

UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS SOCIALES

ESTRUCTURACIÓN GENÉTICA EN
SANTIAGO DE ACUERDO AL ESTRATO
SOCIO-ECONÓMICO

MEMORIA PRESENTADA PARA OBTENER EL TÍTULO DE
ANTROPÓLOGO FÍSICO

TOMÁS BENJAMÍN GONZÁLEZ ZARZAR
PROFESOR GUÍA: SERGIO FLORES

2013

Departamento de Antropología

Agradecimientos

A todos y cada uno de los que se hicieron parte y colaboraron en la realización de esta memoria. A mis padres y hermanas, por su constante apoyo, resguardo y comprensión, y por mostrar siempre un genuino interés en los temas que me apasionan.

A quienes ayudaron en el arduo proceso de toma de muestras: A mis amigas y compañeras, Francisca Concha, Ofelia Meza, Ana Buchi y Constanza Silva, por colaborar siempre desinteresadamente. Al Cultural Pablo de Rokha de Puente Alto, por darme el espacio en sus eventos.

Con quienes compartí toda mi formación de educación superior, a mis compañeros y amigos Gabriel Cavieres, Francisca Concha, Natalia Marabolí, Ofelia Meza, Victoria Montero y Thomas Püschel. Gracias a ustedes por todos esos años de aprendizaje.

A quienes me guiaron en este proceso de formación, Eugenio Aspillaga y Rodrigo Retamal, por su amor a la disciplina, que supieron compartir conmigo.

Al profesor Sebastian Krapivka, por su apoyo y su inmensa ayuda en los momentos críticos.

A mi guía de memoria, profesor Sergio Flores, por su constante apoyo, por su confianza en mis capacidades y por su amistad.

A todos quienes donaron muestras para esta investigación

Resumen

En la población de Santiago se ha identificado un cline socio-genético estructurado (CSE), en donde los distintos estratos socio-económicos (ESE) comparten distintas frecuencias alélicas en los grupos sanguíneos ABO y Rh, evidenciando una mayor proporción de componente genético amerindio en ESE más bajos. Se ha postulado una hipótesis socio-cultural (H1) en donde españoles y amerindios (y sus genomas) habrían ocupado segmentos sociales separados, que se habrían mantenido con un bajo flujo génico durante las generaciones sucesivas. Esta hipótesis se hace plausible en el momento que Santiago ha sido caracterizada como una ciudad segregada socio-económicamente y en donde fenómenos de auto-identificación refuerzan estas nociones. Con lo anterior se ha levantado la propuesta de un Marco de Referencia Sociogenético en Salud Pública como una para-propuesta de las variables etnia y raza. En la presente investigación se evalúa la existencia de un CSE a partir de nuevas tecnologías y métodos. Se compararon las frecuencias alélicas de 5 marcadores genéticos autosomales (2 *Alu* y 3 SNP) en 5 ESE distintos. Las frecuencias de cada marcador se mantienen constante entre los ESE y se muestran consistentes a lo esperado en una población mestiza entre europeos y amerindios. Sin embargo, el componente amerindio se ha estimado mucho mayor (50 %) que en estudios anteriores. La población de Santiago, a pesar de su alto nivel de segregación socio-económica, se ha mantenido cohesionada en el ámbito reproductivo. Por otro lado, la creación de propuestas tangibles (políticas públicas) en torno a segmentos de la población

mediante nociones tipológicas de la variación humana conforman un serio riesgo a la hora de biologizar las diferencias sociales, o la generación de estereotipos.

Índice general

Agradecimientos	2
Resumen	3
1. Introducción	9
2. Antecedentes	12
2.1. Gradiente sociogenético o cline socio-genético estructurado (CSE) . . .	12
2.2. Estructuración poblacional	15
2.3. Historia poblacional y segregación	17
2.3.1. La conformación del mestizo	17
2.3.2. Segregación	23
2.4. Estratos y estratificación	30
2.5. Marcadores genéticos y ancestría	31
2.6. Relevancia	34
2.7. Formulación del problema, objetivos e hipótesis	35
3. Materiales y métodos	37
3.1. Muestras	37
3.2. Metodología	38

4. Resultados	46
4.1. Obtención de datos	46
4.2. Estructuración genética	52
4.3. Búsqueda de SNPs	55
5. Discusión	62
6. Conclusiones	71
Bibliografía	77

Índice de figuras

2.1. Tamaño poblacional de Santiago 1570-1779	20
2.2. Mapa AMGS	27
3.1. Frecuencias alélicas poblacionales	40
3.2. Inserción <i>Alu</i> PV92	41
3.3. Gel electroforesis SNP Oca2	41
4.1. Frecuencias de establecimientos educativod por GRUPO	47
4.2. Frecuencias alélicas entre ESE	49
4.3. Frecuencias alelicas poblacionales	50
4.4. Análisis STRUCTURE	54
4.5. Estimación mezcla génica de Wang (2003)	56
4.6. Estimación mezcla génica de Chikhi et al. (2001)	57
4.7. Frecuencia de SNP analizados por autosoma	58
4.8. Frecuencia y proporción de AIM por autosoma	59
4.9. Frecuencias asiáticas y europeas en 10 nuevos AIM	60

Indice de tablas

2.1. Componente aborigen	14
2.2. Estimación de la población en Chile	19
2.3. Estimación de la población de Santiago	19
3.1. Marcadores utilizados	39
4.1. Equilibrio Hardy-Weinberg	51
4.2. Probabilidades LD	52
4.3. AMOVA	53
4.4. AMOVA locus-por-locus	53
4.5. Estimación de proporción de mezcla genética	55
1. Frecuencias alélicas 5 AIM	73
2. Programas PCR	73
3. Protocolos ER	74
4. SNP AIM seleccionados	74

Capítulo 1

Introducción

Al enfrentarse al análisis de la biología de las poblaciones humanas y su variación se hace necesaria la delimitación de cada una con fines comparativos. Desde este punto de vista, ¿qué puede caracterizar una población humana por sobre otra?; desde la genética de poblaciones esto se responde atendiendo a la cohesión que presenta cada unidad de estudio, tanto temporo-espacialmente, como reproductivamente. Desde otras perspectivas pueden ser relevante variables socio-culturales como la religión, la lengua y las clases sociales, entre otras, para delimitar y comparar las poblaciones de estudio. Sin embargo, ambos ámbitos se cruzan muchas veces para delimitar la existencia de una población. La cohesión reproductiva habla de poblaciones como unidades panmícticas, es decir, en donde cada individuo tiene la misma probabilidad de reproducción con cualquier otro del sexo opuesto. Cuando existe una probabilidad diferencial entre ciertos conjuntos de una población, se establece la existencia de un apareamiento selectivo, teniendo que evaluar la pertinencia del concepto de población para el conjunto entero. Muchas veces en poblaciones humanas, este apareamiento selectivo responde a factores socio-culturales, en donde es más probable la reproducción entre individuos de la misma lengua, religión o clase social, incluso dentro de un mismo espacio geográfico.

Desde la década de 1970 en Chile se han llevado a cabo estudios genéticos en donde la variable a comparar han sido los estratos socio-económicos (ESE). A través de los sistemas sanguíneos ABO y Rh se ha podido determinar que los ESE más bajos tienen un mayor componente genético amerindio, al contrario de los ESE más altos, lo que ha llevado a proponer la existencia de un gradiente socio-genético o cline socio-genético estructurado (CSE), en donde distintos componentes sociales comprenden distintos componentes genéticos. Su importancia ha radicado en la propuesta de un Marco de Referencia Sociogenético en Salud Pública como una para-propuesta de las variables etnia y raza (Vanegas et al., 2008), poniendo el acento en las diferencias genéticas de los distintos ESE, condicionando diferencias en cuanto a morbilidad, patologías y respuestas a drogas.

La distribución del sistema sanguíneo ABO varía de acuerdo a la región. En términos globales el alelo O es el más común, yendo de un 40 % a un 100 %, seguido por el alelo A (0 %-55 %) y el alelo B (0 %-40 %) (Cavalli-Sforza et al., 1994). El alelo O se ha estimado en un 100 % en población indígena de América (Cavalli-Sforza et al., 1994) y en alrededor de un 65 % en población española (Campillo, 1976). Dos hipótesis han emergido para explicar el comportamiento de las frecuencias alélicas de los sistemas sanguíneos ABO y Rh en Chile. La primera (H1) incluye tres procesos: (1) el mestizaje entre amerindios y españoles hace aproximadamente 20 generaciones atrás; (2) la distribución en los segmentos ocupacionales intelectuales o con más poder, prestigio e ingresos por parte de los españoles, mientras que mestizos y amerindios se distribuirían en ocupaciones manuales con menos poder, prestigio e ingresos; y (3) apareamiento selectivo positivo. La segunda hipótesis (H2) establece la existencia de una probabilidad diferencial de ciertas variantes génicas del sistema sanguíneo ABO para ocupar distintas posiciones de poder y prestigio dentro del sistema social chileno (Valenzuela, 2011). Mientras H1 propone una causa cultural cuya consecuencia es la conformación actual biológica de la población, H2 establece

la relación inversa.

La pregunta de investigación específica parte de la necesidad de poner a prueba H1 a través de otros marcadores y nuevas herramientas. No es de extrañar que esta hipótesis sea aceptada, incluso dentro del sentido común, dado el alto nivel de segregación socio-económica de Santiago.

Capítulo 2

Antecedentes

2.1. Gradiente sociogenético o cline socio-genético estructurado (CSE)

La población chilena actual se ha originado mayoritariamente a partir de la mezcla entre población española y amerindia, así muchos estudios han tratado de establecer las proporciones de los componentes genéticos en distintos segmentos y poblaciones del país. Desde la década de 1970 se han llevado a cabo estudios a partir del sistema sanguíneo ABO. Este sistema es probablemente el marcador genético mejor conocido y más estudiado, se ubica en el cromosoma 9 y se compone de tres alelos (A, B y O), dentro de los cuales A y B son codominantes y O es recesivo (Mielke et al., 2006). Sin embargo, el sistema ABO es más complejo, conteniendo subgrupos de los alelos A, B y O.

En 1971 en población urbana de Valparaíso Pinto-Cisternas et al. encuentran frecuencias génicas del alelo O que van desde 67,4% al 84,8% y del alelo A entre un 11,7% a un 28%. En 1983, en Valdivia Cohn encuentra frecuencias que fluctúan entre 61,1% a 82,9% en el alelo O y entre un 12,34% y un 29,92% en el alelo A.

En 1987, Valenzuela et al. encuentran frecuencias que van de 57,7% a 82,4% en el alelo O y entre un 11,4% y un 42,2% en el alelo A. Todos ellos coinciden en que los ESE más bajos presentan mayor frecuencia del alelo O mientras que los ESE altos presentan mayores frecuencias del alelo A (Tabla 2.1).

La distribución del sistema sanguíneo ABO varía de acuerdo a la población. En términos globales el alelo O es el más común, yendo de un 40% a un 100%, seguido por el alelo A (0%-55%) y el alelo B (0%-40%) (Cavalli-Sforza et al., 1994). El alelo O se ha estimado en un 100% (no existirían, por ende, los alelos A y B) en población indígena de América (Cavalli-Sforza et al., 1994), mientras que en población española se ha estimado en un 65% el alelo O y un 29% el alelo A (Campillo, 1976). Así, la estimación de la proporción de mezcla genética amerindia es mayor en los estratos más bajos, utilizando el método de Bernstein 1931. Con estos datos se ha postulado la existencia de un CSE, en donde distintos estratos poseen distintos componentes genéticos. Por lo menos en Santiago se ha propuesto la existencia de al menos 3 estratos socio-genéticos (ESG): el SI (de mayor prestigio, poder e ingresos) que comprende al 5% de la población con una virtual ausencia de mezcla aborígen, el SII que comprende el 23% de la población con un 18% de mezcla aborígen, y el SIII (de menor prestigio, poder e ingresos) que comprende un 72% de la población con un 37% de mezcla aborígen (Valenzuela et al., 1987).

Dos hipótesis han emergido para explicar el comportamiento de las frecuencias alélicas del sistema sanguíneo ABO en la población chilena. La primera (H1) establece tres procesos: (1) el mestizaje entre amerindios y españoles hace aproximadamente 20 generaciones atrás; (2) la distribución en los segmentos ocupacionales intelectuales o con más poder, prestigio e ingresos por parte de los españoles, mientras que mestizos y amerindios se distribuirían en ocupaciones manuales con menos poder, prestigio e ingresos; y (3) apareamiento selectivo positivo. La segunda hipótesis (H2) establece la existencia de una probabilidad diferencial de ciertas variantes génicas

Ciudad	Muestra	Es- trato	Frec. génica ABO*O	Mezcla aborigen m	Estudio
Valpa- raíso u	Población urbana	I	0,714	0,194	Pinto-Cisternas y cols., (1971)
Valpa- raíso u	"	II	0,729	0,239	"
Valpa- raíso u	"	III	0,765	0,349	"
Valpa- raíso u	"	IV	0,811	0,485	"
Santiago .	Donantes Banco de Sangre	I	0,627	-0,065	Valenzuela y cols., (1987)
Santiago .	"	II	0,711	0,186	"
Santiago .	"	III	0,729	0,240	"
Santiago .	"	IV	0,773	0,372	"
Santiago .	"	V	0,780	0,393	"
Valdivia v	Colegio	I	0,611	0,119	Cohn P. (1983)
Valdivia v	"	II	0,802	0,457	"
Valdivia v	"	III	0,788	0,417	"
Valdivia v	"	IV	0,829	0,540	"

Tabla 2.1: Componente aborigen según ESE en Santiago, Valparaíso y Valdivia. Modificado de Harb, 2004

u: clasificado de acuerdo a zona de residencia

.: donantes de sangre con estratificación socioeconómica

v: estudiantes de colegio con estratificación socioeconómica

del sistema sanguíneo ABO y Rh para ocupar distintas posiciones de poder y prestigio dentro del sistema social chileno. Esta última establece que H1 no es suficiente ya que no lograría explicar: (1) la coincidencia de los CSE entre Chile e Inglaterra (Beardmore y Karimi-Booshehri, 1983); (2) las frecuencias de los alelos A y Rh(-) en los ESE altos son significativamente más altos que los encontrados en Europa; (3) el alelo B se comporta en contradicción a los esperado por H1; y (4) la existencia de una interacción entre CSE y el sexo, tanto en Chile como Inglaterra. En cambio, propone que debe existir una función psico-social que asocie ambas variables, postulando para esto 3 hipótesis alternativas: (1) efecto pleiotrópico: los alelos del sistemas sanguíneos ABO y Rh están involucrados directamente en los aspectos psico-socio-culturales que condicionan la probabilidad de pertenecer a cierto ESE; (2) desequilibrio de ligamiento: los alelos del sistema sanguíneos ABO y Rh están ligados a los genes responsable de esta asociación; y (3) interacción génica: los alelos del sistemas sanguíneos ABO y Rh interactúan con otros genes para generar esta asociación (Valenzuela, 2011).

2.2. Estructuración poblacional

Dentro de la genética de poblaciones el cline socio-genético estructurado (CSE) es una forma de estructuración genética poblacional, definida como la heterogeneidad dentro de una población en la probabilidad de reproducción entre dos individuos tomados al azar (Hamilton, 2009). El resultado de esta estructuración poblacional genera que diversos conjuntos de la población se diferencien genéticamente a lo largo del tiempo por deriva génica ¹, o que se diferencien genotípicamente producto de un efecto de endogamia.

En el contexto del presente trabajo el concepto de población es esencial. A lo largo

¹Cambio en las frecuencias génicas entre las generaciones producto de un fenómeno estocástico derivado de un error en el muestreo.

de la historia éste ha tenido una serie de definiciones enfatizando ciertos aspectos por sobre otros. Todos, sin embargo, destacan un punto fundamental: la unión y cohesión entre los individuos. Dentro de éstos hay dos énfasis que destacan, el ecológico y el evolutivo. El primero hace relación con las fuerzas cohesivas demográficas, como la ocurrencia en un mismo tiempo y espacio (teniendo la oportunidad de interacción entre los distintos individuos), mientras que el segundo hace relación con el aspecto genético, enfatizando el factor reproductivo en cuanto interacción (Waples y Gaggiotti, 2006). De esta forma las poblaciones son entendidas como unidades panmícticas, es decir, en donde cada individuo tiene la misma probabilidad de reproducción con cualquiera del sexo opuesto. Cuando existe una probabilidad diferencial entre distintos individuos tomados al azar se establece la existencia de un apareamiento selectivo y de una estructuración genética, teniendo que evaluar la pertinencia del concepto de población para el conjunto entero. Este apareamiento selectivo puede ser positivo (cuando las parejas son fenotípicamente similares), negativo (cuando son fenotípicamente distintas) o puede ser entre parientes, denominado endogamia (Hartl, 1981). La consecuencia de un apareamiento selectivo positivo o endogamia es que tenderán a incrementar la frecuencia de los genotipos homocigotos a expensas de los heterocigotos (incrementando la varianza fenotípica). Por otro lado, al existir un bajo flujo génico entre los conjuntos, las frecuencias génicas tenderán a cambiar a lo largo del tiempo sólo por deriva de manera diferencial. Generalmente se propone la existencia de barreras geográficas que impiden el contacto fluido entre dos conjuntos de una población o barreras conductuales, los cuales disminuyen el flujo génico. En el caso de poblaciones humanas es necesario tener una tercera explicación como lo es la cultura; fenómenos culturales variados pueden y provocan diversas formas de estructuración poblacional. En estudios anteriores se ha comprobado la existencia de una disminución del flujo génico en poblaciones humanas como consecuencia de la distancia geográfica y factores culturales como el lenguaje y la religión (Mielke et al.,

2006).

2.3. Historia poblacional y segregación

Para el CSE se ha postulado clásicamente la H1: por un lado una historia poblacional particular, en donde la masa indígena y mestiza ocupa durante el período de la conquista y en adelante, los ESE más bajos, mientras que los conquistadores castellanos ocupan la cúspide de la pirámide social. Por otro lado, la alta segregación y mantención de estos segmentos dentro de la estructura social de la población lleva a pensar la existencia de un nulo o bajo flujo génico entre los segmentos, aumentando o al menos manteniendo la diferenciación genética inicial. En esta sección, pasaremos a revisar someramente esta explicación, centrándonos primero en la conformación del mestizo (en cuanto entidad biológica) y posteriormente en los aspectos segregativos de la ciudad

2.3.1. La conformación del mestizo

En enero de 1541 arriba a la ribera norte del río Mapocho la segunda expedición castellana dirigida por Pedro de Valdivia. Ya en febrero del mismo año, esta población había incorporado a una buena cantidad de “naturales”, fundamentales en el levantamiento de las primeras construcciones de la ciudad (de Ramón, 2007). De la misma forma, fueron incorporados al sistema productivo agrícola del imperio español bajo la institución de las encomiendas, como “peones”, “inquilinos”, “artesanos” o “trabajadores agrícolas”, siendo integrados a la sociedad colonial mediante la aculturación y el mestizaje (Waldman, 2004). En Santiago, ya para 1550 “convivían españoles e indígenas, aunque estos últimos se habían instalado con sus “rancheríos” en los alrededores y en solares desocupados en el interior de la ciudad como se denunció en 1568” (de Ramón, 2007). A poco andar, el descenso de la población indígena

provocada por la “guerra sostenida (...) contra Pedro de Valdivia en la década de 1540 y al servicio personal establecido por el sistema de encomiendas”, resultó en la decisión de trasladar población indígena desde otras zonas (como los huilliches del sur del país o los indios guarpes desde Cuyo). Con todo lo anterior, el crecimiento natural de la población de Santiago se conformó fundamentalmente en el nacimiento de mestizos mientras no hubo familias constituidas. Sin embargo, al momento de la llegada de las primeras mujeres, a partir de 1555, se conformaron las familias católicas occidentales y legítimas, coexistiendo con las familias compuestas por hijos sin padre y sin legitimidad (Waldman, 2004).

Aunque hay muchos datos con respecto al número poblacional de las primeras décadas de la conquista, la fiabilidad de estas ha sido puesta en duda, tanto para Chile como para la jurisdicción de Santiago (de Ramón, 2007). Sin embargo, es posible ver algunas tendencias como el aumento de la población mestiza y criolla, y la disminución de la población indígena en los primeros años por las razones mencionadas anteriormente (Tablas 2.2 y 2.3). A comienzos del S. XVII el 18,84 % de la población de Santiago (por lo menos de su traza urbana) era de origen español, un 7,41 % mestizo, un 67,57 % indígena y un 6,18 % de origen africano. Por otro lado, a fines del mismo siglo la población de Santiago ha sido estimada en unas 12.000 personas (de Ramón, 2007). Sólo para el año 1779 se realiza el primer censo del Obispado de Santiago arrojando un total de 40.607 habitantes, entre los cuales un 52,49 % eran españoles, 15,43 % mestizos, 13,43 % indios y 18,64 % mulatos y negros (Figura 2.1). Esta casi cuadruplicación de la población de Santiago en solo 100 años, ha sido explicada por una constante migración hacia la ciudad, más que por un crecimiento vegetativo (de Ramón, 2007). Para el siglo XX, el crecimiento de Santiago verá proporciones inéditas, yendo desde casi 1 millón de habitantes en 1940, hasta más de 6,5 millones para el año 2012.

La estratificación social de Santiago se conformará desde el inicio a partir de un

Año	Vecinos, españoles europeos y criollos	Mestizos blancos	Negros y mestizos de color	Indios de paz en encomiendas y otros	Indios rebeldes	Total
1570	7000	10000	7000	450000	150000	624000
1590	9000	17000	20000	420000	120000	582000
1600	10000	20000	19000	230000	270000	549000
1620	15000	40000	22500	230000	250000	557500

Tabla 2.2: Estimación de la población chilena desde 1570 a 1620 dividido entre diversas categorías étnicas. Fuente: Mellafe (1959)

Año	Cronista	Es- pañoles	Indios	Ne- gros	Total
1570	Juan López de Velasco	350-400	70000-80000	-	70350-80400
1610	Antonio Vásquez de Espinoza	1676	-	-	1676
1610	Hernando Machado	1717	8600	300	10617

Tabla 2.3: Estimación de la población santiaguina por diversos cronistas desde 1570 a 1620 dividido entre diversas categorías étnicas. Fuente: de Ramón (2007)

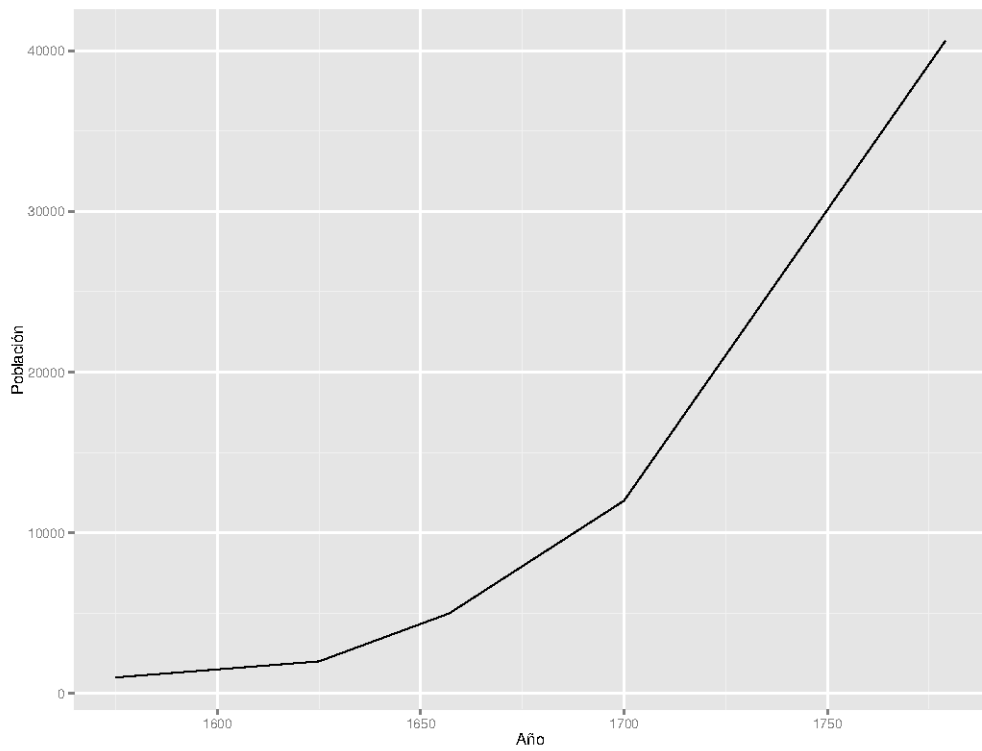


Figura 2.1: Incremento de la población santiaguina total entre los años 1570 a 1779. Fuente de Ramón, 2007

criterio étnico (de Ramón, 2007, INE, 2009), con “una población blanca colonizadora que recibía del conquistador porciones territoriales que debía trabajar; una población indígena encomendada, utilizada en el trabajo a cambio de protección, alimentos y evangelización; y una población negra esclava de número inferior a las otras, pero no por ello despreciable, que llegó sometida a este territorio como mano de obra y servicio personal del español” (INE, 2009). En otras investigaciones también se ha constatado que el aporte de la población africana habría sido marginal dentro de la conformación poblacional de Chile y Santiago (Arcos-Burgos et al., 2004, Mellafe, 1959). Posteriormente, la estratificación social tendió a complejizarse con la constitución de las primeras familias cristianas, el casamiento entre mestizas y españoles que, sin embargo, conformaron importantes familias, la llegada de nuevos contingentes desde España o Perú conformando estratos intermedios, y la gran masa de indígenas conformando a través del sistema de encomiendas los estratos bajos, también constituidos, aunque en menor porcentaje, por familias españolas pobres (de Ramón, 2007).

Desde la genética

El origen mestizo de la población santiaguina se evidencia claramente en la composición genética de los individuos actuales. Bajo la hipótesis de una población chilena actual mestiza podemos suponer que toda variante genética que tenga frecuencias diferenciales entre poblaciones europeas y amerindias, presentará valores intermedios entre ambas en las muestras actuales. De esta manera, los estudios con marcadores genéticos clásicos han encontrado valores intermedios entre las frecuencias europeas y amerindias (Cohn, 1983, Pinto et al., 1980, Pinto-Cisternas et al., 1971, Valenzuela et al., 1987, Valenzuela y Harb, 1977). Otros estudios han detectado otras variantes genéticas en la población de Santiago que se encuentran, de la misma manera, en frecuencias intermedias entre poblaciones europeas y amerindias, con una semejanza

mayor a las europeas (Vieyra et al., 2003). Arcos-Burgos et al. (2004), a través del análisis de 5 marcadores genéticos llegan a estimar, para ESE bajos y medio bajos, un 60 % de componente europeo y un 40 % de componente amerindio. Sin embargo, el reciente análisis del *loci TSER*² ha mostrado ser incompatible con la hipótesis de mestizaje, encontrando valores que se alejan de lo esperado para una población mestiza (Acuña et al., 2006).

La segunda hipótesis que se maneja dentro del mestizaje chileno (y también latinoamericano) es lo que se ha definido como “cruzamiento al derecho” en donde la reproducción habría sido en el comienzo de la conquista mayormente entre mujeres amerindias y hombres europeos. De esta forma, Cifuentes et al. (2004) utilizando los *loci* DYS19 y DYS199 del cromosoma Y³ para evaluar la hipótesis de una composición eminentemente española en la contribución genética paterna de la población mestiza chilena, llegan a calcular un 15,83 % de componente amerindio en ESE bajos, y de 6,91 % en ESE altos. Utilizando el DNA mitocondrial (mtDNA)⁴, Rocco et al. (2002) encontraron una proporción cercana al 84 % de componente genético amerindio, lo que es calificado de un componente “significativamente superior a la componente indígena calculada a partir de marcadores nucleares que es de aproximadamente 40 %, dependiendo del estrato socioeconómico (Rothhammer, 1987), así como de la obtenida a partir de marcadores moleculares del Y que alcanza el 22 % (manuscrito en preparación). En otras palabras, 84 % de las mujeres que dieron origen a la población actual de Santiago habrían sido indígenas(…)” (Rocco et al., 2002). De esta forma, estos datos dan cuenta de lo siguiente: (1) un mayor componente genético europeo, incluso en ESE bajos, para marcadores autosomales (biparentales); (2) un bajo componente amerindio para el cromosoma Y (patrilineal), mayor en ESE bajos (15,83 %) que en altos (6,91 %); y (3) un alto componente ame-

²Thymidylate synthase enhancer region

³Marcador uniparental, cuya herencia es de padre a hijo.

⁴Marcador uniparental, cuya herencia es por vía materna.

rindio en general para el mtDNA (matrilineal).

Hay que tener en cuenta que la conformación actual de la población de Santiago ha tenido cambios durante toda su historia. Uno de los más impresionantes en cuanto explosión demográfica se refiere ha sido el ocurrido a lo largo del siglo XX, denominado migración campo-ciudad, que hizo aumentar la población santiaguina casi 12 veces entre 1895 y 1992. Esta gran migración tiene efectos considerables dentro de la población receptora al momento que ingresan nuevos alelos. Se ha encontrado que durante el siglo XX se ha incorporado un componente genético “sureño” a la población santiaguina (Leiva, 2010).

2.3.2. Segregación

El concepto de segregación hace referencia a un fenómeno de separación de una cosa de otra, no siendo negativo ni positivo *a priori*. Este tipo de fenómenos han sido usuales en toda conformación socio-cultural humana, lo que se puede ver actualmente desde la segregación etaria en el sistema educativo, hasta el acceso segregativo del uso de servicios higiénicos. Los estudios sobre segregación han tenido su mayor desarrollo en el campo de los estudios urbanos y generalmente se la ha entendido como la diferencia distributiva de diversos grupos sociales entre distintas unidades de organización y/o asociadas a un determinado territorio o zona geográfica (James & Taeuber, 1985)”(Valenzuela et al., 2010). En este sentido, se establecen dos dimensiones de la segregación: una social, referida a la ausencia de interacción entre grupos sociales, y una geográfica, que se refiere a una distribución desigual de ciertos grupos en el espacio (Rodríguez, 2001). Sin embargo, se ha definido que cuando la variable que se utiliza para segregar es de índole racial o socio-económica, se pueden llegar a provocar fenómenos negativos de estigmatización, oferta deficiente de servicios, limitaciones para acceder a redes y contactos, y carencia de socialización.

Al contrario del caso estadounidense, en donde la variable que explica en mayor

medida los fenómenos segregativos es racial, tanto en América Latina y el Caribe, la segregación se estructura en torno a variables socio-económicas como los activos sociales, que se expresan en una minoría que los posee en abundancia y una mayoría que los presenta de manera escasa (Rodríguez, 2001). Así, la segregación socio-económica (SSE) puede entenderse como la homogeneidad o escasez de mezcla socio-económica en cualquier unidad que estemos observando. En el caso de Chile, fenómenos de SSE son usuales en varios niveles, siendo ejemplares (en cuanto a la cantidad de información al respecto) el caso residencial y del sistema educativo.

Segregación residencial

La segregación residencial se la ha entendido como la concentración de residentes de un ESE en áreas particulares de la ciudad. Ésta ha sido tratada de diversas formas en la literatura especializada, desde el análisis espacial *per se*, las causas de la segregación, sus consecuencias y relaciones con otros problemas de las actuales urbes (como la desigualdad y la pobreza), hasta su impacto en las relaciones sociales. Como se mencionó, generalmente se establece que cuando la variable que segrega es de índole socio-económico, se generan consecuencias negativas, existiendo amplia evidencia apoyando esta noción en los ámbitos de las oportunidades laborales, aspectos de la salud y la eficacia política de los segmentos segregados (Valenzuela et al., 2010). En Chile, se ha encontrado que el efecto de la segregación residencial tiene repercusiones negativas en los segmentos más pobres en los ámbitos de asistencia preescolar, deserción escolar, atraso en los estudios y la inactividad económica de los más jóvenes (Larrañaga y Sanhueza, 2007). También ha habido intentos de estudiar y analizar los efectos positivos de la segregación residencial en las ciudades, como la llegada de residentes de mayores ingresos a barrios pobres, a través de los “barrios cerrados”, como los condominios (Juárez, 2006).

En la historia de Santiago, la distribución de la heterogeneidad social y étnica

es desde un comienzo desigual entre los distintos territorios. Para fines del S. XVII en el centro de la ciudad residía un 44 % de mestizo, indios y africanos, mientras que un 35 % en San Isidro (sur de la Cañada), un 10,1 % en la parroquia de Santa Ana (abarcando los arrabales del sector occidental) y un 55 % en la de Renca (de Ramón, 2007). Por la década de 1870 el intendente de aquel entonces, Benjamín Vicuña Mackenna se refiere a dos sectores dentro de la ciudad, “el Santiago propio, la ciudad ilustrada, opulenta, cristiana” y la ciudad de los arrabales, “inmensa cloaca de infección y de vicio, de crimen y de peste, un verdadero potrero de la muerte” (de Ramón, 2007). En 1979, las autoridades de la dictadura militar, mediante el decreto supremo número 420 modifican el plan intercomunal en base a los principios de que el suelo urbano no es un recurso escaso y la planificación urbana deberá tener el mínimo de intervención por parte del Estado (De Mattos, 2002, de Ramón, 2007). Conjunto con lo anterior, y en el mismo año, mediante el decreto número 2552, el “Programa de viviendas básicas o programa de erradicación de campamentos” cadastró y movilizó a unas 37.597 familias viviendo en campamentos (un 72,6 % del total viviendo en campamentos en ese entonces). Ambos fenómenos profundizaron el aspecto segregativo de la ciudad de Santiago, ya apreciable en la composición étnicas en los censos de fines del S. XVII (De Mattos, 2002, de Ramón, 2007).

En la actualidad en Chile y específicamente en su capital Santiago, la segregación residencial ha sido altamente estudiada y debatida. Sin embargo, la discusión está lejos de ser zanjada, no existiendo un consenso sobre su intensidad, sus efectos ni sus causas. Esta confusión es básicamente producto del uso de bases de datos y de análisis estadísticos distintos entre los autores. Mientras unos llegan a la conclusión de que en Santiago existe una distribución espacial desigual de acuerdo a los distintos ESE (Vargas, 2006), otros lo tildan de una segregación moderada o no especialmente alta y que ha tendido a disminuir con el tiempo (Agostini, 2010, Rodríguez, 2001)

Algunos estudios han indicado que la segregación de los grupos sociales dentro

de las metrópolis se ha ido incrementando a una escala geográfica cada vez mayor, identificando en la globalización la causa principal de dicho proceso, en cuanto el principal medio por el cual las diferencias sociales se han incrementado (Juárez, 2006). Sin embargo, otros han establecido que la escala geográfica, o en otras palabras, el nivel de agregación de la segregación residencial, está teniendo dos tendencias principales en las ciudades chilenas y también latinoamericanas: (1) la disminución de la escala geográfica en aquellos lugares de gran dinamismo inmobiliario privado⁵, y (2) el aumento de la escala en aquellos sectores de asentamiento de las nuevas familias de bajos recursos. En este caso, ambos fenómenos se explican principalmente por la liberalización de los mercados de suelos, generando un mayor dinamismo en el mercado privado inmobiliario, por una parte, y por el otro, un aumento persistente de los precios de los suelos (Sabatini et al., 2001).

Para la mayoría de las ciudades de América Latina se ha postulado un patrón de segregación residencial similar, en donde las familias de las élites se van concentrando en sectores de crecimiento bien definidos en términos geométricos, conformando un cono que va desde el centro histórico hacia un sector de la periferia, mientras que los grupos sociales más pobres se concentran generalmente en las zonas más lejanas y peor equipadas (Sabatini et al., 2001). Según el índice de pobreza, es posible ver claramente agregados comunales que presentan valores similares: las cinco comunas que presentan menores valores (Vitacura, Las Condes, Providencia, Lo Barnechea y Santiago) conforman un cono con base en el sector nor-oriental y cuyo vértice llega a la comuna de Santiago. Por otro lado, se conforma un arco de comunas con valores moderados que va desde Quilicura en el norte, hasta Puente Alto en el sur, y se concentran una serie de aglomeraciones comunales que presentan valores altos dentro de la ciudad (Figura 2.2).

⁵Por ejemplo, una gran heterogeneidad socio-económica dentro de las comunas, pero barrios (dentro de las comunas), bien diferenciados socio-económicamente unos de otros.

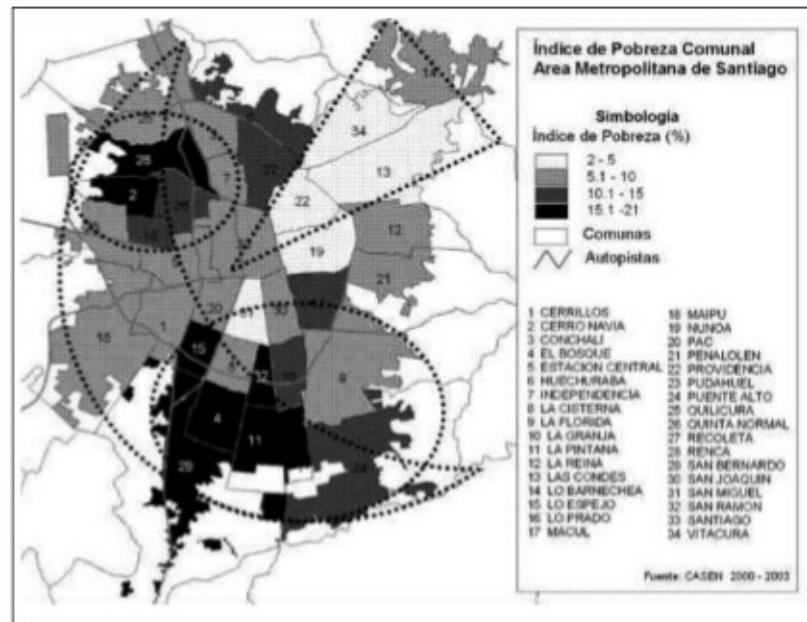


Figura 2.2: Mapa AMGS, tomado de Orellana y Fuentes (2008)

La configuración actual de Santiago es producto de los distintos movimientos migratorios desde su fundación en 1541. En términos interurbanos, se ha establecido en la actualidad la existencia de una dinámica expulsora del centro de la ciudad hacia comunas de la periferia (Rodríguez, 2008). Estos procesos recientes de migración interurbana han llevado a una consolidación de los patrones segregativos existentes, en el momento en que estas migraciones son procesos selectivos. De esta forma, se afirma que los segmentos más ricos tienden a la autosegregación producto de su capacidad de elección de la ubicación, mientras que los segmentos más pobres sólo pueden hacerlo donde sus recursos se los permiten, categorizando este fenómeno como de segregación forzada o involuntaria (Ortiz y Aravena, 2002).

Segregación educativa

La segregación educativa se la ha definido como la distribución desigual de alumnos de distintos ESE entre las escuelas. A diferencia de los estudios sobre la se-

gregación residencial, la segregación educativa sólo ha sido tratada en los últimos años en Chile, existiendo aún mucha confusión sobre su magnitud y relevancia. La confusión en general ha sido producto de dos problemas: (1) por la falta de estudios empíricos, sólo llevados a cabo en los últimos años, y (2) la existencia de resultados contradictorios. Sin embargo, estudios han dado cuenta de una alta segregación educativa (Valenzuela et al., 2010), incluso en el panorama internacional (OECD, 2011). Algunos han criticado el enfoque comparativo entre sectores educacionales (público y particular-subvencionado), haciendo énfasis en la importancia de la escuela como unidad de análisis, ya que, aunque el sistema de *vouchers* (particular-subvencionado) abarca una gran diversidad de ESE, cada escuela en particular abarca una población homogénea desde el punto de vista socio-económico (Mizala y Torche, 2012). Esto ha llevado a algunos a establecer que “la alta segregación del sistema escolar chileno hace que los resultados de aprendizaje de los estudiantes sean altamente predecibles dependiendo de la escuela en que se encuentran estudiando” (Bellei, 2007).

Algunas causas que se han establecido de este fenómeno son: (1) la segregación residencial, (2) la elección de los padres, (3) la libertad de escoger los establecimientos por preferencias culturales más que por factores espaciales, y (4) el sistema de financiamiento basado en *vouchers* o subvención a la demanda (Valenzuela et al., 2010). A pesar de que no existen estudios empíricos sobre las consecuencias efectivas que un sistema educacional segregado pueda generar en los estudiantes en Chile, se han definido tres argumentos de por qué debiera generar preocupación: (1) empobrecimiento de la experiencia de socialización, al no convivir con personas de diversos estratos socio-económicos, (2) una serie de efectos directos e indirectos sobre la calidad de la educación en desmedro de los segmentos más pobres, y (3) la eficacia de las políticas sociales sobre un conjunto de individuos vulnerables, es decir, que a la vulnerabilidad individual se suma la colectiva, haciendo aún más compleja la implementación de estas políticas (Valenzuela et al., 2010).

Barreras reproductivas

Algunos estudios han constatado la existencia de una disminución del flujo génico en poblaciones humanas como consecuencia de la distancia geográfica, y factores culturales como el lenguaje y la religión (Mielke et al., 2006). Este fenómeno adquiere matices distintos que son población-específico, de acuerdo a causas históricas que radican en la elección de parejas reproductivas. Esta elección, individual y afectada por expectativas socio-culturales, tiene un claro impacto en la composición genética de la población como un todo, dando como resultado un fenómeno de apareamiento selectivo. La pregunta pues, de ¿quién se reproduce con quien?, puede ser respondida de forma amplia como: (1) entre iguales, dándose fenómenos como la homogamia y endogamia; (2) entre distintos, con fenómenos como heterogamia y exogamia; y (3) la compleja realidad entre las opciones anteriores. En Chile se han encontrado diferencias en las frecuencias alélicas entre dos ESE distintos, mientras que las frecuencias entre los progenitores no difería de manera significativa, lo que lleva a postular la existencia en Santiago de fuerzas socio-económicas que provocan un apareamiento selectivo (Valenzuela y Harb, 1977).

El análisis del nivel educacional de las parejas ha sido un tópico de interés para el estudio de la desigualdad y su mantención a lo largo de las generaciones, en cuanto cómo influye el nivel educacional, como estimador del ESE, en la elección de parejas. En Chile se ha encontrado un alto índice de homogamia educativa (70%), mayor que el promedio en América latina (66%), tendencia que ha disminuido con el tiempo y en donde el quintil de mayores ingresos presenta el mayor índice de homogamia educativa (De la Maza, 2008). Similarmente, se ha encontrado una alta asociación entre los niveles educativos extremos de las parejas (entre 0 y 6 años de escolaridad y con 16 o más años de escolaridad), en donde el 91% de los individuos casados, lo hacen con parejas en el mismo grupo educacional o alguno adyacente. La tendencia temporal de este fenómeno es que disminuye para quienes tienen educación básica

incompleta y ha aumentado para quienes tienen 16 o más años de escolaridad, “lo que concuerda con lo observado internacionalmente y respalda la hipótesis de una élite que se distancia cada vez más del resto de la sociedad.” (Serrano, 2010). De la misma manera, se ha encontrado la existencia de altas barreras matrimoniales en los ESE altos y barreras más fluidas en los ESE medios y bajos (Torche, 2010). Conjunto con lo anterior, se ha encontrado que Chile posee, a pesar de sus altos índices de desigualdad, una gran movilidad social comparado con el de países industrializados, salvo en la parte superior de la distribución económica (Torche, 2005). En otras palabras, las élites nacionales tienden a presentar mayores barreras matrimoniales y de movilidad social.

Hay que tener en cuenta, sin embargo, que la institución del matrimonio es demasiado limitada para el fenómeno reproductivo que finalmente es el que afecta la composición genética de una población a lo largo del tiempo. En otras palabras, la pregunta de ¿quién se casa con quien? no responde completamente la pregunta ¿quién se reproduce con quien?.

2.4. Estratos y estratificación

Para la estratificación social, entendida como la división y agrupamiento del conjunto de los individuos de una sociedad de acuerdo a variables socio-económicas, vemos dos formas de estratificar, que creemos no agotan en absoluto la extensa discusión sobre el tema. En primer lugar, consideramos los esquemas de clases sociales, que hacen referencia a divisiones estructurales y relacionales dentro de un sistema social particular. En este sentido, en la actualidad se han definido una serie de esquemas de clasificación que se dividen en tres grandes grupos: (1) categorías ocupacionales convencionales; (2) escala subjetiva de prestigio socio-ocupacional; y (3) esquemas de clases con bases teóricas, con referencia a Marx y Weber (Atria, 2004).

La segunda manera de estratificar la sociedad es a través de nociones gradacionales. A través de variables como el ingreso, consumo y escolaridad, entre otros, se generan agregados de individuos que comparten similitudes en esas variables. Al contrario de los anteriores, se caracterizan por la distribución cuantitativa de una variable, y no por la relación que establecen los individuos dentro de la sociedad (Wright, 1985).

2.5. Marcadores genéticos y ancestría

El término “marcador genético” se ha utilizado para describir las variantes genéticas o polimorfismos, siendo este último término referido, por convención, a un rasgo genético que se presenta en a lo menos dos formas dentro de una población, siendo la más infrecuente no menor de un 1% (Mielke et al., 2006). La variación genética humana entre distintas poblaciones fue inicialmente demostrada por un estudio de 1919 de Hirszfled y Hirszfled sobre el grupo sanguíneo ABO (Cavalli-Sforza y Feldman, 2003), el cual sería el primer polimorfismo por describir y uno de los más importantes por largos años (Landsteiner, 1900 en Mielke et al., 2006). Con la identificación de nuevos marcadores proteicos de grupos sanguíneos (p.e. MNS y Rh), los trabajos posteriores de genética de poblaciones hicieron uso de una mayor gama de productos génicos, proteínas del plasma y enzimas. En la actualidad, la descripción y comparación de secuencias directas de DNA es cada vez más usual gracias a los avances tecnológicos, como los métodos más rápidos para la extracción del DNA, el descubrimiento de las enzimas de restricción (ER) y sus potencialidades en la constatación de polimorfismos, el uso de técnicas derivadas de la hibridación del DNA, la invención de la reacción en cadena de polimerasa (PCR) y la secuenciación automática del DNA (Rubicz et al., 2007).

Toda variación genética es producto de mutaciones, las cuales pueden ir desde el

cambio de un sólo nucleótido hasta la inserción o delección de segmentos mayores. La más común y la más útil para muchas investigaciones son los polimorfismos de un solo nucleótido o SNP, los cuales son una variante de la secuencia del DNA, en donde un nucleótido (A, C, T o G) en una misma posición, difiere entre distintos individuos. La mayor parte de la variación en las secuencias de DNA se atribuye a los SNP, teniendo una ocurrencia en promedio de entre 1000 a 2000 bases. A diferencia de los microsatélites, los SNP poseen una menor tasa de mutaciones recurrentes, haciéndolos buenos indicadores de la historia poblacional. Su gran representación dentro del genoma, y su potencialidad para la detección de un gran número de éstos, los han hecho los preferidos para los estudios de asociación. Una de las primeras tecnologías para la detección de SNP fueron los polimorfismos de longitud de fragmentos de restricción o RFLP ⁶. La tecnología subyacente a los RFLP es la existencia de las ER, las cuales tienen la capacidad de reconocer y cortar una secuencia específica de DNA. De esta forma, el polimorfismo será la presencia o ausencia de un sitio de restricción específico, es decir, el cambio de un sólo nucleótido; en donde esté presente el sitio de restricción, la ER cortará el DNA y el resultado serán dos fragmentos cortos, mientras que su ausencia dará como resultado un sólo gran fragmento de DNA. Los RFLP actúan como marcadores co-dominantes, ya que es posible distinguir los heterocigotos de los homocigotos.

Por otro lado, las inserciones *Alu* son parte de los elementos transponibles del grupo SINE ⁷, existiendo 1.500.000 de copias dentro del genoma humano, y siendo los elementos móviles más abundantes dentro de éste. Son un tipo de marcador cuyo carácter ancestral es conocido, sus alelos son idénticos por descendencia, ya que han sido originados por un único evento mutacional, y se mantienen estables en términos filogenéticos (Batzer y Deininger, 2002, Batzer et al., 1994). Algunas inserciones han

⁶Restriction Fragment Length Polymorphism

⁷Elementos intercalados cortos, por su sigla en inglés

ocurrido de manera tan reciente que mantienen un carácter polimórfico dentro de la población humana. Estas características han promovido el uso de las inserciones *Alu* para analizar el origen y relaciones filogenéticas entre grupos humanos modernos. El tamaño de las inserciones *Alu* son de 300pb (pares de base) y se encuentran regularmente en intrones, regiones 3' sin traducir de genes y en regiones intergénicas

En poblaciones misciginadas, para estimar el porcentaje de ancestría con respecto a una de las poblaciones parentales en un individuo o conjunto de individuos, se utilizan los denominados marcadores informativos de ancestría (AIM), que se caracterizan por tener frecuencias diferenciales entre las poblaciones parentales asumidas (europeos y amerindios en nuestro caso). Siendo p_1 la frecuencia de p en la población parental 1 y p_2 la frecuencia de p en la población parental 2, la diferencia estará dada por $delta = |p_1 - p_2|$. En términos ideales un valor $d = 1$ otorgará información precisa sobre la ancestría de un individuo; mientras menor sea este valor, menor información entregará sobre la ancestría de un individuo con respecto a las dos poblaciones parentales. Se ha establecido 0,3 como un valor arbitrario umbral para la elección de AIM (Collins-Schramm et al., 2002). El primero en realizar un modelo para la estimación de la contribución de las poblaciones parentales a una población híbrida fue Bernstein (1931). Para una población híbrida con dos poblaciones parentales (dihíbrida) la contribución de una de ellas (m_1) será establecida por la relación entre las frecuencias de un alelo (p_1) entre las poblaciones parentales e híbridas: $m_1 = \frac{(p_h - p_2)}{(p_1 - p_2)}$. Para el uso de múltiples loci y múltiples poblaciones, Roberts y Hiorns (1965) propusieron una generalización de este modelo a través de mínimos cuadrados, asumiendo que las frecuencias génicas de la población mestiza son una combinación lineal de aquellas con las que contribuyen las poblaciones parentales al momento del mestizaje. Sin embargo, la mayoría de las veces las frecuencias de las poblaciones parentales e híbridas al momento del mestizaje no son conocidas, si no que inferidas a través de muestras de poblaciones actuales, por lo que la estimación puede ser influenciada

por: (1) error de muestreo; (2) cambios en las frecuencias a través del tiempo por efectos de deriva génica; (3) el grado de diferenciación entre las poblaciones parentales previo al mestizaje; y (4) mutaciones que cambien las frecuencias génicas (Wang, 2003). Modelos más actuales han incorporado estas variables en distintos grados. Long (1991) considera un modelo de mínimos cuadrados que permite estimar tanto el error muestral como el efecto de la deriva génica en la población mestiza, pero sólo el error muestral en las poblaciones parentales. El modelo de Bertorelle y Excoffier (1998) toma en consideración tanto mutaciones como la diferenciación genética entre las poblaciones parentales previo el evento de mestizaje. El modelo de Wang (2003) estima, mediante un modelo de máxima verosimilitud, el error muestral, deriva génica, la diferenciación genética de las poblaciones parentales y mutaciones. Finalmente Chikhi et al. (2001) desarrolla un modelo de verosimilitud basado en coalescencia que toma en cuenta la deriva génica y el error muestral en todas las poblaciones.

2.6. Relevancia

El presente estudio se plantea como una expansión de los estudios anteriores que han postulado la existencia de un CSE en Santiago y otras ciudades de Chile, específicamente para evaluar la H1. Expansión en tres sentidos diferenciadores:

1. Utilizar marcadores neutrales y un mayor número de estos: los estudios anteriores han utilizado los sistemas sanguíneos ABO y Rh, ambos productos génicos que pueden estar sometidos a presiones selectivas. Esa condición no los hace buenos candidatos para evaluar la H1 sobre la existencia de un CSE. Al utilizar marcadores neutrales y un mayor número de estos se disminuye la probabilidad de que otros factores evolutivos (salvo deriva y flujo génico) actúen en la modificación de las frecuencias alélicas y/o genotípicas de una generación a otra.

2. Trabajar con genotipos: es posible visualizar fenómenos que puedan quedar ocultos ante una inspección de solo los fenotipos.
3. Nuevos análisis estadísticos: uso de nuevos análisis de diferenciación poblacional y agrupamientos no desarrollados o masificados en los tiempos de las primeras investigaciones.

Algunos puntos que creemos relevantes de las proyecciones e implicancias que puede tener la siguiente investigación son:

1. El conocimiento de la composición genética, como base de la biología de las poblaciones, puede generar cambios en cómo percibimos nuestra capital desde el punto de vista epidemiológico y de la farmacogenética, en donde es posible que no exista la población de Santiago, si no una diversidad que implica diferencias en torno a la morbilidad y respuestas a droga en distintos estratos de nuestra sociedad, como ha sido propuesto anteriormente (Vanegas et al., 2008).
2. Como estudio antropológico que busca conocer la diversidad de nuestras poblaciones, intentamos aportar al conocimiento de nuestra propia realidad, acotado al tema de la segregación, siempre con la finalidad de la reflexión para la construcción de una sociedad más integrada e inclusiva.

2.7. Formulación del problema, objetivos e hipótesis

La presente memoria de título pretende analizar el fenómeno de estructuración genética mediante el uso de marcadores genéticos autosomales ⁸, para responder la pregunta básica que motiva la investigación, a saber:

⁸Los sets de cromosomas pueden clasificarse en cromosomas sexuales y autosomales, en donde los primeros determinan el sexo biológico del individuo.

¿Existe una estructuración genética en la población de Santiago, de acuerdo al ESE?

De esta forma, tenemos como objetivo general:

- Determinar la presencia de estructuración genética poblacional en Santiago, según el ESE

Para lo cual, tendremos los siguientes objetivos específicos:

- Genotipificar las muestras para 5 marcadores autosomales
- Agrupar las muestras de acuerdo al ESE
- Identificar estructuración genética mediante análisis estadísticos de diferenciación poblacional
- Discutir las implicancias de los resultados, a nivel científico y social, así como evaluar las concepciones epistémicas de éste estudio de antropología genética

Nuestra hipótesis de trabajo es que las frecuencias de los distintos marcadores tendrán diferencias significativas entre los distintos ESE de Santiago, cuya causa serían el flujo génico y la deriva. En otras palabras, nuestra hipótesis es H1.

Capítulo 3

Materiales y métodos

3.1. Muestras

Se obtuvieron muestras de saliva de 200 personas no emparentados residentes de Santiago en 3 instituciones de educación superior: Universidad de Chile, Universidad del Desarrollo y UTEM, así como en eventos del Centro Cultural Pablo de Rokha de Puente Alto. Para controlar la variable ESE, se encuestó a cada individuo sobre los establecimientos de educación primaria y secundaria en los que ha estudiado, y las comunas en las que ha residido su familia al momento de su nacimiento y en la actualidad. Esto se realizó en base a la constatación de que la segregación socio-económica (SSE) en Santiago se da fuertemente en el nivel educacional. En otras palabras, es poco probable encontrar personas de ESE distintos dentro de un mismo tipo de establecimiento educacional. Lo anterior tiene limitaciones metodológicas, sobre todo al no considerar esta escasa probabilidad como relevante, y si consideramos que el muestreo de cada unidad no satisface un tamaño muestral en el que se logre estimar certeramente el parámetro poblacional.

La estimación del ESE se realizó de acuerdo a la variable GRUPO descrita por el Sistema de Medición de Calidad de la Educación (SIMCE), el que se realiza mediante

un análisis de clusters, basado en cuatro variables que indican el nivel socio-económico de las familias de los alumnos que ingresan al establecimiento: (1) nivel educacional de la madre; (2) nivel educacional del padre; (3) el ingreso económico total de un mes en el hogar; y (4) el grado de vulnerabilidad del establecimiento, proporcionado por la JUNAEB ¹. La variable está compuesta por 5 factores, desde A a E, yendo desde el ESE bajo al alto, respectivamente.

En la investigación se resguardó la privacidad de las muestras así como su adecuado uso, sólo con los fines expuestos a los mismos donantes, quienes tuvieron que firmar el consentimiento informado correspondiente, aprobado por el Comité de Ética de la Investigación en Ciencias Sociales y Humanidades de la Universidad de Chile.

3.2. Metodología

Fueron utilizados 5 AIM autosomales disponibles en la literatura (Martinez et al., 2004) ². De esta manera fue posible comparar nuestros resultados con otra muestra mestiza urbana de Latinoamérica (Bueno Aires, Argentina). Dos de ellos se caracterizan por la inserción de un segmento *Alu*, mientras que los restantes tres por el cambio de un nucleótido específico (SNP). En la Tabla 3.1 se puede ver el detalle del tipo de marcador, su ubicación cromosómica y los primers utilizados, mientras que en la Figura 3.1 se observan las frecuencias en poblaciones geográficas distintas. El detalle de las frecuencias se puede ver en la Tabla 1 en los Anexos.

La genotipificación se llevó a cabo en el laboratorio de Antropología Genética en las dependencias de la Facultad de Ciencias Sociales de la Universidad de Chile. Los pasos a seguir fueron los siguientes:

1. Extracción del DNA de las muestras de saliva a partir de un protocolo de sales

¹Junta Nacional de Auxilio Escolar y Becas.

²PV92, WI14319, DR2DI, OCA2 y SB19.3.

Poli- mor- fismo	Tipo de marcador	Ubicación cromosómica ^a	Primers
PV92	Inserción <i>Alu</i>	16q24.1, rs3138523	F: 5' TGA GTT CTC AAC TCC TGT GTG TTA G 3' R: 5' AAC TGG GAA AAT TTG AAG AGA AAG T 3'
WI14319	Transición C/T, <i>RsaI</i>	15q14, rs2862	F: 5' TGT TTT TGA TTG AAG AAA CAT CTC T 3' R: 5' AAT TCT GTA CTG TCC CAC CCC 3'
DR2DI	Transición C/T, <i>BclI</i>	Región intrónica de gen DR2D, 11q23.2, rs1079598	F: 5' CCT CCT CTC CGT TCT CCA 3' R: 5' AGG GCC TTA AAA GTC A 3'
OCA2	Transición sinónima G/A, <i>HaeIII</i>	15q11.2–q12, rs1800404	F: 5' CTT TCG TGT GTG CTA ACT CC 3' R: 5' ACC TCT AGC ATG GTT CTT GGG C 3'
SB19.3	Inserción <i>Alu</i>	19p13.11, rs3138524	F: 5' TCT AGC CCC AGA TTT ATG GTA ACT G 3' R: 5' AAG CAC AAT TGG TTA TTT TCT GAC 3'

Tabla 3.1: Tabla con los nombres de los marcadores, el tipo de marcador, su ubicación cromosómica y el set de primers utilizado. Fuente: Martínez et al. (2004)
a. rs = número dbSNP

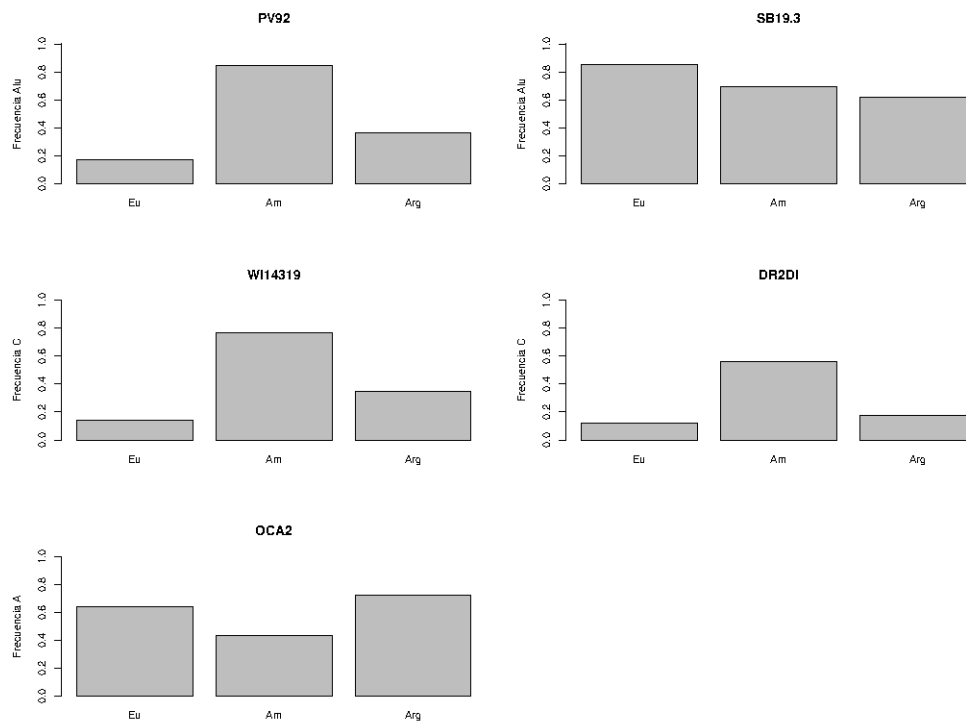


Figura 3.1: Frecuencias génicas de los 5 marcadores utilizados en población amerindia, europea y de una población mestiza urbana de Argentina.

Fuente: Allele Frequency Database (ALFRED) y Martinez et al., 2004.

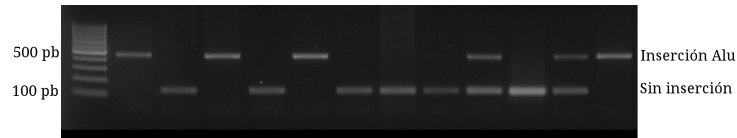


Figura 3.2: Fotografía de gel de agarosa al 1,8 % del marcador PV92 para 12 muestras. Primer carril es el peso molecular

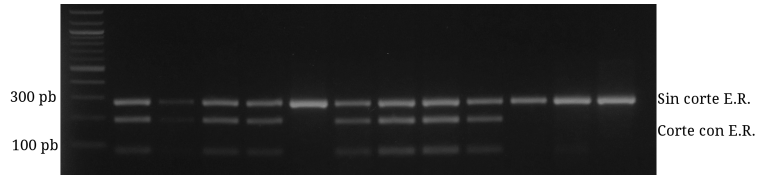


Figura 3.3: Fotografía de gel de agarosa al 2,5 % del marcador Oca2 luego de corte con E.R. para 12 muestras. Primer carril es el peso molecular

(modificado de Quinque et al., 2006).

2. Amplificación del segmento de DNA mediante PCR. Los programas de PCR utilizados se pueden ver en la Tabla 2 en el apartado de Anexos.
3. Para los segmentos de inserción *Alu*, genotipificación mediante electroforesis con gel de agarosa al 1,8 % (Figura 3.2).
4. Para los SNP's, digestión con ER de acuerdo al protocolo establecido, que se puede ver en la Tabla 3 en los Anexos.
5. Genotipificación mediante electroforesis con gel de agarosa al 2,5 % (Figura 3.3).

Los análisis estadísticos utilizados fueron los siguientes:

1. Análisis de frecuencias: se realizaron tablas de contingencia con la variante genética y el ESE para cada marcador, poniendo a prueba la hipótesis de independencia entre las variables a través de un test de ji-cuadrado.

2. Evaluación del supuesto del equilibrio Hardy-Weinberg (EHW): se utilizó un test análogo al test de Fisher, siguiendo una versión modificada del algoritmo de camino aleatorio de cadenas de Markov (Guo y Thompson, 1992). Se realizó para el total de las muestras así como para cada ESE.
3. Desequilibrio de ligamiento (LD): entre pares de loci evaluado usando un test de verosimilitud, cuya distribución de probabilidades es obtenida mediante un procedimiento de permutación (Slatkin y Excoffier, 1996). Se realizó para el total de la muestra.
4. Estructuración poblacional:
 - a) Análisis de la varianza molecular (AMOVA, Excoffier y Lischer, 2010): análisis jerárquico de la varianza que particiona la varianza total en los componentes de la co-varianza producto de diferencias intra-individuales, inter-individuales e inter-poblacionales. Los componentes de la co-varianza son utilizados para computar los estadísticos F . A partir de la definición *a priori* de poblaciones y grupos de poblaciones, se pone a prueba la estructuración genética dada. Se realizó tanto un análisis AMOVA como AMOVA *locus-por-locus*. El primero realiza los análisis a partir de haplotipos, mientras que el segundo lo hace individualmente por cada *locus*. En este último, los distintos componentes de la varianza en distintos niveles jerárquicos son combinados para producir estimadores de los estadísticos F , sumando los componentes de la varianza estimados en el nivel determinado en el numerador y denominador para producir estadísticos F como proporciones de los componentes de la varianza.
 - b) Test exacto de diferenciación poblacional: análogo a la prueba exacta de Fisher en tablas de contingencia, con el uso de un camino aleatorio de cadena de Markov similar a la utilizada en LD, como es descrito por

- Raymond y Rousset (1995). Se estima la probabilidad de encontrar una tabla igual o menos probable que la muestra observada bajo la hipótesis nula de panmixia.
- c) Estadísticos F : como son descritos por Sewall Wright y modificados por Weir y Cockerham (1984) y computados a partir del análisis AMOVA, en donde los componente de la co-varianza son utilizados para determinar los estadísticos F .
- d) STRUCTURE: software que aplica un método basado en un modelo de agrupamiento para datos de genotipos multilocus (Pritchard et al., 2000). Se utilizaron dos modelos con un numero de poblaciones asumidas de $K=1$ a $K=5$. El primero es un modelo de mestizaje, que se basa en que los individuos pueden tener ancestría mixta; y el segundo un modelo de ligamiento, el cual trabaja con desequilibrio de ligamiento de mestizaje, o sea, la asociación que emerge entre *loci* ligados en poblaciones recientemente misceginadas. Para el primero se utilizó la información del ESE de cada individuo como información adicional (modelo LOCPRIOR) para la construcción de los clusters, utilizado generalmente cuando la señal de estructuración es relativamente baja. Los parámetros de duración de quemado (cuanto correr la simulación antes de recolectar datos para minimizar el efecto de la configuración inicial) y de ejecución (cuanto correr el programa luego de la duración de quemado) fueron de 100.000.
5. Estimación de mezcla genética: Se evaluaron 5 modelos distintos. El modelo más simple (Roberts y Hiorns, 1965) asume que las frecuencias génicas de la población mestiza son una combinación lineal de aquellas con las que contribuyen las poblaciones parentales al momento del mestizaje. Long (1991) considera un modelo de mínimos cuadrados que permite estimar tanto el error muestral

como el efecto de la deriva génica en la población mestiza, pero sólo el error muestral en las poblaciones parentales. El modelo de Bertorelle y Excoffier (1998) toma en consideración tanto mutaciones como la diferenciación genética entre las poblaciones parentales previo el evento de mestizaje. El modelo de Wang (2003) estima, mediante un modelo de máxima verosimilitud, el error muestral, deriva génica, la diferenciación genética de las poblaciones parentales y mutaciones. Finalmente Chikhi et al. (2001) desarrolla un modelo de verosimilitud basado en coalescencia que toma en cuenta la deriva génica y el error muestral en todas las poblaciones. Las frecuencias de las poblaciones parentales fueron obtenidas de HapMap (PV92 y SB19.3) y de Ensembl (OCA2 y DR2DI).

Para el análisis 1 se utilizó el software y ambiente de programación estadística R (R Development Core Team, 2011). Para los análisis 2, 3 y 4(a)(b)(c) se utilizó el software Arlequin 3.5 (Excoffier y Lischer, 2010). Para el análisis 4(d) se utilizó el programa STRUCTURE (Pritchard et al., 2000). Para el análisis 5 se utilizaron los programas LEADMIX (Wang, 2003) y LEA (Chikhi et al., 2001).

A partir de los resultados preliminares de la investigación se decidió llevar a cabo una búsqueda en las bases de datos de HapMap y 1000 Genomes por AIM entre las poblaciones asiáticas³ y europeas, es decir, marcadores cuyas frecuencias en ambas poblaciones muestran mayores diferencias. El Proyecto Internacional HapMap es un proyecto colaborativo entre científicos en Japón, Reino Unido, Canadá, China, Nigeria y EE.UU. iniciado oficialmente en octubre de 2002 para ayudar en la identificación de polimorfismos (SNP) asociados a salud, enfermedad y respuestas a drogas y factores ambientales (Altshuler et al., 2010). Dependiendo de la población la base de datos cuenta con una cantidad entre 1 millón y más de 4 millones de SNP. El pro-

³En las bases de datos de HapMap no hay datos con respecto a poblaciones amerindias, por esta razón se utilizaron los datos de las poblaciones asiáticas como *proxy* de las amerindias

yecto 1000 Genomes, también desarrollado en un panorama internacional, iniciado el año 2008 y aún activo, establece objetivos similares al anterior (Abecasis et al., 2010). Dada la completitud del proyecto HapMap, se utilizaron estas bases de datos para realizar la búsqueda inicial, para luego los SNP seleccionados ser corroborados en las bases de datos del proyecto 1000 Genomes. Para lo anterior, se diseñó un programa en lenguaje de programación Python 3.2, cuya función es tomar dos archivos de frecuencias de SNP de poblaciones europeas y asiáticas y, si están presentes en ambas, generar un tercer archivo con ambas frecuencias y la diferencia de éstas. De esta manera, el tercer archivo contendrá el delta o la diferencia de las frecuencias que servirá para discriminar entre los SNP más informativos de ancestría. Las frecuencias alélicas obtenidas fueron de las poblaciones Han en Beijing, China, como representantes de población asiática y residentes de Utah con ancestría del norte y oeste de Europa, como representantes de la población europea.

Capítulo 4

Resultados

4.1. Obtención de datos

Sólo se analizaron 174 individuos de los cuales 12 pertenecen al grupo B, 22 al C, 52 al D y 88 al E. Las restantes 26 se omitieron de la extracción y del análisis final por ser todos pertenecientes al ESE E, que está bien representado. Dentro de la ciudad de Santiago existe un total de 706 establecimientos educacionales, en donde el ESE más representado es el E con aproximadamente 24 %, y el menos representado es el A, con un 12 % (Figura 4.1). La mayoría (94 %) son urbanos. El tipo de establecimiento se asocia significativamente con el ESE estimado ($p = 2,2^{-16}$), en donde el ESE más alto está compuesto en su mayoría (96 %) de establecimientos particulares pagados y el más bajo se compone en su totalidad de establecimientos municipales o particulares subvencionados (datos no mostrados). El muestreo realizado refleja esta distribución, en donde el ESE A fue el único no representado por la dificultad de acceder a los estratos más bajos y la baja representación del GRUPO dentro de Santiago. Los ESE más altos fueron más fáciles de obtener, sobretodo producto de la metodología utilizada para la recolección de las muestras, ya que es más probable que los estratos más altos accedan a la educación superior.

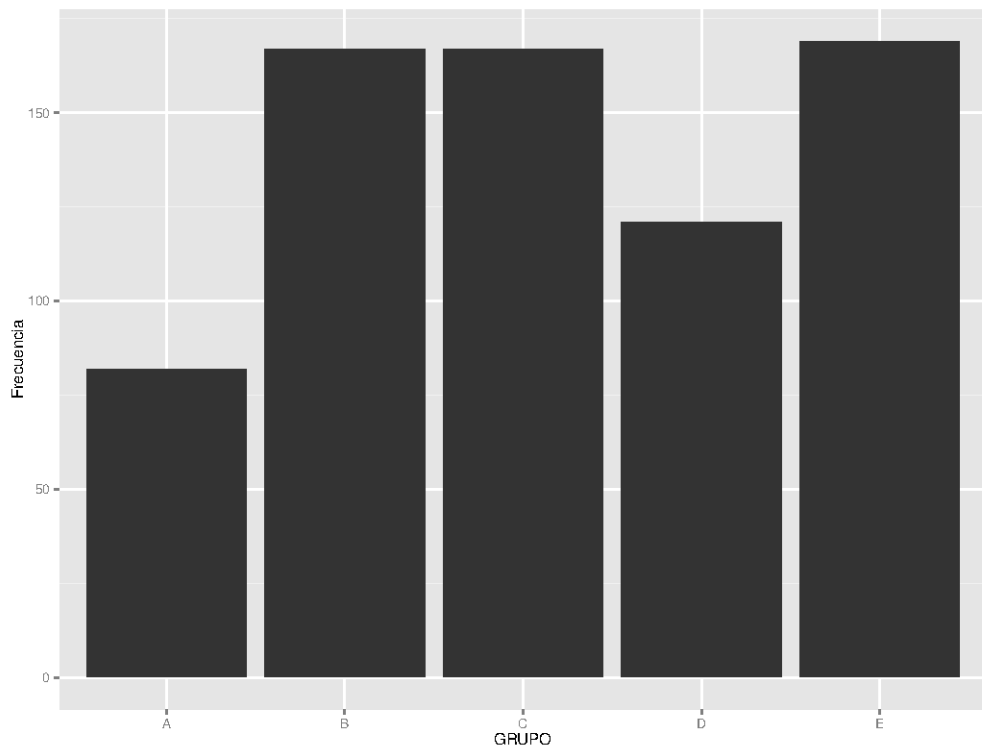


Figura 4.1: Frecuencia de establecimientos en los distintos ESE estimado por el SIMCE el año 2010 en los segundos medio de la ciudad de Santiago

De los 5 marcadores utilizados, finalmente sólo fueron analizados 4 de ellos (PV92, SB19.3, OCA2 y DR2DI), dejando de lado el marcador WI14319 debido a la dificultad de su detección por gel de electroforesis, lo que probablemente condujo a la presencia de un gran porcentaje de falsos positivos y negativos. Por otra parte, al comparar las secuencias de los primers publicados con anterioridad (Martinez et al., 2004) y las bases de datos actuales, se observaron diferencias nucleotídicas que probablemente generaron anomalías en el PCR. Algunas muestras fueron secuenciadas para verificar la presencia de las distintas variantes. Las muestras que se asignaron con la presencia de inserción *Alu* mostraron un segmento de 300 bp, mientras que las que fueron asignados como ausentes no presentaban este segmento. Las variantes SNP se muestran destacadas a continuación, al igual que la diferencia entre los primers publicados y las bases de datos actuales para el marcador WI14319.

Marcador	Muestra	Secuencia
DR2DI	ac09	GCGAATCTGATCATGTGGTTCCTG
	ac11	GCGAATCCGATCATGTGGTTCCTG
	udd04	GCGAATCCGATCATGTGGTTCCTG
	udd27	GCGAATCCGATCATGTGGTTCCTG
	Ensembl ¹	GCGAATCCGATCATGTGGTTCCTG
OCA2	udd30	GAACTCTGGCAGCCATGCTGG
	ac02	GAACTCTTGCAGCCATGCTGG
	udd29	GAACTCTGGCGGCCATGCTGG
	ac03	GAACTCTGGCGGCCATGCTGG
	Ensembl	GAACTCTGGCGGCCATGCTGG
WI14319	Primer Forward	TGTTTTGATTGAAGAAACATCTCT
	dbSNP	TG-TTTTGATTGAAGAAACATCTCT
	Primer Reverse	CCCCACCCTGTCATGTCTTAA
	dbSNP	CCCCACCCTGTCATGTCTTAA

Las frecuencias de los 4 marcadores entre los distintos ESE no mostraron mayores diferencias, sobre todo para los marcadores *Alu*(p\1) (Figura 4.2). Los SNP muestran una leve tendencia hacia un componente genético más europeos en ESE más altos

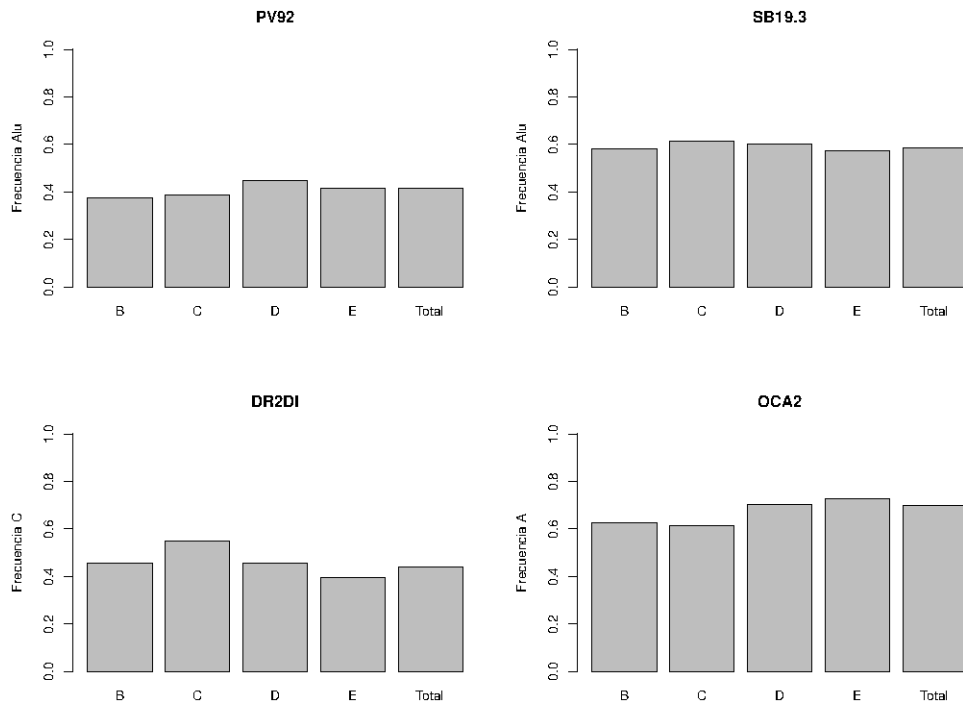


Figura 4.2: Frecuencias alélicas para los 4 marcadores en los distintos ESE y el total de la muestra de Santiago.

y más amerindio en ESE más bajos, no siendo estadísticamente significativo ($p > 0,05$). Todos los marcadores mostraron ser consistentes con una población mestiza (Figura 4.3) con niveles intermedios entre las poblaciones española y amerindia para PV92 y DR2DI, levemente “más amerindio” para SB19.3 y “más europeo” para OCA2. Todas las frecuencias son similares con la población mestiza de Argentina, salvo para el marcador DR2DI en donde Santiago aparece con frecuencias más cercanas a las amerindias.

Los marcadores PV92 y DR2DI muestran un desvío del EHW en el conjunto de la población y en el ESE E. En los ESE C y D solo DR2DI muestra valores significativos, mientras que en el ESE B solo el marcador PV92. Estos desvíos del EHW se deben en su totalidad a una disminución de los heterocigotos (Tabla 4.1)

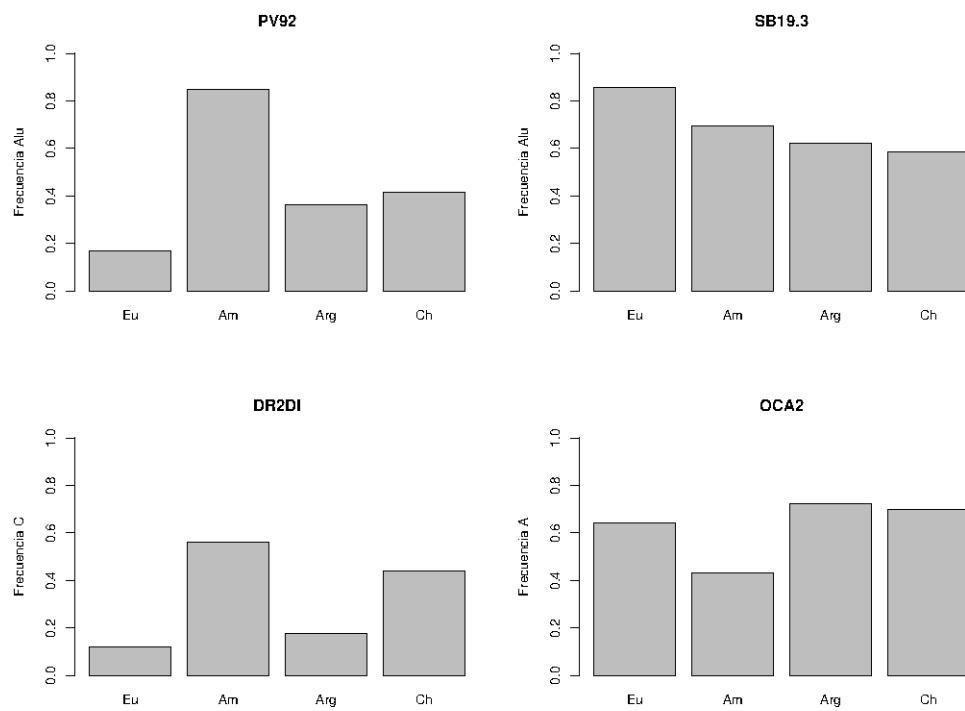


Figura 4.3: Frecuencias alélicas de los 4 marcadores en muestras europeas, amerindias y mestizas urbanas de Argentina y Chile.

Locus	Total			B			C			D			E		
	Esp	Obs	p	Esp	Obs	p	Esp	Obs	p	Esp	Obs	p	Esp	Obs	p
PV92	0,488	0,347	0,000	0,489	0,083	0,006	0,485	0,409	0,653	0,5	0,38	0,097	0,489	0,349	0,012
SB19.3	0,486	0,422	0,135	0,507	0,5	1	0,485	0,5	1	0,485	0,386	0,214	0,493	0,408	0,161
OCA2	0,423	0,395	0,455	0,489	0,417	1	0,485	0,318	0,173	0,423	0,468	0,726	0,398	0,370	0,576
DR2DI	0,495	0,271	0,000	0,518	0,25	0,103	0,508	0,238	0,025	0,502	0,244	0,001	0,481	0,303	0,004

Tabla 4.1: Cuadro con las probabilidades del análisis del EHW en el total de la muestra y particionada por los distintos ESE. Valores significativos destacados en negrilla.

	PV92	SB19.3	OCA2	DR2DI
PV92	-			
SB19.3	0,139	-		
OCA2	0,000	0,044	-	
DR2DI	0,005	0,002	0,002	-

Tabla 4.2: Probabilidades del análisis de LD de a pares por marcador, en la muestra total.

De las 6 comparaciones de a pares entre los marcadores, 5 mostraron ser significativas para LD. El único par que no mostró estar en LD fue el par PV92 y SB19.3 (Tabla 4.2).

4.2. Estructuración genética

Para el análisis AMOVA, el 85,1 % de la diferencia corresponde a la existente entre los distintos individuos, y un 16,53 % corresponde a la diferencia existente entre los individuos dentro de un mismo ESE. Virtualmente nada es explicado por la diferencia entre los grupos ni entre los ESE dentro de un grupo (Tabla 4.3). Por otra parte, para el análisis AMOVA locus-por-locus, el 75,78 % de la variación es explicada por la diferencia entre los individuos, mientras que un 24,87 % es explicado por la diferencia entre los individuos dentro del mismo ESE. Al igual que el AMOVA, nada es explicado por la diferencia entre los grupos ni entre los ESE dentro de un grupo (Tabla 4.4).

Los estadísticos F fueron computados a través del análisis AMOVA y AMOVA locus-por-locus. Para el modelo AMOVA el estadístico F_{st} (excesos de homocigotos explicados por diferencias entre grupos) fue de -0,01624 ($p > 0,05$) y el estadístico F_{is} (exceso de homocigotos explicados por diferencias dentro de los grupos) mostró un valor de 0,163 ($p < 0,05$). Para el modelo AMOVA locus-por-locus, el estadístico F_{st} fue de -0,00661 y el estadístico F_{is} 0,24707. El test de diferenciación genética basado en las frecuencias genotípicas (Raymond y Rousset, 1995) no mostró valores signifi-

Fuente de la variación	Grados de libertad	Suma de cuadrados	Componente de la varianza	Porcentaje de la variación
Entre Grupos	1	0,057	-0,00225	-0,29
Entre ESE dentro de Grupos	2	0,158	-0,01017	-1,33
Entre individuos dentro de ESE	170	153,286	0,12613	16,53
Entre individuos	174	113	0,64943	85,1
Total	347	266,5	0,76313	

Tabla 4.3: Análisis AMOVA con los distintos niveles de partición jerárquica de los elementos de covarianza, los grados de libertad, la suma de cuadrados, el componente de la varianza y el porcentaje de variación de cada uno.

Fuente de la variación	Suma de cuadrados	Componente de la varianza	Porcentaje de la variación
Entre Grupos	1,156	0,00656	0,69321
Entre ESE dentro de Grupos	0,508	-0,01274	-1,34493
Entre individuos dentro de ESE	180,641	0,23549	24,86842
Entre individuos	113	0,71764	75,7833
Total	295,305	0,94696	

Tabla 4.4: Cuadro del análisis AMOVA locus por locus, con los distintos niveles de partición jerárquica de los elementos de covarianza, los grados de libertad, la suma de cuadrados, el componente de la varianza y el porcentaje de variación de cada uno.

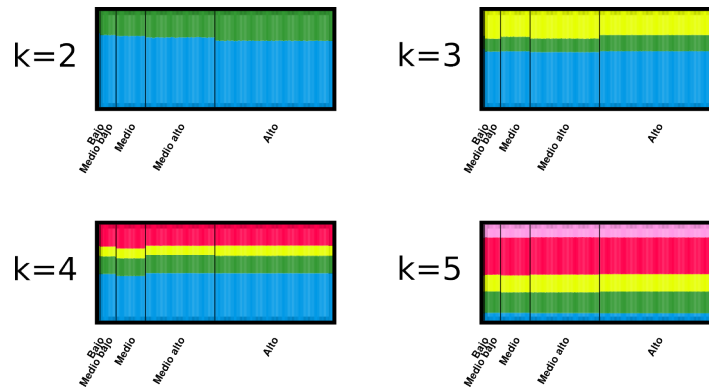


Figura 4.4: análisis de STRUCTURE, con distintos modelos de número poblacional ($K=2$ a $K=5$).

cativos ($p=0,321$, con 100000 pasos de Markov), no pudiendo rechazar la hipótesis nula de no-diferenciación, tampoco al evaluar los ESE de a pares.

Para el análisis de STRUCTURE se utilizaron dos modelos distintos:

- Para el modelo de mestizaje los mayores valores para la estimación del logaritmo natural de los datos, dados diferentes modelos de estructura ($K=1$ a $K=5$), están asociados con un modelo de $K=1$ y algunos pocos se asocian a valores de K mayores. La distribución de los individuos de cada ESE a los clusters inferidos es consistente con lo anterior, mostrando que la proporción de asignación de cada individuo a un cluster es prácticamente igual entre los distintos ESE (Figura 4.4). Esta consistencia indica que el modelo más parsimonioso sería el de $K=1$ o nula estructuración.
- Para el modelo de ligamiento los mayores valores para la estimación del logaritmo natural de los datos, dados diferentes modelos de estructura ($K=1$ a $K=5$), están asociados con un modelo de $K=1$ y algunos pocos se asocian a valores de $K=2$. Consistente con los resultados del modelo anterior (datos no mostrados).

El componente genético europeo fue estimado siguiendo distintos modelos todos bordeando el 50% (Tabla 4.5). Para el estimador de Wang (2003) se puede ver el

Porcentaje de componente genético europeo			
ESE	Roberts y Hiorns (1965)	Long (1991)	Bertorelle y Excoffier (1998)
Total	54,55	54,11	48,18
B	55,19	54,38	47,93
C	54,14	56,59	47,11
D	50,46	50,32	44,85
E	58,05	57,71	51,42

Tabla 4.5: Estimación de proporción de mezcla europea en la muestra de población mixta de Santiago total, y particionada por el ESE, utilizando distintos estimadores.

gráfico de verosimilitud, encontrándose entre un 40 % y 60 % (Figura 4.5), mientras que para el estimador de Chikhi et al. (2001) estaría entre 50 % y 70 % (Figura 4.6)

4.3. Búsqueda de SNPs

El código fuente del programa se puede ver en el Anexo. Se analizaron 3.618.144 SNP's en los 22 cromosomas autosomales de las bases de datos de HapMap, en donde la cantidad de variantes concuerda con el tamaño de los cromosomas (Figura 4.7). Como una prueba inicial, se tomaron los SNP con una diferencia mayor a 0,7 entre ambas poblaciones (AIM), evaluando su frecuencia y su proporción con respecto al número de SNP encontrados. Así, el cromosoma 2 aparece con la mayor cantidad de AIM (506) y el cromosoma 21 con la menor cantidad (16). En términos de proporciones, el cromosoma 18 aparece con el menor valor (0,0001442169) y el cromosoma 15 con el mayor (0,0018285993) (Figura 4.8).

Se seleccionaron los SNP's con mayores diferencias dentro de cada cromosoma y se buscó la variante en las bases de datos de Ensembl para constatar la diferencia a partir de los datos del proyecto 1000 Genomes. A partir de lo anterior, pudimos observar que muchas variantes de HapMap que muestran diferencias cercanas a 1 (es decir, fijación de una variante en alguna de las dos poblaciones) no muestran

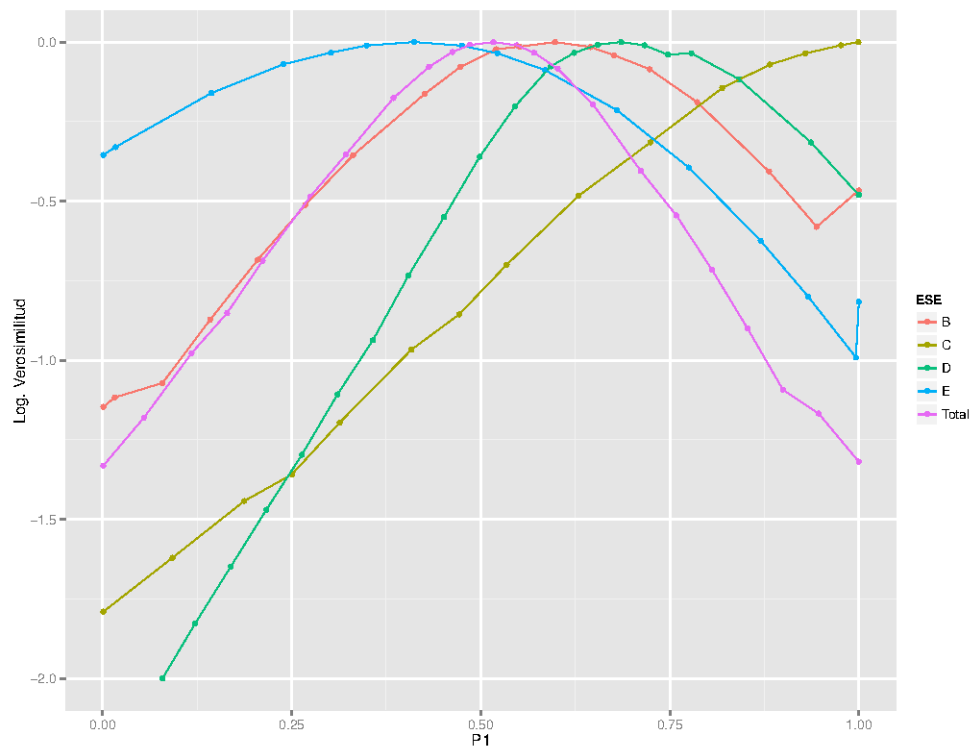


Figura 4.5: Verosimilitud de la estimación del parámetro P1 o proporción de mezcla europea en la población mixta de Santiago, a partir del modelo de Wang (2003). Las diversas líneas representan o bien la muestra total, o las particiones por cada ESE.

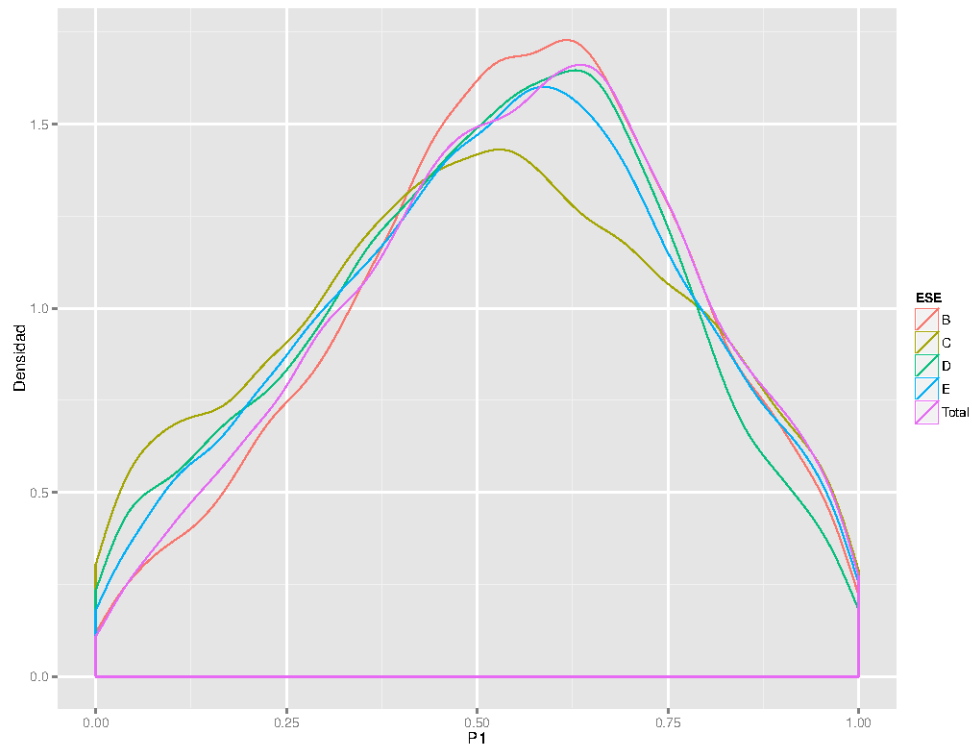


Figura 4.6: Estimación del parámetro P1, o proporción de mezcla europea en la población mixta de Santiago, a partir del modelo de Chikhi et al. (2001). Las diversas líneas representan o bien la muestra total, o las particiones por cada ESE.

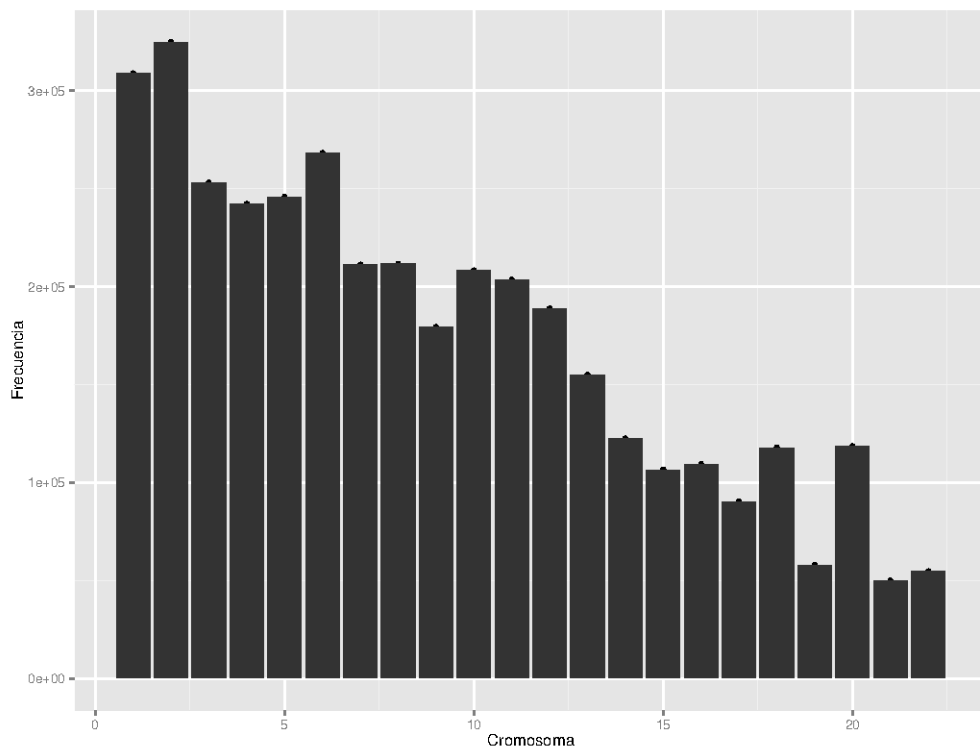


Figura 4.7: Cantidad de SNP analizados por cada uno de los 22 autosomas (en orden correlativo de izquierda a derecha).

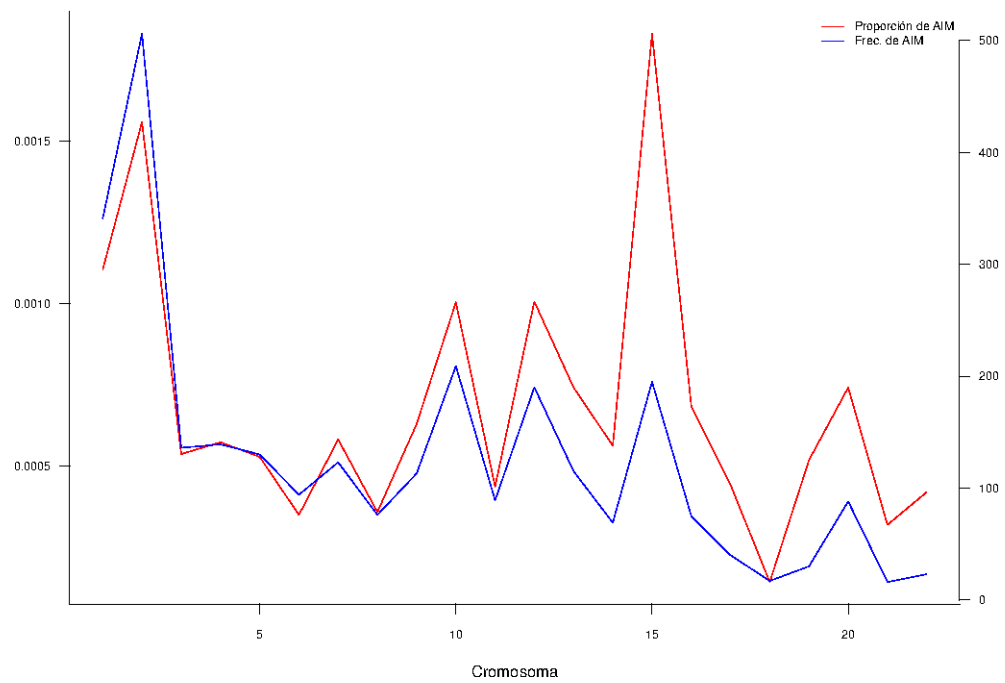


Figura 4.8: Frecuencia y proporción (relativo al tamaño del cromosoma) de AIM en cada autosoma (en orden correlativo de izquierda a derecha).

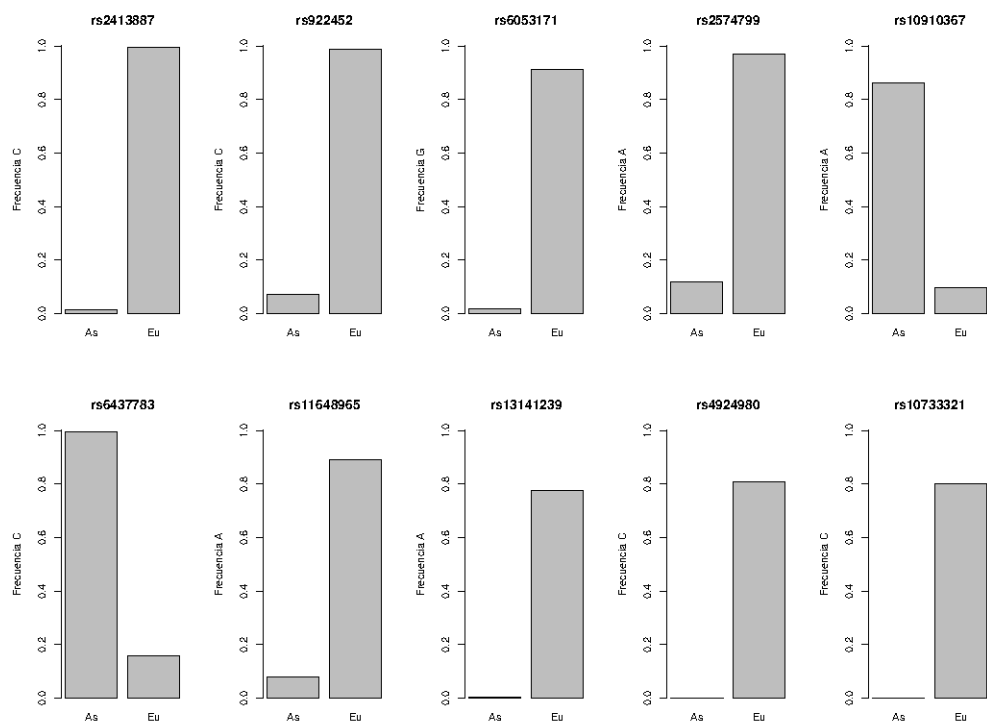


Figura 4.9: Frecuencias alélicas de las variantes de 10 SNP AIM entre poblaciones asiáticas y europeas

mayores variaciones de las frecuencias en las bases de datos de Ensembl, lo que podría indicar un error en alguna de ellas. Finalmente, se promediaron las diferencias encontradas en ambas bases de datos para obtener un promedio de las diferencias. Se seleccionó un SNP de cada cromosoma que cumpliera con los siguientes criterios: que mostrara la mayor diferencia promedio, se encontrara en regiones intergénicas o intrones, que no estuviera ligada a un gen sometido a selección y que sus variantes no tuvieran efectos regulatorios conocidos. De estos, los 10 con mayores diferencias entre poblaciones asiáticas y europeas se muestran en la Figura 4.9. Mayores detalles se pueden ver en la Tabla 4 en los Anexos.

Capítulo 5

Discusión

A partir de la década de los 80 se instala en la literatura el concepto de cline socio-genético estructurado (CSE), en donde componentes genéticos distintos se asocian a distintos componentes sociales. En primera instancia, para explicar esto se genera una hipótesis socio-cultural (H1), en donde se asocia la conformación mestiza (pero desigual entre los distintos estamentos de la sociedad) de la población de Santiago en la conquista castellana, con la mantención de esta desigualdad a través de una disminución del flujo génico; hipótesis que ha sido puesta en duda recientemente (Valenzuela, 2011). Por la misma década de los 80, la ciudad de Santiago vería la profundización de los fenómenos de segregación a partir de las políticas de erradicación y las reformas al sistema educativo llevadas a cabo por la dictadura militar. Una generación ha pasado desde entonces y se hace necesario volver sobre esta hipótesis con las nuevas tecnologías y herramientas estadísticas que nos ofrecen nuestros tiempos.

Las frecuencias alélicas totales son concordantes con lo esperado en una población mestiza, no habiendo diferencias significativas entre los distintos ESE. Sin embargo, mientras las inserciones *Alu* PV92 y SB19.3, establecidos como buenos marcadores para estudios de diferencias poblacionales y de mestizaje (Batzer y Deininger, 2002,

Batzer et al., 1994, ver Sección 2.5) no muestran valores diferentes entre los ESE, los SNP OCA2 y DR2DI muestran una leve diferenciación. Esto puede ser explicado por efecto de mutaciones recurrentes en los SNP analizados que hayan generado esta leve diferencia en las frecuencias, mucho más probable que en inserciones *Alu*, o por error del muestreo, aunque esto cae sólo en el ámbito hipotético. Por otro lado, los marcadores SB19.3 y OCA2, los cuales muestran menos diferencia entre las poblaciones parentales, muestran extremos en la población mestiza santiaguina, con SB19.3 siendo en la población total “más amerindia” y con OCA2 siendo en los ESE medio alto y alto “más europeos”. Lo anterior puede deberse a un efecto fundador en las poblaciones que dieron origen a la actual población mestiza de Santiago, es decir, asumimos frecuencias de las poblaciones parentales a partir de muestras actuales, las que pueden no ser las mismas que en el momento del mestizaje. Por otra parte, la comparación con otra muestra mestiza de América Latina da resultados consistentes, salvo por el marcador DR2DI, en donde la población de Santiago es más cercana a la amerindia. Los marcadores PV92 y DR2DI, que muestran mayor diferencia entre las poblaciones parentales, son los que muestran un desvío del EHW, lo que concuerda con lo esperado, dado que no ha existido el tiempo suficiente para que ambos loci segregen.

No se ha encontrado una clara estructuración genética dentro de la población de Santiago. El análisis mediante STRUCTURE ha mostrado que el modelo más parsimonioso es de $K=1$, consistente con una mayor probabilidad de este modelo y una proporción de asignación de cada individuo a un cluster prácticamente igual entre los distintos ESE (Figura 4.4). El análisis AMOVA locus-por-locus muestra que un 25 % de la variación es explicado por la diferencia existente entre individuos dentro de los ESE que hemos construido y por otro lado, el estadístico F_{is} muestra un valor de 0,25. Los resultados anteriores muestran: (1) existe un exceso de homocigotos aportado fundamentalmente por los dos marcadores con mayor delta entre las

poblaciones parentales (PV92 y DR2DI); (2) este exceso de homocigotos se produce probablemente por el efecto de apareamiento selectivo o la subdivisión de la población total en varios demes; (3) estas subdivisiones no corresponden a los ESE que hemos establecido *a priori*; (4) la diferencia entre individuos del mismo ESE explica el 25 % de la variación total.

La existencia de un CSE en la población de Santiago se ha establecido clásicamente en base a dos marcadores genéticos, los grupos sanguíneos ABO y Rh (Harb, 2004, Valenzuela et al., 1987, Valenzuela y Harb, 1977, Valenzuela, 2011). Recientemente, la H1 ha sido puesta en duda para explicar el comportamiento de las frecuencias en la población chilena (Valenzuela, 2011). En el presente estudio no se ha encontrado una clara estructuración genética, sin embargo, sin constituir resultados concluyentes. Podemos considerar, por ende:

1. H1 verdadera: los estudios con el sistema ABO y Rh han hallado un CSE en varias poblaciones de Chile, siendo en Santiago la más estudiada. El presente estudio, dado su bajo tamaño muestral y la metodología de estratificación, no ha sido capaz de detectarla.
2. H1 falsa: el CSE se ha encontrado sólo en los grupos sanguíneos ABO y Rh debido a causas que no se explican por la hipótesis socio-cultural, si no por una función psico-social que relaciona ambos caracteres (Valenzuela, 2011), o por otros fenómenos selectivos, tanto del mismo sistema ABO y Rh, como de otros genes con los que se encuentre en LD. En marcadores neutrales, el CSE no se logra evidenciar, debido al restringido alcance del cline en sólo algunos *loci*.

Creemos que son necesarias ciertas consideraciones a la hora de abordar la cantidad de datos y posibles hipótesis sobre la estructuración genética en la población de Santiago de acuerdo al ESE, sobre todo con miras a futuros estudios. Muchas de ellas ya han sido señaladas (Valenzuela, 2011):

1. Importancia de los marcadores: dependiendo de las características propias de los marcadores, estudios distintos pueden no ser directamente comparables. En este caso, el uso de marcadores neutrales, de genes o de incluso del propio fenotipo. En el primer caso, se asume una nula o relajada presión selectiva, siendo afectada las frecuencias alélicas por fenómenos como deriva, apareamiento selectivo o migraciones. En el caso de genes, estos pueden estar afectados, en mayor probabilidad e intensidad que los anteriores, a presiones selectivas determinadas. En el último caso, los fenotipos no solo pueden estar afectados por presiones selectivas, si no que puede no ser conocido el genotipo subyacente o, por otro lado, la emergencia del mismo puede no estar asociado directamente con una variante genética específica. Es también de suma importancia que la batería de marcadores pueda llegar a conformar una representación del genoma. Con todo lo anterior, la mejor solución para llegar a conclusiones definitivas sobre el tema sería el uso exclusivo de marcadores neutrales AIM y una mayor cantidad de ellos. La búsqueda en bases de datos internacionales (HapMap y Ensembl) ha mostrado la existencia de buenos marcadores AIM debido a su gran diferencia entre población asiática y europea, su ubicación en el genoma en regiones intergénicas o intrones y el nulo (o hasta el momento no identificado) efecto regulatorio de sus variantes (Tabla 4 en Anexos y Figura 4.9). Lo anterior debe tomarse con precaución debido a que, por una parte, el concepto de población asiática y europea conlleva una minimización de la variación interna de cada agrupación y, por otra, la población amerindia ancestral no puede ser directamente comparable con la población asiática actual, ni tampoco la población europea actual con la ancestral. Sin embargo, y a pesar de lo anterior, sólo el análisis mas exhaustivo de nuestra propia población puede dar la razón o no a los lineamientos trazados anteriormente. Los marcadores utilizados en la presente investigación tienen diferencias entre las poblaciones parentales que

van de 0,318 (SB19.3) a 0,582 (PV92), siendo ideal para obtener información más precisa sobre la ancestría de un individuo, una diferencia cercana a 1. En este sentido, la elección de marcadores pudo no haber sido la adecuada. Cabe la pregunta, sin embargo, de por qué se ha podido encontrar un CSE con el sistema sanguíneo ABO, en donde $d=0,35$ y no con los marcadores actuales.

2. Metodología de estratificación: de importancia clave, ya que a través de ella se construyen los grupos para corroborar su significación biológica. Es, por ende, necesario saber que se está midiendo con la estructuración que se realiza; estratificar en términos de consumo no es lo mismo que en términos estructurales como el poder o prestigio, entre otros. En la presente investigación se ha optado por una forma de estratificación gradacional, que da cuenta sobre todo de aspectos de consumo e ingresos. Como se mencionó anteriormente, los esquemas de clases toman en consideración aspectos estructurales y relacionales de los individuos dentro de un sistema social particular, y creemos que son estas las que pueden dar mejores indicios sobre las relaciones de reproducción entre los individuos de una sociedad. Indicadores de estratificación de tipo gradacional, que no den cuenta de la estructura de las relaciones, si no de agregados de una variable cuantificable, pueden no dar luces sobre la existencia o no de un CSE. Lo anterior, sin embargo, fue subsanado por el uso de construcciones *a posteriori* de agrupaciones, implementadas en el software STRUCTURE y de análisis de correspondencia (datos no mostrados), en donde no se obtuvieron agrupamientos clarificadores y en donde los ESE parecen actuar como una población más o menos homogénea.
3. Tamaños muestrales: dado el gran tamaño de la población residente en Santiago es necesario realizar un estudio lo más representativo posible del mismo. La estrategia de muestreo es también crucial, ya que al reducir el número de

lugares muestreados, se aumenta la probabilidad de muestreo sesgado.

Estudios anteriores (Arcos-Burgos et al., 2004) han estimado porcentajes de mezcla europea mayores en la población de Santiago, mientras que en el presente estudio estos valores bordean el 50 %, indistintamente en cada ESE y con diversos métodos de estimación. Ahora bien, las frecuencias de los distintos marcadores se asemejan a las frecuencias obtenidas de población urbana mestiza de Buenos Aires, Argentina, salvo por el marcador DR2DI que muestra valores más amerindios en la población de Santiago. El uso de una mayor cantidad de marcadores y nuevos modelos que toman en cuenta la deriva y el error de muestreo dan a los resultados actuales una mayor solvencia, pese a su reducido tamaño muestral. Por otro lado, el sistema sanguíneo ABO, el cual se ha utilizado en muchas estimaciones de ancestría, puede estar afectado por fenómenos selectivos, lo cual lo hace un mal marcador para estimaciones de ancestría.

Para explicar estos resultados, tomamos en consideración que ya para febrero de 1541, a sólo un mes de la llegada de la segunda expedición castellana, tanto españoles como amerindios convivían en un mismo espacio y tiempo. Esta conjunción se dará a lo largo de la historia de la ciudad, ya que dentro de los límites urbanos siempre residirán no sólo las élites, si no también los estratos más bajos, elementos de suma importancia dentro del sistema económico-productivo (de Ramón, 2007). Tomando en consideración los antecedentes recopilados, el flujo génico probablemente haya sido mayor del que suponemos en ESE medios y bajos, mientras el extremo superior, no captado por este estudio ¹, probablemente sea el de mayor diferenciación. El hecho de una ciudad segregada no contradice en absoluto estos resultados, si no que otorga otro contexto en donde los distintos componentes sociales han convivido dentro de los márgenes de la ciudad e inevitablemente su interacción ha tenido resultados con-

¹Considerando que la mayor concentración de la riqueza en Chile, es por parte de un porcentaje muy marginal de la población

cretos, una interacción probablemente instalada dentro de una lógica de ilegitimidad y de no reconocimiento de la descendencia. La negación del mestizaje tendría directa relación con el fenómeno de “blanqueamiento” de la población mestiza chilena, con la creación de un imaginario de las clases altas en donde se niega esta unión reproductiva y un imaginario de las clases medias y bajas desde donde emerge la idea del “blanqueamiento” (Waldman, 2004). No es raro que hipótesis como la existencia de un CSE, por un lado, y un alto componente genético europeo, por el otro, no sean extrañas ni alejadas de la realidad para el sentido común de muchas personas. Este imaginario doble debe ser evaluado ante la evidencia de: (1) no disminución del flujo génico entre los distintos ESE; y (2) mayor componente genético amerindio en la población de Santiago en general.

Finalmente creemos necesario discutir sobre los temas de trasfondo que implica o que podría llegar a implicar este estudio. Siguiendo a Duster (2005) se hace urgente reflexionar sobre las categorías que aplicamos a las poblaciones. Sobre todo en genética y en cualquier investigación sobre la variación de la biología de las poblaciones, el concepto de raza o categorías similares siempre rondan nuestras formas de clasificar. Lo complejo de esto es caer en lo que se denomina la falacia de reificación o la falacia de la concreción fuera de lugar, en donde se trata a una entidad abstracta como un evento real o una entidad concreta (Duster, 2005). Esta falacia es tal en el momento que la categorización tipológica cae sobre un grupo de individuos. Estas abstracciones contienen, en términos del proceso investigativo, los siguientes pasos: (1) la búsqueda de una categorización biológicamente significativa; (2) las particularidades de estas categorías en contraposición de las demás, en términos de promedios y no de dispersión (o sea, de variación); y (3) formas de actuar distintas para cada una de ellas (ya sea en políticas públicas y legislación, entre otras). El punto (3) finalmente es donde ocurre la falacia de reificación, mientras que los puntos anteriores refieren a la forma tipológica de clasificar. La idea de un CSE contribuye a la idea

tipológica de estratos socio-genéticos (ESG) asociando variables sociales y biológicas en un conjunto, en donde la variación interna se deja de lado. El último paso es lo que se ha propuesto en Chile de un Marco de Referencia Sociogenético para estudios en Salud Pública como una para-propuesta al uso de variables como etnia y raza (Vanegas et al., 2008), en donde los distintos ESE debieran tener distintos patrones de morbilidad e inmunidad, entre otras.

Es necesario recordar que la narrativa científica tiene un lugar privilegiado dentro de la sociedad y, por ende, su actividad tiene una responsabilidad mayor de la que hacerse cargo. A lo menos dos fenómenos pueden llegar a ser consecuencia de la reificación de los ESG, ya que al identificar una variable social (ESE) con una biológica (componente genético) es posible: (1) caer en la “biologización” de un concepto social, en donde las diferencias sociales pasen a ser explicadas por diferencias biológicas; y (2) agregar al imaginario de las diferencias sociales, las diferencias biológicas. La responsabilidad científica se hace patente en el momento del reconocimiento de una compleja interacción entre fenotipos y prácticas sociales particulares asociadas contra ese fenotipo (independiente de su valoración), las cuales también pueden actuar en la emergencia de patologías y de patrones de morbilidad e inmunidad en los conjuntos. En otras palabras, las patologías no sólo emergen por propiedades biológicas particulares, si no por ambientes diferenciales, accesos diferenciales a los sistemas de salud, estigmatizaciones y expectativas, entre otros. Creemos que la “biologización” de los ESE y la suma de características biológicas a las diferencias sociales tienen directa relación en como se entiende la variación humana, es decir, con una visión tipológica de los conjuntos. Por otro lado, creemos también que existe un costo social de la tipologización, como la creación de nociones que rescaten los promedios y no las dispersiones. Finalmente, que nos puede decir uno o cinco *loci* respecto de los patrones de morbilidad o de inmunidad de un conjunto: probablemente muy poco.

En la presente investigación el riesgo que corremos en tipologizar biológicamente

los distintos ESE de la población es enorme en términos de estigmatización. El problema, al igual que la raza, no es su existencia o no, si no su utilización social para la discriminación o lo que ha sido denominado como antiracial y antiracismo; mientras unos se preocupan de la destrucción del concepto, otros de su uso como factor discriminatorio. Estamos lejos de proponer una solución, más allá de la generación de mayor reflexión sobre los alcances que cualquier investigación científica pueda tener. Tanto si tratamos con restos esqueléticos o saliva, nuestra posición dentro de un sistema valórico que confiere a la narrativa científica un lugar privilegiado, nos hace interlocutores válidos, muchas veces, para aquellos cuerpos que no tienen voz.

Capítulo 6

Conclusiones

En Chile se ha postulado la existencia de un cline socio-genético estructurado (CSE) en donde distintos estratos sociales poseen distintos componentes genéticos. En una primera instancia se propuso una hipótesis socio-cultural (H1) para explicar este fenómeno, en donde las poblaciones parentales que darían origen a la actual población mestiza chilena habrían ocupado distintos segmentos de la sociedad. La gran cantidad de barreras entre estos segmentos habría provocado la mantención de esta diferencia a lo largo de las generaciones. En la presente investigación nos propusimos abordar esta hipótesis con metodologías que nos parecen más adecuadas.

Los resultados actuales no muestran diferencias significativas en las frecuencias de los marcadores de acuerdo al estrato socio-económico (ESE), refutando así la H1. Sin embargo esto no constituye resultados concluyentes debido al bajo número muestral ($n=174$), la nula representación del ESE bajo y la poca diferencia de las frecuencias entre las poblaciones parentales. Hemos podido trazar los lineamientos generales para continuar con las investigaciones, identificando una buena cantidad de marcadores SNP AIM que pueden ser puestos a prueba en la población mestiza chilena. Por otro lado, se ha estimado el componente genético amerindio en 50%, un valor más alto de los encontrados en estudios anteriores. Por otro lado, hemos

trazado someramente las complejidades que emergen en la auto-identificación de una población respecto a sus orígenes y por otro lado, la complejidad de la labor científica a la hora de generar tipologías biológicas.

Anexos

Muestra	N	PV92 <i>Alu</i>	SB19.3 <i>Alu</i>	WI14319*C	DR2DI*C	OCA2*A
Promedio nativos americanos ^a	110	0,714	0,509	0,736	0,614	0,391
Promedio europeo ^a	133	0,132	0,827	0,214	0,124	0,756
La Plata, Argentina ^b	87	0,362	0,621	0,345	0,178	0,724

Tabla 1: Frecuencias poblaciones de los 5 marcadores utilizados en amerindios, europeos y una población mestiza urbana de Argentina.

a. Obtenido de Allele Frequency Database (ALFRED)

b. Obtenido de Martinez et al. (2004)

	Tiempo	Denaturalización		Alineamiento	Elongación	
		5 min	30 seg	30 seg	30 seg	7 min
Partidor	PV92			54,85 °C		
	SB19.3			54,2 °C		
	OCA2	94 °C	94 °C	56,45 °C	72 °C	72 °C
	WI14319			54,05 °C		
	DR2DI			51 °C		
				35 ciclos		

Tabla 2: Información con la temperatura, tiempo y cantidad de ciclos de los programas de PCR utilizados para cada marcador.

ER	PCR	Volúmenes (μ)			ER	Total	Temp. óptima
		ddH2O	Buffer 10x	BSA			
<i>Hae III</i>	10	3.1	1.5	0.1	0.3	15	37°C
<i>Bcl I</i>	10	3.2	1.5	-	0.3	15	55°C
<i>Rsa I</i>	10	3.2	1.5	-	0.3	15	37°C
Fast Digest	10	3.2	1.5	-	0.3	15	37°C

Tabla 3: Información con el protocolo para corte con E.R. con los volúmenes de los reactivos y la temperatura de incubación

SNP	Crms	Delta Map	Hap- sembl	En- Promedio	Ubicación
rs2413887	15	0,989	0,98	0,9845	intron CTXN2
rs922452	2	0,989	0,9	0,9445	intron EDAR
rs6053171	20	0,918	0,89	0,904	intergenica
rs2574799	10	0,868	0,85	0,859	intron KCNMA1
rs10910367	1	0,838	0,86	0,849	intron SLC35F3
rs6437783	3	0,839	0,84	0,8395	intron MYH15
rs11648965	16	0,812	0,82	0,816	intergenica
rs13141239	4	0,857	0,76	0,8085	intergenica
rs4924980	17	0,805	0,81	0,8075	intron EPN2
rs10733321	9	0,808	0,8	0,804	intron BNC2

Tabla 4: Tabla con los 10 SNP seleccionados. Se muestra su ubicación cromosómica y su ubicación específica, la diferencia (Delta) entre poblaciones europeas y asiáticas a partir de las bases de datos de HapMap y Ensembl y el promedio entre ambas.

```
Codigo fuente Programa
```

```
'''
```

```
Created on 31/08/2012
```

```
@author: Tomas B. Gonzalez Z.
```

```
'''
```

```
import os
```

```
print('Programa para evaluar diferencias alelicas  
entre 2 poblaciones \n')
```

```
#ingresar primer archivo y su separador
```

```
nombre = input('Ingrese nombre del primer archivo: ')
```

```
separador = input('Ingrese el separador de las
```

```

columnas del primer archivo: ‘)

#Verificar si existe el fichero
if os.path.exists(nombre):
    print(‘Carga exitosa ‘)
else: #Si no existe, aparece siguiente linea
    print(‘Siga participando ‘)

#ingresar segundo archivo
nombre2 = input(‘\nIngrese nombre del primer archivo: ‘)
separador2 = input (‘Ingrese el separador de las
                    columnas del primer archivo: ‘)

#Verificar si existe el segundo fichero
if os.path.exists(nombre2):
    print(‘Carga exitosa ‘)
else: #Si no existe, aparece siguiente linea
    print(‘Siga participando ‘)

#Abrir el archivo por columnas
with open(nombre) as f1, open(nombre) as f2:
    a = {}
    for line in f1:
        a_snp, a_freq = line.split(separador)
        a[a_snp] = a_freq
    b = {}
    for line in f2:
        b_snp, b_freq = line.split(separador)
        b[b_snp] = b_freq

#Se escribe el primer archivo (output.txt) con los snp’s
# coincidentes y sus dos frecuencias poblacionales
with open(‘output.txt’, ‘w’) as out1:
    for outline in [a[0]+‘\t’+a[1]+‘\t’+b[1] \
                    for a in [line.strip(‘\n’).split(‘ ‘) \
                    for line in open(‘a.txt’, ‘r’).readlines()] \
                    for b in [line.strip(‘\n’).split(‘ ‘) \
                    for line in open(‘b.txt’, ‘r’).readlines()] \
                    if a[0] == b[0]]:

```

```
        out1.write(outline+'\n')

#Verificar si existe el archivo output.txt
if os.path.exists('output.txt'):
    input('\nFichero output.txt creado, presione enter
          para cerrar')
else: #Si no existe, aparece siguiente linea
    input('\nOcurrio un error, presione enter
          para cerrar')

#Se escribe el archivo final, con la interseccion y
#      la resta de las frecuencias de ambas poblaciones
#Se lee el archivo creado antes, se restan columnas
#      y se genera nuevo archivo
with open('output.txt', 'r') as finput, open('outputf.txt',
        'w') as output:
    for linea in finput:
        columnas = linea.strip().split('\t')
        snp = columnas[0]
        freq1 = columnas[1]
        freq2 = columnas[2]
        delta = (float(columnas[1]) - float(columnas[2]))
        output.writelines(snp+'\t'+freq1+'\t'
                          +freq2+'\t'+str(delta)+'\n')

finput.close()
output.close()

#Se comprueba si archivo final fue creado
if os.path.exists('outputf.txt'):
    input('\nFichero outputf.txt creado, presione
          enter para cerrar')
else: #Si no existe, aparece siguiente linea
    input('\nOcurrio un error, presione
          enter para cerrar')
```

Bibliografía

G R Abecasis, D Altshuler, A Auton, L D Brooks, R M Durbin, R A Gibbs, M E Hurles, y G A McVean. A map of human genome variation from population-scale sequencing. *Nature*, 467(7319):1061–73, 2010.

M Acuña, L Eaton, L Cifuentes, y D Massardo. Genetic variants in the enhancer region of the thymidylate synthase gene in the Chilean population. *British journal of clinical pharmacology*, 61(6):778–782, 2006.

C Agostini. Pobreza, desigualdad y segregación en la Región Metropolitana. *Estudios Públicos*, (117):219–268, 2010.

D M Altshuler, R A Gibbs, L Peltonen, E Dermitzakis, S F Schaffner, F Yu, P E Bonnen, P I W de Bakker, P Deloukas, S B Gabriel, R Gwilliam, S Hunt, M Inouye, X Jia, A Palotie, M Parkin, P Whittaker, K Chang, A Hawes, L R Lewis, Y Ren, D Wheeler, D M Muzny, C Barnes, K Darvishi, M Hurles, J M Korn, K Kristiansson, C Lee, S A McCarroll, J Nemes, A Keinan, S B Montgomery, S Pollack, A L Price, N Soranzo, C Gonzaga-Jauregui, V Anttila, W Brodeur, M J Daly, S Leslie, G McVean, L Moutsianas, H Nguyen, Q Zhang, M J R Ghorri, R McGinnis, W McLaren, F Takeuchi, S R Grossman, I Shlyakhter, E B Hostetter, P C Sabeti, C A Adebamowo, M W Foster, D R Gordon, J Licinio, M C Manca, P A Marshall, I Matsuda, D Ngare, V O Wang, D Reddy, C N Rotimi, C D Royal,

- R R Sharp, C Zeng, L D Brooks, y J E McEwen. Integrating common and rare genetic variation in diverse human populations. *Nature*, 467(7311):52–8, 2010.
- M Arcos-Burgos, P Herrera, J Pandey, y C Valenzuela. Análisis de mezcla genética. En Francisco Rothhammer y Elena Llop, eds., *Poblaciones chilenas Cuatro décadas de investigaciones bioantropológicas*, cap. 15, págs. 241–267. Editorial Universitaria, Santiago, primera ed^{ón}., 2004.
- R Atria. *Estructura ocupacional, estructura social y clases sociales*. CEPAL, 2004.
- M A Batzer y P L Deininger. Alu repeats and human genomic diversity. *Nature reviews. Genetics*, 3(5):370–9, 2002.
- M A Batzer, M Stoneking, M Alegria-Hartman, H Bazan, D H Kass, T H Shaikh, G E Novick, P A Ioannou, W D Scheer, R J Herrera, y P L Deininger. African origin of human-specific polymorphic Alu insertions. *Proceedings of the National Academy of Sciences of the United States of America*, 91:12288–12292, 1994.
- J A Beardmore y F Karimi-Booshehri. ABO genes are differentially distributed in socio-economic groups in England. *Nature*, 303(5917):522–524, 1983.
- C Bellei. Expansión de la educación privada y mejoramiento de la educación en Chile. Evaluación a partir de la evidencia. *Pensamiento Educativo*, 40(1):285–311, 2007.
- F. Bernstein. Die geographische Verteilung der Blutgruppen und ihre anthropologische Bedeutung. En *Comitato italiano per lo studio dei problemi della popolazione*, págs. 227–243. Istituto poligrafico dello stato, Roma, 1931.
- G Bertorelle y L Excoffier. Inferring admixture proportions from molecular data. *Molecular biology and evolution*, 15(10):1298–311, 1998.

- F Campillo. Estudio de los grupos sanguíneos en la población española. *An Real Acad Nac Med (España)*, Tomo XCIII(Cuad 3):1–22, 1976.
- L L Cavalli-Sforza y M W Feldman. The application of molecular genetic approaches to the study of human evolution. *Nature genetics*, 33 Suppl(march):266–75, 2003.
- L L Cavalli-Sforza, P Menozzi, y A Piazza. *The History and Geography of Human Genes*. Princeton, New Jersey, 1 ed^{ón}., 1994.
- L Chikhi, M W Bruford, y M A Beaumont. Estimation of admixture proportions: a likelihood-based approach using Markov chain Monte Carlo. *Genetics*, 158(3):1347–62, 2001.
- L Cifuentes, R Morales, D Sepúlveda, H Jorquera, y M Acuña. DYS19 and DYS199 loci in a Chilean population of mixed ancestry. *American journal of physical anthropology*, 125(1):85–89, 2004.
- P Cohn. *Estructura genética de una población urbana. El caso de Valdivia. Chile*. Tesis de magister en ciencias en ecología, Universidad Austral, 1983.
- H E Collins-Schramm, C M Phillips, D J Operario, J S Lee, J L Weber, R L Hanson, W C Knowler, R Cooper, H Li, y M F Seldin. Ethnic-difference markers for use in mapping by admixture linkage disequilibrium. *American journal of human genetics*, 70(3):737–50, 2002.
- V De la Maza. *Un estudio exploratorio sobre la homogamia educacional en Chile*. Seminario, Universidad de Chile, 2008.
- C A De Mattos. Santiago de Chile de cara a la globalización: ¿Otra ciudad? *Revista de Sociología e Política*, (19):31–54, 2002.
- A de Ramón. *Santiago de Chile. Historia de una sociedad urbana*. Catalonia Ltda, Santiago, 1 ed^{ón}., 2007.

- T Duster. Medicine. Race and reification in science. *Science (New York, N.Y.)*, 307(5712):1050–1, 2005.
- L Excoffier y H Lischer. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3):564–567, 2010.
- S Guo y E Thompson. Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles. *Biometrics*, (48):361–372, 1992.
- M Hamilton. *Population Genetics*. Wiley-Blackwell, primera ed^{ón}., 2009.
- Z Harb. Grupos sanguíneos, enzimas eritrocitarias y proteínas séricas. En Francisco Rothhammer y Elena Llop, eds., *Poblaciones chilenas Cuatro décadas de investigaciones bioantropológicas*, cap. 9, págs. 153–164. Editorial Universitaria, Santiago, primera ed^{ón}., 2004.
- D L Hartl. *A Primer of Population Genetics*. Sinauer Associates, Sunderland, Massachusetts, 1981.
- INE. *Retratos de nuestra identidad: Los Censos de Población en Chile y su evolución histórica hacia el Bicentenario*. Santiago, 1 ed^{ón}., 2009.
- M L Juárez. Segregación urbana y sus implicaciones en las ciudades. Una aproximación teórica. *Palapa*, 1(2):45–50, 2006.
- O Larrañaga y C Sanhueza. Residential Segregation Effects on Poor’s Opportunities in Chile, 2007.
- X Leiva. *Efectos de las migraciones recientes en la composición genética de la población de Santiago de Chile*. Tesis Doctoral, Universidad de Chile, 2010.

- J C Long. The genetic structure of admixed populations. *Genetics*, 127(2):417–28, 1991.
- V L. Martinez, B Bertoni, E J Parra, y N O Bianchi. Characterization of Admixture in an Urban Sample from Buenos Aires, Argentina, Using Uniparentally and Biparentally Inherited Genetic Markers. *Human Biology*, 76(4):543–557, 2004.
- R Mellafe. *La introducción de la Esclavitud Negra en Chile*. 1959.
- J Mielke, L Konigsberg, y J Relethford. *Human Biological Variation*. Oxford University Press, Oxford, primera ed^{ón}., 2006.
- A Mizala y F Torche. Bringing the schools back in: the stratification of educational achievement in the Chilean voucher system. *International Journal of Educational Development*, 32(1):132–144, 2012.
- OECD. Education at Glance. Inf. téc., OECD Publishing, 2011. doi: <http://dx.doi.org/10.1787/eag-2011-en>.
- A Orellana y L Fuentes. El arco, el cono y los clusters: geometrías espaciales para la gobernabilidad metropolitana y local de Santiago. En G. Yáñez, A. Orellana, O. Figueroa, y F. Arenas, eds., *Ciudad, poder y gobernanzas*, págs. 111–132. Editorial LOM, Santiago, 2008.
- J Ortiz y E Aravena. Migraciones intraurbanas y nuevas periferias en Santiago de Chile: Efectos en la sociogeografía de la ciudad. *GeoFocus*, (2):49–60, 2002.
- G Pinto, P Ilic, L Paredes, T González, F Marinovic, y Carlos Valenzuela. Frecuencias fenotípicas y génicas de los sistemas sanguíneos ABO y Rhesus en Clínica Alemana (Santiago, Chile). *Revista médica de Chile*, 109:1209–1211, 1980.
- J. Pinto-Cisternas, H. Figueroa, B. Lazo, C. Salinas, y C. Campusano. Genetic Structure of the Population of Valparaiso. *Human Heredity*, 21(5):431–439, 1971.

- J K Pritchard, M Stephens, y P Donnelly. Inference of population structure using multilocus genotype data. *Genetics*, 155(2):945–59, 2000.
- D Quinque, R Kittler, M Kayser, M Stoneking, y I Nasidze. Evaluation of saliva as a source of human DNA for population and association studies. *Analytical biochemistry*, 353(2):272–7, 2006.
- R Development Core Team. R: A Language and Environment for Statistical Computing. 2011.
- M Raymond y F Rousset. An Exact Test for Population Differentiation. *Evolution*, 49(6):1280–1283, 1995.
- D F Roberts y R W Hiorns. Methods of Analysis of the Genetic Composition of a Hybrid Population. *Human biology*, 37:38–43, 1965.
- P Rocco, C Morales, M Moraga, J F Miquel, F Nervi, E Llop, P Carvallo, y F Rothhammer. Composición genética de la población chilena: Distribución de polimorfismos de DNA mitocondrial en grupos originarios y en la población mixta de Santiago. *Revista médica de Chile*, 130(2):125–131, 2002.
- J Rodríguez. Segregación residencial socioeconómica: ¿Qué es?, ¿Cómo se mide?, ¿qué está pasando?, ¿importa? *CEPAL*, 16, 2001.
- J Rodríguez. Dinámica sociodemográfica metropolitana y segregación residencial: ¿qué aporta la CASEN 2006? *Revista de geografía Norte Grande*, (41):81–102, 2008.
- R Rubicz, P Melton, y M H Crawford. Molecular Markers in Anthropological Genetic Studies. En Michael H. Crawford, ed., *Anthropological Genetics. Theory, Methods and Applications*, cap. 6, págs. 141–186. Cambridge University Press, Cambridge, New York, primera edición., 2007.

- F Sabatini, G Cáceres, y J Cerda. Segregación residencial en las principales ciudades chilenas: Tendencias de las tres últimas décadas y posibles cursos de acción. *EURE (Santiago)*, 27(82):21–42, 2001.
- B D Serrano. *Homogamia Educacional en Chile: Influencia Educacional Paterna y Tendencia en el Tiempo*. Tesis Doctoral, Universidad de Chile, 2010.
- M Slatkin y L Excoffier. Testing for linkage disequilibrium in genotypic data using the Expectation-Maximization algorithm. *Heredity*, 76(4):377–383, 1996.
- F Torche. Unequal But Fluid: Social Mobility in Chile in Comparative Perspective. *American Sociological Review*, 70:422–450, 2005.
- F Torche. Educational assortative mating and economic inequality: a comparative analysis of three latin american countries. *Demography*, 47(2):481–502, 2010.
- C Valenzuela, M Acuña, y Z Harb. Gradiente Sociogenetico en la población chilena. *Revista médica de Chile*, 115:295–299, 1987.
- C Valenzuela y Z Harb. Socioeconomic Assortative Mating in Santiago, Chile: a demonstration using stochastic matrices of mother-child relationships applied to ABO blood groups. *Society for the Study of Social Biology*, 24:225–233, 1977.
- C Y Valenzuela. Human sociogenetics. *Biological research*, 44(4):393–404, 2011.
- J P Valenzuela, C Bellei, y D De los Ríos. Segregación Escolar en Chile. En Sergio Martinic y Gregory Elacqua, eds., *¿Fin de Ciclo? Cambios en la gobernanza del sistema educativo*, cap. 8, págs. 209–229. Pontificia Universidad Católica de Chile, Santiago, primera edición., 2010.
- J Vanegas, M Villalón, y C Valenzuela. Consideraciones acerca del uso de la variable etnia/raza en investigación epidemiológica para la Salud Pública: A propósito de investigaciones en inequidades. *Revista médica de Chile*, (136):637–644, 2008.

- M Vargas. Causes of Residential Segregation The Case of Santiago, Chile. págs. 1–23, 2006.
- G Vieyra, M Moraga, H Henríquez, F Aboitiz, y F Rothhammer. Distribución de alelos de los genes DRD4 y DAT1 del sistema dopaminérgico en la población mixta de Santiago de Chile. *Revista médica de Chile*, 131(2), 2003.
- G Waldman. Chile: indígenas y mestizos negados. *Política y Cultura*, (21):97–110, 2004.
- J Wang. Maximum-likelihood estimation of admixture proportions from genetic data. *Genetics*, 164(2):747–65, 2003.
- R S Waples y O Gaggiotti. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular ecology*, 15(6):1419–39, 2006.
- B. S. Weir y C Cockerham. Estimating F-statistics for the Analysis of Population Structure. *Evolution*, 38(6):1358–1370, 1984.
- E O Wright. *Classes*. Verso, primera ed^{ón}., 1985.