QUANTITATIVE DETERMINATION OF *SMN2* COPY NUMBER USING NEXT-GENERATION SEQUENCING AND CORRELATION TO DISEASE SEVERITY

Jody Westbrook[1], Dan Kvitek[1], Erik Gafni[1], Joshua Paul[1], Michael Kennemer[1], Thomas L. Winder[1], Perry B. Shieh[2]

1. Invitae Corporation, San Francisco, CA
2. University of California Los Angeles

INTRODUCTION

Loss of *SMN1* gene function results in spinal muscular atrophy (SMA), a debilitating neuromuscular disorder characterized by the loss of motor neurons in the spinal cord. *SMN1* and its nearly identical gene copy, *SMN2*, are located ~800 kb apart on chromosome 5. The coding regions of *SMN1* and *SMN2* differ by a single nucleotide at c.840 in exon 7, commonly referred to as the gene-determining variant. The alternate nucleotide at c.840 in *SMN2* causes inefficient mRNA processing and consequently reduced protein levels. About 95%–98% of individuals with SMA have zero copies of *SMN1* and about 2%–5% are compound heterozygotes, with a deletion of *SMN1* on one chromosome and a pathogenic sequence variant in *SMN1* on the other chromosome. The number of *SMN2* copies is variable within the population and this number influences the SMA phenotype, with severity decreasing and age of onset increasing as *SMN2* copy number increases. *SMN2* quantitation is also important because it is required for qualification for Spinraza™, an FDA approved treatment for SMA.

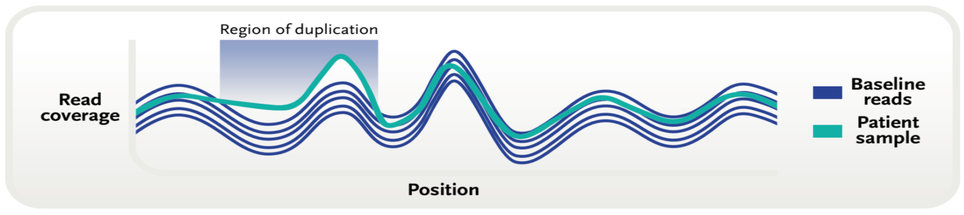
METHODS

SMA is traditionally diagnosed by performing multiplex ligation-dependent probe amplification or quantitative PCR (qPCR) to identify loss of *SMN1* exon 7. These approaches have significant technical limitations and are difficult to efficiently integrate into broader testing. To address these limitations, we developed a next-generation sequencing (NGS)-based approach with a customized bioinformatics solution for simultaneous sequence and copy number variant analysis of *SMN1* and *SMN2*.

Invitae uses a custom, validated algorithm to detect deletions and duplications with NGS. This algorithm detects exon-level deletions and duplications by calculating the statistical likelihood of each copy number state through comparison of the depth of sequence coverage at targeted exons with the depth measured from a set of baseline samples. Due to the high degree of sequence similarity between *SMN1* and *SMN2*, NGS reads derived from both genes are aligned to *SMN1*. The combined *SMN1/2* copy number is determined using Invitae's copy number variant detection algorithm. *SMN1*- and *SMN2*-specific copy number is determined based on the ratio of reads with and without the gene determining variant in exon 7. Sequence variants are also detected by NGS, but their location cannot be unambiguously resolved between *SMN1* or *SMN2*.

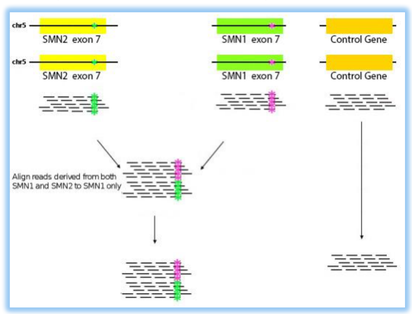
To further validate our method and study *SMN2* genotype-phenotype correlations, 68 patients were tested with our NGS-based method. All patients were followed in the same neuromuscular clinic and had previously been tested in a different commercial laboratory using qPCR.

Read-depth approach to deletion/duplication analysis by NGS (example of a duplication)



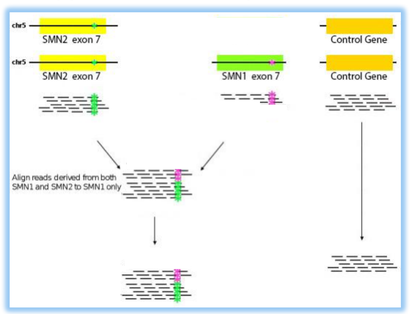
NGS-based method for detecting SMN1 and SMN2 copy number variants

2 copies of *SMN1* & 2 copies of *SMN2*



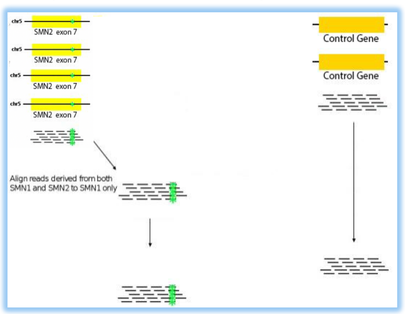
- 4 total copies of *SMN1* and *SMN2*
- 1:1 (*SMN1*:*SMN2*) ratio of the gene-determining variant

1 copy of *SMN1* & 2 copies of *SMN2*



- 3 total copies of *SMN1* and *SMN2*
- 1:2 (*SMN1*:*SMN2*) ratio of the gene-determining variant

0 copies of *SMN1* & 4 copies of *SMN2*



- 4 total copies of *SMN1* and *SMN2*
- 0:1 (*SMN1*:*SMN2*) ratio of the gene-determining variant

RESULTS

Concordance between NGS and qPCR methods for detection of *SMN1* and *SMN2* copy number

Concordance of SMN1 copy number results

SMN1 copy number	Number of patients detected by NGS	Discrepancies with qPCR	Concordance
0	66	0	100%
1	2	0	100%
		Total concordance	100%

Concordance of SMN2 copy number results

SMN2 copy number	Number of patients detected by NGS	Discrepancies with qPCR	Concordance
1	2	0	100%
2	14	1	93%
3	45	4	91%
4	7	0	100%
		Total concordance	93%

- 100% concordance was found for *SMN1* copy number.
- 5 out of 68 patients had discrepancies in *SMN2* copy number
- In discrepant cases, the NGS results were more consistent with the patients' phenotypes.

Details of discrepant results

Case 1
- Clinical phenotype: SMA type 1
- NGS *SMN2* copy number: 2
- qPCR *SMN2* copy number: atypical (between 2-3 copies)

Case 2
- Clinical phenotype: SMA type 2
- NGS *SMN2* copy number: 3
- qPCR *SMN2* copy number: 2

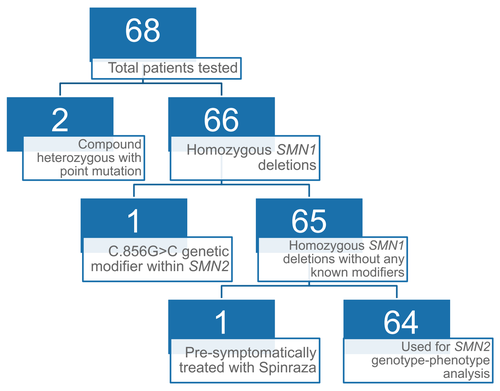
Case 3
- Clinical phenotype: SMA type 2
- NGS *SMN2* copy number: 3
- qPCR *SMN2* copy number: 2

Case 4
- Clinical phenotype: SMA type 2
- NGS *SMN2* copy number: 3
- qPCR *SMN2* copy number: atypical (between 2-3 copies)

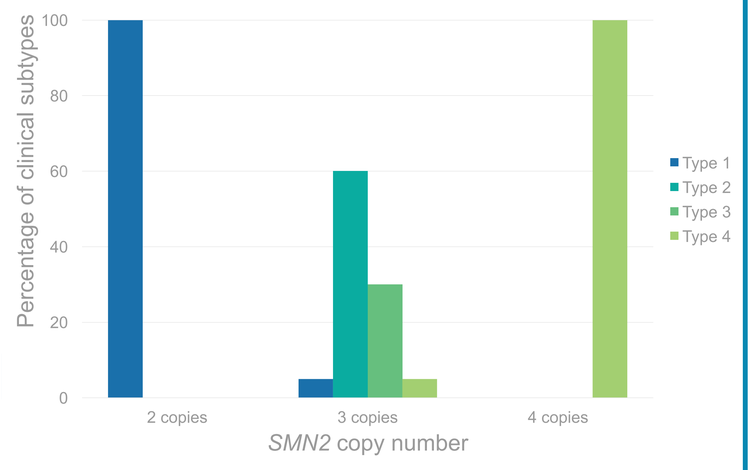
Case 5
- Clinical phenotype: SMA type 2
- NGS *SMN2* copy number: 3
- qPCR *SMN2* copy number: >3

Correlation between *SMN2* copy number and SMA clinical subtype

Cohort for analysis of clinical subtype and SMN2 copy number



Predictive value of SMN2 copy number



CONCLUSIONS

- We found a concordance rate of 100% for *SMN1* copy number and 93% for *SMN2* copy number between our NGS-based method and qPCR performed at a different commercial laboratory. In cases where there was disagreement, the clinical phenotypes of the patients were more consistent with the *SMN2* results from the NGS method than from qPCR testing.
- In addition to accurately determining *SMN1* and *SMN2* copy number, the NGS-based method can also detect sequence variants and genetic modifiers within the *SMN1* and *SMN2* genes.
- Because clinical phenotypes of SMA will become blurred by disease-modifying therapies, *SMN2* copy number may emerge as an important tool for classifying SMA as well as predicting disease progression and therapeutic response.

References: 1. Mailman MD *et al*. Molecular analysis of spinal muscular atrophy and modification of the phenotype by *SMN2*. *Genet Med*. 2002;4:20–6. PMID: 11839954
2. Swoboda KJ *et al*. Natural history of denervation in SMA: Relation to age, *SMN2* copy number, and function. *Ann Neurol*. 2005;57:704–12. PMID: 15852397

Disclosures: *JW, DK, EG, JP, MK, and TW are employees and shareholders of Invitae.*