



# Instituto Politécnico Nacional Escuela Superior de Cómputo



## Bioinformatics

### Lab Session 9. SNPs

González Bocio Erik Alexander 2020630163

Jorge Luis Rosas Trigueros

Fecha de realización: 2 de mayo de 2022

Fecha de entrega: 16 de mayo de 2022

## Marco Teórico:

Los humanos tenemos diferencias y similitudes, sin embargo, a nivel genético los seres humanos nos parecemos a todos los seres vivos: todos utilizamos el ADN como material hereditario. Si comparáramos el ADN de un mono con el de un humano, encontraríamos que son iguales en más de un 96%, y si comparamos el genoma de una persona con el de otra, será igual en más del 99.99%. Aun así, no somos idénticos. Los SNPs (snips) son polimorfismos de nucleótidos individuales, es decir, son variaciones en el ADN que nos diferencian unos de otros a nivel genético. El orden o secuencia de nucleótidos dentro del ADN determina la estructura de los genes, y grandes secuencias de estos determinan la estructura de un genoma. Si comparamos los genomas de muchas personas, a lo largo de ellos iríamos encontrando piezas distintas en algunos. Estas variaciones en una sola pieza de un genoma son a lo que se le llama SNP. Los SNPs se presentan en más del 1% de la población, son la fuente de variación más común entre humanos y en general en miembros de la misma especie. Los SNPs se presentan en una de cada 300 pares de bases en nuestro ADN, lo que nos da un valor aproximado de 10 millones de SNPs para un genoma humano. Muchos de ellos determinan nuestras características físicas, así como la predisposición a ciertas enfermedades, mientras que algunos no cumplen ninguna función. Es posible descubrir SNPs al comparar los genomas de una población. Todos estos análisis se llevan a cabo por computadora y buscan las diferencias entre el mayor número de personas posible. Mediante estos estudios se puede saber si un SNP está asociado a un fenotipo. Ejemplo de ello son los SNPs que sólo están presentes en personas que padecen una enfermedad, por lo que su presencia puede indicar una susceptibilidad. Los SNPs que sólo se encuentran en personas con reacciones adversas a algún medicamento también pueden ser señal de que se debe utilizar un fármaco diferente. Hay SNPs que se encuentran con mayor frecuencia en algunas poblaciones y son heredados de generación en generación; por esto los SNPs también nos pueden decir mucho sobre nuestros orígenes y antepasados. En general los SNPs son muy útiles e informativos. Su estudio permite a los científicos conocer más sobre las enfermedades genéticas y sobre la estructura de las poblaciones tanto humanas como de otros seres vivos. La combinación de nuestros SNPs es como una huella digital genética; conocerla nos ofrece información útil tanto para cuidar nuestra salud (prevenir y tratar enfermedades) como para saber qué nos hace genéticamente únicos.

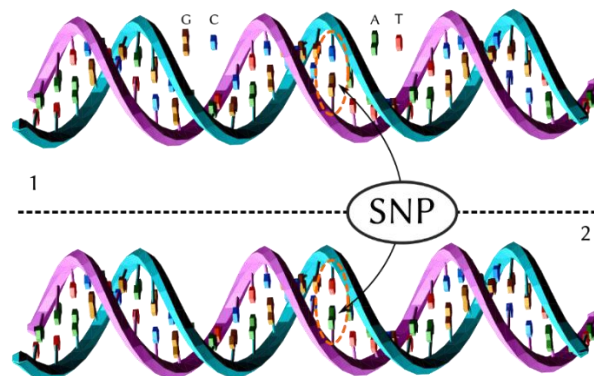


Fig. 1. Imagen sobre SNP

Los polimorfismos de un solo nucleótido (SNP) representan a las variantes genéticas más comúnmente encontradas en el genoma humano. Debido a su amplia distribución, estos polimorfismos se localizan en cualquier parte de la estructura de los genes y el genoma. Los SNP que tienen implicaciones funcionales sobre los niveles de expresión génica se denominan SNP reguladores (rSNP), mientras que los que alteran la traducción de los ARN mensajeros (ARNm), el corte y empalme, la eficiencia para potenciar o inhibir el corte y empalme, la estabilidad de los ARNm y la función de las proteínas (sin alterar su estructura) se denominan SNP ARN estructurales (srSNP). Diversos estudios han documentado la importancia funcional de los rSNP y srSNP en el desarrollo de enfermedades comunes como hipertensión arterial (HTA), obesidad, artritis reumatoide, enfermedad arterial coronaria, entre otras. El objetivo de este artículo es hacer una revisión bibliográfica actualizada de los SNP funcionales (rSNP y srSNP) que tiene efecto en la función del gen, ARNm, proteínas y que se asocian con diversas enfermedades complejas.

Los polimorfismos de un solo nucleótido, frecuentemente llamados SNP (pronunciados “snips”), son el tipo más común de variación genética entre las personas. Cada SNP representa una diferencia en un solo bloque de construcción de ADN, llamado nucleótido. Por ejemplo, un SNP puede reemplazar el nucleótido citosina (C) con el nucleótido timina (T) en un cierto tramo de ADN.

Los SNP se producen normalmente en todo el ADN de una persona. Ocurren casi una vez cada 1000 nucleótidos en promedio, lo que significa que hay aproximadamente de 4 a 5 millones de SNP en el genoma de una persona. Estas variaciones ocurren en muchos individuos; para ser clasificado como un SNP, se encuentra una variante en al menos el 1 por ciento de la población. Los científicos han encontrado más de 600 millones de SNP en poblaciones de todo el mundo.

Los SNP difieren de las variantes de sustitución, que reemplazan un bloque de construcción de ADN (nucleótido) con otro. Las variantes de sustitución generalmente causan enfermedades y generalmente no se encuentran en el 1 por ciento de ninguna población. Además, los SNP se diferencian de las variantes del número de copias (CNV), que se producen cuando se duplica o elimina un gen completo (u otra gran parte del ADN). Más comúnmente, los SNP se encuentran en el ADN entre los genes. Pueden actuar como marcadores biológicos, ayudando a los científicos a localizar genes asociados con enfermedades. Cuando los SNP ocurren dentro de un gen o en una región reguladora cerca de un gen, pueden desempeñar un papel más directo en la enfermedad al afectar la función del gen.

La mayoría de los SNP no tienen ningún efecto sobre la salud o el desarrollo. Sin embargo, algunas de estas diferencias genéticas han demostrado ser muy importantes en el estudio de la salud humana. Los SNP ayudan a predecir la respuesta de un individuo a ciertos medicamentos, la susceptibilidad a factores ambientales como las toxinas y el riesgo de desarrollar enfermedades. Los SNP también se pueden usar para rastrear la herencia de variantes genéticas asociadas a enfermedades dentro de las familias. Se están realizando investigaciones para identificar SNP asociados con enfermedades complejas, como enfermedades cardíacas, diabetes y cáncer.

## Material y Equipo:

- Equipo de Computo
- Red de Internet

## Desarrollo de la Práctica:

Para llevar a cabo esta práctica trabajamos con 4 diferentes SNPs, los cuales elegimos de la página SNPedia y posteriormente analizamos en la página del National Library of Medicine

Primero comenzamos con la rs17822931 que determina el cerumen (cera en los oídos), sudoración y olor corporal

### rs17822931

**rs17822931**, also known as c.538G>A or G180R, is a SNP in the ATP-binding cassette, sub-family C (CFTR/MRP), member 11 [ABCC11](#) gene. The ABCC11 protein helps transport small molecules across apical membranes such as those in apocrine secretory cells.

[PMID 16444273] This SNP determines wet vs dry *earwax* as well as sweat production, and it is also associated with lipid secretion. It is commonly (T;T) for East Asians and (C;C) for Europeans and Africans.

[PMID 23325016] A study of deodorant usage in a large (17,000+) population showed that **rs17822931**(T;T) individuals were at least 5-fold less likely to use deodorant, consistent with them being "genotypically nonodorous". Although not statistically significant, there was some evidence for a heterozygote effect, implying that individuals with a (C;C) genotype for this SNP were likely to use deodorant at a higher frequency than (C;T) genotypes.

gnxp indicates this SNP influences risk of breast cancer in Japanese women. This does not hold true in Caucasian women, though.[PMID 21655989]

[reddit discussion](#)

[PMID 23970085] Strong association of the SNP rs17822931 with wet earwax and bromhidrosis in a Chinese family

Orientation	plus		
Stabilized	plus		
Geno	♦	Mag	♦ Summary
(C;C)	2	Wet earwax. Normal body odour. Normal colostrum.	
(C;T)	2	Wet earwax. Slightly better body odour.	
(T;T)	2.5	Dry earwax. No body odour. Likely Asian ancestry. Reduced colostrum.	
Reference	GRCh38 38.1/141		
Chromosome	16		
Position	48224287		
Gene	ABCC11		
is a	snp		
is	mentioned by		
dbSNP	rs17822931		
dbSNP (classic)	rs17822931		
ClinGen	rs17822931		

Fig. 2. SNP rs17822931 en la SNPedia donde podemos ver un pequeño resumen

Vemos cual cromosoma es donde esta, y los buscamos en el otro servidor, luego de eso vemos un resumen de este

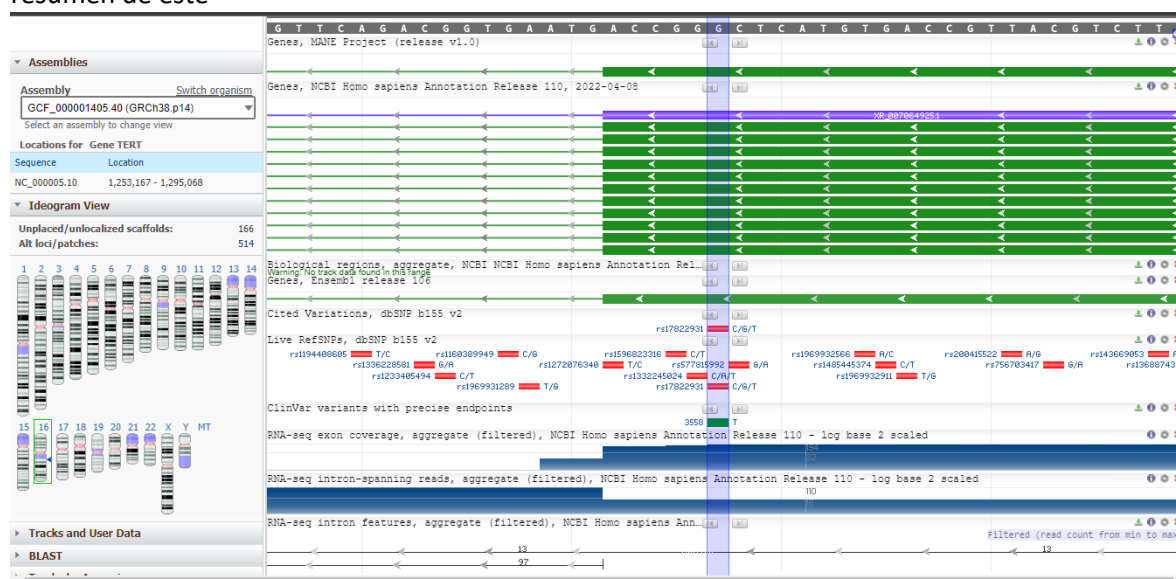


Fig. 3. Búsqueda en el National Library of Medicine, aquí lo buscamos por el cromosoma que le corresponde (16) y nos muestra esta pantalla

Finalmente nos da las enfermedades que conlleva tal SNP

Frequency	Variant Details	Clinical Significance	HGVS	Submissions	History	Publications	Flanks
Allele: T (allele ID: 18597)							
ClinVar Accession	Disease Names		Clinical Significance				
RCV000003737.3	Apocrine gland secretion, variation in		Benign				
RCV000003738.2	Axillary odor		Benign				
RCV000003739.2	Colostrum secretion		Benign				

Fig. 4. Clinical Significance, podemos observar que nos habla sobre la cera de los oídos, el olor de la axila y los calostros

En este caso como recordamos que nos hablaba sobre secreción de cera de los oídos y lo podemos ver en la significación clínica donde habla sobre la secreción de glándulas apocrinas, también de la variación en olor axilar y la secreción de calostro, estos 3 son cosas normales en todos nosotros, por lo que se tienen en una clasificación benigna.

El segundo SNP elegido es el rs10757274 el cual puede duplicar el riesgo de enfermedad cardíaca. Se cree que aproximadamente uno de cada cuatro caucásicos porta las variantes genéticas.

## rs10757274

**rs10757274** is a SNP located in chromosomal region 9p21; to date (2009), it is the strongest genetic predictor of early myocardial infarction (heart attack) discovered so far. SNPs in this region are also associated with increased risk of stroke, [abdominal aortic aneurysm](#) (AAA), and intracranial aneurysm.

Based on a large (20,000+) study primarily of Caucasian females, it appears that on its own, this single SNP does not improve on the discrimination or classification of predicted [heart disease](#) risk achieved by traditional risk factors, high-sensitivity C-reactive protein, and a family history of premature myocardial infarction. [PMID 19153409]

**rs10757274** and **rs2383206** can double the risk of [heart disease](#)[1]. About one in every four Caucasians are thought to carry the gene variants. **rs10757278** in the same region has been linked to [diabetes](#) [2][3]

a [blog post](#) about investigating **rs10757274** and **rs2383206**

part 2 in the series mentions **rs10811661**

part 3 in the series mentions **rs1333040**, **rs2383207** and **rs10116277** **rs10811661**

[PMID 18048766] This SNP was also associated with increased risk for coronary artery disease in a Korean population.

[PMID 18066490] Also found to be significant in a study of 416 Italian myocardial infarction patients

Orientation	plus		
Stabilized	plus		
Geno	Mag	Summary	+
(A:A)		normal	
(A:G)		~1.2x increased risk for heart disease	
(G:G)		~1.3x increased risk for heart disease	
Reference	GRCh38 38.1/141		
Chromosome	9		
Position	22096056		
Gene	CDKN2B-AS1		
is a	snp		
is	mentioned by		
dbSNP	rs10757274		
dbSNP (classic)	rs10757274		
ClinGen	rs10757274		
ebi	rs10757274		
HLI	rs10757274		

Fig. 5. SNP rs10757274 en la SNPedia donde podemos ver un pequeño resumen

Entonces lo buscamos en el segundo servidor y vemos el resumen de este SNP

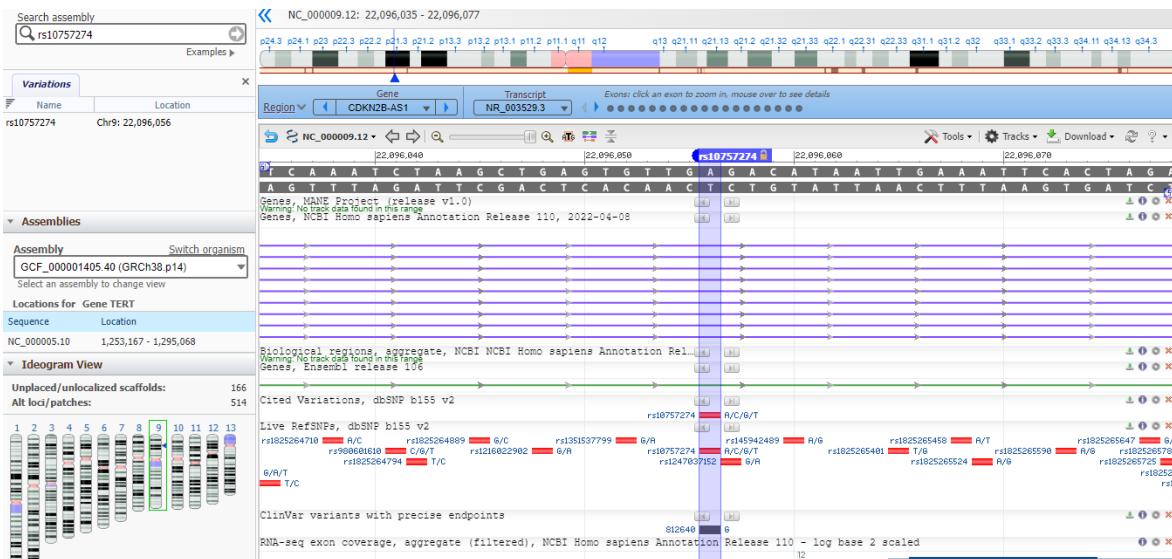


Fig. 6. Búsqueda en el National Library of Medicine, aquí lo buscamos por el cromosoma que le corresponde (9) y nos muestra esta pantalla

Podemos observar que nos habla de una enfermedad coronaria de tres vasos que corresponde a las enfermedades cardiacas de las que se nos hablaba en el resumen de la SNPedia y es un factor de riesgo.

Frequency

Variant Details

Clinical Significance

HGVS

Submissions

History

Publications

Flanks

Allele: G (allele ID: 800885)

ClinVar Accession	Disease Names	Clinical Significance
RCV001003458.1	Three Vessel Coronary Disease	Risk-Factor

Genomic regions, transcripts, and products

Choose placement GRCh38.p13 ( NC\_000009.12 )

See rs10757274 in Variation Viewer

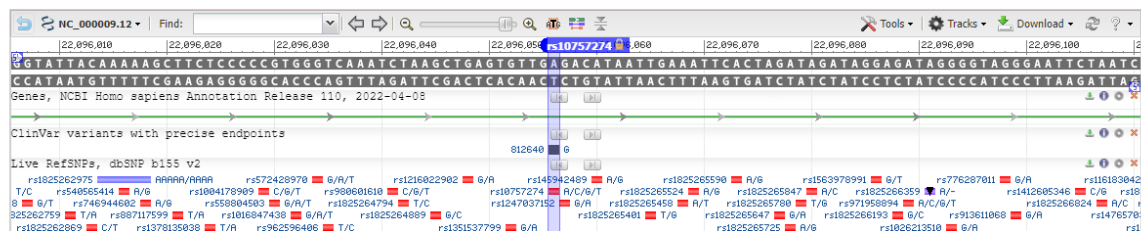


Fig 7. Clinical Significance, podemos observar que nos habla sobre una enfermedad coronaria de tres vasos

El tercer SNP elegido es el rs2273535 el cual desempeña un papel en varios tipos de cáncer

## rs2273535

SNP **rs2273535**, also known as F31I or Phe31Ile, has been associated with increased risk for several cancers, in most cases when individuals are homozygous for the risk allele, **rs2273535(T)**, as oriented to the dbSNP entry.

A meta-analysis of almost 10,000 cases of breast, colon, ovarian, prostate, lung, esophageal and non-melanoma skin cancer, compared to an equal number of Caucasian controls, determined the following risks (i.e., odds ratios, OR) [PMID 15802297]:

- For **colorectal cancer**: OR for homozygotes of 1.5 (CI: 1.14-1.99)
- For **breast cancer**: OR for homozygotes of 1.35 (CI: 1.12-1.64)
- For any of the cancer types studied: OR for heterozygotes of 1.10 (CI: 1.03-1.18), OR for homozygotes of 1.40 (CI: 1.22-1.59)

In a Chinese population, breast cancer risk for **rs2273535(T;T)** homozygotes compared to the other two genotypes led to an odds ratio of 1.66 (CI: 1.29-2.12), and appeared to be more pronounced for younger patients. [PMID 15271856]

However, for **lung cancer** (among Caucasians), **rs2273535(T;T)** homozygotes have been reported to be at lower risk; specifically, an odds ratio of 0.63 (CI: 0.41-0.96) has been reported. [PMID 16926177]

[PMID 18431743] showed **no association** with ovarian cancer risk 4,624 invasive epithelial ovarian cancer cases and 8,113 controls of white non-Hispanic origin **ovarian cancer**

Orientation	plus	
Stabilized	plus	
Geno	Mag	Summary
(A;A)	0	common
(A;T)	2	slightly higher (~1.5x or less) risk for certain cancers
(T;T)	2.1	slightly higher (~1.5x or less) risk for certain cancers

Reference	GRCh38 38.1/141
Chromosome	20
Position	56386485
Gene	AURKA
is a	snp
is	mentioned by
dbSNP	rs2273535
dbSNP (classic)	rs2273535
ClinGen	rs2273535
ebi	rs2273535
HLI	rs2273535

Fig. 8. SNP rs2273535 en la SNPedia donde podemos ver un pequeño resumen

Esta se encuentra en el cromosoma 20, así que lo buscamos en el segundo servidor, dando el siguiente resultado

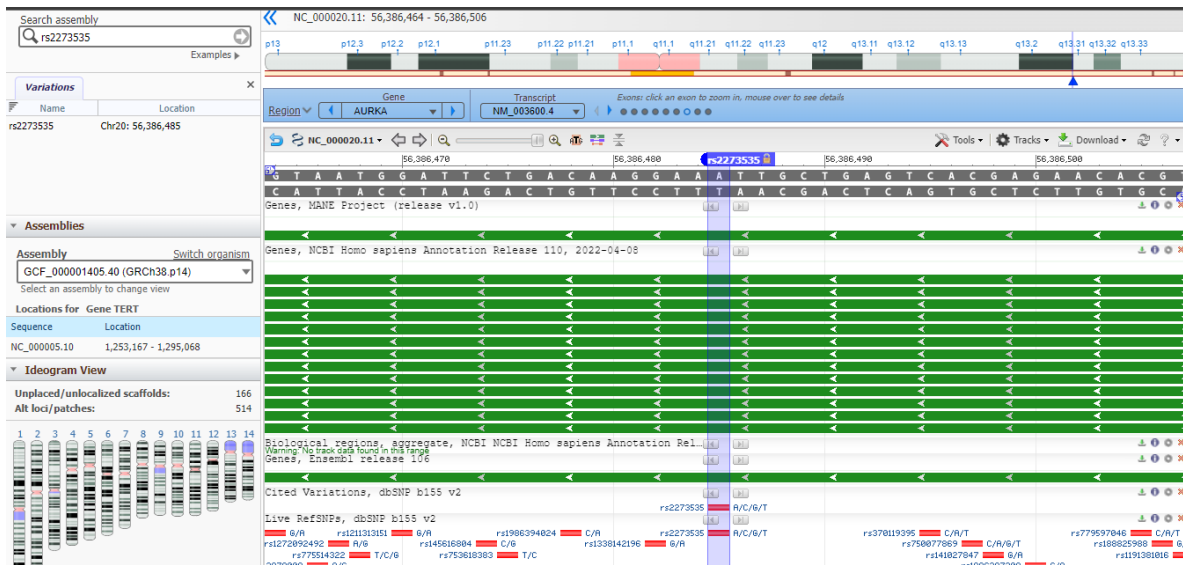


Fig. 9. Búsqueda en el National Library of Medicine, aquí lo buscamos por el cromosoma que le corresponde (20) y nos muestra esta pantalla

Y una vez viendo el resumen de este SNP y la significancia medica podemos ver que está relacionado con una susceptibilidad al cáncer de colon, el cual es un factor de riesgo.

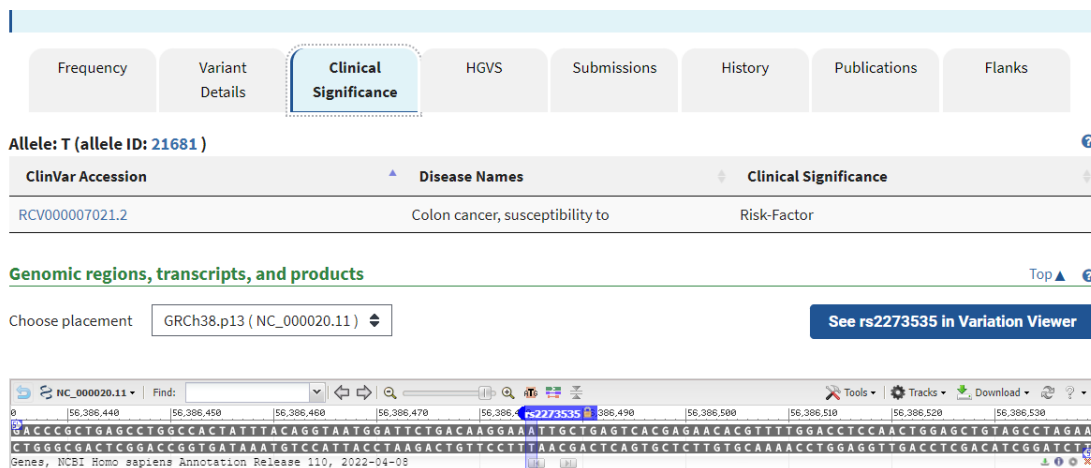


Fig. 10. Clinical Significance, podemos observar que nos habla sobre cáncer de colon

Por último, tenemos el rs7903146 con el cual en el 3% de la población aumenta considerablemente el riesgo de diabetes tipo 2



## rs7903146

This SNP in *TCF7L2* influences the risk of *Type-2 diabetes* (T2D). This SNP is also known as IVS3C>T.

[PMID 17671651] rs7903146(C;T) rs7903146(T;T) strongly predicted future type-2 diabetes. Considered in context with **rs7903146** rs12255372 rs10885406.

Note: this is one of two SNPs within the *TCF7L2* gene that have been reported to be associated with type-2 diabetes, the other being rs4506565. They have approximately equal power to estimate risk for type-2 diabetes, and the results from one correlate 92% of the time with the other. [PMID 17554300]

Full text of the paper is available from Plos Medicine.

Or from NCBI as [PMID 17020404].

[PMID 16855264] *TCF7L2* polymorphisms and progression to diabetes in the Diabetes Prevention Program.

[PMID 17668382] reconfirmed in a diverse population

[PMID 17971425] Associated in a study of ~1000 Hispanic-Americans with reduced insulin secretion as measured by acute insulin response and adjusted for the degree of insulin sensitivity ( $p = 0.032$ ).

[PMID 17671651] the CT/TT genotypes strongly predicted future type-2 diabetes. The risk T

Associated with T2D.

Orientation plus

Stabilized plus

Geno + Mag + Summary

(C;C)	2	Normal (lower) risk of Type 2 Diabetes and Gestational Diabetes.
(C;T)	2.1	1.4x increased risk for diabetes (and perhaps colon cancer).
(T;T)	3.5	2x increased risk for Type-2 diabetes

Reference GRCh38 38.1/141

Chromosome 10

Position 112998590

Gene *TCF7L2*

is a snp

is mentioned by

dbSNP rs7903146

Fig. 11. SNP rs7903146 en la SNPedia donde podemos ver un pequeño resumen

Lo buscamos en el segundo servidor con el cromosoma 10 y nos muestra el siguiente resultado

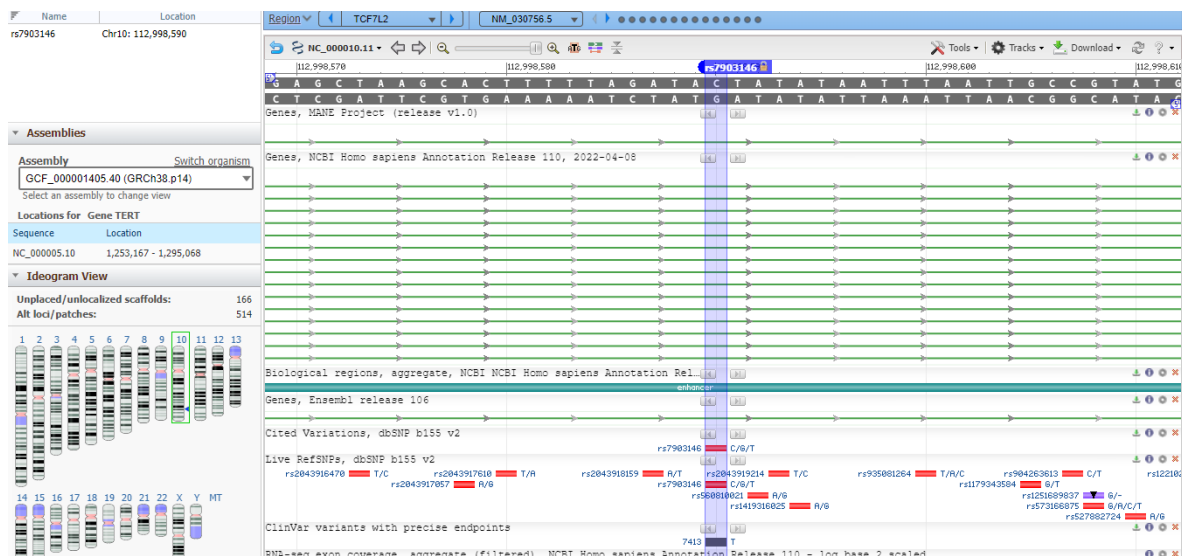


Fig. 12. Búsqueda en el National Library of Medicine, aquí lo buscamos por el cromosoma que le corresponde (10) y nos muestra esta pantalla

Finalmente podemos ver que efectivamente como se nos decía en la SNPedia tiene que ver con la susceptibilidad a diabetes tipo 2 y es un factor de riesgo, de igual manera tiene que ver con la respuesta de derivados de urea con sulfonamidas, la cual es de tipo respuesta a un fármaco

Allele: T (allele ID: 22452)		
ClinVar Accession	Disease Names	Clinical Significance
RCV000007838.5	Diabetes mellitus type 2, susceptibility to	Risk-Factor
RCV000211424.1	sulfonamides, urea derivatives response - Efficacy	Drug-Response

Genomic regions, transcripts, and products Top ▲ ?

Ilustración 13. Clinical Significance, podemos observar que nos habla sobre diabetes tipo 2 y una respuesta a fármacos relacionado con las sulfonamidas

## Conclusiones y Recomendaciones:

En esta práctica pudimos aprender más acerca de los SNPs y con ayuda de los servidores ver como es que afectan algunos, en algunos casos pues conllevan cosas normales, como lo fue con el primer SNP que elegí, el cual solo era sobre la cera y las axilas, mientras que en otros casos tienen que ver con respuestas a fármacos o lamentablemente con factores de riesgo como es el caso de los ejemplos de cáncer de colon y la diabetes tipo 2.

Me pareció una buena práctica ya que la SNPedia tiene una gran cantidad de SNPs y es muy interesante como estos tienen que ver con muchas cosas, además de que al ser libre la elección pues ayuda a investigar un poco más.

## Bibliografía:

- Alvarado, J. (2022). *SNP*. Conogasi. Retrieved 13 May 2022, from <https://conogasi.org/articulos/snp/>.
- Castro, D. (2022). *Curiosidades geneticas: ¿Qué son los SNPs? | ADNTRO*. ADNTRO. Retrieved 13 May 2022, from <https://adntro.com/es/blog/curiosidades-geneticas/que-son-los-snps/>.
- Genetics, H., & Research, G. (2022). *What are single nucleotide polymorphisms (SNPs)?*: MedlinePlus Genetics. Medlineplus.gov. Retrieved 13 May 2022, from <https://medlineplus.gov/genetics/understanding/genomicresearch/snp/>.
- Ramírez, J. (2022). *Polimorfismos de un solo nucleótido (SNP): implicaciones funcionales de los SNP reguladores (rSNP) y de los SNP-ARN estructurales (srSNP) en enfermedades complejas*. Anmm.org.mx. Retrieved 13 May 2022, from [https://www.anmm.org.mx/GMM/2013/n2/GMM\\_149\\_2013\\_2\\_220-228.pdf](https://www.anmm.org.mx/GMM/2013/n2/GMM_149_2013_2_220-228.pdf).