
Article:
Smith, S.L., Lones, M.A., Bedder, M. et al. (2015) Computational approaches for understanding the diagnosis and treatment of Parkinson's disease. IET Systems Biology, 9 (6). pp. 226-233.

https://doi.org/10.1049/iet-syb.2015.0030

Published in final edited form as:
IET Syst Biol. 2015 December ; 9(6): 226–233. doi:10.1049/iet-syb.2015.0030.

Computational approaches for understanding the diagnosis and treatment of Parkinson's disease

Stephen L. Smith[1,✉], **Michael A. Lones**[2], **Matthew Bedder**[3], **Jane E. Alty**[4], **Jeremy Cosgrove**[4], **Richard J. Maguire**[5], **Mary Elizabeth Pownall**[6], **Diana Ivanoiu**[6], **Camille Lyle**[6], **Amy Cording**[6], and **Christopher J.H. Elliott**[6]

[1]Department of Electronics, University of York, Heslington, York Y010 5DD

[2]School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh EH14 4AS

[3]Department of Computer Science, University of York, Heslington, York Y010 5GW

[4]Neurology Department, Leeds Teaching Hospitals NHS Trust, Leeds LS1 3EX

[5]Department of Biomedical Science, University of Sheffield, Sheffield S10 2TN

[6]Department of Biology, University of York, Heslington, York Y010 5DD

Abstract

This study describes how the application of evolutionary algorithms (EAs) can be used to study motor function in humans with Parkinson's disease (PD) and in animal models of PD. Human data is obtained using commercially available sensors via a range of non-invasive procedures that follow conventional clinical practice. EAs can then be used to classify human data for a range of uses, including diagnosis and disease monitoring. New results are presented that demonstrate how EAs can also be used to classify fruit flies with and without genetic mutations that cause Parkinson's by using measurements of the proboscis extension reflex. The case is made for a computational approach that can be applied across human and animal studies of PD and lays the way for evaluation of existing and new drug therapies in a truly objective way.

1 Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disease and its prevalence is likely to increase dramatically over the next decade as people live longer [1]. Conventional diagnostic approaches for PD are based on clinical observation and this subjective approach can be unreliable, especially at early stages of disease. Monitoring of response to treatment, where the rate of improvement or deterioration may be modest, is difficult to objectively quantify.

Tools that provide straightforward, objective measurement of disease progression would allow better tailoring of treatments to individuals with PD, improving quality of life and maximising health system resources. When disease-modifying drugs become available for

✉stephen.smith@york.ac.uk.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

PD the need for early, accurate diagnosis and effective measurement of response to treatment will become even more important.

This paper describes computational approaches using evolutionary algorithms (EAs) that provide clinically relevant objective measures to identify and to quantify PD, both in humans and animal models. We begin by providing an overview of the diagnosis and monitoring of PD in humans. Thereafter two animal models of PD, the fruit fly (*Drosophila melanogaster*) and the zebrafish (*Danio rerio*), will be discussed. An overview of EAs will be provided and then a description of how we have used EAs to study motor function in humans and animal models not only to provide effective classifiers for discrimination between disease and controls, but also between disease types.

2 Parkinson's disease

PD is a progressive, incurable neurodegenerative condition characterised by distinct pathological changes including a loss of dopamine containing brain cells. A lack of dopamine in the brain causes a movement disorder, or 'motor dysfunction', characterised by slowness ('bradykinesia'), stiffness ('rigidity'), shaking of the body ('tremor') and impaired balance ('postural instability'). Bradykinesia is a complex clinical sign and consists of a number of separable components such as the speed, frequency and rhythm of a movement. As well as slowness of movement, those with PD often exhibit a reduction in the amplitude of movements ('hypokinesia') or absence of movement altogether ('akinesia').

The cause of the majority of cases of PD is unknown, and clinicians often refer to this sporadic form as 'idiopathic PD'. Only 5 to 15% of PD cases are familial and a number of causative autosomal dominant and recessive genetic mutations have been identified. All ethnic groups and countries are affected by PD [2]; approximately 1 in 500 of the whole population and 1 in 100 of those aged over 60 years [3]. The mean age of onset is 60 years, and incidence increases with age, but 5% of cases present before the age of 40.

2.1 Diagnosis of PD

The diagnosis of PD is primarily based on clinical interpretation of symptoms and signs elicited through history taking and examination. Sometimes it is straightforward for a doctor to make a confident diagnosis of PD, especially if there are several striking abnormal findings. However there are a number of other neurological conditions that can manifest with the same, or very similar, abnormal clinical signs, for example essential tremor (ET) or progressive supranuclear palsy (PSP). These conditions have most overlap during the early stages and it may be difficult to make an accurate diagnosis of PD based solely on clinical assessment, especially if the clinical findings are subtle or few.

Community studies suggest that at least 15–26% of people with a diagnosis of PD have been misdiagnosed [4, 5] and pathological studies broadly support these findings with 10–24% of clinical PD cases not confirmed at post mortem [6–8]. Even consultant neurologists with specific expertise in PD have been shown to misdiagnose PD for other tremor disorders with sensitivity/specificity in the range of 0.72–0.93/0.79–0.86 [9].

Europe PMC Funders Author Manuscripts

If there is some diagnostic uncertainty after the initial clinical evaluation a period of 'watchful waiting' may be recommended, typically for 6 to 12 months. This allows the symptoms and signs to progress further in order for the diagnosis to hopefully be made with more confidence. In cases where this method is unacceptable to the patient or clinician, or the interval review has still not helped, ancillary tests may be employed to supplement clinical assessment and aid diagnosis making.

Over the last decade, single photon emission computed tomography (SPECT) scans that quantify the number of dopamine containing brain cells have played an important role in aiding accurate diagnosis of PD. The scans are abnormal in PD but they are normal in ET and some other differential diagnoses such as dystonic tremor. However, they are not specific for PD and may also be abnormal in other conditions such as PSP, thus limiting their usefulness in aiding PD diagnosis. SPECT scans are expensive, typically costing at least £600 each; they involve administering ionising radiation to the patient and are not available in all healthcare settings. There are no routinely used blood or spinal fluid tests to diagnose PD and standard CT and MRI brain scans are not usually helpful in the early stages of PD or its mimics.

In summary, the fallibility of clinical opinion, taken together with the limitations of radionuclide SPECT scans, means that there remains a real need for new tests that can help accurately diagnose and monitor PD.

2.2 Treatment and monitoring of PD

Treatment of PD usually involves patients taking oral medications several times a day. The drugs work by increasing brain dopamine levels and this improves motor function for several hours. Unfortunately though, after several years, most patients develop two troublesome drug side effects called 'dyskinesia' and 'wearing off'. Dyskinesia means that body parts will writhe or jerk involuntarily and is often caused by drug levels being too high in the body. 'Wearing off' means the duration of each drug dose reduces so that the motor symptoms return unpredictably. Both of these side effects, taken together with the progressive nature of PD, results in increasingly more complicated drug regimens being required to keep the symptoms under control.

Patients with PD vary in terms of how they respond to the therapeutic and adverse effects of medications, how their symptoms progress, how other illnesses interact with the disease and its treatment, and how well they can recognise and report the signs and symptoms of PD. Thus it is very important that they are monitored to optimise their clinical management and to minimise functional disability. Typically clinical monitoring is undertaken during a 15 min clinic appointment every 6 months or so. This method provides only a brief 'snapshot' of the patient's status and may not accurately reflect their functional impairment on a day-to-day basis, particularly as motor dysfunction in PD can fluctuate. Patient-completed symptom diaries are sometimes used to gather more clinical data to supplement the clinic appointment review. They may be useful in a small subset of patients but generally are considered unhelpful as many patients simply find them too onerous to fill out regularly and the data from diaries has been shown to correlate poorly with clinician assessments [10].

Europe PMC Funders Author Manuscripts

This latter point may be due to the fact that many PD patients are unaware of their own motor symptoms [11, 12].

The second important reason why PD patients require monitoring is for research studies. This typically involves serial detailed clinical measurements of impairment and disability using formal clinical rating scales. Whilst such scales allow a degree of standardisation across studies, they have a number of pertinent drawbacks such as length of time to complete the various assessments, limitations of using coarse-grained scales of severity, and the necessary subjective interpretation that results in inter-rater variability. Thus there is a real need for an accurate objective measure of PD clinical signs to improve the quality of monitoring for clinical and research purposes

2.3 Objective approaches to measuring PD

The scope for improving diagnosis of PD using automated and objective methods have been explored for many years and much progress has been made using a range of technologies and data processing approaches. Previous work can be summarised in terms of the particular symptoms of interest, the sensors employed to measure these symptoms and the methods used to analyse the resulting data.

Abnormal movement of the limbs and trunk in PD can be measured with accelerometers [13] gyroscopes [14], electromagnetic (EM) tracking sensors [15], line of sight tracking systems [16] and video motion capture [17]. Speech problems in PD can be measured using audio recordings and specialised devices such as the Laryngograph [18]. Cognitive problems in PD can be measured using conventional clinical and specialised tasks such as figure copying tests (using digitising tablets) [19], memory tasks [20], measuring reaction time [21] and completing questionnaires.

The subsequent analysis and classification of the data resulting from these sensors include conventional statistical approaches as well as machine learning including neural networks, support vector machine and EAs. EAs are the focus of the work presented in this paper and are considered in more detail in Section 4.

3 Animal models of PD

As already discussed, idiopathic PD is the most common form of PD in humans. Although the familial forms of PD are less common, studying the genes that are disrupted in these patients is beginning to provide a better understanding of what causes idiopathic PD. Elucidating the molecular function of the products of the genes that are mutated in some people with familial PD will shed light on the fundamental cellular and molecular processes underlying the disease. Understanding human pathogenesis at the cellular and molecular level relies on the use of animal models [22]. Because mouse models of PD have been disappointing [22] we will discuss the use of two genetic models for studying PD, the fruit fly (*D. melanogaster*) and the zebrafish (*D. rerio*).

Europe PMC Funders Author Manuscripts

3.1 Drosophila

Drosophila (*D. melanogaster*) is the best understood of all model organisms; this is largely due to its sophisticated genetics [23]. Because of the small size of Drosophila and its short life cycle of only 2 weeks, it has been used extensively in genetic research for more than 100 years and over this time vast numbers of mutants have been collected and accurate cytological and genetic maps of the chromosomes have been made.

Robust methods have been developed for mutating genes in specific tissues, at specific times or in response to a stimulus and used to investigate gene function. Importantly, it is known that many of the genes that control processes in the fruit fly are the same genes that control the same processes in vertebrates, including humans [24]. Indeed, flies share over 75% of genes linked to human disease. The highly conserved nature of the genetic control of biological processes means that what we learn from studies of the fruit fly is relevant to understanding human biology and disease.

Flies have provided an excellent model of genetic forms of PD [25], reflecting many features of the disease including loss of dopaminergic neurons, oxidative stress, mitochondrial abnormalities and reduced movement. Traditionally, movement has been recorded by timing the speed at which flies walk up a glass cylinder in response to a sharp tap [26–28]. This conflates their response to the startle stimulus, gravity and the central pattern generator for walking with the motor neuron and muscular physiology. This assay also fails to discriminate between the different possible movement defects (akinesia, hypokinesia and bradykinesia). Nor is it clear which of the sensory, motor and CNS processes contributing to the climbing response are influenced by dopamine and/or the expression of PD related genes. The multifaceted ‘bradykinesia’ seen in these mutants, and the difficulty of relating it to neuronal function, suggest we need another, simpler assay system.

The requirement for a simpler assay is reinforced by the difficulty of determining which dopaminergic neurons are critical for controlling locomotion. The fly brain contains ~125 dopaminergic neurons [29, 30], mostly grouped into clusters [31]. Most manipulations of PD-related transgenes result in a relatively small number of dopaminergic neurons dying, with many clusters unaffected. In a recent study of the LRRK2-G2019S mutation, the protocerebral posterior medial cluster dropped from 14 to 12 dopaminergic neurons without loss in others, for example, the protocerebral anterior lateral cluster [32]. Throughout the literature, the multiple processes involved in bradykinesia combine with the small loss of dopaminergic neurons to obscure the functional relationship. To progress, we need to link a precise measurement of movement with the physiology of a few specific dopaminergic neurons.

The discovery of dopaminergic modulation of the proboscis extension response (PER) presents an exciting way forward [33]. When a walking fly encounters a sweet solution, the chemosensory cells on its front legs respond (Fig. 1, step 1). Their axons signal to the sub-oesophageal ganglion (SOG; the part of the brain responsible for taste, Fig. 1, step 2). Within the SOG, the chemosensory inputs activate a single dopaminergic neuron (Fig. 1, step 3) [35]. Action potentials in the single dopaminergic neuron are sufficient to excite the

pharyngeal E49 motoneurons [36] (Fig. 1, step 4), leading to contraction of a proboscis muscle (Fig. 1, step 5, muscle M3). This well-defined pathway results in the extrusion of the proboscis towards the food (Fig. 1, step 6). This simple reflex circuit allows the fly to feed on the sweet solution. A key observation is that genetic silencing of the single dopaminergic neuron prevents the PER, while 'thermogenetic' activation of this cell elicits the full PER [35].

The fly GAL4/UAS system [37, 38], a biochemical method used to study gene expression, provides the tools for neuron-specific expression of PD related genes – for example, the tyrosine hydoxylase (TH) GAL4 can be used to express the *LRRK2-G2019S* mutation, which is the most common cause of PD in all dopaminergic neurons, while the Gr5a GAL4 will express the transgene in just the sugar-sensitive neurons, or the GAD GAL4 in just the GABAergic inhibitory neurons. Equally, other PD related mutations (e.g. *LRRK2-R1441C*), or the wild-type *hLRRK2*, can be expressed by using the TH GAL4, by using the appropriate UAS sequence. Fly genetics also provides advanced tools for expression in subsets of neurons, for example, just a proportion of the dopaminergic neurons. In addition, the use of genetically encoded voltage or calcium sensing dyes [39] offers opportunities to record neural action potentials and/or synaptic activity during the sensory reflex.

In Drosophila, we therefore tested if expression of the *LRRK2-G2019S* transgene in dopaminergic neurons (*TH > G2019S*) affected the PER. We found 31% of these flies responded ($N = 180$), compared with 50% of flies expressing the wild-type *hLRRK2* transgene in the dopaminergic neurons (*TH > hLRRK2*, $N = 160$, flies 3 days old). This difference was significant ($\chi^2_{1df} = 10.2$, $P = 0.0013$). In older flies, (kept for 14 days at 29° C), the same result was found. Only 49% of the *TH > G2019S* (PD- mimic flies, $N = 109$) responded, compared with 66% of *TH > hLRRK2* flies ($N = 130$, $\chi^2_{1df} = 7.49$, $P = 0.006$). This demonstrates the PER in the PD-mimic (*TH > G2019S*) have a form of akinesia. In addition, the PER seems slower in old *TH > G2019S* flies compared with controls of the same age. We recorded video under a Zeiss Stemi microscope, and found that the *TH > G2019S* PER took approximately 1.5 as long as the control *TH > hLRRK2* response (0.52 ± 0.05 s v 0.34 ± 0.03 s; $F_{1,19df} = 8.3$ $P = 0.01$). These observations suggest that the PER displays both akinesia and hypokinesia in flies expressing the PD-related mutation, *TH > G2019S*.

In summary, Drosophila provides a powerful genetic toolbox, it mimics many of the features of PD, and has a simple nervous system which can be explored to show dopaminergic expression of PD related genes generates a movement disorder.

3.2 Zebrafish

One major problem with using Drosophila as an animal model for human disease is that it is an invertebrate and its anatomy and physiology is therefore different from ours. Humans are vertebrates characterised by having a dorsal spinal cord and a similar organisation of neurons within the central nervous system. Harnessing the power of genetic analysis within a vertebrate animal model would improve the prospects for understanding human genetic disease. Over the last 25 years, the zebrafish (*D. rerio*) has been used as a vertebrate genetic

model to study development and disease. Very recent advances in genomic editing, and the simplicity of using these new tools in zebrafish, will provide a powerful method for modelling human genetic disease.

Zebra fish produce thousands of embryos when they spawn and these are useful because they develop outside the mother, are translucent, are easily injected, and are genetically tractable. One example of a human disease that has a zebrafish model is Duchenne's muscular dystrophy (DMD) [40]. There are zebrafish with a mutation in the dystrophin gene, the same gene that is defective in human patients with DMD, and in both humans and fish with this genotype the muscles degenerate and die. The zebrafish mutant for dystrophin was identified in a large-scale genetic screen [41] and these fish are currently being used for *in vivo* drug screens and real time analysis of muscle fibre loss. The identification of the dystrophin mutant was random and fortunate, but it was not directed.

One zebrafish mutant for a PD-associated gene has been identified in Pink 1 [42]. The pink1−/− zebrafish was found to have characteristics associated with human PD, such as loss of dopaminergic neurons and mitochondrial impairment. However, the method used to identify this mutation (targeting induced lesions in genomes) also relies on random mutation of the genome and a PCR based screen for mutations in specific genes [43]. It is now possible to undertake directed mutagenesis of genomes that until very recently was only possible using homologous recombination in mouse embryonic stem cells [43, 44]

There are significant new technologies that are making gene targeting not only possible but also practical. transcription activator-like effector nucleases (TALENs) are synthetic restriction endonucleases, comprising a DNA binding domain fused to one part of an endonuclease; specificity is ensured by a second TALEN that binds to an adjacent sequence and provides the other part of the nuclease. Each TALEN fusion protein can be specifically designed to bind to a DNA sequence flanking the region of the genome to be targeted resulting in cleavage of genomic DNA at the target site. Subsequent non-homologous end joining repair at the cleavage site results in the excision or insertion of a variable number of base pairs. If the TALENS are targeted to a protein coding exon, typically exon 1, the result is likely to be a mutant gene coding for a non-functional truncated protein [45].

Another new technology for gene editing is the clustered regularly interspersed short palindromic repeat (CRISPR)-Cas9 system [43]. Similar to TALENs, the CRISPR-Cas9 system is another method to induce double stranded breaks in genomic DNA. The breaks are repaired by a cellular process called non-homologous end-joining. This process is error prone and leads to insertions or deletions that can disrupt the coding region of the targeted gene. The advantage of the CRISPR system is that it is driven by a guide strand RNA that undergoes standard base-pairing with the endogenous genomic target. The cas9 nuclease is recruited to the guide strand RNA and therefore to the target sequence. The simplicity of this system makes it practical for most labs experienced in molecular biology.

The key feature of these new technologies is that unlike many antisense knock-down and silencing strategies that have been widely used over the last 15 years to post-transcriptionally inhibit genes, TALENs and CRISPR-cas9 produce heritable gene

Europe PMC Funders Author Manuscripts

disruption and germ line transmission of targeted mutation. This technology is effective in zebrafish and allows the efficient production of mutation targeted to genes of interest. To date, approximately 16 loci have been associated with PD [46]. Creating zebrafish mutants to model for each of these genes is now a practical and achievable goal.

4 Evolutionary algorithms

EAs [47] are a form of evolutionary computation and members of the artificial intelligence family, or more precisely computational intelligence, as they depend on a form of learning inspired by Darwinian evolution. They are in effect a number (or *population*) of candidate solutions (*individuals*) to a classification problem that are repeatedly refined (or *evolved*) over a number of iterations (*generations*) until a suitably accurate classifier algorithm is obtained or the computational resources have been exhausted.

The procedure for finding a classifier, for example to discriminate PD patient data from healthy control data, can be summarised as follows: A population of individuals (candidate solutions) is randomly initialised. The effectiveness or *fitness* of each individual to correctly classify data previously obtained from PD patients and healthy controls is determined using a *fitness function*. The fittest individual (the one with the highest fitness score determined by the fitness function) is retained and the others discarded. Copies (or *clones*) of this fittest individual are then generated and subtly modified (or *mutated*) to form a new population of individuals. The fitness of this new population of individuals is then evaluated in the same way using the fitness function and the process is repeated over a number of generations until a sufficiently fit classifier is obtained or the number of predetermined generations has been reached.

Many different types of evolutionary algorithm have been developed which specify not only the characteristics of the evolutionary process but also the representation of the individual candidate solutions. For our work Cartesian genetic programming (CGP) [48] is used, which does not adopt the tree structure representation of traditional genetic programming. In CGP each individual is represented by a network of processing nodes that are arranged in a non-cyclic directed graph (two-dimensional grid).

A simplified example of a CGP network is shown in Fig. 2 with four central processing nodes arranged in a 2 × 2 grid. Two inputs, I/P 0 and I/P 1, provide input values to the network and an output, O/P 0, its result. Each node within the network comprises a function taken from a predefined set (an example of which is given in Table 1). Outputs from each of the nodes are connected to inputs of nodes to the right or the output. Nodes in the network are numbered consecutively starting at zero with the first of the input nodes, as shown in the top right-hand side corner of each node in Fig. 2. This allows the nodes within the network to be represented by a string of integers (or *chromosome*), an example of which is given in Fig. 3.

The chromosome representing any CGP network consists of triplets (referred to as *genes*), providing values for each respective (non-input) node in the network, for example, the first triplet relates to node number 2, the second to the node number 3, and so on. The first two genes of each triplet specify preceding nodes in the network that provide values to the

node's two inputs. The third value specifies the index of the function in the function set (shown in Table 1) to be applied to the values presented at the inputs. The final integer in the chromosome specifies the node that provides the value to be presented at the output node O/P 0. A CGP network comprising three columns of 10 rows as shown in Fig. 4 is more representative of that typically used.

A number of these chromosomes form the individuals of a population, which are initialised with random values. Each chromosome is then used to configure the network and subsequently calculate a result for the problem under consideration. The result presented at the output of this network is compared with the desired result, and a fitness score derived, which is then associated with the respective individual's chromosome. After all the individuals in the population have been evaluated in this manner, the fittest is retained and the remaining discarded. The subsequent generation of individuals is then constructed from this fittest individual and its clones, which are subjected to a variation function, typically a conventional mutation operation according to a predefined probability. Mutation is simply achieved by randomly altering the integer values of the chromosome within a valid range.

The training of the CGP networks is continued until either perfect fitness is achieved (according to the predetermined fitness function) or other criteria is met, such as a maximum number of generations. The fittest network can then be evaluated by applying conventional testing and validation stages with data not used in the training stage.

There are two ways in which CGP provides an advantage over other processing techniques. First of all, for highly non-linear, complex data sets such as that found in measurement of human movement disorders, CGP has been shown to evolve high performance classifiers. Second, unlike many processing techniques, once a high performing classifier has been evolved, a mathematical expression defining this classifier can be easily obtained by decoding the resulting CGP network. This can provide valuable insight into which data obtained from the physical measurements of movement have been used in the evolved classifier and are therefore an important distinguishing feature of the condition [49].

4.1 Application of EAs to measure movement disorders in humans

CGP EAs have been used by our group to process data obtained from a range of sensors to diagnose and monitor PD in humans:

Diagnosis of PD (Fig. 5a)—We attached EM tracking sensors to the finger and thumb of 49 PD patients and 41 age-matched healthy controls whilst undertaking a finger tapping task for 30 s – a common conventional clinical evaluation. The majority (80%) of the patients had only clinically mild bradykinesia. The movement data collected by the EM sensors was analysed by custom-written EAs and compared with clinical diagnosis using receiver operating characteristic (ROC) curves. The best classifiers had an area under the ROC curve (AUC) of 0.9 corresponding to predictive accuracies of 80–90% depending on choice of threshold. This suggests that automated tests of bradykinesia could have a role to play in providing objective information to support a tentative clinical diagnosis of PD [15].

Recognition and monitoring of levodopa-induced dyskinesia (Fig. 5b)— Matchbox-sized accelerometer/gyroscope devices were worn on the limbs, head and trunk of PD patients for periods of up to 24 hours and movement data was analysed using EAs. Provisional results show that the device has excellent accuracy (AUC 0.9) for monitoring severe and moderate degrees of dyskinesia. Providing this kind of objective clinical information to clinicians would enable them to make more informed decisions regarding administration of medication and also could be used to evaluate new drugs for its treatment [50].

A unified test for neurodegenerative disorders (Fig. 5c)—Data gloves and motion tracking sensors are used to assess PD patients and healthy controls as they reach and grasp a cylinder under several different conditions, for example using an auditory cue, a visual cue and with eyes closed. Data analysis is currently underway. EAs will be applied to the movement data to classify healthy controls and PD patients as well as further classifying PD patients into those with and without cognitive impairment. Previous research using standard statistical analysis has shown differences in reaching and grasping between patients with PD and other neurodegenerative conditions such as Alzheimer's disease and corticobasal ganglionic degeneration [51, 52].

In all of these cases the processing of data resulting from the patient measurements is treated in the same way. The data, obtained from any of the sensors considered above, comprises a stream of co-ordinates that describe the motion of the patient, whether it is in response to a specific task (such as finger tapping, drawing or reach and grasp task) or resulting from unconstrained movement, as in the case of the wearable accelerometer/gyroscope devices. The data is preprocessed and then presented to each of the individual CGP networks in the population through a moving windowing operation that presents 10 data points at a time until the entire data stream has been presented.

4.2 Application of EAs to measure movement disorders in animal models

Once a genetically altered animal is generated, it is important to determine how good a model it is for the human disease being studied. For animal models of PD, there are methods to identify and measure numbers of dopamine producing neurons as well as physiological methods to assess the function of these neurons [53, 54]. An impaired diving behaviour has been noted in zebrafish lacking dopamine [55]. However, it would be an over-interpretation to align this aspect of fish behaviour to symptoms of PD patients. A hallmark of humans with PD is bradykinesia, which can be measured in humans and quantified using EAs as described above.

In zebrafish models, a high frame rate video of fish swimming has been analysed using computer vision techniques to generate a minimal set of data characterising how the body of the fish flexes, as shown in Fig. 6. We are currently developing methods to measure movement of the fish, in specially designed tanks, to enable EAs to be evolved. Adapting the protocols used to diagnose PD in humans for use in genetic models will provide an assessment of the effects of disrupting a specific PD gene. A zebrafish model of PD has the

potential for high throughput *in vivo* drug screening that is not economically viable using other vertebrate models.

In Drosophila models, we have analysed the videos of the PER by importing them into MATLAB, defining the frames in which the proboscis is extended, and then tracking its movement. This has allowed us to adopt the same strategy for analysis of movement in the fly as in the human. The results presented in Fig. 7 show the ability of an EA to train classifiers that discriminate (a) wild-type controls (46 flies) from *TH > hLRRK2* (47 flies) with an AUC of 0.78, and (b) *TH > LRRK2-G2019S* mutation (45 flies) from *TH > hLRRK2* (47 flies) with an AUC of 0.78 according to this captured movement.

In conclusion, the fly model of PD recapitulates PD in a very simple sensory reflex, where activity of a single dopaminergic neuron is manipulated by a PD-related mutation to produce both akinesia and hypokinesia. It is amenable to genetic dissection, to identify the impact of *G2019S* expression on functioning neuronal components, and to analysis by the same EAs as used in patients. The results presented here demonstrate an exciting opportunity to relate the change in kinase activity in *G2019S* to subsequent cellular, neuronal and movement disorders. This offers the powerful potential to test the impact of existing drugs (e.g. L-DOPA) and of first-in-vivo testing of novel tool compounds (e.g. BMPPB-32 [53]) in a quantifiable manner.

5 Summary and future work

This paper has reviewed how EAs can be successfully applied to the assessment of movements in humans to classify PD patients from healthy controls and classify severity of dyskinesia in PD patients. We have presented new results that demonstrate the ability of EAs to classify wild-type Drosophila from those with PD related genetic mutations. The ability to use CGP EAs in both human and animal models of PD represents an exciting development, as decoding of the CGP network will allow distinguishing movement features to be analysed and compared across species. This has great potential for increasing understanding of PD as well as the effects of new and existing drug therapies.

Work to apply EAs to zebrafish models of PD and to use EAs to classify degrees of cognitive impairment in humans with PD by analysing movement data is under way. Ultimately we hope that the application of EAs to movement data in animals and humans will provide a unifying model for motor dysfunction in PD.

6 Acknowledgments

We acknowledge the support of the following organisations: The Wellcome Trust [ref: 097829] through the Centre for Chronic Diseases and Disorders (C2D2) at the University of York, the Royal Academy of Engineering Enterprise Fellowship Scheme, Parkinson's UK and the National Institute for Health Research, through the Dementias and Neurodegenerative Diseases Research Network.

7 References

1. Parkinson's UK. Parkinson's prevalence in the United Kingdom. 2009. Parkinson's UK; London: 2012. p. 1-13.

2. de Lau LM, Breteler MM. Epidemiology of Parkinson's disease. The Lancet Neurology. 2006; 5(6): 525–535. [PubMed: 16713924]
3. Guttmacher AE, Collins FS, Nussbaum RL, et al. Alzheimer's disease and Parkinson's disease. N. Engl. J Med. 2003; 348(14):1356–1364. [PubMed: 12672864]
4. Meara J, Bhowmick BK, Hobson P. Accuracy of diagnosis in patients with presumed Parkinson's disease. Age Ageing. 1999; 28(2):99–102. [PubMed: 10350403]
5. Schrag A, Ben-Shlomo Y, Quinn N. How valid is the clinical diagnosis of Parkinson's disease in the community? J. Neuro. Neurosurg. Psychiatry. 2002; 73(5):529–534. [PubMed: 12397145]
6. Rajput A, Rozdilsky B, Rajput A. Accuracy of clinical diagnosis in Parkinsonism – a prospective study. Can. J. Neurol. Sci.. Le J. Can. des Sci. Neurolog. 1991; 18(3):275–278. [PubMed: 1913360]
7. Hughes AJ, Daniel SE, Kilford L, et al. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J. Neurol. Neurosurg. psychiatry. 1992; 55(3): 181–184. [PubMed: 1564476]
8. Hughes AJ, Daniel SE, Lees AJ. Improved accuracy of clinical diagnosis of lewy body Parkinson's disease. Neurology. 2001; 57(8):1497–1499. [PubMed: 11673599]
9. Bajaj NP, Gontu V, Birchall J, et al. Accuracy of clinical diagnosis in tremulous Parkinsonian patients: A blinded video study. J. Neurol. Neurosurg. Psychiatry. 2010; 81(11):1223–1228. [PubMed: 20547625]
10. Golbe LI, Pae J. Validity of a mailed epidemiological questionnaire and physical self-assessment in Parkinson's disease. Movement Disorders. 1988; 3(3):245–254. [PubMed: 3193962]
11. Vitale C, Pellecchia M, Grossi D, et al. Unawareness of dyskinesias in Parkinson's and huntington's diseases. Neurolog. Sci. 2001; 22(1):105–106. [PubMed: 11487181]
12. Amanzio M, Monteverdi S, Giordano A, et al. Impaired awareness of movement disorders in Parkinson's disease. Brain Cogn. 2010; 72(3):337–346. [PubMed: 19914762]
13. Tripoliti EE, Tzallas AT, Tsipouras MG, et al. Automatic detection of freezing of gait events in patients with Parkinson's disease. Comput. Methods Programs Biomed. 2013; 110(1):12–26. [PubMed: 23195495]
14. Tsipouras MG, Tzallas AT, Rigas G, et al. An automated methodology for levodopa-induced dyskinesia: assessment based on gyroscope and accelerometer signals. Artif. Intell. Med. 2012; 55(2):127–135. [PubMed: 22484102]
15. Lones MA, Smith SL, Alty JE, et al. Evolving classifiers to recognize the movement characteristics of Parkinson's disease patients. IEEE Trans. Evol. Comput. 2014; 18(4):559–576.
16. Foreman KB, Singer ML, Addison O, et al. Effects of dopamine replacement therapy on lower extremity kinetics and kinematics during a rapid force production task in persons with Parkinson disease. Gait Posture. 2014; 39(1):638–640. [PubMed: 23968974]
17. Rao AS, Dawant BM, Bodenheimer RE, et al. Validating an objective video-based dyskinesia severity score in Parkinson's disease patients. Parkinsonism Related Disord. 2013; 19(2):232–237. [PubMed: 23182314]
18. Mate MA, Cobeta I, Jiménez-Jiménez FJ, et al. Digital voice analysis in patients with advanced Parkinson's disease undergoing deep brain stimulation therapy. J. of Voice. 2012; 26(4):496–501. [PubMed: 21704492]
19. Smith SL, Gaughan P, Halliday DM, et al. Diagnosis of Parkinson's disease using evolutionary algorithms. Genet. Programming Evolvable Machines. 2007; 8(4):433–447.
20. Wesnes KA, Burn DJ. The prevalence and nature of mild cognitive impairment in Parkinson's disease (PD-MCI) identified using automated cognitive tests. Mov. Disord. 2013; 28(Suppl 1):568. [abstract].
21. Pfeiffer HCV, Løkkegaard A, Zoetmulder M, et al. Cognitive impairment in early-stage non-demented Parkinson's disease patients. Acta Neurol. Scand. 2014; 129(5):307–318. [PubMed: 24117192]
22. Dawson TM, Ko HS, Dawson VL. Genetic animal models of Parkinson's disease. Neuron. 2010; 66(5):646–661. [PubMed: 20547124]
23. Cauchi RJ, Van Den Heuvel M. The fly as a model for neurodegenerative diseases: is it worth the jump? Neurodegenerative Dis. 2006; 3(6):338–356. [PubMed: 17192723]

24. Reiter LT, Potocki L, Chien S, et al. A systematic analysis of human disease-associated gene sequences in drosophila melanogaster. Genome Res. 2001; 11(6):1114–1125. [PubMed: 11381037]
25. Whitworth AJ. Drosophila models of Parkinson's disease. Adv. Genet. 2011; 73:1–50. [PubMed: 21310293]
26. Barone MC, Bohmann D. Assessing neurodegenerative phenotypes in drosophila dopaminergic neurons by climbing assays and whole brain immunostaining. J Visualized Experiments: JoVE. 2013; (74) [PubMed: 23644755]
27. Gargano JW, Martin I, Bhandari P, et al. Rapid iterative negative geotaxis (Ring): a new method for assessing age-related locomotor decline in drosophila. Exp. Gerontol. 2005; 40(5):386–395. [PubMed: 15919590]
28. Phillips M, Roberts S, Kladt N, et al. An automated, high-throughput climbing assay for behavioral screening in drosophila. Front. Behav. Neurosci. :357.
29. Drobysheva D, Ameel K, Welch B, et al. An optimized method for histological detection of dopaminergic neurons in drosophila melanogaster. J. Histochem. Cytochem. 2008; 56(12):1049–1063. [PubMed: 18574253]
30. Mao Z, Davis RL. Eight different types of dopaminergic neurons innervate the drosophila mushroom body neuropil: anatomical and physiological heterogeneity. Front. Neural Circuits. 2009; 3
31. Friggi-Grelin F, Coulom H, Meller M, et al. Targeted gene expression in drosophila dopaminergic cells using regulatory sequences from tyrosine hydroxylase. J. Neurobiol. 2003; 54(4):618–627. [PubMed: 12555273]
32. Martin I, Kim JW, Lee BD, et al. Ribosomal protein S15 phosphorylation mediates Lrrk2 neurodegeneration in Parkinson's disease. Cell. 2014; 157(2):472–485. [PubMed: 24725412]
33. Inagaki HK, de-Leon SB-T, Wong AM, et al. Visualizing neuromodulation *in vivo*: tango-mapping of dopamine signaling reveals appetite control of sugar sensing. Cell. 2012; 148(3):583–595. [PubMed: 22304923]
34. Scott K, Brady R, Cravchik A, et al. A Chemosensory gene family encoding candidate gustatory and olfactory receptors in drosophila. Cell. 2001; 104(5):661–673. [PubMed: 11257221]
35. Marella S, Mann K, Scott K. Dopaminergic modulation of sucrose acceptance behavior in drosophila. Neuron. 2012; 73(5):941–950. [PubMed: 22405204]
36. Gordon MD, Scott K. Motor control in a drosophila taste circuit. Neuron. 2009; 61(3):373–384. [PubMed: 19217375]
37. Brand AH, Perrimon N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development. 1993; 118(2):401–415. [PubMed: 8223268]
38. Duffy JB. Gal4 system in drosophila: a fly geneticist's swiss army knife. Genesis. 2002; 34(1–2): 1–15. [PubMed: 12324939]
39. Looger LL, Griesbeck O. Genetically encoded neural activity indicators. Curr. Opin. Neurobiol. 2012; 22(1):18–23. [PubMed: 22104761]
40. Berger J, Currie PD. Zebrafish models flex their muscles to shed light on muscular dystrophies. Dis. Models Mech. 2012; 5(6):726–732. [PubMed: 23115202]
41. Granato M, Van Eeden F, Schach U, et al. Genes controlling and mediating locomotion behavior of the zebrafish embryo and larva. Development. 1996; 123(1):399–413. [PubMed: 9007258]
42. Flinn LJ, Keatinge M, Bretaud S, et al. Tigarb causes mitochondrial dysfunction and neuronal loss in pink1 deficiency. Ann. Neurol. 2013; 74(6):837–847. [PubMed: 24027110]
43. Harrison MM, Jenkins BV, O'Connor-Giles KM, et al. A crispr view of development. Genes Dev. 2014; 28(17):1859–1872. [PubMed: 25184674]
44. Lei Y, Guo X, Liu Y, et al. Efficient targeted gene disruption in xenopus embryos using engineered transcription activator-like effector nucleases (Talens). Proc. of the National Academy of Sciences. 2012:201215421. [PubMed: 23045671]
45. Bedell VM, Wang Y, Campbell JM, et al. *In vivo* genome editing using a high-efficiency talen system. Nature. 2012; 491(7422):114–118. [PubMed: 23000899]

46. Corti O, Lesage S, Brice A. What genetics tells us about the causes and mechanisms of Parkinson's disease. Physiol. Rev. 2011; 91(4):1161–1218. [PubMed: 22013209]
47. Chiong, R.; Weise, T.; Michalewicz, Z. Variants of evolutionary algorithms for real-world applications. Springer; 2012.
48. Miller, JF. Cartesian genetic programming. Springer; 2011.
49. Lones M, Smith SL, Harris AT, et al. Discriminating normal and cancerous thyroid cell lines using implicit context representation cartesian genetic programming. 2010 IEEE Congress on Evolutionary Computation (CEC). 2010
50. Lones, MA.; Alty, JE.; Duggan-Carter, P., et al. Classification and characterisation of movement patterns during levodopa therapy for Parkinson's disease. Proc. of the 2014 Conf. Companion on Genetic and Evolutionary Computation Companion; 2014;
51. Caselli RJ, Stelmach GE, Caviness JN, et al. A kinematic study of progressive apraxia with and without dementia. Mov. Disord. 1999; 14(2):276–287. [PubMed: 10091622]
52. Castiello U, Stelmach GE, Lieberman AN. Temporal dissociation of the prehension pattern in Parkinson's disease. Neuropsychologia. 1993; 31(4):395–402. [PubMed: 8502374]
53. Afsari F, Christensen KV, Smith GP, et al. Abnormal visual gain control in a Parkinson's disease model. Hum. Mol. Genet. 2014; 23(17):4465–4478. [PubMed: 24718285]
54. Wade AR, Elliott CJ. Could the detection of visual disturbances associated with Parkinson's disease genes in flies lead to new treatments for the disease? Neurodegenerative Dis. Manag. 2014; 4(4):291–293. [PubMed: 25313984]
55. Formella I, Scott EK, Burne TH, et al. Transient knockdown of tyrosine hydroxylase during development has persistent effects on behaviour in adult zebrafish (*Danio Rerio*). PloS one. 2012; 7(8):e42482. [PubMed: 22879998]

Europe PMC Funders Author Manuscripts

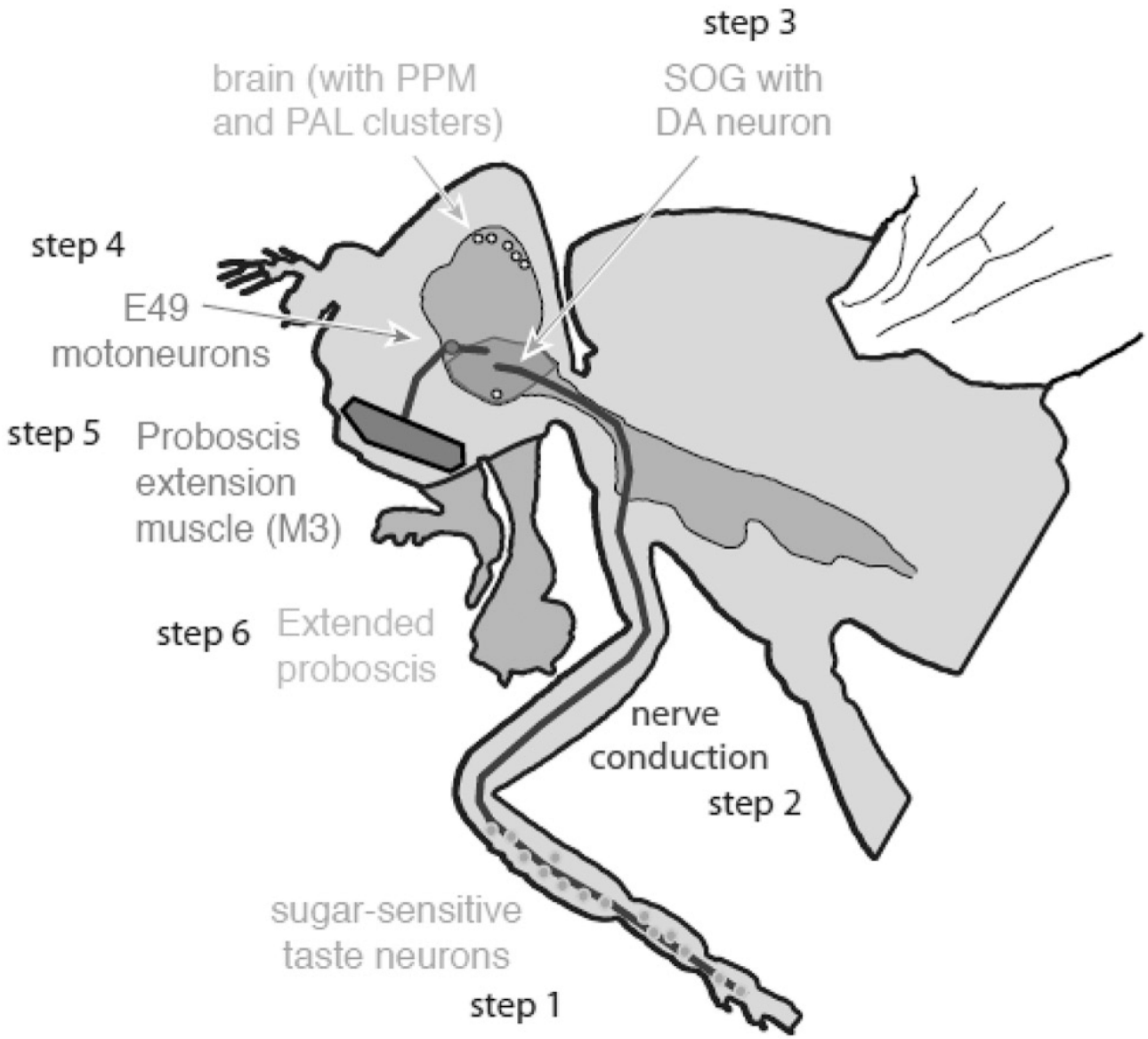

Fig. 1.

Outline schematic of the PER pathway, showing the essential neuronal circuit. Sugar applied to the legs (step 1) stimulates the chemosensory neurons which project to the SOG. This contains a single dopaminergic neuron (yellow, DA, step 3), which modulates the connection between the sensory endings and the E49 motoneurons. When the sensory - motor relay is permitted by dopaminergic activity, the E49 neurons fire (step 4), causing contraction of the muscle, M3 (step 5), and extension of the proboscis (step 6) [After Scott et al., [34]]

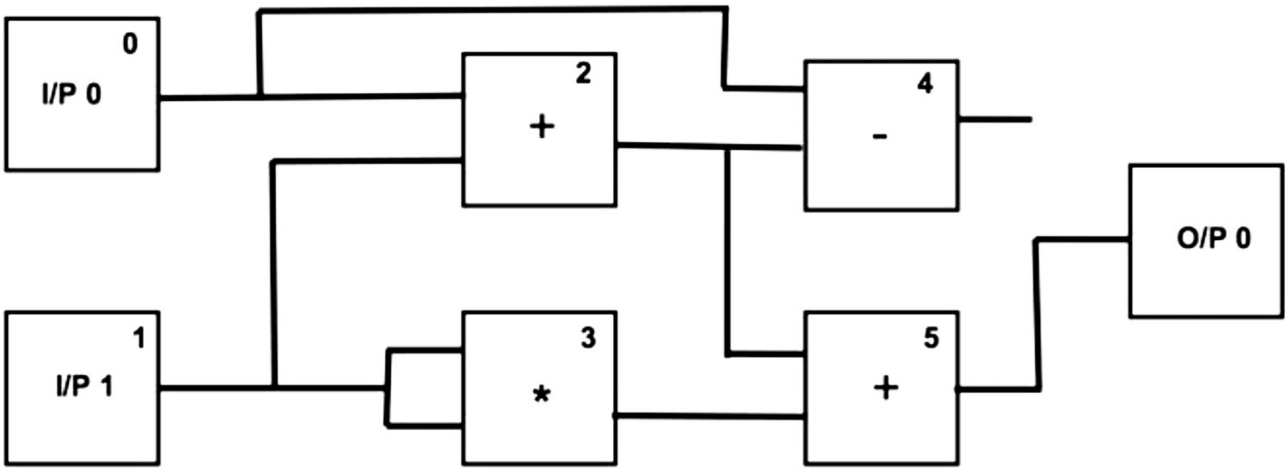

Fig. 2.
Example Cartesian Genetic Program. Node number is specified in the top right-hand side corner of each node, the node function is specified in the centre

Europe PMC Funders Author Manuscripts

Fig. 3.
Example chromosome for configuration of the CGP network

Europe PMC Funders Author Manuscripts

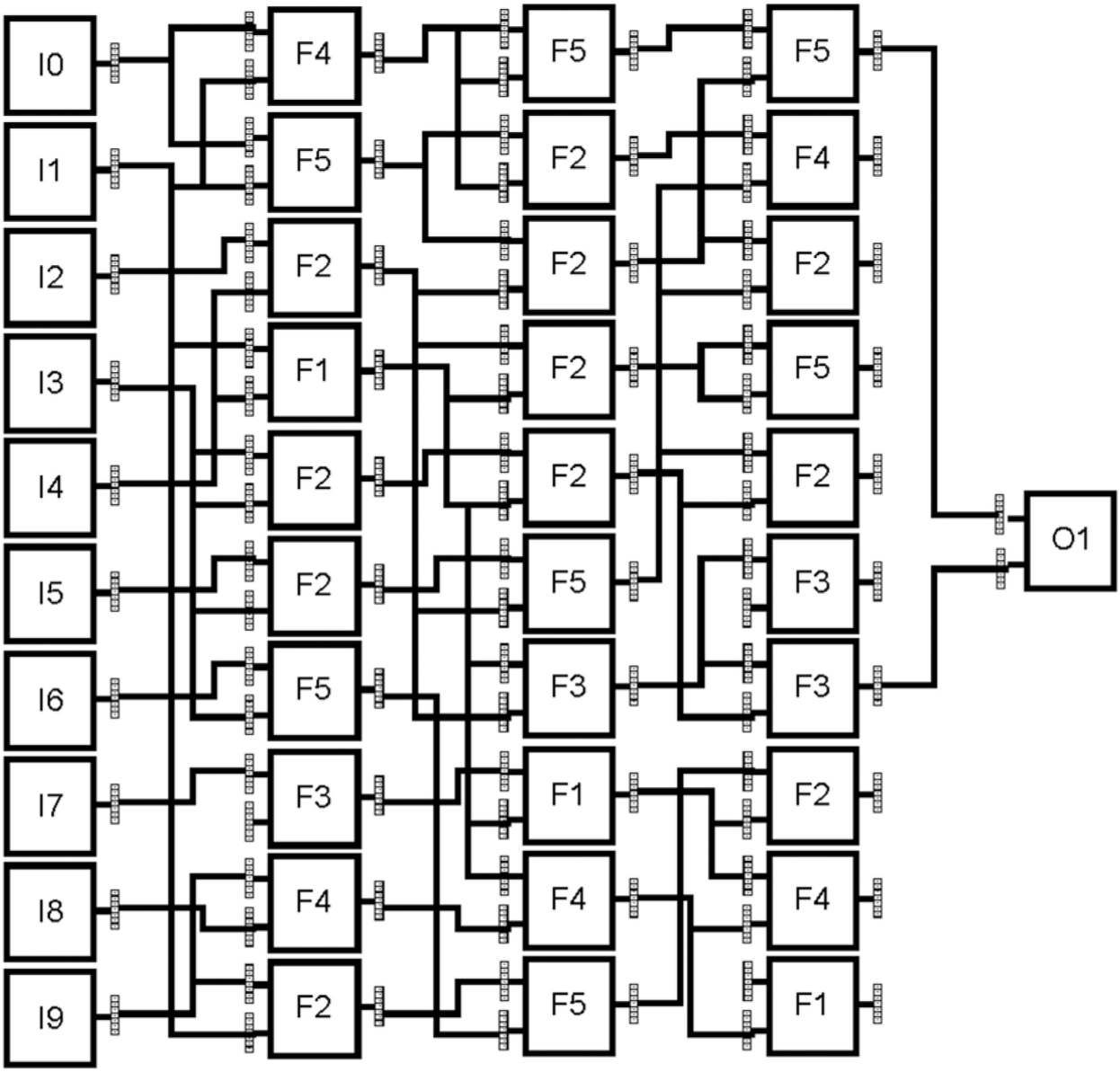

Fig. 4.
Typical CGP network configuration with 3 columns and 10 rows of nodes

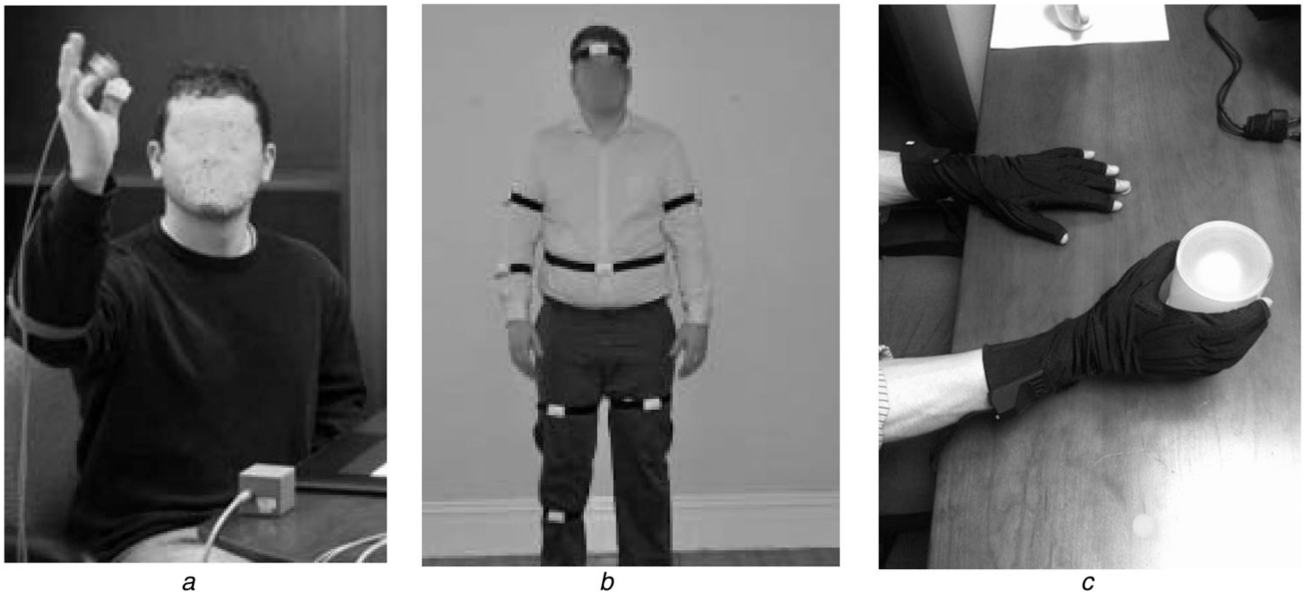

Fig. 5.

Use of sensors for diagnosis, monitoring and differentiation of PD

a Finger tapping task using EM tracking sensors

b Unconstrained movement measurement using accelerometer/gyroscopes

c Reach and grasp task using computer data gloves

Europe PMC Funders Author Manuscripts

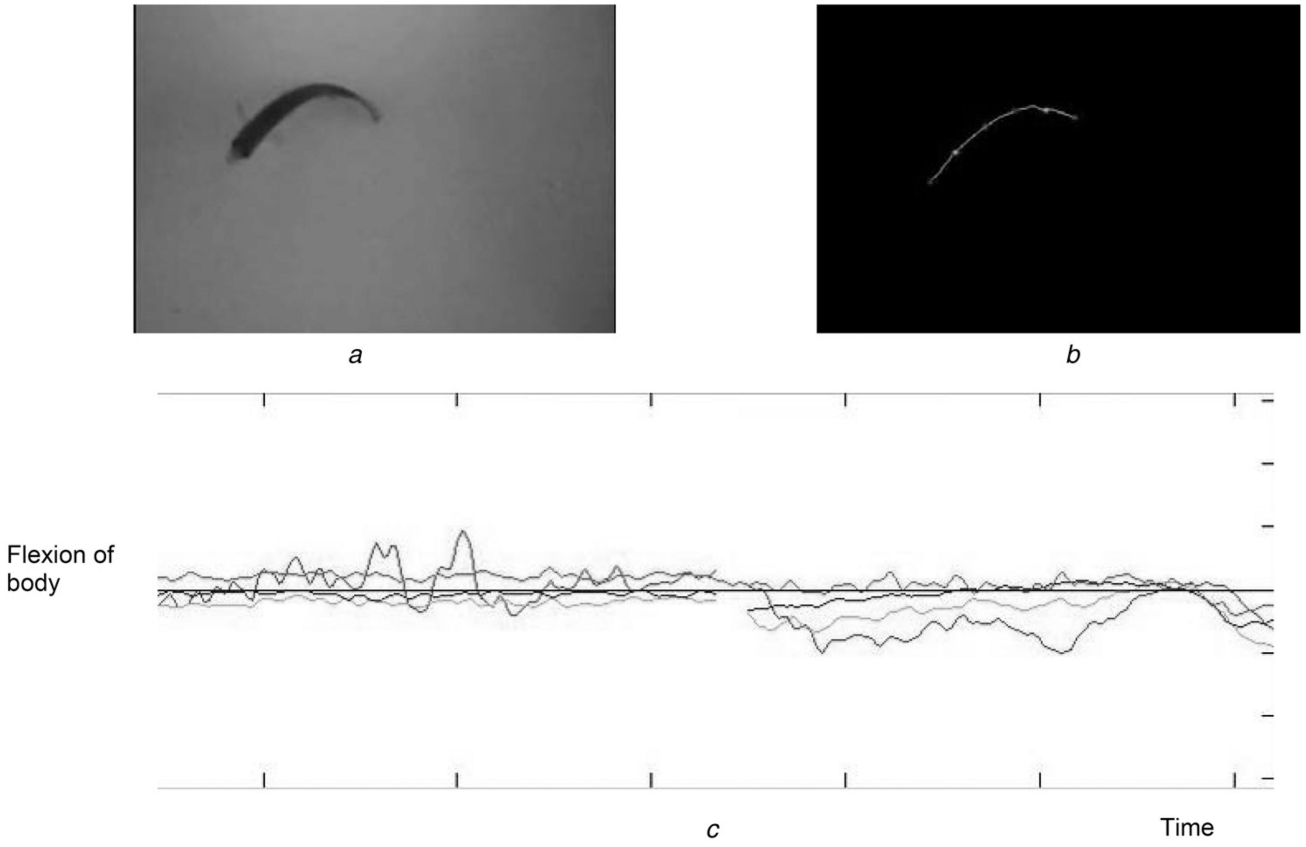

Fig. 6.

Automated tracking of the body flexion of the zebra fish

a Is the source video of the zebra fish

b Result of processing, with coloured knots representing segments of the body

c Plot of the movement of the fish represented by these knots with respect to the direction of travel

Europe PMC Funders Author Manuscripts

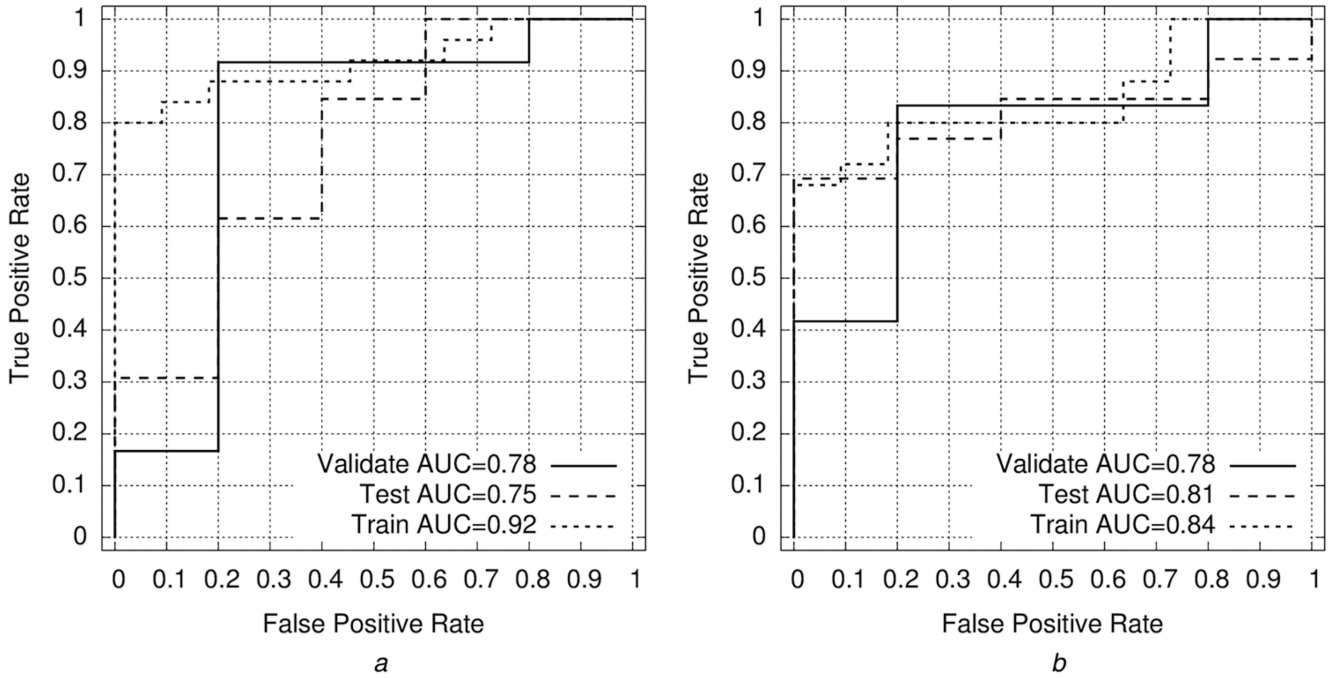

Fig. 7.

ROC curves demonstrating the evolutionary algorithm to correctly classify *a* Wild-type (*CS* × w^{1118} outcross) and *TH* > *hLRRK2* with an AUC of 0.78 *b TH* > *LRRK2-G2019S* mutation and wild-type *TH* > *hLRRK2* with an AUC of 0.78 (flies aged 8 days)

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Table 1

Example function set

Function reference	Function
1	+
2	−
3	*
4	/