NGS-based carrier screen for Gaucher's disease calls variants and detects large rearrangements between *GBA* and *GBAP1*

Kevin D'Auria, Mark Theilmann, Kevin Iori, Clement S. Chu, Imran S. Haque, Eric A. Evans, H. Peter Kang, Jared Maguire, Dale Muzzey

Counsyl

All Counsyl posters available online at **research.counsyl.com**

GBA variants can't be called correctly without deconvolving the gene and pseudogene. We do it with NGS.

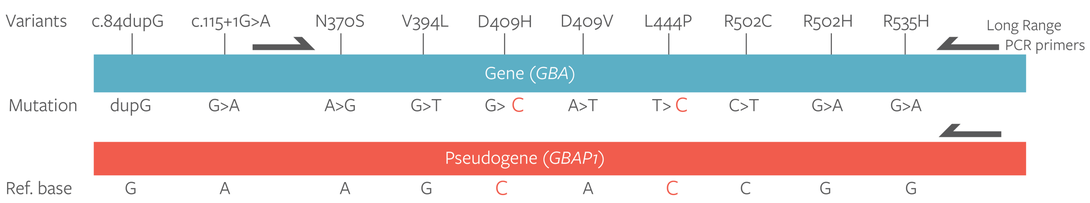
Introduction

With a carrier rate of 1 in 18 in the Ashkenazi Jewish population, Gaucher's disease is included on most carrier-screening panels, yet 95% homology between the disease-associated gene (*GBA*) and its pseudogene (*GBAP1*) complicates genotyping by NGS alone. Variant L444P, which is present in 30% of the affected population, is particularly difficult to assess because the mutant base in *GBA* is the normal base in the pseudogene. As such, low-throughput and costly assays — including allele-specific qPCR or long-range PCR followed by Sanger sequencing — have been used to resolve this challenging variant and others, such as D409H, that are also pseudogene-derived. Here we present an entirely NGS-based method that detects ten of the most common deleterious *GBA* variants, giving a 97% carrier detection rate.

Methods

Hybrid-capture probes were designed to enrich for DNA fragments spanning the variants of interest. To ensure that only GBA-derived reads contributed to variant calling in highly homologous regions, reads were discarded if they did not contain a sufficient number of GBA-derived bases for unambiguous mapping by the alignment software. Rearrangements in the locus were detected by combining a depth-based copy-number analysis and an assessment of allele balance at each variant position. NGS results for variants L444P and D409H were validated using a combination of long-range PCR and allele-specific, endpoint TaqMan PCR.

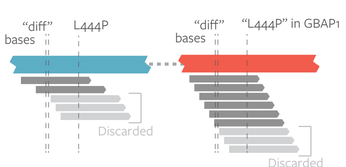
GBA variants



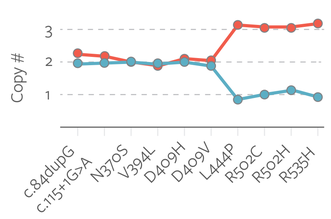
Example with a *GBA-GBAP1* fusion gene

NGS screen

1. Hybrid capture of regions of interest
2. Discard ambiguous alignments



3. Copy number analysis



Indicates:

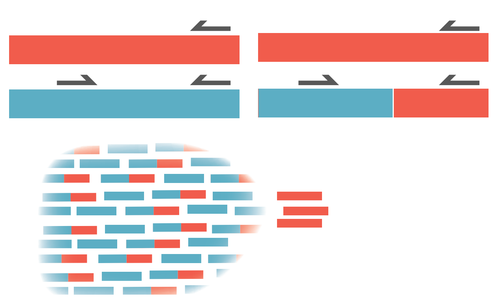
2 *GBAP1* copies,

1 *GBA* copy, &
1 fusion gene

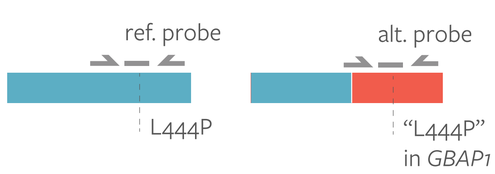
4. Check allele balance for SNPs and false positive SNPs (see right column)

Long Range PCR + TaqMan (validation method)

1. Enrich for *GBA* or *GBA-GBAP1* fusions with long range PCR



2. Endpoint, allelic discrimination TaqMan PCR of long range PCR products produces a "mutant" signal from the fusion gene and a "reference" signal from *GBA*



Results

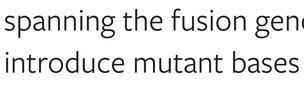
Relative to the PCR assay, our NGS test correctly genotyped the L444P and D409H variants in all 73,339 samples tested.

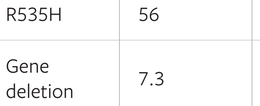
	L444P			D409H		
	FP	TP	FN	FP	TP	FN
Our method	0	121	0	0	10	0
GBA-only endpoint PCR*	0	84	31	0	9	1
Standard SNP caller (NGS)	434	84	31	9	9	1

**expected | Concordance with validation method*

Remarkably, 28% of L444P carriers resulted from fusion genes and would have been false negatives using either a standard SNP caller on NGS data or a qPCR assay with *GBA*-derived primer sequences**, underscoring the importance of our custom rearrangement-detection algorithm.

*** Our downstream LR-PCR primer could hybridize GBA and GBAP1 to detect fusion genes. These and the following rates are based off 109,244 samples using our validated assay.*

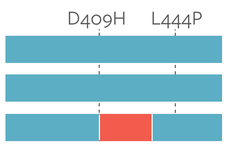
	Variant	
Carrier type	L444P	D409H
 Deleterious fusion gene	28%	25%*

	Variant	
Carrier type	L444P	D409H
 Deleterious SNP	72%	75%

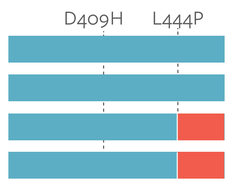
**The breakpoint for these fusion genes caused L444P + D409H carrier status. Our data suggests but cannot definitively prove the mutations are in* cis.

Our assay greatly improved the specificity of calling L444P and D409H, identifying 77% and 33% (respectively) of heterozygous calls from a standard SNP caller as false positives.

Examples that would be called as carriers using allele balance alone:



D409H allele balance ~67%
L444P allele balance ~100%



D409H allele balance ~100%
L444P allele balance ~50%

Since the mutant base of D409H and L444P is pseudogene-derived, reads spanning the fusion gene breakpoints introduce mutant bases into the *GBA* pileups.

Our assay further improved carrier detection by identifying partial or full *GBA* deletions.

Variant	Carriers per 100k	Carriers per 100k Exome Sequencing Project (All)
c.84dupG	20	N/A
c.115+1G>A	34	N/A
N370S	770	400 (24/6502)
V394L	5.4	N/A
D409H	26	30 (2/6503)
D409V	none found	N/A
L444P	160	60 (4/6453)
R502C	16	20 (1/6503)
R502H	2.7	N/A
R535H	56	30 (2/6428)
Gene deletion	7.3	N/A

Conclusions

NGS sequencing alone, with proper probe design and analysis, can accurately genotype all common *GBA* variants, even those that are pseudogene-derived and nested in highly homologous regions.