Mis-splicing and disease 1
Natural sources of errors in eukaryotic protein synthesis:

Additional mutations are introduced randomly by the environment.
Splicing mutations are found in any intron-containing gene. The frequency depends on overall length and susceptibility.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Disease/phenotype</th>
<th>Splicing/total mutation* (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATM</td>
<td>Ataxia telangiectasia</td>
<td>18</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Breast cancer predisposition</td>
<td>9</td>
</tr>
<tr>
<td>CADM</td>
<td>Medium chain acyl CoA dehydrogenase deficiency</td>
<td>10</td>
</tr>
<tr>
<td>CFTR</td>
<td>Cystic fibrosis</td>
<td>14</td>
</tr>
<tr>
<td>DMD</td>
<td>Duchenne muscular dystrophy</td>
<td>9</td>
</tr>
<tr>
<td>HBA1/2</td>
<td>Blood disorders (thalassaemias, anaemia etc)</td>
<td>3</td>
</tr>
<tr>
<td>HBB</td>
<td>Blood disorders (thalassaemias, anaemia etc)</td>
<td>10</td>
</tr>
<tr>
<td>HPRT</td>
<td>Hypoxanthine-guanine phosphoribosyltransferase 1 deficiency</td>
<td>15</td>
</tr>
<tr>
<td>IKBKAP</td>
<td>Dysautonomia, familial</td>
<td>33</td>
</tr>
<tr>
<td>MAPT</td>
<td>Frontotemporal dementia and Parkinsonism</td>
<td>33</td>
</tr>
<tr>
<td>MLH1</td>
<td>Colorectal cancer</td>
<td>18</td>
</tr>
<tr>
<td>MSH2</td>
<td>Colorectal cancer</td>
<td>9</td>
</tr>
<tr>
<td>NF1</td>
<td>Neurofibromatosis type 1</td>
<td>19</td>
</tr>
<tr>
<td>NF2</td>
<td>Neurofibromatosis type 2</td>
<td>22</td>
</tr>
<tr>
<td>RHO</td>
<td>Retinitis pigmentosa</td>
<td>3</td>
</tr>
<tr>
<td>SMN1/2</td>
<td>Spinal muscular atrophy</td>
<td>4</td>
</tr>
<tr>
<td>WT1</td>
<td>Wilms tumour</td>
<td>11</td>
</tr>
</tbody>
</table>

*Data calculated from the public Human Gene Mutation Database (24/11/2008).
Experimental approaches in the study of alternative splicing and disease can be divided in two major classes:

Detection and validation of splicing mutations:

- Direct amplification of spliced transcripts.
- Minigene splicing systems.
- In vitro splicing systems.

Experimental methods to define the functional effects of the mutation:

- RNA-protein interactions.
- snRNP-RNA interactions.
- RNA secondary structures.
Classical disease mutations

Basic trans-acting factor(s) bind to splicing regulatory sequences at the acceptor, donor, and branch-site regions.

Nucleotide substitutions can change the binding properties of these basic sequences both directly or through changes in RNA secondary structure.

Nucleotide substitutions can further regulate trans-acting factor binding by determining creation or loss of intronic/exonic SRE elements.

External stimuli, spliceosomal rearrangements due to processing of upstream introns/exons, RNA Pol II processivity, and gene architecture can further modify the outcome.

Splicing mutations

Inclusion

Skipping

Altered inclusion levels

Disease (hydrops fetalis)
Types of splicing mutations

5’ss, 3’ss or SRE inactivating mutation

de novo splicing site creation within introns

de novo 5’ss

cryptic 3’ss

cryptic 5’ss

Exon skipping (single or multiple)

cryptic site activation

Full intron retention

Fig. 1
What is the involvement of splicing mutations in human disease?

**In general,** in a survey where the mutations considered consisted ONLY of those directly affecting the standard consensus splice sequences 15% of point mutations resulted in human disease.

**In the NF-1 and ATM genes,** where analysis was performed both at the RNA and DNA level almost 50% of the patients were found to have disease due to mutations that resulted in aberrant splicing.
How can splicing alterations then be detected?
(functional splicing assays)
Total RNA is extracted from patient tissue and splicing products are analyzed by RT-PCR/Northern blot.
Minigenes

Advantages:
1) the RNA is processed “in vivo”, taking into account the complexity of the cellular environment.
2) is recommended to analyze complex exonic/intronic regions.

Drawbacks:
1) it does not provide (easily) information regarding splicing intermediates or kinetics.
2) it is rather difficult to act on the cellular environment and remove/replace selected factors.

In vitro splicing

Advantages:
1) splicing intermediates can be seen and kinetics followed easily.
2) splicing factors can be removed and replaced (also with mutants) to assess exactly their activity.

Drawbacks:
1) the RNA is transcribed “in vitro” before the splicing reaction.
2) there is a limit to the length of the pre-mRNAs that can be used.
Northern/RT-PCR direct detection

Total RNA is extracted from patient tissue and splicing products are analyzed by RT-PCR/Northern blot.
Direct detection:

Advantages:

1) the RNA is analyzed directly from the patient.

Drawbacks:

1) patients or RNA samples from affected tissues are not always available.

2) the gene of interest may not be expressed in easily available tissues.

3) often, unless the suspected splicing mutation is present in both alleles, care must be taken when evaluating results.
Minigene systems

CMV/Globin promoter

Poly-A signal

Transfected in Eukaryotic cells for 24-48 hours.

Total RNA is extracted and splicing products are analyzed by RT-PCR.

Minigenes
The minigene system is based on a plasmid constructs that contain all the elements required for the formation of a “spliceable” mRNA.

Hybrid minigenes transfection in different cell types

Cotransfection with plasmids codifying for splicing factors

0.5 μg DNA 300,000 cells

24 hours

RT-PCR analysis of splicing products using specific oligonucleotides
A PATHOLOGIC VARIATION OR A POLYMORPHISM?

DEFINITION OF A NF1 DISEASE CAUSING MUTATION BY A FUNCTIONAL SPLICING ASSAY
G > C causes Exon 3 aberrant splicing

Transfection of hybrid minigenes in different cell lines

0.5ug DNA
250,000 cells

Cotransfection with splicing factors coding plasmids

24hr

RT-PCR analysis of splicing products with specific primers

NF1Ex 3
NF1Ex 3 G>C control
Exon 3 WT

TCTTGCTGGG/gtaatgtaaa

Exon 3 +5G>C

TCTTGCTGGG/gtaactaaa

-4-3-2-1 +1+2+3+4+5

NF1 ex 3 (wt)  NF1 ex 3 (+5 G>C)  NF1 ex 3 (+5 G>C)

344bp
298bp
220bp

U1snRNP (C>G)

NF-1 wt

NF-1 ex3 +5G>C

U1snRNA  ggUCCAUCAUA

● base-pair match
1. In vitro transcription of RNA and treatment with sodium-m-periodate

2. Binding to adipic acid dehydrazide beads

3. Incubation with protein extract

4. Repeated cycles of centrifugation and washing

SDS-PAGE analysis (and Coomassie staining)

Nanospray Mass Spectrometric Analysis
Pull down analysis of the wt and mutated (+5G>C) exon 3 sequence revealed a great variety of protein complexes of potential interest. Beside U1snRNP protein signatures (which disappear in the mutant) and the KSRP and p54nrb proteins (which have been characterized in independent studies on splicing complex assembly) the presence of hnRNP H is particularly interesting.

<table>
<thead>
<tr>
<th>No.</th>
<th>Protein Name</th>
<th>Molecular Weight (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P/Q Splicing Factor (P23246)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Far Upstream Bdg Protein (Q96AE4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mc1: KSRP (O00301)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mc2: MLL septin-like fusion (Q9Y5W4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mc3: hnRNP M (P52272)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>U1-70K (P08621)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>p54nrb (Q15233)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mc1: Paraspeckle-Protein (Q8WXF1)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>hnRNP H (P31943/P55795)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mc1: Nuclear RNA helicase (O00148)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mc2: hypothetical protein FJL10849</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>hnRNP D-like (O14979)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mc1: hnRNP B1 (P22626)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mc2: hnRNP A3 (P51991)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>U1-A (P09012)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>SmRNP B/B1 (P14678)</td>
<td></td>
</tr>
</tbody>
</table>
NF-1 exon 3 (wt)  UCUUGCUGGGquauaquaa  U1 SnRNA ggUCCAuuUCAUA  

NF-1 exon 3 (+5G>C)  UCUUGCUGGquauaquaa

NF-1 exon 3 (-2G>A)  UCUUGCUGGquauaquaa

NF-1 exon 3 (-2G>A,+5G>C)  UCUUGCUGGquauaquaa  

Base-pair match
NF-1 exon 7 donor site also carries a c in +5. This donor site can also be made to bind hnRNP H by introducing a A>G mutation in position -2 and this correlates with loss of exon recognition:

Exon 7 WT  \text{CUUAAG} \text{guaacaugc}

Exon 7 -2A>G  \text{CUUAGG} \text{guaacaugc}
However, siRNA knock-down of hnRNP H is not sufficient to recover splicing activity.

Are there additional factors involved that can compensate for hnRNP H absence?
Correlation between the number of Gs in positions -4 to +1 and conservation on the intronic G nucleotide in position +5 as calculated in 23,461 donor sites (AltExtron database) available at http://www.ebi.ac.uk/asd/altertron/index.html.
The long QT syndrome (LQTS) is a heart condition associated with longer recovery time following depolarization (excitation) of the cardiac ventricles. It is associated with fainting and sudden death. Hereditary long QT syndrome (LQTS) is caused by over 250 mutations in five genes:

- KCNQ1 (KVLQT1, LQT1)
- KCNH2 (HERG, LQT2)
- KCNE1 (mink, LQT5)
- KCNE2 (MiRP1, LQT6)
- SCN5A (LQT3)

<table>
<thead>
<tr>
<th>Normal QT (390-410 msec)</th>
<th>Prolonged QT (&gt; 440 msec)</th>
</tr>
</thead>
</table>

Genes involved in Long QT syndrome:

- KCNQ1 (KVLQT1, LQT1)
- KCNH2 (HERG, LQT2)
- KCNE1 (mink, LQT5)
- KCNE2 (MiRP1, LQT6)
- SCN5A (LQT3)

We have studied a patient that presented an IVS7+6T>C mutation in the HERG gene.
KCNH2 Exon 7

ATCTGCCTCATGCTCATTTGGCT

gtgagtgtggccccaggggccccgg

KCNH2 IVS 7

\[
\times
\]

C

[Genetic diagram with symbols indicating family members and genetic status]
One of the most important events in 5’ splice site definition is represented by base pairing of the U1snRNA component of U1snRNP with the 5’ splice site consensus sequence.

A “T” in the +5 position is rather loosely conserved:

```
-3  -2  -1  +1  +2  +3  +4  +5  +6
A 32  56  8  0  0  38  70  5  13
C 38  15  4  0  0  4  9  5  21
G 19  15  80 100 0  56 14  86  25
U 11  14  8  0  100 2  7  4  41
```

C  A  G  G  U  R  A  G  U

Is a +6T>C change capable of affecting splicing?
The +6T>C is a mutation that perturbs the snRNP U1 interaction with the 5’ss and in a minigene system causes intron retention.
5’ splice site complexes are tailor made

Mutations that disrupt U1 snRNP binding to 5’ss

NF1 exon3

HERG exon7

Exon skipping

Intron retention
# Classical and non classical CF

<table>
<thead>
<tr>
<th>Classic cystic fibrosis (no functional CFTR protein)</th>
<th>Nonclassical cystic fibrosis (some functional CFTR protein, providing survival advantage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic sinusitis</td>
<td>Chronic sinusitis</td>
</tr>
<tr>
<td>Severe chronic bacterial infection of airways</td>
<td>Chronic bacterial infection of airways (later onset, but variable)</td>
</tr>
<tr>
<td>Severe hepatobiliary disease (5–10% of cases)</td>
<td>Adequate pancreatic exocrine function (usually); pancreatitis (5–20% of cases)</td>
</tr>
<tr>
<td>Pancreatic exocrine insufficiency</td>
<td></td>
</tr>
<tr>
<td>Meconium ileus at birth (15–20% of cases)</td>
<td></td>
</tr>
<tr>
<td>Sweat chloride value usually 90–110 mmol/liter; sometimes 60–90 mmol/liter</td>
<td>Sweat chloride value usually 60–90 mmol/liter; sometimes normal (&lt;40 mmol/liter)</td>
</tr>
<tr>
<td>Obstructive azoospermia</td>
<td>Obstructive azoospermia</td>
</tr>
</tbody>
</table>
hCFTR exon 9 region present traces of an ancient transposon hit. This region is also amplified and over 20 copies are spread through the genome, some of them are transcriptionally active

\[ mCFTR \]

Exon 8  |  2.2kb |  Exon 9  |  6kb  |  Exon 10

\[ hCFTR \]

Exon 8  |  6.5kb  |  Exon 9  |  10.6kb  |  Exon 10
Skipping of exon 9 is associated to monosymptomatic forms of CF (Congenital Bilateral Aplasia Vas Deferens, bronchiectasia, pancreatitis) and produce a non functional protein

Presence of polymorphic variants in humans at the polypirimidine tract at the 3’ end of intron 8 (TG repeats and T9, T7 T5)

Not evolutionary conserved

Amplification of exon 9 related sequences in the human genome (Rozmahel et al 1997)
EXON 9 WITH THE INTRONIC SEQUENCES (IVS8 AND IVS9)

1. AN ENHANCER IN EXON 9
2. A SILENCER IN EXON 9
3. AN INTRONIC SILENCER IN IVS9
4. A POLYMORPHIC TRACT (TG)m(T)n IN IVS8

TTTTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG...
CFTR gene variants associated with different phenotypic expression

**CFTR function**

- Normal
- Non-classical CF
- Cystic Fibrosis

**Clinical phenotype**

- Normal
- Non-classical CF
- Cystic Fibrosis

**CFTR genotype**

- ex9
  - G424S
  - V456F
  - I444S
  - A455E
  - Q414X
  - N418S
  - Q542P
  - D443Y

- IVS9
  - T9
  - T7
  - T5

- T3
Minigene system applied to CFTR exon 9 splicing allows to replicate the effects of the UGmUn polymorphisms observed in patients affected by monosymptomatic forms of CF (CVABD).

Niksic et al. HMG, 1999, 8:2339-2349.
Identification of RNA binding proteins by UV-crosslinking

UV-crosslinking assay

1. In vitro transcription of labelled RNA

2. Incubation with protein extract

3. UV crosslinking

4. RNase digestion

SDS-PAGE analysis + autoradiography
The protein was identified as TDP-43 and was shown to be responsible for the inhibition of CFTR exon 9 splicing.

Buratti et al. EMBO J 2001, Buratti and Baralle, JBC 2001
TAR DNA Binding Protein (TDP 43) is a splicing factor of the hnRNP family that plays a role in many aspects of RNA metabolism. Its aggregation/dysfunction is central to ALS and FTLD pathogenesis.

The TDP-43 protein

TDP 43 is structurally and functionally conserved from fly to man
The first well documented function of TDP-43 was the inhibition of CFTR exon 9 splicing

Buratti et al. EMBO J 2001, Buratti and Baralle, JBC 2001
TDP 43 controls its cellular levels by a self regulation mechanism.
Transgene TDP-43 expression in stably transfected HEK293 cell lines cause reduction in endogenous TDP 43 gene expression

TDP-43 mRNA species in stably transfected HEK293 cell lines are downregulated following transgene TDP 43 cDNA induction:

TDP-43 RNA binding function is essential for self regulation
The splicing process in the 3’UTR may be recognized by the cell as anomalous and marked as an intron incompletely processed that must be degraded. Why the RNA processing is blocked?

The same mRNA displays different cellular distribution according to its biosynthesis process: transcription only or transcription followed by splicing.
A wide variety of neurodegenerative diseases are characterized by the accumulation of intracellular or extracellular protein aggregates.

Ubiquitinated, misfolded and hyper phosphorylated TDP 43 was identified as the major component of the pathological inclusions found in the brain of FTLD and ALS patients

Neumann et al., 2006 Science 314: 130-136
GAIN AND LOSS OF FUNCTION MODELS

Gain of function model

Residual nuclear TDP

Absence nuclear TDP

Neuronal death/inactivation

Gain of function model

Loss of function model

Toxic aggregates

Lack of TDP-43 function

Adapted from Chen-Plotkin et al, 2010
How similar are flies and humans?
TBPH is the *Drosophila melanogaster* orthologue of TDP-43

**High degree of functional conservation**

- Preferential binding to (UG)n rich sequences
- Can replace TDP 43 in vivo and viceversa
- TDP 43 and TBPH interact with the same set of proteins such as hnRNP 38-hnRNP A1/A2

**Drosophila models to study TDP-43-dependant neurodegeneration**
Deletion of the Drosophila homologue of human TDP-43 (TBPH) leads to a paralytic phenotype

But cannot get out of the pupal cage without external help, are deficient in locomotion and have a very short life span.

Flies apparently develop normally.

Feiguin et al. 2009 FEBS Lett
Expression of Human TDP-43 in neural cells can rescue fly motility.
1- Drosophila TDP-43 in vivo has a short half life and its synthesis is permanently required for larval motility and for NMJs assembly.

2- Late rescue of Drosophila TDP-43 function recovers NMJ’s structure.

3- Acute silencing of TDP-43 in the adult fly down-regulates specific presynaptic proteins and affects muscle innervation.

4- Drosophila TDP-43 binds Syntaxin mRNA and down-regulates its intracellular levels (40%), other synaptic proteins/functions are also affected.

G. Romano et al NBD
A hypothetical approach to gene therapy of TDP 43 aggregation proteinopathies could be the design of TDP 43 molecules that can avoid being sequestered by the aggregates but are fully functional in the splicing process. We have tested this approach using our knowledge of the structural features critical for aggregation
Structural definition of TDP-43 N-terminal domain

WT 31V-32T, Captures endogenous TDP 43

31V/R-32T/R, Does not capture endogenous TDP 43

MSEYIRVTEDEPIEIPSEDDGTVLLSTVTAAQFGACGLRYRNPVSQCMRGVRLVEGILHAPDAGWGNLNYVYNYPKDNKRKMDETDAASAVKVKRAV

TDP-43 N-terminal domain

31V/R-32T/R
NTD structural integrity is essential for endogenous TDP-43 entrapment

<table>
<thead>
<tr>
<th>TET</th>
<th>TDP12X</th>
<th>TDP12X-31V/R-32T/R</th>
<th>POLDIP3 RT-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>+</td>
<td></td>
<td>exon 3 inclusion</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>exon 3 skipping</td>
</tr>
</tbody>
</table>

**WB anti-flag**

<table>
<thead>
<tr>
<th>TDP12X</th>
<th>TDP12X-31V/R-32T/R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-TDP-43</td>
<td>Anti-TDP-43</td>
</tr>
<tr>
<td>- TET</td>
<td>+ TET</td>
</tr>
<tr>
<td>+ TET</td>
<td>+ TET</td>
</tr>
<tr>
<td>+ TET</td>
<td>+ TET</td>
</tr>
</tbody>
</table>

Romano et al HMG 2014, Romano et al submitted 2016
Defining the C-terminal Q/N rich region involved in aggregation

TDP-43 C-terminal domain

SNAEPKHSNRQLERSGRGFNNPGGFGNNQGGFNSRGGGAGLGGNNQGSNMGGGMNFHAFSINPAMMAAAQAALQASSWGMMSMGLASQQQNGPSGNNQNQGNNMQR
EPNQAFGSNGNSSYSGNSSGAAGWGASNASNSAGGSFGNGFGSSMDKSSGWGM

Flag-TDP-WT Flag-TDP-Δ342-345

GFP-F4L 12xQ/N - + - + - +
CFTR + + + + + +
siTDP + + + + + +

exon 9 inclusion
exon 9 skipping

Merge (GFP/anti-Flag/TO-PRO3)

CFTR exon 9 RT-PCR

Flag-TDP-WT Flag-TDP-Δ342-345
Climbing assay elav 12xQ/N
**siRNA TDP-43**

**POLDIP3 splicing**

<table>
<thead>
<tr>
<th>sTDP</th>
<th>TET</th>
<th>sTDP</th>
<th>TET</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

**Top scoring (%)**

- W1118
- bph102, elavG4/+
- UAS-Dcr-2, TBPH-RNAi/UAS-GFP

NBD, 2014

---

**EGFP-12xQ/N model**

**POLDIP3 splicing**

<table>
<thead>
<tr>
<th>EGFP/α-TDP-43</th>
</tr>
</thead>
<tbody>
<tr>
<td>TET - +</td>
</tr>
</tbody>
</table>

**Top scoring (%)**

- ELAV-Ga4/+
- UAS-EGFP/+ 

ELAV-Ga4/+
- UAS-EGFP-12xQ/N+

---

**TDPΔ1-ΔC-RRM2F/L-12xQ/N**

**POLDIP3 splicing**

| TET - + |

**Top scoring (%)**

- elav-Ga4+/+
- elav-Ga4/UAS_Egfp
- elav-Ga4/UAS_2B
- elav-Ga4/UAS_5A

JBC, 2011; HMG, 2014
TDP 43 levels drop during aging, this phenomena correlates with the onset of the locomotion defect.

Reduction of TDP 43 levels with aging in mouse

**Brain**

<table>
<thead>
<tr>
<th>Mouse ID#</th>
<th>Age (Days)</th>
<th>Days</th>
<th>Reduction of TDP 43 levels drop during aging, this phenomena correlates with the onset of the locomotion defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10, 90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10, 90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10, 90</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Liver**

<table>
<thead>
<tr>
<th>Mouse ID#</th>
<th>Age (Days)</th>
<th>Days</th>
<th>Reduction of TDP 43 levels drop during aging, this phenomena correlates with the onset of the locomotion defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10, 90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10, 90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10, 90</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Skeletal muscle**

<table>
<thead>
<tr>
<th>Mouse ID#</th>
<th>Age (Days)</th>
<th>Days</th>
<th>Reduction of TDP 43 levels drop during aging, this phenomena correlates with the onset of the locomotion defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10, 90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10, 90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10, 90</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Cragnaz et al Neuroscience 2015**

De Conti et al submitted 2016
Defective locomotion phenotype is linked to a physiological, evolutionary conserved and age related fall in brain TBH level

Aging/reduction in TDP43 transcription

Efficient clearing system

Persistent aggregation

Newly produced TDP-43 will be less sequestered
chronic stress or other factors

Newly synthetized TDP-43

TDP-43

TDP 43

aggregation

TDP-43

transcript

Newly synthetised TDP-43

self regulation of cellular levels

recovery of neurological phenotype

degradation of aggregates

neurological phenotype

chronic stress or other factors

TDP-43 aggregation

Newly synthetised TDP-43 sequestration

age related decay

levels
Acknowledgments

Simona Langellotti, Chiara Chiavelli, Maureen Akinyi, Jeremias Herzog

Doug Laurens
Madrid

Fabian Feiguin
Raffaella Klima
Giulia Romano

Valentina Romano/Lucia Cragnaz

Marco Baralle/Laura De Conti

Emanuele Buratti
Finding drugs to clear TDP-43 aggregates

One possible approach could be to enhance degradation of the aggregated protein.

Soluble monomeric → Oligomeric → Aggregate = “TDP-43 sink”

Newly produced TDP-43

Will be less sequestered
Selective screen for compounds that enhance aggregate clearance in our EGFP 12Q/N and TDPΔ1-ΔC-RRM2F/L-12xQ/ models

<table>
<thead>
<tr>
<th>Compound</th>
<th>Drug class</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desipramine</td>
<td>Antidepressant</td>
<td><img src="image1.png" alt="Structure" /></td>
</tr>
<tr>
<td>Nortriptyline</td>
<td>Antidepressant</td>
<td><img src="image2.png" alt="Structure" /></td>
</tr>
<tr>
<td>Clomipramine</td>
<td>Antidepressant</td>
<td><img src="image3.png" alt="Structure" /></td>
</tr>
<tr>
<td>Promethazine</td>
<td>Antihistaminic</td>
<td><img src="image4.png" alt="Structure" /></td>
</tr>
<tr>
<td>Fluphenazine</td>
<td>Antihistaminic</td>
<td><img src="image5.png" alt="Structure" /></td>
</tr>
<tr>
<td>Thioridazine</td>
<td>Antipsychotic</td>
<td><img src="image6.png" alt="Structure" /></td>
</tr>
<tr>
<td>Chlorpromazine</td>
<td>Antipsychotic</td>
<td><img src="image7.png" alt="Structure" /></td>
</tr>
<tr>
<td>Cyclobenzaprine</td>
<td>Muscle relaxant</td>
<td><img src="image8.png" alt="Structure" /></td>
</tr>
</tbody>
</table>

- Structurally related tricyclic compounds have similar actions in autophagy induction (Tsvetkov et al 2010)
- FDA approved drugs
- Can pass the BBB
- Have being shown to be neuroprotective
Nortriptyline ameliorates GFP-12xQ/N Drosophila phenotype

Protocol

Start feeding with compound at Day 1

ELAV-Gal4/+; UAS-EGFP-12xQ/N/+ Day 8 untreated

Test the flies for climbing

Day 8 untreated

Day 8 Nor 7mM

Protocol Diagram:

- ELAV-Gal4/+; UAS-EGFP/+ (gray)
- ELAV-Gal4/+; UAS-EGFP-12xQ/N/+ (black)

Graphs:
- Day 8 untreated: ***
- Day 8 Nor 7mM: ns
Acknowledgments

Frederick Allain
Peter Lukavsky
Zurich CH

Fabian Feiguin
Raffaella Klima
Giulia Romano

Simona Langellotti
Maurizio Romano
Cristiana Stuani
Jeremias Herzog

Valentina Romano/Lucia Cragnaz

Marco Baralle/Laura De Conti

Emanuele Buratti