Next NGS approaches to the unsolved: Telethon Undiagnosed Program

vincenzo nigro

Prof medical genetics

Università degli Studi della Campania “Luigi Vanvitelli”, Naples, Italy
Telethon Institute of Genetics and Medicine (TIGEM), Pozzuoli, Italy
funnel of hope
recognize the disease?

The Starry Night, 1889
Vincent van Gogh
no idea
#nonmiarrendo
thousands of unknown genetic disorders

- Without name
- Without diagnosis
- Without genetic testing
- Without prognosis
- Without scientific research
- Without targeted therapy

international initiatives for undiagnosed diseases
The pediatrics and clinical genetics TUDP network
Vincenzo Leuzzi, Policlinico Umberto I, Rome
Marcella Zollino, Policlinico Gemelli, Rome
Renzo Guerrini, A.O.Uni. Meyer, Florence
Sabrina Giglio, A.O.Uni. Meyer, Florence
Maria Donati, A.O.Uni. Meyer, Florence
Corrado Romano, Oasi Maria SS., Troina
Teresa Mattina, Policlinico V. Emanuele, Catania
Giovanni Chillemi, CINECA, Bologna, Chiara Pantaleoni, Istituto Neurologico Besta, Milano

Bruno Dalla Piccola, Marco Tartaglia, Ospedale Bambino Gesù, Rome
NovaSeq6000

6,000,000,000,000 bp/40 hours
= 40 seconds/ billion DNA bp
7$/ billion DNA bp
Four different flow cells S1-S4
may be a new genetic disease?

clinical check up

sequencing and matching

data sharing validation

Patient expectations: accessibility, diagnosis, sharing, communication

genetic disease is well known and patients receive diagnosis

genetic disease can be diagnosed in specific centers

diagnosis of a mutation in a disease gene

return of results and counselling
Informed consent for analyzing whole exome/genome sequence data from all family members and for sharing deidentified results with other Undiagnosed Diseases networks

Incidental findings (known, validated actionable genetic counseling)
cases recruitment using a web form
Which responses?

- Disease is known or not genetic, an indication for a referral center is provided
- Additional information required
- Patient is recalled for deep phenotyping and priority scoring at clinical plenary meeting
standardization of patients

phenotypic data capture and storage using the standardized vocabulary defined by the Human Phenotype Ontology (HPO)
Clinical fellows meetings (14 days)

- **Aim:** monitor the flow of cases submitted to triage
- **Attendance:** clinical geneticists hired for the projects, Project Manager

Clinical plenary meetings (30 days)

- **Aim:** final selection of cases to enter the project
- **Attendance:** mandatory for clinical geneticists hired for the project and their supervisors, Project Coordinator, Project Manager, members of a Board of Independent Clinicians from the three participating Centers. The Board of Independent Clinicians can be expanded to include clinical investigators with key expertise in the cases to be discussed (selections to be made beforehand by the three clinical centers).
Clinical Plenary meeting prioritization criteria

Priority is higher

• severity and complexity
• pediatric age
• accurate negative testing
• large pedigree
• new peculiar phenotype

Priority is lower

• array-CGH uncertain
• incomplete testing
• nonspecific phenotype
• mild disease
very rare heterozygous carriers, ultra-rare patients

two mutations in the same gene !!

1/3,000

1/36,000,000 = 1-2 patients in Italy
general population, new patients

absent

one crucial \textit{de novo} mutation

absent
**Trio analysis** will be useful
- to improve variant calling and correct errors
- to discover *de novo* variants
- to get phase and haplotype

Whole exome target enrichment performed by **SureSelect Clinical Research** Exome 54Mb

sequencing 150x2 nt, av coverage 250x, 20x >95%

Proband has a **double coverage, parents are run separately**

a WES pipeline has been optimized and shared between OPBG (Rome) and TIGEM
Progress of WES and WGS quality (from GWE Santen)

2011 genome 25X

2011 exome 30x

2014 exome 60x

2014 exome 160x

2015 genome (X10)
WES CRE Agilent (>20x) in 84 subjects
bases covered

<table>
<thead>
<tr>
<th>COL6A1</th>
<th>100</th>
<th>100</th>
<th>100</th>
<th>100</th>
<th>100</th>
<th>100</th>
<th>100</th>
<th>100</th>
<th>100</th>
<th>100</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99.7</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>COL6A2</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99.6</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99.9</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99.6</td>
<td>100</td>
</tr>
<tr>
<td>COL6A3</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
NGS meetings (30 days)

- Aims: comparative NGS analyses. Discussion of genetic data.
- Attendance: mandatory for all those involved in the NGS pipelines and their supervisors, Project Coordinator.

Standardized pipeline run on CINECA cluster (20min)
Data will be mainly shared through **PhenomeCentral** ([http://phenomecentral.org](http://phenomecentral.org))

This has great potential for identifying **additional cases** of phenotypically similar patients required to validate the identified putative disease causative variant(s)

>350 users worldwide, such as the Canadian (**Care for Rare**), US (**Undiagnosed Diseases Network**) and European (**Neuromics**, **AnDDI-rare**) rare disease sequencing programs
PhenomeCentral, a repository for secure data sharing targeted at clinicians and scientists working in the rare disorder community.

The PhenomeCentral collaboration model allows to learn about the existence of similar cases world-wide, which helps to improve understanding of disorder manifestations and confirm underlying causes.

The matching algorithms -the semantic model- are sensitive to atypical phenotypic manifestations of disorders

https://www.phenomecentral.org/
<table>
<thead>
<tr>
<th></th>
<th>Patient Details</th>
<th>Genetic Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17 years ID glaucoma, corneal opacity, polycystic kidney, spleen, skin and ovarian cysts, arachnoid cysts, corpus callosum agenesis, anosmia, macrocephaly</td>
<td>de novo PLAG1, de novo PPP6R1, comp het MAP4K2</td>
</tr>
<tr>
<td>2</td>
<td>12 years intellectual disability, obsessive compulsive behaviour, hypotonia, prognathia, primary teeth, decreased body weight</td>
<td>de novo SH3BP4, de novo BUB1B</td>
</tr>
<tr>
<td>3</td>
<td>13, 11 years ID Leber congenital amaurosis, microcephaly, ataxia, bone dysplasia, femoral head and lumbar vertebral dysplasia, cerebral hypotrophy</td>
<td>No shared variants, Shared haplotype</td>
</tr>
<tr>
<td>4</td>
<td>10 years microcephaly, short stature, cerebellar atrophy, mesencephalon-synapsis, aqueductal stenosis, C4-C5 and C6-C7 defects, brother deceased</td>
<td>comp het PRG4, comp het THOC6 (LoF)</td>
</tr>
<tr>
<td>5</td>
<td>8 years severe hypotonia, l.limb hypertonia, neuro-axonal dystrophy, epilepsy, optical nerves and chiasma atrophy, obesity, microcephaly</td>
<td>de novo KIF1A C92R</td>
</tr>
</tbody>
</table>
Homozygous p.Gly46Arg, in THOC6
<table>
<thead>
<tr>
<th>Age</th>
<th>Description</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>6 years severe cognitive impairment deafness hypotonia anomaly of globus pallidus vomiting</td>
<td>de novo GRIN2B S555N</td>
</tr>
<tr>
<td>7</td>
<td>6 years tetralogy of Fallot corectopia esotropia microcephaly short stature</td>
<td>de novo SMAD4 I550V</td>
</tr>
<tr>
<td>8</td>
<td>15 years blindness spastic tetraparesis seizures lissencephaly ocular proptosis macroglossia accessory spleen conical teeth intestinal malrotation</td>
<td>autosomal recessive WDR81 W28X + R619X</td>
</tr>
<tr>
<td>9</td>
<td>20-30 years ataxia hypotonia seizures severe ID speech impairment strabismus</td>
<td>autosomal recessive GAMT A22T + A22T</td>
</tr>
<tr>
<td>10</td>
<td>5 years microcephaly pontocerebellar atrophy severe ID speech absence epileptic encephalopathy</td>
<td>autosomal recessive RARS2 fshift + -2A&gt;G</td>
</tr>
<tr>
<td>Age</td>
<td>Diagnosis</td>
<td>Genes/variants</td>
</tr>
<tr>
<td>-----</td>
<td>---------------------------------------------------------------------------</td>
<td>-----------------------------------------------------</td>
</tr>
<tr>
<td>6</td>
<td>6 years severe cognitive impairment deafness hypotonia anomaly of globus pallidus vomiting</td>
<td>de novo GRIN2B S555N</td>
</tr>
<tr>
<td>7</td>
<td>6 years tetralogy of Fallot corectopia esotropia microcephaly short stature</td>
<td>de novo SMAD4 I550V</td>
</tr>
<tr>
<td>8</td>
<td>15 years blindness spastic tetraparesis seizures lissencephaly ocular proptosis macroglossia accessory spleen conical teeth intestinal malrotation</td>
<td>autosomal recessive WDR81 W28X + R619X</td>
</tr>
<tr>
<td>9</td>
<td>20-30 years ataxia hypotonia seizures severe ID speech impairment strabismus</td>
<td>autosomal recessive GAMT A22T + A22T</td>
</tr>
<tr>
<td>10</td>
<td>5 years microcephaly pontocerebellar atrophy severe ID speech absence epileptic encephalopathy</td>
<td>autosomal recessive RARS2 fshift + -2A&gt;G</td>
</tr>
<tr>
<td>No.</td>
<td>Age</td>
<td>Symptoms</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>---------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>severe cognitive impairment, deafness, hypotonia, anomaly of globus pallidus, vomiting</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>tetralogy of Fallot, corectopia, esotropia, microcephaly, short stature</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>blindness, spastic tetraparesis, seizures, lissencephaly, ocular proptosis, macroglossia</td>
</tr>
<tr>
<td>9</td>
<td>20-30</td>
<td>ataxia, hypotonia, seizures, severe ID, speech impairment, strabismus</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>microcephaly, pontocerebellar atrophy, severe ID, speech, absence, epileptic encephalopathy</td>
</tr>
</tbody>
</table>
15 years
• Born at term by dystocic delivery (facial presentation) after uncomplicated pregnancy
• Healthy non-consanguineous parents
• At birth: weight 2.9 kg (34°pc), length 46 cm (8° pc), head circumference 28 cm (<3°pc)
• Microcephaly and facial dysmorphisms, (prominence of coronal ridge, ocular proptosis; bulbous nose, macroglossia, crowded and conical teeth)
• Generalized seizures (tonic-clonic) with neonatal onset
• Bilateral glaucoma, blindness
• Recurrent pneumonia
• Feeding problems
• Cyclic vomiting
• Spastic tetraparesis
• Profound ID, no milestone acquisition
• Abdominal US and Rx study: intestinal malrotation, accessory spleen

• Blood tests: hypercholesterolemia, ALT, AST, GGT, CPK, amilase and lipase mildly elevated

• RX skeletal at 5yrs: bowing radius, congenital hip dysplasia, cocxix agenesis

• Karyotype: 46,XX

• Array CGH: negative

• PTCN, CASK, ARX genes: no mutations

• Cerebellar and brainstem malformation NGS panel: negative
- EEG: slow ea, undifferentiated nor reactive
- Brain MRI: Lyssencephaly, corpus callosum agenesis, cerebellar and brainstem hypoplasia
- ABR and VEP: absent
New phenotype

WDR81  chr17:1639471

Maternal NM_001163673:c.84G>A:p.W28X (exon 2)
Paternal NM_001163673:c.1855C>T:p.R619X (exon 9)

WDR81 encodes a multi-domain transmembrane protein, predominantly expressed in the brain
Family may provide evolution clue

Five siblings from Turkey who walk on all fours could provide science with an insight into human evolution, researchers have said.

The four sisters and one brother could yield clues to why our ancestors made the transition from four-legged to two-legged animals, says a UK expert.

But Professor Nicholas Humphrey rejects the idea that there is a "gene" for bipedalism, or upright walking.

A BBC documentary about the family will be shown on Friday 17 March.

Professor Humphrey, from the London School of Economics (LSE), says that our own species' transition to walking on two feet must have been a more complex process that involved many changes to the skeleton and to the human genetic make-up.

However, a German group says a genetic abnormality does seem to be involved in the siblings' gait.

Coordination problem
autosomal recessive cerebellar ataxia, mental retardation, and dysequilibrium syndrome-2 with homozgyous p.P856L
<table>
<thead>
<tr>
<th>Number</th>
<th>Age</th>
<th>Diagnosis Details</th>
<th>Genotype Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>6 yrs</td>
<td>Severe cognitive impairment, deafness, hypotonia, anomaly of globus pallidus, vomiting</td>
<td>De novo GRIN2B S555N</td>
</tr>
<tr>
<td>7</td>
<td>6 yrs</td>
<td>Tetralogy of Fallot, corectopia, esotropia, microcephaly, short stature</td>
<td>De novo SMAD4 I550V</td>
</tr>
<tr>
<td>8</td>
<td>15 yrs</td>
<td>Blindness, spastic tetraparesis, seizures, lissencephaly, ocular proptosis, macroglossia, accessory spleen, conical teeth, intestinal malrotation</td>
<td>Autosomal recessive WDR81 W28X + R619X</td>
</tr>
<tr>
<td>9</td>
<td>6 yrs</td>
<td>Cerebral atrophy, skeletal anomalies, hearing loss, mild ID, dysphagia</td>
<td>De novo NOTCH3 L2137fs (ex 33)</td>
</tr>
<tr>
<td>10</td>
<td>5 yrs</td>
<td>Microcephaly, pontocerebellar atrophy, severe ID, speech absence, epileptic encephalopathy</td>
<td>Autosomal recessive RARS2 fshift + -2A&gt;G</td>
</tr>
</tbody>
</table>
#donotgiveup

de novo
NOTCH3
L2137fs (exon 33)
**TUDP, an update**

The goal: 350 undiagnosed families in 3 years

<table>
<thead>
<tr>
<th>High priority cases discussed at plenary meeting</th>
<th>263</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ongoing/complete WES analysis</td>
<td>156</td>
</tr>
</tbody>
</table>

Diagnosis success rate 51%

Among solved cases, *de novo* mutations 65%

Recessive mutations (homozygous + compound het.) 35%
solved
undetected
unrecognized
non Mendelian
If >90% of genetic cause is known, **why is the NGS diagnostic yield below 50%?**

**Undetected mutations**
- in regions with low coverage
- insertions/deletions >20-50nt
- long repeat expansions
- copy number variations
- translocations
- inversions

**Increase NGS diagnostic yield**

**Unrecognized effects of mutations**
- deep intronic or elusive splicing defects
- in enhancers and promoters
- in genes with unknown function
- in regions with duplicated sequences
- in long genes that are variable in the population
- disease gene has one/more pseudogenes

**Nonmendelian disorders**
- Digenic/polygenic
- Immune-related diseases
- environmental
**patient selection worsens:**
- poor clinical studies
- isolated cases
- mild/late onset

**knowledge improves:**
- number of new disease genes
- data sharing
- RNA studies
- functional studies
- WGS of large populations

**detection rate improves:**
- improved coverage/uniformity
- pipeline of analysis
- WGS
- new technologies for nonsubstitutions
To find information everywhere in many newspapers...
but you have paper strips to be read..
Phased sequencing (10x Genomics)

Standard sequence

hap 1 (mother allele)

hap 2 (father allele)
no connection
no phase information
ultra-exome = capture of polymorphic sequences within long introns (>30kb)
The Undiagnosed Diseases Network International (UDNI)

2017 entries:
- Sweden
- South Africa
UDNI
2018 Meeting
Naples
19-21 June