTA□n repeticiones, donde “n” puede ser 6, 7, 8, 9, 10, 11, 12, 13, 14. (Ver anexos tabla 1, 3 y cuadro2). Las frecuencias alélicas del DYS391 para las poblaciones a nivel mundial, han sido ya reportadas según anexo Tabla 06DYS391 Global allele frequency distribution of DYS391 (4).

El STR DYS392 (Gen Bank G09867, **DYS392-16, 264 bp.** que presenta 16 repeticiones) localizado en el cromosoma Y, marcador molecular en estudio, contiene □TAT□n repeticiones, donde “n” puede ser 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17.. (Ver anexos tabla 1,4 y cuadro3). Las frecuencias alélicas del DYS392 para las poblaciones a nivel mundial, han sido ya reportadas según anexo Tabla 07DYS392 Global allele frequency distribution of DYS392 (4).

"Programa Cybertesis PERÚ - Derechos son del Autor"

MATERIALES Y MÉTODOS

POBLACIÓN EN ESTUDIO

Mediante campañas de salud, que proporcionaron evaluación médica general por personal médico especialista de la Facultad de Medicina de la Universidad Nacional Mayor de San Marcos, se seleccionó un grupo de personas para participar en el estudio (Provincias de Juliaca y Puno, departamento de Puno) (13). A estos participantes se les dio una capacitación y encuesta para la participación en el presente estudio (Ver Formato de Consentimiento en anexos), en la cual los participantes de manera libre, voluntaria, sin remuneración se ofrecen como donantes de 5ml de sangre.

TAMAÑO DE MUESTRA

La muestra está constituida por 20 personas no emparentadas entre si, aparentemente sanas y que cumplan los criterios de inclusión.

Criterios de Inclusión:

- Varones mayores de 18 años (para legalidad de su participación).
- Personas cuya ascendencia dos generaciones atrás provenga de localidades del departamento de Puno.
- Personas que acepten participar en el estudio y firmar el compromiso de participación.

Criterios de Exclusión:

- Personas de sexo femenino así como varones menores de 18 años.
- Personas cuya ascendencia dos generaciones atrás no provenga de localidades del departamento de Puno.
- Personas que no acepten participar en el estudio.

EXTRACCIÓN Y CUANTIFICACIÓN

La colección se realizó por extracción de sangre venosa mediante sistema de vacío usando tubos vacutainer Venoject II EDTA(K2) estériles, y por centrifugación se obtuvo los leucocitos previa hemólisis. Para purificar el DNA, se utilizó el método de cloroformo - alcohol isoamílico (12). Cada solución de DNA fue diluida 1:10 en agua milliQ. A estas diluciones se les midió la absorbancia en un espectrofotómetro (LKB), empleando celdas de 1 ml de cuarzo y tomando lecturas a 260 y 280 nm.

AMPLIFICACIÓN

Se usó la técnica de PCR con los primers DYS390, DYS391 y DYS392 según el CHLC (Cooperative Human Linkage Center, publicados por Murray, J., Sheffield, V, Weber, J. L., Duyk, G. and Buetow, K. H.) para flanquear la región del STR DYS390, DYS391 y DYS392 en el cromosoma Y. Estos cebadores amplifican un fragmento de 191 hasta 227 pares de bases (TABLA 13), 268 hasta 300 pares de bases (TABLA 14), 234 hasta 267 pares de bases (TABLA 15), respectivamente. Luego de calcular las temperaturas de denaturación teóricos para los primers utilizados en este trabajo (2, 3, 5), se procedió a realizar las amplificaciones con tres diferentes temperaturas de alineamiento en el perfil térmico, estandarizando una sola temperatura que sea común, para luego amplificar estas regiones siguiendo el protocolo base descrito a continuación:

Preparación del Pre-Mix PCR (x 4 Muestras)

Buffer 10x	10 ul
MgCl2 (1.5mM)	08 ul
dNTP´s (200uM)	08 ul
Taq 3U	0.8 ul
H2O	1.2 ul.
VOLUMEN TOTAL	28 ul.

Preparación del Mix- PCR (x Muestra)

Pre-Mix	7 ul
Primers(20pmoles/ul)	1 ul
DNA (80 ng)	1 ul
H2O	16 ul
VOLUMEN TOTAL	25 ul.

Programa de Amplificación o Perfil Térmico

Inicialización	95°C 11 minutos
Denaturación	94°C 01 minuto
Hibridación	59.1°C 01minuto 15 segundos
Extensión	72°C 01 minuto
Extensión Final	72°C 07 minutos
Hold	04°C

Como primer paso se prepara el pre-mix según el número de muestras que se desea amplificar para luego proceder a preparar el mix PCR en donde para mayor rapidez así como para evitar posibles errores de cálculo se trabaja con una concentración aproximada de 80 ng de DNA, esta aproximación es técnicamente válida ya que se está aprovechando el amplio rango de sensibilidad de la técnica PCR. En cuanto a los primers se preparó un mix para cada uno (primer forward y primer reverse) a concentraciones equimolares utilizando así una concentración de 20pmoles/ul, disminuyendo también la probabilidad de error al preparar el mix o mezcla especialmente cuando se va a amplificar un número considerable de muestras a la vez. Este Protocolo también es válido para realizar una amplificación simultánea con estos marcadores (2, 3, 5), pero teniendo que realizar algunas variaciones en el perfil térmico.

VISUALIZACIÓN

Como primer paso se realizaron corridas electroforéticas en geles de poliacrilamida de 8x12 cm al 10% 7M UREA y luego al 12% 7M UREA (acrilamida-bis 19:1) (1) logrando una visualización de fragmentos de distintos tamaños siendo importante también la determinación de calidad de amplificado, es decir sin productos inespecíficos así como de poca nitidez de los fragmentos buscados. Para una tipificación más exacta con la cual podamos asignar fragmentos, es decir determinar que alelos estarían presentes en las diferentes muestras, se prepararon geles de poliacrilamida de 38x30 cm al 4% 7M UREA (acrilamida-bis 19:1) (1), seguida por una tinción con nitrato de plata que hizo visible dicha migración.

La asignación de fragmentos se realizó por comparación de bandas de las muestras trabajadas con muestras de referencia conocidas, es decir fragmentos ya tipificados con los marcadores en estudio, proporcionadas por el laboratorio Genes- Medellín Colombia.

VALORACIÓN ESTADÍSTICA Y ANÁLISIS.

La valoración estadística fue realizada mediante el programa estadístico ARLEQUÍN (Software for Population Genetics Data Analysis versión 2.000) con el cual se realizaron test de índices de diversidad (por ejemplo Diversidad Molecular), estructura génica como es el caso de AMOVA y otros relacionados a comparaciones entre poblaciones. Una vez calculados estos parámetros estadísticos se procedió a utilizar el programa estadístico MEGA (11) con el cual se pudieron realizar las aproximaciones de comparación de poblaciones al generar cladogramas con los mismos.

Para realizar los análisis correspondientes, primero se compararon las frecuencias alélicas de la población puneña y lo publicado para poblaciones amerindias según tablas adaptadas de Knijff et al., 1997 (4) para los marcadores utilizados. Además, a partir de dos publicaciones relacionadas con el tema de este trabajo, es decir publicaciones en lo referente a estudios poblacionales con marcadores STR en cromosoma Y, se trabajaron las frecuencias haplotípicas de poblaciones sudamericanas (DYS390 y DYS391) y de diferentes poblaciones a nivel mundial (DYS390, DYS391 y DYS392) pudiendo de esta manera realizar comparaciones inter-poblacionales con la población de ascendencia puneña estudiada. La primera referencia es un trabajo de investigación titulado “Genetic Differentiation in South Amerindians Is Related to Environmental and Cultural Diversity: Evidence from the Y Chromosome” (21), en el cual se trabajan 6 marcadores STR del cromosoma Y incluyendo dos de los marcadores utilizados en el presente trabajo, en 8 poblaciones sudamericanas (ver figura 05) dentro de las cuales se tienen dos poblaciones peruanas; la segunda referencia utilizada titulada “An Extensive Análisis of Y-Chromosomal Microsatellite Haplotypes in Globally Dispersed Human Populations (10),

"Programa Cybertesis PERÚ - Derechos son del Autor"

contiene información más completa ya que trabajaron 20 poblaciones a nivel mundial (ver figura 06) con 7 marcadores STR del cromosoma Y incluyendo los tres marcadores utilizados en el presente trabajo, pudiendo hacer de esta manera comparaciones mucho más precisas con respecto a la población en estudio.

"Programa Cybertesis PERÚ - Derechos son del Autor"

RESULTADOS

AMPLIFICACIÓN

Mediante la técnica de PCR (Reacción en cadena de la Polimerasa), con las condiciones y metodologías descritas anteriormente se obtuvo buena cantidad de amplificado sin productos inespecíficos. Estos amplificados fueron visualizados por tinción con nitrato de plata luego de la electroforésis en geles de poliacrilamida 7M UREA 8x12cm al 10% y 12% (acrilamida-bis 19:1) (1). Los resultados de la amplificación con el STR DYS390 se muestran en las figuras 03 y 04, siendo el gel al 12% el que permite una mejor resolución de los amplificados, visualizándose los diferentes tipos de alelos en las muestras poblacionales trabajadas.

Por medio de corridas electroforéticas en geles de poliacrilamida 38x30cm al 4% (acrilamida-bis 19:1) (1), seguida por una tinción con nitrato de plata se consiguió hacer más visible dicha migración y tipificar las muestras en estudio, obteniéndose lo siguiente:

Tabla 01: Cuadro Comparativo de Frecuencias Alélicas para el Marcador STR DYS390.

STR DYS390					
ALELOS REPORTADOS	**21**	**22**	**23**	**24**	**25**
POBLACIÓN AMERINDIA (n=101)	15.8%	19.8%	26.7%	29.8%	7.9%
POBLACIÓN DE PUNO (n=20)	**0%**	**5%**	**40%**	**45%**	**10%**
	0	1	8	9	2

Tabla 02: Cuadro Comparativo de Frecuencias Alélicas para el Marcador STR DYS391.

STR DYS391				
ALELOS REPORTADOS	**9**	**10**	**11**	**12**
POBLACIÓN AMERINDIA (n=101)	4%	64.3%	29.7%	2%
POBLACIÓN DE PUNO (n=20)	**0%**	**85%**	**10%**	**5%**
	0	17	2	1

Tabla 03: Cuadro Comparativo de Frecuencias Alélicas para el Marcador STR DYS392.

STR DYS392						
ALELOS REPORTADOS	**10**	**11**	**12**	**13**	**14**	**15**
POBLACIÓN AMERINDIA (n=65)	6.2%	52.3%	1.5%	20%	16.9%	3.1%
POBLACIÓN DE PUNO (n=20)	**0%**	**5%**	**25%**	**40%**	**30%**	**0%**
	0	1	5	8	6	0

VALORACIÓN ESTADÍSTICA

Como primera etapa, mediante el programa estadístico ARLEQUÍN (Software for Population Genetics Data Analysis versión 2.000) se realizó un análisis individualizado a nivel de las frecuencias alélicas obtenidas para los marcadores DYS390, DYS391 y DYS392 en la población de Puno, con respecto a lo publicado para poblaciones Amerindias (Tablas 17, 18, 19) encontrándose:

Tabla 04: Cuadro Comparativo a nivel de Frecuencias Alélicas, Puno vs. Amerindia.

	DYS390	**DYS391**	**DYS392**
Test de diferenciación entre Puno (n=20) y Amerindia (n=101)	0.07991+-0.0019	0.14145+-0.0131	0.0000+-0.0000
Diversidad Génica de la población de Puno (n=20)	0.6579+-0.0648	0.2789+-0.1235	0.7211+-0.0506
Diversidad Génica de la población Amerindia (DYS390 y DYS391 n = 101, DYS392 n = 65)	0.7774+-0.0146	0.5006+-0.0395	0.6630+-0.0476

Además, se comparó la población de Puno con 8 poblaciones sudamericanas, construyéndose primero una tabla de frecuencias haplotípicas (datos no mostrados) con la cual se realizaron test de índices de diversidad, estructura génica como es el caso de AMOVA y otros relacionados a comparaciones entre poblaciones, encontrándose lo siguiente:

Tabla 05: Puno vs. Poblaciones Sudamericanas - Cuadro de Índices de Diversidad Génica a nivel de Frecuencias Haplotípicas, calculada como la probabilidad de que dos haplotipos elegidos al azar sean diferentes en la muestra.

		Índice de Diversidad Génica
1	Cayapa-Ecuador (n=26)	0.8154+-0.0466
2	Tayacaja-Perú (n=44)	0.7653+-0.0469
3	Arequipa-Perú (n=15)	0.8190+-0.0636
4	Gaviao-Brasil (n=34)	0.7362+-0.0443
5	Xavante-Brasil (n=5)	0.0000+-0.000
6	Wai Wai-Brasil (n=5)	0.4000+-0.2373
7	Karitiana-Brasil (n=8)	0.250+-0.1802
8	Ticuna-Brasil (n=32)	0.5262+-0.0683
9	Puno-Perú (n=20)	0.7421+-0.0705

F-ST Convencional de frecuencias haplotípicas *(16)*

	1	2	3	4	5	6	7	8	9
1	0.00000								
2	0.08031	0.00000							
3	0.08162	0.00763	0.00000						
4	0.01948	0.02233	0.04719	0.00000					
5	0.32071	0.16560	0.29470	0.19163	0.00000				
6	0.27557	0.12244	0.05344	0.24409	0.75000	0.00000			
7	0.24998	0.11863	0.25441	0.12109	-0.06870	0.62433	0.00000		
8	0.18502	0.23350	0.15579	0.16564	0.44831	0.46141	0.40621	0.00000	
9	0.12048	-0.00687	-0.02187	0.06066	0.26436	0.01752	0.22034	0.26036	0.00000

Tabla 06: Matriz FSTs Poblacional.

(M=N datos haploides, M=2N datos diploides) *(16)*

	1	2	3	4	5	6	7	8	9
1	0.00000								
2	0.08372	0.00000							
3	0.08514	0.00766	0.00000						
4	0.01967	0.02259	0.04834	0.00000					
5	0.38671	0.18104	0.34913	0.21274	0.00000				
6	0.32237	0.13062	0.05493	0.27984	1.38629	0.00000			
7	0.28765	0.12627	0.29358	0.12908	0	0.97906	0.00000		
8	0.20460	0.26592	0.16936	0.18109	0.59476	0.61880	0.52123	0.00000	
9	0.12838	0	0	0.06258	0.30769	0.01763	0.24890	0.30241	0.00000

Tabla 07: Matriz de coeficientes de Coancestros t/M=-ln(1-FST), que calcula las distancias poblacionales en donde estas han divergido t generaciones atrás, pero en tiempos cortos de divergencia

(M=N datos haploides, M=2N datos diploides) *(20)*

	1	2	3	4	5	6	7	8	9
1	0.00000								
2	0.08732	0.00000							
3	0.08887	0.00769	0.00000						
4	0.01987	0.02284	0.04953	0.00000					
5	0.47213	0.19847	0.41783	0.23706	0.00000				
6	0.38039	0.13953	0.05646	0.32291	3.00000	0.00000			
7	0.33329	0.13459	0.34123	0.13778	0	1.65194	0.00000		
8	0.22703	0.30464	0.18454	0.19652	0.81260	0.85669	0.68409	0.00000	
9	0.13699	0	0	0.06458	0.36028	0.01783	0.28261	0.35311	0.00000

Tabla 08: Matriz de Slatkin linearizada FSTs as t/M=FST/(1-FST), que considera un modelo demográfico simple, donde dos poblaciones de tamaño N han divergido hace t generaciones desde una población de tamaño idéntico, estas poblaciones han permanecido siempre separadas sin intercambios y migraciones.

Una vez calculados estos parámetros estadísticos se procedió a utilizar el programa estadístico MEGA (11) con el cual se pudieron realizar las aproximaciones de comparación de poblaciones al generar cladogramas UPGMA con los mismos (figura 13).

Como tercer análisis, se realizó de manera similar una comparación de la población de Puno con 20 poblaciones a nivel mundial, construyéndose primero una tabla de frecuencias haplotípicas (tabla 16) con la cual se realizaron test de índices de diversidad (por ejemplo Diversidad Génica), estructura génica como es el caso de AMOVA y otros relacionados a comparaciones entre poblaciones, se encontró lo siguiente:

Tabla 09: Puno vs. Poblaciones Mundiales -Cuadro de Índices de Diversidad Génica a nivel de Frecuencias Haplotípicas, calculada como la probabilidad de que dos haplotipos elegidos al azar sean diferentes en la muestra.

		Índice de Diversidad Génica
1	Buenos Aires Metropolitans (BAM) (n=100)	0.9345+-0.0144
2	Bascos (BAS) (n=30)	0.7195+-0.0745
3	China (CHI) (n=36)	0.9048+-0.0354
4	Dutch (DUT) (n=88)	0.6094+-0.0136
5	Alemania (GER) (n=88)	0.9425+-0.0096
6	India (IND) (n=25)	0.6167+-0.0304
7	Indonesia (INO) (n=69)	0.9096+-0.0186
8	Artico (INU) (n=62)	0.7314+-0.0487
9	Italia (ITA) (n=100)	0.9147+-0.0135
10	Mongolia (MON) (n=40)	0.9141+-0.0220
11	Nativos Sudamericanos (NSA) (n=46)	0.9237+-0.0198
12	Papua Nueva Guinea (PNG) (n=26)	0.7723+-0.0778
13	Pigmeos de Africa Central (CAP) (n=31)	0.8882+-0.0346
14	Roro de la Costa Sur de Nueva Guinea (RPN) (n=12)	0.8636+-0.0786
15	Surinam (SUR) (n=54)	0.9294+-0.0215
16	Suiza (SWI) (n=64)	0.9390+-0.0103
17	Aborígenes Taiwaneses (TAW) (n=30)	0.7908+-0.0672
18	Tolai de Nueva Bretaña (TNB) (n=16)	0.9583+-0.0363
19	Isleños de Trobiand (TRO) (n=59)	0.9077+-0.0165
20	Samoa (WSA) (n=10)	0.7778+-0.1374
21	Puno-Perú (PUNO) (n=20)	0.8842+-0.0535

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0.00000															
2	0.04629	0.00000														
3	0.04639	0.14480	0.00000													
4	0.02[illegible]05	0.07705	0.06571	0.00000												
5	0.02601	0.11074	0.05903	0.02125	0.00000											
6	0.04759	0.14746	0.06319	0.04690	0.04981	0.00000										
7	0.04804	0.13143	0.01115	0.03384	0.02826	0.02497	0.0000									
8	0.00746	0.18873	0.12571	0.12771	0.13059	0.16455	0.18055	0.00000								
9	0.02583	0.13330	0.05910	0.03773	0.03415	0.06387	0.06104	0.13418	0.00000							
10	0.05[illegible]85	0.16652	0.04616	0.02590	0.01547	0.07364	0.05007	0.14430	0.05766	0.00000						
11	0.04[illegible]30	0.13303	0.03005	0.03452	0.02237	0.06754	0.05397	0.13000	0.05821	0.06429	0.00000					
12	0.05508	0.10605	0.10214	0.06453	0.10756	0.11286	0.05568	0.02901	0.11305	0.12583	0.12477	0.00000				
13	0.07560	0.10406	0.10882	0.08720	0.08407	0.07390	0.08006	0.10960	0.08130	0.06316	0.08532	0.15181	0.00000			
14	0.07560	0.18893	0.09998	0.08005	0.07552	0.02403	0.05913	0.20950	0.08828	0.10458	0.08624	0.17032	0.12348	0.00000		
15	0.04804	0.15740	0.07908	0.04439	0.02878	0.06415	0.05375	0.15705	0.04261	0.04300	0.00857	0.11076	0.01347	0.05257	0.00000	
16	0.01804	0.08348	0.04282	0.00507	0.00312	0.02556	0.03577	0.09555	0.00820	0.02329	0.04289	0.09951	0.05747	0.00533	0.02662	0.00000
17	0.00947	0.20878	0.00212	0.12781	0.11873	0.13892	0.04671	0.16724	0.12393	0.09834	0.08059	0.14724	0.15489	0.17489	0.13043	0.10289
18	0.01231	0.13715	0.03029	0.03868	0.02874	0.01373	0.02424	0.11304	0.03561	0.04365	0.04340	0.06056	0.07807	0.00815	0.04494	0.01035
19	0.05[illegible]02	0.14721	0.01388	0.00925	0.00159	0.02671	0.02475	0.08522	0.07489	0.06895	0.04132	0.07973	0.10104	0.11157	0.07131	0.05170
20	0.1[illegible]021	0.24389	0.13282	0.13810	0.12125	0.14120	0.12715	0.21360	0.13544	0.13514	0.13214	0.18756	0.15007	0.17791	0.15225	0.11027
21	0.04507	0.18004	0.03136	0.07485	0.06682	0.05274	0.00765	0.13162	0.06093	0.05057	0.05821	0.07500	0.11230	0.12177	0.05059	0.05040

	17	18	19	20	21
17	0.00000				
18	0.10629	0.00000			
19	0.03704	0.03384	0.00000		
20	0.15517	0.10054	0.13300	0.00000	
21	0.00818	0.04351	0.01359	0.14490	0.00000

Tabla 10: Matriz FSTs Poblacional. F-ST Convencional de frecuencias haplotípicas (16)

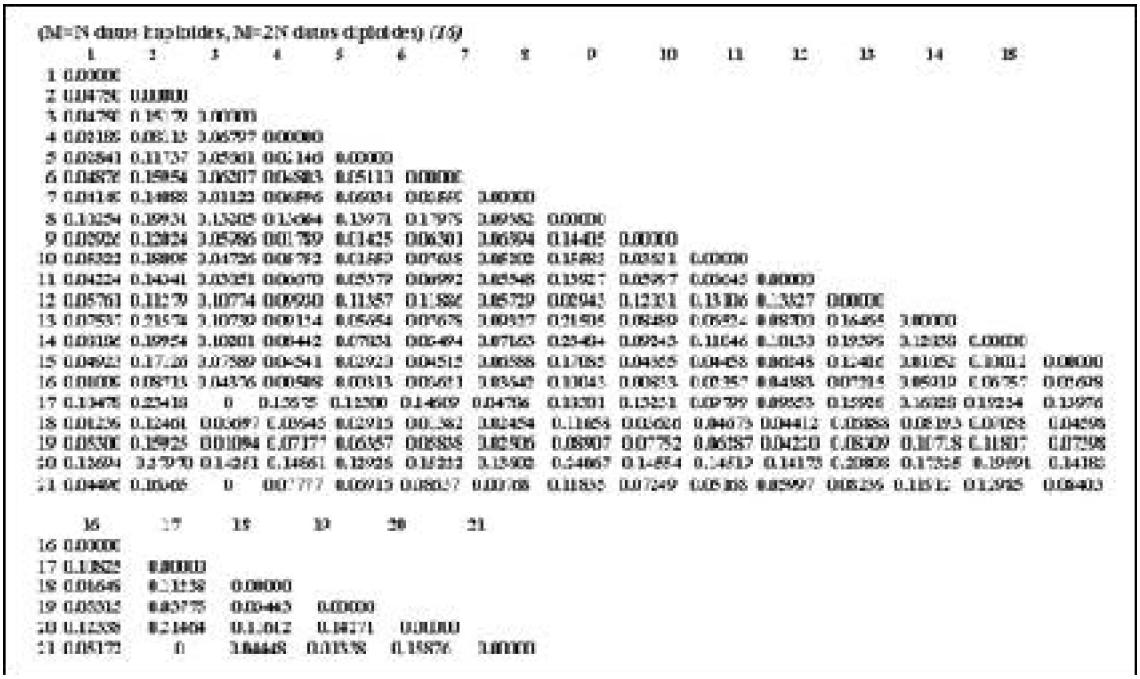

Tabla 11: Matriz de coeficientes de coancestros t/M=-ln(1-FST), que calcula las distancias poblacionales en donde estas han divergido t generaciones atrás, pero en tiempos cortos de divergencia.

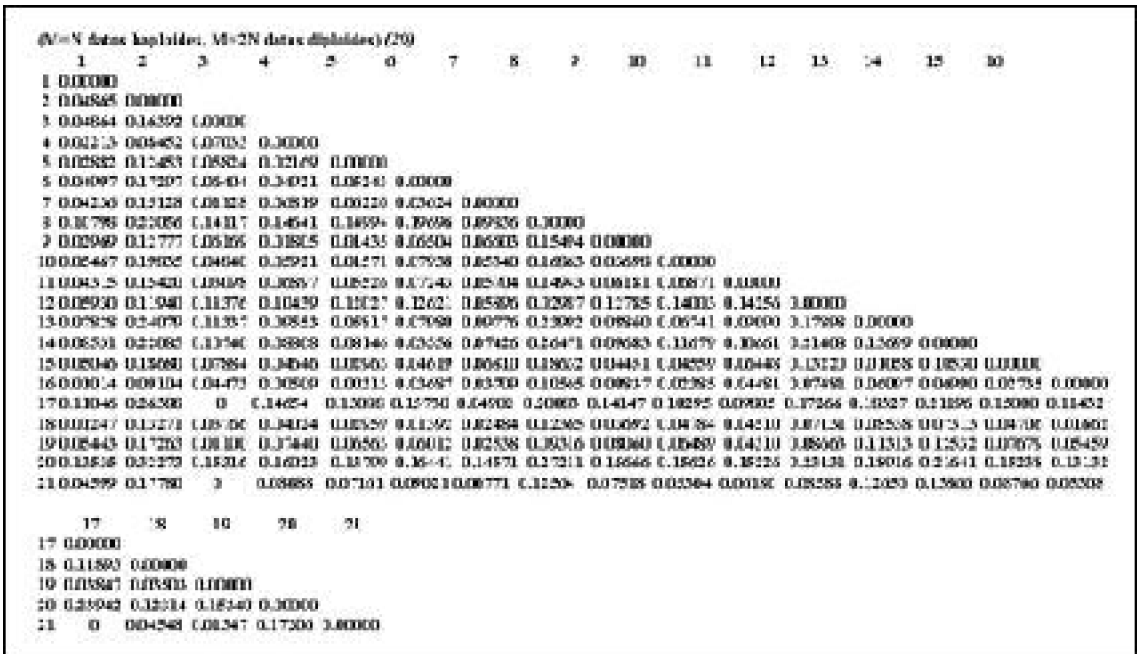

Tabla 12: Matriz de Slatkin linearizada, t/M=FST/(1-FST), considerando un modelo demográfico simple, donde dos poblaciones de tamaño N han divergido hace t generaciones desde una población de tamaño idéntico, estas poblaciones han permanecido siempre separadas sin intercambios y migraciones.

Una vez calculados estos parámetros estadísticos se procedió a utilizar el programa estadístico MEGA (11) con el cual se pudieron realizar las aproximaciones de comparación de poblaciones al generar cladogramas UPGMA con los mismos (figura 14).

"Programa Cybertesis PERÚ - Derechos son del Autor"

DISCUSIÓN

En cuanto al análisis de frecuencias alélicas, al comparar la población de Puno y la población Amerindia, según el test de diferenciación (ARLEQUÍN Software for Population Genetics Data Analysis version 2.000), se tiene que los marcadores DYS390 y DYS391 no presentan diferencias significativas (Figura 07) es decir la presencia alélica así como la distribución de frecuencias alélicas entre estas dos poblaciones estudiadas siguen un patrón similar de distribución para los alelos más frecuentes (ver figura 08 y 09); caso contrario el observado para el marcador DYS392 el cual sí presenta patrones diferentes más marcados de distribución, por tener según el test de diferenciación, un valor inferior al 5%. Esta diferencia se debe por ejemplo a que el alelo más frecuente para este marcador en la población Amerindia es el menos frecuente en la población de Puno (alelo 11) entre otras como se puede apreciar en la figura 10. Los patrones de distribución para estos casos no se ven fuertemente afectados por la diferencia en tamaños poblacionales ya que en la población de Puno, que es la de menor número poblacional estudiado, se ha encontrado un nivel de variabilidad consistente con respecto a lo publicado (tablas 04 y 09), además se puede apreciar fácilmente la consistencia de distribución alélica en los otros dos marcadores analizados (ver figuras 08 y 09); por lo tanto si aumentáramos el número de muestra es decir el tamaño poblacional de la población puneña en estudio, la probabilidad de encontrar el mismo patrón de distribución sería alta.

Según los índices de diversidad calculados (ARLEQUÍN Software for Population Genetics Data Analysis versión 2.000) para la población de Puno y otras poblaciones definidas (8 poblaciones sudamericanas y 20 poblaciones dispersas a nivel mundial), la población de Puno se encuentra en el rango promedio de diversidad esperada aceptable,

en comparación con las demás poblaciones (figura 11 y 12), definiendo esta diversidad genética calculada como la probabilidad de que dos haplotipos elegidos al azar sean diferentes en la muestra.

Además, para realizar los análisis de estructura genética con el programa estadístico ARLEQUÍN (Software for Population Genetics Data Analysis versión 2.000), especialmente análisis de comparaciones poblacionales como es el caso de cálculos de matrices Fst por pares de poblaciones según diversos modelos, a partir de la matriz de distancias Fst obtenida, se construyeron otras matrices por ejemplo la matriz de distancias de Reynold's (16) que calcula las distancias poblacionales en donde estas han divergido t generaciones atrás, pero en tiempos cortos de divergencia. Otra matriz presentada es la matriz de Fst's linearizada de Slatkin's (20) con el cual se esta considerando un modelo demográfico simple, donde dos poblaciones de tamaño N han divergido hace t generaciones desde una población de tamaño idéntico, estas poblaciones han permanecido siempre separadas sin intercambios y migraciones. Para hacer un análisis más visual de estos resultados se diseña por medio del programa MEGA (11), Cladogramas UPGMA de aproximaciones interpoblacionales, el cual asume una misma tasa de mutación para todos los linajes evolucionantes para poder generar árboles enraizados (15), tomando como base la matriz de Fst's linearizada de Slatkin's (20), tanto para comparaciones a nivel de Sudamérica (figura 13) como a nivel de poblaciones dispersas a nivel mundial (figura 14).

A nivel de Sudamérica se puede observar un encaje perfecto de la población puneña con las otras poblaciones peruanas estudiadas en la publicación de referencia la de Arequipa y la de Tayacaja, (4), que a pesar de solamente tener en el presente trabajo coincidencia con únicamente los STR DYS390 y DYS391 (ya que no se trabajó con el marcador DYS392 en dicha publicación) se puede corroborar lo expuesto por los autores en cuanto a que las frecuencias alélicas presentes en las regiones andinas son relativamente las mismas a las de la zona del oeste de Sudamérica, pero que ellas son diferencialmente distribuidas con respecto a las poblaciones del este de Sudamérica, además de tener las poblaciones andinas una gran similitud de medio ambiente y desarrollo cultural, es decir ya visto desde otro ángulo, el área geográfica ha estado involucrada en un único proceso cultural en los últimos 12000 años (4), que los ha guiado a una relativa homogeneidad lingüística y cultural comparados con las poblaciones nativas del este de Sudamérica.

A nivel mundial la población peruana de Puno estaría muy bien relacionada con poblaciones de la zona asiática (figura 14), lo que nos confirmaría una vez más lo ya sabido en cuanto a las teorías de poblamiento americano, en el cual hubo una migración de poblaciones asiáticas, apoyados por estudios de eras geológicas que habrían favorecido a tales supuestos, dándonos de esta manera una excelente correlación entre lo posiblemente acontecido en la historia con la información genética, teniendo un apoyo cada vez más científico para la comprensión de los orígenes e historia de las poblaciones humanas especialmente historia evolutiva.

"Programa Cybertesis PERÚ - Derechos son del Autor"

CONCLUSIONES

1. Se ha logrado establecer una base de datos de frecuencias alélicas para los marcadores microsatélites del cromosoma Y estudiados, luego de haber estandarizado las metodologías aplicadas y determinado las variaciones alélicas existentes en la población de Puno. Cumpliendo con los objetivos planteados en el presente trabajo.

2. Según el análisis de frecuencias alélicas el marcador DYS392, al compararlo con la base de datos Amerindia, indica que esta base de datos no refleja el pool génico de la población estudiada.

3. Al comparar las poblaciones sudamericanas con la población de Puno, en cuanto a frecuencias haplotípicas, Puno está diferenciada como población alto andina manteniendo similitud con la población de Arequipa que geográfica y culturalmente tienen similares condiciones.

4. Con respecto a las poblaciones nativas brasileras existe un nivel de diferenciación con la población de Puno en cuanto a sus frecuencias haplotípicas.

5. A través del análisis de frecuencias haplotípicas, la población de Puno mantiene una similitud con poblaciones asiáticas como es el caso de China y Taiwán, con lo cual se estaría aportando a teorías de poblamiento americano, orígenes e historia evolutiva humana.

6. Se ha logrado esquematizar correctamente lo esperado con un número reducido de marcadores, en contraste con los trabajos que utilizaron un número mayor de marcadores.

7. Los marcadores STR de cromosoma Y demuestran una vez más el buen grado de efectividad para la construcción de genealogías informativas para estudios poblacionales, debido al tipo de segregación y la no recombinación de estos.

"Programa Cybertesis PERÚ - Derechos son del Autor"

RECOMENDACIONES

1. Por los resultados expuestos en el presente trabajo de investigaciones es recomendable el diseño y ejecución de trabajos relacionados a la genética de poblaciones humanas especialmente en poblaciones sudamericanas y de manera mas precisa en poblaciones peruanas, tanto para reflejar la identidad cultural como para todas las aplicaciones que inciden de manera beneficiosa a nuestras poblaciones, ya que actualmente se han realizado escasos trabajos relacionados al tema, teniendo de esta manera vacíos de información.

2. Metodológicamente es siempre recomendable trabajar con una mayor cantidad de marcadores, tratando de tipificar el haplotipo mínimo requerido y también los diversos marcadores diseñados para el cromosoma Y, identificando luego cuáles de estos vendrían a ser los más informativos para nuestra población.

"Programa Cybertesis PERÚ - Derechos son del Autor"

ILUSTRACIONES

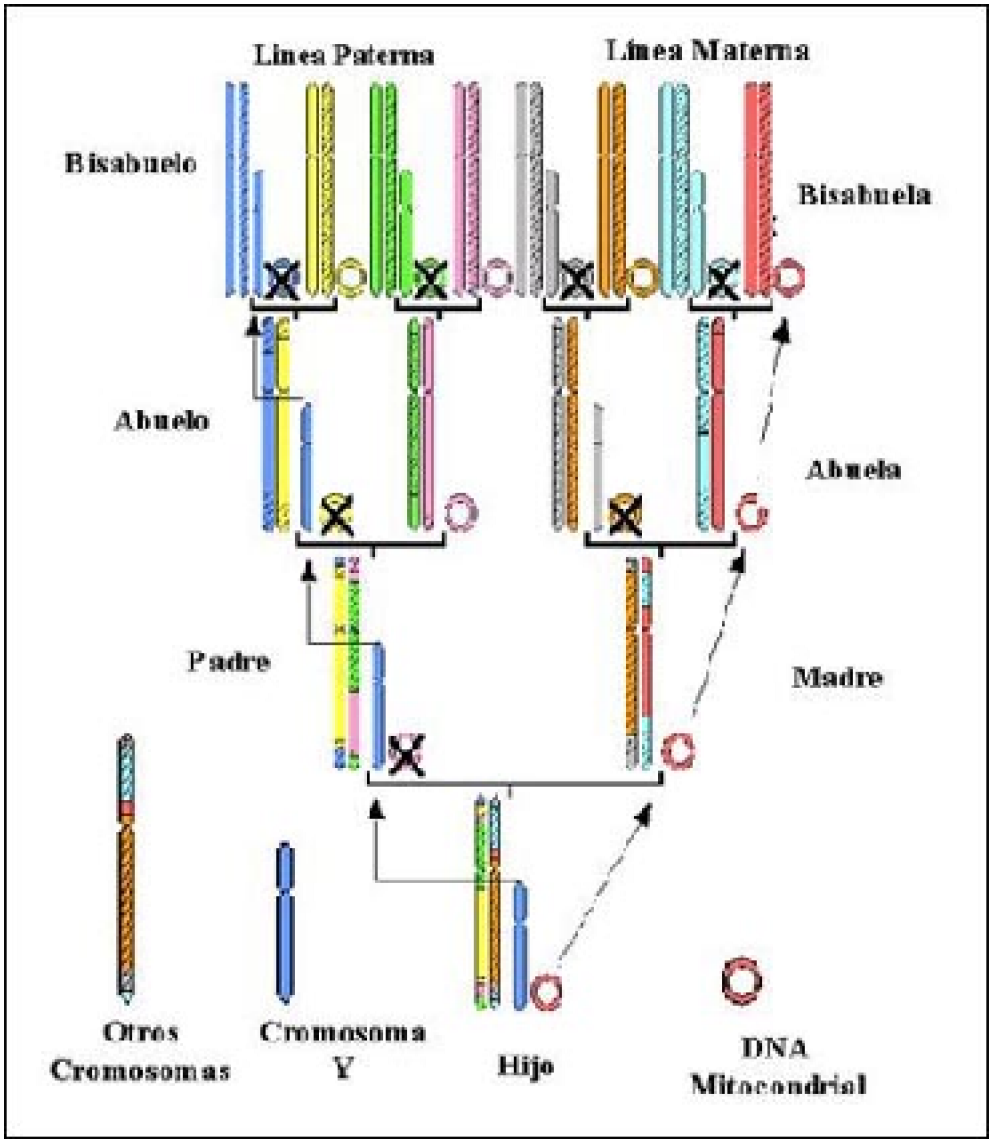

FIGURA 01: Esquema de segregación de los distintos cromosomas incluyendo el mitocondrial, Leonor Gusmao, Marcadores del cromosoma Y Aspectos Generales y Aplicaciones (comunicación personal, 2003)

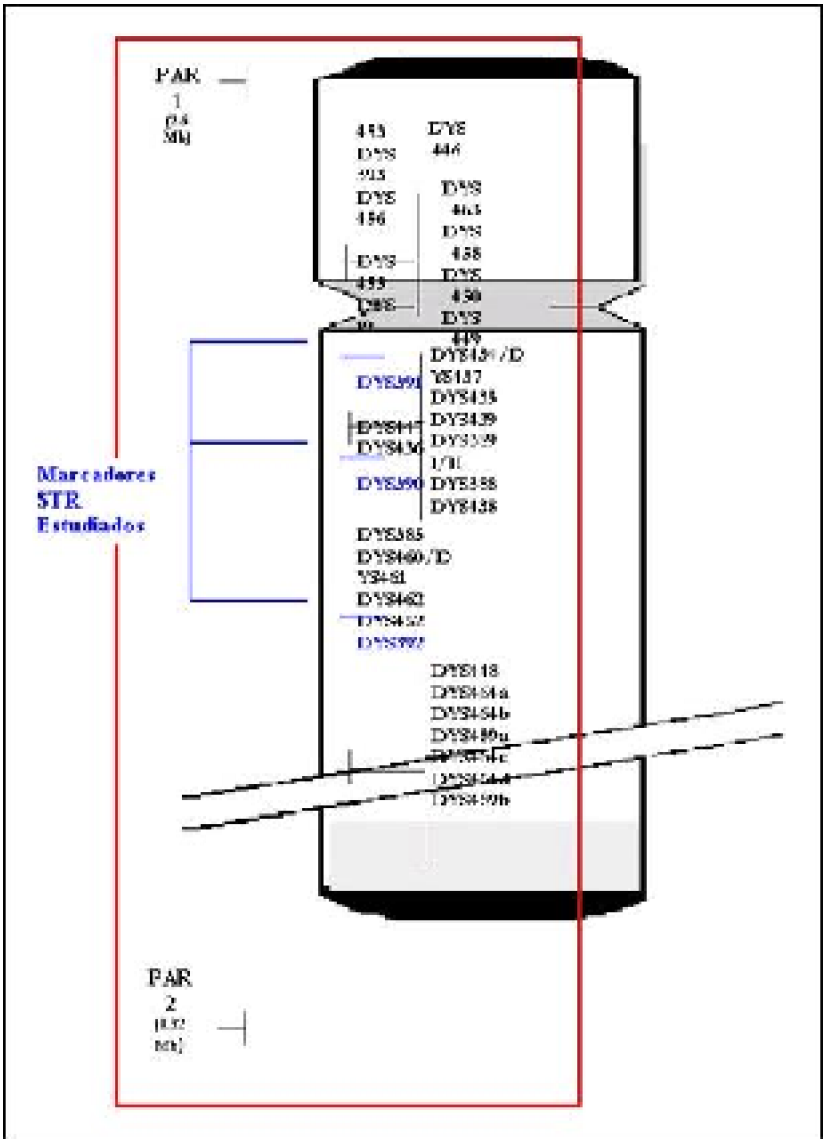

FIGURA 02: Esquema de los Diversos Marcadores STR del Cromosoma Y Leonor Gusmao, Marcadores del cromosoma Y Aspectos Generales y Aplicaciones (comunicación personal, 2003)

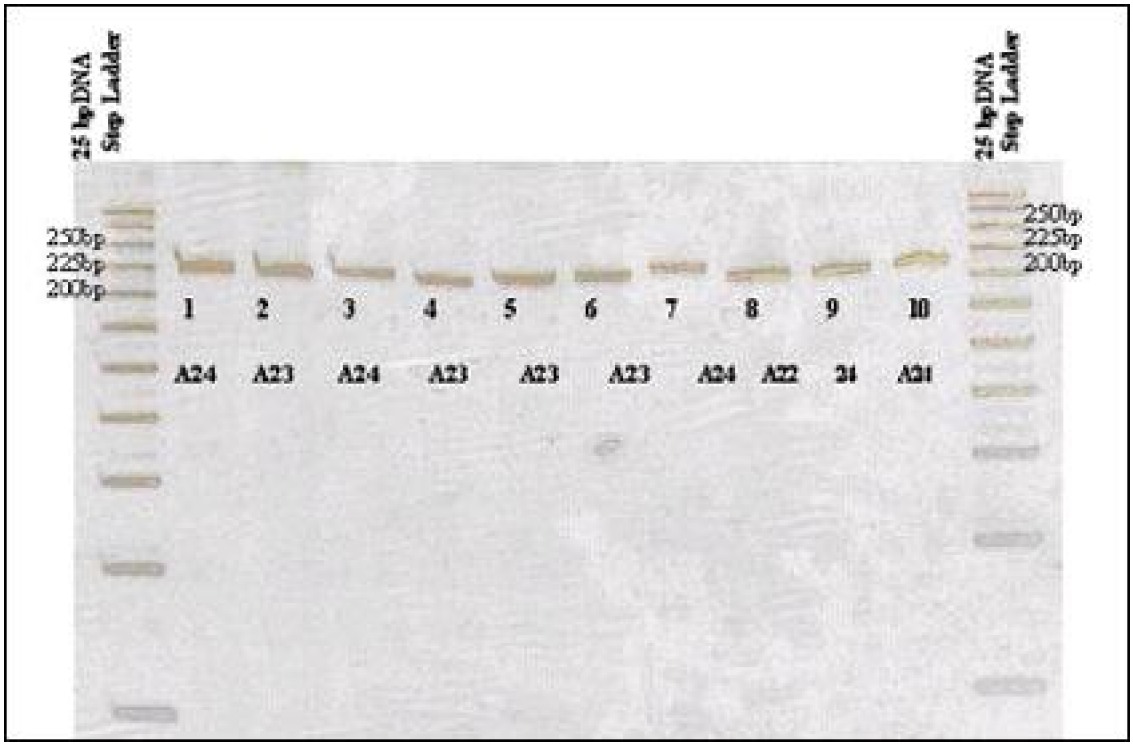

FIGURA 03: Resultados, PAGE - 7M UREA 10% DYS390 en donde se aprecia las diferencias de migración poco marcada de los alelos A22, A23 y A24 en las muestras numeradas del 1 al 10.

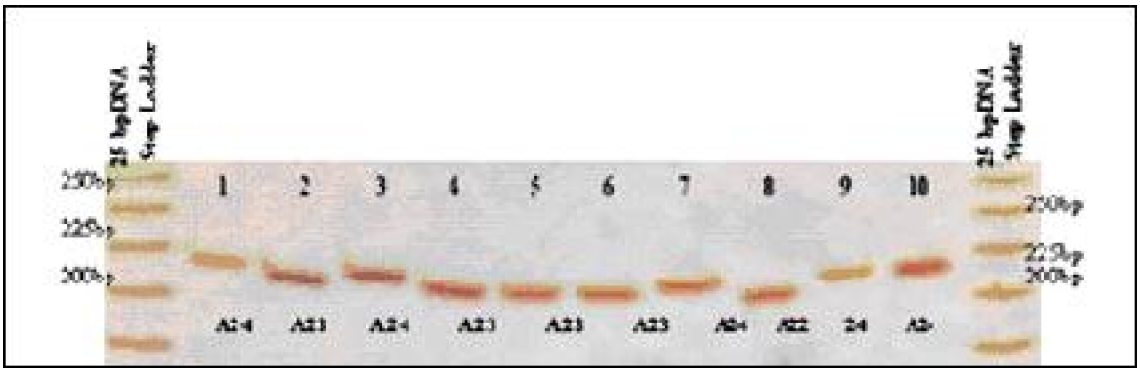

FIGURA 04: Resultados, PAGE - 7M UREA 12% DYS390 en donde se aprecia mejor las diferencias de migración de los alelos A22, A23 y A24 en las muestras numeradas del 1 al 10.

FIGURA 05: Ubicación de las poblaciones sudamericanas utilizadas para los estudios comparativos con la población de Puno. Eduardo Tarazona-Santos et al.: Genetic Differentiation in South Amerindians Is Related to Environmental and Cultural Diversity: Evidence from the Y Chromosome.

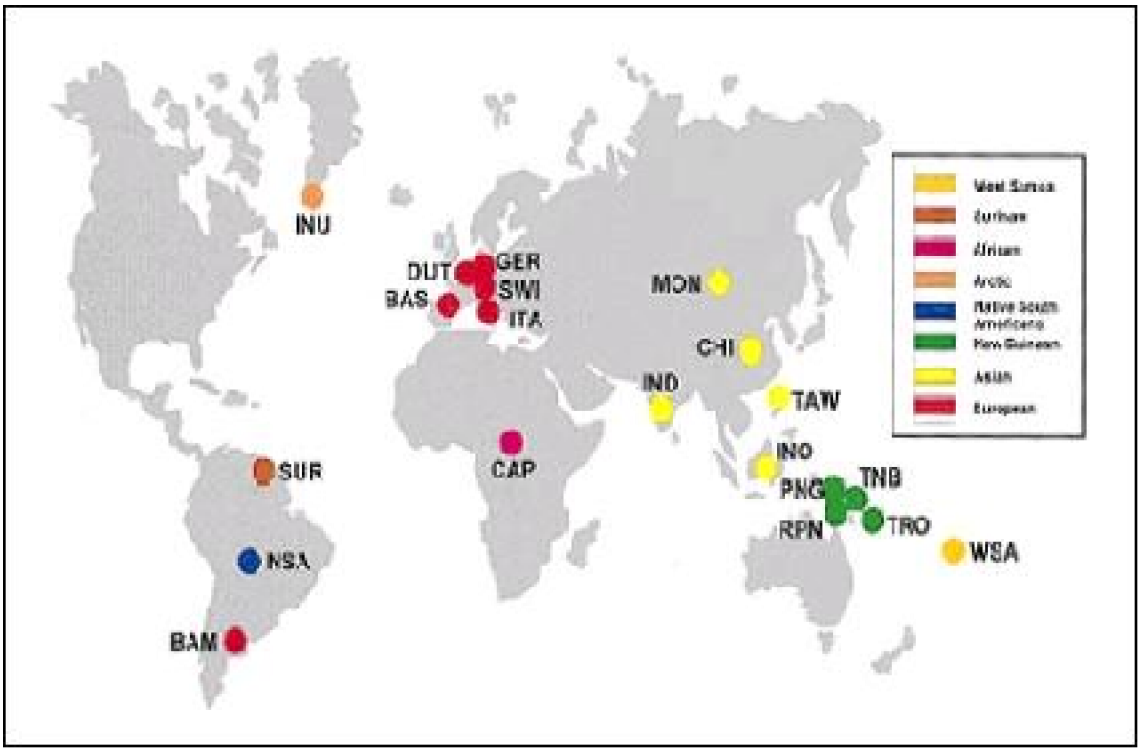

FIGURA 06: Ubicación de poblaciones mundiales utilizadas para los estudios comparativos con la población de Puno. Kayser et al.: Globally Dispersed Y Microsatellite Haplotypes.

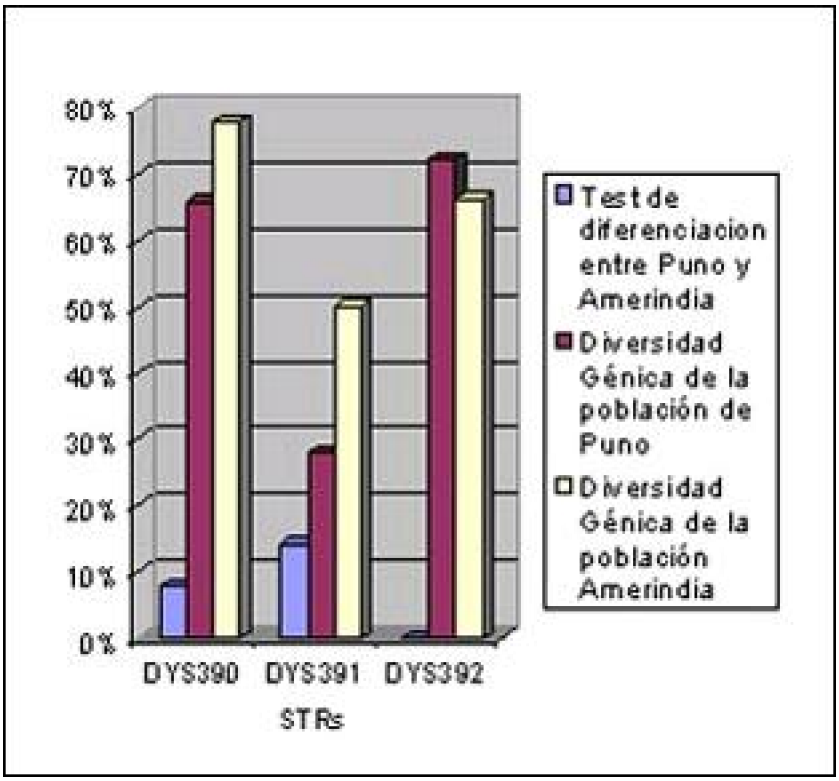

FIGURA 07: Gráfico Comparativo entre Población Amerindia y Población de Puno, observándose diferenciación inferior al 5% entre las poblaciones y patrón similar de distribución de diversidad génica.

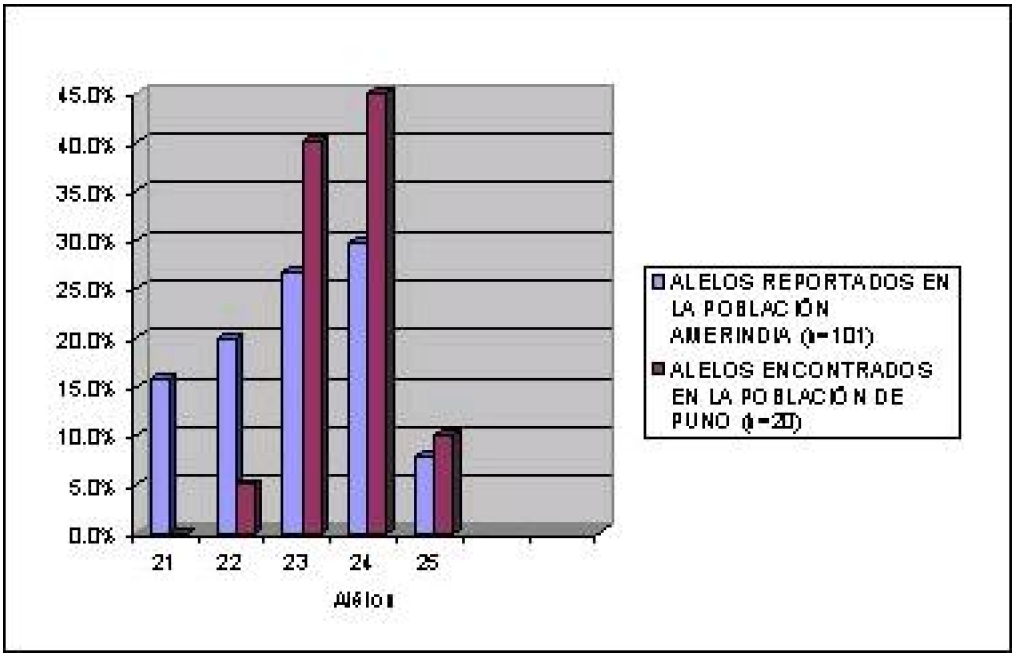

FIGURA 08: Gráfica de Distribución de Frecuencias Alélicas del STR DYS390 Población de Puno vs. Población Amerindia, observándose patrones similares de distribución.

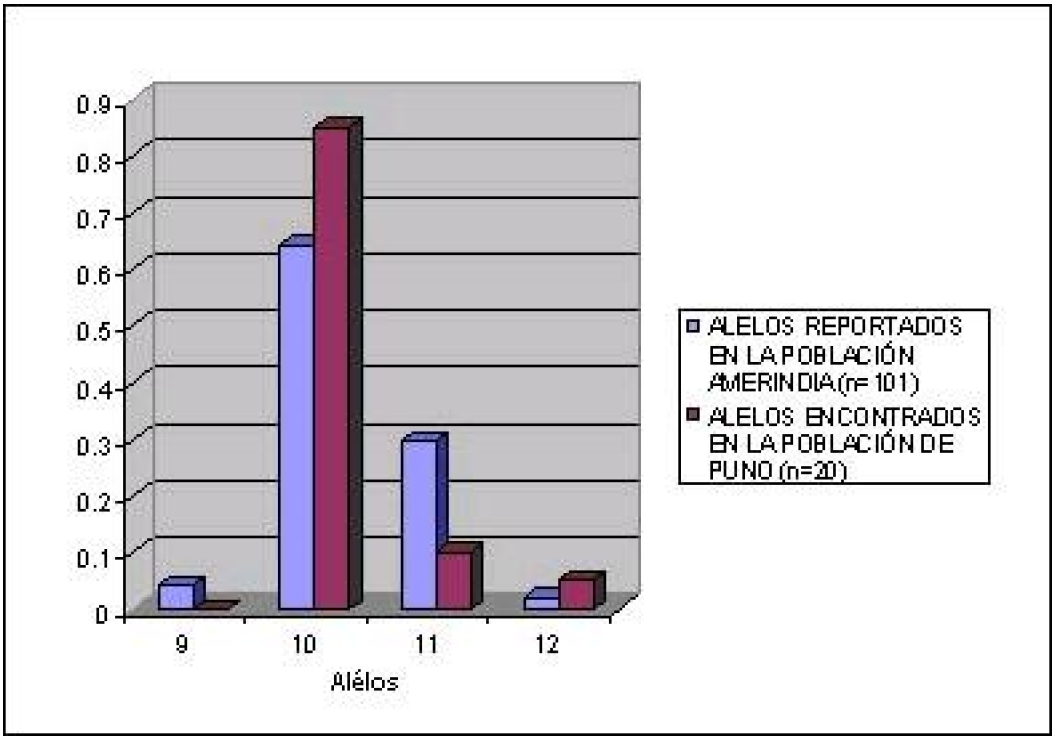

FIGURA 09: Grfica de Distribucin de Frecuencias Allicas del STR DYS391 Poblacin de Puno vs. Poblacin Amerindia, observndose patrones similares de distribucin.

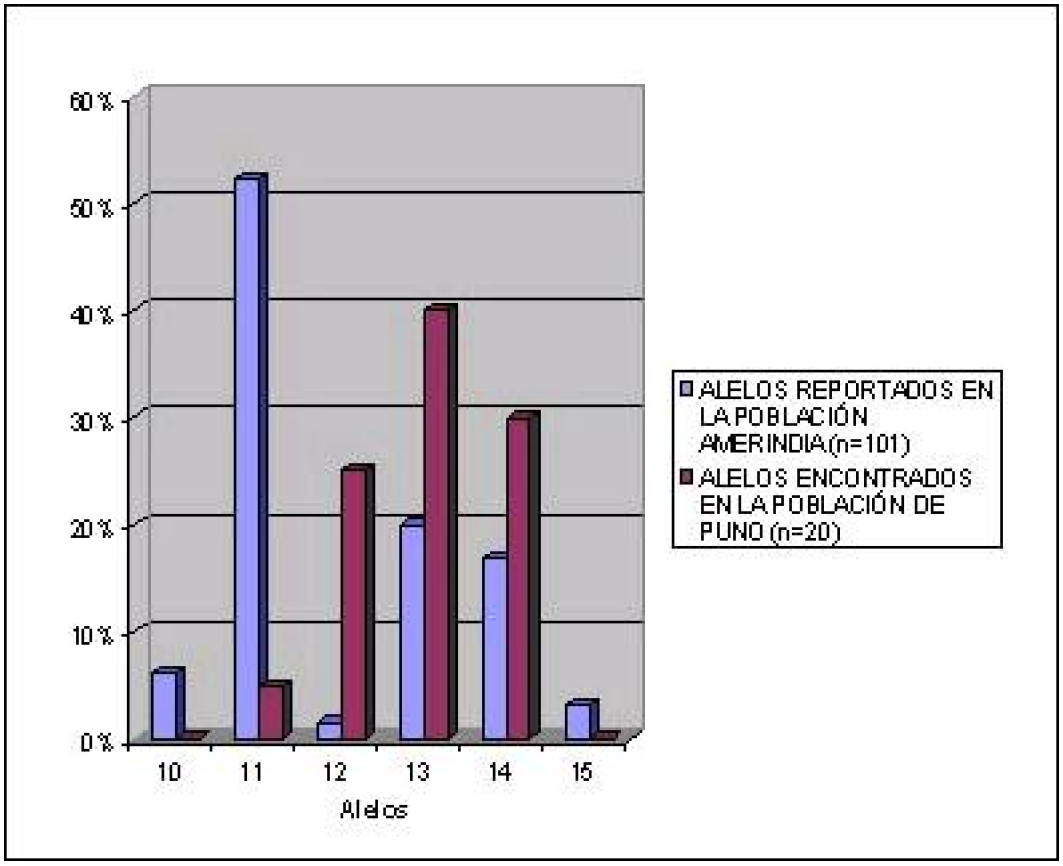

FIGURA 10: Gráfica de Distribución de Frecuencias Alélicas del STR DYS392 Población

de Puno vs. Población Amerindia observándose patrones diferentes de distribución.

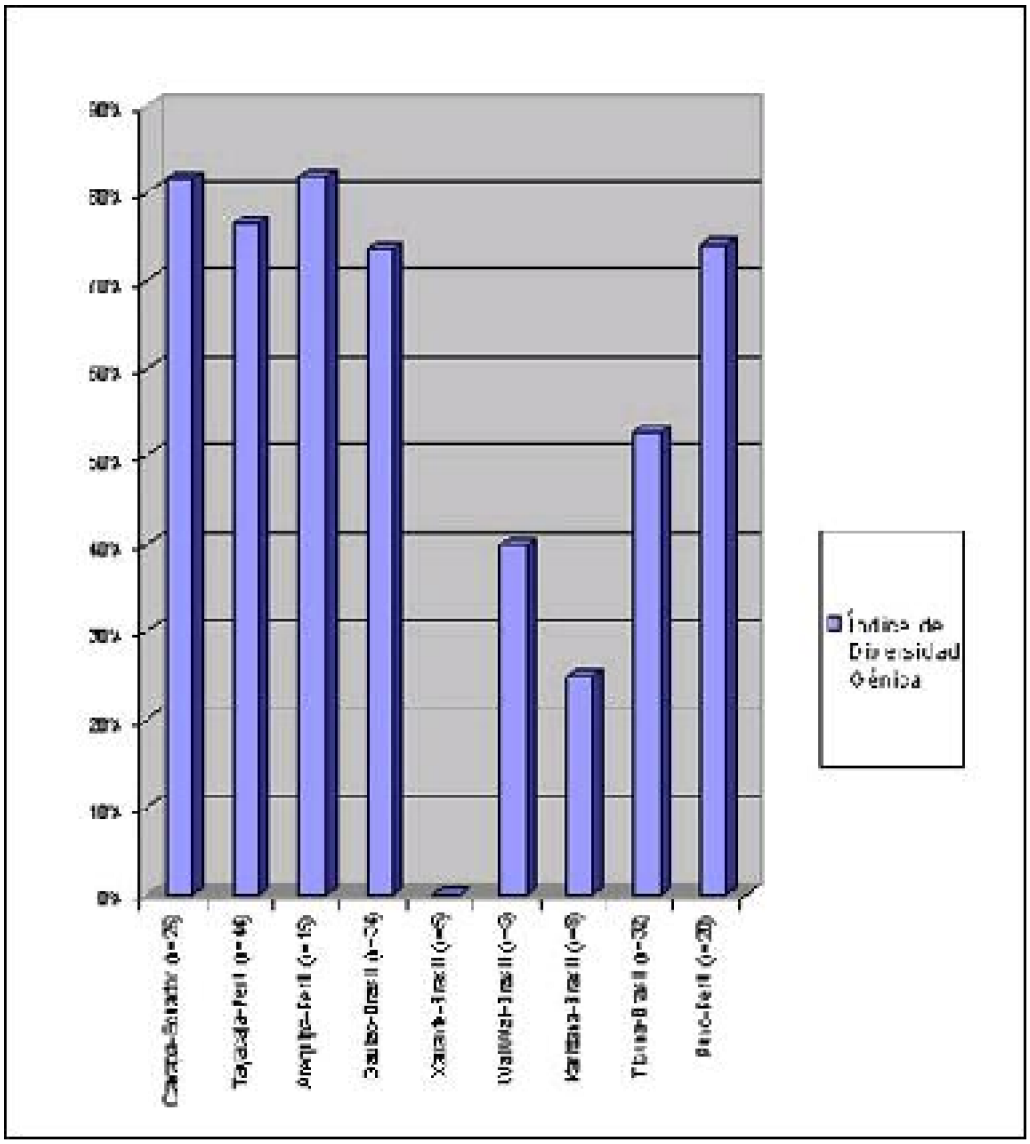

FIGURA 11: Índice de Diversidad Génica Población de Puno vs. Poblaciones Sudamericanas, observándose nivel aceptable de diversidad génica de Puno según tamaños poblacionales.

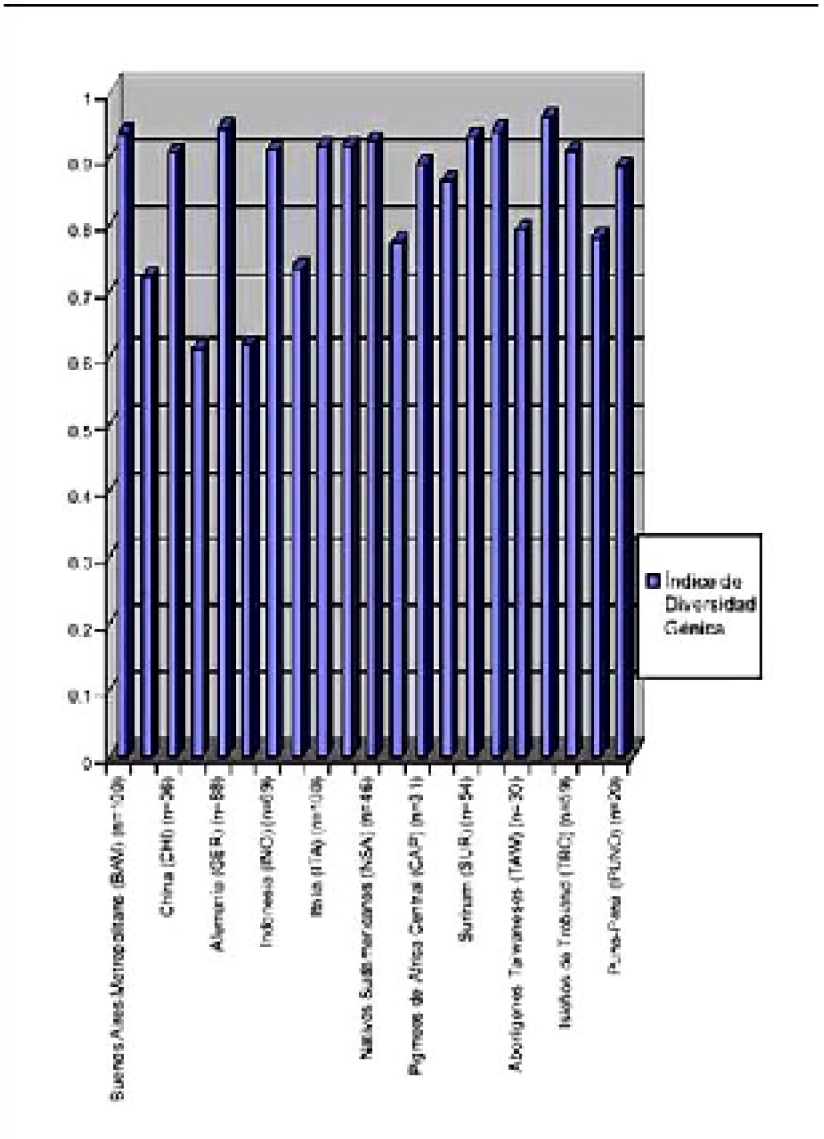

FIGURA 12: Índice de Diversidad GénicaPoblación de Puno vs. Poblaciones Mundiales, observándose nivel aceptable de diversidad génica de Puno según tamaños poblacionales.

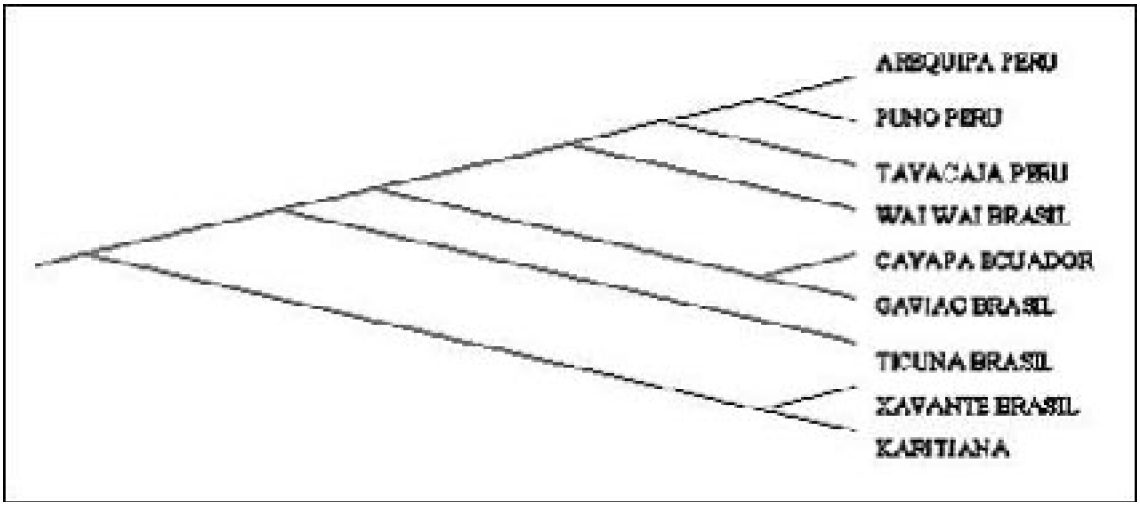

FIGURA 13: Arbol UPGMA Población de Puno vs. Poblaciones Sudamericanas, observándose mayor cercanía genética de la población de Puno con las poblaciones peruanas de Arequipa y Tayacaja. Arequipa (n = 15), Puno (n = 20), Tayacaja (n = 44), Wai Wai (n = 5), Cayapa (n = 26), Gaviao (n = 34), Ticuna (n = 32), Xavante (n = 5), Karitiana (n = 8).

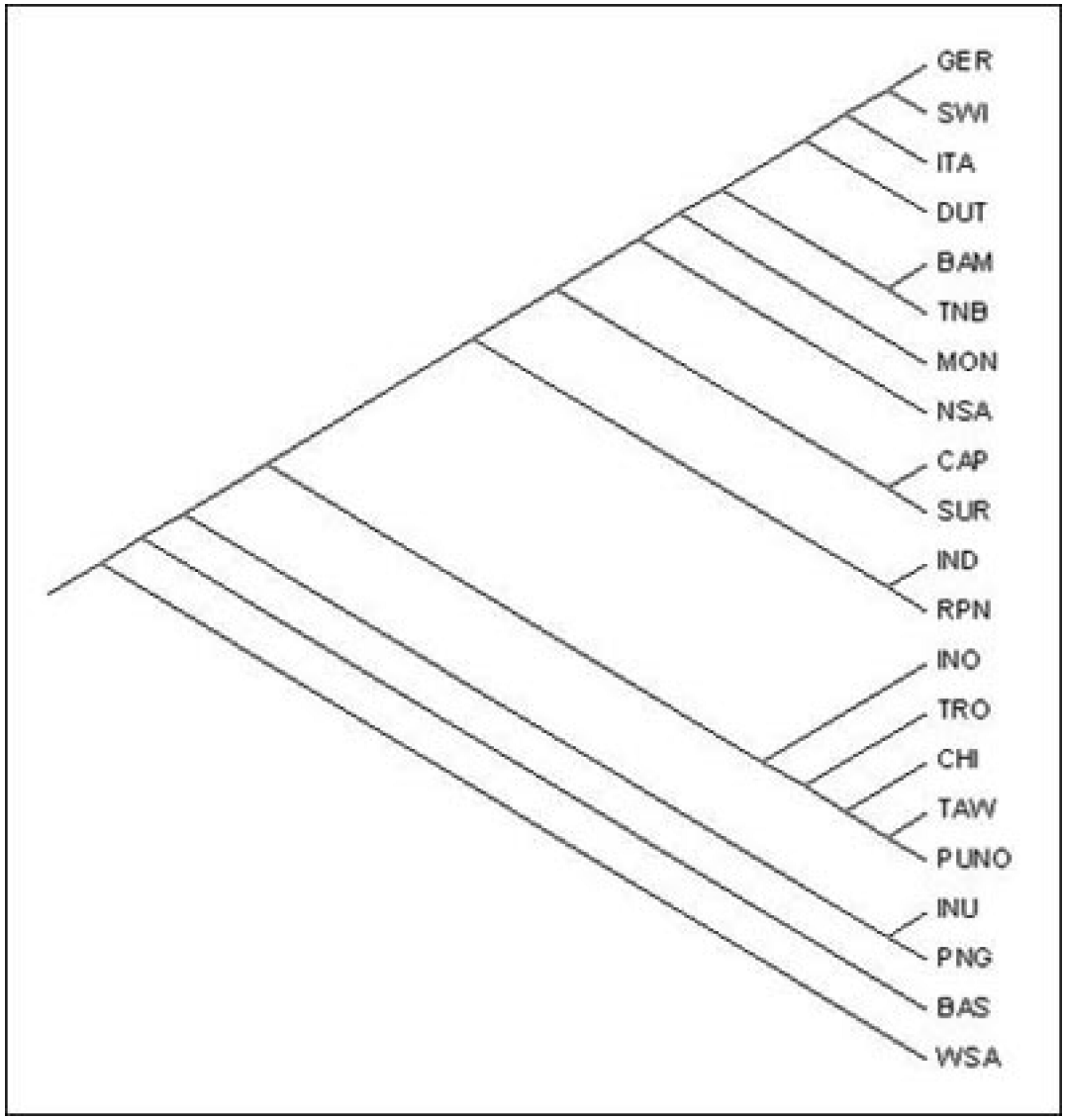

FIGURA 14: Arbol UPGMA Población de Puno vs. Poblaciones Mundiales, observándose mayor cercanía genética de la población de Puno con las poblacionesasiáticas de Taiwán y China. Buenos Aires Metropolitans (BAM) (n=100), Bascos (BAS) (n=30), China (CHI) (n=36), Dutch (DUT) (n=88), Alemania (GER) (n=88), India (IND) (n=25), Indonesia (INO) (n=69), Artico (INU) (n=62), Italia (ITA) (n=100), Mongolia (MON) (n=40), Nativos Sudamericanos (NSA) (n=46), Papua Nueva Guinea (PNG) (n=26), Pigmeos de Africa Central (CAP) (n=31), Roro de la Costa Sur de Nueva Guinea (RPN) (n=12), Surinam (SUR) (n=54), Suiza (SWI) (n=64), Aborígenes Taiwaneses (TAW) (n=30), Tolai de Nueva Bretaña (TNB) (n=16), Isleños de Trobiand (TRO) (n=59), Samoa (WSA) (n=10), Puno-Perú (PUNO) (n=20).

"Programa Cybertesis PERÚ - Derechos son del Autor"

BIBLIOGRAFÍA

Barton E. and Albrigth L., (1994), L. DNA sequencing.Current Protocols in Molecular Biology Massachusetts General Hospital & Harvard Medical School; I,7.6.1-7.6.13.

Bosch E, Lee AC, Calafell F, Arroyo E, Henneman P, de Knijff P, Jobling MA. (2002). High resolution Y chromosome typing: 19 STRs amplified in three multiplex reactions. Forensic Science International. 125:42:51.

Butler JM, Schoske R, Vallone PM, Kline MC, Redd AJ, Hammer MF. (2002). A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers. Forensic Science International 129 (1):10-24.

De Knijff P, Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Piccinini A, Perez-Lezaun A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, Roewer L. (1997). Chromosome Y microsatellites: population genetic and evolutionary aspects. *International Journal of Legal Medicine* ; 110: 134-149.

González-Neira A, Elmoznino M, Lareu MV, Sánchez-Diz P, Gusmao L, Prinz M, Carracedo A. (2001). Sequence structure of 12 novel Y chromosome microsatellites and PCR amplification strategies. Forensic Science International 122:19-26.

Gusmão L, Alves C, Amorin A.(2001). Molecular characterisation of four human Y-specific microsatellites (DYS434, DYS437, DYS438, DYS439) for population and forensic studies. American Journal of Human Genetics. 65:285-291.

Gusmão L, Alves C. Beleza S. Amorim A.(2002). Forensic evaluation and population data on the new YSTRs DYS434, DYS437, DYS438, DYS439 and GATA A10. *International Journal of Legal Medicine* ,116:139-147.

Henke J., Henke L., Chatthopadhyay P., Kayser M., Dülmer M., Cleef S., Pöche H., Felske-Zech H.,(2001), Application of Y chromosomal STR Haplotypes to forensic genetics, Croatian Medical Journal of Legal Medicine, 42(3):292-297.

Izagirre N., Alonso S., De la Rúa C., (2001). DNA Analysis and the evolutionary history of the Basque Population: a review. Journal of Anthropology Research, vol. 57.

Kayser M., Krawczak M., Exccoffier L., Dieltjes P., D. Corach, V. Pascali, C. Gehrig, L.F. Bernini, J. Jespersen, E. Bakker, L. Roewer, P, de Knijff, (2001) An Extensive Analysis of Y-Chromosomal Microsatellite Haplotypes in Globally Dispersed Human Populations, American Journal of Human Genetics. 68:990-1018.

Kumar S., Tamura K.,. Jakobsen I., and Nei M. (2001) MEGA2: Molecular Evolutionary Genetics Analysis Software, Arizona State University, Arizona, USA.

Marmur, J. (1961). A procedure for the isolation of deoxyribonucleic acid from microorganisms. Journal of Molecular Biology. 3: 208-218.

Medina A. (2003), Tesis para optar el Grado de Magíster: Distribución de los alelos del triplete repetido del gen de la miotonin kinasa en muestras de pobladores de los departamentos de Puno y Lima, UNMSM, Perú.

Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E, et al (Mar 27,1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235(4796):1616-22.

Nei M., Kumar S. (2000), Molecular evolution and phylogenetics, Oxford Univ. Press.

Reynolds, J., Weir, B.S., and Cockerham, C.C. (1983).Estimation for the coancestry coefficient: basis for a short term genetic distance. Genetics 105:767-779.

Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Corte-Real HB, de Knijff P, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Golge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskas V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar AV, Mitchell RJ, Nafa K, Nicholson J, Norby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previdere C, Roewer L, Rootsi S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Villems R, Tyler-Smith C, Jobling MA (2000).Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. American Journal of Human Genetics. Dec;67(6):1526-43. Epub 2000 Nov 09

Sajantila A., Salem A-H, Savolainen P., Bauer K., Gierig C., and Pääbo S. (1996), Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. Proceedings of the National Academy of Science of the United States of America, Vol. 93, Issue 21, 12035-12039, October 15.

Santos F. R., Tyler-Smith C., (1996) Reading the human Y chromosome: the emerging DNA markers and human genetic history. Brazilian Journal of Genetics.19:665-670.

Statkin,M.(1995) A measure of population subdivision based on microsatellite allele

frecuencies. Genetics 139: 457-462.

Tarazona-Santos E., Carvalho-Silva D., Pettener D., Luiselli D., De Stefano G. F., Martinez C., Rickards O., Tyler-Smith C., Pena S. D. J. and Santos F. R. (2001). Genetic Diferentiation in South Amerindians Is Related to Evironmental and Cultural Diversity: Evidence from the Y chromosome. American Journal of Human Genetics. 68:1485-1496.

Wells R., Yuldasheva N., Ruzibakiev R., Underhill P, Evseeva I., Blue-Smith J., Jin L., Su B., Pitchappan R., Sadagopal Shanmugalakshmi, Karuppiah Balakrishnan, Mark Read, Pearson N., Zerjal T., Webster M., Zholoshvili I., Jamarjashvili E., Gambarov S., Nikbin B., Dostiev A., Aknazarov O., Zalloua P., Tsoy I., Kitaev M., Mirrakhimov M., Chariev A., and Bodmera W.(2001). The Eurasian Heartland: A continental perspective on Y-chromosome diversity. Proceedings of the National Academy of Science of the United States of America, 98(18): 10244-10249.

Zerjal, T., Dashnyam, B., Pandya, A., Kayser, M., Roewer, L., Santos, F. R., Schiefenhövel, W., Fretwell, N., Jobling, M. A., Harihara, S., (1997). Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis. American Journal of Human Genetics. 60, 1174-1183.

"Programa Cybertesis PERÚ - Derechos son del Autor"

ANEXOS

TABLA 12: Secuencia de Primers Utilizados; Genbank Accession G09611(DYS390), G09613 (DYS391), G09867 (DYS392)

MARCADOR STR	ORIENTACIÓN	OLIGONUCLEÓTIDO
DYS390	F (Forward	5´TAT ATT TTA CAC ATT TTT GGG CC 3´
	R (Reverse)	5´TGA CAG TAA AAT GAA CAC ATT GC 3´
DYS391	F (Forward	5´CTATTCATTCAATCATACACCCA3´
	R (Reverse)	5´GATTCTTTGTGGTGGGTCTG 3´
DYS392	F (Forward	5´TCATTAATCTAGCTTTTAAAAACAA3´
	R (Reverse)	5´AGACCCAGTTGATGCAATGT3´

TABLA 13: Tamaños de los Productos de PCR y los Respectivos Alelos para el STR DYS390 Genbank Accession G09611, AC011289.

Alelos Repeat #	Fragmento Amplificado	Estructura Repetida
18	191 bp	No reportado
19	195 bp	No reportado
20	199 bp	No reportado
21	203 bp	$[TCTG]_8[TCTA]_5ACTA[TCTA]_2$ $[TCTG]1\ [TCTA]_4$
22	207 bp	$TCTG]_8\ [TCTA]_9[TCTG]_1[TCTA]_4$
23	211 bp	$[TCTG]_8[TCTA]_{10}[TCTG]_1[TCTA]_4$
24	215 bp	$[TCTG]_8\ [TCTA]_{11}[TCTG]_1[TCTA]_4$
25	219 bp	$[TCTG]_8\ [TCTA]_{12}[TCTG]_1[TCTA]_4$
26	223 bp	
27	227 bp	$[TCTG]_8\ [TCTA]_{14}\ [TCTG]_1\ [TCTA]_4$

TABLA 14: Tamaños de los Productos de PCR y los Respectivos Alelos para el STR DYS391 Genbank Accession G09613.

Alelos Repeat #	Fragmento Amplificado	Estructura Repetida
6	268 bp	$[TCTA]_6$
7	272 bp	$[TCTA]_7$
8	276 bp	$[TCTA]_8$
9	280 bp	$[TCTA]_9$
10	284 bp	$[TCTA]_{10}$
11	288 bp	$[TCTA]_{11}$
12	292 bp	$[TCTA]_{12}$
13	296 bp	$[TCTA]_{13}$
14	300 bp	$[TCTA]_{14}$

TABLA 15: Tamaños de los Productos de PCR y los Respectivos Alelos para el STR DYS392. Genbank Accession G09867.

Alelos Repeat #	Fragmento Amplificado	Estructura Repetida
6	234 bp	$[TAT]_6$
7	237 bp	$[TAT]_7$
8	240 bp	$[TAT]_8$
10	246 bp	$[TAT]_{10}$
11	249 bp	$[TAT]_{11}$
12	252 bp	$[TAT]_{12}$
13	255 bp	$[TAT]_{13}$
14	258 bp	$[TAT]_{14}$
15	261 bp	$[TAT]_{15}$
16	264 bp	$[TAT]_{16}$
17	267 bp	$[TAT]_{17}$

TABLA 16: Puno vs. Poblaciones Mundiales -Cuadro de ocurrencia haplotípica diseñada a partir de lo publicado Kayser et al.: Globally Dispersed Y Microsatellite Haplotypes.

Haplot	DYS39	BA	BA	CH	DU	GE	INI	IN	INU	ITA	MO	NS	PN	CA	RP	SU	SV	TA	TN	TR	WSA	PU	Total
1	19.10.1	1																	1				1
2	19.10.1	2																			1		1
3	20.09.1	1																		3			3
4	20.10.1	1											1							2			3
5	20.10.1	2																			5		5
6	21.09.1	2						2															2
7	21.09.1	4																		1			1
8	21.10.1	1			1		2	2		1		1	2	8		12	2						32
9	21.10.1	2																			1		1
10	21.10.1	3													1								1
11	21.10.1	4		1															1				2
12	21.11.1	1						4				1		2		2	1						11
13	22.08.1	1												1									1
14	22.09.1	1																					1
15	22.09.1	2								1													1
16	22.10.1	0														1							1
17	22.10.1	3			15	5	2		2	10	2			1	1	3	7						51
18	22.10.1	2				1					1												2
19	22.10.1	4						1		1						2				2		1	11
20	22.10.1	4		2				2				9											14
21	22.10.1	4						1															5
22	22.10.1	6								1													1
23	22.11.1	1				1				2						1							5
24	22.11.1	3			1				1							2							4
25	22.11.1	5										1											1
26	23.09.1	2			1	2				2	1			2									10
27	23.09.1	3						1		1			1										3
28	23.09.1	4										1						1					2
29	23.10.0	7				1																	1
30	23.10.1	3					1	1								3							8
31	23.10.1	4	2	2	8	10			2	18	3			1		3	4						57
32	23.10.1	2		1	4	3		2	1	6						1	4	1				1	25
33	23.10.1	5		2	4	4	4	1		4		2	1			3	3	1	2	4		1	41
34	23.10.1	4		10				11		1	4	4						13		11		6	61
35	23.10.1	5		1								1							1	2			6
36	23.10.1	2						1															3
37	23.11.1	4								1		1		6		2	1	1					16
38	23.11.1	2								1													1
39	23.11.1	3	3		14	6			1	4			1			4	3			2			41
40	23.11.1	4		3		1				1	1	2					2	3			1		14
41	23.11.1	6									1												1
42	23.12.1	3			3	1											1						5
43	23.12.1	4										1											1

Haplot	DYS39	BA	BA	CH	DU	GE	INI	INC	INU	ITA	MC	NS	PN	CA	RP	SU	SW	TA	TN	TR	WSA	PU	Total
44	24.09.12	2			2		1	1			5			2									13
45	24.09.12	2				1																	1
46	24.09.13	3	1					1							1								3
47	24.09.14	4											1			1							2
48	24.09.16	6																	1		1		2
49	24.10.10	0																					1
50	24.10.14	4		1	5	6		3	1	17	6	2				3	6		1			1	56
51	24.10.12	2			1																	2	5
52	24.10.13	32	5	4	6	3	1	14	29	5	3	1	12			1	7	5	2	9	1	4	124
53	24.10.14	4		2	1	1			14			6					2	1		8			35
54	24.10.15	5				1		1	1			1					2	1		1			10
55	24.10.16	6																	3				4
56	24.11.15	5				3			1	3	1								1				14
57	24.11.12	2				1																	1
58	24.11.13	20	15	2	13	6	1	2	4	9		4	4		1		8	1	1			1	92
59	24.11.14	4				1			3	4		5				1				2			16
60	24.11.15	5										3											6
61	24.12.11	1								1						1							2
62	24.12.13	3	1			2												1				1	7
63	25.09.11	1														1							1
64	25.09.13	3						2								1							3
65	25.10.11	1			1	11					8			3		5	5						33
66	25.10.12	2						1						1								1	3
67	25.10.13	3	3	1	3	1	4	4					2						1	6			26
68	25.10.14	4																1		5			7
69	25.10.15	5		1																			2
70	25.11.10	0							1														2
71	25.11.11	1				10			1	3	2			3		1	3						23
72	25.11.12	2		1																		1	2
73	25.11.13	3		2	2	1	5	10		2					3		3		1				29
74	25.11.14	4									1		1										2
75	25.11.15	5																		1			1
76	25.12.11	1												1									1
77	26.10.11	1			1	3																	4
78	26.10.12	2								1													1
79	26.10.13	3			1		2																3
80	26.11.11	1				1	1								1								3
81	26.11.13	3				1	1	1							4								7
82	27.11.11	1									1												1
83	27.11.13	3			1																		1
		100	30	36	88	88	25	69	62	100	40	46	26	31	12	54	64	30	16	59	10	20	

Haplot	DYS39	BA	BA	CH	DU	GE	INI	IN	INU	ITA	MO	NS	PN	CA	RP	SU	SW	TA	TN	TR	WSA	PU	Total
	DYS390	BAMBA	BAMBA	CHILCA	DUNAS	GEORG	INDIN	INDIN	INUCA	ITAPA	MONTA	NSA	PNCA	CARPA	RPNS	SUPE	SWITA	TAWIN	TNBTR	TRUJ	WSA	PUNO	
Total		100	30	36	88	88	25	69	62	100	40	46	26	31	12	54	64	30	16	59	10	20	

GDB Locus name: DYS390

Genbank Accession G09611
Alias CHLC.GATA31E10.P14963.
GDB: G00-365-248, accession id GDB:365248
Consensus sequence structure: Allele DYS390-27 (8-14-1-4), 227 bp.

5 tatatttacacatttttgggccctgcattttggtacccccataatatattctatc
tatctgtctgtctgtctgtctgtctgtctgtctatctatctatctatctatctatct
ctatctatctatctatctatctatctatctgtctatctatctatctatcatctatctatcttt
ccttgtttctgagtatacac att gcaatgtgttcattttactgtca 3'

CUADRO 1. Acceso al Banco de Genes y Secuencia de Consenso del Marcador DYS390.

GDB Locus name: DYS391

Genbank Accession G09513
Alias CHLC.GATA32C10.P14988
GDB: GDB: G00-355-251, accession id GDB:365251
Consensus sequence structure: Allele DYS391-11, 288 bp.

5.. 1 ..et attcattcaa tcctacacce atatctgtct gtctgtctat
61 ctatctatct atctatctat ctatctatct atctatctgc ctatctgcct gcctatcctat
121 ccctctatgg caattgcttg caccaggcag attttattcc caggagatat tggctatgt
181 ctgacatcaa ttttttggt tgtcacaaat gggatgaatg ttactggcat ctggtgggtg
241 gagtctcagag atgctgcttca atatcctatca gtgcatcaaga cagatccacc aaaaagaatc
301 . 3'

CUADRO 2. Acceso al Banco de Genes y Secuencia de Consenso del Marcador DYS391.

GDB Locus name: DYS392

Genbank Accession G09867
Alias CHLC.ATA25F04.P16979.
GDB: GDB: G00-455-698, accession id GDB: 455698
Consensus sequence structure: Allele DYS392-16, 264 bp.

5' 1 tcattaatct agcttttaaa aacaactaat ttgatttcaa gtgtttgtta tttaaaagcc
61 aagaaggaaa acaaattttt tcttgtatc accattatt tattattatt attattatta
121 ttattattat tattattnnn rcctaaggaa tgggattggt aggtttaatg atcctctgt
181 tttgacttct ttgagatatt tccagactac tttccacttt gactgtagga attacattgcatcaactgg
gtct ..3'

CUADRO 3. Acceso al Banco de Genes y Secuencia de Consenso del Marcador DYS392.