Greenwood Genetic Center
Where Compassion Inspires Progress
www.GGC.org

The Development and Validation of a Next Generation Sequencing Panel for Skeletal Dysplasias

Monica J. Basehore, Stephen McGee, Katharine Kubiak, Kellie King, Jamie Butler, Jennifer A. Lee, Julie R. Jones, and Michael J. Friez
Greenwood Genetic Center, Greenwood, South Carolina, 29646 USA

INTRODUCTION

- Skeletal dysplasias, as a whole, represent a large and diverse group that consists of more than 450 recognized conditions. Since there is a great deal of phenotypic overlap and genetic heterogeneity among the skeletal dysplasias, a diagnosis is often made based on a combination of clinical and radiological findings that is further supported by molecular and, in some cases, biochemical findings.

- Using RainDance™ microdroplet enrichment and the SOLiD™ Next Generation Sequencing (NGS) platform, we have developed and validated an assay that can simultaneously analyze ten genes known for their association with various skeletal dysplasias. Collectively, these genes account for more than 30 distinct clinical phenotypes, and an estimated 90% of individuals with a skeletal dysplasia have a mutation in one of these ten genes. The current design of our skeletal dysplasia panel includes the coding and flanking intronic regions for all ten genes.

- **Genes included on the panel**:

COL1A1	***COL2A1***	***SLC26A2***	***FLNA***	***SOX9***
COL1A2	***COMP***	***FGFR3***	***HSPG2***	***TRPV4***

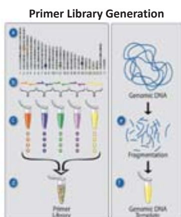
MATERIALS AND METHODS

RainDance™ Enrichment:

- 5 µg of each genomic DNA sample was fragmented to 4-6 kb using a Covaris instrument (Woburn, MA). RainDance™ enrichment for the 10 genes was performed on the RDT1000 instrument (RainDance™ Technologies, Lexington, MA).

Figure 1. Enrichment of Targeted Regions Using RainDance™ Technology

Primer Library Generation



(a) Identify targeted sequences of interest in the genome.

(b) Design and synthesize forward and reverse primer pairs for each targeted sequence.

(c) A microfluidic chip is used to encapsulate the aqueous PCR primers in inert fluorinated carrier oil with a block-copolymer surfactant to generate the equivalent of a picoliter scale test tube compatible with standard molecular biology.

(d) Primer pair droplets are mixed together so that each library element has an equal representation.

(e) Genomic DNA is fragmented and purified.

(f) Purified genomic DNA is mixed together with all of the components of the PCR reaction except the PCR primers.

Primer-Template Merge and PCR

PCR Library
Droplet PCR
Break Emulsion
DNA Clean-Up
Sequence

(g) Primer Library droplets (~8 pL) are dispensed to the microfluidic chip.

(h) Genomic DNA Template is delivered as an aqueous solution, and template droplets (~18 pL) are formed within the microfluidic chip. The primer pair droplets and template droplets are then paired together in a 1:1 ratio.

(i) Paired droplets flow through the channel of the microfluidic chip to pass through a merge area where an electric field induces the two discrete droplets to coalesce into a single PCR droplet (~26 pL). Up to 2 million PCR droplets are collected into a single 0.2 ml PCR tube. The collection of PCR droplets (PCR Library) is processed in a standard thermal cycler for targeted amplification, followed by breaking the emulsion of PCR droplets to release the PCR amplicons into solution for purification and next generation sequencing.

Next Generation Sequencing on a SOLiD™ 5500xL Platform:

- Standard fragment libraries for analysis by SOLiD™ next generation sequencing (NGS) were prepared for each sample. Library amplification was performed using emulsion PCR, and the products of the emulsion PCR were purified (AMPure) and then deposited onto a glass slide for analysis by the 5500xL system (Applied Biosystems, Foster City, CA).

MATERIALS AND METHODS (CONTINUED)

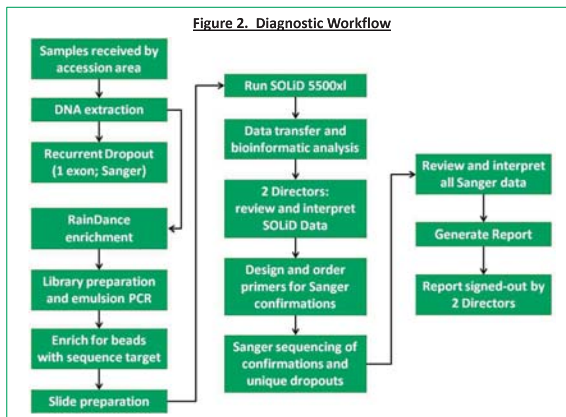
Bioinformatic Analysis and Filtering:

- Primary and secondary analyses were performed on the 5500xL. Subsequent tertiary analysis was performed using NextGENe® software (SoftGenetics, State College, PA). The data was then reviewed with an initial emphasis on novel alterations and those reported in the Human Gene Mutation Database (HGMD). The alterations detected were cross-referenced to the other samples within the same run as well as to a cumulative database of all the sample data from previous runs. Filtering out the known SNPs allowed us to determine the changes that warranted further consideration and follow-up testing.

Sanger Sequencing:

- All changes deemed of potential clinical relevance were confirmed by Sanger sequencing. Additionally, a single amplicon of the *HSPG2* gene (exon 1) routinely failed to amplify or had poor coverage by NGS, so this region was Sanger sequenced in each patient (Recurrent Dropout; see Figure 2).

Figure 2. Diagnostic Workflow



RESULTS: TABLE 1

Summary of Validation Phase of Study

Patient	Suspected Condition/Gene	Status prior to NGS Testing	NGS Testing Results
1	Stickler Syndrome/*COL2A1*	Known molecular diagnosis	*COL2A1* c.2926_2927ins4
2	Campomelic Dysplasia/*SOX9*	Known molecular diagnosis	*SOX9* c.508C>G (p.P170A)
3	Osteogenesis Imperfecta/*COL1A2*	Known molecular diagnosis	*COL1A2* c.2422G>A (p.G808S)
4	Thanatophoric Dysplasia Type 1/*FGFR3*	Known molecular diagnosis	*FGFR3* c.742C>T (p.R248C)
5	Thanatophoric Dysplasia Type 2/*FGFR3*	Known molecular diagnosis	*FGFR3* c.1948A>G (p.K650E)
6	Spondyloepiphyseal Dysplasia/*COL2A1*	Known molecular diagnosis	*COL2A1* c.3627_3638del12
7	Achondroplasia/*FGFR3*	Known molecular diagnosis	*FGFR3* c.1138G>A (p.G380R)
8	Spondyloepiphyseal Dysplasia/*COL2A1*	Unknown†	*COL2A1* c.1537G>A (p.G513S) + 1 VUS
9	Pseudoachondroplasia/*COMP*	Unknown†	*COMP* c.1450T>G (p.C484G)
10	Pseudoachondroplasia/*COMP*	Unknown†	*COMP* c.1051T>C (p.C351R) + 1 VUS
11	Odontochondrodysplasia	Unknown	*HSPG2* c.3303-7C>T (VUS)
12	Unknown	Unknown	*COL1A1* c.1461+13G>T (VUS)
13	Unknown	Unknown	*HSPG2* c.8848G>A (p.G2950R; VUS) *HSPG2* c.8929C>T (p.R2977W; VUS)
14	Unknown	Unknown	*COL2A1* c.2680-9C>T (VUS) *FLNA* c.7174T>C (p.F2392L; VUS) *HSPG2* c.8545G>T (p.V2849L; VUS)
15	Schmid Metaphyseal Chondrodysplasia	Unknown	Normal (no changes to report)
16	Unknown	Unknown	*COL1A1* c.4018G>A (p.G1340S; VUS)

Sixteen patients were analyzed during the validation phase of this assay. Seven of these patients had a known molecular diagnosis prior to analysis, and the finding detected by our NGS panel was concordant with the previous molecular result. The remaining nine patients had an unknown suspected skeletal dysplasia. †For three of these patients, the NGS finding was consistent with the clinical phenotype and is, therefore, considered to be likely pathogenic. All changes were detected in the heterozygous state. VUS = variant of unknown clinical significance.

RESULTS: TABLE 2

Summary of Findings in Patients with Undiagnosed Skeletal Dysplasia (n=42)

Interpretation Category	Number	Percent
Patients with a Pathogenic Alteration(s)	8	19.0
Patients with a Likely Pathogenic Alteration(s)	11	26.3
Patients with $\geq$ 1 Variant of Unknown Clinical Significance (VUS)	15	35.7
Patients with a Normal Result (no molecular findings)	8	19.0

To date, 42 patients with undiagnosed skeletal dysplasia have been tested and analyzed by our NGS panel. This cohort includes the nine unknown patients tested during the validation phase of the study.

RESULTS: TABLE 3

Pathogenic and Likely Pathogenic Alterations Detected

Gene	Inheritance Pattern	Patient	Alteration Detected	Type	Pathogenicity
COL1A1	Autosomal Dominant	12	c.3641G>A (p.R1214H)	Missense	Pathogenic*
		18	c.2155G>A (p.G719S)	Missense	Likely Pathogenic
COL1A2	Autosomal Dominant	14	c.650G>A (p.G217E)	Missense	Likely Pathogenic
COL2A1	Autosomal Dominant	3	c.1537G>A (p.G513S)	Missense	Likely Pathogenic
		4	c.2710C>T (p.R904C)	Missense	Likely Pathogenic
		6	c.3121G>A (p.G1041S)	Missense	Likely Pathogenic
		7	c.1933G>A (p.G645S)	Missense	Likely Pathogenic
		8	c.2114G>A (p.G705D)	Missense	Likely Pathogenic
		9	c.1151G>A (p.G384D)	Missense	Pathogenic*
COMP	Autosomal Dominant	1	c.1051T>C (p.C351R)	Missense	Likely Pathogenic
		2	c.1450T>G (p.C484G)	Missense	Likely Pathogenic
		10	c.925G>C (p.G309R)	Missense	Pathogenic£°
FGFR3	Autosomal Dominant	15	c.742C>T (p.R248C)	Missense	Pathogenic‡
		16	c.742C>T (p.R248C)	Missense	Pathogenic‡
HSPG2	Autosomal Recessive	5	c.1219_1220insC	Frameshift	Pathogenic†
			c.12899G>A (p.R4300Q)	Missense	Likely Pathogenic
		17	c.1057C>T (p.R353X)	Nonsense	Pathogenic†£
			c.5014+1G>A	Splice site	Pathogenic†£
		19	c.5755C>T (p.R1919C)	Missense	Likely Pathogenic
			c.8848G>A (p.G2950R)	Missense	Likely Pathogenic
SLC26A2	Autosomal Recessive	11	c.428C>A (p.T143K)	Missense	Likely Pathogenic
			c.835C>T (p.R279W)	Missense	Likely Pathogenic
SOX9	Autosomal Dominant	13	c.1180C>T (p.R394X)	Nonsense	Pathogenic*†£

Of the 42 patients with undiagnosed skeletal dysplasia, 19 (45.2%) had at least one pathogenic or likely pathogenic alteration. These alterations are spread across eight of the ten genes included in the panel. Rationale for classifying the nine pathogenic alterations, as such, is as follows: ‡previously reported, known mutation; °different alteration(s) at the same codon reported in HGMD; †expected to cause disease based on ACMG guidelines; *proven to be *de novo* after parental testing; or £consistent with the family history and/or patient's reported phenotype. All changes were detected in the heterozygous state. Variants of unknown clinical significance, if identified in addition to the above findings, are not shown. Patients 1-3 shown here were tested during the validation phase of the study.

SUMMARY AND CONCLUSIONS

- **This diagnostic NGS panel has a high detection rate for patients suspected of having a skeletal dysplasia, as a pathogenic or likely pathogenic alteration was detected in 19/42 patients (45.2%).**

- **Panel performance:**
  - The average coverage of the targeted regions = 824X
  - The average span (exon) coverage = 99.63%
  - Size of panel = 45,141 basepairs

- **Parental testing is necessary to determine if alterations are *de novo* (AD) or *in trans* (AR), which helps confirm, or rule out, pathogenicity.**

- **These results support the notion that NGS testing has advantages over Sanger sequencing, even for a small panel comprised of only ten genes.**

REFERENCES

- Warman et al. (2010). Nosology and Classification of Genetic Skeletal Disorders: 2010 Revision. *Am J Med Genet Part A* 155:943-968.