Abstract No. 775

Gene Expression Profiling of Children with Growth Hormone Deficicency Prior to Treatment with Recombinant Human Growth Hormone is Associated with Growth Response Over 5 Years of Therapy

Adam Stevens[1], Philip Murray[1], Ekaterina Koledova[2], Pierre Chatelain[3], Peter Clayton[1]

[1]University of Manchester and Royal Manchester Children's Hospital, Manchester, UK; [2]Merck KGaA, Darmstadt, Germany; [3]Université Claude Bernard, Lyon, France

INTRODUCTION

- The relationship between pre-treatment gene expression and long-term growth response in growth hormone deficiency (GHD) is unknown.
- Prediction of long-term responses to recombinant human growth hormone (r-hGH) therapy would enable better decision-making about the start and maintenance doses and, hence, improve the cost-benefit ratio of r-hGH therapy.

OBJECTIVES

- To investigate the relationship between baseline gene expression and response to r-hGH over 5 years of therapy in children with GHD.

METHODS

- Patient population
  - Pre-pubertal children with GHD (N=50) were enrolled from the PREDICT (NCT00256126) and PREDICT long-term follow-up (NCT00699855) studies.
- Treatment
  - Children started with a 35 µg/kg/day r-hGH dose (all same brand, Saizen®) for the first month.
  - During the long-term follow-up period, patients could use any available r-hGH at a dose recommended by the physician.
- Genomic analysis
  - Baseline whole-blood gene expression was determined from peripheral blood mononuclear cells using Affymetrix U133 v2.0 microarray and Gene Expression Barcode 3.0.[1]
  - Gene expression data were normalised for Tanner stage.
  - Analysis of network modules was performed using Moduland algorithm.[2]
- Auxological analyses
  - Height velocity (cm/year) on r-hGH over 5 years was used as the marker for growth response.
  - Two groups of patients were defined according to growth response over 5 years of treatment.
    - Always above the median (G1, n=9).
    - Always below the median (G2, n=10).
  - The effect of age, gender and distance to target height were also assessed.
- Statistical analyses
  - A Random Forest algorithm was tested for prediction of growth response (based on normalised gene expression data, age and sex).
  - Predictive capacity was assessed using Area Under the Receiver Operating Characteristic Curve (AUC-ROC).
  - For prediction of growth response, as the data were unbalanced, a synthetic minority over-sampling technique (SMOTE[3]) was used to rebalance the dataset prior to Random Forest prediction.
  - The robustness of the gene expression markers was assessed using a one-way permutation test (1000 permutations) in R 3.3.1.

RESULTS

Patient Characteristics and Height Velocity

- The patient characteristics and the height velocity for the complete PREDICT LTFU cohort (n=125) are shown in **Tables 1 and 2**.
- There was no difference in age, gender and distance to target height between the G1 and G2 height velocity groups (data not shown).

Table 1. Patient Characteristics of Whole PREDICT LTFU Cohort

Characteristics	GHD (N=125)
Gender (male:female)	78 (62.4%):47 (37.6%)
Age at baseline (years)	9.6 (6.3, 11.2)
Baseline height SDS	–2.2 (–2.7, –1.7)
Baseline weight SDS	–1.4 (–2.1, –0.8)
Baseline BMI SDS	–0.3 (–1.0, 0.5)
Bone age (years)	7.0 (3.5, 9.5)
Basal height velocity (cm/year)	4.0 (3.0, 6.0)
Mid-parental height SDS	–0.8 (–1.7, –0.1)
GH peak response (µg/L)	4.1 (2.4, 5.6)
1-year r-hGH height velocity (cm/year)	8.4 (7.1, 10.1)

Data are n (%) or median (Q1, Q3).
BMI, body mass index; **GH**, growth hormone; **SDS**, standard deviation score.

Table 2. Height Velocity (cm/year) Throughout Treatment

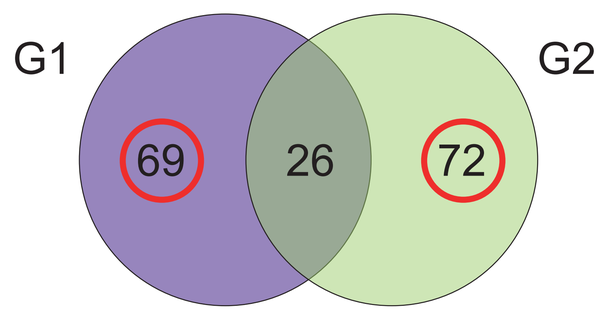
	GHD (N=125)		
Height velocity at year of treatment	**Mean (±SD)**	**Median (min, max)**	**n**
Year 1	8.9 (±2.1)	8.7 (4.7, 14.3)	75
Year 2	7.4 (±1.6)	7.1 (3.4, 12.2)	68
Year 3	6.7 (±2.0)	6.5 (2.0, 11.4)	68
Year 4	6.1 (±2.3)	6.2 (0.9, 11.6)	62
Year 5	5.1 (±2.3)	5.2 (0.0,10.8)	55

Data are mean (±SD) or median (min, max).

Genomic analyses

- 69 uniquely expressed genes ($p<1x10^{-5}$) were identified in the patients in group G1 (**Figure 1**).
- 72 uniquely expressed genes ($p<1x10^{-5}$) were identified in the patients in G2 (**Figure 1**).

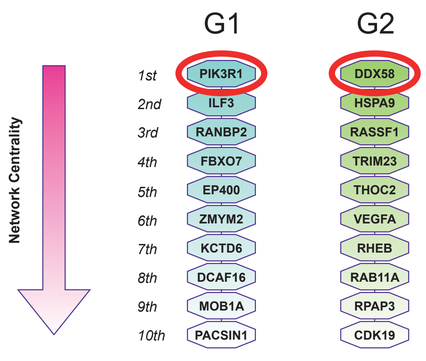
Figure 1. Overlap of Gene Expression in Treatment Groups G1 and G2



G1 (n=9): patients always above the median throughout the study. G2 (n=10): patients always below the median throughout the study. Intersection signifies genes common to both groups.

- Network models prioritised 94 of these 141 genes. The hierarchical results (network centrality) for the top ten genes associated with G1 and G2 are shown in **Figure 2**.

Figure 2. Hierarchy of Gene Modules Identified by Network Analysis



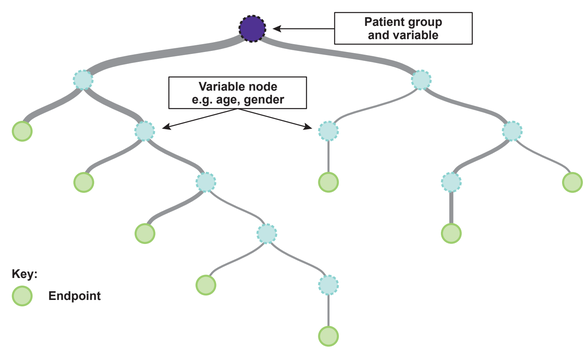
Hierarchy of network modules shown in order of centrality score. Each hexagon represents a module of interacting genes, with the most central gene named.

- Phosphatidylinositol 3-kinase regulatory subunit alpha (PIK3R1) expression was related to consistently good height velocity (G1) over 5 years ($p=1.2x10^{-9}$).
  - PIK3R1 encodes a protein involved in signalling that is important for cell growth, division and movement, and hormonal regulation.[4]
- DEXD/H-box helicase 58 (DDX58) expression was related to consistently poor height velocity (G2) over 5 years ($p=2.2x10^{-10}$).
  - DDX58 encodes a putative RNA helicase; these enzymes are implicated in RNA binding and alteration of RNA secondary structure.[5]

Random Forest Analysis

- Random Forest analysis (**Figure 3**) of baseline gene expression consistently predicted growth response above and below the median over 5 years in the genes selected by network analysis.

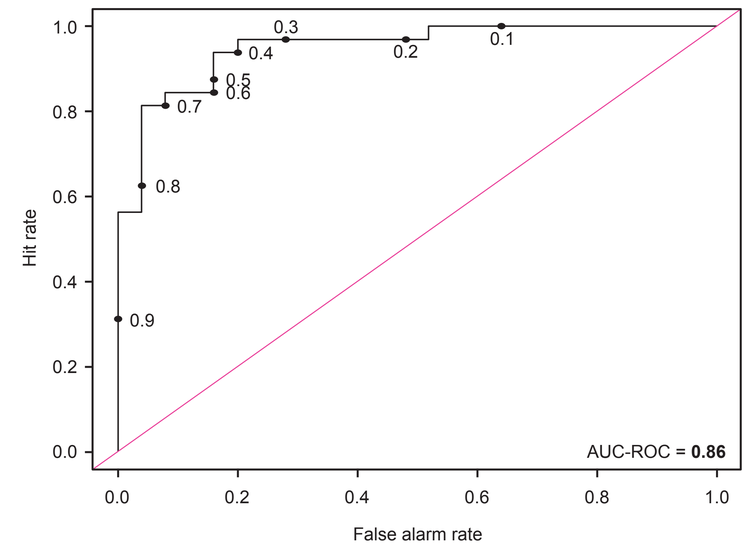
Figure 3. Random Forest Classification Schematic Decision Tree



Random Forest Classification is a machine-learning method based on use of an ensemble of decision trees (i.e., a forest); a schematic representation of the available alternatives and their possible consequences are useful for sequential decision-making analyses.

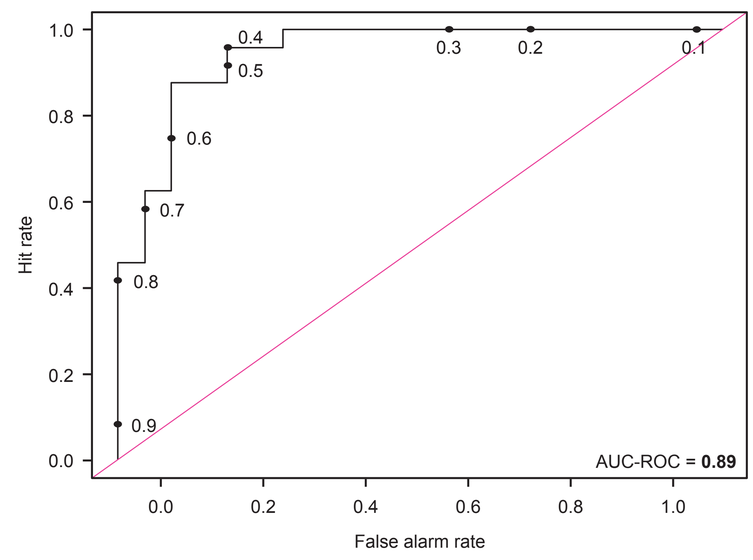
- G1 versus G2: SMOTE AUC-ROC was 0.86 (95% confidence interval 0.76–0.95) (**Figure 4**).

Figure 4. Predictive Value of Gene Expression for Good Growth Response



- G2 versus G1: SMOTE AUC-ROC was 0.89 (95% confidence interval 0.80–0.98) (**Figure 5**).

Figure 5. Predictive Value of Gene Expression for Poor Growth Response



CONCLUSIONS

- **We have identified genes uniquely expressed before treatment in 50 pre-pubertal patients with GHD that are associated with quality of growth response (responsiveness) over 5 years of therapy**
- **Responsiveness to r-hGH therapy seems to be genetically controlled in GHD, which may have implications for personalised therapy**
- **These gene expression markers may be used prior to r-hGH treatment to identify which patients will be good or poor long-term responders**
- **Further assessment is required to validate the predictive value and determine the functional significance of the gene subsets we have identified**

REFERENCES

1. McCall, et al. *Nucleic Acids Research* 2014;42:D398-D946.
2. Szalay-Beko M, et al. *Bioinformatics* 2012;28:2202-04.
3. Chawla N, et al. *J Artificial Intell Res* 2002; 16:321–57.
4. Hoyle J, et al. *Genomics* 1994;24:400-02.
5. Morosky SA, et al. *J Biol Chem* 2011;286:28574-83.

ACKNOWLEDGMENTS

The authors would like to thank the patients who participated in the PREDICT trial and their parents and guardians. This trial was sponsored by Merck KGaA, Darmstadt, Germany. Medical writing support was provided by Steven Goodrick of inScience Communications, London, UK, sponsored by Merck KGaA, Darmstadt, Germany.

DISCLOSURES

AS, PM, PCh, PCl have all received honoraria and/or research support from Merck KGaA, Darmstadt, Germany. **EK** is an employee of Merck KGaA, Darmstadt, Germany. This study was sponsored by Merck KGaA, Darmstadt, Germany.

GET POSTER PDF

Presented at ESPE 2016; September 10-12; Paris, France

55ESPE 14--5-RFC

Growth: Mechanisms

Adam Stevens

DOI: 10.3252/pso.eu.55ESPE.2016

Poster presented at: ESPE 2016 PARIS