DOSAGE ANALYSIS BY NEXT GENERATION SEQUENCING AND MICROARRAY CGH INDICATES PUTATIVE PROCESSED PSEUDOGENES IN *SMAD4* AND *NBN*

Debora Mancini-DiNardo, PhD; Thad Judkins, BS; Maria C. Elias, PhD; Courtney Daniels, BS; Ammon Marshall, BS; Jayson Holladay, BS; Yaping Ryan Qian, PhD; Benjamin Roa, PhD
Myriad Genetic Laboratories, Inc., Salt Lake City, UT

BACKGROUND

- The presence of pseudogenes has the potential to confound the interpretation of genetic testing results.
- The focus of this analysis was to investigate the presence of processed pseudogenes in the *SMAD4* and *NBN* genes not present in the reference genome in patients who underwent genetic testing with a pan-cancer panel.
- Genetic testing for *SMAD4* is performed to identify patients with Juvenile Polyposis Syndrome (JPS) and Hereditary Hemorrhagic Telangiectasia. Patients with JPS have a high risk for cancer as a result of hamartomatous polyps in the gastrointestinal system, particularly in the colon, rectum and stomach, as well as an elevated risk for small bowel and pancreatic cancer.
- Individuals with mutations in the *NBN* gene carry an increased risk for breast cancer in women, and prostate cancer in men.

METHODS

- A commercial laboratory database was queried to identify patients who underwent testing with a pan-cancer multi-gene panel, which included *SMAD4* and *NBN*, between September 2013 and March 2015.
- All patient data regarding clinical history was obtained by health care provider report on test requisition forms.
- Large rearrangement (LR) analysis was performed using Next Generation Sequencing (NGS), and targeted microarray CGH for *SMAD4* and *NBN* as part of the panel test. Additional testing by MLPA was used to investigate unusual dosage results for *SMAD4*.

RESULTS

- The collective evidence from MLPA, microarray, and NGS are consistent with the presence of a novel pseudogene showing high sequence homology to the exonic regions of the *SMAD4* gene.
- This pseudogene appears to be processed because:
  - All elevated probes on MLPA are completely exonic (Figure 1).
  - Only exonic probes show an increase in dosage on microarray (Figure 2, left).
  - Elevated amplicons on NGS LR analysis utilize primer pairs that are completely exonic (Figure 3, top).
  - NGS sequencing was designed to detect a 95bp deletion, reflecting the absence of intron 5 sequence in the pseudogene (Figure 4).
- 200 patients tested showed evidence of the putative *SMAD4* pseudogene.
  - 7 reported a colon cancer diagnosis, but none reported a diagnosis of JPS.
  - 12 carried pathogenic mutations in genes other than *SMAD4*.
  - There was no common ancestry among these 200 patients.
- We also detected aberrant gene dosage patterns in the *NBN* gene for 2 patients, which is suggestive of a putative *NBN* pseudogene by the same criteria as *SMAD4*.
- Microarray and NGS LR results for *NBN* (Figure 2, right and Figure 3, bottom) support the presence of a processed pseudogene, since only exonic regions are affected. This is similar to the evidence observed for *SMAD4*.

Figure 1. *Representative example of the SMAD4 anomalous results by MLPA analysis. Red boxes represent exons showing an increase in dosage. Green boxes represent exons that are present at normal dosage levels. Blue boxes represent control probes.*

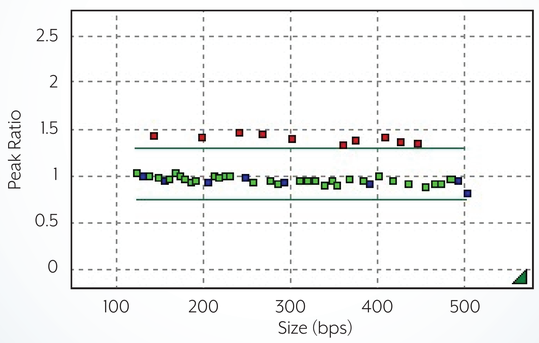


Figure 2. *Representative examples of microarray CGH evidence of duplications in SMAD4 (left) and NBN (right). A normal exon (yellow bar) is represented by a cluster of probes (black circles) centered at zero. Enlarged views of duplicated exons show that exonic probes within the region denoted by the yellow bar are elevated while the flanking intronic probes contained within the gray flanking bars are within the normal range.*

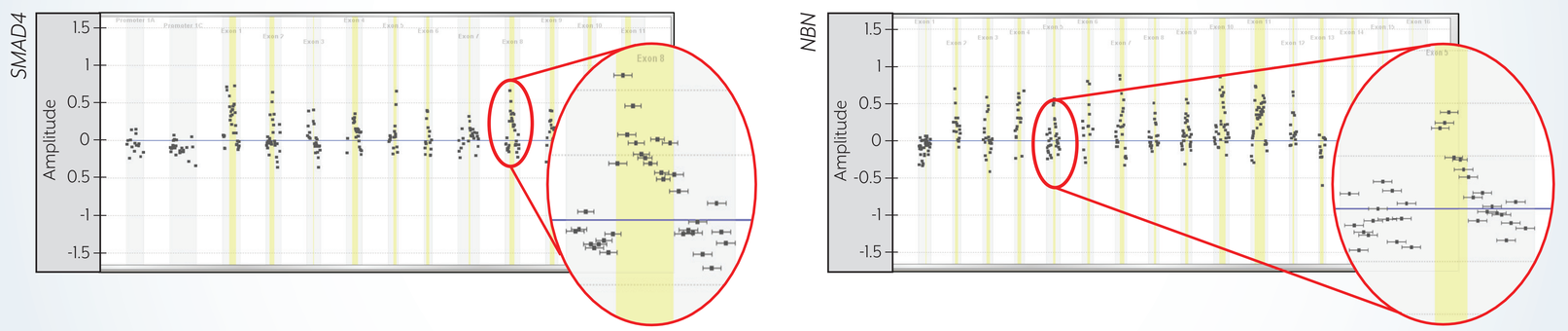


Figure 3. *Representative examples of elevated dosage detected in SMAD4 (top) and NBN (bottom) using NGS. Red squares represent the average dosage of several overlapping amplicons for each exon. Exons at normal dosage align to 2 on the Y-axis, 1 for deleted exons, and 3 for duplicated exons. Enlarged views show increased dosage in amplicons that utilize primers pairs located in exonic regions.*

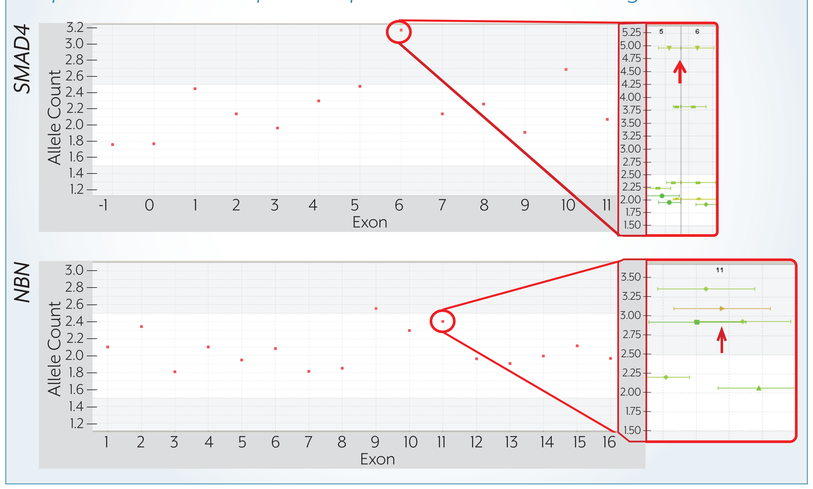
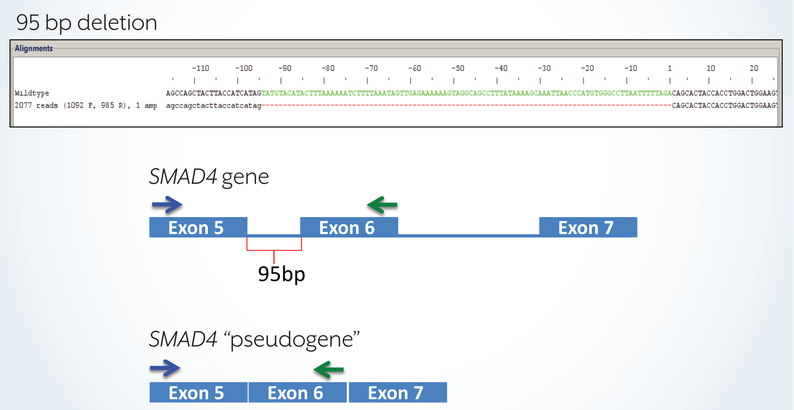


Figure 4. *NGS sequencing results for SMAD4. A 95bp deletion in the region spanning exons 5-6 of SMAD4 (top) is derived from the putative pseudogene, which is devoid of intronic sequences (bottom).*



CONCLUSIONS

- Available evidence strongly supports the presence of pseudogenes in *SMAD4* and *NBN*.
- The presence of such pseudogenes has the potential to confound the interpretation of genetic testing results, as only mutations in the native gene are clinically significant.
- Studies are ongoing to confirm the structure and locations of these putative pseudogenes.

Presented at ACMG - March 26, 2015