Greenwood Genetic Center FDNA

VALIDATION STUDY OF THE INTEGRATION OF NEXT-GENERATION PHENOTYPING IN EXOME ANALYSIS

A collaboration between Greenwood Genetic Center (SC, USA))and FDNA (MA, USA)

Summary

Next-generation phenotyping (NGP) technologies capture, structure, and analyze complex phenotypic information to produce actionable genomic insights. DeepGestalt is an NGP technology that currently supports more than 300 specific genetic syndromes and syndrome groups, representing 45% of cases solved by WES. Next-generation sequencing (NGS) augmented by NGP results in more efficient and accurate diagnoses.

In this study, we show that:

- Integration of NGP in the variant analysis workflow dramatically increases diagnostic yield and efficiency as compared to traditional methods of variant prioritization, placing causative genes in ≤ 5th rank in 50% of cases;
- NGP increases the number of top-ranked causative genes in both "easily solved" and "challenging" cases, with an overall improvement in ranking across all cases;
- NGP ranking among cases supported by DeepGestalt is even more impressive, showing 67% ranked in the top 5, compared to 11% based on GGC and CADD ranks;
- NGP may enable prioritization of causative variants independent of parental data, decreasing costs by reducing the need to complete parental WES and increasing diagnostic yield by enabling diagnosis of cases where parental samples are unavailable;
- Expanding NGP analysis beyond the face (brain MRIs, skeletal radiographs, fundoscopy) will increase the amount of syndromes supported.

Introduction and background

Exome sequencing results in a vast set of possibly-relevant variants which must be filtered and prioritized. With the recent introduction and increasing use of clinical genome sequencing, this challenge will only increase (Lionel et al, 2018, Smith et al, 2019). In some settings, NGS is accessible enough that it has become a good initial testing option (avoiding the costs associated with a series of stepwise tests), but variant interpretation and clinical correlation require the integration of rich phenotypic data. The use of patient phenotypic data is an essential annotation to support variant prioritization and filtering, enabling variant analysts to more efficiently identify possibly relevant variants. More than ever, next-generation sequencing (NGS) requires next-generation phenotyping (NGP) (Hennekam, Biesecker 2012).

Meanwhile, in settings where exome or genome analysis is not easily accessible, the ability to target and prioritize analysis of specific genes based on phenotypic presentation for targeted gene panels is even more important.

In 2014, FDNA launched the Face2Gene (F2G) application (FDNA, Inc., USA), which leverages an NGP analysis of various phenotypic signals. The facial analysis technology is called DeepGestalt. As described by Gurovich et al (2019), the DeepGestalt technology uses deep learning convolutional neural networks to measure similarities between patient photos and hundreds of syndrome models. Based on an input of 2D frontal facial photographs, clinical notes, and/or HPO phenotypic features (Robinson et al 2014), F2G presents clinicians with a list of the 30 best-matched syndromes for

consideration. The use of F2G's automated facial analysis is useful in different stages of the diagnostic workup, including establishing a differential or clinical diagnosis (Basel-Vanagaite et al. 2016; Martinez-Monseny et al., 2018), ordering the most relevant genetic test (Liehr et al., 2017), and making a diagnosis following exome analysis (Gripp et al. 2016). Its value has been demonstrated in varying ethnicities (Vorravanpreecha et al, 2018; Pantel et al, 2018; Mishima et al., 2019).

DeepGestalt is currently trained to support about 45% of the cases diagnosed in a series of consecutive exome cases (as listed in "Supplemental Table 1" in Retterer et al. 2016). In cases for which DeepGestalt has not been trained yet, syndrome suggestions can be derived through other tools, such as feature matching based on HPO feature extraction from free text. Additional NGP tools utilizing MRI, radiographs, and fundoscopy are under development (Hanani et al. 2019).

The results of NGP analyses such as F2G's facial analysis can be integrated efficiently into the variant interpretation process, enabling prioritization of variants and genes with the strongest phenotypic relevance. The PEDIA algorithm (Hsieh et al. 2019) is the first framework to integrate NGP scores with molecular data for variant prioritization. A recent retrospective study showed the PEDIA algorithm added value to the diagnostic yield of a cohort of 679 patients with 105 different monogenic disorders (Hsieh et al. 2019), improving the top-1 accuracy rate for the disease causing gene to 89%. The top-10 accuracy rate is 99%. This increased accuracy demonstrates that NGP results can be used similarly to deleteriousness scores on the molecular level. This could allow the "matching phenotype" PP4 criterion in the ACMG variant interpretation guidelines to move from a dichotomous assessment to a quantifiable score.

Integration of NGP into the molecular diagnostic process greatly impacts diagnostic efficiency and yield. In this study, we evaluated the potential impact of using NGP in the analysis of NGS data. We focused on a retrospective set of 126 clinical WES cases from the Greenwood Genetic Center (GGC) Molecular Diagnostic Laboratory (SC, USA) which were either diagnosed or remain undiagnosed based on conventional variant analysis processes using a proprietary in-house pipeline. We applied NGP to each case and compared the traditional methods of variant prioritization to NGP-enriched variant prioritization. Based on this comparison, we assess how often the inclusion of NGP promotes improved diagnostic efficiency.

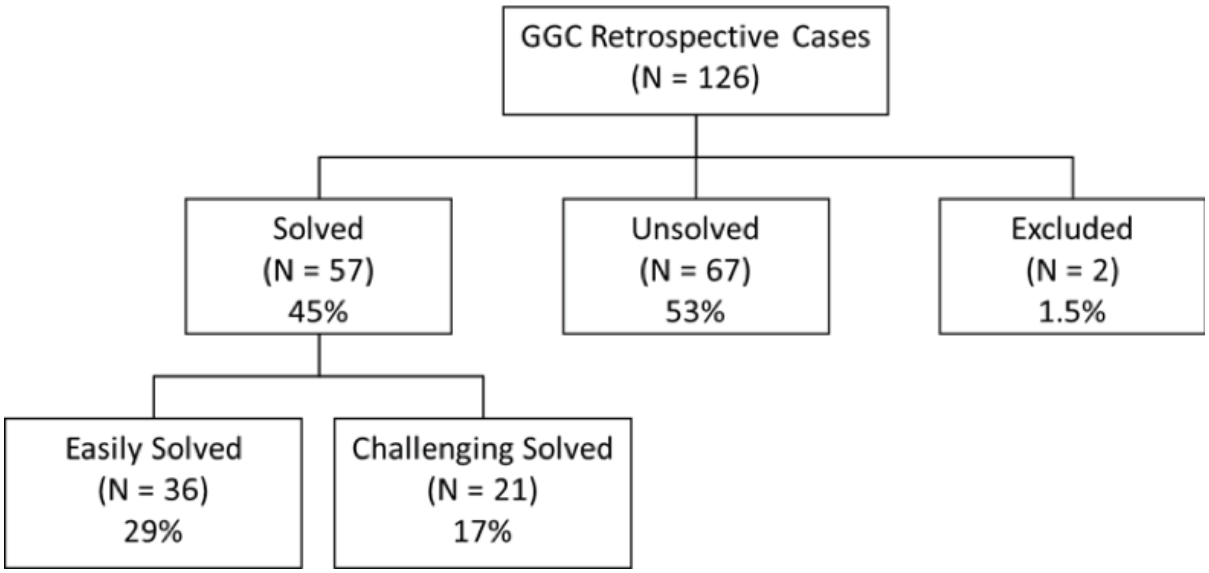
Methods

Cohorts

Retrospective data were provided for analysis by the Greenwood Genetic Center. Each case had to include a frontal facial photo consented for analysis in the F2G app, a selection of clinician-described clinical features (mapped to HPO), and a VCF for analysis. Cases were stratified into "solved" (N=57), "unsolved" (N=67), and 2 cases were excluded due to a diagnosis being made through means other than variant analysis from WES. Cases in the "solved" cohort included those for which the causative variant was identified using GGC's conventional variant analysis process (GGC rank). Cases were characterized based on whether they were easily solved or challenging. Among the 57 diagnosed cases, 36 cases were considered "easily solved", meaning that the diagnostic variant was identified easily and 21 cases were considered "challenging". Within the subset of "challenging" cases, 15 cases were defined as "challenging without analysis of parental samples", meaning the variant was initially

identified as a result of analysis of parental samples and would not be easily identified if parental data was unavailable. Cases in the "unsolved" cohort were those which did not yet have a known causative variant.

Figure 1. Study cohort flow chart



Of the 57 diagnosed cases, 28 were diagnosed with a syndrome represented by a validated facial model in the Face2Gene app.

For cases in the "solved" cohort, known causative variants were provided. For 40 of the 57 solved cases, data indicating the number of variants evaluated by the GGC team before arriving at the causative variant (representing a GGC "rank") was available. For the remaining 17 "solved" cases, GGC rank data was not available.

Analyses

We generated NGP scores for the genes and variants in each "solved" case. All cases were analyzed with Face2Gene algorithms, including the DeepGestalt and Feature Match algorithm. These algorithms enabled ranking and scoring of syndromes based on facial gestalt (G score) and feature similarity (F score), resulting in a ranked list of syndromes and a ranked list of genes based on known gene associations for each syndrome. Upon review of the VCF file, each variant within a given gene was assigned the associated G and F scores and ranks.

Additionally, a CADD (C) score (Kircher et al., 2014) was generated for each variant. The CADD score combines several types of variant-specific data, and is often used as a gauge of the variant's potential pathogenicity. However, it does not consider relevance to the patient's specific phenotypic presentation. The CADD score per gene was calculated using the highest CADD score of the variants in that gene.

Finally, VCFs for each case were also analyzed using the PEDIA score (Hsieh et al, 2019), as an example for possible integration of G and F scores with molecular data. The PEDIA score integrates several datasets for each variant, including G and F scores, BOQA and Phenomizer scores (Koehler et al, 2009), and a CADD score. A PEDIA score is generated for each gene, using the CADD score from the variant in that gene with the highest CADD score.

Genes for each case were sorted according to the PEDIA score described above, giving each gene a PEDIA rank. We then compared the PEDIA rank for the known causative gene to the original GGC rank for the causative variant (when available) and the CADD rank for the causative gene. (See Appendix 1 for example cases with their associated ranks.)

Results

Comparing PEDIA rank to GGC and CADD rank

PEDIA rank placed the causative gene in the first position in 11/40 cases (27.5%), ≤ 5th position in 20/40 cases (50%). This was notably improved compared to the ranks based on GGC and CADD. Improvement was more dramatic for cases diagnosed with a syndrome represented by a facial model validated in the Face2Gene app (see Tables 1 and 2 and Figures 2 and 3).

Figure 2. ROC curves showing ranking of causative genes based on GGC, CADD, and PEDIA ranks for 40 cases

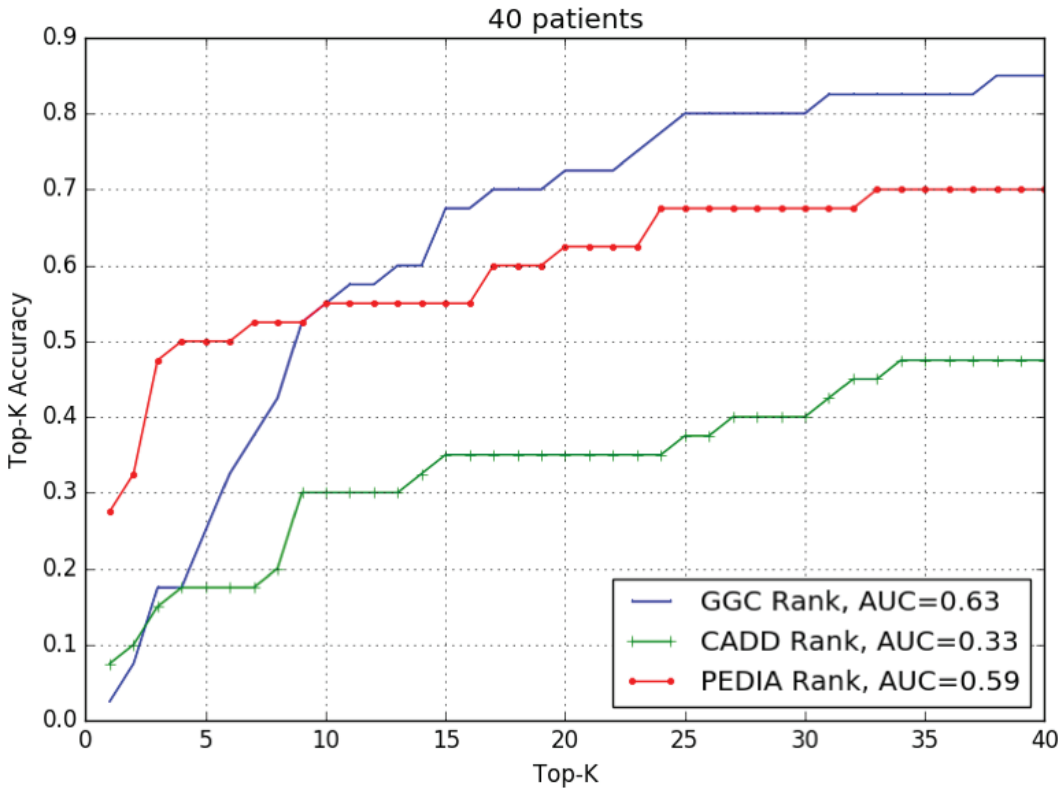


Figure 3. ROC curves showing ranking of causative genes based on GGC, CADD, and PEDIA ranks for 18 cases with diagnoses supported by F2G facial models.

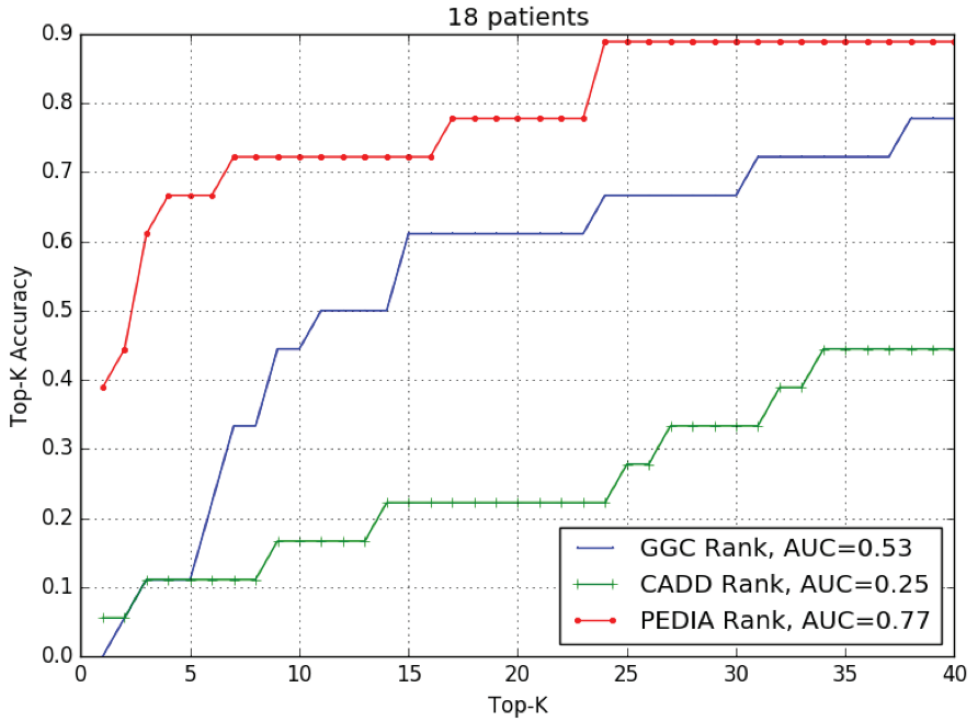


Table 1. Case ranking categories for 40 cases based on GGC, CADD, and PEDIA ranks.

	GGC rank	CADD rank	PEDIA rank
Ranked 1st	1 (2.5%)	3 (7.5%)	11 (27.5%)
Ranked ≤ 5th	10 (25%)	7 (17.5%)	20 (50%)
Ranked ≤ 10th	22 (55%)	12 (30%)	22 (55%)
Ranked ≤ 25th	32 (80%)	15 (37.5%)	27 (67.5%)

Table 2. Case ranking categories for 18 cases with diagnoses supported by F2G facial models.

	GGC rank	CADD rank	PEDIA rank
Ranked 1st	0 (0%)	1 (5.5%)	7 (39%)
Ranked ≤ 5th	2 (11%)	2 (11%)	12 (67%)
Ranked ≤ 10th	8 (44%)	3 (16.7%)	13 (72%)
Ranked ≤ 25th	12 (67%)	5 (27.8%)	16 (89%)

Of the 40 "solved" cases for which GGC rank was available, PEDIA rank showed an improvement of ≥ 1 position compared to the GGC rank in 19/40 cases (47.5%), and rank improved by ≥ 15 positions in 7/40 cases (17.5%). Among these 19/40 cases with improved PEDIA rank, the average improvement was 24.2 positions.

Of the total set of 57 "solved" cases, PEDIA rank versus CADD rank showed an improvement of ≥ 1 position in 32/57 cases (56%) and rank improved by ≥ 15 positions in 19/57 cases (33%). Among these 32/57 cases with superior PEDIA rank, the average improvement was 56.8 positions.

PEDIA Ranking for Cases Originally Requiring Analysis of Parental Data

For cases originally solved only after analysis of parental samples (N=15), PEDIA successfully placed the causative gene in the top 5 in 5/15 cases (33%) without parental data. Compared to the GGC rank (available for 12 of these cases), PEDIA rank showed improvement in 3/12 (25%), with an average improvement of 27.7 positions.

Discussion

These preliminary results clearly demonstrate the value of incorporating NGP in the variant analysis and prioritization process. The use of NGP dramatically increases diagnostic efficiency, as demonstrated by the increase in number of top-ranked causative genes in both "easily solved" and "challenging" cases, and the overall improvement in ranking across cases.

Even for cases with relatively fast and easily identified candidate variants (as represented by the "easily solved" cohort), the ability to improve ranking slightly can have a significant impact on the efficiency of the variant analysis process. In fact, the implementation of incremental time savings will provide substantial additive value, as this cohort represents 25% of all WES cases. Additionally, this demonstrates NGP's strong ability to facilitate clinical correlation by quantifying the relationship between candidate variants and the individual patient's phenotypic presentation.

The potential value-add from NGP is even more striking for challenging cases requiring more time and effort during variant analysis. Using NGP, variant analysis teams may dramatically decrease the time required to review each case, arriving at the causative variant more quickly.

Finally, in cases resolved through analysis of parental samples, NGP may enable prioritization of causative variants independent of parental data. In this manner, laboratories may decrease costs by reducing the need to complete parental WES and increase diagnostic yield, by enabling diagnosis of cases for which parental samples are unavailable.

The impact of NGP, resulting in improved ranking of causative genes, is much greater among cases diagnosed with a syndrome that is supported by a facial model in Face2Gene. This demonstrates the strong impact and value-add gained from the use of facial analysis. The DeepGestalt facial analysis

technology currently supports more than 300 specific genetic syndromes and syndrome groups, representing 45% of cases solved by WES. Consistent with this, in this cohort of diagnosed cases from GGC, 45% had syndromes supported by DeepGestalt. However, DeepGestalt is just one form of NGP, and as additional forms of NGP analysis are developed, such as analysis of brain MRIs, skeletal radiographs, and fundoscopy images, the proportion of patients supported by these NGP technologies will increase dramatically.

Conclusions

Integration of NGP into the variant prioritization process for a representative case cohort demonstrates significantly improved ranking of disease-causing genes and variants.

- PEDIA ranked the disease-causing variant first in 27.5% of cases, compared to 2.5% based on GGC rank and 7.5% based on CADD rank.
- Compared to GGC rank and CADD rank, PEDIA resulted in a significantly improved rank (≥ 15 positions) in 17.5% and 33% of cases, respectively.
- Among cases which are diagnosed with syndromes supported by DeepGestalt facial analysis, PEDIA ranking is even more impressive, showing 67% ranked in the top 5, compared to 11% based on GGC and CADD ranks.
- Even for cases easily solved by WES, incorporation of NGP increases diagnostic efficiency and facilitates prompt and accurate clinical correlation and reporting.
- Cases originally requiring parental samples for diagnosis benefited from NGP, potentially allowing diagnosis for such cases, even in the absence of parental samples.

NGP will dramatically increase efficiency in the molecular diagnostic process.

Appendices

Appendix 1: Example cases showing improved rank with PEDIA

Cohort	Causative Gene	Diagnosis	GGC rank	CADD rank	PEDIA rank
Challenging without parental samples	*SMARCB1*	Coffin-Siris syndrome	38	14	1
Easy	*ACTB*	Baraitser-Winter syndrome 1	11	47	3
Challenging without parental samples	*SMARCA2*	Nicolaides Baraitser	44	34	1
Easy	*KIAA0586*	Joubert syndrome 23	77	43	3

References

Basel-Vanagaite, L. Wolf L, Orin M, et al. (2016). Recognition of the Cornelia de Lange syndrome phenotype with facial dysmorphology novel analysis. Clin. Genet. 89,557–563.

Bauer, S., Köhler, S., Schulz, M. H. & Robinson, P. N. Bayesian ontology querying for accurate and noise-tolerant semantic searches. Bioinformatics 28, 2502–2508 (2012).

Gripp KW, Baker L, Telegrafi A, Monaghan KG.(2016) The role of objective facial analysis using DNA in making diagnoses following whole exome analysis. Report of two patients with mutations in the BAF complex genes. Am J Med Genet A. 2016 Apr 26.

Gurovich Y, Hanani Y, Bar O, Nadav G, Fleischer N, Gelbman D, Basel-Salmon L, Krawitz PM, Kamphausen SB, Zenker M, Bird LM, Gripp KW. (2019) Identifying rare genetic syndromes using deep learning. Nature Medicine 25, pages 60–64 DOI:10.1038/s41591-018-0279-0

Hennekam, R. & Biesecker, L. G. Next-generation sequencing demands next-generation phenotyping. Hum. Mutat. 33, 884–886 (2012).

Hsieh TC, Mensah MA, Pantel JT, PEDIA consortium, Krawitz P (2019) PEDIA: Prioritization of Exome Data by Image Analysis. Genet in Med. http://doi.org/10.1038/s41436-019-0566-2

Köhler, S. et al. Clinical diagnostics in human genetics with semantic similarity searches in ontologies. Am. J. Hum. Genet. 85, 457–464 (2009).

Liehr T, Acquarola N, Pyle K, St-Pierre S, Rinholm M, Bar O, Wilhelm K, Schreyer I. (2017) Nextgeneration phenotyping in Emanuel and Pallister Killian Syndrome using computer-aided facial dysmorphology analysis of 2D photos. Clin Genet. 2017 Jun 29. doi: 10.1111/cge.13087.

Lionel AC, Costain G, Monfared N, et al. Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. Genet Med. 2018;20(4):435–443. doi:10.1038/gim.2017.119

Mishima H, Suzuki H, Doi M, Miyazaki M, Watanabe S, Matsumoto T, Morifuji K, Moriuchi H, Yoshiura K, Kondoh T, Kosaki K (2019) Evaluation of Face2Gene using facial images of patients with congenital dysmorphic syndromes recruited in Japan. J Human Genet. 2019 May 29. doi: 10.1038/s10038-019-0619-z.

Pantel, J.T., Zhao, M., Mensah, M.A. et al. Advances in computer-assisted syndrome recognition by theexample of inborn errors of metabolism. J Inherit Metab Dis (2018) 41: 533. https://doi.org/10.1007/s10545-018-0174-3

Smith HS, Swint JM, Lalani SR, Yamal JM, de Oliveira Otto MC, Castellanos S, Taylor A, Lee BH, Russell HV. Clinical Application of Genome and Exome Sequencing as a Diagnostic Tool for Pediatric Patients: a Scoping Review of the Literature. Genet Med. 2019 Jan;21(1):3-16. doi: 10.1038/s41436-018-0024-6. Epub 2018 May 14.

Vorravanpreecha, N, Lertboonnum, T, Rodjanadit, R, Sriplienchan, P, Rojnueangnit, K. Studying Down syndrome recognition probabilities in Thai children with de-identified computer-aided facial analysis. Am J Med Genet Part A. 2018; 176A: 1935– 1940. https://doi.org/10.1002/ajmg.a.40483.

Wright CF, McRae JF, Clayton S, et al. (2018) Making new genetic diagnoses with old data: iterative reanalysis and reporting from genome-wide data in 1,133 families with developmental disorders. Genet Med http://dx.doi.org/10.1038/gim.2017.246.