Next generation sequencing in research and diagnostics of ocular birth defects

Gordana Raca, Craig Jackson, Berta Warman, Tom Bair, Lisa A. Schimmente

The advent of NGS technologies is expected to transform the practice of medical genetics [1–3]. With the high throughput and decreased sequencing costs achieved by NGS, it is no longer impossible to sequence hundreds or even thousands of exons and other genomic sequences in an individual with a suspected genetic disease. It is predicted that in the near future NGS might replace array based target enrichment followed by re-sequencing on the Roche Genome Sequencer FLX (GS FLX) system could be used for novel mutation identification in more than 1000 exons representing 100 candidate genes for ocular birth defects, and as a control, whether these methods could detect two known mutations in the PAX2 gene. We assayed two samples with heterozygous sequence changes in PAX2 that were previously identified by conventional Sanger sequencing. These changes were a c.527G > C (S176T) substitution and a single basepair deletion c.77delG. The nucleotide substitution c.527G > C was easily identified by NGS. A deletion of one base in a long polyG stretch (c.77delG) was not registered initially by the GS Reference Mapper, but was detected in repeated analysis using two different software packages. Different approaches were evaluated for distinguishing false positives (sequencing errors) and benign polymorphisms from potentially pathogenic sequence changes that require further follow-up. Although improvements will be necessary in accuracy, speed, ease of data analysis and cost, our study confirms that NGS can be used in research and diagnostic settings to screen for mutations in hundreds of loci in genetically heterogeneous human diseases.

Keywords:
Next generation sequencing
Sequence capture
GS FLX
Anophthalmia
Microphthalmia
Coloboma

NGS could be particularly advantageous in research and testing for genetically heterogeneous hereditary conditions. Common disorders evaluated by clinical geneticists are caused by heterogeneous Mendelian loci and lend themselves to enrichment strategies followed by NGS. Examples include intellectual disability [6], deafness [7], familial cardiomyopathy [8] and retinitis pigmentosa [9]. In these conditions there are often very subtle phenotypic differences between affected patients to guide molecular diagnostics by indicating which gene is likely to be mutated in a particular individual. Current diagnostic evaluation proceeds by sequencing a series of genes, individually or in small sets, based on the relative frequency of the mutations and the sensitivity of available assays. If there is no predominant mutation(s) causing the disease, the pathogenic change often remains unknown even after very extensive and expensive molecular testing. With enrichment strategies followed by NGS, sequencing of all genes implicated in a particular genetic disorder could be performed simultaneously, efficiently and at low cost.

While clearly superior to traditional Sanger sequencing, NGS has had little impact on clinical testing to date. There are very few examples of successful application of NGS in translational research and diagnostics. Clinical testing using NGS is currently offered for Hypertrophic Cardiomyopathy (HCM), Dilated...
Cardiomyopathy (DCM) and Long QT syndrome (http://www.genedx.com/). NGS has also been explored as a method to perform rapid human leukocyte antigen (HLA) typing for high resolution allele identification [10,11], and to develop assays for Neurofibromatosis Type 1 [12], autosomal recessive ataxia [13] and mitochondrial disorders caused by mutations in the mitochondrial genome and 362 nuclear genes controlling mitochondrial function [14]. Ng et al. applied targeted sequencing of all coding regions (“exome”) to show the presence of causative mutations in four unrelated individuals with a rare dominantly inherited disorder, Freeman–Sheldon syndrome (FSS) [15] and to discover the gene for a rare recessive disorder of previously unknown cause, Miller disease [16]. Exciting applications have also been described in cancer research, where NGS has been applied in discovering new candidate genes for acute myeloid leukemia [17,18], glioblastoma [19] and other malignancies [20].

We describe the first study which showed the feasibility of using genomic enrichment by sequence capture followed by NGS to investigate genetic causes of the ocular birth defects anophthalmia, microphthalmia and coloboma. These eye anomalies are among the most prevalent causes of childhood blindness, affecting annually ~2 per 10,000 newborns worldwide [21]. Although they can be of different origins, the majority are caused by defects in genes which regulate normal eye development [22–25]. There is increased evidence that mutations in large numbers (possibly hundreds) of different genes can cause congenital eye malformations, but no single gene is responsible for a high percentage of cases [23–25]. Anophthalmia, microphthalmia and coloboma therefore represent disorders where simultaneous sequencing of large numbers of candidate genes by NGS is an ideal approach to study genetic causes and demonstrate the feasibility of NGS for clinical diagnostics.

We showed that the combination of array based SCE re-sequencing on the GS FLX instrument using Titanium chemistry allows concurrent sequencing of more than 100 candidate genes for anophthalmia, microphthalmia and coloboma. However, improvements will be necessary in several areas including accuracy, speed, ease of data analysis and cost to allow successful diagnostic implementation of NGS for simultaneous mutation testing in hundreds of genes in genetically heterogeneous human diseases.

Materials and methods

We tested whether two known sequence changes in the renal-coloboma syndrome (a.k.a. Papillorenal syndrome, OMIM #120330) associated gene PAX2, which were previously characterized by Sanger sequencing, can reliably be detected by NGS. The first variant was a missense change in exon 5, c.527G > C, which resulted in serine to threonine amino-acid change S176T. This base substitution was identified in one of our previous studies in a father of a proband with ocular birth defects, but not in his affected child. Since detailed clinical information for the parent was not available, the change was described as a variant of unknown clinical significance.

Since short stretches of mono-, di-, and trinucleotide repeats represent hotspots for disease causing frameshift mutations in genomic DNA, we wanted to determine if this mutation type is detectable by NGS. Therefore, the second sequence change selected for the study was a deletion of one base in a polyG stretch in exon 2 of the PAX2 gene (c.77delG), which has previously been described by Schimenti et al. [26].

We designed a custom 385,000 probe SCE array with more than 100 candidate genes for eye malformations. The list of selected candidate genes is provided in Table 1. The genes were chosen based on reports of mutations identified in patients with coloboma, microphthalmia and anophthalmia [22,23,25]. Additionally, a comprehensive literature search was performed for published mutations associated with ocular phenotypes in animal models [27,28]. Some genes were included based on their role in signaling and developmental pathways which are known as important for eye formation and function [24].

Appropriate SCE probes for the regions of interest were chosen in collaboration with Roche-NimbleGen design team (Roche-NimbleGen, Madison, WI). Only protein coding regions (coding exons) of the 112 candidate genes were targeted on the array. We selected 385,000 long oligonucleotide probes (>60 bp) to tile the exons of the genes of interest with a very high density. All the probes had the uniqueness score of one (defined as having no match in the genome other than itself longer than 38 bp, allowing up to five insertions/deletions/mismatches in that match), to exclude repetitive regions from probe selection and avoid capturing pseudogene sequences [29]. Upon completion of the design and manufacturing of the array, the SCE on samples of genomic DNA from two patients with known mutations in the PAX2 gene was performed at the NimbleGen service laboratory (Roche-NimbleGen, Madison, WI), following previously described protocols [30–32]. Briefly, ~20 µg of each patient’s genomic DNA were randomly fragmented by nebulization to an average size of 500 bp. Linkers were ligated to the DNA fragments to provide a priming site for post-enrichment amplification of the eluted fragment pool. The fragments were then amplified and hybridized to the custom SCE array. After a 72-h hybridization, unbound material was removed by stringent washing. The arrays were transferred to the NimbleGen elution system, and the enriched fragment pool was eluted and recovered from the array. The enriched fragments were amplified with 22mer linkers to generate enough DNA template for downstream applications. After amplification, the amount of captured DNA was measured by spectrophotometry and the product was tested for enrichment level by quantitative PCR with four proprietary QC control loci. These QC loci are conserved in both human and mouse genomes and have been empirically determined to accurately predict enrichment with several different array designs.

Sequencing of the two SCE prepared samples was performed on the GS FLX instrument using Titanium chemistry, at the University of Iowa DNA Facility following standard protocols. Two samples, separated by gaskets, were sequenced independently on two regions of the picotiter-plate. Briefly, amplified fragments from SCE were end-repaired and ligated to adapter oligonucleotides. The library was diluted based on the results of a previously performed titration, so that upon denaturation single DNA fragments hybridized to individual beads containing sequences complementary to adapter oligonucleotides. The beads were compartmentalized into water-in-oil microvesicles to allow clonal expansion of separate DNA molecules bound to the beads by emulsion PCR. After amplification, the emulsion was disrupted, and the beads containing clonally amplified template DNA were enriched. The beads were again separated by limiting dilution, deposited into individual picotiter-plate wells, and combined with sequencing enzymes. Iterative pyrosequencing was performed on the picotiter-plate by successive flow addition of the four dNTPs. A nucleotide-incorporation event in a well containing clonally amplified template produced pyrophosphate release and picotiter-plate well-localized luminescence, which recorded by a charge-coupled device (CCD) camera. With the flow of each dNTP reagent, wells were imaged, analyzed for their signal-to-noise ratio, filtered according to quality criteria, and subsequently algorithmically translated into a linear sequence output [33–35].

Data analysis was performed using the Roche proprietary software package for the GS FLX system. Image acquisition, image processing and signal processing were performed during the run. Post run analysis was conducted using the GS Reference Mapper.
Sequence runs were mapped against the human reference genome (hg18), using the default software settings (minimum base overlap of 40 bp and minimum overlap identity 90%). Sequence variations were detected automatically during mapping, and were annotated with known gene (reSeq genes from http://genome.ucsc.edu/) and SNP information (dbSNP129 from http://genome.ucsc.edu/). Variants were determined as high quality differences (HCdiffs) if the change was present in at least three non duplicate reads which included at least one read from each direction (forward and reverse). Additional analysis was performed with the CLC Genomic Workbench software (CLCbio, Aarhus, Denmark) and the NextGENe™ software (SoftGenetics, State College, PA).

Results

We performed target enrichment by array based SCE and sequencing using GS FLX instrument on two DNA samples with known mutations in the PAX2 gene. The sequencing was performed simultaneously for 112 candidate genes for ocular birth defects which were selected based on extensive literature search. A custom SCE array designed for target enrichment contained probes for all coding regions (total of 1017 exons) of the selected 112 genes. The size of the entire target region was 373,083 bp. Since 385,000 probe SCE arrays are able to capture up to 5 Mb of target sequence [32], our target region of ~0.37 Mb only used a portion of the capacity of the 385 K array. This allowed the use of a large number of probes per each targeted region and helped in obtaining good enrichment. DNA yield after elution from the SCE arrays and PCR amplification was approximately 10^{-3} g for each sample; completely lacked coverage were in both cases the first coding exons of the genes BMP4, MITF, HESX1, HESX1, SIX3, RAB3GAP1, ZFHXB1, ALG2, PITX2, SHH, PRKCN, PPP1R11, CREBBP, SALL1 NDP, MKS1, SALL4, PDBP1, ICBR, PORCN, IGBP1, FLNA, CHD7, CC2D2A, HMX1, CLDN19, LRPI, TAP2A, IGBP1, IKBKG, BSGALTL, GDF6, SX2, OTX2, JAG1, BMP4, MITF, HESX1.

Table 1
Candidate genes for anophthalmia, microphthalmia and coloboma represented on the SCE array.

<table>
<thead>
<tr>
<th>Implicated in</th>
<th>Implicated in</th>
<th>Code for proteins on the</th>
<th>Code for proteins on the</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>syndromic anophthalmia, microphthalmia and coloboma</td>
<td>non-syndromic anophthalmia, microphthalmia and coloboma</td>
<td>the SHH</td>
<td>the WNT signaling pathway</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2**
GS FLX sequencing summary.

<table>
<thead>
<tr>
<th>Sample 1</th>
<th>Sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence yield</td>
<td>214 Mb</td>
</tr>
<tr>
<td>Bases on target</td>
<td>76.1 Mb</td>
</tr>
<tr>
<td>Average depth</td>
<td>206</td>
</tr>
<tr>
<td>Average coverage</td>
<td>99.24%</td>
</tr>
<tr>
<td>% Covered ≥15x</td>
<td>94.87</td>
</tr>
<tr>
<td># Regions with &lt;100% coverage</td>
<td>54</td>
</tr>
<tr>
<td># Regions without coverage</td>
<td>3</td>
</tr>
<tr>
<td>Total # of detected sequence changes</td>
<td>376</td>
</tr>
<tr>
<td># Changes in coding regions</td>
<td>110</td>
</tr>
<tr>
<td># Changes not corresponding to known SNPs</td>
<td>116</td>
</tr>
<tr>
<td># Changes in coding regions not corresponding to known SNPs</td>
<td>26</td>
</tr>
<tr>
<td># Changes likely to be real sequence alterations</td>
<td>5</td>
</tr>
</tbody>
</table>

* a Bases on target were determined based on rigorous criteria, so that bases that are a single-nucleotide outside of the captured region are not included in the count, even if they map to regions flanking the targets.
* b Average depth is defined as average number of reads covering individual targeted nucleotides.
* c Average coverage is the average proportion of bases within targeted exons which were covered by reads.
* d These sequence variants met the following selection criteria: (1) they do not represent synonymous substitutions, (2) there is at least 30-fold coverage at the variant site and (3) the percentage of reads showing the variant allele exceeds 30%.

(out of a total of 1017) 15× coverage was not obtained for all the bases within a region. Most regions with low coverage (44/51) were identical between the two samples. The regions which completely lacked coverage were in both cases the first coding exons of the genes WNT4, WNT9A and LRPS, which were noted to have a very high GC-content (83%, 81% and 83%, respectively). Additionally, low complexity GC rich repetitive elements were present in the vicinity (100 bp upstream or downstream) of all the three exons without coverage. Unusual sequence characteristics with high GC-content and a presence of repetitive elements are well known causes of poor enrichment by SCE and decreased sequencing efficiency [12].

Studies have shown that the coverage in the 10–15-fold range may be sufficient for re-sequencing applications, but higher coverage depths (50–60-fold) provide better alignment, assembly and accuracy [36]. Therefore, most of our target exons have higher coverage than needed for a robust re-sequencing assay, allowing to expand our target region by incorporating additional genes and to sequence our samples as a pool.

The missense change in Sample 1 (c.527G > C, S176T) was easily identified by NGS. A total of 151 reads covered the variant site, with 52% of the reads showing the wild-type base and 48% showing the variant base, as would have been expected for a heterozygous allele. The c.527G > C change identified by Sanger sequencing is
shown in Fig. 2A, while the same change detected by NGS is presented in Fig. 2B.

A deletion of one base in a long polyG stretch was observed. A total of 76 reads covered the region, with ~38% showing the wild-type allele (7 Gs), ~54% showing the mutation (6 Gs) and the remaining ~8% showing different numbers of Gs which did not correspond to either allele present in the sample (Fig. 3). Surprisingly, the c.77delG mutation was not registered by the GS Reference Mapper as a high quality difference. We therefore performed additional analysis with the CLC Genomic Workbench.
and the NextGENe™ analysis programs, focusing only on sequence changes in the coding region of the PAX2 gene. The c.77delG was clearly identified both by the CLC Genomic Workbench and the NextGENe™ analysis. Reasons for the discrepancy between different analysis tools are unclear, however, it is well known that pyrosequencing based methods (like GS FLX) tend to erroneously interpret long stretches (>6) of the same nucleotide [33,37]. To compensate for the resulting “noise” at all homonucleotide sites the GS Reference Mapper may be less likely to register base deletions and insertions at such regions.

We performed detailed analysis of sequence changes detected in targeted genes other than PAX2, to determine how well can SCE–NGS be applied to analysis of unknown samples from patients with eye anomalies. Since NGS applications usually identify a large amount of genetic variation, we wanted to determine if false positive findings (sequencing errors) and benign polymorphisms can readily be distinguished from possibly pathogenic changes which require further follow-up. The summary of this analysis is provided in Table 2. To consider clinically important variants, both in the SCE array design and in the analysis, we focused only on coding sequences. Both homozygous and heterozygous changes were included, since both recessive and dominant models of inheritance are considered for the studied disorders. We filtered out all known nonpathogenic single-nucleotide polymorphisms reported in the dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/). This left the total of 26 changes in coding regions not corresponding to known SNPs in both Sample 1 and Sample 2 (Table 2). We than applied three additional previously described parameters to further reduce the number of candidate mutations in tested samples. We focused only on changes detected in the regions with sufficient coverage (>30x) with the rationale that those are less likely to represent false positives [36]. We also required that the percentage of reads showing the variant allele exceeds 30% [12]. Using these filters we reduced the number of detected coding alterations which were likely to represent real sequence changes (rather than sequencing errors) to five (two frameshift and three missense mutations) in Sample 1 and three (two frameshift and one missense mutation) in Sample 2 (listed in Table 3). Only these eight alterations were selected for confirmation by Sanger sequencing, although the presence of other real sequence changes (which did not meet our selection criteria) could not be excluded. The base substitutions (missense mutations) detected by NGS were also identified by Sanger sequencing, but the frameshift mutations were not confirmed (Table 3). This is not surprising, however, since all the frameshift mutations detected by NGS (deletion “CGA” in the FGFR2 gene, deletion “T” in the CHD2 gene, deletion “A” in the APC gene and deletion “A” in the LRP2 gene) mapped to long homonucleotide stretches, where sequencing errors are known to occur with GS FLX technology. Although the frameshift changes listed in Table 3 meet stringent selection criteria, they still represent false positives, showing the difficulty in accurate detection of frameshift mutations by pyrosequencing based NGS technologies.

PolyPhen (http://genetics.bwh.harvard.edu/pph/) and SIFT (http://sift.jcvi.org/) analyses were performed to examine the likelihood that the observed missense changes alter the functions of the encoded proteins. Both analysis tools predicted one missense mutation (R942C in the RAB3GAP1 protein) to be probably damaging, and one (I744V in ZEB2) to be likely benign (Table 3). However, for the remaining two amino-acid changes (I3389V in LRP2 and L1129S in APC) PolyPhen and SIFT gave different predictions, thus showing that an accurate determination of functional significance of amino-acid changes in proteins cannot be achieved based solely on computer analysis. In summary, although confirmation by Sanger sequencing and the use of bioinformatics tools may have decreased the number of sequence variants which would have needed further studies in unknown samples, our approach also required that the percentage of reads showing the variant allele exceeds 30% [12]. Using these filters we reduced the number of detected coding alterations which were likely to represent real sequence changes (rather than sequencing errors) to five (two frameshift and three missense mutations) in Sample 1 and three (two frameshift and one missense mutation) in Sample 2 (listed in Table 3). Only these eight alterations were selected for confirmation by Sanger sequencing, although the presence of other real sequence changes (which did not meet our selection criteria) could not be excluded. The base substitutions (missense mutations) detected by NGS were also identified by Sanger sequencing, but the frameshift mutations were not confirmed (Table 3). This is not surprising, however, since all the frameshift mutations detected by NGS (deletion “CGA” in the FGFR2 gene, deletion “T” in the CHD2 gene, deletion “A” in the APC gene and deletion “A” in the LRP2 gene) mapped to long homonucleotide stretches, where sequencing errors are known to occur with GS FLX technology. Although the frameshift changes listed in Table 3 meet stringent selection criteria, they still represent false positives, showing the difficulty in accurate detection of frameshift mutations by pyrosequencing based NGS technologies.
examples illustrate that appropriate evaluation of detected changes in patient specimens will often require parental testing and functional analysis. This requirement may represent a significant challenge in the clinical settings.

Multiple known single-nucleotide polymorphisms were detected, providing reassurance that both SCE and GS FLX sequencing performed as expected. For example, Sample 2 showed presence of the previously described polymorphism in the exon 8 of the PAX2 gene 1521A > C, P326P (rs1800898) [38].

Discussion

SCE and GS FLX sequencing allowed us to simultaneously sequence 112 candidate genes for ocular birth defects in two independent DNA samples, and to detect two known mutations in the PAX2 gene. Although whole genome sequencing represents the most comprehensive approach to identifying disease causing mutations in patients with hereditary disorders, such sequencing and analysis is not yet possible [39]. We therefore focused our studies on genes which are known to be associated with ocular birth defects, as well as numerous candidate genes which play a role in regulating early eye development.

To overcome the limitations of target enrichment by PCR, we used high-density SCE microarrays for selecting and capturing relevant exons [29–32]. The GS FLX instrument with Titanium chemistry was chosen based on its ability to produce long reads which are advantageous for sequencing complex genomes. Additionally, protocols for target enrichment by array based SCE were originally developed and optimized for the GS FLX system. However, our methods can be easily adapted for use with other sequencing platforms, like the Genome Analyzer IIx from Illumina (Illumina, Inc., San Diego, CA) and the SOLiD System from Applied Biosystems (Life Technologies, Carlsbad, CA).

In our pilot study “hands on” and instrument run times for sequencing 112 genes in two independent specimens were measured in days, while the cost was less than $8000 per sample ($3000 for SCE and ~$5000 for sequencing). Performing enrichment by traditional PCR and Sanger sequencing for the same number of genes would have been a daunting task, even for facilities with high levels of automation and high throughput. Although our costs for performing SCE and NGS were high compared to savings that can potentially be achieved with these technologies, they were still modest since the price for diagnostic Sanger sequencing for only one gene can be thousands of dollars. Furthermore, there are multiple ways to further optimize our approach. Significant increase in sequencing quality and throughput and decrease in cost can be achieved by: (1) running multiple samples in the same run either by separating them physically (with gaskets) or by adding short oligonucleotide adapters as “barcodes”, and running the samples as a pool [40,41], (2) optimizing the custom SCE array to achieve more uniform coverage, by adding additional probes for

Fig. 3. A screenshot from the CLCbio analysis software showing the frameshift mutation c.del177G. The mutated base is shown between the two vertical grey lines. The reads showing the mutant allele (6 Gs) are grouped at the top, while the reads generated from the wild-type allele (7 Gs) are shown at the bottom.
Table 3

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Reference position</th>
<th>Variant name</th>
<th>Variant nucleotide</th>
<th>Variant percentage (%)</th>
<th>Reference AA</th>
<th>Nucleotide (frameshift)</th>
<th>PolyPhen prediction</th>
<th>SIFT prediction</th>
<th>Variant percentage (%)</th>
<th>Reference AA</th>
<th>Nucleotide (frameshift)</th>
<th>PolyPhen prediction</th>
<th>SIFT prediction</th>
<th>Variant percentage (%)</th>
<th>Reference AA</th>
<th>Nucleotide (frameshift)</th>
<th>PolyPhen prediction</th>
<th>SIFT prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr2</td>
<td>169,750,453</td>
<td>FGFR2</td>
<td>99</td>
<td>FG</td>
<td>Not confirmed</td>
<td>Not confirmed</td>
<td>Tolerated</td>
<td>Tolerated</td>
<td>Not confirmed</td>
<td>Not confirmed</td>
<td>Not confirmed</td>
<td>Tolerated</td>
<td>Tolerated</td>
<td>Not confirmed</td>
<td>Not confirmed</td>
<td>Not confirmed</td>
<td>Tolerated</td>
<td>Tolerated</td>
</tr>
<tr>
<td>chr15</td>
<td>9,135,594</td>
<td>RAB3GAP1</td>
<td>155</td>
<td>A</td>
<td>Not confirmed</td>
<td>Not confirmed</td>
<td>Tolerated</td>
<td>Tolerated</td>
<td>Not confirmed</td>
<td>Not confirmed</td>
<td>Not confirmed</td>
<td>Tolerated</td>
<td>Tolerated</td>
<td>Not confirmed</td>
<td>Not confirmed</td>
<td>Not confirmed</td>
<td>Tolerated</td>
<td>Tolerated</td>
</tr>
</tbody>
</table>

* This SIFT prediction was described as "low confidence". The amino-acid change may have been predicted to affect the protein function just because other protein sequences used for comparison were not diverse enough.
Data storage and analysis are large challenges for individual laboratories considering use of NGS. The issue of data storage is required for a long time to confirm data generated by NGS. Even large sample sets will need to be studied and specific, objective, evidence based guidelines and quality controls developed to address main technical issues and enable use of SCE and NGS technologies to reliably detect pathogenic mutations in research and diagnostics laboratories. Routine Sanger sequencing will likely be required for a long time to confirm data generated by NGS. Even when technical limitations get resolved, serious test reporting issues and ethical considerations are likely to remain. For example, each person is likely to carry multiple clinically relevant mutations which can be detected by genome-wide sequencing. It is unclear should participants in NGS based studies receive all their sequencing information (even when it is not possible to implement effective medical treatments for clinically relevant variants) and what should this information be delivered and interpreted to patients [45].

In conclusion, our study showed that even with existing limitations, individual researchers working with core sequencing facilities can use NGS as a research tool with confidence and ease. In the near future, thanks to continuing improvements in throughput, accuracy, cost and ease of data analysis, it will also become feasible to apply NGS in the clinical setting.

Acknowledgments

We gratefully acknowledge Einat Snir and Jennifer Bair from the University of Iowa DNA Facility for technical assistance with GS FLX sequencing and manuscript review, and Eric Cabot from the University of Wisconsin-Madison Biotechnology Center for help with reviewing the sequencing data.

References


G. Raca et al. / Molecular Genetics and Metabolism xxx (2010) xxx–xxx 9

Please cite this article in press as: G. Raca et al., Next generation sequencing in research and diagnostics of ocular birth defects, Mol. Genet. Metab. (2010), doi:10.1016/j.ymgme.2010.03.004