EBioMedicine xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Paper

Rapid Targeted Next-Generation Sequencing Platform for Molecular Screening and Clinical Genotyping in Subjects with Hemoglobinopathies

Xuan Shang [a,z,aa,1], Zhiyu Peng [b,1], Yuhua Ye [a,z,aa,1], Asan [s,t,1], Xinhua Zhang [c], Yan Chen [d], Baosheng Zhu [e], Wangwei Cai [f], Shaoke Chen [g], Ren Cai [h], Xiaoling Guo [i], Chonglin Zhang [j], Yuqiu Zhou [k], Shuodan Huang [l], Yanhui Liu [m], Biyan Chen [n], Shanhuo Yan [o], Yajun Chen [p], Hongmei Ding [q], Xiaolin Yin [c], Liusong Wu [d], Jing He [e], Dongai Huang [f], Sheng He [g], Tizhen Yan [h], Xin Fan [g], Yuehong Zhou [r], Xiaofeng Wei [a,z,aa], Sumin Zhao [s,t], Decheng Cai [a,z,aa], Fengyu Guo [s,t], Qianqian Zhang [a,z,aa], Yun Li [u], Xuelian Zhang [a,z,aa], Haorong Lu [u], Huajie Huang [a,z,aa], Junfu Guo [s,t], Fei Zhu [a,z,aa], Yuan Yuan [s,t], Li Zhang [a,z,aa], Na Liu [u], Zhiming Li [a,z,aa], Hui Jiang [b], Qiang Zhang [a,z,aa], Yijia Zhang [a,z,aa], Wan Khairunnisa Wan Juhari [v], Sarifah Hanafi [v], Wanjun Zhou [a,z,aa], Fu Xiong [a,z,aa], Huanming Yang [b,y], Jian Wang [b,y], Bin Alwi Zilfalil [v], Ming Qi [w,ab], Yaping Yang [x], Ye Yin [b], Mao Mao [b,*], Xiangmin Xu [a,z,aa,**]

[a] Department of Medical Genetics, School of Basic Medical Sciences, Southern Medical University, Guangzhou, Guangdong, China
[b] BGI-Shenzhen, Bei Shan Industrial Zone, Yantian District, Shenzhen, Guangdong, China
[c] Department of Hematology, 303rd Hospital of the People's Liberation Army, Nanning, Guangxi, China
[d] The Second Department of Pediatrics, Affiliated Hospital of Zunyi Medical College, Zunyi, Guizhou, China
[e] Genetic Diagnosis Center, First People's Hospital of Yunnan Province, Medical School of Kunming University of Science and Technology, Kunming, Yunnan, China
[f] Department of Biochemistry and Molecular Biology, Hainan Medical College, Haikou, Hainan, China
[g] Department of Genetic and Metabolic Laboratory, Guangxi Zhuang Autonomous Region Women and Children Health Care Hospital, Nanning, Guangxi, China
[h] Department of Medical Genetics, Liuzhou Municipal Maternity and Child Healthcare Hospital, Liuzhou, Guangxi, China
[i] Maternity and Child Health Care Hospital of Foshan City, Foshan, Guangdong, China
[j] Guilin Women and Children health care hospital, Guilin, Guangxi, China
[k] Department of Clinical Laboratory, Zhuhai Municipal Maternal and Child Healthcare Hospital, Zhuhai Institute of Medical Genetics, Zhuhai, Guangdong, China
[l] Maternal and Child Health Hospital in Meizhou, Meizhou, Guangdong, China
[m] Department of Prenatal Diagnosis Center, Dong Guan Maternal and Child Health Hospital, Dongguan, Guangdong, China
[n] Baise Women and Children Care Hospital, Baise, Guangxi, China
[o] Genetic Laboratory, Qinzhou Maternaland Child Health Hospital, Qingzhou, Guangxi, China
[p] Women and Children's Health Hospital of Shaoguan, Shaoguan, Guangdong, China
[q] Department of Gynecology and Obstetrics, The People's Hospital of Yunfu City, Yunfu, Guangdong, China
[r] Pingguo Women and Children Care Hospital, Baise, Guangxi, China
[s] Tianjin Medical Laboratory, BGI-Tianjin, BGI-Shenzhen, Tianjin, China
[t] Binhai Genomics Institute, BGI-Tianjin, BGI-Shenzhen, Tianjin, China
[u] BGI Clinical Laboratories-Shenzhen, BGI-Shenzhen, Shenzhen, China
[v] Department of Paediatric, School of Medical Sciences, Universiti Sains Malaysia, Kelantan, Malaysia
[w] School of Basic Medical Sciences, Zhejiang University, Hangzhou, Zhejiang, China
[x] Departments of Molecular and Human Genetics, Baylor College of Medicine, Houston, USA
[y] James D. Watson Institute of Genome Sciences, Hangzhou, Zhejiang, China
[z] Guangdong Technology and Engineering Research Center for Molecular Diagnostics of Human Genetic Diseases, Guangzhou, Guangdong, China
[aa] Guangdong Key Laboratory of Biological Chip, Guangzhou, Guangdong, China
[ab] Center for Genetic & Genomic Medicine, Zhejiang University Medical School 1st Affiliated Hospital, James Watson Institute of Genome Sciences, Hangzhou, Zhejiang, China

ARTICLE INFO

Article history:
Received 29 June 2017
Received in revised form 15 August 2017
Accepted 15 August 2017
Available online xxxx

ABSTRACT

Hemoglobinopathies are among the most common autosomal-recessive disorders worldwide. A comprehensive next-generation sequencing (NGS) test would greatly facilitate screening and diagnosis of these disorders. An NGS panel targeting the coding regions of hemoglobin genes and four modifier genes was designed. We validated the assay by using 2522 subjects affected with hemoglobinopathies and applied it to carrier testing in a cohort of 10,111 couples who were also screened through traditional methods. In the clinical genotyping analysis of 1182

* Correspondence to: Mao Mao, BGI-Shenzhen, Bei Shan Industrial Zone, Yantian District, Shenzhen, Guangdong 518000, China.
** Correspondence to: Xiangmin Xu, Department of Medical Genetics, School of Basic Medical Sciences, Southern Medical University, Guangzhou, Guangdong 510515, China.
E-mail addresses: maomao@genomics.cn (M. Mao), xixm@smu.edu.cn, gzxuxm@pub.guangzhou.gd.cn (X. Xu).
[1] These authors contributed equally to this study.

http://dx.doi.org/10.1016/j.ebiom.2017.08.015

Keywords:
Hemoglobinopathy
Next-generation sequencing
Molecular screening
Clinical genotyping

β-thalassemia subjects, we identified a group of additional variants that can be used for accurate diagnosis. In the molecular screening analysis of the 10,111 couples, we detected 4180 individuals in total who carried 4840 mutant alleles, and identified 186 couples at risk of having affected offspring. 12.1% of the pathogenic or likely pathogenic variants identified by our NGS assay, which were undetectable by traditional methods. Compared with the traditional methods, our assay identified an additional at-risk 35 couples. We describe a comprehensive NGS-based test that offers advantages over the traditional screening/molecular testing methods. To our knowledge, this is among the first large-scale population study to systematically evaluate the application of an NGS technique in carrier screening and molecular diagnosis of hemoglobinopathies.

1. Introduction

Hemoglobinopathies, including sickle cell anemia and α-/β-thalassemia, are the most common monogenic diseases worldwide (Higgs et al., 2012). Approximately 340,000 children with significant hemoglobin (Hb) disorders are born worldwide each year, 90% of whom are in developing and low-income countries, where they represent a massive public health burden (Williams and Weatherall, 2012). Prevention programs based on carrier screening and prenatal diagnosis have resulted in a continuous decline in rates of thalassemia major at birth in the Mediterranean region (Cao and Kan, 2013). Such prospective screening has recently been introduced in at-risk countries in Asia (Ahmed et al., 2002; Cao and Kan, 2013). Because the identification of high-risk couples by carrier testing is necessary for effective prevention, prevention programs must determine how to perform comprehensive carrier screening in large populations.

A traditional routine strategy for thalassemia carrier screening first identifies people with phenotypic traits associated with thalassemia by using hematological and biochemical tests and subsequent molecular genetic testing in this selected group to generate definitive diagnoses (Cao and Kan, 2013; Traeger-Synodinos et al., 2015). However, a phenotypic screening approach in carrier testing would not detect individuals with "silent" forms of thalassemia, most of whom may be missed because they have normal or borderline red cell indices and/or HbA_2 levels (Piel and Weatherall, 2014; Traeger-Synodinos et al., 2015). In addition, the characterization of disease-causing defects in samples from individuals suspected of having thalassemia may require various labor-intensive methodologies (Cao and Kan, 2013; Traeger-Synodinos et al., 2015). At least 1530 mutations causing thalassemia or abnormal hemoglobin variants, ranging from single-base changes to large rearrangements, have been characterized to date (HbVar database for human hemoglobin variants and thalassemia mutations, n.d). Moreover, several variants that modify β-thalassemia phenotypes have been identified at various loci (Giardine et al., 2011; Sankaran and Weiss, 2015; Thein, 2013). The application of high-throughput molecular approaches to direct mass screening and accurate molecular diagnosis of hemoglobinopathies raises several challenges.

Next-generation sequencing (NGS) has been shown to allow rapid, multiplex and high-throughput detection of genetic variants (Korf and Rehm, 2013). NGS technologies—applied to the whole genome, the exome, or targeted gene panels—have been effectively used in research settings, as well as in clinical testing and diagnosis of genetic disorders (Stark et al., 2016; Yang et al., 2013). Recently, target capture and NGS have been validated in patient-based and population-based carrier testing of Mendelian recessive diseases, thereby yielding high-quality genotype calls and acceptable false-positive and false-negative rates and cost-effectiveness (Chong et al., 2012; Haque et al., 2016).

Here, to evaluate whether NGS might be suitable for application in clinical management of monogenic diseases in a large-scale population, we retrospectively analyzed 2522 subjects with various hemoglobinopathies by using a targeted NGS approach and validated this assay for carrier testing of multiple genetic defects in a cohort of 20,222 individuals pre-typed by routine screening methods. A population-based study of hemoglobinopathies in southern China was also performed. Our results demonstrated that NGS, compared with traditional methods, can detect pathogenic or likely pathogenic variants in a more precise and general manner, thus providing an effective platform for molecular screening and clinical genotyping in subjects with hemoglobinopathies.

2. Materials and Methods

2.1. Study Participants

The study was approved by the Medical Ethics Committee in accordance with the Declaration of Helsinki. The study methodology was designed with guidance by STROBE (von Elm et al., 2007).

Samples from two groups of subjects were collected for this study. In Group I, 2522 subjects included 2087 samples used for study methodology of NGS assay and the sample data of 435 Thalassemia Intermedia (TI)/Thalassemia Major (TM) patients from our previous cohort (Fig. 1) (Liu et al., 2014). The 2087 subjects were from China and Malaysia and included 49 normal controls and 2038 clinical samples with various hemoglobinopathies (Table S1 in the Supplementary appendix). They were used for assessing the performance of the NGS-based assay in detecting hemoglobinopathy-causing mutations in a retrospective study. The 49 controls had been sequenced previously, and they did not carry any pathogenic or likely pathogenic variants. The 2038 clinical samples had been analyzed by traditional molecular genotyping techniques (See the "Genotypic Analysis Using Traditional Methods" section). NGS analysis was performed on these pre-typed results with blinding, and the results were subsequently compared with the pre-typed results to obtain the concordance rate. Sanger sequence analysis and multiplex ligation-dependent probe amplification (MLPA) were performed to verify samples with discrepancies. Because some modifier genes had been shown to be responsible for modifying β-thalassemia phenotypes, we also analyzed whether genotypes at these loci might assist in β-thalassemia diagnosis in 1182 β-thalassemia patients with thalassemia major (TM) or thalassemia intermedia (TI; 747 patients of the 2087 samples described above and 435 patients from the cohort described in a previous study by our group (Fig. 1; Liu et al., 2014). Among the 747 samples, 510 were β^0/β^0, and 237 were β^0/β^+. Among the 435 samples, 300 were β^0/β^0 and 135 were β^0/β^+. In total, the 1182 samples included 810 β^0/β^0 and 372 β^+/β^0.

Group II, which included a total of 10,111 couples (either before or during pregnancy) were randomly selected from five provinces (Guangxi, Guandong, Yunnan, Guizhou and Hainan) in southern China (Fig. 2). The age range of these samples was 18–48 years old. We applied the assay in a blind analysis for carrier testing on these couples, who were also enrolled in prevention programs for thalassemias using traditional screening methods (See the "Hematological Analysis" and "Genotypic Analysis Using Traditional Methods" sections). They were also included in an epidemiological survey of hemoglobinopathies.

2.2. Thalassemia Genotype Definition and Clinical Diagnosis in Thalassemia Patients

HBB genotype categories are defined as follows: (β^0): HBB:c.124_127delTTCT, HBB:c.52A>T, HBB:c.316-197C>T,