

# SCIENTIFIC REPORTS



OPEN

## Disease-causing mutations in subunits of OXPHOS complex I affect certain physical interactions

Gilad Barshad<sup>1</sup>, Nicol Zlotnikov-Poznianski<sup>1</sup>, Lihi Gal<sup>2</sup>, Maya Schuldiner<sup>1</sup> & Dan Mishmar<sup>1</sup>

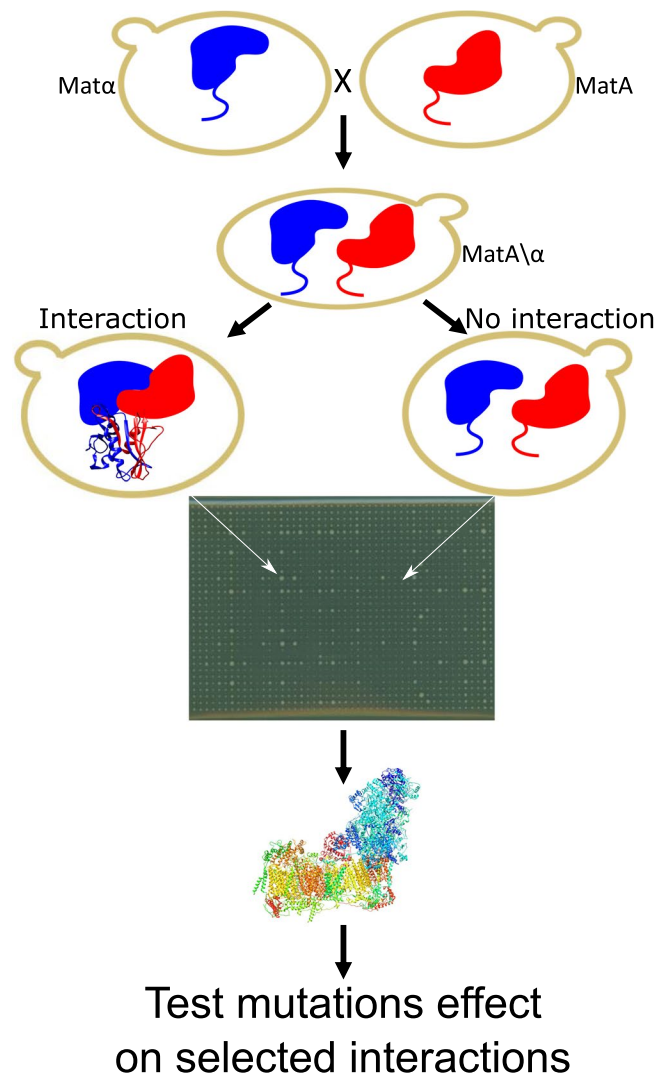
Mitochondrial complex I (CI) is the largest multi-subunit oxidative phosphorylation (OXPHOS) protein complex. Recent availability of a high-resolution human CI structure, and from two non-human mammals, enabled predicting the impact of mutations on interactions involving each of the 44 CI subunits. However, experimentally assessing the impact of the predicted interactions requires an easy and high-throughput method. Here, we created such a platform by cloning all 37 nuclear DNA (nDNA) and 7 mitochondrial DNA (mtDNA)-encoded human CI subunits into yeast expression vectors to serve as both 'prey' and 'bait' in the split murine dihydrofolate reductase (mDHFR) protein complementation assay (PCA). We first demonstrated the capacity of this approach and then used it to examine reported pathological OXPHOS CI mutations that occur at subunit interaction interfaces. Our results indicate that a pathological frame-shift mutation in the *MT-ND2* gene, causing the replacement of 126 C-terminal residues by a stretch of only 30 amino acids, resulted in loss of specificity in ND2-based interactions involving these residues. Hence, the split mDHFR PCA is a powerful assay for assessing the impact of disease-causing mutations on pairwise protein-protein interactions in the context of a large protein complex, thus offering a possible mechanistic explanation for the underlying pathogenicity.

The mitochondrial oxidative phosphorylation (OXPHOS) system is the main ATP production machinery in eukaryotic cells. The OXPHOS system comprises five multi-subunit protein complexes, of which NADH-ubiquinone oxidoreductase (complex I, CI) is a major electron entry point into the electron transport chain (ETC) that is central to mitochondrial ATP synthesis. CI dysfunction leads to devastating mitochondrial diseases<sup>1</sup> and is a hallmark of mitochondrial dysfunction in Parkinson's disease<sup>2,3</sup>, dementia<sup>4</sup> and the renal cancer oncocyoma<sup>5,6</sup>. Moreover, CI subunits are the most frequent gene targets for mitochondrial disease-causing mutations (reviewed by Koopman, *et al.*<sup>7</sup>), some of which affect CI assembly and stability in cultured cells and in patients<sup>8,9</sup>. Hence, deciphering the molecular basis of CI dysfunction is of major interest.

Mammalian CI is the largest of the OXPHOS complexes (~1 MDa), comprising 44 subunits<sup>10,11</sup>. These include 14 core subunits conserved from bacteria to man<sup>12</sup>, and 30 'super-numerary' (accessory) subunits that were recruited during the course of evolution after the radiation of eukaryotes. Recently, high-resolution structural models of complex I from ovine<sup>13</sup> and bovine<sup>14</sup> hearts became available, as well as from human embryonic kidney cell line (HEK293)<sup>15</sup>. Additionally, a deep view into the assembly process and assembly intermediates of CI were recently reported<sup>16</sup>. These studies enabled better understanding of the network of CI subunit interactions. Nevertheless, most studies and experimental approaches taken to study CI subunit interactions focused either on the holoenzyme or on assembly intermediates<sup>8,10,17–21</sup>. As such, these studies allowed for assessing the impact of CI pathological mutations only on the assembly process or on the overall function of CI in certain cells<sup>22</sup>, yet overlooked the possible impacts of mutations on specific subunit interactions. For this, a high-throughput screening method for considering all CI pairwise interactions, is needed.

To create such an experimental platform, we cloned the entire set of human CI protein subunits (N = 44) into the two complementary fragments of murine dihydrofolate reductase (mDHFR) and performed a split-mDHFR protein complementation assay (PCA) in the yeast *Saccharomyces cerevisiae*<sup>23,24</sup>. This experimental system enabled the detection of pairwise interactions between CI subunits in a rigorous and experimentally facile manner. This strategy, moreover, provided a robust method to assess the impact of mutations on such interactions. To demonstrate the power of this approach, we first identified subunit interactions supported by available

<sup>1</sup>Department of Life Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel. <sup>2</sup>Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel. Correspondence and requests for materials should be addressed to D.M. (email: [dmishmar@bgu.ac.il](mailto:dmishmar@bgu.ac.il))



**Figure 1.** Workflow of the split-mDHFR-based screen for testing pairwise interactions between complex I subunits. Upper panel shows the construction of the haploid yeast split mDHFR libraries which harbor complex I subunit sequences, followed by mating. This allows to identify candidate interacting proteins, which lead to increase in yeast colony size after seeding on MTX+ plates. Then, the screen results were compared to published structural data to identify testable interactions. Notably, mammalian CI structure shown is based on (PDB ID: 5XTD)<sup>15</sup>.

mammalian CI structural models<sup>13–15</sup> and screened for pathological mutations that localize to subunit interaction interfaces. We then chose three changes, namely two SNPs and one frame shift mutation, all associated with mitochondrial dysfunction, and experimentally tested their effects. Our results indicate that a frameshift mutation in the mtDNA-encoded ND2 subunit, which underlies CI dysfunction in a patient with severe exercise intolerance<sup>25</sup>, led to a loss of specificity of ND2 subunit interactions. Our results provide proof-of-concept supporting the use of our experimental system as a resource to assess the impact of mutations on diverse protein-protein interactions in human CI.

## Results

**The split mDHFR protein complementation assay offers a robust platform for following protein-protein interactions in human complex I.** To assess pairwise interactions between CI subunits, we fused the coding sequences (CDSs) of all nuclear DNA (nDNA)- and recoded mitochondrial DNA (mtDNA)-encoded human complex I subunits (N = 44) to the 5' end of that part of mDHFR gene encoding a C-terminal fragment (termed the CI-F3 constructs). In addition, 37 CI subunits were fused to the 5' end of that part of mDHFR gene encoding a N-terminal fragment (termed the CI-F1,2 constructs) (Fig. 1). Notably, unlike the traditional yeast-two-hybrid system which utilizes a split transcription factor, the split mDHFR system was chosen as it is better adapted to screen for interactions between membrane-bound proteins<sup>26</sup>. Secondly, since the baker's yeast lost the multi-subunit CI during evolution our assay is less prone to false negative results that could stem from interactions between our human proteins with endogenous yeast proteins. In our hands, 8 subunits were not

successfully cloned into the latter constructs (i.e., ND6, NDUFA9, NDUFA12, NDUFB9, NDUF51, NDUF54). While plasmids containing the CI-F3 constructs were inserted into MAT $\alpha$  yeast, those plasmids containing the CI-F1,2 constructs were transformed into MAT $\alpha$  yeast strains (Fig. 1, Supplementary Table S1) to enable rapid mating and the creation of diploids. The coding sequences of the mtDNA-encoded subunits (i.e., ND1-ND6 and ND4L) were edited to comply with yeast cytosolic codon usage. The construct harboring the mDHFR F1,2 fragment also harbored a YFP reporter gene inserted between the CI subunit sequence and the F1,2 fragment; this enabled tracing intracellular localization of the resultant fusion protein in yeast. Approximately 23 out of the 37 human CI subunits cloned in this plasmid showed mitochondrial associated signals (Supplementary Material, Fig. S1). Five additional subunits were localized to membrane-embedded organelles such as the endoplasmic reticulum or the plasma membrane (Supplementary Material, Fig. S1). These results suggest that although the experiment was performed in the baker's yeast which totally lacks an endogenous CI, most human complex I subunits retained their mitochondrial localization, or membrane association, thus supporting the potential physiological relevance of our experiments. Notably, to avoid false positive results we focused our analysis only on interactions that (A) share the same/adjacent subcellular compartment, and (B) are consistent with the published mammalian complex I structures. To further reduce false-positive interactions, we excluded two highly over-expressed subunits (i.e., NDUFA2 and NDUFA5) from further analyses. To assess interactions between the various subunits, the haploid yeast strains were mated in a high-throughput manner (see Materials and Methods), thus creating 1,628 independent CI-F1,2/CI-F3 diploid strains. The resultant colony sizes of these diploid strains were measured as a proxy for growth rate in the presence of methotrexate (MTX), thus selecting for yeast growth only when complementation of mDHFR activity had occurred (Fig. 1). This experiment revealed 94 significantly positive (feasible) pairwise interactions (Fig. 2). In seeking candidate subunit interactions able to report on the impact of pathological mutations, we focused on 19 interactions that were consistent with the CI structural models from bovine and ovine heart<sup>13,14</sup> (Fig. 1). Such interactions were also consistent with recently published human CI structures<sup>15</sup>.

### Two disease-associated point mutations, affecting residues at the interaction interface of ND2 and ND5, had no effect on the pairwise subunit interaction.

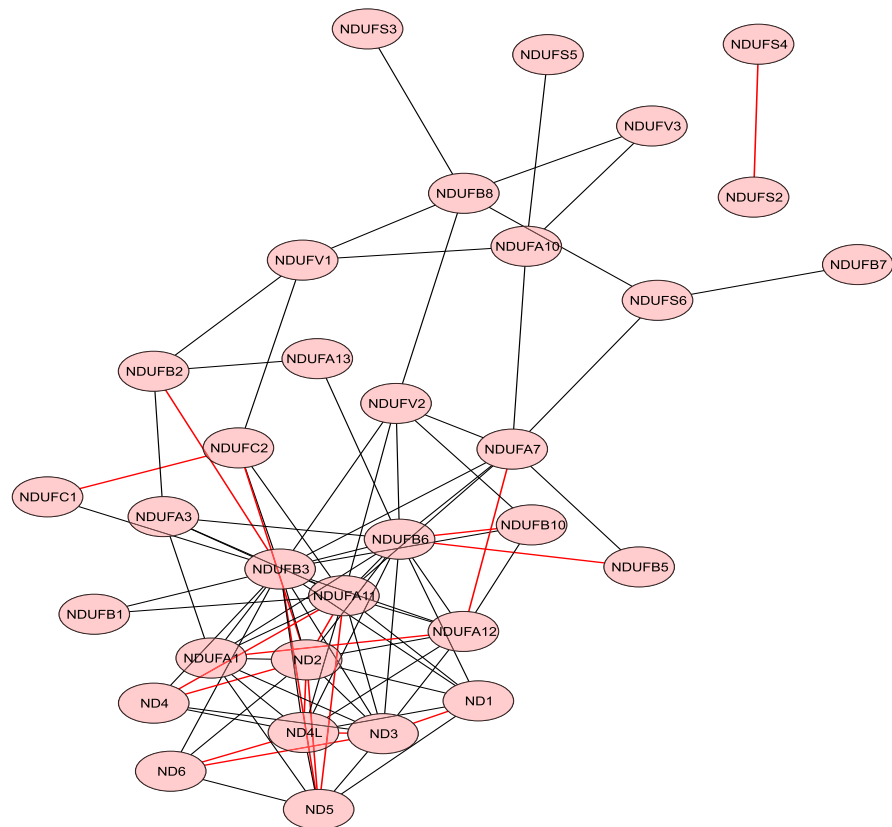
Next, we employed our assay to assess the impact of disease-associated human gene variants on physical interactions between the encoded proteins. First, we screened the literature for CI disease-associated mutations that affect residues found at physical interaction interfaces of two CI subunits, according to the available structures (Table 1, Supplementary File S1)<sup>27,28</sup>. Two missense mutations at nucleotide positions 4917G and 4640A are associated with Leber's hereditary optic neuropathy (LHON) in separate families<sup>27,29</sup>. Both of the resulting amino acid substitution (i.e. N150D and I57M, respectively) occur in ND2 trans-membrane  $\alpha$ -helixes (TMHs), which in the structure are adjacent to a ND5 subunit TMH, and hence potentially alter the interaction interface (Fig. 3A). Furthermore, the 4460A mutation was found in a Russian family from Novosibirsk who exhibited poor oxygen utilization in cybrids<sup>29</sup>. The 4917G nucleotide substitution, an ancient mutation that defines the mtDNA genetic background haplogroup T, was associated with LHON only in the presence of the R2-JT haplogroup cluster-defining 4216C mutation<sup>27</sup>. We introduced these two SNPs either separately or together into the ND2 re-coded sequence within the F3 mDHFR fragment-containing construct. We then mated the MAT $\alpha$  yeast strains harboring these mutant constructs with a MAT $\alpha$  strain harboring the ND5-F1,2 construct and conducted a dilution drop-test to assess the impact of these mutations on the ND2-ND5 interaction. Neither of the two single mutants showed any significant change in terms of yeast growth on selective media (containing MTX), suggesting that no significant impact on the ND2-ND5 interaction had occurred (Fig. 3B). Notably, this result does not reflect reduced resolution of our protein complementation assay, as it was previously successfully used to assess the impact of point mutations on protein-protein interactions<sup>30</sup>. In addition, the double mutant construct was not expressed to a detectable level in yeast, thus precluding any analysis of this combination of mutations (Fig. 3C, Supplementary Material Fig. S2).

### A two base-pair deletion causing an *MT-ND2* frame shift mutation leads to the emergence of interaction promiscuity.

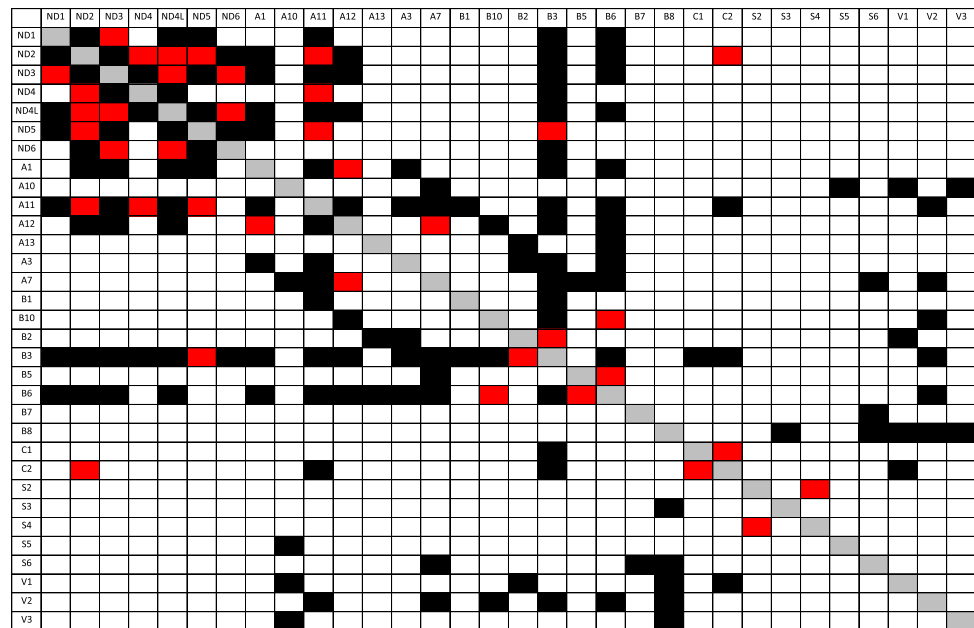
We next examined the impact of a double adenine deletion within an adenine triplet at mtDNA nucleotide positions 5132 to 5134, causing a frame shift in the *MT-ND2* sequence. This deletion was reported in a 28-year-old patient with mitochondrial myopathy and severe exercise intolerance<sup>31</sup>. Such frameshift led to replacement of the 126 C-terminal residues harboring three TMHs in ND2 by 30 highly hydrophilic residues (Fig. 4A). We cloned the ND2 deletion mutant into the ND2-F3 construct and introduced the resulting plasmid into the MAT $\alpha$  yeast strain, which was in turn mated with MAT $\alpha$  strains containing one of four ND2 interactors (i.e., NDUFC2, ND4, ND4L and ND5) (Fig. 4B) fused to the F1,2 mDHFR fragments. Here too, we performed dilution drop-test experiments for all four mating-pairs and compared the relative growth rates of the diploid strains containing the ND2 mutants with that of diploid strains harboring wild type ND2 (Fig. 5A). Our results indicated unexpected increased growth rates of the ND2 mutant strains when mated with NDUFC2- and ND4-expressing strains, while no significant change in growth rate was evident when the ND2 mutant strain was mated with ND4L or ND5 (Fig. 5B). However, while inspecting the control diploid strains, harboring mutated ND2 and an empty (i.e., not encoding any CI subunits) F1,2 construct, we noticed much higher growth as compared to the diploid strain harboring wild type ND2 and the empty F1,2 vector (Fig. 5B). This suggests that the ND2 mutant had become a promiscuous interacting protein.

However, promiscuity did not explain the increase in interaction strength in all cases; ND4L exhibited a stronger interaction with the mutant ND2, as compared to the empty vector control (Fig. 5B). A close inspection of the CI structure revealed ND4L to be the only subunit tested that interacts with the N-terminal portion of ND2 (Fig. 4B). This suggests that the specific interaction of ND2 and ND4L, which is dictated by their interaction interface, is not affected by the deletion mutation in ND2.

A



B



**Figure 2.** Network of identified CI subunit interactions using the split mDHFR PCA. **(A)** A network scheme of the interactions identified by the screen - edge length is positively correlated with the p-value for an interaction by a Mann Whitney U (MWU) test (see Materials and Methods). Red edges represent physical interactions that are consistent with the available 3D structural models of mammalian CI. **(B)** A matrix representation of the screen interactions - red squares represent physical interactions that are consistent with the available 3D structural models of mammalian CI. Black squares - additional identified interactions.

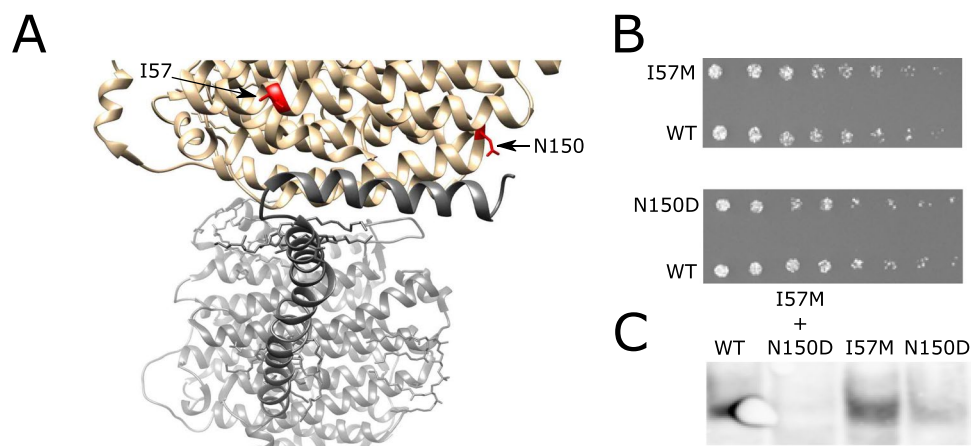
## Discussion

The availability of structures of multi-subunit protein complexes enables predicting the potential impact of mutations on subunit interactions. Such predictions can serve to design experiments that test the impact of such mutations. Nevertheless, such experiments cannot easily employ classical structural biology methodologies, which

Subunit name	Mutation (residue # in the mature nuclear-encoded CI subunit)
NDUFS1	Val71Asp (48)
	Arg241Trp (218)
	Asp619Asn (596)
NDUFV1	Pro122Leu (102)
	Arg199Pro (179)
	Tyr204Cys (184)
	Arg386Cys (366)
NDUFS2	Asp110Val (77)
	Arg118Gln (85)
	Arg138Gln (105)
	Ala224Val (191)
	Arg323Gln (290)
NDUFS3	Arg199Trp (163)
	Pro223Leu (187)
NDUFS8	Arg54Trp (20)
	Glu63Gln (29)
	Arg77Trp (43)
	Pro79Leu (45)
	Pro85Leu (51)
	Arg94Cys (60)
	Arg138His (104)
	Ala159Asp (125)
NDUFS7	Arg145His (107)
ND6	Ile26Met
	Gly36Ser
	Tyr59Cys
	Leu60Ser
	Met63Val
	Met64Val
	Met64Ile
	Ala72Val
	Ala74Val
	Asn117Asp
	Ser132Leu
	ND5
Tyr159His	
Ala171Val	
Gln434Arg	
Gly465Glu	
ND4L	Cys32Arg
	Val65Ala
ND4	Ile165Thr
	Arg340Ser
	Arg340His
	Tyr409His
ND3	Ser34Pro
	Ser45Pro
	Ala47Thr
	Ile60Thr
	Glu66Asn
ND2	Ile57Met
	Leu71Pro
	Asn150Asp
Continued	

Subunit name	Mutation (residue # in the mature nuclear-encoded CI subunit)
ND1	Val11Met
	Glu24Lys
	Leu28Met
	Tyr30His
	Met31Val
	Ala52Thr
	Met53Ile
	Glu59Lys
	Ser110Asn
	Ala112Thr
	Gly131Ser
	Ala132Thr
	Glu143Lys
	Ala164Thr
	Val208Leu
	Glu214Lys
	Tyr215His
	Leu285Pro
Leu289Met	
TYR304HIS	
NDUFS4	Asp119His (77)
NDUFS6	Cys115Tyr (87)
NDUFA9	Arg360Cys (325)
NDUFA1	Gly8Arg (8)
	Arg37Ser (37)
NDUFA10	Gly99Glu (64)
NDUFB9	Leu64Pro (63)

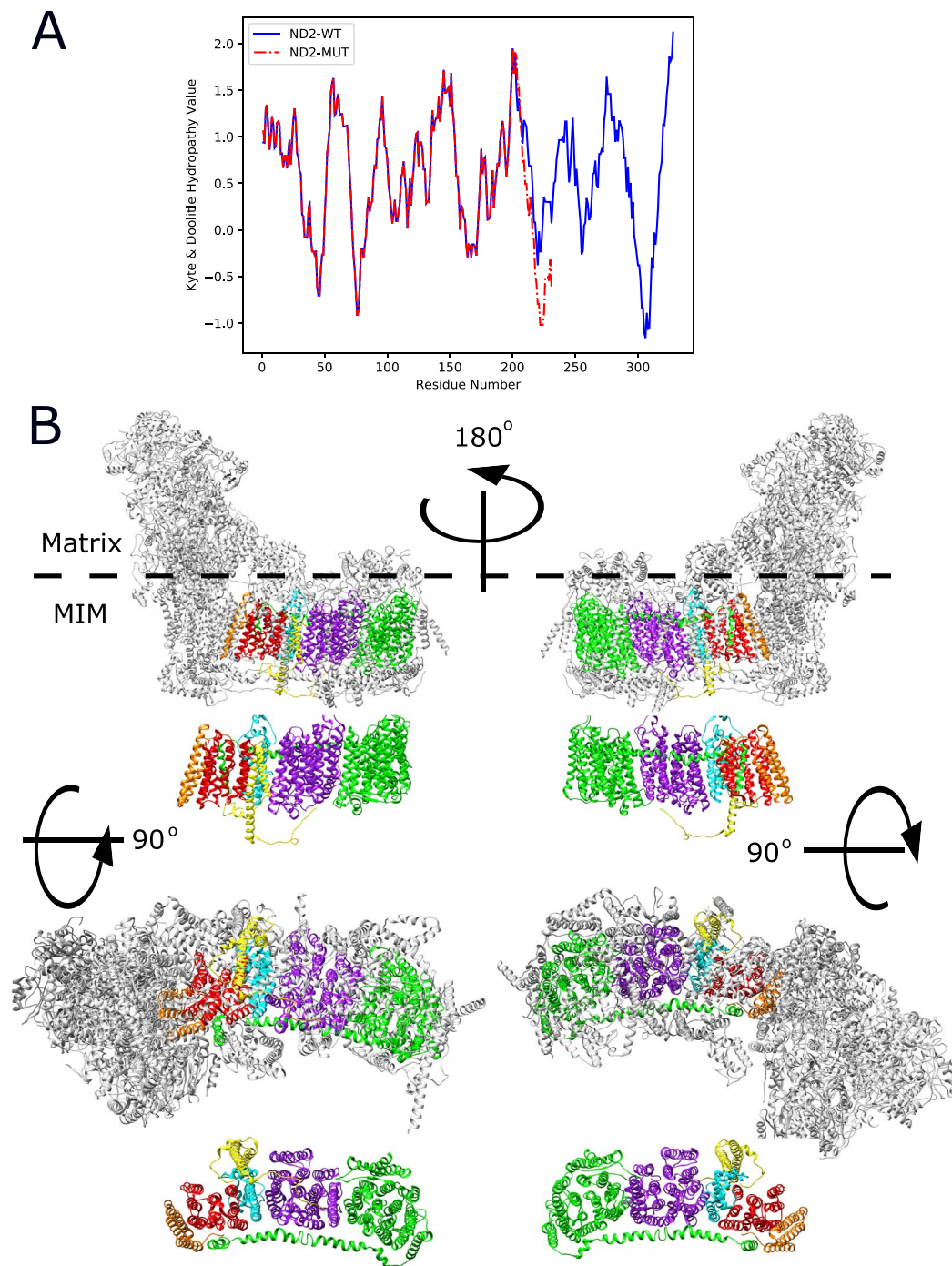
**Table 1.** Single amino acid substitutions associated with CI deficiency and potentially affecting protein-protein interactions between CI subunits. The listed mutations were mostly extracted from<sup>28</sup> in addition to a complementary literature search<sup>25,29,32,33</sup>.



**Figure 3.** Ancient mtDNA SNPs which associate with LHON show no effect on the ND2-ND5 interaction. (A) The location of changes in the ND2 sequence (red) in two TMHs (gray) and at the interaction interface with ND5 (black). The structure shown is based on (PDB ID: 5XTD)<sup>15</sup>. (B) Dilution drop-test results for the two ND2 mutants against wild type (WT) ND2. (C) Western blot demonstrating expression of the two single mutants and WT ND2 but not of the double mutant.

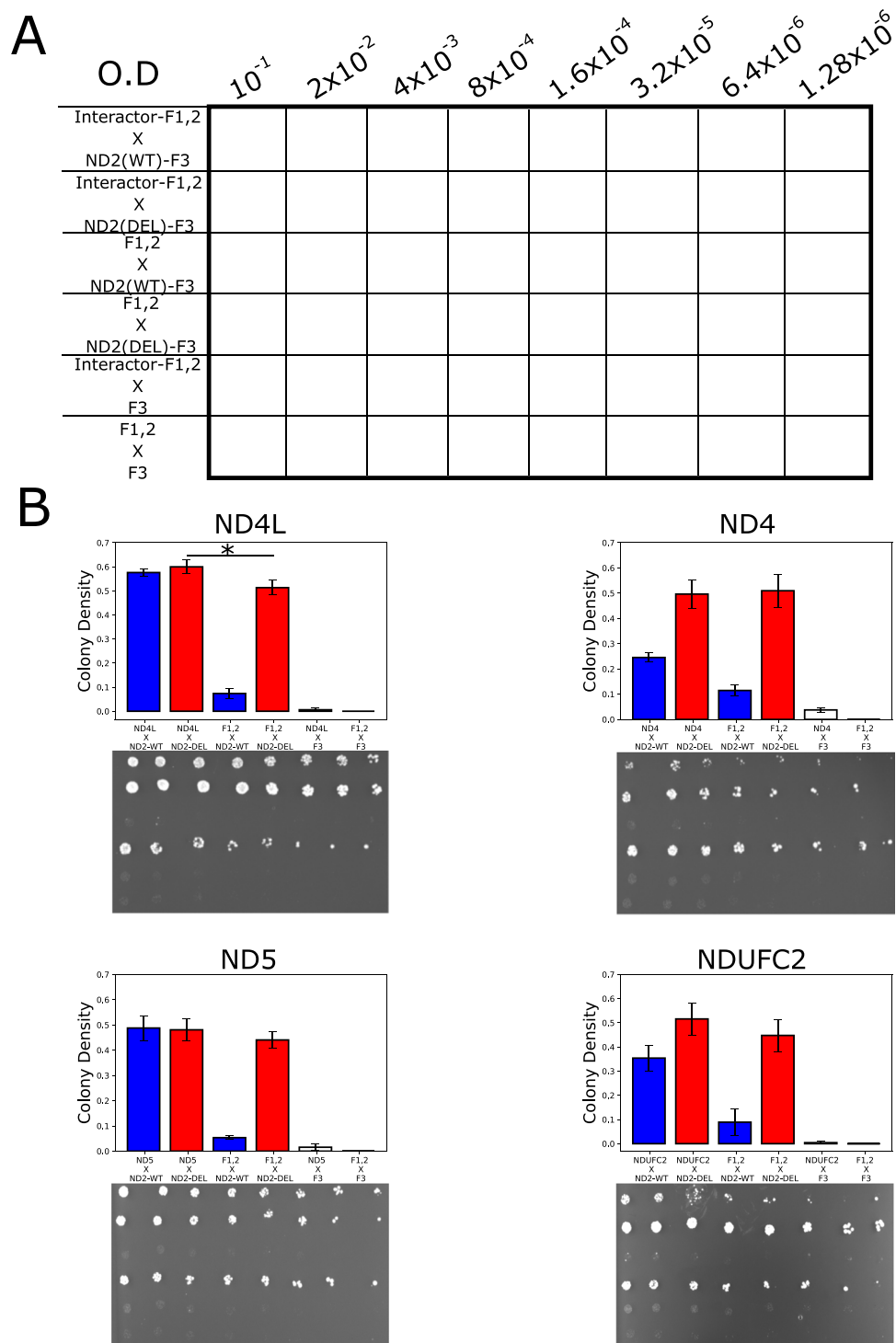
are not compatible with high-throughput approaches and are less accessible to most molecular biologists. Hence, we established an assay to enable such assessment to be carried out easily. Specifically, we used the split mDHFR protein complementation assay (PCA) to assess the impact of mutations at interaction interfaces between CI subunits on protein-protein interactions. First, we cloned all known human CI subunits (N = 44) into constructs





**Figure 4.** A disease-causing ND2 frameshift mutant results in a shorter, more hydrophilic C-terminus. **(A)** Kyte-Doolittle hydropathy plot comparing the hydrophobicity of the wild type (blue) and the ND2 frameshift mutant (red). **(B)** Four viewpoints of human CI showing ND2 and its four tested interactors within the full complex (upper panel of each view point) and a “naked” presentation (lower panel of each view point): Red - the N-terminal region of ND2, cyan - the deleted ND2 C-terminal region, yellow - NDUF2, purple - ND4, green - ND5 and orange - ND4L. The structure shown is based on (PDB ID: 5XTD)<sup>15</sup>. Location of the matrix arm (Matrix) and Mitochondrial inner membrane (MIM) are indicated.

harboring either of both parts of split mDHFR, thus establishing a resource which enables testing for all pairwise interactions. Such a resource offers the obvious advantage of accessibility and relative ease-of-use, relying on an assay based on yeast mating and measurement of growth rate on selective medium (containing methotrexate). The main disadvantage, especially in the case of a multi-subunit protein, is that the PCA is limited to assessment of pairwise interactions. Additionally, in OXPHOS CI, several subunits are simultaneously added during each stage of the assembly process<sup>32,33</sup>. This suggests that many interactions between subunit pairs likely depend on the presence of additional subunits; naturally, an assay designed to test pairwise interactions is blind to such



**Figure 5.** The ND2 frameshift mutation results in promiscuous C-terminal interactions. (A) The yeast growth plate designed for the dilution drop-test. ND2-DEL – frameshift mutant; ND-WT – wild type ND2. (B) Representative drop-test plates and bar-plots summarizing the yeast colony intensity obtained in three biological replicates. \*Indicates that the p value < 0.05.

interactions, which will generate false negative results. Moreover, some of the observed interactions may reflect transient interactions that occur during the assembly process of complex I, yet cannot be currently verified due to relative data sparsity. Additionally, we cannot exclude that false negative rate of subunit interactions was also influenced by protein misfolding. These may explain, at least in part, why only 19 interactions in our assay were consistent with the 3D structures. Once human CI will be reconstituted in yeast at least some of these caveats could be addressed, while also taking into account assembly factors. Regardless, our observed interactions



contained core subunits, which play a major role in the activity of CI and reflect the usefulness of our tool for assessing the impact of mutations on such interactions.

As proof-of-concept, we focused our experimental assessment on three changes, considering two disease-associated SNPs and one frameshift mutation, all occurring in the mtDNA-encoded ND2 subunit. The first SNP, 4640A, was originally found in a Russian family from Novosibirsk. The 4640A mutation resulted in poor oxygen utilization by cybrids, although these experiments did not reveal any reduction in the specific activity of CI, relative to control lymphoblast cells<sup>29</sup>. Close inspection of this mutation showed that it altered an amino acid position with poor evolutionary conservation. Additionally, the 4640A mutation appears to associate with the U3b sub-haplogroup in the human mtDNA phylogenetic tree. Similarly, the 4917G mutation associates with both the T and N1b haplogroups. Haplogroup T associates with LHON<sup>34</sup>, and with reduced sperm motility in the Spanish population<sup>35</sup> and haplogroup N1b altered the susceptibility of Ashkenazi Jewish type 2 diabetes patients to develop complications<sup>36</sup>. Notably, haplogroup T cybrids possess a higher capability to cope with oxidative stress in the form of an H<sub>2</sub>O<sub>2</sub> challenge<sup>37</sup>. Taken together, the impact of SNPs that co-occur with certain mtDNA genetic backgrounds cannot be easily distinguished from the linked set of mutations. One should thus take into consideration the fact that our observed lack of impact of both the 4640 and 4917 mutations on CI subunit interactions may either attest to their minor functional effect, suggesting that their impact is not manifested at the level of protein-protein interactions, or that this reflects the resolution limitations of the PCA. These possibilities cannot be resolved at this point and requires additional experiments.

Unlike the tested point mutations, a pathological *MT-ND2* frameshift mutation had a dramatic impact on the pattern of subunit interactions. Our results suggest that the *MT-ND2* frameshift mutant generated a promiscuous interacting protein, specifically in interactions involving the region encoded downstream to the frameshift. Such a molecular phenotype might explain the appearance of the disease at the relatively high (90%) heteroplasmy level seen in the patient affected by the deletion mutation<sup>25</sup>. It is also possible that at high heteroplasmy levels, the translation product of the mutated *MT-ND2* gene interacts with many members of the mitochondrial proteome, thus interfering with the proper ND2 interaction pattern, thereby causing the mitochondrial dysfunction observed in the clinic. Additionally, our results suggest that interaction of the N-terminal region of ND2 with the ND4L subunit was not affected, suggesting that the promiscuous interaction phenotype is limited to interactions involving the C-terminal region of ND2, thus excluding the possibility of a general mis-folding effect of ND2. These results provide, for the first time, a possible molecular mechanistic explanation for the pathophysiology of the ND2 frameshift mutation.

As CI structure and assembly pathway are very complex, studies of the phenotypic impact of human disease-causing mutations and disease-associated polymorphisms have largely relied on previous measurements of CI enzymatic activity<sup>2,29,38</sup> or on earlier work on the assembly of the holoenzyme. Still, little is known to date of the effects of human mutations on specific pairwise interactions between CI subunits. We speculate that protein-protein interactions between CI subunits that occur independently of additional subunits are important for the structure and proper assembly of CI.

A recent RNA-seq analysis from multiple human tissues revealed that certain CI subunits show tissue-dependent expression patterns<sup>39</sup>, thus questioning the generality of structural analysis of the complex from specific tissues, such as the heart or liver. With this in mind, our analysis of CI subunit interactions is not linked to a particular tissue, and hence is less influenced by tissue-dependent subunit composition considerations.

In summary, we provide a resource of clones for all known human CI subunits ready to be tested for pairwise protein-protein interactions using the split mDHFR PCA. Our experimental assessment of the impact of mutations on pairwise CI subunit interactions serve as a proof-of-concept and may encourage using this approach to study the impact of specific mutations on other multi-subunit protein complexes, such as the ribosome. Our finding of a profound impact of a *MT-ND2* frameshift mutation on interactions with partners revealed in the crystal structure of the complex provides the first direct explanation of the molecular mechanism underlying the phenotypic impact of this mutation. This discovery increases the motivation to recruit the system established in the current study to screen for the impact of other mutations of interest on confirmed subunit interactions.

## Materials and Methods

**Yeast growth media.** The complete recipes for yeast media used in the current study are detailed in Supplementary Material, Table S2.

**Preparation of split mDHFR libraries.** Two 2-micron plasmids were used for human CI subunit library preparation. Plasmid p340 includes amino acid sequence-coding motifs at the cloning site for a linker peptide (GlyGlyGlyGlySer)X2 followed by the ORF of the yellow fluorescent protein (YFP)-coding gene, followed by another (GlyGlyGlyGlySer)X2 peptide linker and DNA encoding mDHFR N' terminal fragments (F1,2). Expression of this fusion construct was under the transcriptional control of the *GAP* promoter and *CYC1* terminator. Plasmid pMS141 includes amino acid sequence-coding motifs at the cloning site for a linker peptide (GlyGlyGlyGlySer)X2 followed by DNA encoding a mDHFR C' terminal fragment (F3), and is expressed under the transcriptional control of the *TEF* promoter and *ADH1* terminator. In both plasmids, a *SpeI* restriction site was fused to the 5' codon of the constructs (Supplementary Material, Fig. S3).

ORFs of nDNA-encoded human CI subunits were amplified from commercial human cDNA samples (See Supplementary Material, Table S3 for the list of sources of human RNA). ORFs of the mtDNA-encoded subunits were recoded to match the cytoplasmic translation code (Supplementary Material, Table S4), generated commercially according to our design and cloned into pUC57 plasmids (OriGene). To create the two libraries, namely a F3 library in strain BY4741 [*MATa*, *his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ0*] and a YFP-F1,2 library in strain BY4742 [*MATα*, *his3Δ1*, *leu2Δ0*, *lys2Δ0*, *ura3Δ0*], the ORF encoding each subunit was amplified using primers harboring 5' tails of ~50 bp homology to either plasmid p340 or pMS141 (primers are listed in Supplementary Material,

Table S5). PCR amplification of the inserts was performed using Phusion DNA polymerase (Thermo-Scientific), with the following amplification protocol: Two min at 98 °C followed by 35 cycles of 30 sec at 98 °C, 30 sec at the melting temperature of the primers used and 72 °C for an elongation time calculated as template length in kilo base pairs multiplied by 30 sec. To perform homologous recombination-based cloning in yeast, 15 µL of the PCR reaction were used for co-transformation into yeast haploid strains (plasmids p340 and pMS141 were used for transformation into the BY4742 and BY4741 strains, respectively). To improve the efficiency of the cloning process in yeast, the plasmids were linearized by digestion with *SpeI*, and 200 ng of the digested plasmids were used for co-transformation with the PCR fragments. The yeast transformations were performed as previously described<sup>40</sup>. Following co-transformation of the inserts and the linearized plasmids, the yeast were grown on selective plates, namely SD–His agar plates for plasmid p340-transformed cells and YPD agar plates with cloNAT for plasmid pMS141-transformed cells, for a minimum of 48 h at 30 °C. Colonies (N = 8 to 10) were picked and streaked onto a similar second selective plate and grown overnight at 30 °C.

To validate successful cloning, a crude extract containing plasmid DNA was prepared by boiling small patches of the streaked colonies in 50 µL of 20 mM NaOH at 100 °C for 25 minutes and centrifugation for 2 min at 2,500 × g. The supernatant was used as template in a 10 µL PCR amplification using primers from the plasmid sequence flanking the insert (see Supplementary Material, Table S6). PCR products of appropriate size were Sanger-sequenced at the Ben-Gurion University sequencing unit. The sequences obtained were compared with the RefSeq sequences of the subunit-encoding ORFs (see Supplementary Material, Table S6 for CI subunits RefSeq protein accession numbers).

**Split mDHFR screen.** All 44 human CI subunits were cloned into plasmid pMS141 (mDHFR fragment F3) and 37 of the subunits were also cloned into plasmid p340 (mDHFR fragments F1,2). To test for all possible pairwise interactions between CI subunits, seven haploid strain selection plates were prepared, harboring 1,536 colonies each. Three SD–His agar plates were prepared for plasmid p340-transformed BY4742 cells and four YPD agar plates were prepared with cloNAT for plasmid pMS141-transformed BY4741 cells. The haploid strain libraries were seeded onto haploid selection plates and grown at 30 °C for 48 h, replicated on a second identical haploid selection plate and grown for an additional 24 h to reduce contamination. The libraries were then mated so that each of the plasmid p340-transformed BY4742 clones (three 1,536 clone plates) could be crossed with each of the four plasmid pMS141-transformed BY4741 plates, and vice-versa, yielding a total of 12 mating plates. Mating was performed while growing the yeast on non-selective YPD agar plates at 30 °C overnight to allow for the differently selected haploid strains to mate. These plates were subsequently replicated onto diploid selection SD–His agar plates with cloNAT. The resulting yeast colonies were grown overnight at 30 °C and then, to reduce haploid contamination, were replicated onto a second diploid selection plate and once again grown overnight at 30 °C. Each of the 12 diploid selection plates from the second round of diploid replication were replicated onto SD agar plates containing MTX and placed in a 30 °C incubator for 5 days (which in our hands gave clearer results than previously suggested 4 days). All yeast plate replicates and mating experiments were performed using a RoToR bench top colony arrayer (Singer Instruments, UK). The split mDHFR PCA results are semi quantitative, as protein-protein interactions are supported by significantly larger colony size from our negative control, similar to our positive control.

Plate distribution of the haploid strains was designed so that crosses of the BY4741 and BY4742 strains would result in a minimum of 9 replicates of each of tested subunit A–F1,2 × subunit B–F3 interactions, with an additional 15 colonies representing the control population for each tested interaction. Here, 9 subunit A–F1,2 × F3 colonies did not contain an insert and 6 colonies were F1,2 without insert × subunit B–F3 crosses. Each haploid plate harbored control colonies in which the construct did not include an insert.

**Validation of F1,2 fragment expression and human CI subunits localization in yeast.** BY4742 yeast clones with human CI subunits that were cloned into the p340 construct were grown over night in 30 °C SD–His medium in 15 ml falcons with shaking, back diluted ( $OD_{600} = 0.1$ ) into 96 well plate and grown for 4 hours in 30 °C SD–His medium with shaking. Notably, the yeast were grown in 30 °C rather than in 37 °C to avoid increased membrane propensity to be leaky, which likely influence growth and mask the influence of protein complementation on such growth. Then, yeasts were transferred into a 384 well plate for imaging in a Zeiss 710 Confocal Microscope. YFP expression was validated and the pattern of YFP distribution within the yeast cells were compared with the distribution patterns observed at Breker *et al.*<sup>41</sup> to determine protein localization within the yeast.

**Statistical analysis of the yeast colonies.** Colony area size was measured using Balony<sup>42</sup>. Firstly, colony sizes were normalized to the median size of all colonies within their row and column on each plate. Then, a one-tailed Mann Whitney U (MWU) test was performed with the alternative hypothesis of the form  $P(\text{case} > \text{control}) > P(\text{control} > \text{case})$ . An interaction was considered positive only if the ‘case’ group was significantly larger than both the ‘control’ groups described above (i.e., colonies resulting from mating of subunit A–F1,2 × F3 without insert and from F1,2 without insert × subunit B–F3). When colony area size was equal to zero, the colony was excluded from further statistical analysis to avoid technical bias.

**Drop test for evaluating the effect of the ND2 frameshift mutation.** Haploid strains were seeded onto haploid selection plates (SD–His agar plates for strain BY4742 containing the F1,2 constructs with DNA encoding ND4, ND4L, ND5 or NDUFC2 or insert-less and YPD agar plates with cloNAT for strain BY4741 containing the F3 constructs with DNA encoding ND2–WT, ND2–DEL (frameshift mutant), ND2–I57M, ND2–N150D or ND2–I57M + N150D or insert-less), grown at 30 °C for 48 h, replicated onto a second identical haploid selection plate and grown for an additional 24 h to reduce contamination. Mating was performed while

growing the yeast on non-selective YPD agar plates at 30 °C overnight. These plates were subsequently replicated on diploid selection SD- His agar plates with cloNAT. The resulting yeast colonies were grown overnight at 30 °C and then, to reduce haploid contamination, were replicated onto a second diploid selection plate and once again grown overnight at 30 °C. After validating expression of the ND2 variants, the diploid strains were grown in SD liquid medium containing MTX (to reduce nucleotide pools) and SD liquid medium containing DMSO (control) for 24 h at 30 °C with shaking. These yeast strains were diluted to  $OD_{600} = 0.1$  in fresh SD liquid medium containing MTX or DMSO, respectively, and each of the diluted strains was further diluted 7 times at a 1:5 ratio. Five  $\mu\text{L}$  of each diluted diploid strain were placed on a SD agar plate containing MTX or DMSO and placed in a 30 °C incubator for 5 days. To avoid reaching a growth asymptote, images were taken at days 2–5. Growth was measured as colony color intensity (which equals biomass density)– the intensity of the white color given by the yeast colony, over background. The overall highest differences between colonies at the same dilution resulting from the different matings were spotted on day 3 with the smallest intra-mating variations at the seeding density of  $OD_{600} = 0.1$ . As the only measurable background was upon mating the insert-less-F1,2 control with either ND2-WT- or ND2-DEL-expressing cells, we tested for significant growth over background by comparing the growth of the interactor-F1,2 X ND2-WT/ND2-DEL cross with the appropriate control mating of the insert-less-F1,2 X ND2-WT/ND2-DEL cross, respectively. Such pairwise comparisons were made using a two-tailed, unpaired Student's t-test with equal variance.

**Validation of F3 fragment expression by western blot.** Diploid strain colonies representing all matings performed for the drop-test were picked from the second diploid selection plates (see above) and grown overnight in SD-His liquid medium supplemented with cloNAT at 30 °C with shaking. The strains were diluted to  $OD_{600} = 0.1$  and grown overnight in SD-His liquid medium supplemented with cloNAT until mid-log phase ( $OD_{600} = 0.6$ – $0.8$ ). Approximately 2.5  $OD_{600}$  units (3.125–4.167 ml) of the cells were centrifuged at 1000 g for 5 min, resuspended in 100 mM NaOH and incubated for 5 min at room temperature. The resulting cell extracts were centrifuged at 16,000  $\times$  g for 3 min and the protein-containing pellet was resuspended in SDS-containing loading buffer and incubated at 70 °C for 10 min. Thirty  $\mu\text{g}$  aliquots were loaded onto 4–20% polyacrylamide gels and the separated proteins were transferred to a polyvinylidene difluoride membrane at  $\sim 330$  mA for 1 h. The membrane was then subjected to a standard blotting protocol with primary rabbit anti-mDHFR(F3) antibodies (Sigma-Aldrich D0942) diluted 1:1000 and horseradish peroxidase-conjugated goat anti-rabbit secondary antibodies. Antibody binding was visualized with an enhanced chemiluminescent solution (WesternBright Sirius, advansta, BGU) for 5 min and image analysis (LAS-3000 luminescent image analyzer, Fujifilm). Such analysis revealed an additional  $\sim 10$  kDa band in the ND2-WT strains that was absent from the ND2-DEL strains, probably reflecting some level of unfused F3 expression product (Supplementary Material, Fig. S4).

## References

- Distelmaier, F. *et al.* Mitochondrial complex I deficiency: from organelle dysfunction to clinical disease. *Brain* **132**, 833–842 (2009).
- Morais, V. A. *et al.* Parkinson's disease mutations in PINK1 result in decreased Complex I activity and deficient synaptic function. *EMBO molecular medicine* **1**, 99–111 (2009).
- Smigrodzki, R., Parks, J. & Parker, W. D. High frequency of mitochondrial complex I mutations in Parkinson's disease and aging. *Neurobiology of aging* **25**, 1273–1281 (2004).
- Gatt, A. P. *et al.* Dementia in Parkinson's disease is associated with enhanced mitochondrial complex I deficiency. *Mov Disord* **31**, 352–359, <https://doi.org/10.1002/mds.26513> (2016).
- Baracca, A. *et al.* Mitochondrial Complex I decrease is responsible for bioenergetic dysfunction in K-ras transformed cells. *Bba-Bioenergetics* **1797**, 314–323, <https://doi.org/10.1016/j.bbabi.2009.11.006> (2010).
- Kurelac, I. *et al.* Somatic complex I disruptive mitochondrial DNA mutations are modifiers of tumorigenesis that correlate with low genomic instability in pituitary adenomas. *Human molecular genetics* **22**, 226–238 (2013).
- Koopman, W. J., Willems, P. H. & Smeitink, J. A. Monogenic mitochondrial disorders. *New England Journal of Medicine* **366**, 1132–1141 (2012).
- Ugalde, C., Janssen, R. J., van den Heuvel, L. P., Smeitink, J. A. & Nijtmans, L. G. Differences in assembly or stability of complex I and other mitochondrial OXPHOS complexes in inherited complex I deficiency. *Human molecular genetics* **13**, 659–667 (2004).
- Gershoni, M. *et al.* Disrupting mitochondrial-nuclear co-evolution affects OXPHOS complex I integrity and impacts human health. *Genome Biol Evol* **6**, 2665–2680, <https://doi.org/10.1093/gbe/evu208> (2014).
- Balsa, E. *et al.* NDUFA4 is a subunit of complex IV of the mammalian electron transport chain. *Cell metabolism* **16**, 378–386 (2012).
- Carroll, J., Fearnley, I. M., Shannon, R. J., Hirst, J. & Walker, J. E. Analysis of the subunit composition of complex I from bovine heart mitochondria. *Molecular & Cellular Proteomics* **2**, 117–126 (2003).
- Leif, H., Sled, V. D., Ohnishi, T., Weiss, H. & Friedrich, T. Isolation and Characterization of the Proton-translocating NADH: ubiquinone Oxidoreductase from Escherichia coli. *European Journal of Biochemistry* **230**, 538–548 (1995).
- Fiedorczuk, K. *et al.* Atomic structure of the entire mammalian mitochondrial complex I. *Nature* **538**, 406–410 (2016).
- Zhu, J., Vinothkumar, K. R. & Hirst, J. Structure of mammalian respiratory complex I. *Nature* (2016).
- Guo, R., Zong, S., Wu, M., Gu, J. & Yang, M. Architecture of human mitochondrial respiratory megacomplex I2III2IV2. *Cell* **170**, 1247–1257. e1212 (2017).
- Guerrero-Castillo, S. *et al.* The assembly pathway of mitochondrial respiratory chain complex I. *Cell metabolism* **25**, 128–139 (2017).
- Carroll, J. *et al.* Bovine complex I is a complex of 45 different subunits. *Journal of Biological Chemistry* **281**, 32724–32727 (2006).
- Guarani, V. *et al.* TIMMDC1/C3orf1 functions as a membrane-embedded mitochondrial complex I assembly factor through association with the MCI1A complex. *Molecular and cellular biology* **34**, 847–861 (2014).
- Rak, M. & Rustin, P. Supernumerary subunits NDUFA3, NDUFA5 and NDUFA12 are required for the formation of the extramembrane arm of human mitochondrial complex I. *FEBS letters* **588**, 1832–1838 (2014).
- Vinothkumar, K. R., Zhu, J. & Hirst, J. Architecture of mammalian respiratory complex I. *Nature* **515**, 80–84 (2014).
- Zhu, J. *et al.* Structure of subcomplex I $\beta$  of mammalian respiratory complex I leads to new supernumerary subunit assignments. *Proceedings of the National Academy of Sciences* **112**, 12087–12092 (2015).
- Fassone, E. & Rahman, S. Complex I deficiency: clinical features, biochemistry and molecular genetics. *Journal of medical genetics* **49**, 578–590 (2012).
- Remy, I., Campbell-Valois, F. & Michnick, S. W. Detection of protein–protein interactions using a simple survival protein-fragment complementation assay based on the enzyme dihydrofolate reductase. *Nature protocols* **2**, 2120–2125 (2007).
- Tarassov, K. *et al.* An *in vivo* map of the yeast protein interactome. *Science* **320**, 1465–1470 (2008).

25. Schwartz, M. & Vissing, J. Paternal inheritance of mitochondrial DNA. *New England Journal of Medicine* **347**, 576–580 (2002).
26. Tarassov, K. *et al.* An *in vivo* map of the yeast protein interactome. *Science* **320**, 1465–1470 (2008).
27. Johns, D. R. & Berman, J. Alternative, simultaneous complex I mitochondrial DNA mutations in Leber's hereditary optic neuropathy. *Biochemical and biophysical research communications* **174**, 1324–1330 (1991).
28. Fiedorczuk, K. & Sazanov, L. A. Mammalian Mitochondrial Complex I Structure and Disease-Causing Mutations. *Trends in cell biology* (2018).
29. Brown, M. D. *et al.* Novel mtDNA mutations and oxidative phosphorylation dysfunction in Russian LHON families. *Human genetics* **109**, 33–39 (2001).
30. Diss, G. & Lehner, B. The genetic landscape of a physical interaction. *Elife* **7**, <https://doi.org/10.7554/eLife.32472> (2018).
31. Schwartz, M. & Vissing, J. Paternal inheritance of mitochondrial DNA. *N Engl J Med* **347**, 576–580 (2002).
32. Hoefs, S. J. *et al.* NDUFA2 complex I mutation leads to Leigh disease. *The American Journal of Human Genetics* **82**, 1306–1315 (2008).
33. Ugalde, C. *et al.* Impaired complex I assembly in a Leigh syndrome patient with a novel missense mutation in the ND6 gene. *Annals of neurology* **54**, 665–669 (2003).
34. Torroni, A. *et al.* Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of the primary mutations 11778 and 14484. *American journal of human genetics* **60**, 1107 (1997).
35. Ruiz-Pesini, E. *et al.* Human mtDNA Haplogroups associated with high or reduced spermatozoa motility. *American journal of human genetics* **67**, 682–696 (2000).
36. Feder, J. *et al.* Differences in mtDNA haplogroup distribution among 3 Jewish populations alter susceptibility to T2DM complications. *BMC Genomics* **9**, 198 (2008).
37. Mueller, E. E. *et al.* Functional differences between mitochondrial haplogroup T and haplogroup H in HEK293 cybrid cells. *PLoS One* **7**, e52367, <https://doi.org/10.1371/journal.pone.0052367> (2012).
38. Loeffen, J. *et al.* The first nuclear-encoded complex I mutation in a patient with Leigh syndrome. *The American Journal of Human Genetics* **63**, 1598–1608 (1998).
39. Barshad, G., Blumberg, A., Cohen, T. & Mishmar, D. Human primitive brain displays negative mitochondrial-nuclear expression correlation of respiratory genes. *Genome Res* **28**, 952–967, <https://doi.org/10.1101/gr.226324.117> (2018).
40. Gietz, R. D. & Woods, R. A. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods in enzymology* **350**, 87–96 (2002).
41. Breker, M., Gymrek, M. & Schuldiner, M. A novel single-cell screening platform reveals proteome plasticity during yeast stress responses. *J Cell Biol* **200**, 839–850 (2013).
42. Young, B. P. & Loewen, C. J. Balony: a software package for analysis of data generated by synthetic genetic array experiments. *BMC bioinformatics* **14**, 1 (2013).

## Acknowledgements

The authors wish to thank Prof. Raz Zarivach (BGU) for critical discussion during early stages of the manuscript. This study was funded by Grants 372/17 and 610/12 from the Israeli Science Foundation, a Binational Science Foundation grant number 2013060 and a US army life sciences division grant (6793LS) awarded to D.M. The Schuldiner lab is supported by a DIP Grant (P17516). M.S. is an incumbent of the Dr. Gilbert Omenn and Martha Darling Professorial Chair in Molecular Genetics.

## Author Contributions

G.B. performed most of the experiments and heavily participated in writing the paper; N.Z.P. did the experiments demonstrated in Figures S2 and S4; L.G. assisted in yeast growth and initial screen for complex I subunit interactions; M.S. assisted in designing the yeast experiments, provided the plasmids for cloning and contributed to writing the manuscript text; D.M. conceived the study and wrote the paper.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-46446-8>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019