تحديد التباين الوراثي لمتلازمة هارادا في المرضى السعوديين باستخدام تسلسل الاكسوم

علياء محمود علي البلوي

اشراف

د. نجلاءعلي محمد البرعي

المستخلص

متلازمة هارادا هو مرض مناعي ذاتي للخلايا المنتجة للصبغة. يعتبر المرض نادرًا جدًا في السكان ويؤثر بشكل أساسي على الأشخاص الذين لديهم بشرة غامقة او سمراء. متلازمة هارادا هو مرض متعدد العوامل والسبب الدقيق لمرض متلازمة هارادا غير معروف. جميع الدراسات السابقة استخدمت دراسة الارتباط (ارتباط الجينات ذات الصلة / حالة الأليل المرشح بالإضافة إلى الارتباط على نطاق الجينوم) لفك دور العناصر الجينية في مسببات متلازمة هارادا. ومع في هذه الدراسة، تم استخدام نهج تسلسل عميق .ذلك، فإن علم الوراثة الكامن وراء متلازمة هارادا لا يزال غير واضح) من الجينوم البشري في مرضى متلازمة WESلتسلسل جميع جينات التشفير بالكامل (نهج تسلسل إكسوم بالكامل - نهج هارادا. تم جمع ١٧ عينة من مرضى متلازمة هارادا من منطقة المدينة المنورة بالمملكة العربية السعودية. تم فحص باستخدام مجموعة DNAهؤلاء المرضى سريريًا من قبل استشاري طب العيون وتم التشخيصهم سريريا. تم استخلاص باستخدام WESلـ DNA . من أصل ١٧ مريضًا، تم إخضاع ٥ عينات من DNAمتاحة تجاريًا وتم قياس تركيز Illumina متبوعة بتوليد الكتلة وتسلسل النهايات المتقطعة على جهاز whole exome مجموعة أدوات إعداد مكتبة / GRCh38. تم إنشاء قراءات عالية الجودة والتي تم محاذاتها ومقارنتها بالجينوم البشري المرجعي (NextSeq500 BWA (Burrows-Wheeler Alignment). تم إجراء المقارنات وتحديد المتغيرات باستخدام برامج hg38 (مجموعة أدوات تحليل الجينوم). تم الحصول على قائمة تضم حوالي ٩٠,٠٠٠ متغير في المتوسط عبر كل GATKو عينة. تم وضع علامة على هذه المتغيرات باستخدام أدوات قياسية متبوعة بالترشيح باستخدام قواعد بيانات التكرار في السكان. فقط المتغيرات النادرة ستحتفظ بتكرار السكان أقل من ٠,٠٠٥ (<٠,٠٠٥) والمتغيرات غير المترادفة مثل تغيير إطار القراءة، خطأ القراءة، ومتغيرات موقع لصق اثناء عملية النسخ والترجمة. أسفر هذا التحليل عن ١٠٠ متغير تقريبًا في المتوسط لكل عينة.الكشف عن المتغيرات المشتركة في جميع العينات حيث وجد ٥٥ متغيرًا مشتركا في ٢٢ جينًا. تم (جامعة كاليفورنيا، UCSCتحديد أولويات هذه الجينات بناءً على تعبيرها وبياناتها الوظيفية المتاحة في متصفح جينوم سانتا كروز) وتم اختيار فقط تلك الجينات ذات التعبير المعروف في الجلد والعين والدماغ والأذن. كما تم النظر في الجينات غير المميزة. هذا أدى إلى تحديد ١٥ متغيرًا في ٦ جينات مرشحة ذات صلة بالمرض. الجينات الستة المرشحة و *FAM86B2* و *ZNF717* و *PRIM2* و *ANKRD20A1* و*FLJ22184*لـ متلازمة هارادا المحددة في هذه الدراسة هي طويل غير مشفر ، و RNA يشفر *FLJ22184*، والطفرات الوراثية قد تكون سببا في التغير الجيني لها. *RFPL4A* وعوامل النسخ في متلازمة هارادا بالفعل من خلال دراسات lncRNAs يشفر عامل النسخ. تم وصف دور *ZNF717* بترميز وحدة فرعية أساسية وتم ربط المتغيرات في هذا الجين باضطراب التصبغ. *PRIM2*مختلفة. يقوم هي جينات غير مميزة وهناك حاجة إلى دراسات وظيفية لتحديد دورها *RFPL4A* و *FAM86B2* و *ANKRD20A1* الدقيق في صحة الإنسان والمرض.

Identifying Genetic Variation Underlying Vogt-Koyanagi-Harada (VKH) in Saudi Patients Using Exome Sequencing

By
Alia Mahmoud Ali Albalawi

Supervised By

Dr. Najla Ali Alburae

ABSTRACT

Vogt-Koyanagi-Harada (VKH) is an autoimmune disease of the pigment-producing cells. The disease is considered rare in the population and predominantly affects people with darker skin. VKH is a multifactorial disease, and the precise cause of the VKH disease is unknown. All previous studies had used an association study design to decipher the role of genetic elements in the etiology of VKH. However, the genetics underlying VKH is still unclear. In this study, an in-depth sequencing approach was used to completely sequence all coding genes of the human genome in VKH patients. For this, 17 VKH patients were collected from the Madinah region of Saudi Arabia. The consultant ophthalmologist clinically examined the patients. Genomic DNA was extracted using the commercially available kit, and the DNA was quantified using a spectrophotometer. Out of 17 patients, 5 DNA samples were subjected to WES using the SureSelect^{XT} library preparation kit followed by cluster generation and paired-end sequencing on an Illumina NextSeq500 instrument. A high-quality reads were generated which were aligned to the human reference genome assembly (GRCh38/hg38). Alignment and variant calling were performed using BWA (Burrows-Wheeler Alignment) and GATK (genome analysis toolkit) softwares. A list of approximately 90,000 variants on average was obtained across each sample. These variants were annotated using standard tools followed by filtration using population frequency databases. Only rare variants with a population frequency of less than 0.005 (<0.005) and non-synonymous variants such as nonsense, frameshift, missense, and splice site variants were retained. This analysis yielded approximately 124 variants on an average per sample. Detecting shared variants across all samples yielded 55 variants in 22 genes. These genes were prioritized further based on their expression, and functional data available in the UCSC (University of California, Santa Cruz) genome browser, and only those genes with known expression in skin, eye, brain, and ear were selected. Also, uncharacterized genes were considered. This led to the identification of 15 variants in 6

candidate genes. The six VKH candidate genes identified in this study are *FLJ22184, ANKRD20A1*, *PRIM2*, *ZNF717, FAM86B2,* and *RFPL4A*. The variants in these genes include insertions, deletion, nonsense, and missense variants. *FLJ22184* encodes a long noncoding RNA, and *ZNF717* encodes a transcription factor. Different studies have already described the role of lncRNAs and transcription factors in VKH. *PRIM2* encodes a primase subunit and variants in this gene have been associated with a depigmentation disorder. *ANKRD20A1, FAM86B2,* and *RFPL4A* are uncharacterized genes, and functional studies are needed to establish their precise role in human health and disease.