The Heritable Nature of rs2910164 and Behçet's Disease

Andrew Hill[1†], Kristen Steenbergen[1†], Rohan Shetty[1†], Theresa Boyer[1†*],

1 Advanced Academic Programs Division, Johns Hopkins University, Baltimore, Maryland, USA

†These authors contributed equally to this work.
¤Current Address: Johns Hopkins University, Zanvyl Krieger School of Arts & Sciences, 3400 N. Charles St, Baltimore, MD 21218, USA
* tboyer3@jhmi.edu

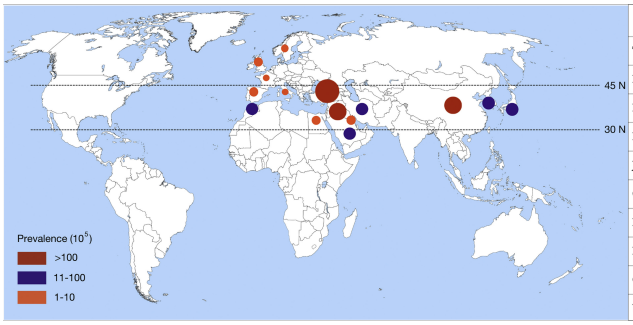
Abstract

Behçet's disease is an autoimmune disease that manifests as an inflammation of blood vessels leading to ulcers, bleeding, and multi-system failure. The single nucleotide polymorphism (SNP) rs2910164 occurring at position 60 of the miR146A gene is associated with an increase in expression leading to an increase in Bechet's disease incidence. This SNP is heritable and has been reported in different populations and at varying rates. We hypothesized that there are differential frequencies of rs2910164 alleles in distinct populations. Publicly available NGS sequence datasets were surveyed for individuals of varying ethnicities and were queried for the presence of the SNP. Our search landed upon a familial trio of Ashkenazi Jewish individuals (mother, father, and son) all of whom contained the SNP. These results confirm the expected genotype of the individuals and also reveal the hereditary nature of rs2910164. Further analysis of datasets may be done with these methods to help understand the prevalence of this SNP genotype and track its ancestry.

Introduction

Behçet's disease is a rare autoimmune condition characterized by inflammation of blood vessels in the body leading to ulcers in the mouth and/or genitals, inflammation in the eyes, joint pain, and later multi-system failure. Behçet's disease has shown differential incidence across various ethnic groups and geographic locations and to have at least some degree of inheritance (Fig 1). Several studies have shown a relatively high incidence of familial cases in Turkey and Japan [1], with the former being shown to be as high as 31% familial cases [2]. Behçet's disease has also shown a high prevalence in Israel and Germany, as well as in Jewish and Ashkenazi Jewish populations. While the precise model of inheritance remains unknown, there is evidence to suggest that it is autosomal dominant [3]. Furthermore, an association between the miR146A SNP rs2910164 and the increased risk for Behçet's disease has been suggested [4]. The inheritance pattern of Behçet's disease patients, proposes that rs2910164 possesses a confounding factor affecting its interaction with disease models which may be influenced by both inheritance and environment.

Due to the lack of data correlating the frequency of rs2910164 SNP with specific risk groups, we investigate whether the difference in the incidence of Behçet's disease in various groups arises due to the difference in frequency of specific alleles of rs2910164.



Country	Prevalence (1/100,000)
Turkey [8, 9]	20–602
Israel [14, 15]	50–185
Iran [12, 16]	16.7–100
Iraq [17]	17
Japan [9]	7–13.5
China [18]	2.62
Kuwait [12]	2.1
Italy [19, 20]	4.1–15.9
Germany [21]	0.9
USA [12]	0.33

A B

Fig 1. A.) Geographic distribution and prevalence of Behçet's disease [5] B.)Prevalence of Behçet's disease demonstrating the high regional prevalence in Turkey, Israel, Japan, and Germany [6]

The miR146A gene

miR146A gene is a mammalian, intronless intragenic miRNA that is located on the 2nd exon of its host lncRNA gene miR3142HG on the human chromosome 5 (Fig 2).

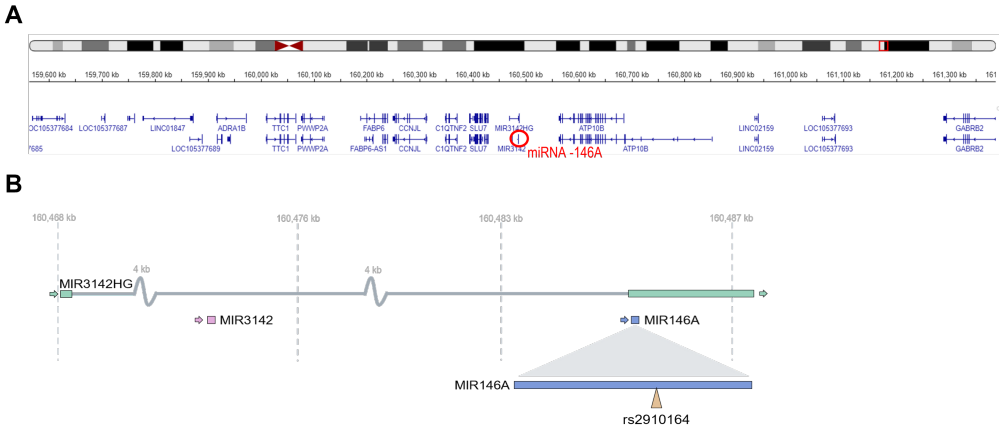


Fig 2. The miR146A locus. A.) hg19 Chromosome 5 map showing the location of miR146A highlighted in red at the right end of the map. B.) Map and structure of miR-3142HG and miR146A with exons shown in green. miR146A transcript shown in blue with approximate location of rs2910164 SNP.

miR146A serves as a negative feedback loop inhibitor of NF-κB, a nuclear transcription factor innately tied to the inflammatory response. NF-κB triggers the expression of miR146A which subsequently inhibits the activity of IRAK1 and TRAF6. IRAK1 and TRAF6 are key adaptors for NF-κB. Therefore, miR146A is critical in the modulation of inflammatory responses thus preventing hypersensitivity reactions and tissue damage caused by inflammation or the immune response. Due to its role in inflammatory regulation, miR146A has been correlated with the development of a variety of autoimmune conditions such as rheumatoid arthritis, Behçet's disease, and lupus, as well as in the development of certain cancers such as pancreatic cancer and leukemias [7–9]. Autoimmune conditions develop as a consequence of excessive and uncontrolled inflammatory response, whereas the development of cancer has been correlated with persistent inflammatory states further implicating miR146A as a key regulator of NF-κB.

Regulation of miR164A

Epigenetic modifications are known to affect the miR146A gene promoter, which was identified 16kb upstream of the gene [10]. CpG methylation and histone acetylation are involved in altering the miR146A expression [11](Fig 3).

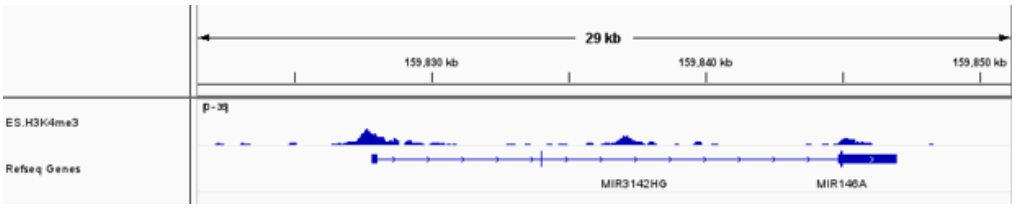


Fig 3. H3K4me3 methylation. H3K4me3 methylation in the MIR3142HG and miR146A regions showing the approximate location of the miR146A promoter within the miR3142HG intron (IGV, hg18 Chr5).

Furthermore, evidence has shown that intragenic miRNAs are co-regulated with their host gene promoters [12]. Therefore, the regulation of miR3142HG also influences miR146A expression. miR3142HG is a lncRNA gene, which is involved in the regulation and reduction of the inflammatory response, and miR146A can be co-expressed with its host. For example, Interleukin 1β (IL-1β) has been shown to induce the expression of miR3142HG which metabolizes to miR146A and promotes the release of IL-6, IL-8, and CCL2 [13–15]. Dysregulation leads to the pathogenesis of autoimmune diseases such as rheumatoid arthritis, and psoriasis. This is notable since the pathology of interest, Behçet's disease, is an autoimmune disease.

rs2910164 polymorphism and Bechet's disease

One specific SNP, rs2910164, has been associated with the development of a variety of illnesses. In the mature miRNA, the rs2910164 SNP occurs in the 3' antisense product and an alteration in the miRNA sequence can affect the hybridization to mRNA targets (Fig 4A and B). Within rs2910164, there are two common alleles - G and C (Fig 4C). The G allele has been shown to increase expression of miR146A thus increasing inhibition of NF-κB-associated pathways which is believed to increase the risk of developing cancer and autoimmune illnesses.

Since the prevalence of publicly available sequencing data with linked Behçet's disease phenotype is quite limited, we assessed the allele frequency from a population approach. We investigated whether the difference in the incidence of Behçet's disease in various groups arises due to the difference in the frequency of specific alleles of rs2910164. We hypothesize that there are differential frequencies of rs2910164 alleles in various populations. Given that Behçet's disease is most commonly seen in areas such as Turkey, Asia, Japan, and the Mediterranean ancestral groups from these regions were targeted (Fig 1A).

Methods

Three human genome files were selected from the NIST's Genome in a Bottle (GIAB) database [16]. The samples selected are a mother, father, and son 'trio' which were sequenced using the Illumina HiSeq platform. The raw data files were sourced from GIAB's GitHub repository of data indexes [17]. Raw FASTQ data had been pre-analyzed and annotated variant call files (.vcf) were available from the repository.

For this analysis, .vcf files were downloaded from GitHub for the Ashkenazim trio: son (HG002), father (HG003), and mother (HG004). All three of these genomes were

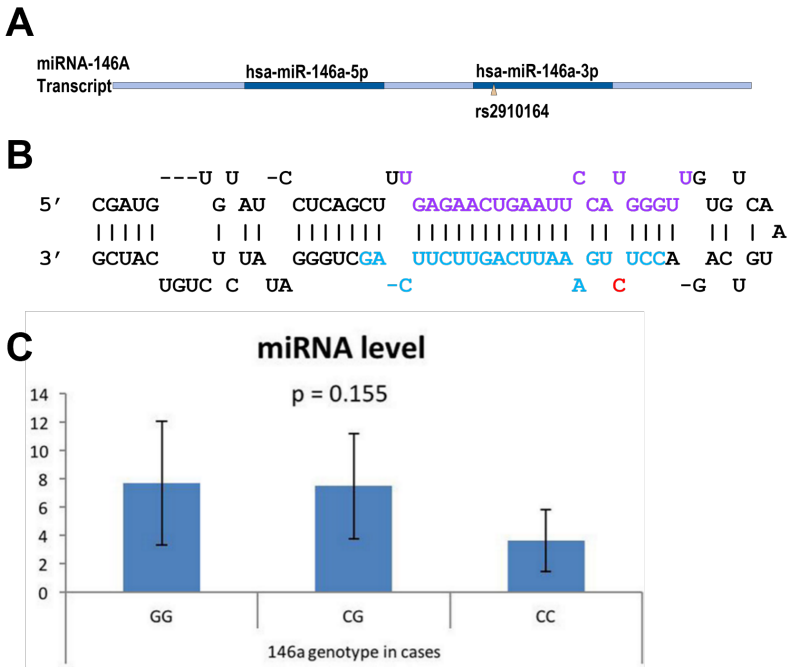


Fig 4. rs2910164 SNP and miR146A A) miR146A transcript and rs2910164 location. B) Secondary structure of miR146A with polymorphism site highlighted in red. C allele represented. Purple indicates 5' product, and blue indicates 3' antisense product. C) Prevalence of miR146A allele genotypes.

previously aligned to the hg19 assembly (GRCh37.p13) by GIAB and this assembly was used for establishing SNP coordinates. The VCFfilter tool was utilized in Galaxy to filter the whole-genome .vcf file down to the region chr5:159911585-159913000 to screen specifically for rs2910164. Variants covered to a depth greater than 10 reads were selected. The resulting filtered .vcf file displayed the SNP status of each individual for rs2910164 in miR146A.

Results

Our analysis revealed that all 3 individuals had the rs2910164 SNP. Furthermore, the reported allele frequency from the population of 1000 genomes project data is .62 (Fig 5. Given that the trio analyzed are of Ashkenazi Jew ancestry, gnomAD reports the allele frequency of the SNP to be .7763 [18].

Additionally, the lowest second most frequent allele (reported as the global minor allele frequency) for this location is cytosine, at a frequency of .3814. This is notable since cytosine is the nucleotide reported by the reference genome as seen in Fig 5.

Discussion

By examining the allele frequency and identifying the presence of the rs2910164 SNP in the Ashkenazi trio, it further confirms the hypothesis that this SNP may be a driver of inflammatory-mediated diseases within this ethnic group and geographic region.

The results suggest that those of Ashkenazi Jewish descent have an increased incidence of the G allele which has been correlated with the development of a variety of

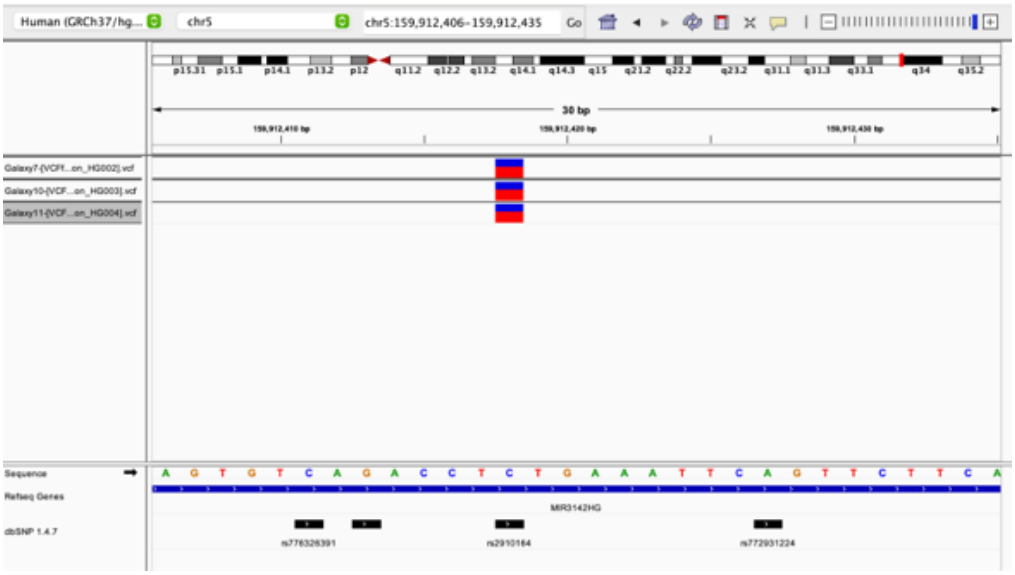


Fig 5. Allele Frequency Allele frequency of the son (HG002), father (HG003), and mother trio (HG004) at the chromosomal location of rs2910164 SNP within IGV. The dark blue bars indicate that the allele frequency is heterozygous at this location with a frequency of 0.62.

cancers and autoimmune conditions suggesting that this ethnic group may present an increased risk for illnesses associated with this SNP. Jewish populations have been noted to have increased incidence of Behçet's disease lending this theory credence.

Furthermore, there appears to be a strong heritability of the SNP given that the mother, father, and son all have the SNP. While it has been shown that a large proportion of Behçet's disease cases are sporadic in nature, a notable portion of cases have been noted to be familiar as well. As noted earlier, familial Behçet's disease has shown a higher incidence in specific geographic areas and ethnicities. Data collected has shown the high prevalence in Turkey and among Jewish and Ashkenazi Jewish populations as discussed above(Fig 1B). Additionally, it has been suggested that Behçet's disease follows an inheritance pattern of autosomal dominance [3]. This suggests that rs2910164 may possess a confounding factor influencing its effect on the development of Behçet's and possibly other conditions, including both genetic and geographic drivers of the disease. The transcriptome is largely unexplored and it is possible that other unidentified ncRNA's play an impact on the regulation of immune pathways.

The limitations of our study include the small size sample of the available dataset as well as the ability to only investigate fixed inheritance patterns. A lack of additional information about disease phenotype and exposures to other factors that may lead to Behçet's Disease, including lifestyle and environmental exposures, limits the interpretation of our findings.

Conclusion

Our findings demonstrate the heritable nature of rs2910164 and its implication in the pathogenesis of Behçet's Disease. Further research would be required to identify the specific molecular mechanisms by which this SNP affects the etiology and incidence of the disease within geographically and genetically high-risk groups. The rs2910164 SNP

has the potential to serve as a useful biomarker for identifying groups at risk for the development of Behçet's and related illnesses.

References

1. Srivastava N, Chand S, Bansal M, Srivastava K, Singh S, et al. Familial Behcet's disease. Indian journal of dermatology venereology and leprology. 2007;73(4):260.
2. Ozyurt K, Colgecen E, Baykan H. Does Familarity Or Family History Of Recurrent Oral Ulcers Effect Clinical Characteristics of Behçet's Disease? Acta Dermatovenerologica Croatica. 2013;21(3):168–168.
3. Shigemura T, Kaneko N, Kobayashi N, Kobayashi K, Takeuchi Y, Nakano N, et al. Novel heterozygous C243Y A20/TNFAIP3 gene mutation is responsible for chronic inflammation in autosomal-dominant Behçet's disease. RMD Open. 2016;2(1):e000223.
4. Ibrahim W, Sakr BR, Obaya E, Ghonem H. MicroRNA-146a expression and microRNA-146a rs2910164 polymorphism in Behcet's disease patients. Clinical rheumatology. 2019;38(2):397–402.
5. Greco A, De Virgilio A, Ralli M, Ciofalo A, Mancini P, Attanasio G, et al. Behçet's disease: new insights into pathophysiology, clinical features and treatment options. Autoimmunity reviews. 2018;17(6):567–575.
6. Oguz I, Hizli P, Gonul M. The epidemiology of Behçet's disease. Behcet's disease: InTech. 2017; p. 15–26.
7. Abdelaleem O, Fouad N, Shaker O, Hussein H, Ahmed F, Ali D, et al. Association of miR-146a rs57095329 with Behçet's disease and its complications. British Journal of Biomedical Science. 2021;78(2):63–66.
8. Pei JS, Chang WS, Hsu PC, Chen CC, Chin YT, Huang TL, et al. Significant Association Between the MiR146a Genotypes and Susceptibility to Childhood Acute Lymphoblastic Leukemia in Taiwan. Cancer Genomics Proteomics. 2020;17(2):175–180.
9. Tavasolian F, Hosseini AZ, Soudi S, Naderi M. miRNA-146a improves immunomodulatory effects of MSC-derived exosomes in rheumatoid arthritis. Current Gene Therapy. 2020;20(4):297–312.
10. Taganov KD, Boldin MP, Chang KJ, Baltimore D. NF-κB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. Proceedings of the National Academy of Sciences. 2006;103(33):12481–12486.
11. Iacona JR, Lutz CS. miR-146a-5p: Expression, regulation, and functions in cancer. Wiley Interdisciplinary Reviews: RNA. 2019;10(4):e1533.
12. Liu B, Shyr Y, Cai J, Liu Q. Interplay between miRNAs and host genes and their role in cancer. Briefings in functional genomics. 2019;18(4):255–266.
13. Xia M, Sui Z. Recent developments in CCR2 antagonists. Expert opinion on therapeutic patents. 2009;19(3):295–303.
14. Nishimoto N. Interleukin-6 in rheumatoid arthritis. Current Opinion in Rheumatology. 2006;18(3):277–281.

15. Hadjicharalambous MR, Roux BT, Feghali-Bostwick CA, Murray LA, Clarke DL, Lindsay MA. Long Non-coding RNAs Are Central Regulators of the IL-1β-Induced Inflammatory Response in Normal and Idiopathic Pulmonary Lung Fibroblasts. Frontiers in Immunology. 2018;9:2906.

16. Zook JM, Catoe D, McDaniel J, Vang L, Spies N, Sidow A, et al. Extensive sequencing of seven human genomes to characterize benchmark reference materials. Scientific data. 2016;3(1):1–26.

17. Genome-In-A-Bottle. Genome-in-A-bottle/giab_data_indexes: GitHub repository containing data indexes from NIST's genome in a bottle project.; 2021. Available from: `https://github.com/genome-in-a-bottle/giab_data_indexes`.

18. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581(7809):434–443.