Supplementary Methods:

- Definition of areflexia and total reflex score:

  The following limb reflex items used to define reflex-related measures were scored as observed in the neurological examination: Bicep, tricep, brachioradialis, patella, and ankle, scored separately for the left and right side as Absent = 0, Decreased (trace or only with reinforcement) = 1, Normal = 2, Increased (brisk) = 3, or Hyperreflexic (clonus) = 4. If all 10 items were scored = 0, then areflexia was considered present; if each of the 10 items was assigned an integer-valued code 1-4, then areflexia was considered absent; otherwise, areflexia was considered missing, i.e. neither present nor absent. Finally, if each of the 10 items was assigned an integer-valued code 0-4, then the codes were summed to calculate a total reflex score with possible range 0 to 40; if any item among the 10 did not have an integer-valued code 0-4, then the total reflex score was coded as missing for the subject.

- Definition of Rostrocaudal Gradient, Face > Arm > Leg:

  The following rating scale was used to code severity of separate symptoms, as observed by IH as part of the neurological examination, for the subject's mouth, speech/swallowing, left arm, right arm, left leg, and right leg: No dystonia = 0, Slight = 1, Mild = 2, Moderate = 3, and Severe = 4. If each of the six items was assigned an integer-valued code 0-4, then the codes were summed to calculate a face score (mouth + speech/swallowing), an arm score (left + right), and a leg score (left + right); these were used, in turn, to assign presence or absence of the gradient based on a composite of strict inequalities: If (face score > arm score, and arm score > leg score) is true, then gradient = present, and if (face score > arm score, and arm score > leg score) is not true, then

gradient = absent. If any item among the six did not have an integer-valued code 0-4, then the gradient was coded as missing for the subject, i.e., neither present, nor absent.

- Definition of Absent, Possible/Probable, and Definite Parkinsonism:

  The following parkinsonism rating scale (1-4) was used to code parkinsonism, as observed as part of the neurological examination: Unaffected/no parkinsonism = 1, possible parkinsonism = 2, probable parkinsonism = 3, and definite parkinsonism = 4. If the parkinsonism rating scale was unscored, then parkinsonism was coded as missing.

- Definition of Diagnostic History of Dystonia, and Dystonia on Examination:

  All subjects or other informants were asked whether the subject had ever been diagnosed with a movement disorder, and, if they had, then they were asked what was the diagnosis or diagnoses. If the response contained "rapid-onset dystonia-parkinsonism," "RDP," "dystonia," or "dysphonia" then the subject was considered history-positive for a diagnosis of dystonia. The Burke-Fahn-Marsden Dystonia Rating Scale was used to define the presence of dystonia on examination (BFMDRS>0).

- Definition of Area Most Severely affected:

  Severity was determined with reference to the severity subscores for each body region in the BFMDRS. Each subjectively symptomatic carrier was considered separately, and all areas yielding the highest severity score for that individual were considered ‘areas most severely affected’ This was then totaled across all carriers. Percentages per area were derived from the ratio of total number of carriers indicating that area as a ‘most affected’ to the total number of carriers.

- Definition of dystonia distribution:

Each patient with dystonia was assigned to either ‘generalized’, ‘multifocal’, ‘segmental’, or focal on the basis of their ‘provoking factor’ Burke-Fahn-Marsden Dystonia Rating Scale provoking factor subscale scores. Provoking factors are graded as: 0 - No dystonia at rest or with action; 1 - Dystonia only with particular action; 2 - Dystonia with many actions; 3 - Dystonia on action of distant part of body or intermittently at rest; 4 - Dystonia present at rest. Subjects were assigned categories on the following bases: generalized – trunk involved along with two other regions; multifocal – more than one noncontiguous region involved; segmental – multiple areas involved but all contiguous; focal – only a single region involved.

- Definition of Bulbar Symptoms:

Eight items were used to code bulbar symptoms, as observed as part of the neurological examination. The first three items were related to speech: dysarthric, breathy, and voice breaks. Each of these was coded as: Yes = 1, No = 2, and Not testable = 3. The next three items were related to cranial nerves: palate, shoulder shrug, and tongue. Each of these was coded as: Normal = 1, Abnormal = 2, and Not testable = 3. The last two items were dystonia in the mouth region, and speech/swallowing difficulty, which are Provoking Factor items from the Burke-Fahn-Marsden Dystonia Rating Scale observed on the subjects. The mouth item was coded as: No dystonia at rest or with action = 0, Dystonia only with particular action = 1, Dystonia with many actions = 2, Dystonia on action of distant part of body or intermittently at rest = 3, and Dystonia present at rest = 4. Speech/swallowing was coded as: No difficulty with either = 0, Occasional either or both = 1, Frequent either = 2, Frequent one and occasional other = 3, Frequent both = 4. Bulbar symptoms were considered present if one or more of the eight was scored

affirmatively (even if one or more than one was unscored): dysarthric speech, breathy, or voice breaks = 1; palate, shoulder shrug, or tongue = 2; or mouth or speech/swallowing 1-4. Bulbar symptoms were considered absent if all of the eight items were scored in the negative: dysarthric speech, breathy, and voice breaks = 2; palate, shoulder shrug, and tongue = 1; and mouth and speech/swallowing = 0. Otherwise, if all of the items were either negative, not testable, or unscored, then bulbar symptom status was considered missing, i.e., neither present, nor absent.

- Definition of Cerebellar Dysfunction:

  The following items and rating scales from the International Cooperative Ataxia Rating Scale (ICARS) were used to define cerebellar dysfunction, as observed by IH as part of the neurological examination: Knee-tibia test (decomposition of movement and intention tremor) scored separately for the left and right side as Normal = 0, Lowering of heel in continuous axis, but the movement is decomposed in several phases, without real jerks, or abnormally slow = 1, Lowering jerkily in the axis = 2, Lowering jerkily with lateral movements = 3, and Lowering jerkily with extremely strong lateral movements or test impossible = 4; finger-to-nose test (decomposition and dysmetria) scored separately for the left and right finger as No trouble = 0, Oscillating movement without decomposition of the movement = 1, Segmented movement in more than 2 phases and/or moderate dysmetria in reaching nose = 2, Segmented movement in more than 2 phases and/or considerable dysmetria in reaching nose = 3, and Dysmetria preventing the patient from reaching nose = 4; gaze-evoked nystagmus, scored as Normal = 0, Transient = 1, Persistent but moderate = 2, and Persistent and severe = 3; and Dysmetria of the saccade, scored as Absent = 0, and Bilateral clear overshoot or undershoot of the saccade = 1. If

at least one of the six items was scored > 0, then cerebellar dysfunction was considered present (even if one or more than one was unscored). If all six were scored = 0, then cerebellar dysfunction was considered absent. Otherwise, if all of the items were either = 0 or unscored, then cerebellar dysfunction status was considered missing, i.e., neither present, nor absent.

- Methodology for Figure 2: 'region where first symptom was felt' was gathered from the patient questionnaire, which allowed 'face', 'mouth', 'arm', or 'leg'. Regions involved at initial exam and dystonia score distributions were derived from subjects' Burke Fahn Marsden Dystonia Rating Scale regional subscores for eyes and mouth (face), speech/swallowing, arms, and legs (scored 0-4, with no dystonia = 0, slight = 1, mild = 2, moderate = 3, and severe = 4). When both right and left severity scores were available for a region, the higher value was selected.

**Supplementary Table 1: Statistical analysis of key motor features in those with and without ATP1A3 mutations, with missing data**

| Characteristics | | Carriers N=44 | Noncarriers N=44 | p-Value |
|---|---|---|---|---|
| Rapidity of onset | Rapid onset | 35 (79.5%) | 0 (0.0%) | 0.014 |
| | Non-rapid onset | 9 (20.5%) | 3 (6.8%) | |
| | No onset | 0 (0.0%) | 41 (93.2%) | |
| | Missing data | 0 (0.0%) | 0 (0.0%) | |
| F > A > L Gradient | Yes | 3 (7.0%) | 0 (0.0%) | 0.241 |
| | No | 40 (93.0%) | 43 (100.0%) | |
| | Missing data | 1 (2.3%) | 1 (2.3%) | |
| Bulbar Symptoms | Yes | 41 (95.3%) | 2 (5.0%) | < 0.001 |
| | No | 2 (4.7%) | 38 (95.0%) | |
| | Missing data | 1 (2.3%) | 4 (9.1%) | |
| Headaches | Yes | 26 (68.4%) | 17 (39.5%) | 0.031 |
| | No | 12 (31.6%) | 26 (60.5%) | |
| | Missing data | 6 (13.6%) | 1 (2.3%) | |
| Seizures | Yes | 12 (30.8%) | 3 (7.1%) | 0.001 |
| | No | 27 (69.2%) | 38 (90.5%) | |
| | I don't know | 0 (0.0%) | 1 (2.4%) | |
| | Missing data | 5 (10%) | 2 (4.6%) | |
| Parkinsonism | Absent | 1 (2.3%) | 40 (97.6%) | < 0.0001 |
| | Possible/Probable | 9 (20.9%) | 1 (2.4%) | |
| | Definite | 33 (76.7%) | 0 (0%) | |
| | Missing data | 1 (2.3%) | 3 (6.8%) | |
| Tremor | Any | 13 (29.5%) | 6 (14.0%) | 0.118 |
| | Missing data | 0 (0.0%) | 1 (2.3%) | |
| | Action & rest tremor | 3 (6.8%) | 0 (0.0%) | 0.103 |
| | Action only | 10 (22.7%) | 6 (14.0%) | 0.118 |
| | Rest only | 0 (0.0%) | 0 (0.0%) | |
| | Missing data | 0 (0.0%) | 0 (0.0%) | |
| | None | 31 (70.5%) | 37 (86.0%) | |
| BFMDRS | | 53.2 (29.4) | 0.2 (0.6) | < 0.001 |
| | Missing data | 1 | 1 | |
| UPDRS | | 42.6 (18.6) | 4.5 (0.6) | 0.002 |
| | Missing data | 35 | 40 | |
| Cerebellar Dysfunction | Yes | 7 (70.0%) | 0 (0.0%) | 0.070 |
| | No | 3 (30.0%) | 4 (100.0%) | |
| | Missing data | 34 (77.3%) | 40 (90.9%) | |

| | | | | |
|---|---|---|---|---|
| Reflexes | | 22.7 (7.0) | 18.8 (4.0) | 0.004* |
| | Missing data | 7 (14.0%) | 5 (10.0%) | |

Statistical analysis of key motor features in those with and without *ATP1A3* mutations. Percentages are given with respect to total assessed subjects; where there are missing data due to patient-specific issues, information is given in supplementary data. Standard deviations are given in parentheses (SD) where appropriate. BMFDRS = Burke Fahn Marsden Dystonia Rating Scale, UPDRS = Unified Parkinson's Disease Rating Scale motor subscale, Y-BOCS = Yale Brown Obsessive Compulsive Symptom Scale. * reflex scores are given as a sum of limb reflex scores (bilateral biceps, triceps, brachioradialis, patellar, Achilles' tendons with each scored 0-4, total range 0-40). Further detail on calculation methodology is given in Supplemental Methods above.

**Supplementary Table 2: Statistical analysis of key nonmotor features in those with and without ATP1A3 mutations, with missing data**

| Characteristics | | Carriers N=44 | Noncarriers N=44 | p-Value |
|---|---|---|---|---|
| Cognitive Function | Trailmaking B | 35.8 (10.3) | 51.2 (12.0) | < 0.001 |
| | Missing data | 18 | 15 | |
| | Linguistic Fluency | 33.9 (8.5) | 46.9 (9.2) | < 0.001 |
| | Missing data | 24 | 16 | |
| | Semantic Fluency | 37.1 (13.1) | 51.4 (9.6) | < 0.001 |
| | Missing data | 23 | 15 | |
| Psychiatric symptoms | SCID Mood Disorder | 16 (50%) | 7 (22%) | 0.036 |
| | Missing data | 15 | 17 | |
| | SCID Anxiety Disorder | 13 (41%) | 12 (38%) | 0.080 |
| | Missing data | 15 | 17 | |
| | SCID Psychosis | 7 (22%) | 0 (0%) | 0.011 |
| | Missing data | 15 | 17 | |
| | SCID Substance Abuse | 11 (34%) | 8 (26%) | 0.590 |
| | Missing data | 15 | 17 | |
| | Hamilton Anxiety | 7.2 (7.3) | 5.2 (8.7) | 0.240 |
| | Missing data | 16 | 15 | |
| | Hamilton Depression | 10.5 (9.2) | 6.3 (7.7) | 0.051 |
| | Missing data | 16 | 15 | |
| | YBOCS | 3.2 (6.3) | 0.9 (3.1) | 0.385 |
| | Missing data | 15 | 16 | |

Statistical analysis of key nonmotor features in those with and without *ATP1A3* mutations. Mean cognitive scores in the mutation-positive group were in the Mild Impairment range based on the tests' normative values (Heaton *et al.*, 2004), while mutation negative patients all exhibited Average range performance on these measures. Numeric values are counts (column percentages) and means (±SD). SCID = Structured Clinical Interview for DSM-IV, Y-BOCS = Yale Brown Obsessive Compulsive Symptom Scale.