

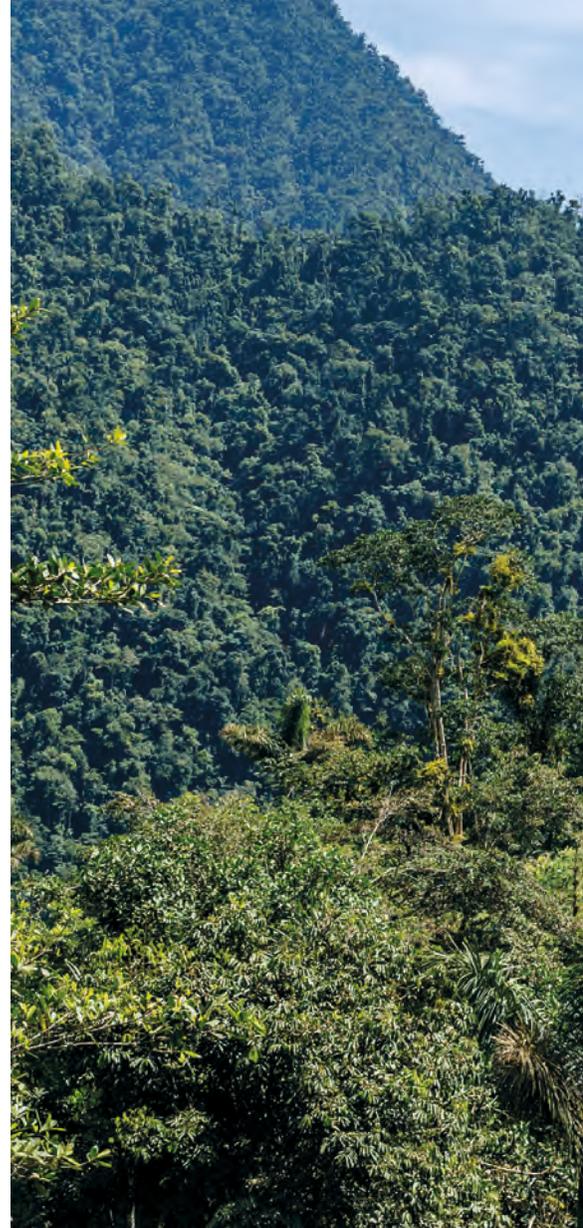
# CÓDIGO-Colombia: diversidad genómica, ancestría y salud en Colombia

**Augusto E. Valderrama-Aguirre**

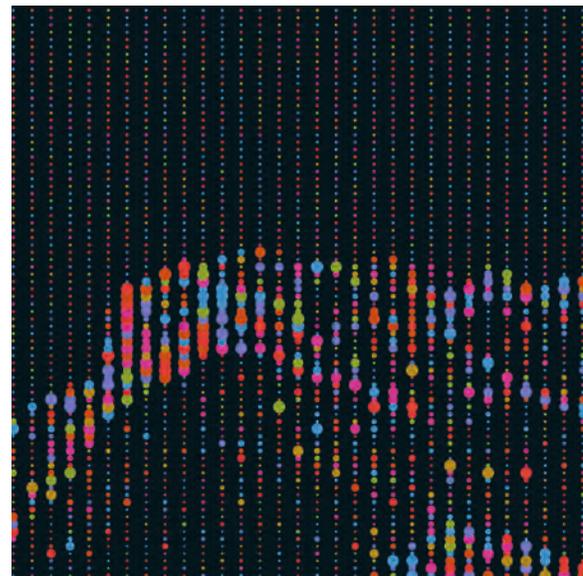
*Profesor asistente, Departamento de Ciencias Biológicas (DCB)*

Las metodologías para la captura de información genómica han permitido recoger de manera eficiente grandes cantidades de información, a la vez que permiten una disminución en los costos de recolección. Lo anterior ha permitido que en la actualidad existan decenas de iniciativas alrededor del mundo dedicadas a almacenar datos masivos de múltiples especies y de diversas características, incluyendo entre ellas genómica, transcriptómica, proteómica. En el caso de los seres humanos, el uso de esta información tiene al menos dos perspectivas principales, almacenar datos de individuos afectados por enfermedades o de la población general, sin que eso signifique que son mutuamente excluyentes.

El análisis de los datos genómicos almacenados permite la realización de varios ejercicios aplicables a la salud humana. Convencionalmente, en un ejercicio llamado diagnóstico, ha sido posible detectar la presencia de polimorfismos que afecten el estado actual de la salud de un individuo o un subgrupo poblacional.

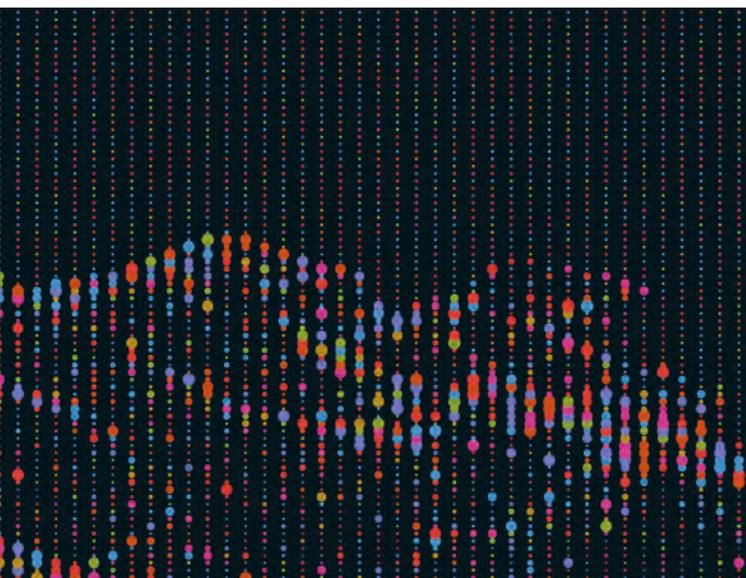


*Ciudad Perdida, Santa Marta, Magdalena, Colombia - <https://codig>*





[s://codigo.biosci.gatech.edu/](https://codigo.biosci.gatech.edu/)



Sin embargo, durante las últimas décadas han ganado fuerza dos ejercicios adicionales que se pueden realizar con esta información. En primera instancia, la información genómica permite analizar los riesgos que existen para desarrollar enfermedades, sobre todo las complejas no transmisibles, que puedan afectar a un individuo o a un subgrupo poblacional en el futuro; este ejercicio se denomina pronóstico.

*Visualización de datos genómicos*

*Fuente: Freepik*

En segunda instancia, en un ejercicio llamado predicción, es posible analizar la posible respuesta a un medicamento al que individuos o comunidades sean expuestos.

Como se infiere de lo anterior, estos ejercicios pueden realizarse de forma individual, lo que se ha denominado medicina personalizada, o de forma comunitaria, lo que se ha denominado, salud pública de precisión.

La puesta en marcha de las bases de datos genómicas ha permitido comprender que la mayor parte de la información disponible procede de individuos localizados en países desarrollados, cuya ancestría es predominantemente caucásica y de países cuya estructura poblacional es de un carácter homogéneo [1,2].

Esto implica un reto para generar conocimiento genómico a partir de poblaciones con ancestrías diferentes y cuyas estructuras poblacionales son más heterogéneas [3].

La población colombiana está compuesta de al menos tres tipos de ancestría, derivadas del mestizaje generado durante el encuentro colombino: europea, africana y nativa americana [4]. Estas ancestrías no se distribuyen de manera homogénea sobre el territorio colombiano, de tal manera que tenemos poblaciones con alto nivel de mestizaje y otras que aún conservan en una muy alta proporción sus ancestrías originales.

En el año 2014 se creó el proyecto ChocoGen [4,5], con objeto de estudiar los patrones de diversidad genómica y ancestría existentes en una muestra de individuos del departamento del Chocó y conocer su efecto sobre diversos fenotipos, patológicos o fisiológicamente normales.

Este proyecto generó un impacto alto, no solo por su novedad y producción intelectual, sino por su difusión por medios masivos y su red de colaboradores que poco a poco se fue ampliando [6].

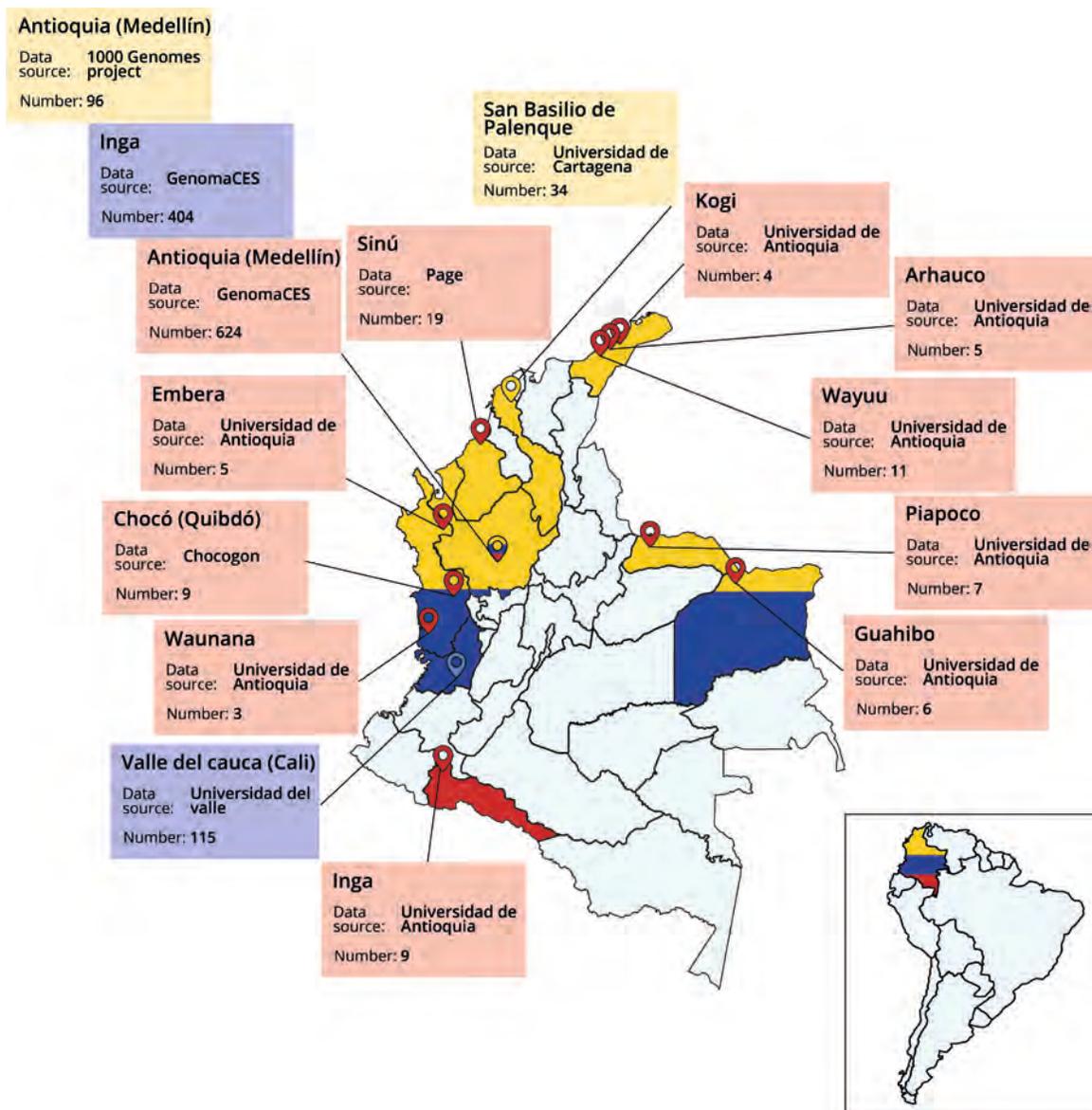
ChocoGen permitió la creación de tuberías de análisis bioinformático adaptadas a la población colombiana, conocer las huellas genómicas del mestizaje en Colombia y explorar las posibles implicaciones en los perfiles de salud-enfermedad de diversas subestructuras poblacionales del país [7-14].



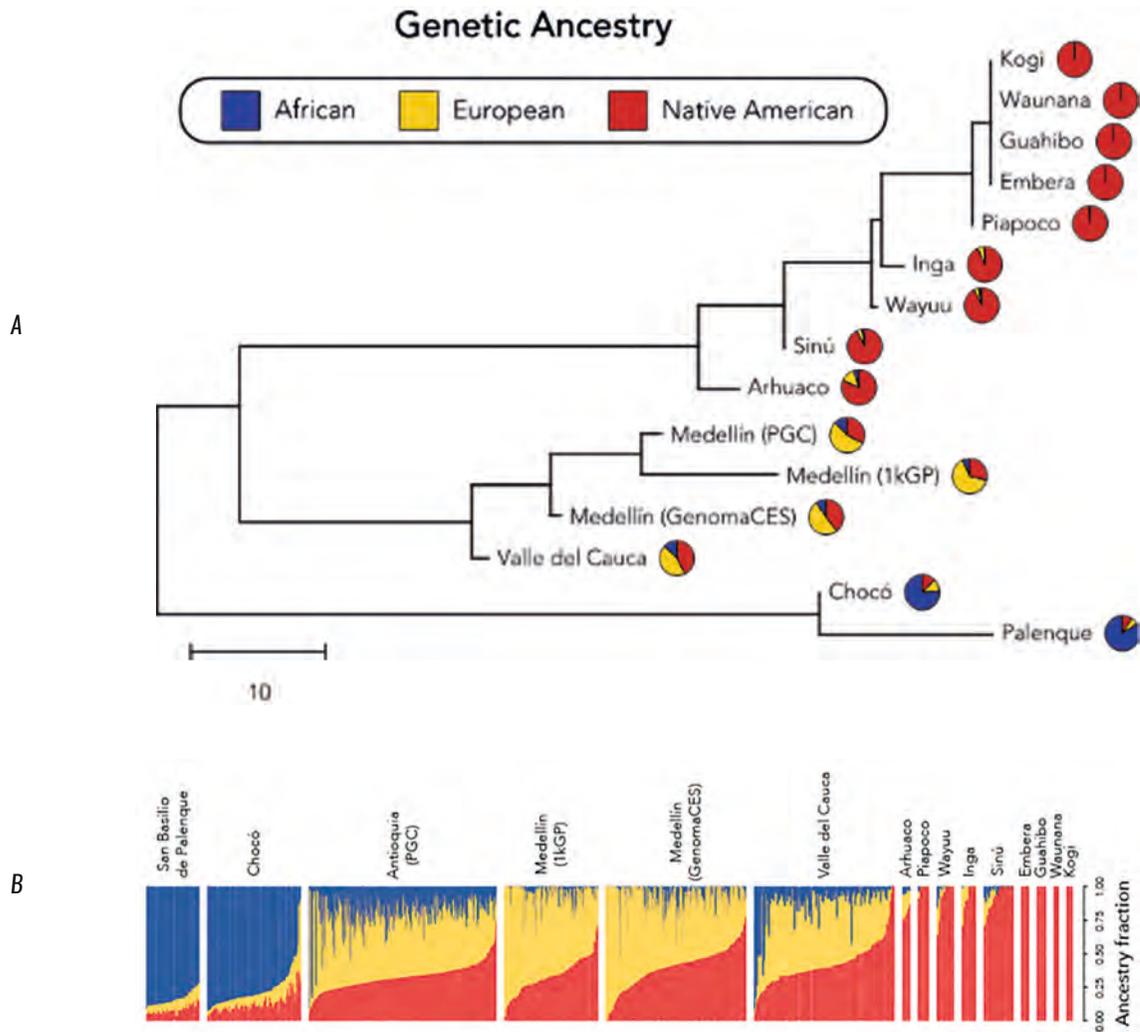
*Logo CÓDIGO-Colombia, representa la ganancia de conocimiento genómico a lo largo y ancho del país.*

Con la llegada del Profesor Augusto Valderrama al Departamento de Ciencias Biológicas (DCB) se dio fin al proyecto ChocoGen y se puso en marcha el consorcio CÓDIGO-Colombia. CÓDIGO-Colombia nace como una consecuencia de ChocoGen y reúne las contribuciones de varios investigadores colombianos.

Durante su desarrollo, ChocoGen fue ganando adeptos y, en consecuencia, se fue generando comunidad alrededor de la genómica humana poblacional en el país, de tal forma que la primera versión de la base de datos cuenta con una colección de cerca de dos mil colombianos y desde ya se vislumbra una colección de al menos tres mil para la segunda versión.



*CÓDIGO-Colombia, muestra el origen geográfico de las muestras incluidas en la primera versión de la plataforma.*



**Caracterización de ancestría en los muestreos poblacionales incluidos en CÓDIGO-Colombia. El panel A es un dendrograma que representa la separación de las poblaciones incluidas en CÓDIGO-Colombia, según su ancestría global. En el panel B se muestra la ancestría de cada individuo incluido en cada muestreo poblacional. Cada columna representa un individuo.**

CÓDIGO-Colombia representa una semilla para lo que podría ser la primera base de datos de genómica poblacional del país. Esta es una iniciativa importante en el sentido de que en Colombia existe una gran atomización de los esfuerzos en genómica poblacional. Con CÓDIGO-Colombia lo que se persigue entonces es generar comunidad y que, con unas normas de interacción claras, los diversos actores de la genómica del país logren involucrarse y aportar al proyecto.

A la fecha contamos con al menos ocho miembros contribuyentes. Cada uno de los miembros contribuyentes representa un profesor investigador, vinculado a instituciones de educación superior o centros de investigación, que tras realizar proyectos de genómica humana en poblaciones colombianas, decide aportar sus datos a CÓDIGO-Colombia.



# CÓDIGO

Consortium for Genomic Diversity, Ancestry, and Health in Colombia

Search for Genetic Variants with RSID, Position, GeneName, or Drug

GRCh38

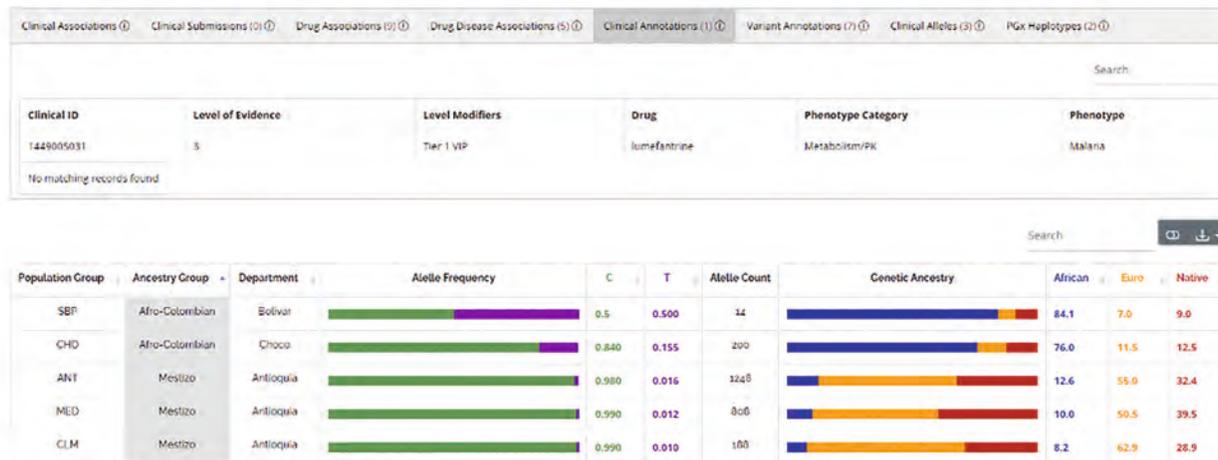
Search by RSID, Position, GeneName, or Drug

Search

## Homepage del sitio web CÓDIGO-Colombia (<https://codigo.biosci.gatech.edu/>)

Los miembros contribuyentes reciben el crédito público en la página web de CÓDIGO, mantienen el control sobre sus datos y participan en la coautoría de los manuscritos en los que se utilicen sus datos. Adicionalmente, la comunidad de miembros de CÓDIGO-Colombia tiene como opción desarrollar sus proyectos de investigación con el apoyo del consorcio. Una tercera forma de interactuar es que otros investigadores del país soliciten acceso a datos específicos y este se les otorgue; sin embargo, esta última aún se encuentra en evaluación con objeto de diseñar un protocolo de acceso que regule con detalle el uso de los recursos genómicos de los miembros contribuyentes. Todas estas interacciones están precedidas de la firma de acuerdos de transferencia de datos y de memorandos de entendimiento; así como la obtención de avales a fin de cumplir a cabalidad con la normatividad y los aspectos bioéticos a los que haya lugar.

CÓDIGO-Colombia es una plataforma de libre acceso que presenta estadísticos de resumen de diversos muestreos poblacionales en Colombia. La plataforma está particularmente enfocada en SNP (en inglés, *Single Nucleotide Polymorphisms*) y muestra datos genómicos como localización cromosómica, alelo de referencia, alelo alterado, gen y enlaces hacia sus descripciones en bases de datos globales como PharmGKB y ClinVar. Al ingresar a la plataforma, el usuario encuentra, al inicio, una barra de búsqueda que le permite explorar la base de datos según el código RSID del SNP, su posición en el genoma, el nombre del gen o el nombre genérico de algún medicamento. Una vez se introduce un criterio de búsqueda, por ejemplo, un RSID para un SNP, la plataforma arroja datos acerca de la frecuencia del SNP en los diversos muestreos poblacionales y las fracciones de ancestría en cada muestreo.



**Ejemplo de resultado en CÓDIGO-Colombia. El criterio de búsqueda fue RS10264272. Este pantallazo está sobre las anotaciones clínicas**

El resultado muestra varias propiedades adicionales de la variante que dan cuenta de su importancia clínica o farmacogenómica. Existen varias funcionalidades adicionales, pero en aras del espacio, se invita al lector que las explore en <https://codigo.biosci.gatech.edu/>.

CÓDIGO-Colombia cuenta con varios elementos diferenciadores, entre ellos se destaca, primero, que tiene un objetivo altruista, es decir, que este proyecto busca generar comunidad alrededor de la genómica humana en Colombia. Segundo, que no opera bajo costos; ni los miembros, ni los usuarios pagan por los beneficios obtenidos. Tercero, que provee acceso libre a estadísticos de resumen de múltiples muestreos poblacionales.

Cuarto, que provee datos poblacionales de variantes de interés en la salud pública de precisión que pueden finalmente traducirse a la medicina personalizada.

El consorcio cuenta con un grupo de fundadores que han coordinado las actividades iniciales de CÓDIGO-Colombia. Entre estos se encuentra el profesor King Jordan, PhD del Georgia Institute of Technology en Atlanta, Estados Unidos, Juan Esteban Gallo y Leonardo Mariño del National Institute on Minority Health and Health Disparities (NIH) en Bethesda, Maryland, Estados Unidos, y el profesor Augusto Valderrama del DCB de la Universidad de Los Andes en Bogotá.

Como se mencionó, el objetivo es consolidar una comunidad científica, académica y clínica entre los que se dedican a la genómica humana en Colombia.

Esperamos que el futuro cercano, CÓDIGO-Colombia gane adeptos y se logre aportar de una manera más consistente al conocimiento de la diversidad genómica y la ancestría de las diferentes subestructuras poblacionales del país, así como la aplicación de estos conceptos en la salud pública de precisión y la medicina personalizada en Colombia.

### Comparte este artículo



### Referencias

- [1] Popejoy AB, Fullerton SM. Genomics is Failing on Diversity. *Nature*. 2016; 538:161-164 .
- [2] Hindorff LA, Bonham VL, Brody LC, Ginoza MEC, Hutter CM, Manolio TA, *et al.* Prioritizing Diversity in Human Genomics Research. *Nat. Rev. Genet.* 2018; 19: 175-185.
- [3] McGuire AL, Gabriel S, Tishkoff SA, Wonkam A, Chakravarti A, Furlong EEM, *et al.* The Road Ahead in Genetics and Genomics. *Nat. Rev. Genet.* 2020; 21: 581-596.
- [4] Rishishwar L, Conley AB, Wigington CH, Wang L, Valderrama-Aguirre A, Jordan IK. Ancestry, Admixture and Fitness in Colombian Genomes. *Scientific Reports*. 2015; 5: 12376.
- [5] Medina-Rivas, MA, Norris ET, Rishishwar L, Conley AB, Medrano-Trochez C, Valderrama-Aguirre A, *et al.* Chocó, Colombia: A Hotspot of Human Biodiversity. *Revista Biodiversidad Neotropical*. 2016; 6(1): 45-54.
- [6] ChocoGen [Homepage, Internet]. 2019 [citado el 30 de junio del 2023]. Disponible en: <https://www.chocogen.com/>.
- [7] Norris ET, Wang L, Conley AB, Rishishwar L, Mariño-Ramírez L, Valderrama-Aguirre A, *et al.* Genetic Ancestry, Admixture and Health Determinants in Latin America. *BMC Genomics*. 2018; 19(8): 861.
- [8] Norris, ET, Rishishwar L, Wang L, Conley AB, Chande AT, Dabrowski AM, *et al.* Assortative Mating on Ancestry-Variant Traits in Admixed Latin American Populations. *Front. Genet.* 2019; 10: 359 .
- [9] Norris ET, Rishishwar L, Chande AT, Conley AB, Ye K, Valderrama-Aguirre A, *et al.* Admixture-enabled Selection for Rapid Adaptive Evolution in the Americas. *Genome Biol.* 2020; 21(1): 29.
- [10] Chande AT, Rishishwar L, Conley AB, Valderrama-Aguirre A, Medina-Rivas MA, Jordan IK. Ancestry Effects on Type 2 Diabetes Genetic Risk Inference in Hispanic/Latino Populations. *BMC Medical Genetics*. 2020; 21(2): 132. .
- [11] Chande AT, Nagar SD, Rishishwar L, Mariño-Ramírez L, Medina-Rivas MA, Valderrama-Aguirre AE, *et al.* The Impact of Ethnicity and Genetic Ancestry on Disease Prevalence and Risk in Colombia. *Front. Genet.* 2021;12: 690366.
- [12] Chande AT, Rishishwar L, Ban D, Nagar SD, Conley AB, Rowell J, *et al.* The Phenotypic Consequences of Genetic Divergence between Admixed Latin American Populations: Antioquia and Chocó, Colombia. *Genome Biol. Evol.* 2020; 12(9):1516–1527.
- [13] Nagar SD, Moreno AM, Norris ET, Rishishwar L, Conley AB., O'Neal KL. Population Pharmacogenomics for Precision Public Health in Colombia. *Front. Genet.* 2019; 10: 241.
- [14] Chande AT, Rowell J, Rishishwar L, Conley AB, Norris ET, Valderrama-Aguirre A, *et al.* Influence of Genetic Ancestry and Socioeconomic Status on Type 2 Diabetes in the Diverse Colombian Populations of Chocó and Antioquia. *Sci. Rep.* 2017; 7: 17127.