The genetic basis of undiagnosed muscular dystrophies and myopathies

Results from 504 patients

Marco Savarese, PhD
Giuseppina Di Fruscio, PhD
Annalaura Torella, PhD
Chiara Fiorillo, MD
Francesca Magri, MD
Marina Fanin, PhD
Lucia Ruggiero, MD
Giulia Ricci, MD
Guja Astrea, PhD
Luigia Passamano, MD
Alessandra Ruggieri, PhD
Dario Ronchi, PhD
Giorgio Tasca, MD
Adele D'Amico, MD
Sandra Janssens, MD
Olimpia Farina, MD
Margherita Mutarelli, PhD
Veer Singh Marwah, MSc
Arcomaria Garofalo, MSc
Teresa Giugliano, PhD
Simone Sanpaolo, MD
Francesca Del Vecchio Blanco, PhD
Gaia Esposito, BSc
Giulio Piluso, PhD
Paola D'Ambrosio, MD
Roberta Petillo, MD
Olimpia Musumeci, MD
Carmelo Rodolico, MD
Sonia Messina, MD
Anni Evilä, PhD
Peter Hackman, PhD
Massimiliano Filosto, MD
Giuseppe Di Iorio, MD
Gabriele Siciliano, MD
Marina Mora, PhD
Lorenzo Maggi, MD
Carlo Minetti, MD
Sabrina Sacconi, MD
Lucio Santoro, MD
Kathleen Claes, PhD
Liliana Vercelli, MD
Tiziana Mongini, MD

Author list continued on next page

ABSTRACT

Objective: To apply next-generation sequencing (NGS) for the investigation of the genetic basis of undiagnosed muscular dystrophies and myopathies in a very large cohort of patients.

Methods: We applied an NGS-based platform named MotorPlex to our diagnostic workflow to test muscle disease genes with a high sensitivity and specificity for small DNA variants. We analyzed 504 undiagnosed patients mostly referred as being affected by limb-girdle muscular dystrophy or congenital myopathy.

Results: MotorPlex provided a complete molecular diagnosis in 218 cases (43.3%). A further 160 patients (31.7%) showed as yet unproven candidate variants. Pathogenic variants were found in 47 of 93 genes, and in more than 30% of cases, the phenotype was nonconventional, broadening the spectrum of disease presentation in at least 10 genes.

Conclusions: Our large DNA study of patients with undiagnosed myopathy is an example of the ongoing revolution in molecular diagnostics, highlighting the advantages in using NGS as a first-tier approach for heterogeneous genetic conditions. *Neurology*® 2016;87:71–76

GLOSSARY

CM = congenital myopathy; **LGMD** = limb-girdle muscular dystrophy; **MD** = muscular dystrophy; **NGS** = next-generation sequencing; **NMD** = neuromuscular disorder.

Muscular dystrophies (MDs)[1,2] and congenital myopathies (CMs)[3] represent the majority of inherited neuromuscular disorders (NMDs).[4,5]

Until the development of next-generation sequencing (NGS), routine molecular diagnosis of NMDs had been based on a gene-by-gene approach.[1,6–8] This time-consuming and expensive approach[9] failed to identify causative variants in more than 40% of cases, detection rates varied with different genes or conditions, and the size of some large genes, such as *TTN* or *NEB*, hampered routine analysis.

Targeted NGS, which focuses only on specific genes of interest,[10] has recently been proposed as a cost-effective strategy for the molecular diagnosis of heterogeneous disorders.[11]

Enzo Ricci, MD
Francesca Gualandi, MD
Rossella Tupler, MD
Jan De Bleecker, MD
Bjarne Udd, MD
Antonio Toscano, MD
Maurizio Moggio, MD
Elena Pegoraro, MD
Enrico Bertini, MD
Eugenio Mercuri, MD
Corrado Angelini, MD
Filippo Maria Santorelli,
    MD
Luisa Politano, MD
Claudio Bruno, MD
Giacomo Pietro Comi,
    MD
Vincenzo Nigro, MD

Correspondence to
Dr. Nigro:
vinnigro@gmail.com

Although different targeted approaches analyzing genes causing NMDs have been described in literature,[12–14] screening for mutations in all known muscular disease genes in a large NMD patient cohort is rare.

Herein, we describe the results obtained in an extensive study of 504 genetically undiagnosed patients presenting clinical signs of MD, CM, or other conditions affecting the muscles. We report the different diagnostic rate for each clinical condition studied and discuss the advantages and disadvantages of our strategy, illustrating some unexpected results as well as its main limitations.

METHODS **Patients and study design.** With the help of the Italian Network of Congenital Myopathies and the Italian Network of Limb-Girdle Muscular Dystrophy (LGMD), we collected DNA samples from patients with a clinical diagnosis

Figure 1   Characteristics of recruited patients
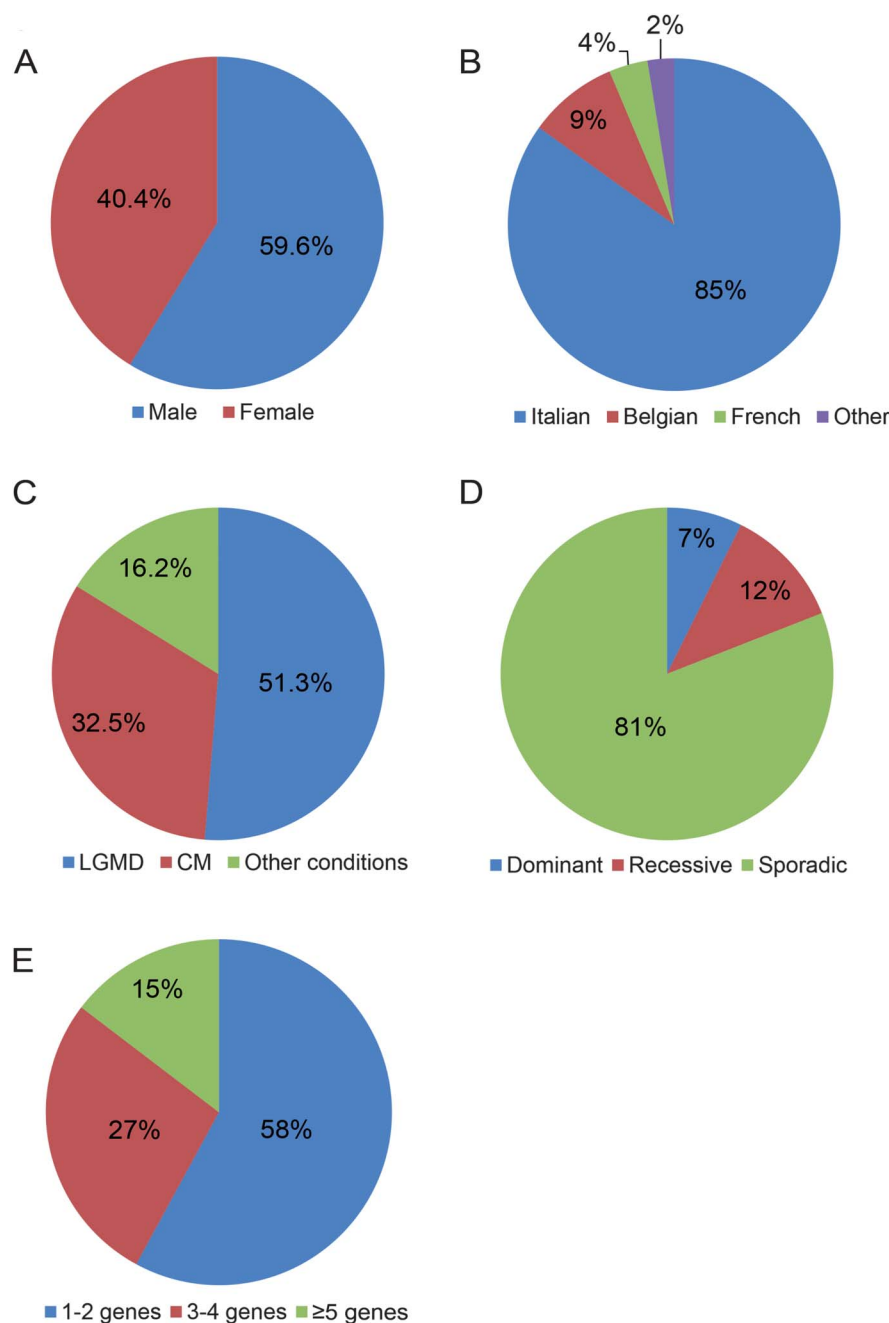
(A) Percentage of male and female patients. (B) Geographical origin of patients. (C) Percentage of patients in 3 different categories (LGMD, CM, other conditions). (D) Mode of inheritance for the studied conditions. (E) Previous unsuccessful molecular tests performed in prescreened patients. CM = congenital myopathy; LGMD = limb-girdle muscular dystrophy.

of a nonspecific myopathy, congenital myopathy, proximal muscle weakness, or LGMD. Genes that are universally considered as genetic causes of nonsyndromic myopathies (table e-1 on the *Neurology*® Web site at Neurology.org) were investigated by using a custom NGS panel, named MotorPlex.[15]

For each patient, we collected clinical and laboratory data as well as the results of familial segregation analyses and previous genetic tests.

We included in this study a total of 588 samples from 504 patients and 84 unaffected relatives and analyzed them using 4 different strategies according to genetic transmission and availability of further samples from relatives. Specifically, in 400 cases (mainly sporadic), only the proband's DNA samples were analyzed; in 50 sporadic cases, trio (n = 31) or duo (n = 19) analysis was performed. In 54 familial cases, additional samples from other affected and unaffected family members were analyzed.

Standard protocol approvals, registrations, and patient consents. All the patients and families provided written informed consent to the different clinical centers involved in the Project GUP11006 following the guidelines of Telethon Foundation, Italy, according to the Declaration of Helsinki. The Ethics Committee of The Second University of Naples approved the NGS study protocol.

NGS workflow and sequencing analysis. We prepared a library according to the manufacturer's instructions (HaloPlex Target Enrichment System for Illumina Sequencing; Agilent Technologies, Santa Clara, CA). Bioanalyzer High Sensitivity DNA Assay kit (Agilent Technologies) was used to validate and quantify library preparation. Twenty individual samples were run in a single lane of a HiSeq1000 system (Illumina Inc., San Diego, CA), generating 100–base pair paired end reads. Analysis of NGS data was performed using an in-house pipeline described elsewhere.[15,16]

RESULTS Patients under study. We collected and sequenced 588 samples from 504 patients and 84 unaffected relatives. Specifically, 85% of the patients were Italian (figure 1A), and sex distribution showed that male patients were predominantly affected (59.6%) (figure 1B). All patients were classified, according to their clinical phenotype, as affected by LGMD (51.2%), CM (32.5%), or other clinical conditions (16.3%), including, among others, distal myopathy (3.8%), isolated hyperCKemia (3.4%), and metabolic myopathy (1.2%) (figure 1C, table e-2).

Most patients were sporadic cases (81%), while 96 were familial, including 37 dominant and 59 recessive forms (figure 1D).

The vast majority (>90%) of samples collected had previously been tested unsuccessfully (figure 1E), according to the observed phenotype.

Molecular findings and clinical evaluation. In view of the high number of variants identified in each patient, bioinformatic filters were applied to prioritize variant type (missense, indel, stopgain, or stoploss), frequency in public and internal databases, and annotation as disease-associated variants were used. In addition, the correspondence of molecular findings with clinical presentation, age at onset, and segregation in familial cases was critically reevaluated. On the basis of these procedural steps, 218 cases obtained a diagnosis, showing known disease-associated variants, or variants of a likely pathogenicity, or variants predicted to affect function in genes corresponding to clinical suspicions. The majority of these patients (76%) showed already described or clear loss-of-function mutations, and most of the variants identified (65%) were missense (figure 2). Only one mutation causing the loss of the stop codon was identified and a single homozygous large deletion in *SGCB* was detected.

In 115 patients, LGMD-related genes were considered causative of the observed phenotype (table 1). Specifically, the *CAPN3* gene was responsible for 22 cases, the sarcoglycan genes for 20 cases, and the *DYSF* and *ANO5* genes for 15 cases, each.

Among the remaining genes (table 2), *RYR1* caused the highest number of cases (n = 25), followed by *NEB*, *LAMA2*, and *MYH7*.

A total of 160 cases (83 LGMD, 50 CM, and 27 with other conditions) are still under investigation since a clear and complete explanation of the observed

Figure 2   Diagnostic rate and molecular results
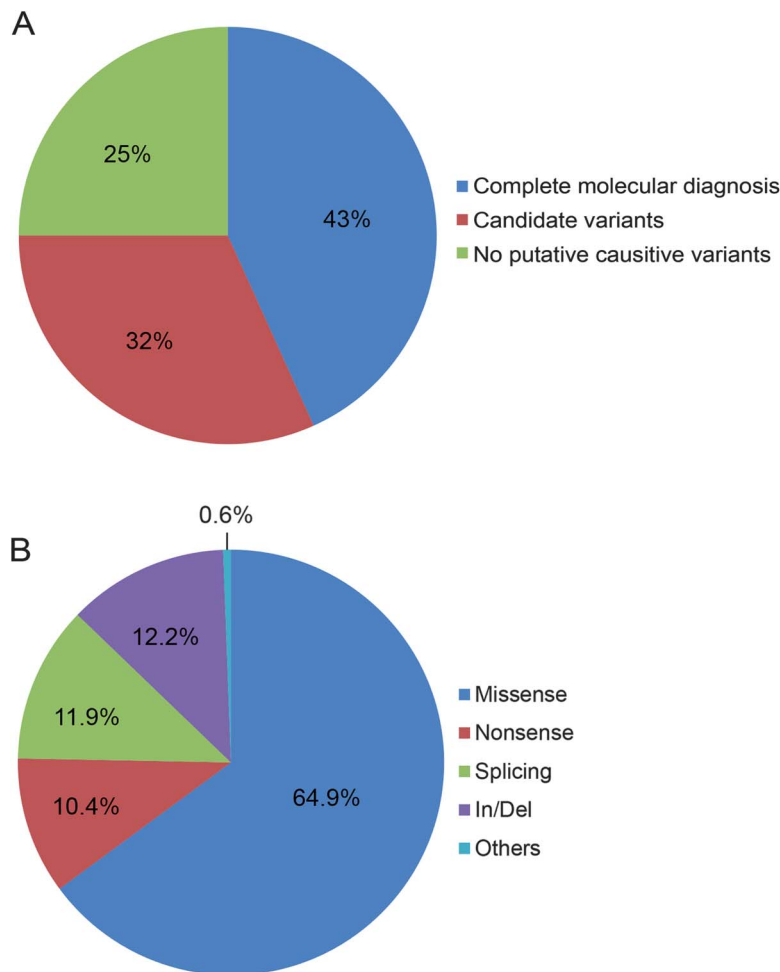
(A) Definitive diagnosis was obtained in 43.2% of patients. In addition, 31.8% of patients showed variants that required further characterization. (B) Type of causative mutations identified in diagnosed patients: missense (64.9%), small indels (12.2%), splice-site variants (11.9%), and nonsense mutations (10.4%).

Table 1	LGMD genes[22]		
Disease	Locus	Gene	No. of patients
LGMD1B	1q22	LMNA	3
LGMD1C	3p25.3	CAV3	2
LGMD2A	15q15	CAPN3	22
LGMD2B	2p13.2	DYSF	15
LGMD2C	13q12	SGCG	4
LGMD2D	17q21	SGCA	10
LGMD2E	4q12	SGCB	6
LGMD2G	17q12	TCAP	1
LGMD2H	9q33.1	TRIM32	1
LGMD2I	19q13.3	FKRP	7
LGMD2J	2q24.3	TTN	5
LGMD2K	9q34.1	POMT1	1
LGMD2L	11p13	ANO5	15
LGMD2M	9q31	FKTN	2
LGMD2N	14q24	POMT2	6
LGMD2R	2q35	DES	1
LGMD2S	4q35.1	TRAPPC11	2
LGMD2T	3p21	GMPPB	2
LGMD2V	17q25	GAA	10

Abbreviation: LGMD = limb-girdle muscular dystrophy.

Table 2	Other myopathy genes	
Locus	Gene	No. of patients
1q42.13	ACTA1	2
1p21	AGL	2
21q22.3	COL6A2	4
2q37	COL6A3	1
11q22.3–q23.1	CRYAB	1
Xp21.2	DMD	7
19p13.2	DNM2	5
Xq28	EMD	1
7q32	FLNC	4
17q25.2–q25.3	GLA	1
3p12	GNE	3
3p22.1	GTDC2	1
3q24	GYG1	1
12q13.2	ITGA7	2
6q22–q23	LAMA2	8
Xq28	MTM1	5
17p13.1	MYH2	1
14q12	MYH7	8
5q31	MYOT	1
2q23.3	NEB	9
11q12–q13.2	PYGM	3
20p13	RYR1	25
1p36.13	SEPN1	2
18p11.32	SMCHD1	1
6q25	SYNE1	1
14q23.2	SYNE2	1
9p13	TPM2	1
1q21.2	TPM3	2

phenotype has not been identified. In most of these cases, a single variant in at least one gene fitting the clinical conditions of the patients was detected: more than 250 variants were identified, including 128 previously described variants discovered in 98 patients.

In the 126 cases with no molecular findings fitting the observed phenotype, a large number of variants of uncertain significance were identified. Although we cannot exclude a possible role for some of these, other additional and undetectable genetic changes may cause the phenotype.

Of note, only 19% of the patients with LGMD showed no putative deleterious variants. In contrast, almost 31% of patients with CM and 32% of patients with other clinical conditions presented no putative causative variants.

DISCUSSION In recent years, NGS has totally transformed the approach to the study of NMDs,[17] and a number of diagnostic platforms have been developed for the simultaneous analysis of several different genes.[13,14,17,18] However, all previously published reports have shown a proof of concept on a limited number of patients only.

Herein, we report the results of a large study of an NMD cohort. The use of stringent bioinformatic filters to obtain relevant molecular data[15,16] and a careful reevaluation of clinical presentation and laboratory findings allowed us to identify putative causative mutations in 218 cases, representing slightly more than 43% of our cohort.

This ratio reflects the large number of prescreened samples in our cohort. As previously mentioned, samples from more than 90% of our patients had already been tested by Sanger sequencing or other laboratory techniques. Focusing on LGMD cases, for example, the number of patients with dysferlinopathy or calpainopathy, accounting for more than 40% of LGMDs in Italy,[19,20] is small because of extensive studies previously performed on these genes. In our screening analysis, 37 cases, corresponding to 14% of patients with LGMD recruited, showed causative mutations in the DYSF and CAPN3 genes. This large difference (40% vs 14%) reflects the filtering out of patients with DYSF and CAPN3 mutations. Further molecular tests, including segregation analyses or functional tests, on the 160 cases

Figure 3 Distribution of the 504 patients tested according to clinical phenotype and molecular findings
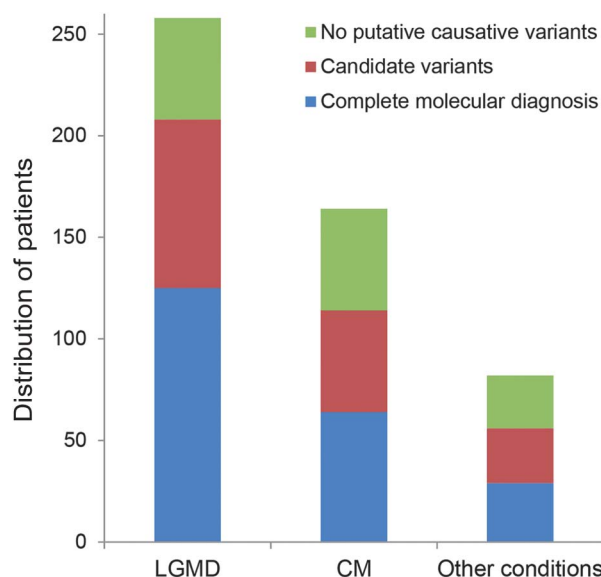
A total of 125 patients (48.4%) with LGMD received a molecular diagnosis, 83 (32.2%) are still under consideration, and in 50 patients (19.4%), no clear causative variants were detected. A total of 64 patients (39%) with CM received a molecular diagnosis, 50 (30.5%) are still under investigation, and 50 (30.5%) showed no causative variations. A total of 29 patients (35.4%) with other clinical conditions received a molecular diagnosis, 27 (32.9%) are to be further investigated, and in 26 patients (31.7%), no causative variants were identified. CM = congenital myopathy; LGMD = limb-girdle muscular dystrophy.

still under investigation could increase the detection rate by at least 10% to 15%, according to our conservative estimate.

One key finding from our study was the identification of patients with an LGMD phenotype carrying variants in causative genes not typically associated with LGMD. In particular, late-onset Pompe disease, resulting from mutations in the *GAA* gene,[21] was diagnosed in 10 patients. In our screening assay, *GAA* represents the fourth most common cause of a recessive limb-girdle myopathy that might mimic an LGMD,[22] suggesting that the difficulty in making a differential diagnosis between LGMD and late-onset Pompe disease may be underestimated.

Similarly, causative variants in patients with LGMD were identified in genes causing congenital MD (*GMPPB* or *LAMA2*), Bethlem myopathy (*COL6*), and other metabolic myopathies (*AGL*, *GYG1*, and *PYGM*).

We also detected seven causative variants in the *DMD* gene, involved in Becker muscular dystrophy.[23] The enormous size of the *DMD* gene had always hampered analysis by Sanger sequencing[24] and the prevalence of Becker muscular dystrophy is probably underestimated.

Of note, the diagnostic rate was higher for the LGMD cohort of patients than for both the CM group and for patients with other clinical conditions (figure 3). The molecular diagnosis of CM may have been hindered by the significant clinical overlap between CM and other NMDs including congenital MD, congenital myasthenic syndromes, metabolic myopathies, spinal muscular atrophy, and Prader-Willi syndrome, which can all be characterized by marked muscle weakness and/or hypotonia.[8] Motor-Plex is in fact unable to provide a molecular diagnosis for most of these disorders, and a subsequent diagnostic step is necessary to exclude them.

Interpreting the molecular data obtained in patients with "other clinical conditions" is much more complex as some of them present with a mild clinical phenotype. For example, of the 17 patients with isolated hyperCKemia, a genetic diagnosis was only obtained in 2 cases, as expected. However, 7 of them had at least one well-known heterozygous mutation in a gene that might possibly be responsible for the disorder.

We conducted a critical reevaluation of all the molecular data, looking for possible "double trouble," i.e., 2 independent molecular events, each responsible for a specific phenotype, causing the observed clinical condition. No clear "double trouble" was identified. However, in 72 patients with a complete diagnosis, a further, already described, causative variant in a recessive gene was detected, suggesting that it is quite a common finding.

Our study represents an important first step toward the identification of novel disease-causing genes. Some of the unsolved cases could in fact be attributable to previously undescribed causative genes.[25] Whole exome sequencing or whole genome sequencing may therefore be crucial tools to discover novel NMD genes.[26,27]

The combined use of other tools, such as CGH-array[28,29] or RNA-Seq,[30] to detect copy number variations and/or other variants missed by NGS is a necessary subsequent step to increase the diagnostic rate of NMDs.

Our extensive study confirms the importance of NGS screening as a powerful tool in the diagnostic workflow for NMD. In light of the speed, efficacy, and ever-decreasing cost of this approach, we predict that it will become a universal first-tier test for heterogeneous genetic conditions.

DISCLOSURE

The authors report no disclosures relevant to the manuscript. Go to Neurology.org for full disclosures.

REFERENCES

1. Bonnemann CG, Wang CH, Quijano-Roy S, et al. Diagnostic approach to the congenital muscular dystrophies. Neuromuscul Disord 2014;24:289–311.
2. Wicklund MP. The muscular dystrophies. Continuum 2013;19:1535–1570.
3. Iannaccone ST, Castro D. Congenital muscular dystrophies and congenital myopathies. Continuum 2013;19:1509–1534.
4. Reilly MM, Hanna MG. Genetic neuromuscular disease. J Neurol Neurosurg Psychiatry 2002;73(suppl 2):ii12–ii21.
5. Kaplan JC, Hamroun D. The 2015 version of the gene table of monogenic neuromuscular disorders (nuclear genome). Neuromuscul Disord 2014;24:1123–1153.
6. Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nat Genet 2003; (suppl 33):228–237.
7. Mercuri E, Muntoni F. Muscular dystrophies. Lancet 2013;381:845–860.
8. North KN, Wang CH, Clarke N, et al. Approach to the diagnosis of congenital myopathies. Neuromuscul Disord 2014;24:97–116.
9. Xue Y, Ankala A, Wilcox WR, Hegde MR. Solving the molecular diagnostic testing conundrum for mendelian disorders in the era of next-generation sequencing: single-gene, gene panel, or exome/genome sequencing. Genet Med 2015;17:444–451.
10. Mamanova L, Coffey AJ, Scott CE, et al. Target-enrichment strategies for next-generation sequencing. Nat Methods 2010;7:111–118.
11. Rehm HL, Bale SJ, Bayrak-Toydemir P, et al. ACMG clinical laboratory standards for next-generation sequencing. Genet Med 2013;15:733–747.
12. Lim BC, Lee S, Shin JY, et al. Genetic diagnosis of Duchenne and Becker muscular dystrophy using next-generation sequencing technology: comprehensive mutational search in a single platform. J Med Genet 2011;48:731–736.
13. Valencia CA, Rhodenizer D, Bhide S, et al. Assessment of target enrichment platforms using massively parallel sequencing for the mutation detection for congenital muscular dystrophy. J Mol Diagn 2012;14:233–246.
14. Vasli N, Laporte J. Impacts of massively parallel sequencing for genetic diagnosis of neuromuscular disorders. Acta Neuropathol 2013;125:173–185.
15. Savarese M, Di Fruscio G, Mutarelli M, et al. MotorPlex provides accurate variant detection across large muscle genes both in single myopathic patients and in pools of DNA samples. Acta Neuropathol Commun 2014;2:100.
16. Mutarelli M, Marwah V, Rispoli R, et al. A community-based resource for automatic exome variant-calling and annotation in mendelian disorders. BMC Genomics 2014;15 (suppl 3):S5.
17. Nigro V, Piluso G. Next generation sequencing (NGS) strategies for the genetic testing of myopathies. Acta Myol 2012;31:196–200.
18. Ankala A, da Silva C, Gualandi F, et al. A comprehensive genomic approach for neuromuscular diseases gives a high diagnostic yield. Ann Neurol 2015;77:206–214.
19. Magri F, Del Bo R, D'Angelo MG, et al. Frequency and characterisation of anoctamin 5 mutations in a cohort of Italian limb-girdle muscular dystrophy patients. Neuromuscul Disord 2012;22:934–943.
20. Piluso G, Politano L, Aurino S, et al. Extensive scanning of the calpain-3 gene broadens the spectrum of LGMD2A phenotypes. J Med Genet 2005;42:686–693.
21. Musumeci O, la Marca G, Spada M, et al. LOPED Study: looking for an early diagnosis in a late-onset Pompe disease high-risk population. J Neurol Neurosurg Psychiatry 2016;87:5–11. doi: 10.1136/jnnp-2014-310164.
22. Nigro V, Savarese M. Genetic basis of limb-girdle muscular dystrophies: the 2014 update. Acta Myol 2014;33:1–12.
23. Flanigan KM. Duchenne and Becker muscular dystrophies. Neurol Clin 2014;32:671–688, viii.
24. Bonnal RJ, Severgnini M, Castaldi A, et al. Reliable resequencing of the human dystrophin locus by universal long polymerase chain reaction and massive pyrosequencing. Anal Biochem 2010;406:176–184.
25. Bamshad MJ, Ng SB, Bigham AW, et al. Exome sequencing as a tool for mendelian disease gene discovery. Nat Rev Genet 2011;12:745–755.
26. Gilissen C, Hoischen A, Brunner HG, Veltman JA. Unlocking mendelian disease using exome sequencing. Genome Biol 2011;12:228.
27. Torella A, Fanin M, Mutarelli M, et al. Next-generation sequencing identifies transportin 3 as the causative gene for LGMD1F. PLoS One 2013;8:e63536.
28. Piluso G, Dionisi M, Del Vecchio Blanco F, et al. Motor chip: a comparative genomic hybridization microarray for copy-number mutations in 245 neuromuscular disorders. Clin Chem 2012;57:1584–1596.
29. Savarese M, Piluso G, Orteschi D, et al. Enhancer chip: detecting human copy number variations in regulatory elements. PLoS One 2012;7:e52264.
30. Chandrasekharappa SC, Lach FP, Kimble DC, et al. Massively parallel sequencing, aCGH, and RNA-Seq technologies provide a comprehensive molecular diagnosis of Fanconi anemia. Blood 2013;121:e138–e148.