Enabling routine β-Thalassemia Prevention and Patient Management by Scalable, Combined Thalassemia and Hemochromatosis mutation analysis

Ghazala Hashmi(Former Corresponding Author)
Biomolecular analytics
Asim Qidwai
Afzal Memorial Foundation, Karachi
Kristopher Fernandes
Biomolecular Analytics
Michael Seul(New Corresponding Author) (michael@biomolecularanalytics.com)

Research article

Keywords: beta (β)-thalassemia, hemochromatosis, iron overload, inherited disorders, LeanSequencing, patient management, prevention

Posted Date: February 6th, 2020

DOI: https://doi.org/10.21203/rs.2.13323/v2

Version of Record: A version of this preprint was published on May 15th, 2020. See the published version at https://doi.org/10.1186/s12881-020-01017-x.

Abstract

Background : Beta (β)-thalassemia is one of the most common inherited disorders worldwide, with high prevalence in the Mediterranean, the Middle East and South Asia. Over the past 40 years, many countries have implemented awareness and prevention campaigns that have greatly reduced the incidence of affected child births. In contrast, much remains to be done in South-Asia. Thus, for Pakistan, current estimates of the number of children born with thalassemia are at ~7,000 per annum, with no sign of improvement. Although there is good agreement that intermarriage of carriers is a principal source of the high prevalence of the disorder, effective tools for screening and diagnosis on which to base prevention programs are not readily available. Methods : Here, we present results for a novel “LeanSequencing” process to identify 18 β-thalassemia and related sickle cell anemia, and simultaneously a set of 3 hemochromatosis mutations in a multi-ethnic population of 274 pediatric and adolescent patients treated at Afzaal Memorial Thalassemia Foundation hospital in Karachi, Pakistan. Results : We found substantial differences in the abundance of disease-causing mutations among the principal ethnic groups in our cohort. We also found the hemochromatosis mutation H63D C > G in 61 (or 22.1%) of our patients including 7 (or 2.6%) homozygotes. Conclusions : To our knowledge, this is the first screen combining β-thalassemia and hemochromatosis mutations in a single test, so as to facilitate the early identification of patients who may be at increased potential risk for complications from iron overload and thereby to improve the prospective management of thalassemia patients.

Background

Beta (β)-thalassemia is a genetic disorder that remains a public health challenge in many countries around the world, with a high prevalence especially in the Mediterranean, Middle East as well as South and South-East Asia. Its molecular origin lies in abnormalities of the beta (β)-globin gene which leads to the limited synthesis or completes lack of correctly folded hemoglobin [“HBB”; 1, 2, 3]. Individuals who inherit 2 abnormal copies of the beta-globin gene develop anemia and often require life-long red cell transfusion support to maintain adequate hemoglobin levels. The nature and severity of clinical symptoms are known to reflect the specific underlying mutation(s), and even individuals with just a single abnormal copy of the gene may be affected, though generally with less severe clinical manifestations [3, 4]. Given the autosomal recessive pattern of inheritance, pre-marital carrier screening can be effective in prevention [5], and mandatory programs have been introduced in several of the affected countries [6]. In addition, the early identification of the specific underlying pathogenic mutations by pre-natal and neo-natal screening, likewise implemented in many countries [7], represents a critical element of patient management.

Thalassemia remains prevalent in Pakistan where, according to current estimates, 5,000-7,000 children are born each year with thalassemia major [8]. Several organizations provide information to raise awareness and offer programs to diagnose the disorder [8, 9, 10]. However, while to date at least two provincial governments, Sindh and Punjab, have passed bills to support awareness campaigns and premarital carrier screening, a national program remains to be implemented. To a large extent, diagnosis

currently relies on electrophoresis which, while relatively inexpensive, does not provide the differential diagnosis available only by the reliable identification of the underlying mutations. In fact, many patients are misdiagnosed for anemia. At present, the requisite genetic testing is performed only sporadically, as it is perceived to be costly, and in any case, is currently available only at a few hospital laboratories [8, 10]. Thus, there is growing agreement that limited access to genetic testing and reliable disease diagnosis in fact is contributing to a gradual increase in the number of thalassemia patients in the country.

When diagnosed, clinical treatment, in addition to addressing acute clinical complications , includes the transfusion of red cells [11,12]. However, transfusion, often on chronic protocols, frequently leads to the formation of allo-antibodies as well as the accumulation of excess iron [13]. Individuals afflicted with hemochromatosis absorb excess iron which accumulates in tissues and organs and may lead to major organ damage, and patients with HFE ("high-Fe") gene mutations that cause hemochromatosis may be at increased risk [14]. At present, even the scarce genetic analysis performed on thalassemia patients in Pakistan and elsewhere generally is limited to mutations in the HBB gene, and does not routinely include mutations in the HFE gene.

Here, we describe a novel approach to simultaneously detect a set of 18 HBB gene mutations, as well as 3 HFE mutations in chronically transfused pediatric and adolescent thalassemia patients, including the determination of the abundance of these mutations in several ethnic groups in Pakistan. Our ultimate objective is that of providing a cost-effective screening process that holds the promise of improving thalassemia prevention and patient management.

Methods

Patient Cohort – Our analysis included samples from a total of 288 (mostly) pediatric and adolescent patients of diverse (self-identified) ethnic background, most abundant among them Urdu-speaking (North India), Sindhi, Saraiki, Punjabi, Pathan and Balochi, all treated at the Afzaal Memorial Thalassemia Foundation ("AMTF") hospital in Karachi. All patients were on chronic transfusion protocols.

The majority of these patients (n=275) had a diagnosis of β-thalassemia, established by standard clinical methods including blood work, hemoglobin electrophoresis and HPLC; in addition, 13 blinded samples from patients with diagnoses other than β-thalassemia also were included as "negative controls", namely: hereditary spherocytosis (5), immune thrombocytopenic purpura (2), hemolytic anemia (2), autoimmune hereditary anemia (1), Fanconi anemia (1), bone marrow failure syndrome (1) and alpha (α)-thalassemia (1). One sample failed to produce data, both by LeanSequencing and Sanger sequencing, and was excluded from further analysis. Of the remaining 274 patients with β-thalassemia, 132 were female, with a mean age of 8.7 years, and 142 were male, with a mean age of 8.2 years (Table 1); the age distribution for both is positively skewed (with a mode below the mean, near ~6 years).

Sample Collection & Processing – Duplicate barcoded buccal swab samples were collected, and crude extracts prepared using the Phusion Human Specimen Direct PCR kit (ThermoFisher Scientific, Waltham, MA) by a protocol that, following preparation of the lysate, requires only simple spinning (but no high-

speed centrifugation) and takes ~10 min to complete for a batch of 8 samples. These extracts were processed by a multiplexed PCR reaction comprising both HBB and HFE genes, followed by an allele-specific labeling reaction and analysis by capillary electrophoresis (the latter performed by send-out to a service provider, Genewiz, South Plainfield, NJ), in accordance with LeanSequencing^{TM} ("LSQ").

LSQ Protocol: Amplify & Discriminate – LeanSequencing is a novel process for analyzing sequence variants, developed at BioMolecularAnalytics (Warren, NJ). The two analytical steps in LeanSequencing are amplification and discrimination (Fig. 1). Amplification, in a single multiplex PCR reaction, produces a set of amplicons comprising all sequence variants of interest; each amplicon bears a molecular tag (aka "barcode") that identifies the sample of origin and permits amplicons from multiple samples to be combined ("pooled"). Following pooling, discrimination, in a second multiplex PCR reaction, produces labeled allele-specific amplicons, from, in this case, 4 samples per well, for analysis in a standard capillary sequencer, in this case the 96-channel ABI 3730xl.

The protocol omits extraneous steps including DNA purification, "normalization" and "clean-up" reactions.. The discrimination reaction is configurable so as to select specific marker ("SNP") sets according to ethnicity or geographic area of interest, and to accommodate individual or multiple (2 or 4) "pooled" samples per well. The entire process takes less than 1h of hands-on time per 96-well plate, and it is readily automated using inexpensive laboratory pipetting instrumentation. The process achieves a very high data rate – for example, in the "pools of 4" configuration used here, a single run on a standard sequencer with a 96-capillary array produces complete molecular HBB and HFE profiles for 384 samples.

Selection of HBB and HFE Variants - To select mutations of interest, we started with the set most commonly observed in Pakistan and Middle Eastern countries [15], namely: –del 619; IVS I-1 G > T or G > A; IVS I-5 G > C; IVS I-6 T > C, IVS I-110 G > A, –88 C > T, –29 A > G, cd 8/9 +G, cd 41/42 –TTCT, as well as 3 structural mutations: cd 6 A > T (the sickle cell anemia mutation, "HbS"), cd 6 A > G ("HbC") and cd 26 G > A ("HbE"). Following initial testing of this design, we selected 64 samples with at most one of these most commonly observed mutations for Sanger sequencing of HBB gene exon 1, partial intron 1 and exon 2 to check for any additional mutations or variants,and thereupon expanded our initial set to include these five mutations: cd 5 –CT, cd 15 G > A, cd 16 –C, cd 30 G > C ("Monroe") ; and the rare –90 C > T mutation (aka rs34999973 C > T [16]) as well as these two variants: rs713040 c.9 T > C and rs35799536 G > C.

The final selection for our LSQ application comprises 18 HBB and 2 HFE mutations; as initial testing showed all patients to be normal for S65C, this was omitted (Table 2). This selection covers many of the mutations commonly observed in other regions, namely (with reference to Table 1 in reference [2], Figure. 3 in reference [3]): Mediterranean (cd 5 –CT, IVS I–1 G > A, IVS I–6 T > C, IVS I–110 G > A); Central and SE Asian (cd 41/42 –TTCT); East Asian (IVS I–5 G > C); African (–29 A > G, –88 C > T); and Indian (–del 619) and (with reference to Table 1 in [17]): Middle Eastern (IVS I–5 G > C, IVS I-1 G > A, IVS I-6 T > C and cd 5 –CT).

Statistical Analysis – Allele frequencies were determined by “gene counting” from genotypes. All analysis was performed, and data tables and figures were generated, using Microsoft Excel.

Results

Overall Abundance of Mutations – Table.3 lists estimated variant allele frequencies for the subset of HBB and HFE mutations encountered in our cohort. The three most common β-thalassemia mutations overall, in order of decreasing variant allele frequency were: IVS I-5 G > C (42.0%); cd 8/9 +G (28.5%) and cd 41/42 −TTCT (6.8%). Less commonly, we also observed structural mutations , namely cd 26 G > A (HbE) (0.7%, corresponding to 4 heterozygotes) and cd 6 A > T (HbS) (1 heterozygote); we determined one patient to be homozygous for the rare −90 (aka rs34999973) C > T mutation, without any other mutation or variant, and we found one additional patient who was compound heterozygous for rs35133315 del T (a variant detected by Sanger sequencing, but not included in the LSQ design) and IVS I-5 G > C. None of the individuals in our cohort was a carrier for any of four additional mutations (IVS I-6 T > C, IVS I-110 G > A, −88 C > T and −29 A > G, or the HbC mutation,) , though these have been observed in other ethnic groups outside of Pakistan.

Variants of Unknown Significance - The two variants, rs713040 c.9 T > C and rs35799536 G > T, were quite prevalent in β-thalassemia patients as well as in individuals with diagnoses other than β-thalassemia. Thus, 10 of the former had one or both of these variants, in combination with a pathogenic mutation. In contrast, 10 of the latter had one or more of the variants, but none of the mutations.

Overall Mutation Status for Patients and Negative Controls - The configuration of mutations and that of the genotypes observed in our patient cohort are summarized in Table 4. As expected, the majority of the patients in our cohort were homozygous (n=213 of 274 β-thalassemia patients, or 77.7%) or compound heterozygous (n=40 or 14.6%) for β-thalassemia (including the HBB variants HbS, C and E) mutations; an additional 21 individuals (or 7.6%) had 1 of the mutations, namely 11 in isolation, and 10 in combination with one of the non-pathogenic variants (Tb. 4A). Thus, all of the patients had at least 1 of the mutations included in our design. After un-blinding, we confirmed that 10 of the 13 “negative controls” had at least one copy of one or both of the variants rs713040 c.9 T > C and rs35799536, but none of the pathogenic mutations, while 3 had neither.

The classification of observed genotypes in Tb. 4B shows that 246 of our patients, or almost 90%, were homozygous or compound heterozygous for β0 mutations. An additional 26 had one of these mutations, most in combination with SNP’s or structural variants: of these, one patient was compound heterozygous for IVS I-5 G>C and HbS. Of the two patients with “Other” configurations, one was homozygous for the β+ mutation −90 C > T, and one heterozygous for cd 26 G > A (“HbE”) and homozygous for both SNPs.

Predominance of Different Disease-Causing Mutations In Principal Ethnicities - As evident from Figs. 2A and 2B, the prevalence of the predominant disease-causing mutations varies substantially across the principal ethnic groups in our cohort which included: Pathan (74), Urdu-speaking (45), Saraiki (42), Sindhi (30), Balochi (24) and Punjabi (15).

Thus, a remarkable 19 out of the 24 Balochi patients, or almost 80%, were homozygous for IVS I-5 G > C which, in combination with cd 8/9+G, accounted for 90% of the mutations observed in this group; 32 of the 74 Pathan patients, or 42%, were homozygous for cd 8/9+G, and an additional 12 were heterozygous for that mutation; and a significantly higher proportion of the Sindhi patients compared to that in the other groups, namely 2 of 30, were homozygous for IVS I-1 G > T, and an additional 2 were heterozygous.

Hemochromatosis - Of the patients in our cohort, 61 (or 22.1%) had at least 1 copy of the H63D C > G mutation which is strongly associated with this disorder: specifically, 6 (or 2.2%) were homozygous, and 55 (or ~20.1%) heterozygous for that mutation; all patients were normal for C282Y.. The frequency of the H63D variant allele exceeded 10% in all but one of the ethnic groups most abundantly represented in our sample, ranging from 9.6% (Pathan) to 16.7% (Punjabi and Urdu), as summarized in Fig. 2. To our knowledge, the high prevalence of this mutation in a Pakistani population of β-thalassemia patients has not been previously reported.

Discussion

Ours appears to be the first study of mutation prevalence focusing exclusively on patients, and it thus conveys information about the spectrum and prevalence of pathogenic mutations in different ethnic groups. A previous related study also included putative carriers, that is: individuals having an affected family member, along with chorionic villus samples from referrals by thalassemia clinics [18].

As expected, the majority of our patients, namely 253 (or 92.3%), were either homozygous or compound heterozygous for the most prevalent mutations, the majority of which were of type β(0). However, identity and abundance of the predominant mutation(s) differ markedly among the six principal ethnic groups represented in our cohort.

Our “LeanSequencing” analysis indicates that our design, comprising 18 HBB (“hemoglobin”) mutations, as well as 2 HFE (“hemochromatosis”) mutations, provides excellent coverage: all of the patients in our multi-ethnic cohort of 274 (predominantly) pediatric and adolescent Pakistani β-thalassemia patients in fact had at least 1 of the selected mutations (Table. 1). Our design also achieved reliable differential analysis: of the 13 individuals with diagnoses other than β-thalassemia, none had any of the mutations, but only common variants, while 3 (namely 2 diagnosed with hereditary spherocytosis and one diagnosed with ITP) had neither variants nor mutations.

Regarding coverage in other affected countries and regions, our design also contains mutations that, while rare or absent in our cohort, are common in other populations including those listed in the Methods section. As LSQ is readily configurable, at the variant discrimination stage, it may be extended by “reading-out” additional mutations such as cd 39 C > T (abundant in Spain, Sardinia, Middle East and North-Africa).

We identified 6 H63D homozygotes in our patient cohort, corresponding to an estimated prevalence of 2.2%, and a variant allele frequency exceeding 10% in all but one the principal ethnic groups represented

in our study, in contrast, for example, to a previous report concerning a North Indian population [19, 20].

In (otherwise) healthy individuals, H63D homozygotes have been reported to display increased transferrin saturation and serum ferritin, but decreased unsaturated iron binding capacity compared to H63D heterozygotes or normal individuals [21]. Our observation therefore naturally raises questions as to a potential adverse effect of hemochromatosis in β-thalassemia patients, namely to increase the risk of iron overload, and while hemochromatosis may confer elevated risk to all patients, the situation may be exacerbated by chronic transfusion. In fact, a recent report from Egypt does point to a correlation between homozygosity for that mutation and elevated levels of several clinical indicators of iron overload, including serum iron and ferritin[14]; others have found no correlation [22, see also Kaur et al, ref. 19].

A preliminary review of serum ferritin levels, for a subset of 124 of the patients in our cohort, showed high levels in all by the time first readings were taken, suggesting that, regardless of a possible genetic predisposition to iron overload as a result of the H63D mutation, ferritin levels may reflect the cumulative adverse effect of a high number of transfusions, even during the first year of transfusion support. However, by including the most relevant HFE mutations in our β-thalassemia design, we will be able to spot patients with this potential additional risk factor, enabling a prospective study of new β-thalassemia patients with and without HFE mutations as the basis for a systematic evaluation of any statistically significant differences in the rise of ferritin levels and related parameters between patients with β-thalassemia and those with other diagnoses, and, among the former, those receiving and those not receiving periodic transfusions. HFE mutation status would be used as an important factor in early risk assessment, to be taken into account when tailoring the transfusion protocol and the chelation protocol to individual patients so as to minimize the risk of iron overload.

Conclusion

Programs for pre-marital, prenatal and neonatal screening for β-thalassemia which have been shown to yield substantial benefits [7] also hold promise for Pakistan, perhaps with a focus on extended families of patients [9]. A recently published study reports a tendency for couples to abort when disease-causing mutations are detected in the foetus [8]. In addition, neonatal screening would provide the basis for interceding as early as possible, ideally by devising personalized transfusion protocols in such a manner as minimize the systemic adverse consequences of iron accumulation.

In view of the results reported here, we believe that LeanSequencing offers an effective new approach to faithfully providing combined β-thalassemia and hemochromatosis screening and disease diagnosis to the larger population. The SKH and AMT foundations are planning to initiate this service in Pakistan and elsewhere.

Abbreviations

LeanSequencingTM (“LSQ); Hemochromatosis (HFE); Hemoglobin (“HBB”); Afzaal Memorial Thalassemia Foundation (“AMTF”); Sanya Kiran Hashmi Foundation (SKH);

Declarations

Ethics approval and consent to participate

Informed consent was obtained from all individual participants and/or their guardians included in the study.

Consent for publication

All authors read and approved the study for submission/publication

Availability of data and materials

Datasets generated and analyzed for the current study are available from the corresponding author on reasonable request.

Competing interests

Authors GH, KF and MS are employees of BioMolecular Analytics, LLC. AQ is an employee of Afzaal Memorial Thalassemia Foundation Hospital and is one of the treating physicians for the patients.

Funding

NA

Author contribution

GH & MS: Designed the study, analyzed and managed data, wrote the manuscript

AQ: Collected samples and relevant clinical information.

KF: performed LeanSequencing protocol and handled data collection.

Acknowledgement

We acknowledge patients and their guardians included in the study for agreeing to participate, and we also thank the staff of Afzaal Thalassemia Memorial Foundation for collecting and samples.

References

1. Rachmilewitz EA, Giardina PJ. How I treat thalassemia. *Blood*.2011; 118: 3479-3488.

2. Galanello R, Origa R. Beta thalassemia. *Orphanet J Rare Dis* .2010; 5:11.
3. Cao A, Galanello R. Beta-Thalassemia. *Genetics in Medicine*. 2010; 12: 62-76.
4. Thein SL. The Molecular Basis of β-Thalassemia. *Cold Spring Harb Perspect Med*. 2013;4: a011700. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3633182/pdf/cshperspectmed-HMG-a011700.pdf
5. Cao A, Kan YW. The Prevention of Thalassemia. *Cold Spring Harb Perspect*. 2013; 2: a011775.
6. Cousens NE, Gaff CL, Metcalfe SA, Delatycki MB. Carrier screening for Beta Thalassemia: a review of international practice. *Eur J Human Gen.* 2010; 18:1077-1083.
7. Saffi M, Howard N. Exploring the Effectiveness of Mandatory Premarital Screening and Genetic Counselling Programmes for β-Thalassemia in the Middle East: A Scoping Review. *Public Health Genomics*. 2015;18:193-203.
8. Baig SM, Sabih D, Rahim MK, Azhar A, Tariq M, Hussain MS, Naqvi SMS, Raja GK, Khan TN, Jameel M, Iram Z, Noor S, Baig UR, Qureshi JA, Baig SA, Bakhtiar SM. (2012) -Thalassemia in Pakistan: A Pilot Program on Prenatal Diagnosis in Multan. *J Pediatr Hematol Oncol*. 2012; 34: 90-92.
9. Ahmed S, Saleem M, Modell B, Petrou M. Screening extended families for genetic hemoglobin disorders in Pakistan. *N Engl J Med.* 2002; 347:1162-1168.
10. Maheen H, Malik F, Siddique B, Qidwai A. Assessing Parental Knowledge About Thalassemia in A Thalassemia Center of Karachi, *Pak J Genet Counsel*. 2015; 24:945:951.
11. Olivieri NF, Brittenham GM. Iron-chelating therapy and the treatment of thalassemia. *Blood*. 1997; 89:739-761.
12. Weatherall DJ, Clegg JB. The Thalassemia Syndromes, *Fourth ed., Blackwell Sci., Oxford*.2001; p. 846.
13. Vichinsky E, Neumayr L, Trimble S et al. Transfusion complications in thalassemia patients: a report from the Centers for Disease Control and Prevention. *Transfusion*. 2014; 54:972-981.
14. Enein AA, El Dessouky NA, Mohamed KS et al. Frequency of Hereditary Hemochromatosis (HFE) Gene Mutations in Egyptian Beta Thalassemia Patients and Its Relation to Iron Overload. *Open Access Macedonian J of Med Sci.* 2016; 4:226-231.
15. Ahmed S, Petrou M, Saleem M. Molecular genetics of β-thalassemia in Pakistan: a basis for prenatal diagnosis. *Brit J Hematol*. 1996; 94: 476-482.
16. Faustino P, Lavinha J, Marini MG, Moi P. β-thalassemia mutation at -90 C > T impairs the interaction of the proximal CACCC box with both erythroid and non-erythroid factors. *Blood*. 1996; 88: 3248-3249.
17. Zahed L. The spectrum of β-thalassemia mutations in the Arab populations. *J Biomedicine and Biotechnology.* 2001; 3: 129-132.
18. Ansari SH, Shamsi TS, Asjraf M et al. Molecular epidemiology of β-thalassemia in Pakistan: Far reaching implications. Indian J Hum Genet. 2011;18: 193-197.
19. Kaur G, Rapthap CC, Xavier M, Saxena R, Choudhary VP, Reuben SK, Mehra NK. Distribution of C282Y and H63D mutations in the HFE gene in healthy Asian Indians and patients with thalassaemia major.

Natl Med J India. 2003; 16: 309-310.

20. Lok CY, Merryweather-Clarke AT, Viprakasit V, Chinthammitr Y, Srichairatanakool S et al. Iron overload in Asian community. *Blood*. 2009; 114: 20-25.
21. Jackson HA, Carter K, Darke C et al. HFE mutations, iron deficiency and overload in 10 500 blood donors. *Brit J of Haematology*. 2001; 114: 474:484.
22. Garewal G, Das R, Ahluwalla J, Marwaha RK. Prevalence of the H63D mutation of the HFE in north India: its presence does not cause iron overload in beta thalassemia trait. *Eur J Haematol*. 2005; 74: 333-336.

Tables

Table 1: Mutations and Variants Analyzed

NOTE: the HBB variants marked "SNP" and HFE:S65C were not included in the final LSQ selection

β0	β0	β0	β0	β0	β0	β0	β0	β0	β0	β+	β+	β+	β+	β+	HbC	HbS	HbE	SNP	SNP	HFE	HFE	HFE
–del 619	IVS I-1 G > T	IVS I-1 G > A	IVS I-5 G > C	cd 5 –CT	cd 8/9 +G	cd 15 G > A	cd 16 –C	cd 30 G > C ("Monroe")	cd 41/42 –TTCT	–90 C > T	–88 C > T	–29 A > G	IVS I-6 T > C	IVS I-110 G > A	cd 6 G > A	cd 6 A > T	cd 26 G > A	rs713040 c.9 T > C	rs35799536 G > C	H63D C > G	C282Y G > A	S65C A > T

Table 2: Self-declared ethnicities (left) and age distribution by gender (right) in our β-thalassemia patient cohort; NA refers to 9 individuals whose ethnicity was not recorded

Ethnicities in Patient Cohort		
	Number	Percentage
Pathan	74	27.0%
Urdu	45	16.4%
Saraiki	42	15.3%
Sindhi	30	10.9%
Baloch	24	8.8%
Punjabi	15	5.5%
Memon	11	4.0%
Hazarwi	9	3.3%
NA	9	3.3%
Burohi	4	1.5%
Bhatti	3	1.1%
Afreedi	2	0.7%
Bangali	2	0.7%
Hindko	2	0.7%
Mansahra	1	0.4%
Mianwali	1	0.4%

AgeDistribution		
Yrs	F	M
1	3	3
2	7	5
4	19	26
6	33	29
8	17	14
10	20	27
12	14	16
14	6	6
16	5	9
18	1	4
20	2	2
22	2	1
24	1	0
26	0	0
28	1	0
30	0	0
32	1	0

Table 3: Pan-ethnic Frequencies of HBB and HFE Mutations

and Variants Observed in our β-Thalassemia Patient Cohort

Type	Name	< fv >
β(+)	IVS I-5 G > C	**0.420**
β(0)	cd 8-9 +G	**0.285**
β(0)	cd 30 G > C ("Monroe")	**0.223**
β(0)	cd 15 G > A	**0.163**
β(0)	cd 16 delC	**0.141**
β(0)	cd 5 -CT	**0.109**
β(0)	cd 41-42 -TTCT	**0.068**
β(0)	IVS I-1 G > T or G > A	**0.033**
β(0)	del-619	**0.033**
β(+)	-90 C > T	**0.022**
HbE	cd 26 G > A	**0.007**
HbS	cd 6 A > T	**0.002**
SNP	rs713040 c.9 T > C	0.256
SNP	rs35799536 G > C	0.222
HFE	H63D C > G	**0.123**

Table 4A: Configuration of mutations in our β-thalassemia patient cohort

Zygosity	Count	Cum	Cum (%)
Homozygous	**213**	**213**	**77.7%**
Compound Het	**40**	**253**	**92.3%**
Heterozygous	**11**	**264**	**96.4%**
Compound Het Mutation & SNP	**10**	**274**	**100.0%**

Table 4B: Genotypes in our β-thalassemia patient cohort

Genotype	Count	Cum	Cum (%)
β0 / β0	**212**	**212**	**77.4%**
β0 / β0 (compound het)	**34**	**246**	**89.8%**
β0 / SNP	**11**	**257**	**93.8%**
β0 / β S,C,E	**4**	**261**	**95.3%**
β0 / β	**11**	**272**	**99.3%**
Other	**2**	**274**	**100.0%**

Figures

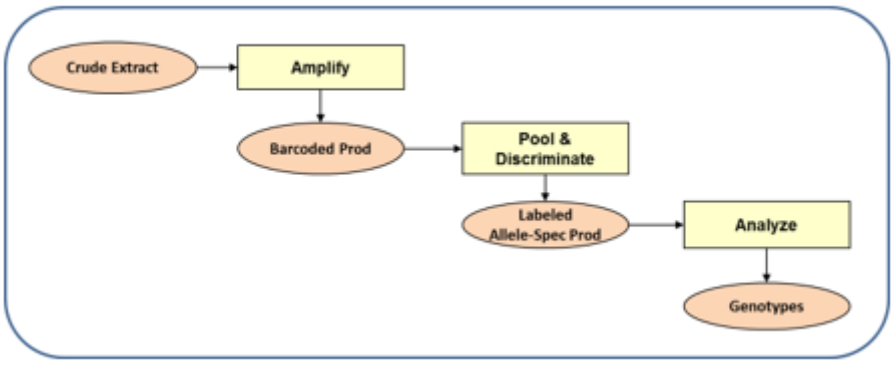

Figure 1

LeanSequencing™ Process Flow

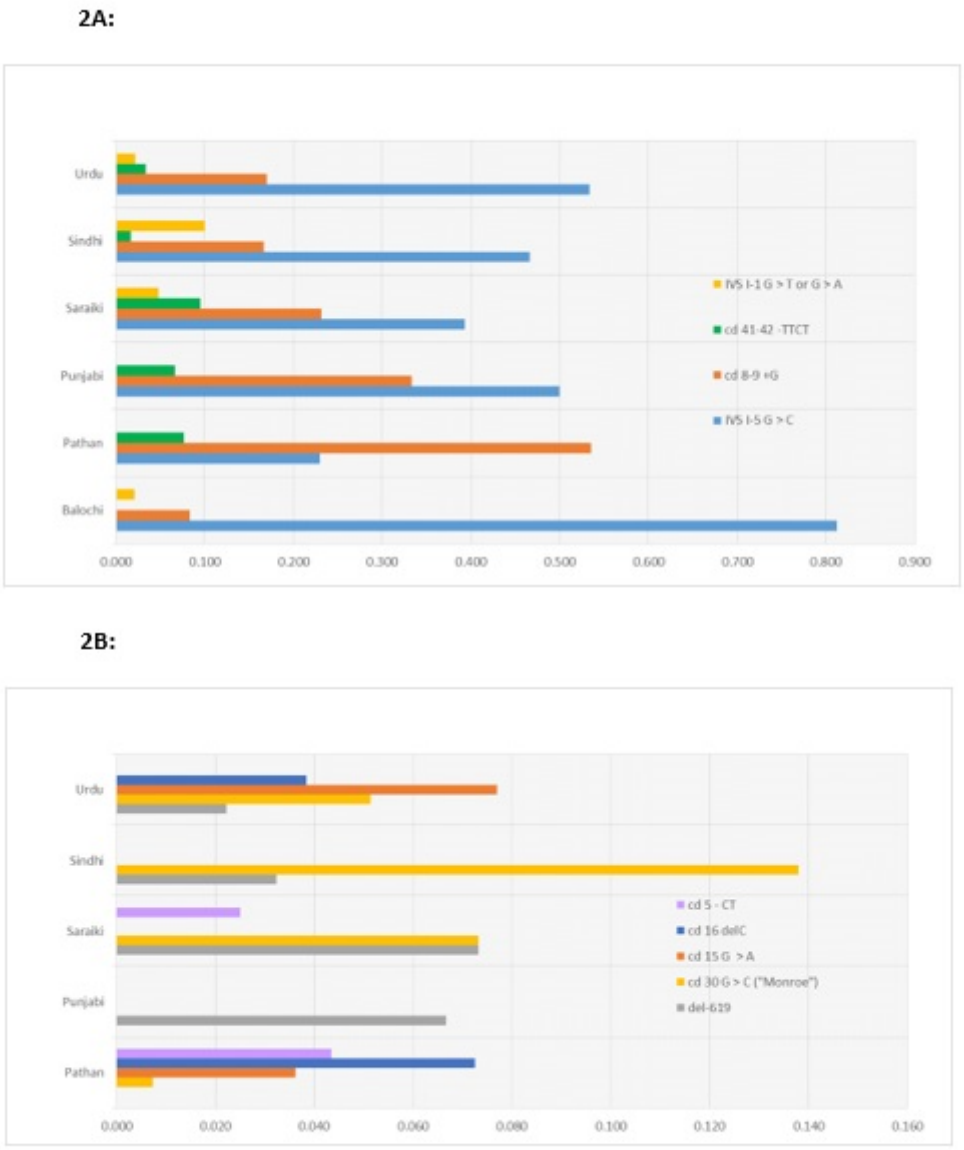

Figure 2

A: Frequencies for the 4 most prevalent disease-causing mutations, for the 6 most prevalent ethnic groups in our cohort. B: Frequencies of less prevalent disease-causing mutations, for the 6 most prevalent ethnic groups in our cohort: none of these mutations were observed in Balochi patients (not shown)

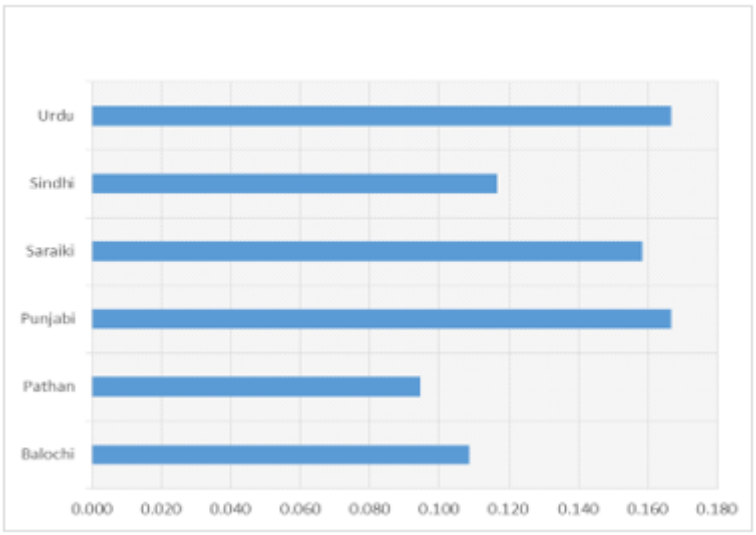

Figure 3

H63D variant allele frequency for principal ethnicities in our cohort