(096L)

Validation of Multigene Targeted Sequencing Approach for the Identification of Genetic Risk Factors for Increased Thrombotic Events

Authors: Alaa Hassan Al-Amoudi[1,2], Adel Abuzenadah[1,2], Ashraf Dallol[1,2]

Center of Excellence in Genomic Medicine Research [1], and Faculty of Applied Medical Sciences [2], King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia

Background

- Thrombophilia is a common chronic disease leading to increase the tendency to clot and increase the risk of thrombosis (1).
- Venous Thromboembolism (VTE) is associated with acquired and/or inherited factors and include deep vein thrombosis (DVT) and pulmonary embolism (PE) (2).
- DVT occurs in two-third of the patients with VTE while PE occurs in about one-third of the patients with VTE (3).
- The inherited thrombophilia is a result of DNA mutation in genes responsible for the production of blood clotting-proteins in which a genetic mutation affects the function or amount of the protein (1).
- While inherited thrombophilia can be caused by a number of mutations, the most common ones are factor V leiden (FVL) and prothrombin (Factor II) that are mutations responsible for the gain of the function (1, 4).
- There are many other factors that can cause thrombophilia such high level of other procoagulants such as factors VII, VIII, and IX (4, 5).
- This study aimed to investigate the feasibility of multigene targeted sequencing approach for the identification of genetic risk factors for increased thrombotic events.

Objectives

- To establish the technical requirements necessary to perform multigene testing for thrombosis.
- To determine the genetic factor other than FVL and FII underline increase the risk for thrombotic events.

Methods

A custom Ampliseq panel targeting 23 genes in the coagulation pathway was designed and manufactured. A total of 14 DNA samples derived from peripheral blood of thrombophilia patients and one control were used to test the panel's coverage and uniformity. The sample were acquired from the Center of Excellence in Genomic Research (CEGMR) for suspected patients with thrombophilia and were tested negative for FVL and prothrombin. The alteration in the 23-thrombophilia gene that are known to harbour variants associated with thrombotic disorders were examined via next generation sequencing using Ion PGM Machine. We follow the strategies that will allow the selection of rare variants of high quality and more likely to be germline.

SYMBOL	NAME	SYMBOL	NAME
KLKB1	kallikrein B1	FGG	fibrinogen gamma chain
KNG1	kininogen 1	F13A1	coagulation factor XIII A chain
F11	coagulation factor XI	F13B	coagulation factor XIII B chain
F9	coagulation factor IX	THBD	thrombomodulin
F8	coagulation factor VIII	PROC	protein C, inactivator of coagulation factors Va and VIIIa
F3	coagulation factor III, tissue factor	PROS1	protein S (alpha)
F7	coagulation factor VII	PLAT	plasminogen activator, tissue type
F10	coagulation factor X	PLG	plasminogen
F5	coagulation factor V	SERPIND1	serpin family D member 1
F2	coagulation factor II, thrombin	CPB2	carboxypeptidase B2
FGA	fibrinogen alpha chain	SERPINC1	serpin family C member 1
FGB	fibrinogen beta chain		

Table 1: Gene List.

Filter Strategy 1:
P Value:
Less than or equal 0.05

Filter Strategy 2:
Frequency:
Greater than 30 %

Filter Strategy 3:
MAF:
Less than or equal 0.05

Figure 1: Filtration Strategies.

Results

- Up to 43 variants was discovered after a stringent filtration strategy. The custom sequencing panel produced libraries with >94% of the target sequenced and with mean depth of >1000, with >94 of uniformity. The 15 samples were negative for both FVL and FII and diagnosed as patients with thrombotic events except the control case.
- Most of the variants were SNVs. Fifteen patients were found on NGS to have different variants. Eight cases contain more than 2 variants. The most common variant rs6023 that appear in three cases include case 2,3, and 6. More than four variants appear in cases, 2,7,8, and case 14, while four variants appear in case 9. In these cases, eight variants were found in F5 gene, five variants found in F13B gene, four variants found in KLKB1 gene, two variants found in F13a1 gene, two variants found in PLG gene, two variants found in F10 gene, one variant found in PROS1 gene, and one variant found in F2 gene. There are five variants in F5 gene similar between case 7 and 4 include rs6027, rs2227243, rs9332609, rs6018, and rs6024, and to do association a large population need to be studied.

Sample	NO. of filtered Variants	% On Target	Mean Depth	Uniformity
1* BB-892-934	5	96.53	1141	96.60
2* BB-914-329	8	97.36	5547	96.57
3* BB-923-849	6	96.46	8513	96.85
4* BB-924-305	5	97.07	9639	96.88
5* BL-91-17	7	97.02	7371	95.19
6* BL-563-18	7	96.44	3015	96.53
7* BL-73-15	13	96.81	1325	96.18
8* BL-106-16	9	96.98	1256	94.84
9* BL-107-15	7	94.82	3284	95.24
10* BL-108-15	4	96.22	1143	96.45
11* BB-872-450	6	97.58	1849	96.31
12* BB-912-628	5	96.28	3203	95.65
13* BB-923-260	5	96.95	9876	95.87
14* BB-924-272	10	96.73	1952	94.62
15* PBG-003-19 (control)	5	96.93	6979	95.63

Table 2: Mean depth and Uniformity.

Gene	dbsnp	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12	Case 13	Case 14	Case 15
FGA	rs1442242 38	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
F9	rs6049	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
F5	rs6023	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0
KLKB1	rs7555298 8	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
KLKB1	rs1483996 02	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
KLKB1	rs1489743 69	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
F10	rs4128660 8	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
SERPINC1	rs1456178 46	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0
PLG	rs4252174	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
KLKB1	rs1482772 06	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
KNG1	rs1864828 77	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
PI4KA|SERPIND1	rs5902	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0
F9	rs3469877 9	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
PROS1	rs7384607 0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
KNG1	rs1441236 48	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
PLG	rs1165737 85	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
F5	rs6027	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
F5	rs2227243	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
F5	rs9332609	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
F5	rs6018	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
F5	rs6024	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
F13B	rs6000	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
F13B	rs5999	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
F13B	rs5994	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
F13B	rs1754959 6	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
F13B	rs2015855 19	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
F5	rs6034	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
F5	rs6011	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
KLKB1	rs4253331	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
F13A1	rs3024481	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
F13A1	rs3024480	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
PLG	rs1835346	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0
PROS1	rs1457115 36	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
PLG	rs4252128	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
F2	rs4562263 5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
F10	rs3211798	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
PLG	rs4252060	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
KNG1	rs7643893 8	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
F10	rs6175326 6	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
KLKB1	rs4253377	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
CPB2-AS1|CPB2	rs1145627 02	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
F2	rs3136436	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
F2	rs5898	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
0	None	1	Hetero	2	Homo											

Table 3: Filtered Variants.

Discussion

- From the many factors that cause thrombophilia, FVL and FII are the most common hereditary factors that cause it.
- FVL was one of the most reason can cause thrombophilia like Bengt, et al study in 2018 that he found 220 patients from 4890 Swedish patients had VTE during follow-up (113 DVT, 78 PE, 29 both) and the incidence of VTE was significantly higher in the patients with homozygous and heterozygous FVL (6).
- Also another study done by Gonzalez et al in 2016 found that the FII is significantly associated with CVT when they compared with the healthy controls (7).
- Many genome-wide studies, and more recently also whole exome sequencing studies have been conducted to identify genetic risk factors for VTE. However, approximately all the variants that have been found to be associated with the risk for thrombophilia are located within known susceptibility genes (8).
- This project demonstrates the utility of using NGS sequencing panels to diagnose and understanding of the thrombotic events and the molecular bases of the genes and centered around identifying the other genetic factor that cause thrombophilia other than these two factors including 23 genes of thrombophilia that play a role in the coagulation pathway where FVL and FII are negative which will help us to do association between the different variants that can cause thrombophilia.
- When we filter the known and most common variants like FVL and FII we could not find a pathogenic disease-causing variant in the coagulation pathway and nothing really stand out.
- By using NGS 23 gene we found other variants that could be associated with thrombosis and we identified variants of uncertain significance, Benign, and Likely benign.

How to improve ?

- More sample size and controls are needed to be studied to do association since small sample size is one of the limitations that we cannot get more result to correlate.
- Using whole exome or whole genome technique in the future will help analysis and diagnosis of the thrombophilia.

Conclusions

- While the present study comprised of small sample size, the findings indicate that more than one variant appear in each patient sample suggesting this accumulating variants in the coagulation pathway is significant for thrombophilia disease.
- The custom panel produced good results in the term of coverage, uniformity, and mapped read and therefore it could be utilized for analyzing a larger cohort of patients with thrombotic events.
- In the future it is probably better to do NGS sequencing using exome or whole genome sequencing in the coagulation pathway genes since nothing appear in the 23 coagulation gene we used except the common variant that could be associated with the disease but confirm the association statistically significant or not by using a large set of samples and controls.

References