Clinical and Molecular Prevalence of Lipodystrophy in an Unascertained Large Clinical Care Cohort

Claudia Gonzaga-Jauregui,[1] Wenzhen Ge,[2] Jeffrey Staples,[1] Cristopher Van Hout,[1] Ashish Yadav,[1] Ryan Colonie,[3] Joseph B. Leader,[3] H. Lester Kirchner,[3] Michael F. Murray,[3] Jeffrey G. Reid,[1] David J. Carey,[3] John D. Overton,[1] Alan R. Shuldiner,[1] Omri Gottesman,[1] Steve Gao,[2] Jesper Gromada,[2] Aris Baras,[1] and Judith Altarejos,[2] on behalf of the Geisinger-Regeneron DiscovEHR collaboration

Diabetes 2020;69:249–258 | https://doi.org/10.2337/db19-0447

Lipodystrophies are a group of disorders characterized by absence or loss of adipose tissue and abnormal fat distribution, commonly accompanied by metabolic dysregulation. Although considered rare disorders, their prevalence in the general population is not well understood. We aimed to evaluate the clinical and genetic prevalence of lipodystrophy disorders in a large clinical care cohort. We interrogated the electronic health record (EHR) information of >1.3 million adults from the Geisinger Health System for lipodystrophy diagnostic codes. We estimate a clinical prevalence of disease of 1 in 20,000 individuals. We performed genetic analyses in individuals with available genomic data to identify variants associated with inherited lipodystrophies and examined their EHR for comorbidities associated with lipodystrophy. We identified 16 individuals carrying the p.R482Q pathogenic variant in LMNA associated with Dunnigan familial partial lipodystrophy. Four had a clinical diagnosis of lipodystrophy, whereas the remaining had no documented clinical diagnosis despite having accompanying metabolic abnormalities. We observed a lipodystrophy-associated variant carrier frequency of 1 in 3,082 individuals in our cohort with substantial burden of metabolic dysregulation. We estimate a genetic prevalence of disease of ∼1 in 7,000 in the general population. Partial lipodystrophy is an underdiagnosed condition. and its prevalence, as defined molecularly, is higher than previously reported. Genetically guided stratification of patients with common metabolic disorders, like diabetes and dyslipidemia, is an important step toward precision medicine.

Inherited lipodystrophies are a group of genetically heterogenous disorders characterized by selective deficiency of adipose tissue in the absence of nutritional deprivation or catabolic state. This marked loss or absence of adipose tissue is commonly accompanied by hormonal and metabolic dysregulation that result in severe comorbidities due to ectopic fat accumulation in other organs such as liver and muscle (1).

Lipodystrophic disorders can be classified based on the extent and areas of fat loss that range from localized and partial lipodystrophy, generally affecting the limbs but sparing the trunk and face, to generalized lipodystrophy where lack of adipose tissue occurs in mostly the entirety of the body. Interestingly, because abnormal fat distribution is a key diagnostic feature of lipodystrophy disorders, this tends to be more recognizable in females versus males, resulting in a higher rate of reported cases for females compared with males, who are generally considered to be more muscular. Additionally, lipodystrophies can have genetic or environmental etiologies, with the latter usually a consequence of antiretroviral therapies in HIV-infected individuals (1–3). However, in the absence of significant family history or clinical signs evidencing a genetic disorder, lipodystrophy in many patients is considered to have an acquired cause or be lipodystrophy of unknown etiology (4). Nevertheless, genetic studies in families and individuals with lipodystrophy have been able to identify several genes and pathogenic variants that cause inherited lipodystrophy disorders.

[1]Regeneron Genetics Center, Regeneron Pharmaceuticals, Inc., Tarrytown, NY
[2]Regeneron Pharmaceuticals, Inc., Tarrytown, NY
[3]Geisinger Health System, Danville, PA

Corresponding authors: Judith Altarejos, judith.altarejos@regeneron.com, and Claudia Gonzaga-Jauregui, clau.gonzagajauregui@regeneron.com

Received 2 May 2019 and accepted 18 November 2019

This article contains Supplementary Data online at http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db19-0447/-/DC1.

Inherited lipodystrophies are classified based on their molecular etiology and adipose tissue loss distribution but can be broadly divided into congenital generalized lipodystrophies, which are mostly autosomal recessive disorders with a near complete absence of adipose tissue, and familial partial lipodystrophies (FPLDs), which are mostly inherited dominantly and present with a progressive selective loss of specific adipose depots (3,5–7). Inherited lipodystrophies are generally considered to be very rare genetic disorders with estimated prevalences ranging from 1 in 10 million for congenital generalized lipodystrophies to 1 in 1 million for FPLDs. Previous studies have estimated that the burden of inherited lipodystrophies in the general population ranges from 1.0 in 1 million (7) to 4.7 in 1 million (8). Mutations in a handful of genes have been associated with inherited lipodystrophy disorders, many of which are involved in adipocyte differentiation (*BSCL2*, *AKT2*, *PPARG*), maintenance (*LMNA*, *ZMPSTE24*, *CIDEC*, *PSMB8*), or function (*AGPAT2*, *CAV1*, *PTRF*, *LIPE*, *PLIN1*) (1,2,5,9–15). A few additional genes have been identified in patients with syndromic presentations where lipodystrophy is one of the clinical manifestations (*SPRTN*, *PIK3R1*, *WRN*). Mutations in *LMNA* and *PPARG* account for >50% of the reported cases of FPLD (16).

Independently from the underlying cause of the disease, the loss of or failure to develop adipose tissue in patients with lipodystrophy generally results in metabolic and endocrine abnormalities that usually encompass insulin resistance, diabetes, hypertriglyceridemia, hyperlipidemia, hypertension, hepatic steatosis, acanthosis nigricans, hormonal imbalance, and polycystic ovaries (3,17). Depending on the genetic defect, patients can also have features of cardiomyopathy, peripheral neuropathy, and skeletal abnormalities (18).

In this report, we examine the prevalence of clinically diagnosed lipodystrophy from de-identified electronic health record (EHR) data available from the Geisinger Health System (GHS) and a large claims database. Whole exome sequencing in 92,455 participants, part of the GHS-Regeneron DiscovEHR collaboration, enabled estimates of the prevalence of molecularly confirmed inherited lipodystrophies. Our findings suggest that inherited lipodystrophies are more common than previously reported.

RESEARCH DESIGN AND METHODS

Samples

DiscovEHR participants are a subset of the GHS MyCode Community Health Initiative. The MyCode Community Health Initiative is a GHS-wide repository of blood, serum, and DNA samples from GHS patients that have consented to participate in research and donate samples for broad research use, including genomic analysis that can be linked to de-identified EHR information. Participants consented in accordance with the GHS Institutional Review Board–approved protocol (study number 2006-0258).

Phenotype Query

We interrogated the EHR information for 1,361,924 adult individuals (age ≥18 years) from the GHS for cases of at least one inpatient diagnosis or two outpatient diagnoses of "lipodystrophy/lipoatrophy" diagnosis codes (ICD-9 272.6/ICD-10 E88.1) excluding individuals with HIV infection diagnosis codes (ICD-9 042/ICD-10 B20). Similarly, we interrogated the Truven Health MarketScan Research Database, a database of health care claims data for 85,688,196 adult individuals enrolled in the period from 1 January 2012–30 September 2017 for the number of cases of at least one inpatient diagnosis or two outpatient diagnoses (on separate calendar days) of lipodystrophy/lipoatrophy (ICD-9 272.6/ICD-10 E88.1) with no instance of HIV infection diagnostic codes (ICD-9-CM 042/ICD-10-CM B20) during the study period divided by the total adult population enrolled during the same period. Prevalence of lipodystrophy was age and sex standardized to the U.S. population in 2017. Comorbidities including hyperlipidemia, diabetes, hypertension, and nonalcoholic fatty liver were defined as having at least one diagnosis code during the study period. Medication use including insulin, metformin, and statins was defined as having at least one dispense of each drug during the study period. A list of ICD codes used for phenotype queries can be found in Supplementary Table 1.

Exome Sequencing

Sample preparation, whole exome sequencing, and sequence data production were performed at the Regeneron Genetics Center (RGC) as previously described (19). In brief, 1 μg high-quality genomic DNA was used for exome capture using the NimbleGen VCRome 2.1 or the IDT xGen target enrichment reagent. Captured libraries were sequenced on the Illumina HiSeq 2500 platform with v4 chemistry using paired-end 75 base pair (bp) reads. Exome sequencing was performed such that >85% of the bases were covered at ≥20×. Mean coverage across all samples included in the DiscovEHR cohort was ~80×. Mean coverage for samples specifically referenced in this study was 86.4×. Raw sequence reads were mapped and aligned to the GRCh38/hg38 human genome reference assembly using BWA-MEM. Single nucleotide and indel variants and genotypes were called using the GATK HaplotypeCaller. Sequencing and data quality metric statistics were captured for each sample to evaluate capture performance, alignment performance, and variant calling. Called variants were annotated using an RGC cloud-based developed pipeline.

Genetic Analyses

For downstream genetic analyses, variants were further annotated and analyzed using an in-house implemented annotation and analysis pipeline and additional customized Perl bioinformatics scripts for data processing.

For ascertainment and survey of pathogenic variants, a union of NCBI's ClinVar pathogenic/likely pathogenic and HGMD's (Human Gene Mutation Database) high-confidence disease-causing mutations (DM-High) reported variants with a phenotype association to "lipodystrophy" was considered. Variants were considered "pathogenic" if they were reported in both databases without conflicting interpretations.

Data and Resource Availability

The Truven Health MarketScan Research Database is available through paid license. The Exome Aggregation Consortium (ExAC) database is publicly available through exac.broadinstitute.org. Genomic variant frequency data for the first 50,000 DiscovEHR participants are available through www.discovehrshare.com. Individual phenotype and genomic data for DiscovEHR participants are not available due to privacy considerations. Genetic variants reported and discussed in this study are disclosed within the main text and figures and in Supplementary Data. Additional information for reproducing the results described in this study is available upon reasonable request and subject to a data use agreement and appropriate consent and privacy considerations.

RESULTS

We interrogated the EHR data for 1,361,924 adult individuals from the GHS for "lipodystrophy/lipoatrophy" diagnosis codes (ICD-9 272.6/ICD-10 E88.1) to assess the prevalence of this phenotype from a clinical perspective in an unascertained EHR-linked clinical cohort. We identified 114 adult patients with a clinical diagnosis code consistent with lipodystrophy; of these, 99 did not have a diagnosis of HIV infection and were therefore considered as likely having an inherited or genetic cause of lipodystrophy versus an acquired etiology. The calculated prevalence for lipodystrophy in the GHS clinical cohort was 7.2 in 100,000 (or 1 in 13,889). In order to replicate our observations, we similarly interrogated the Truven Health MarketScan Research Database for the number of cases of diagnoses of lipodystrophy/lipoatrophy divided by the total adult population enrolled during the same period. We identified 6,055 patients with a lipodystrophy diagnosis, of whom 4,029 did not have a diagnosis of HIV infection, resulting in an estimated prevalence for lipodystrophy of 4.7 in 100,000 (or 1 in 21,277). Consistent with previous reports on lipodystrophy, we observed a sex bias in the clinical diagnosis of lipodystrophy, with 77–80% of patients being female. Additionally, and consistent with disease presentation and known accompanying metabolic abnormalities, we observed an increased proportion of lipodystrophy patients also having diagnosis codes for comorbidities such as hyperlipidemia (ICD-9 272/ICD-10 E78), diabetes (ICD-9 250/ICD-10 E11), hypertension (ICD-9 401/ICD-10 I10), and nonalcoholic fatty liver disease (ICD-9 571.8/ICD-10 K75.8 and K76) in both databases (Table 1). The demographic and clinical characteristics of individuals with a clinical diagnosis of lipodystrophy in these two cohorts are summarized in Table 1. We further standardized the observed prevalence of the disease in the MarketScan database based on age and sex and projected this onto the U.S. population (as of July 2017) to better estimate the prevalence of lipodystrophy in the general population, resulting in an estimate of 47.3 cases per 1 million people (Supplementary Table 2) (or ~1 in 21,142). Based on these data, we estimate the clinical prevalence of inherited lipodystrophy to be ~1 in 20,000 individuals.

Of the 99 individuals with a likely inherited cause of lipodystrophy, 24 had been sequenced as part of the ongoing Geisinger-Regeneron DiscovEHR collaboration that links EHR and genomic data (19). Analyses of the phenotypes in these individuals showed that type 2 diabetes is significantly enriched in patients with a clinical diagnosis of lipodystrophy (odds ratio [OR] 4.28 [95% CI 1.90 – 9.64]), $P = 1.27 \times 10^{-4}$) (Table 2). Next, we performed genetic analyses in the 24 individuals with available genetic data in order to identify the likely causative variants of the condition in these patients. We identified four individuals who are heterozygous for a previously reported pathogenic variant in *LMNA* [hg38.g.chr1:156136985(G>A); c.1445G>A; p.R482Q] associated with Dunnigan FPLD (FPLD2) (11,20–22). Interestingly, an additional 12 individuals were also carriers of this genetic variant; however, they did not have any documented diagnoses of lipodystrophy in their EHR. This variant is present in our unascertained clinical cohort at a frequency of 0.000173 (16 of 92,455 sequenced individuals). The frequency of this variant appears to be higher in DiscovEHR than in other sequenced population cohorts and publicly available databases such as ExAC (minor allele frequency [MAF] = 0.000008324) and gnomAD (MAF = 0.000004063). De-identified EHR review of these 16 LMNA p.R482Q heterozygous individuals showed that they have diagnosis codes for phenotypes consistent with the metabolic abnormalities observed in lipodystrophy and that these abnormal metabolic phenotypes are significantly enriched in these individuals versus the rest of the DiscovEHR cohort, including hyperlipidemia (93.75%) (OR 11.19 [95% CI 1.47–84.76], $P = 3.18 \times 10^{-3}$), hypertension (93.75%) (OR 12.47 [95% CI 1.64–94.45], $P = 1.65 \times 10^{-3}$), diabetes (81.25%) (OR 13.26 [95% CI 3.77–46.54], $P = 1.47 \times 10^{-7}$), and inflammatory liver disease (25%) (OR 3.42 [95% CI 1.10–10.60], $P = 0.02344$) (Fig. 1 and Table 2). Their EHR also shows that they are on medications to treat these conditions, including insulin (68.75%), metformin (75%), and statins (100%). Additionally, we also observed other phenotypes less consistently associated with lipodystrophy including atherosclerosis (56.25%), hypothyroidism (37.5%), and chronic kidney disease (37.5%). Interestingly, other phenotypes of interest that were enriched in this group of patients were heart disease (68.75%) with numerous diagnoses of heart failure, myocardial infarctions and conduction disturbances, and neuropathies (50%) including carpal tunnel syndrome and one patient with a diagnosis of diabetic polyneuropathy (Fig. 1). These accompanying diagnoses would be consistent with the clinical spectrum observed for laminopathies, specifically, *LMNA*-associated disorders (18,23). Of note, identity by descent (IBD) metrics derived from genomic data to estimate relatedness (24) among individuals carrying the LMNA (p.R482Q) variant only identified two of these carriers to be related, a brother-sister sibling pair where, although both are carriers of the variant, only the sister (patient 4 [Fig. 1*A*]) has a clinical diagnosis of lipodystrophy. Examination of the brother's EHR information showed that he has clinical diagnoses of

Table 1—Demographics and comorbidities of lipodystrophy* and control patients from two unascertained clinical population databases

	GHS all adult patients (age ≥18 years) (N = 1,361,924)	GHS lipodystrophy cases, no HIV (age ≥18 years) (N = 99)	OR (95% CI)	MarketScan database, all adult patients (age ≥18 years) (N = 85,688,196)	MarketScan database, lipodystrophy cases, no HIV (age ≥18 years) (N = 4,029)	OR (95% CI)
Age (years)						
Mean (SD)	51.32 (20.67)	54.56 (15.57)		45.26 (17.11)	52.24 (13.72)	
Median	50.9	54.8		45	53	
Sex						
Male	633,536 (46.51)	20 (20.2)		41,282,911 (48.18)	921 (22.86)	
Female	728,388 (53.48)	79 (79.8)		44,405,285 (51.82)	3,108 (77.14)	
Hyperlipidemia[1]						
No	1,052,782 (77.30)	57 (57.57)	2.51 (1.68–3.73)	66,037,692 (77.07)	501 (12.43)	23.66 (21.5–25.9)
Yes	309,142 (22.69)	42 (42.42)	($P = 2.79 \times 10^{-6}$)	19,650,504 (22.93)	3,528 (87.57)	($P < 1 \times 10^{-12}$)
Type 2 diabetes[2]						
No	1,234,278 (90.62)	78 (78.78)	2.60 (1.60–4.21)	78,301,315 (91.38)	2,665 (66.15)	5.42 (5.08–5.79)
Yes	127,646 (9.37)	21 (21.21)	($P = 5.3 \times 10^{-5}$)	7,386,881 (8.62)	1,364 (33.85)	($P < 1 \times 10^{-12}$)
Hypertension[3]						
No	1,034,366 (75.94)	63 (63.63)	1.80 (1.19–2.71)	66,713,916 (77.86)	1,808 (44.87)	4.32 (4.05–4.59)
Yes	327,558 (24.05)	36 (36.36)	($P = 4.15 \times 10^{-3}$)	18,974,280 (22.14)	2,221 (55.13)	($P < 1 \times 10^{-12}$)
Nonalcoholic fatty liver[4]						
No	1,328,261 (97.52)	90 (90.9)	3.94 (1.98–7.83)	84,054,976 (98.09)	3,572 (88.66)	6.58 (5.97–7.25)
Yes	33,663 (2.47)	9 (9.09)	($P = 2.21 \times 10^{-5}$)	1,633,220 (1.91)	457 (11.34)	($P < 1 \times 10^{-12}$)
Medication use						
HMG-CoA reductase inhibitor	272,332 (19.99)	34 (34.34)	2.09 (1.38–3.16) ($P = 3.58 \times 10^{-4}$)	9,472,308 (11.05)	1,050 (26.06)	2.83 (2.64–3.04) ($P < 1 \times 10^{-12}$)
Metformin	96,561 (7.09)	21 (21.21)	3.52 (2.17–5.71) ($P = 4.37 \times 10^{-8}$)	3,401,863 (3.97)	619 (15.36)	4.39 (4.03–4.78) ($P < 1 \times 10^{-12}$)
Insulin	91,599 (6.72)	22 (22.22)	3.96 (3.46–6.36) ($P = 7.44 \times 10^{-10}$)	1,396,592 (1.63)	383 (9.51)	6.34 (5.70–7.04) ($P < 1 \times 10^{-12}$)

Data are n (%) unless otherwise indicated. *Lipodystrophy is defined as having one inpatient or two outpatient diagnosis codes of ICD-9-272.6/ICD-10 E88.1. [1]Hyperlipidemia is defined as having at least one diagnosis code of ICD-9 272/ICD-10 E78; [2]diabetes is defined as having at least one diagnosis code of ICD-9 250/ICD-10 E11; [3]hypertension is defined as having at least one diagnosis code of ICD-9 401/ICD-10 I10; [4]nonalcoholic fatty liver is defined as having at least one diagnosis code of ICD-9 571.8/ICD-10 K75.8 and K76.

Table 2—Demographics and comorbidities of individuals with lipodystrophy and individuals carrying the LMNA (p.R482Q) variant in the DiscovEHR cohort

	GHS, all adult patients (N = 1,361,924)	GHS-RGC DiscovEHR participants (N = 92,455)[a]	OR (95% CI)	GHS-RGC DiscovEHR lipodystrophy cases, no HIV (N = 24)[b]	OR (95% CI)	GHS-RGC DiscovEHR LMNA (p.R482Q) variant carriers (N = 16)[c]	OR (95% CI)
Age (years)							
Mean (SD)	51.32 (20.67)	57.1 (17.7)		56.83 (10.9)		62.5 (9.01)	
Median	50.9	58.7		56.65		60.6	
Sex, *n* (%)							
Male	36,468 (39.44)	36,468 (39.44)		20 (20.2)		3 (18.75)	
Female	55,987 (60.55)	55,987 (60.55)		22 (83.33)		13 (81.25)	
BMI (kg/m^2)							
Mean (SD)	29.03 (7.40)	31.2 (7.3)		29.53 (6.12)		25.7 (3.45)	
Median (min–max)	27.8 (10–79.9)*	30.1 (14.4–57.5)		28.55 (19.8–44.8)		26.0 (18.3–32.8)	
Lipodystrophy[1]							
No	1,361,852 (99.99)	92,431 (99.97)	4.39 (2.77–6.95) ($P = 5.05 \times 10^{-12}$)	0 (0)	NA	12 (75)	1,540.317 (457–5,183) ($P < 1 \times 10^{-12}$)
Yes	99 (0.00007)	24 (0.026)		24 (100)		4 (25)	
Hyperlipidemia[2]							
No	1,052,782 (77.30)	39,509 (42.73)	5.30 (5.22–5.37) ($P < 1 \times 10^{-12}$)	9 (56.25)	1.24 (0.54–2.84) ($P = 0.60423$)	1 (6.25)	11.19 (1.47–84.7) ($P = 3.18 \times 10^{-3}$)
Yes	309,142 (22.69)	52,946 (57.26)		15 (62.5)		15 (93.75)	
Type 2 diabetes[3]							
No	1,234,278 (90.62)	69,678 (75.36)	3.63 (3.57–3.69) ($P < 1 \times 10^{-12}$)	10 (41.66)	4.28 (1.90–9.64) ($P = 1.27 \times 10^{-4}$)	3 (18.75)	13.26 (3.77–46.5) ($P = 1.47 \times 10^{-7}$)
Yes	127,646 (9.37)	22,777 (24.63)		14 (58.33)		13 (81.25)	
Hypertension[4]							
No	1,034,366 (75.94)	41,975 (45.40)	4.30 (4.25–4.36) ($P < 1 \times 10^{-12}$)	10 (41.66)	1.16 (0.51–2.62) ($P = 0.71329$)	2 (12.5)	12.47 (1.64–94.4) ($P = 8.19 \times 10^{-3}$)
Yes	327,558 (24.05)	50,480 (54.59)		14 (58.33)		14 (87.5)	
Nonalcoholic fatty liver[5]							
No	1,328,261 (97.52)	84,242 (91.11)	4.76 (4.64–4.89) ($P < 1 \times 10^{-12}$)	19 (79.16)	2.70 (1.00–7.23) ($P = 0.03958$)	12 (75)	3.42 (1.10–10.6) ($P = 0.02344$)
Yes	33,663 (2.47)	8,213 (8.88)		5 (20.83)		4 (25)	

Continued on p. 254

Table 2—Continued

	GHS, all adult patients (N = 1,361,924)	GHS-RGC DiscovEHR participants (N = 92,455)[a]	OR (95% CI)	GHS-RGC DiscovEHR lipodystrophy cases, no HIV (N = 24)[b]	OR (95% CI)	GHS-RGC DiscovEHR LMNA (p.R482Q) variant carriers (N = 16)[c]	OR (95% CI)
Medication use							
HMG-CoA reductase inhibitor	272,332 (19.99)	44,857 (48.51)	4.31 (4.25–4.37) ($P < 1 \times 10^{-12}$)	13 (54.16)	1.25 (0.56–2.79) (P = 0.57971)	16 (100)	NA
Metformin	96,561 (7.09)	19,451 (21.03)	4.11 (4.04–4.19) ($P < 1 \times 10^{-12}$)	10 (41.66)	2.68 (1.91–6.03) (P = 0.01314)	12 (75)	11.26 (3.63–34.9) ($P = 1.18 \times 10^{-7}$)
Insulin	91,599 (6.72)	19,718 (21.32)	4.51 (4.43–4.59) ($P < 1 \times 10^{-12}$)	10 (41.66)	2.63 (1.17–5.93) (P = 0.01497)	11 (68.75)	8.45 (2.93–24.3) ($P = 2.01 \times 10^{-6}$)

Data are n (%) unless otherwise indicated. min–max, minimum–maximum. *Extreme outlier BMI values <10 or >80 kg/m^2 were considered likely EHR entry errors and excluded from these calculations. [1]Lipodystrophy defined as having clinical diagnosis code for ICD-9 272.6/ICD-10 E88.1; [2]hyperlipidemia is defined as having at least one diagnosis code of ICD-9 272/ICD-10 E78; [3]diabetes is defined as having at least one diagnosis code of ICD-9 250/ICD-10 E11; [4]hypertension is defined as having at least one diagnosis code of ICD-9 401/ICD-10 I10; [5]nonalcoholic fatty liver is defined as having at least one diagnosis code of ICD-9 571.8/ICD-10 K75.8 and K76. ORs (95% CI) and significance were calculated comparing values for the following groups: [a]DiscovEHR all vs. Geisinger all, [b]DiscovEHR lipodystrophy vs. DiscovEHR all, and [c]LMNA p.R482Q carriers vs. DiscovEHR all.

type 2 diabetes, hyperlipidemia, and hypertension (patient 7 [Fig. 1*A* and *B*]). The remaining individuals do not appear to be related up to a fifth-degree relationship, which would be equivalent to unrelated individuals from an outbred population. We further confirmed this by exploring the IBD segments among these individuals and did not observe large shared genomic segments, except for the sibling pair previously identified and beyond the region where the *LMNA* variant is located. We were able to narrow the genomic region in which this variant resides to a 5.332 Mb minimum shared haplotype across all carriers (Fig. 1*C*). Genetic analyses of the remaining patients with a clinical diagnosis of lipodystrophy did not identify a clear pathogenic variant for the disease in known lipodystrophy disease genes.

To further explore whether additional cases of genetically driven lipodystrophy exist in our clinical population, we surveyed the DiscovEHR cohort (N = 92,455 sequenced individuals) for pathogenic and likely pathogenic variants previously reported in the HGMD or the NCBI database of Clinical variants (ClinVar) in lipodystrophy-associated genes. We identified an additional eight individuals with previously reported pathogenic variants in these databases in *LMNA*, *PPARG*, and *PIK3R1*, all without a clinical diagnosis of lipodystrophy (Supplementary Table 3). Similar to our observations in the *LMNA* (p.R482Q) variant carriers, we observed a substantial burden and similar pattern of metabolic dysregulation by EHR review in carriers of these variants consistent with a likely undiagnosed lipodystrophy disorder (Fig. 2*A*). We identified an individual carrying a predicted pathogenic missense variant classified as "likely pathogenic" in ClinVar [hg38.g.chr3:12392714(G>A); c.G497A; p.R166Q] affecting the highly conserved Arg^{166}/Arg^{194} residue in PPARG. A previously reported pathogenic variant (p.R166W/p.R194W) affecting this same residue and shown to disrupt DNA binding activity of PPARG (25) was also found in our cohort (Fig. 2 and Supplementary Table 3). This p.R166Q/p.R194Q variant is predicted to be pathogenic by bioinformatic algorithms, including the MITER classifier tool specifically developed to assess pathogenicity of missense variants in *PPARG* (26). The "probability of causing FPLD3" (FPLD type 3) according to MITER for each of the three missense variants in *PPARG* identified in our cohort (p.R166W/p.R194W, p.R166Q/R194Q, and p.V290M/p.V318M) was 97.2%, 99.9%, and 5.6%, respectively. Based on previously reported pathogenic variants only, the prevalence of lipodystrophy in the DiscovEHR cohort is 1 in 4,020. However, we additionally identified six individuals with heterozygous predicted loss of function and expected pathogenic variants in *PPARG*. Loss-of-function and dominant negative variants in *PPARG* have been associated with FPLD3 (MIM #604367), severe insulin resistance, diabetes, and hypertension (10,14,26). These additional six carriers of expected pathogenic variants had a clinical presentation similar to those of the other pathogenic variant carriers (Fig. 2*A*). Four of these individuals were related and part of the same pedigree segregating a novel nonsense variant in exon 3 of *PPARG*

[hg38.g.chr3:12379922(G>T); c.G217T; p.E73X] (Fig. 2*B*). Interestingly, a progression of disease and additional phenotypes consistent with FPLD3 can be observed in the older individuals of these pedigree versus the two younger sisters (Fig. 2*B*). Altogether, we identified 14 additional individuals who are carriers of a pathogenic, likely pathogenic, or expected pathogenic variant for lipodystrophy in our DiscovEHR cohort. These individuals have hallmarks of metabolic disease in the spectrum of lipodystrophy comorbidities but do not have a documented clinical diagnosis of the disease in their EHR. Overall, among the 92,455 sequenced participants from the DiscovEHR clinical cohort, we identified 30 carriers of pathogenic, likely pathogenic, or expected pathogenic variants in the lipodystrophy-associated genes, *LMNA*, *PPARG*, and *PIK3R1*, which amounts to a genetic variant carrier prevalence of 1 in 3,082. In order to evaluate the genetic prevalence of lipodystrophy in a different data set, we conducted a similar survey of lipodystrophy-associated genetic variants in the ExAC (27) database of 60,706 non–clinically ascertained individuals and identified 8 individuals carrying pathogenic or likely pathogenic variants in lipodystrophy-associated genes. This number accounts for a pathogenic variant carrier prevalence for autosomal dominant FPLD of ~1 in 7,588 individuals in ExAc. Combining these two observations from a large clinical care cohort with available genomic data linked to longitudinal EHR information and a non–clinically ascertained large population genomic data set with broad continental ancestry representation, we estimate the molecular prevalence of lipodystrophy disorders to be 1 in 7,000 individuals.

DISCUSSION

Through clinical and molecular investigations in an unascertained large clinical cohort, our data show that lipodystrophy is an underestimated and underdiagnosed condition. The prevalence of disease is higher than previously estimated both from a clinical and molecular genetics perspective.

From large clinical cohorts, we have derived an estimated clinical prevalence of lipodystrophy disorders of ~1 in 20,000 individuals, which is much higher than previously reported. This was first observed in the GHS database encompassing >1.3 million adult patients and further confirmed in the larger MarketScan database that includes de-identified data from >85 million adult individuals. This estimated prevalence is in contrast with previous reported estimates from the literature of <1.0 in a million (7) and a more recent study that similarly looked at EHR-based databases and literature to achieve an estimated prevalence of 1.3–4.7 cases per million for all types of lipodystrophy (8). Similarly to Chiquette et al. (8), we queried the databases for the corresponding clinical diagnosis codes for lipodystrophy and excluded individuals with HIV. However, instead of using additional inclusion/exclusion criteria, we decided to characterize the clinical features of this disorder by evaluating the fraction of individuals presenting with the

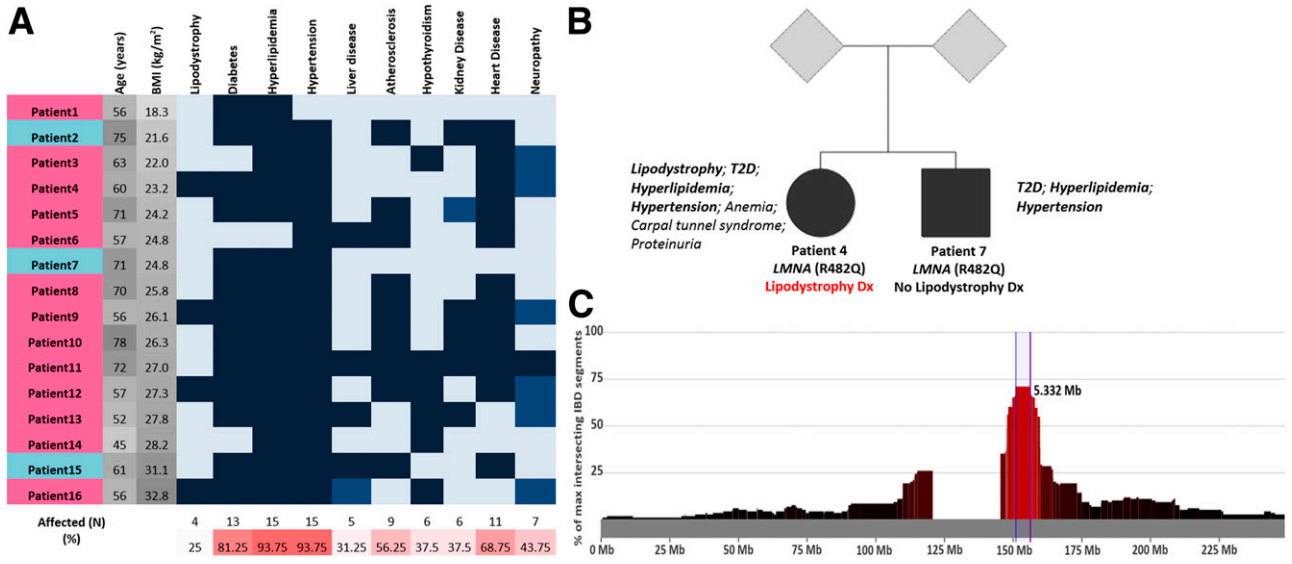

Figure 1—Phenotypic and genetic characterization of carriers of the pathogenic LMNA (p.R482Q) variant in DiscovEHR cohort. *A*: Major phenotypes identified through de-identified EHR review of the encounter diagnoses of DiscovEHR participants identified to be carriers of the pathogenic p.R482Q variant in LMNA. Patients are listed from lower to higher BMI as noted in the third column and colored based on sex: females, pink; males, blue. Age is listed in the second column. For the phenotypes: dark blue, presence of specific diagnosis codes for the particular disease; light blue, absence of diagnosis codes for the disease; intermediate blue, suggestive of disease due to related diagnosis codes but no listing of the specific code for the particular disease. *B*: Inferred pedigree for the two first-degree–related individuals among the 16 p.R482Q variant carriers, a brother-sister pair where only the female has a clinical diagnosis of lipodystrophy, despite the brother having diagnosis codes consistent with comorbidities documented in lipodystrophy patients. *C*: Identification of the common 5.332 Mb haplotype harboring the p.R482Q variant in LMNA by IBD segment analysis in the 16 carrier individuals.

expected associated comorbidities. We observe an enrichment of diagnosis codes for hyperlipidemia, type 2 diabetes, hypertension, and liver disease in individuals that also carry a clinical diagnosis of lipodystrophy (Table 1). This enrichment is consistent and significant both in the broader clinical lipodystrophy cohorts and in the subset of sequenced DiscovEHR participants, despite the high rate of metabolic disease in the larger DiscovEHR cohort as evidenced in Table 2 and as previously reported (19). Although the majority of the DiscovEHR cohort is clinically unascertained, a fraction of the participants is derived from the GHS cardiac catherization laboratory and the bariatric surgery clinic (19), which could contribute to an overrepresentation of metabolic abnormalities in this cohort. Furthermore, individuals who have more frequent interaction with the health care system are more likely to be approached and recruited to consent for the MyCode Community Health Initiative, which could also contribute to the enrichment of chronic diseases represented among the DiscovEHR participants.

Consistent with previous reports in lipodystrophy patients, we observed a 3:1 ratio of females being diagnosed versus males. This has been documented to be a diagnostic bias due to the increased likelihood of detecting abnormal fat distribution in females versus males due to a more muscular appearance in the absence of subcutaneous fat. It is also possible that in our database query, we are capturing cases of localized lipoatrophy or insulin lipoatrophy, which based on ICD diagnoses are not possible to differentiate and may be contributing to an increase in our estimate of clinical prevalence of lipodystrophy. However, because individuals with lipodystrophy may also use insulin as a medication, as evidenced by our data in Table 2 where 11 of the 16 (68.75%) patients with a molecular diagnosis of FPLD2 due to the p.R482Q variant in LMNA have used insulin as a medication, excluding patients on insulin may not be appropriate when assessing the clinical prevalence of this disorder.

Additionally, we performed genetic analyses to better characterize the underlying molecular etiology of lipodystrophy in patients with a clinical diagnosis of disease, as well as to explore the phenotypic associations of pathogenic and likely pathogenic variants in lipodystrophy-associated genes. In our DiscovEHR cohort, we identified a known molecular cause of disease in 16.66% (4 of 24) of our sequenced participants with a clinical diagnosis of lipodystrophy. In addition to four lipodystrophy-diagnosed patients, the p.R482Q variant in LMNA that has been previously reported and characterized as causative of FPLD2, Dunnigan type, was observed in additional patients in our cohort who did not carry a clinical diagnosis code for lipodystrophy. However, de-identified manual chart review of their EHR showed that they had clinical features consistent with disease. The p.R482Q variant was identified in probands from distinct

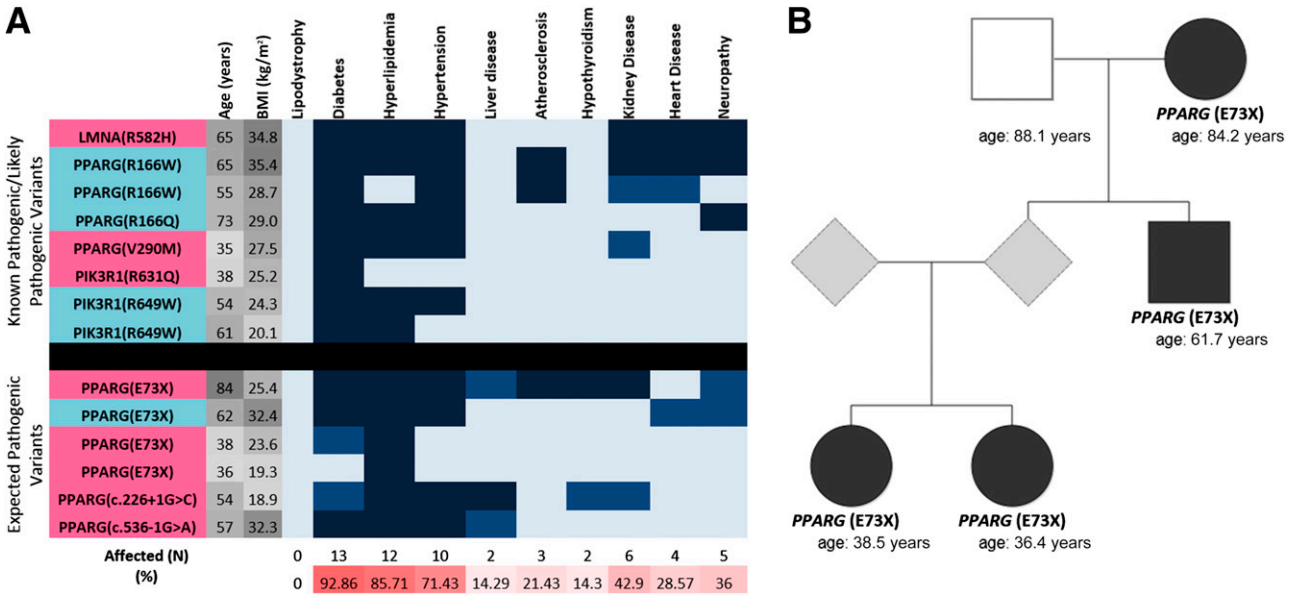

Figure 2—Phenotypes in carriers of pathogenic, likely pathogenic, and expected pathogenic variants for lipodystrophy in the DiscovEHR cohort. *A*: Major phenotypes identified through de-identified EHR review of the encounter diagnoses of DiscovEHR participants found to be carriers of known pathogenic, likely pathogenic, and expected pathogenic variants for lipodystrophy. The gene and identified variant are listed in the first column, age is reported in the second column, and corresponding BMI is noted in the third column. Patients are colored based on sex: females, pink; males, blue. For the phenotypes: dark blue, presence of specific diagnosis codes for the particular disease; light blue, absence of diagnosis codes for the disease; intermediate blue, suggestive of disease due to related diagnosis codes but no listing of the specific code for the particular disease. *B*: Inferred pedigree for four identified individuals carrying a novel nonsense variant (p.E73X) in PPARG expected to be pathogenic and cause lipodystrophy. These individuals correspond to patients 8, 9, 10, and 11 from panel *A* listed in descending order for age. Severity of the phenotype appears to correlate with age with the older female individual being the most affected and with the youngest only having recorded diagnoses for "other and unspecified hyperlipidemia" within a short electronic health record.

families segregating Dunnigan-type FPLD from Canada and the U.K. (11,20). Clinical characterization of patients with this variant has shown that variant carriers have reduced plasma leptin levels by ~60% of noncarriers and increased concentrations of fasting plasma insulin and C-peptide (28). Additionally, they often are diagnosed with dyslipidemia, hyperinsulinemia, diabetes, and hypertension (21). Our observations confirm these previously reported associations with metabolic abnormalities in patients with this variant. Furthermore, they highlight the underdiagnosis of lipodystrophy in these individuals that have all of the metabolic disease hallmarks of FPLD2 due to the p.R482Q variant in LMNA. Additional variants in *LMNA* including other missense variants altering the same codon of the Arg^{482} residue have been associated with FPLD2, suggesting an important role of this residue and the COOH-terminal globular domain of lamin A/C in adipocyte maintenance (11,22). In addition to FPLD2, mutations in *LMNA* have been associated with a range of genetic disorders broadly known as laminopathies, including Charcot-Marie-Tooth disease type 2B1 (MIM #605588), dilated cardiomyopathy (MIM #115200), Hutchinson-Gilford progeria (MIM #176670), and a spectrum of muscular dystrophies, among others (18). The clinical spectrum of these disorders often overlaps and involves other tissues including skeletal and cardiac muscle and the peripheral nervous system. Therefore, it is not surprising that FPLD2 patients due to mutations in *LMNA* often present with additional comorbidities beyond the metabolic complications common to lipodystrophy, including heart disease and peripheral neuropathy, as evidenced by our patients with *LMNA* mutations (23) (Figs. 1 and 2). Variability in the expression of these phenotypes in patients can be partly explained by the progressive and later-onset nature of these traits; however, the possibility of genetic and/or other environmental modifiers that may exacerbate, accelerate, or prevent the onset of additional comorbidities may be further explored in the future as more patients with these conditions are ascertained and phenotyped in detail. While lipodystrophy is not a condition currently covered by the ACMG SF v2.0 gene list (29), *LMNA* is one of the genes screened for clinically actionable secondary findings in genomic sequencing studies due to its association with dilated cardiomyopathy characterized by cardiac dilation and reduced systolic function (29,30). It is possible that as more genomic sequencing efforts start looking for pathogenic variants in this gene, additional cases of molecularly undiagnosed lipodystrophy patients will become evident (31), further informing the prevalence and contribution of *LMNA* pathogenic variation to partial lipodystrophy and metabolic disease in unascertained patients.

In addition to the p.R482Q variant in LMNA that appears to be enriched in our particular cohort compared with other publicly available population databases, we identified additional pathogenic and likely pathogenic variants in *LMNA*, *PPARG*, and *PIK3R1* in patients with metabolic abnormalities in the spectrum of lipodystrophy but without a clinical diagnosis. This suggests that the prevalence of genetic lipodystrophies, as defined molecularly, might be higher than previously reported. Loss of function and pathogenic variation in *PPARG* and *LMNA* appear to be the major molecular contributors to cases of undiagnosed lipodystrophy. In our cohort, we calculate the prevalence of inherited lipodystrophy disorders to be ~1 in 3,082. However, with exclusion of the p.R482Q LMNA variant enriched in the DiscovEHR cohort to correct for an overascertainment of individuals with lipodystrophy in this population, a carrier frequency of ~1 in 6,604 is calculated. This lipodystrophy variant carrier frequency is more in line with the observation in a non–clinical care cohort such as ExAc, where ~1 in 7,588 individuals carry a pathogenic or likely pathogenic variant. The resulting calculated molecular prevalence of ~ 1 in 7,000 individuals for lipodystrophy is much higher than previously estimated. This is particularly relevant, as carriers of lipodystrophy-associated variants show metabolic abnormalities in the spectrum of lipodystrophy disease; however, they are more often diagnosed as having common metabolic diseases, such as type 2 diabetes, dyslipidemias, and unspecified metabolic syndrome (17).

Although the clinical and molecular prevalence of lipodystrophy was higher in our DiscovEHR cohort compared with broader external databases, it is likely that the clinical prevalence observed in DiscovEHR will be typical for other clinical care–ascertained cohorts, while the molecular prevalence estimate will be relevant to the implementation of precision medicine in similar clinical settings. The availability of both clinical and genetic data from a single large clinical care cohort such as DiscovEHR allows us not only to begin to estimate prevalence rates but also to explore issues of underdiagnosis, variant pathogenicity, penetrance, and expressivity of "rare" genetic disorders such as lipodystrophy. Genetic diagnosis of these conditions can inform disease course and potential disease complications in carriers of pathogenic variants. Early treatment and appropriate management can improve prognosis in patients with a molecular diagnosis of lipodystrophy before the onset or exacerbation of metabolic complications and comorbidities. Additionally, other therapies that address the molecular defect, such as the use of leptin analogs, may improve the metabolic profile of these patients. Furthermore, genetically guided stratification of patients with "common" disorders, such as diabetes and dyslipidemia, is an important step toward precision medicine and has the potential to lead to more effective therapies for patients.

Acknowledgments. The authors thank the participants of the Geisinger MyCode Community Health Initiative and the DiscovEHR Geisinger-Regeneron collaboration who contribute their genomic and clinical data for research purposes to improve and implement genomic precision medicine.

Duality of Interest. C.G.-J., W.G., J.S., C.V.H., A.Y., J.G.R., J.D.O., A.R.S., S.G., A.B., and J.A. are full-time employees of the Regeneron Genetics Center or Regeneron Pharmaceuticals, Inc., and receive stock options as part of compensation. O.G. and J.G. are former employees of Regeneron Pharmaceuticals, Inc.,

and received stock options as part of compensation. J.G. has recently become an employee of Exonics Therapeutics. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. C.G.-J. coconceived and designed the study, performed genetic and data analyses, elaborated tables and figures, and wrote and revised the manuscript. W.G. performed phenotype data analyses. J.S. contributed to genetic analyses. C.V.H. contributed to statistical analyses and manuscript editing and revision. A.Y., R.C., J.B.L., H.L.K., and M.F.M. contributed to phenotype data analyses. J.G.R. contributed bioinformatic resources. D.J.C. contributed to study design and manuscript revision. J.D.O. contributed to sequencing of samples, data quality, and manuscript revision. A.R.S. contributed to study design and discussion and manuscript editing and revision. O.G. contributed to study design; phenotype data analyses, resources, and tools; and manuscript revision. S.G. contributed to phenotype data analyses and methods. J.G. and A.B. contributed to study design and manuscript revision. J.A. coconceived and designed the study and wrote and revised the manuscript. C.G.-J. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at the American Society of Human Genetics 67th Annual Meeting 2017 (Orlando, FL, 17–21 October 2017) and European Human Genetics Conference 2018 (Milan, Italy, 16–19 June 2018).

References

1. Agarwal AK, Garg A. Genetic disorders of adipose tissue development, differentiation, and death. Annu Rev Genomics Hum Genet 2006;7:175–199
2. Hussain I, Patni N, Garg A. Lipodystrophies, dyslipidaemias and atherosclerotic cardiovascular disease. Pathology 2019;51:202–212
3. Akinci B, Meral R, Oral EA. Phenotypic and genetic characteristics of lipodystrophy: pathophysiology, metabolic abnormalities, and comorbidities. Curr Diab Rep 2018;18:143
4. Garg A. Acquired and inherited lipodystrophies. N Engl J Med 2004;350:1220–1234
5. Hussain I, Garg A. Lipodystrophy syndromes. Endocrinol Metab Clin North Am 2016;45:783–797
6. Gupta N, Asi N, Farah W, et al. Clinical features and management of non-HIV-related lipodystrophy in children: a systematic review. J Clin Endocrinol Metab 2017;102:363–374
7. Garg A. Clinical review#: lipodystrophies: genetic and acquired body fat disorders. J Clin Endocrinol Metab 2011;96:3313–3325
8. Chiquette E, Oral EA, Garg A, Araújo-Vilar D, Dhankhar P. Estimating the prevalence of generalized and partial lipodystrophy: findings and challenges. Diabetes Metab Syndr Obes 2017;10:375–383
9. Garg A, Agarwal AK. Lipodystrophies: disorders of adipose tissue biology. Biochim Biophys Acta 2009;1791:507–513
10. Barroso I, Gurnell M, Crowley VE, et al. Dominant negative mutations in human PPARgamma associated with severe insulin resistance, diabetes mellitus and hypertension. Nature 1999;402:880–883
11. Shackleton S, Lloyd DJ, Jackson SN, et al. LMNA, encoding lamin A/C, is mutated in partial lipodystrophy. Nat Genet 2000;24:153–156
12. Magré J, Delépine M, Khallouf E, et al.; BSCL Working Group. Identification of the gene altered in Berardinelli-Seip congenital lipodystrophy on chromosome 11q13. Nat Genet 2001;28:365–370
13. Agarwal AK, Arioglu E, De Almeida S, et al. AGPAT2 is mutated in congenital generalized lipodystrophy linked to chromosome 9q34. Nat Genet 2002;31:21–23
14. Savage DB, Tan GD, Acerini CL, et al. Human metabolic syndrome resulting from dominant-negative mutations in the nuclear receptor peroxisome proliferator-activated receptor-γ. Diabetes 2003;52:910–917
15. Gandotra S, Le Dour C, Bottomley W, et al. Perilipin deficiency and autosomal dominant partial lipodystrophy. N Engl J Med 2011;364:740–748
16. Jéru I, Vatier C, Araujo-Vilar D, Vigouroux C, Lascols O. Clinical Utility Gene Card for: familial partial lipodystrophy. Eur J Hum Genet 2017;25
17. Mann JP, Savage DB. What lipodystrophies teach us about the metabolic syndrome. J Clin Invest 2019;130:4009–4021
18. Jacob KN, Garg A. Laminopathies: multisystem dystrophy syndromes. Mol Genet Metab 2006;87:289–302
19. Dewey FE, Murray MF, Overton JD, et al. Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. Science 2016;354:aaf6814
20. Cao H, Hegele RA. Nuclear lamin A/C R482Q mutation in Canadian kindreds with Dunnigan-type familial partial lipodystrophy. Hum Mol Genet 2000;9:109–112
21. Hegele RA, Anderson CM, Wang J, Jones DC, Cao H. Association between nuclear lamin A/C R482Q mutation and partial lipodystrophy with hyperinsulinemia, dyslipidemia, hypertension, and diabetes. Genome Res 2000;10:652–658
22. Speckman RA, Garg A, Du F, et al. Mutational and haplotype analyses of families with familial partial lipodystrophy (Dunnigan variety) reveal recurrent missense mutations in the globular C-terminal domain of lamin A/C. Am J Hum Genet 2000;66:1192–1198
23. Garg A, Vinaitheerthan M, Weatherall PT, Bowcock AM. Phenotypic heterogeneity in patients with familial partial lipodystrophy (dunnigan variety) related to the site of missense mutations in lamin a/c gene. J Clin Endocrinol Metab 2001;86:59–65
24. Staples J, Maxwell EK, Gosalia N, et al. Profiling and leveraging relatedness in a precision medicine cohort of 92,455 exomes. Am J Hum Genet 2018;102:874–889
25. Monajemi H, Zhang L, Li G, et al. Familial partial lipodystrophy phenotype resulting from a single-base mutation in deoxyribonucleic acid-binding domain of peroxisome proliferator-activated receptor-gamma. J Clin Endocrinol Metab 2007;92:1606–1612
26. Majithia AR, Tsuda B, Agostini M, et al.; UK Monogenic Diabetes Consortium; Myocardial Infarction Genetics Consortium; UK Congenital Lipodystrophy Consortium. Prospective functional classification of all possible missense variants in PPARG. Nat Genet 2016;48:1570–1575
27. Lek M, Karczewski KJ, Minikel EV, et al.; Exome Aggregation Consortium. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285–291
28. Hegele RA, Cao H, Huff MW, Anderson CM. LMNA R482Q mutation in partial lipodystrophy associated with reduced plasma leptin concentration. J Clin Endocrinol Metab 2000;85:3089–3093
29. Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. Genet Med 2017;19:249–255
30. Green RC, Berg JS, Grody WW, et al.; American College of Medical Genetics and Genomics. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med 2013;15:565–574
31. Schwartz MLB, McCormick CZ, Lazzeri AL, et al. A model for genome-first care: returning secondary genomic findings to participants and their healthcare providers in a large research cohort. Am J Hum Genet 2018;103:328–337