

Migração do fluxo de trabalho de sequenciamento previamente validado de 4.700 genes associados a fenótipos clínicos para a plataforma MGI DNBSEQ-G400: dados NGS de alta qualidade e concordância com os resultados esperados

Nara Diniz Soares Pessoa¹, Gustavo Barcelos Barra¹, Anderson Coqueiro dos Santos¹, Pedro Góes Mesquita¹, Ilária Cristina Sgardioli¹, Amanda Minafra Reis Lamas¹, Rosenelle de Oliveira Araújo Benício¹, Lorena de Melo Gama¹, Letícia Rodrigues Brito¹, Gabriela Togawa Meguerditchian¹, Alessandra de Freitas Andrade¹, Graciela Ribeiro Martins¹, Rafael Henriques Jácomo¹, Lidia Freire Abdalla Nery¹

¹Sabin Diagnóstico e Saúde, Brasília, Distrito Federal



INTRODUÇÃO

Os avanços no sequenciamento de nova geração (NGS) aprimoraram significativamente o diagnóstico clínico, permitindo uma análise abrangente de genes associados a diversas doenças hereditárias.

OBJETIVO

Este estudo teve como objetivo transferir o fluxo de trabalho de sequenciamento previamente validado no NextSeq-500 (Illumina, Inc.) para a plataforma DNBSEQ-G400 (MGI Tech Co.), avaliando o desempenho desta em termos de rendimento de dados, qualidade e precisão na identificação de variantes e sua concordância com variantes detectadas no NextSeq-500.

METODOLOGIA

A biblioteca de NGS com 12 pacientes (DNA extraído da saliva) foi preparada utilizando o kit xGEN DNA Library Prep e enriquecida com o *xGen Inherited Disease Panel V1* (~4.700 genes com associação gene-doença) e o kit *Hybridization Capture* (Integrated DNA Technologies, Inc.), sequenciada nas plataformas NextSeq-500 (150 ciclos *paired-end*) e DNBSEQ-G400 (300 ciclos *paired-end*). Para o DNBSEQ-G400, as bibliotecas foram convertidas de DNA linear para *nanoballs* com o kit reagente DNBSEQ™ OneStep DNB Make V4.0. O processamento de dados foi realizado utilizando o Dragen v.4.2.4 na plataforma EMEDGENE (Illumina, Inc.). Foram calculadas as principais métricas de desempenho do DNBSEQ-G400 (%>Q30, contagem total de *reads*, % de *reads* mapeadas, profundidade média de cobertura, % de bases com cobertura >20X e duplicatas). Além disso, foi comparado o número de SNVs e INDELs e calculado *F-measure* entre as duas plataformas de sequenciamento.

RESULTADOS

DNBSEQ-G400 gerou uma mediana de 54.143.871 *reads* por amostra (*range*: 31.912.278–72.109.870). Obteve-se uma mediana de 93,06% das bases >Q30 (*range*: 93,03%–93,22%) e 99,1% das *reads* foram mapeadas para o genoma humano de referência hg38 (*range*: 98,02%–99,32%). A mediana de cobertura média foi de 397X (*range*: 230X–537X), a mediana de 99,7% de cobertura >20X (*range*: 99,60%–99,80%) e mediana de 13,52% de taxa de duplicação (*range*: 12,37%–15,09%). Foi identificada uma mediana de 7.709 SNVs (*range*: 7.437–8.028) e 344 INDELs (*range*: 329–367) por amostra. A *F-measure* foi de 99% para SNVs (*range*: 98,5%–99,5%) e 82,3% para INDELs (*range*: 79,3%–84,6%).

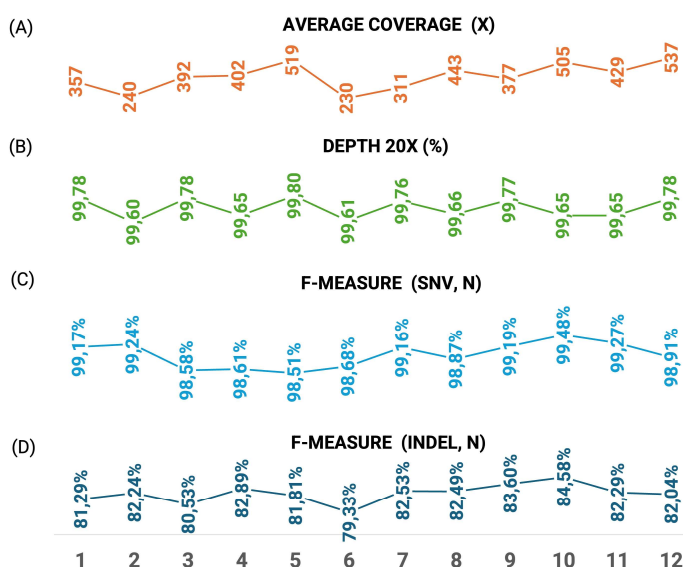


Figura 1. Desempenho de sequenciamento em 12 amostras de saliva na plataforma MGI DNBSEQ-G400. Os gráficos mostram valores específicos para cada métrica: (A) Cobertura média do alvo (x) para o painel de 4.700 genes; (B) Porcentagem de bases com cobertura ≥ 20x; (C) *F-measure* da detecção de SNV (%); (D) *F-measure* da detecção de INDEL (%). Todas as métricas foram calculadas a partir de corridas de 300 ciclos *paired-end*. *F-measure* é a média harmônica da precisão e do recall em relação ao conjunto de variantes esperado.

CONCLUSÕES

Os resultados sugerem que o fluxo de trabalho pode ser migrado do NextSeq-500 para a plataforma MGI DNBSEQ-G400. Esta gerou dados NGS de alta qualidade, evidenciados pela qualidade robusta da base (mediana de 93,06% de bases > Q30), altas taxas de mapeamento (mediana de 99,1%) e forte detecção de SNV (*F-measure*: 99%). A *F-measure* mais baixa para INDELs (82,3%) é esperada, dados os desafios inerentes à detecção de INDEL. Uma limitação deste estudo é a diferença nos números de ciclos de sequenciamento (150 ciclos vs. 300 ciclos), o que pode afetar as comparações diretas. Esses achados corroboram a adoção da plataforma DNBSEQ-G400 em ambientes clínicos.