Russo et al., J Clin Case Rep 2014, 4:11
DOI: 10.4172/2165-7920.1000455

Journal of Clinical Case Reports

Case Report Open Access

Comparative Study of a CGH and Next Generation Sequencing (NGS) for Chromosomal Microdeletion and Microduplication Screening

Claudio Dello Russo[1]*, Gianluca Di Giacomo[1], Alvaro Mesoraca[1], Domenico Bizzoco[1], Megan R McCluskey[2], Arianna Paganelli[3] and Claudio Giorlandino[4]

[1]*Department of Genetics, Altamedica-Artemisia SpA-Rome, Italy*
[2]*Softgenetics LLC, State College, Pennsylvania, USA*
[3]*Diatech Lab Line srl, Jesi, Italy*
[4]*Department of Fetal Maternal Medicine, Altamedica-Artemisia SpA-Rome, Italy*

Abstract

Background: Prenatal genetic diagnosis of rare disorders is undergoing in recent years a significant enhancement through the application of methods of massive parallel sequencing. Despite the quantity and quality of the data produced, just few analytical tools and software have been developed in order to identify structural and numerical chromosomal anomalies through NGS, mostly not compatible with bench top NGS platform and routine clinical diagnosis.

Methods: We developed technical, bioinformatic, interpretive and validation pipelines for Next Generation Sequencing to identify SNPs, indels, aneuploidies, and CNVs (Copy Number Variations).

Results: We show a new targeted resequencing approach applied to prenatal diagnosis. For sample processing we used an enrichment method for 4,813 genes library preparation; after sequencing our bioinformatic pipelines allowed both SNPs analysis for approximately thirty diseases or diseases family involved in fetus development and numerical chromosomal anomalies screening.

Conclusions: Results obtained are compatible with those obtained through the gold standard technique, aCGH array, moreover allowing identification of genes involved in chromosome deletions or duplications and exclusion of point mutation on allele not affected by chromosome aberrations.

Keywords: Next generation sequencing; Copy Number Variation (CNV); Prenatal diagnosis

Introduction

Prenatal genetic diagnosis of rare disorders is undergoing in recent years a significant enhancement through the application of methods of massive parallel sequencing. In recent years Next Generation Sequencing (NGS) has become an important tool not only for gene discovery and research area but also for clinical diagnosis. To date, few studies have described the clinical use of NGS in prenatal diagnosis, most of which have concentrated on the study of single case report [1-7]. However, only a very limited number have evaluated the use of NGS for the identification of chromosome aneuploidies and rearrangements following birth [2] and before birth [3]. Despite the quantity and quality of the data produced, just few analytical tools and software have been developed in order to identify structural and numerical chromosomal anomalies through NGS, mostly not compatible with bench top NGS platform and routine clinical diagnosis. The current gold standard method for chromosomal micro deletions and micro duplications analysis is comparative genomic hybridization microarray (aCGH). The advantage of using NGS for a combined analysis of point mutations (SNPs), indels, aneuploidies, and CNVs (Copy Number Variations) is to increase the analysis resolution and detection rate with one single test. In addition this approach could allows SNPs analysis on locus affected by microdeletion/microduplication on the other allele or on correlated loci, so providing any possible information regarding genomic region and clinical effects.

We show a new targeted resequencing approach applied to prenatal diagnosis. For library preparation we use an enrichment method developed by Illumina; gene panel includes 4,813 genes, a cumulative target region size of 12 Mb, for a total of about 62,000 exons covered. Using a producer validated kit allowed us to avoid the development and validation of library for each gene of interest, obtaining 20X as minimum target coverage value. This strategy is consistent with small amount and quality of DNA extracted from prenatal sample, and especially with timing provided by prenatal diagnosis. After sequencing our bioinformatic pipelines allow both SNPs analysis for approximately thirty diseases or diseases family involved in fetus development and associated to 152 genes included in gene panel and structural and numerical chromosomal anomalies screening.

Here we show results obtained for chromosomal analysis using for NGS data processing Nextgene Software (Softgenetics). For this evaluation trial we compared NGS data to aCGH.

Materials and Methods

Choice of samples to be analysed and their processing

We analyzed 248 samples using both aCGH and NGS. The samples studied were obtained through DNA extraction from amniotic fluid and chorionic villi (QIAamp DNA Blood Mini Kit, Qiagen). Following extraction, the DNA is quantified through the Qubit® 2.0 Fluorometer system (Life technologies) and 2100 Bioanalyzer Instruments (Agilent Technologies).

aCGH: For aCGH analysis, we used BAC-array CytoChip Focus Constitutional following the manufacturer's instructions.

***Corresponding author:** Claudio Dello Russo, Department of Genetics, Altamedica-Artemisia SpA-Rome, Italy, Tel: 06 56 82 798; E-mail: claudio.dellorusso@artemisia.it

Received November 19, 2014; **Accepted** November 28, 2014; **Published** November 30, 2014

Citation: Russo CD, Di Giacomo G, Mesoraca A, Bizzoco D, McCluskey MR, et al. (2014) Comparative Study of a CGH and Next Generation Sequencing (NGS) for Chromosomal Microdeletion and Microduplication Screening. J Clin Case Rep 4: 455. doi:10.4172/2165-7920.1000455

With BAC-array CytoChip Focus Constitutional, it is possible to perform genome analysis through 1Mb resolution and 100-200 Kb resolution for 106 selected syndromic regions.

Next generation sequencing: library preparation and analysis

The targeted resequencing was performed using an illumina kit; a trusight one sequencing panel on a NEXTSEQ500 platform. This kit makes it possible to perform enrichment and final analysis of a panel of approximately 5000 genes. A Trusight one sequencing panel contains all the reagents necessary for the amplification, amplicon enrichment, indexing of the samples and the use of NextSeq 500 without needing any external reagents. Each procedure was realized following the manufacturer's instructions.

The NEXTSEQ500 system provides fully integrated on-instrument data analysis software. Base space Reporter software performs secondary analysis on the base calls and Phred-like quality score (Qscore) generated by Real Time Analysis software (RTA) during the sequencing run. The Trusight one sequencing panel workflow in NEXTSEQ500 Reporter evaluates short regions of amplified DNA (amplicons) for variants through the alignment of reads against a "manifest file" specified while starting the sequencing run. The manifest file is provided by Illumina and contains all the information on the custom assay. The workflow requires the reference genome specified in the manifest file (Homo sapiens, hg19, build 37.2). The reference genome provides variant annotations and sets the chromosome sizes in the BAM file output. The Trusight one sequencing panel workflow performs demultiplexing of indexed reads, generates FASTQ files, aligns reads to a reference, identifies variants, and writes output files for the Alignment folder. SNPs and short indels are identified using the Genome Analysis Toolkit (GATK), by default. GATK calls raw variants for each sample, analyzes variants against known variants, and then calculates a false discovery rate for each variant. Variants are flagged as homozygous (1/1) or heterozygous (0/1) in the variant call file sample column. Because a SNP database dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP) is available in the Annotation subfolder of the reference genome folder, any known SNPs or indels are flagged in the VCF output file. A reference gene database is available in the Annotation subfolder of the reference genome folder and any SNPs or indels that occur within known genes are annotated.

Each single variant reported in the VCF output file has been evaluated for the coverage and the Qscore and visualized via an Integrative Genome Viewer (IGV). Based on the guidelines of the American College of Medical Genetics and Genomics, all regions that have been sequenced with a sequencing depth <30 were considered unsuitable for analysis. Furthermore, we established a minimum threshold in Qscore of 30 (base call accuracy of 99.9%). For variant calling we used Variant Studio software (Illumina). For selection and reporting we used HGMD professional and ClinVar NCBI database.

Copy number variation analysis

Bam file obtained from sequencing were processed by Nextgene software (Sotgenetics) for copy number analysis.

Nextgene software (Sftgenetics) was developed for Copy-Number Variation (CNV) detection from a wide variety of projects, including whole-exome and targeted sequencing panels. Copy number variations are detected by comparing the coverage (RPKM) of specified regions in a "sample" project and a "control" project. The coverage ratio (sample divided by sample plus control) is used as the basis for CNV detection. A beta-binomial model is fit to the coverage ratio (similar to the recently published) ExomeDepth software in order to model the amount of dispersion. Likelihood values are calculated based on the dispersion measurements and coverage ratios. These probabilities are then entered into a Hidden Markov Model (HMM) to make CNV classifications for each region.

The resulting report gives a simple classification for each region-either "Duplication" (increased copy number), "Normal" (little evidence of a CNV), "Deletion", or "Uncalled" (due to low coverage). Additionally, each region receives three phred-scaled probability scores-Deletion, Normal, and Duplication.

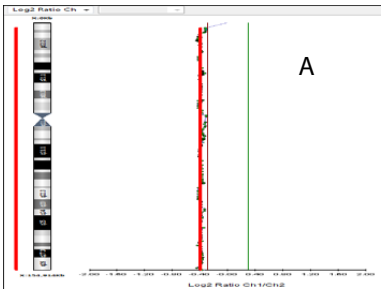
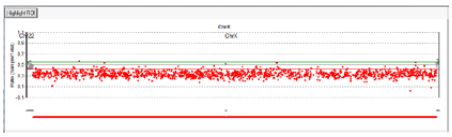
One "sample" project and one "control" project are loaded into the CNV menu. The regions are identified- either by annotation, incremental length, or a BED file. A BED file specifying amplicon locations is created for targeted sequencing projects, and exon locations are useful for whole-exome sequencing. For automatic fitting, the raw data is grouped to generate "fitting points" describing the dispersion at a given level of coverage.

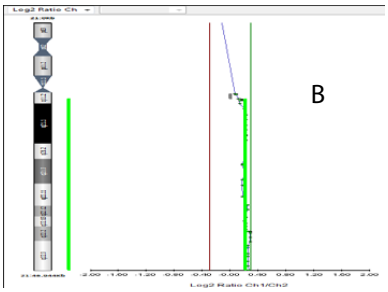
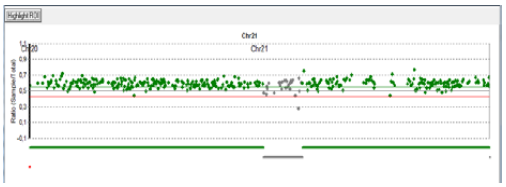
A line is fit to these points and used to calculate the dispersion value for each region. The number of fitting points is automatically set based on the number of regions but it may be set manually instead. As a rule of thumb, there should be at least 4 to 5 fitting points and at least 100 raw data points per fitting point.

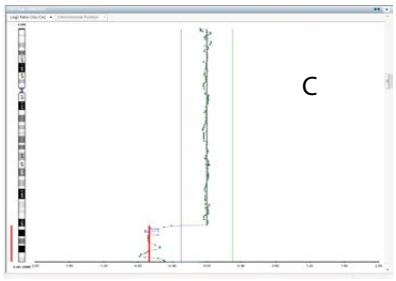
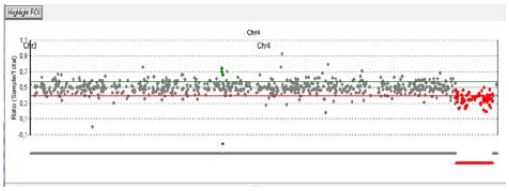
The goal of fitting the equation is to measure the amount of dispersion (noise) present in "normal" regions. The coverage ratio is expected to be equal to 0.5 for regions in the absence of a CNV. There is some randomness expected for this value, with higher-coverage regions showing a tighter distribution around the expected value than lower-coverage regions. The software first splits the data up into groups based on the total coverage, generating a summary "fitting point" for each group based on measured dispersion and the median coverage. A line is fit to these "fitting points" and the equation for this line is used to calculate dispersion for every individual region.

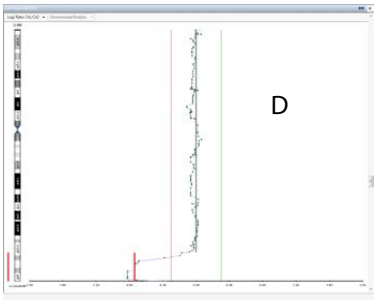
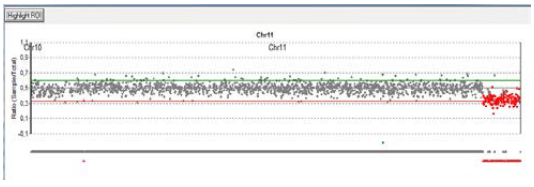
The dispersion value is used to calculate parameters for a beta distribution, which is used to generate a confidence interval. A higher dispersion value gives a broader CI because the ratios are expected to be more widely dispersed. If the expected CNV frequency is 10%, the software will calculate fitting points by incrementing the dispersion value until it produces an appropriate 90% (equal to 100%-10%) Confidence Interval (CI) of ratios. An appropriate confidence interval is one where the lower half of the CI is lower than the 5th percentile ratio of the real data (because Duplication=5% and Deletion=5% in this case), or the upper half of the confidence interval is greater than the 95th percentile. This one-sided fitting allows the software to be tolerant of CNVs that cause the raw data to have an asymmetrical distribution.

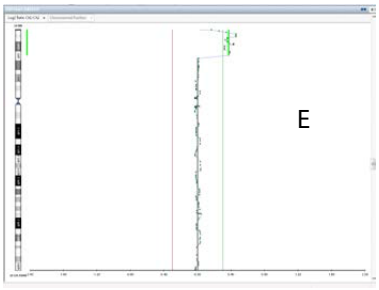
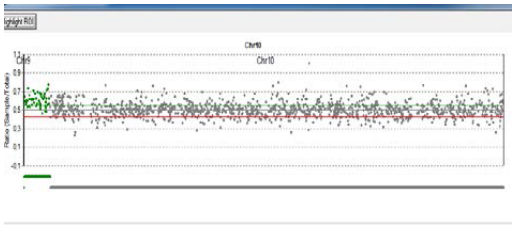
Dispersion values calculated for each region are used to generate normalized (probability of Normal+Duplication+Deletion=1) beta-binomial distributions (Figure 1e). When dispersion in a given region is high, the likelihood for any one call is low except for extreme ratio values (close to 0.0 or 1.0). The HMM used to make CNV calls makes some assumptions. The initial likelihood of each state is related to the expected CNV frequency, as is the probability of transitioning from a "normal" region to a region with a CNV. Once a region is called as a CNV, the next region is assumed to have a 50% chance of continuing that CNV or going back to normal. This transition probability enables the HMM to both ignore possibly erroneous ratios from single regions and also identify long CNVs where no individual region in the call has a very high probability.

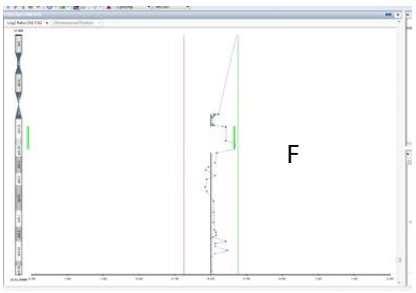
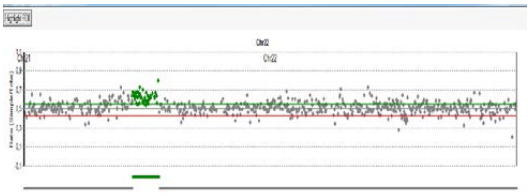
Results of aCGH Sample C1	Results of NGS Sample C1
	
Turner syndrome	Turner syndrome

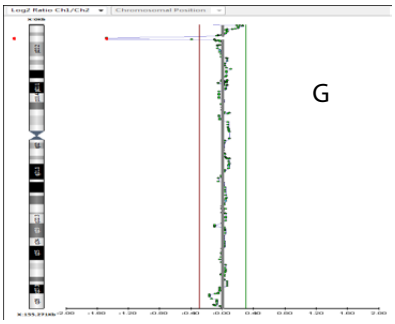
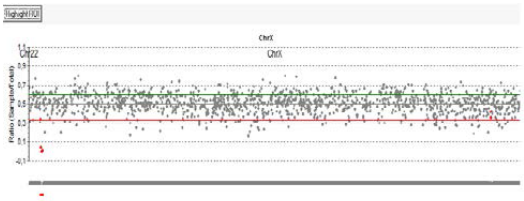
Results of aCGH Sample C2	Results of NGS Sample C2
	
Down syndrome (Trisomy 21)	Down syndrome (Trisomy 21)

Results of aCGH Sample C3	Results of NGS Sample C3
	
arr 4q32.1q35.2 (160,764,741x2, 161,374,901-190,815,481x1)	Region deletion from 162,246,448 bp to 188,455,721 bp (red line)

Results of aCGH Sample C4	Results of NGS Sample C4
	
arr 11q23.3q25 (121,307,099x2, 119,774,967-134,852,671x1)	Region deletion from 121,415,942 bp to 134,131,794 bp (red line)

Results of aCGH Sample C5	Results of NGS Sample C5
	
arr 10p15.3p13 (142,203-14,378,024x3, 16,078,362x2)	Region duplication from 255,819 bp to 13,536,606 bp (green line)

Results of aCGH Sample C6	Results of NGS Sample C6
	
arr 22q11.21q11.23 (19,542,207x2, 19,542,281-24,319,952x3, 25,254,126x2)	Region duplication from 19,753,415 bp to 24,237,141 bp (green line)

	
arr Xp22.31(6,223,645x1, 7,239,742-8,153,286x0, 8,227,525x1)	Region deletion from 6,451,779 bp to 7,268,312 bp (red line)

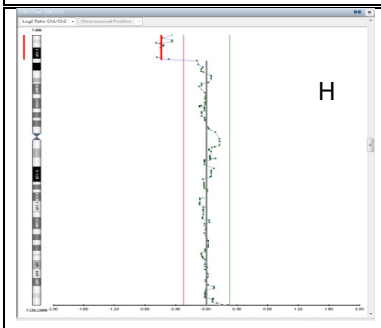
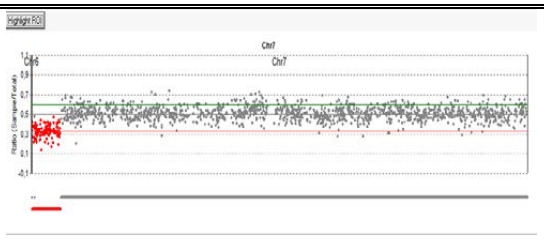
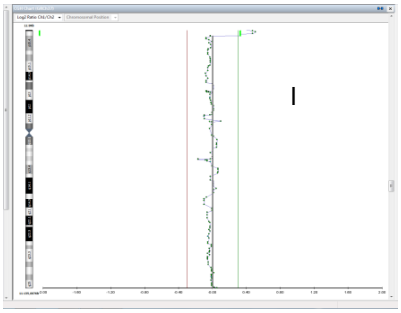
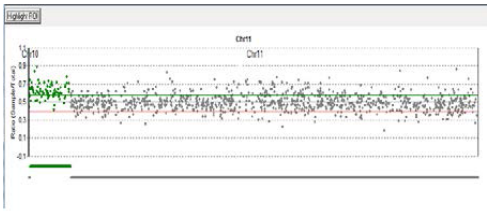
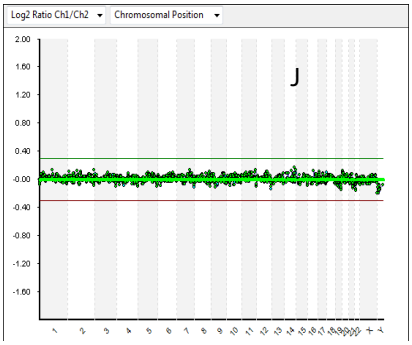
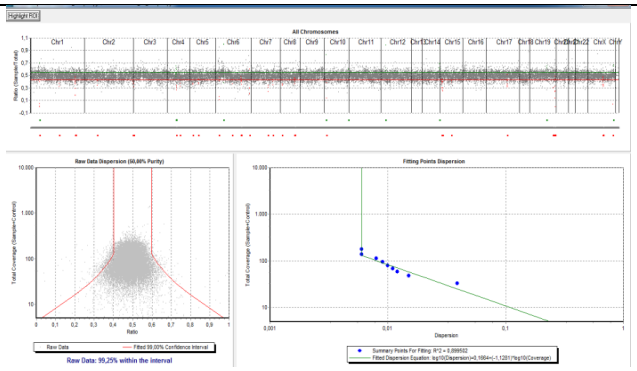
	
arr 7p22.3p21.2 (14,916-14,227,858x1, 14,993,340x2)	Region deletion from 295,805 bp to 11,871,582 bp (red line)
Results of aCGH Sample C9	Results of NGS Sample C9
	
arr 11p15.5 (232,848-2,763,614x3, 2,968,091x2)	Region duplication from 236,038 bp to 2,482,954 bp (green line)
Results of aCGH Sample CNEG	Results of NGS Sample CNEG
	

Figure 1a-1j: Of the 248 samples analyzed using both aCGH (CytoChip Focus Constitutional BAC-array platform) and NGS data through NextGENe Software (Softgenetics), we identified nine samples affected by aneuploidies and chromosomal micro deletions/micro duplications.

Phred scores are also calculated using these likelihoods. They are capped at 80, equivalent to a 99.999999% probability. Phred scores are much lower if the dispersion is high, because there is less certainty about the classifications (Figure 1f). Generally deletion calls can be more confident than duplication calls because the expected heterozygous ratio (0.333) is farther away from the normal ratio (0.5) than the heterozygous duplication ratio (0.6).

Results

We analyzed samples obtained from amniocentesis or chorionic villi sampling. Here we show results obtained through aCGH and NGS for chromosomal analysis.

Of the 248 samples analyzed using both aCGH (CytoChip Focus Constitutional BAC-array platform) and NGS data through NextGENe Software (Softgenetics), we identified nine samples affected by aneuploidies and chromosomal micro deletions/micro duplications. In Table 1 and Figures 1a-1j, we showed results obtained for positive samples and an example for negative sample.

As shown above, for nine positive samples, results were overlapping between aCGH and NGS. Remarkable for samples C6, C7 and C9 the extension of CNVs was identical for both the techniques used, whereas for samples C3, C4, C5 and C8, CNVs extension was found to be less than what was revealed using aCGH, with a difference ranging from 1 Mb (sample C5) and 3 Mb (sample C3). The size difference is probably associated to using an exome-like enrichment for NGS library preparation, carrying out for analysis only coding regions and intron regions flanking exons.

Furthermore, NGS approach was able to identify chromosomal mosaicism on sample C2 (trisomy 21 present in 40% of the metaphases analyzed) and C6 (duplication of region 22q11.2 present in 45% of the metaphases analyzed).

Nextgene Software allowed to identify chromosomal location, genes involved in chromosomal aberration, length and values obtained for deletion or duplication relating to reference comparison used during sample processing (Table 2).

For each case, we performed an analysis for missense mutation, frame shift, splicing, stop codon gained/lost, in frame insertion/in frame deletion through Variant Studio Software.

Discussion

In recent years Next Generation Sequencing (NGS) has become an important tool not only for gene discovery and research area but also for clinical diagnosis. To date, Next Generation Sequencing has been predominantly used for SNPs/Indel diagnosis and only in a few cases for detection of chromosomal aneuploidy [5-7]; however for chromosomal screening were used whole-genomic approach, not compatible to prenatal diagnosis for timing and sample type. Here we showed NGS application in prenatal diagnosis both for SNPs analysis and chromosomal screening.

We used a Next Generation Sequencing method based on the use of an enrichment gene panel library produced by Illumina and including 4,813 genes. After sequencing our bioinformatic pipelines allows SNPs and structural/numerical chromosomal anomalies analysis. For SNPs analysis we selected a genes pool (about 152 genes) associated approximately to thirty diseases or diseases family involved in fetus development, targeted exome-like approach. Using a producer validated

Sample	aCGH	NGS	Consequence
C1	45, X0	45,X0	Turner syndrome
C2	47, XX, +21	47, XX, +21	Down syndrome (mosaicism 40%)
C3	arr 4q32.1q35.2 (161,374,901-190,815,481 x1) 29 Mb deletion	4q32.1q35.2 (162,246,448- 188,455,721 x1) 26 Mb deletion	4q- syndrome
C4	arr 11q23.3q25 (119,774,967-134,852,671 x1) 15 Mb deletion	11q23.3q25 (121,415,942 - 134,131,794 x1) 13Mb deletion	Jacobsen syndrome
C5	arr 10p15.3p13 (142,203-14,378,024 x3) 14,2 Mb duplication	10p15.3p13 (255,819 - 13,536,606 x3) 13,2 Mb duplication	DiGeorge syndrome/velocardiofacial syndrome complex 2 Hypoparathyroidism, sensorineural deafness, and renal disease
C6	arr 22q11.21q11.23 (19,542,281-24,319,952 x3) 4,7 Mb Interstitial mosaicism duplication	22q11.21q11.23 (19,753,415 - 24,237,141 x3) 4,5 Mb Interstitial duplication	Duplication syndrome 22q11.2 (mosaicism through FISH: 40%)
C7	arr Xp22.31 (7,239,742-8,153,286 x0) 900 Kb Interstitial mosaicism deletion	Xp22.31 (6,451,779 bp -7,268,312 x0) 817 Kb Interstitial mosaicism deletion	X-linked ichthyosis
C8	arr 7p22.3p21.2 (14,916-14,227,858 x1) Terminal deletion of 14 Mb on the short arm of chr 7	7p22.3p21.2 (295,805 - 11,871,582 x1) Terminal deletion of 11,6Mb on the short arm of chr 7.	Partial monosomy 7p
C9	arr 11p15.5 (232,848-2,763,614 x1) Terminal deletion of 2,5 Mb on the short arm of chr 11	11p15.5 (da 236,038 bp a 2,482,954 bp x1) Terminal deletion of 2,3 Mb on the short arm of chr 11	Developmental delay/ Intellectual disability/ ASD
CNEG	Normal	Normal	_____

Table 1: We analyzed samples obtained from amniocentesis or chorionic villi sampling. Here we show results obtained through aCGH and NGS for chromosomal analysis.

C8 7p22.3p21.2 (295,805 - 11,871,582 x1) Terminal deletion of 11,6Mb on the short arm of chr 7.									
Description	Chr	Chr End	Length	Ratio	Total RPKM	Dispersion	Normalized Likelihoods	Deletion Score	HMM Calls
FAM20C.chr7.193199.193804	chr7	193814	625	0.299	47.88	0.0251	-0,06;-0,91;-1,82	8.63	Deletion
FAM20C.chr7.195553.195732	chr7	195742	199	0.395	38.985	0.0295	-0,24;-0,52;-0,94	3.78	Deletion
FAM20C.chr7.295814.295995	chr7	296005	201	0.212	39.928	0.0289	-0,03;-1,24;-2,40	12.11	Deletion
FAM20C.chr7.299696.299946	chr7	299956	270	0.277	23.236	0.0443	-0,14;-0,68;-1,24	5.72	Deletion
HEATR2.chr7.766358.766952	chr7	766962	615	0.329	11.366	0.0778	-0,30;-0,50;-0,74	3.04	Deletion
HEATR2.chr7.769300.769484	chr7	769494	205	0.378	24.524	0.0425	-0,26;-0,51;-0,84	3.41	Deletion
HEATR2.chr7.794226.794458	chr7	794468	253	0.225	45.575	0.0261	-0,02;-1,28;-2,50	12.56	Deletion
HEATR2.chr7.796419.796631	chr7	796641	233	0.368	67.714	0.0191	-0,10;-0,73;-1,54	6.7	Deletion
HEATR2.chr7.801390.801533	chr7	801543	164	0.292	64.744	0.0198	-0,03;-1,17;-2,36	11.4	Deletion
HEATR2.chr7.803443.803611	chr7	803621	189	0.402	28.977	0.0372	-0,29;-0,49;-0,80	3.15	Deletion
HEATR2.chr7.810108.810255	chr7	810265	168	0.309	65.434	0.0196	-0,04;-1,06;-2,16	10.26	Deletion
HEATR2.chr7.813685.813835	chr7	813845	171	0.265	93.528	0.0148	-0,01;-1,81;-3,63	18.06	Deletion
HEATR2.chr7.814643.814799	chr7	814809	177	0.37	120.165	0.0121	-0,05;-1,02;-2,28	9.95	Deletion
HEATR2.chr7.819590.819781	chr7	819791	212	0.294	92.001	0.015	-0,01;-1,50;-3,06	14.88	Deletion
HEATR2.chr7.825154.825287	chr7	825297	154	0.357	97.7	0.0143	-0,05;-0,99;-2,16	9.66	Deletion
CYP2W1.chr7.1022848.1023021	chr7	1023031	194	0.529	32.548	0.034	-0,65;-0,41;-0,41	1.11	Deletion
CYP2W1.chr7.1024048.1024210	chr7	1024220	183	0.326	30.765	0.0355	-0,15;-0,64;-1,19	5.3	Deletion
CYP2W1.chr7.1024586.1024735	chr7	1024745	170	0.24	25.341	0.0414	-0,09;-0,81;-1,52	7.31	Deletion
CYP2W1.chr7.1024802.1024959	chr7	1024969	178	0.334	116.745	0.0124	-0,02;-1,36;-2,90	13.49	Deletion
CYP2W1.chr7.1026260.1026433	chr7	1026443	194	0.241	77.808	0.0171	-0,01;-1,79;-3,54	17.84	Deletion
CYP2W1.chr7.1026743.1026881	chr7	1026891	159	0.241	27.772	0.0385	-0,08;-0,85;-1,61	7.78	Deletion
CYP2W1.chr7.1026983.1027167	chr7	1027177	205	0.302	29.115	0.0371	-0,13;-0,68;-1,28	5.87	Deletion
CYP2W1.chr7.1027913.1028054	chr7	1028064	162	0.323	30.556	0.0357	-0,15;-0,64;-1,20	5.35	Deletion
CYP2W1.chr7.1028271.1028455	chr7	1028465	205	0.193	65.681	0.0195	-0,00;-2,00;-3,86	19.93	Deletion
MAD1L1.chr7.1855709.1855864	chr7	1855874	176	0.359	55.422	0.0223	-0,12;-0,70;-1,43	6.28	Deletion
MAD1L1.chr7.1937836.1938026	chr7	1938036	211	0.321	100.321	0.014	-0,02;-1,33;-2,79	13.18	Deletion
MAD1L1.chr7.1976323.1976533	chr7	1976543	231	0.335	46.524	0.0256	-0,11;-0,74;-1,47	6.64	Deletion
MAD1L1.chr7.2054137.2054277	chr7	2054287	161	0.316	43.257	0.0272	-0,09;-0,78;-1,55	7.16	Deletion
MAD1L1.chr7.2108829.2108973	chr7	2108983	165	0.304	49.89	0.0243	-0,06;-0,91;-1,83	8.62	Deletion
MAD1L1.chr7.2255792.2255922	chr7	2255932	151	0.283	50.826	0.0239	-0,05;-1,04;-2,07	9.99	Deletion
MAD1L1.chr7.2262210.2262389	chr7	2262399	200	0.245	66.571	0.0193	-0,01;-1,55;-3,06	15.34	Deletion
MAD1L1.chr7.2265045.2265185	chr7	2265195	161	0.313	41.711	0.0279	-0,09;-0,78;-1,54	7.11	Deletion
MAD1L1.chr7.2269619.2269768	chr7	2269778	170	0.337	40.243	0.0287	-0,13;-0,68;-1,33	5.95	Deletion
NUDT1.chr7.2284197.2284361	chr7	2284371	184	0.317	33.75	0.033	-0,13;-0,69;-1,31	5.96	Deletion
NUDT1.chr7.2289491.2289637	chr7	2289647	166	0.247	42.275	0.0277	-0,04;-1,09;-2,13	10.55	Deletion
NUDT1.chr7.2290463.2290636	chr7	2290646	193	0.346	38.932	0.0295	-0,15;-0,64;-1,24	5.44	Deletion
LFNG.chr7.2552790.2552962	chr7	2552972	192	0.271	79.29	0.0168	-0,01;-1,53;-3,08	15.19	Deletion
LFNG.chr7.2559495.2559927	chr7	2559937	452	0.324	28.011	0.0382	-0,16;-0,62;-1,13	5.04	Deletion
LFNG.chr7.2565047.2565201	chr7	2565211	174	0.339	22.638	0.0452	-0,22;-0,55;-0,95	4.08	Deletion
LFNG.chr7.2565877.2566043	chr7	2566053	186	0.291	58.645	0.0214	-0,04;-1,10;-2,21	10.65	Deletion
BRAT1.chr7.2577706.2578398	chr7	2578408	713	0.309	74.973	0.0176	-0,03;-1,17;-2,40	11.41	Deletion
BRAT1.chr7.2578813.2578985	chr7	2578995	193	0.311	95.662	0.0145	-0,02;-1,39;-2,87	13.71	Deletion
BRAT1.chr7.2580932.2581118	chr7	2581128	207	0.231	43.941	0.0268	-0,03;-1,22;-2,37	11.86	Deletion
BRAT1.chr7.2583224.2583596	chr7	2583606	393	0.327	66.693	0.0193	-0,06;-0,96;-1,98	9.2	Deletion
BRAT1.chr7.2584543.2584690	chr7	2584700	168	0.183	5.421	0.1395	-0,27;-0,53;-0,77	3.3	Deletion
BRAT1.chr7.2586958.2587112	chr7	2587122	175	0.263	64.587	0.0198	-0,02;-1,38;-2,74	13.59	Deletion
AP5Z1.chr7.4820805.4820943	chr7	4820953	158	0.345	102.144	0.0138	-0,04;-1,13;-2,43	11.08	Deletion
AP5Z1.chr7.4821198.4821385	chr7	4821395	207	0.285	97.098	0.0144	-0,01;-1,65;-3,35	16.42	Deletion
AP5Z1.chr7.4822946.4823091	chr7	4823101	165	0.391	56.973	0.0219	-0,17;-0,58;-1,19	4.87	Deletion
AP5Z1.chr7.4823833.4824002	chr7	4824012	189	0.257	52.96	0.0232	-0,03;-1,23;-2,42	12	Deletion
AP5Z1.chr7.4824538.4824679	chr7	4824689	161	0.296	32.179	0.0343	-0,11;-0,74;-1,40	6.5	Deletion

AP5Z1.chr7.4825152.4825315	chr7	4825325	183	0.309	95.407	0.0146	-0,02;-1,39;-2,89	13.81	Deletion
AP5Z1.chr7.4825880.4826059	chr7	4826069	199	0.311	147.104	0.0104	-0,01;-1,93;-4,02	19.25	Deletion
AP5Z1.chr7.4827264.4827407	chr7	4827417	163	0.257	54.088	0.0228	-0,03;-1,24;-2,46	12.16	Deletion
AP5Z1.chr7.4827784.4827925	chr7	4827935	161	0.285	108.52	0.0132	-0,01;-1,81;-3,68	18.06	Deletion
AP5Z1.chr7.4830089.4830222	chr7	4830232	153	0.313	123.222	0.0119	-0,01;-1,65;-3,45	16.48	Deletion
AP5Z1.chr7.4830303.4830518	chr7	4830528	235	0.273	90.102	0.0152	-0,01;-1,68;-3,39	16.75	Deletion
AP5Z1.chr7.4830745.4831016	chr7	4831026	291	0.241	90.306	0.0152	-0,00;-2,01;-3,98	20.07	Deletion
SLC29A4.chr7.5327448.5327616	chr7	5327626	189	0.275	42.929	0.0273	-0,06;-0,96;-1,89	9.12	Deletion
SLC29A4.chr7.5330363.5330494	chr7	5330504	152	0.333	46.207	0.0258	-0,10;-0,75;-1,48	6.73	Deletion
SLC29A4.chr7.5336567.5336829	chr7	5336839	283	0.287	63.759	0.02	-0,03;-1,19;-2,40	11.66	Deletion
SLC29A4.chr7.5338619.5338757	chr7	5338767	159	0.301	50.885	0.0239	-0,06;-0,94;-1,88	8.92	Deletion
SLC29A4.chr7.5338871.5339058	chr7	5339068	208	0.287	40.167	0.0288	-0,07;-0,87;-1,70	8.08	Deletion
SLC29A4.chr7.5340053.5340293	chr7	5340303	261	0.288	41.665	0.028	-0,07;-0,88;-1,74	8.26	Deletion
SLC29A4.chr7.5342428.5342567	chr7	5342577	160	0.306	77.719	0.0171	-0,03;-1,23;-2,52	12.04	Deletion
ACTB.chr7.5567378.5567522	chr7	5567532	164	0.291	102.952	0.0137	-0,01;-1,67;-3,42	16.67	Deletion
ACTB.chr7.5567634.5567816	chr7	5567826	202	0.306	122.052	0.012	-0,01;-1,73;-3,58	17.23	Deletion
ACTB.chr7.5567911.5568350	chr7	5568360	459	0.318	83.434	0.0162	-0,03;-1,19;-2,48	11.7	Deletion
ACTB.chr7.5568791.5569031	chr7	5569041	260	0.347	83.667	0.0162	-0,05;-0,97;-2,07	9.4	Deletion
PMS2.chr7.6012869.6013173	chr7	6013183	324	0.451	143.568	0.0106	-0,27;-0,41;-1,11	3.35	Deletion
PMS2.chr7.6017218.6017388	chr7	6017398	190	0.415	97.186	0.0144	-0,16;-0,58;-1,33	5.05	Deletion
PMS2.chr7.6022454.6022622	chr7	6022632	188	0.381	139.536	0.0108	-0,05;-1,00;-2,32	9.84	Deletion
PMS2.chr7.6026389.6027251	chr7	6027261	882	0.363	95.972	0.0145	-0,06;-0,94;-2,05	9.05	Deletion
PMS2.chr7.6029430.6029586	chr7	6029596	176	0.291	20.863	0.0482	-0,17;-0,62;-1,10	4.96	Deletion
PMS2.chr7.6038738.6038906	chr7	6038916	188	0.309	36.689	0.0309	-0,11;-0,74;-1,43	6.61	Deletion
PMS2.chr7.6042083.6042267	chr7	6042277	204	0.483	64.972	0.0197	-0,50;-0,37;-0,59	1.66	Deletion
PMS2.chr7.6045522.6045662	chr7	6045672	160	0.383	66.371	0.0194	-0,13;-0,65;-1,37	5.74	Deletion
RAC1.chr7.6441499.6441658	chr7	6441668	180	0.256	23.891	0.0434	-0,11;-0,74;-1,37	6.46	Deletion
C1GALT1.chr7.7273951.7274170	chr7	7274180	240	0.477	38.51	0.0298	-0,47;-0,41;-0,57	1.82	Deletion
C1GALT1.chr7.7277886.7278553	chr7	7278563	688	0.393	42.583	0.0275	-0,22;-0,53;-1,00	4.05	Deletion
C1GALT1.chr7.7283155.7283355	chr7	7283365	221	0.419	16.223	0.0588	-0,38;-0,46;-0,63	2.38	Deletion
GLCCI1.chr7.8008982.8009438	chr7	8009448	477	0.316	28.523	0.0377	-0,15;-0,64;-1,19	5.33	Deletion
GLCCI1.chr7.8043538.8043689	chr7	8043699	172	0.304	119.336	0.0122	-0,01;-1,72;-3,56	17.14	Deletion
GLCCI1.chr7.8099726.8099878	chr7	8099888	173	0.351	110.599	0.013	-0,03;-1,14;-2,48	11.25	Deletion
GLCCI1.chr7.8110551.8110761	chr7	8110771	231	0.26	40.499	0.0286	-0,05;-1,00;-1,94	9.49	Deletion
GLCCI1.chr7.8125823.8126165	chr7	8126175	363	0.307	62.964	0.0202	-0,04;-1,05;-2,14	10.16	Deletion
THSD7A.chr7.11418697.11418907	chr7	11418917	231	0.384	128.649	0.0115	-0,06;-0,93;-2,14	9.02	Deletion
THSD7A.chr7.11441422.11441595	chr7	11441605	194	0.347	120.426	0.0121	-0,03;-1,25;-2,72	12.39	Deletion
THSD7A.chr7.11445927.11446101	chr7	11446111	195	0.304	136.592	0.011	-0,01;-1,91;-3,95	19.03	Deletion
THSD7A.chr7.11446537.11446682	chr7	11446692	166	0.422	64.347	0.0199	-0,24;-0,49;-1,00	3.77	Deletion
THSD7A.chr7.11452283.11452427	chr7	11452437	165	0.379	87.083	0.0157	-0,09;-0,77;-1,69	7.23	Deletion
THSD7A.chr7.11457077.11457230	chr7	11457240	174	0.283	83.453	0.0162	-0,01;-1,49;-3,03	14.82	Deletion
THSD7A.chr7.11464323.11464459	chr7	11464469	157	0.255	50.876	0.0239	-0,03;-1,20;-2,37	11.74	Deletion
THSD7A.chr7.11468571.11468752	chr7	11468762	202	0.315	76.957	0.0173	-0,03;-1,14;-2,36	11.18	Deletion
THSD7A.chr7.11485688.11485951	chr7	11485961	284	0.366	63.235	0.0201	-0,11;-0,72;-1,49	6.5	Deletion
THSD7A.chr7.11486857.11487051	chr7	11487061	215	0.328	87.045	0.0157	-0,03;-1,15;-2,40	11.23	Deletion
THSD7A.chr7.11501638.11501770	chr7	11501780	153	0.32	98.736	0.0142	-0,02;-1,33;-2,78	13.14	Deletion
THSD7A.chr7.11513961.11514195	chr7	11514205	255	0.253	68.755	0.0189	-0,01;-1,52;-3,02	15.1	Deletion
THSD7A.chr7.11521415.11521609	chr7	11521619	215	0.355	58.522	0.0214	-0,10;-0,74;-1,52	6.73	Deletion
THSD7A.chr7.11581046.11581258	chr7	11581268	233	0.343	103.076	0.0137	-0,03;-1,16;-2,48	11.35	Deletion
THSD7A.chr7.11582589.11582744	chr7	11582754	176	0.369	116.605	0.0124	-0,05;-1,01;-2,25	9.84	Deletion
THSD7A.chr7.11630087.11630268	chr7	11630278	202	0.345	148.991	0.0103	-0,01;-1,50;-3,26	14.92	Deletion
THSD7A.chr7.11632881.11633129	chr7	11633139	269	0.345	124.121	0.0118	-0,02;-1,31;-2,82	12.94	Deletion
THSD7A.chr7.11675757.11676588	chr7	11676598	852	0.336	97.496	0.0143	-0,03;-1,17;-2,49	11.51	Deletion
THSD7A.chr7.11871383.11871572	chr7	11871582	210	0.301	25.856	0.0407	-0,15;-0,65;-1,20	5.43	Deletion

Table 2: Nextgene Software allowed identifying chromosomal location, genes involved in chromosomal aberration, length and values obtained for deletion or duplication relating to reference comparison used during sample processing.

kit allowed us to avoid the development and validation of library for each gene of interest, obtaining 20X as minimum target coverage value. The advantage of this approach was robustness of experimental design and results obtained reproducibility and speed of execution.

We used dedicated software (Nextgene, Sofgenetetics) to carry out copy number variation analysis of data obtained from NGS. This evaluation was performed through comparison to aCGH, the gold standard technique for the identification of chromosome aneuploidies, microdeletions and microduplications.

The results are comparable with those obtained from aCGH both for chromosomal aneuploidy that for CNVs extension between 10 Mb and less than 1 Mb. This system has a number of advantages compared to the use of microarray. Using a single analytical tool for Mendelian disorders and chromosomal abnormalities screening, makes NGS compatible to prenatal diagnosis. Moreover, with similar resolution level to aCGH, it is possible to obtain clear clinical effects of chromosome anomalies, considering not only chromosome position and size of microdeletion/microduplication but also sequencing analysis of same locus on the other allele. This makes it possible to exclude possible pathogenetic SNPs that cannot be identified through aCGH.

In the future, we will use an enrichment panel similar to the one used in this study, but it will include 19,000 genes and we will compare the results obtained with a higher resolution array CGH platform.

References

1. Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, et al. (2013) Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med 369: 1502-1511.
2. Chen S, Li S, Xie W, Li X, Zhang C, et al. (2014) Performance comparison between rapid sequencing platforms for ultra-low coverage sequencing strategy. PLoS One 9: e92192.
3. Dan S, Chen F, Choy KW, Jiang F, Lin J, et al. (2012) Prenatal detection of aneuploidy and imbalanced chromosomal arrangements by massively parallel sequencing. PLoS One 7: e27835.
4. Talkowski ME, Ordulu Z, Pillalamarri V, Benson CB, Blumenthal I, et al. (2012) Clinical diagnosis by whole-genome sequencing of a prenatal sample. N Engl J Med 367: 2226-2232.
5. Zhang C, Zhang C, Chen S, Yin X, Pan X, et al. (2013) A single cell level based method for copy number variation analysis by low coverage massively parallel sequencing. PLoS One 8: e54236.
6. Hayes JL, Tzika A, Thygesen H, Berri S, Wood HM, et al. (2013) Diagnosis of copy number variation by Illumina next generation sequencing is comparable in performance to oligonucleotide array comparative genomic hybridisation. Genomics 102: 174-181.
7. Carss KJ, Hillman SC, Parthiban V, McMullan DJ, Maher ER, et al. (2014) Exome sequencing improves genetic diagnosis of structural fetal abnormalities revealed by ultrasound. Hum Mol Genet 23: 3269-3277.