



Sequenciamento Completo do Exoma (WES)

Metodologia

O Sequenciamento Completo do Exoma (WES, do inglês: *Whole Exome Sequencing*) é uma abordagem baseada no Sequenciamento de Nova Geração (NGS, do inglês: *Next Generation Sequencing*) utilizada para identificar variantes genéticas ligadas a um distúrbio/doença. Essa metodologia analisa as partes do DNA de um indivíduo denominadas éxons; estas regiões fornecem instruções para a produção de proteínas. Uma vez que a maioria das variantes que causam doenças estão presentes nos éxons, acredita-se que o WES seja uma técnica eficiente para identificar variantes causativas que podem levar a uma determinada doença.

O sequenciamento completo do exoma é realizado utilizando-se DNA extraído do sangue, saliva ou tecido do(a) paciente. A partir de DNA fragmentado, aproximadamente 45Mb do genoma, que correspondem a 99% da Sequência Codificante Consenso, são enriquecidos com sondas desenhadas para o genoma humano e otimizadas para a cobertura de todos os genes clinicamente relevante (Nextera DNA Flex Pre-Enrichment Library Prep, Illumina). A biblioteca gerada é sequenciada na plataforma NovaSeq 6000 (Illumina). Os dados de sequenciamento são processados pelo *pipeline* de bioinformática interno da Igenomix (v1.0). Em resumo, os dados brutos são primeiro demultiplexados, os barcodes moleculares são vinculados à ID da amostra, e os adaptadores e leituras de baixa qualidade são eliminados. Após essa etapa, as leituras provenientes são mapeadas com base na sequência referência do genoma humano. Leituras duplicadas são removidas antes da chamada e anotação das variantes. Métricas de qualidade são observadas para cada amostra, de acordo com os seguintes valores de corte: Q30 >80%, profundidade média de cobertura > 100x e porcentagem de bases com cobertura acima de 20x >95%. Amostras com parâmetros eventualmente abaixo dos valores de corte são novamente sequenciadas.

Cobertura do exoma

- Tamanho do painel: ~45Mb, correspondente a 99% da Sequência Codificante de Consenso (bancos de dados: CCDS, RefSeq, GENCODE, ENSEMBL e ClinVar)
- Profundidade média: 157,96 (>100x)
- Cobertura altamente uniforme de todo o exoma (20.000 genes):
 - >97,0% das regiões cobertas em >10x
 - >96,2% das regiões cobertas em >20x
- Leituras: 300 ciclos (2x150 bp)

Cobertura aprimorada de regiões clinicamente relevantes

- 4.555 genes associados a doenças (OMIM):
 - Profundidade média: 189 (>100x)
 - >97,0% das regiões cobertas em >20x
- Variantes patogênicas e provavelmente patogênicas pelo ClinVar:
 - Profundidade média: 201,4 (>200x)
 - >99,4% das regiões-alvo cobertas em >20x
- Variantes patogênicas e provavelmente patogênicas:
 - Profundidade média: 214 (>200x)
 - >99,3% das regiões-alvo cobertas em >20x

Tipos de variantes incluídos no estudo

- O NGS pode detectar com alta sensibilidade e especificidade:
 - SNVs
 - Indels
 - CNVs: ≥15 éxons com sensibilidade >80%
≥500Kb com sensibilidade >95%
- O NGS não pode detectar: regiões homopoliméricas, variantes em regiões de pseudogenes, fusões gênicas, translocações balanceadas, inversões, alterações de ploidia, dissonia uniparental e regiões de repetição expandidas.

A interpretação das variantes é realizada por profissionais qualificados. Todas as variantes com frequência alélica inferior a 1% em bancos de dados populacionais (ExAC, 1000G, gnomAD) são avaliadas. São priorizadas variantes presentes nas regiões codificantes ou em sítios de splice conservados, variantes intrônicas e em UTRs. Uma combinação de critérios de biologia molecular, função gênica, padrão de herança da doença (OMIM), informações de bancos de dados de doenças (HGMD, ClinVar, ClinGen) e pesquisa bibliográfica (PubMed) é utilizada para correlacionar o fenótipo clínico com variantes genéticas. Além de SNVs e pequenos Indels, as variações de número de cópias (CNVs, do inglês: *Copy Number Variation*) são detectadas a partir dos dados de sequenciamento, utilizando a ferramenta *ExomeDepth* (v1.1.15) (PMID: 22942019). Este algoritmo detecta CNVs raras com base na comparação das profundidades de leitura dos dados do teste com um conjunto de dados de referência. Esta análise permite identificar CNVs ≥ 500Kb com mais de 95% de sensibilidade e CNVs com ao menos 15 éxons com mais de 80% de sensibilidade.

A patogenicidade das variantes é determinada utilizando as recomendações do *American College of Medical Genetics* (ACMG) (PMID: 25741868). Observe que a classificação das variantes pode mudar com o tempo. Para verificar se há alguma alteração na classificação das variantes relatadas, entre em contato com a Igenomix.



GPDx

Genômica precisão
diagnóstico

Igenomix[®]
WITH SCIENCE ON YOUR SIDE

As categorias de variantes estabelecidas pela ACMG são as seguintes: Patogênica, Provavelmente Patogênica, Variante de Significado Incerto (VUS, do inglês: *Variant of Unknown Significance*), Provavelmente Benigna e Benigna. Apenas variantes patogênicas, provavelmente patogênicas e de significado incerto (VUS) relacionados à indicação clínica do paciente serão relatadas neste teste genético. De acordo com as diretrizes do ACMG para relatar achados secundários no sequenciamento de exomas clínicos (PMID: 34012068), variantes patogênicas e provavelmente patogênicas nos seguintes genes são relatadas se o consentimento for indicado no Formulário de Requisição de Teste: *ACTA2, ACTC1, ACVRL1, APC, APOB, ATP7B, BMPRIA, BRCA1, BRCA2, BTBD, CACNA1S, CASQ2, COL3A1, DSC2, DSG2, DSP, ENG, FBN1, FLNC, GAA, GLA, HFE, HNF1A, KCNH2, KCNQ1, LDLR, LMNA, MAX, MEN1, MLH1, MSH2, MSH6, MUTYH, MYBPC3, MYH11, MYH7, MYL2, MYL3, NF2, OTC, PALB2, PCSK9, PKP2, PMS2, PRKAG2, PTEN, RB1, RET, RPE65, RYR1, RYR2, SCN5A, SDHA2, SDHB, SDHC, SDHD, SMAD3, SMAD4, STK11, TGFBR1, TGFBR2, TMEM127, TMEM43, TNNI3, TNNT2, TP53, TPM1, TRDN, TSC1, TSC2, TTN, VHL e WT1.*

Aplicações

O WES é utilizado no diagnóstico ou avaliação de uma doença/distúrbio genético em que os resultados de um teste genético podem influenciar o manejo médico e os resultados clínicos de um paciente ou uma família direta ou indiretamente. Com o advento da tecnologia, o sequenciamento tornou-se um processo rotineiro no diagnóstico clínico. Em situações em que a apresentação clínica não é clara e a condição em questão é desconhecida, o sequenciamento e a análise de um pequeno número de genes por vezes é um processo caro e demorado. Isso pode atrasar ainda mais o diagnóstico, o que pode ter impacto na qualidade de vida do paciente.

O WES é uma solução de diagnóstico com melhor custo-benefício, que permite o sequenciamento de dados de ~24.000 genes a partir de uma simples coleta de sangue, saliva ou outras amostras biológicas (sob consulta). O WES examina uma gama mais ampla de genes e variantes, o que é válido especialmente para diagnosticar doenças genéticas associadas a um grande conjunto de genes ou cujas características clínicas e o conhecimento científico atual não permitem direcionar a análise para um conjunto de genes específico.

Limitações

As sondas utilizadas para este teste são projetadas para detectar genes conhecidos no genoma humano. Portanto, este teste é incapaz de detectar genes não contemplados na sequência de referência GRCh37 disponibilizada pelo NCBI ou sequências de genoma não-humano, incluindo sequências virais ou DNA não nuclear. Além disso, devido às limitações das tecnologias de NGS, as seguintes variantes não podem ser prontamente detectadas: regiões homopoliméricas, variantes em regiões de pseudogenes, fusões gênicas, translocações balanceadas, inversões, alterações de ploidia, dissomia uniparental e regiões de repetição expandidas. Além disso, variantes presentes fora dos éxons (região não codificadora) podem ser perdidas; essas variantes podem afetar a atividade gênica e a produção de proteínas, o que pode levar a distúrbios genéticos. Essa técnica não cobre a extensão completa do exoma (a porcentagem de bases com cobertura acima de 20x é de aproximadamente 97%) Certos detalhes sobre as variantes, como mosaïcismo, fases ou ambiguidade de mapeamento também podem não ser determinados com precisão. Limitações analíticas podem ocorrer devido às informações clínicas fornecidas. São necessárias informações clínicas precisas e minuciosas do(s) paciente(s) e familiares, pois informações incompletas podem levar a resultados falso-positivos ou falso-negativos.