Elsevier Editorial System(tm) for Journal of Neuroscience Methods

Manuscript Draft

Manuscript Number: JNEUMETH-D-16-00606R1

Title: Analysis of Locomotor Behavior in the German Mouse Clinic

Article Type: SI:Measuring Behavior 2016

Section/Category: Basic Neuroscience

Keywords: CatWalk; Gait; Principal Component Analysis; Mouse; Locomotion; Activity; Phenotyping

Corresponding Author: Dr. Annemarie Zimprich,

Corresponding Author's Institution: Helmholtz Zentrum München

First Author: Annemarie Zimprich

Order of Authors: Annemarie Zimprich; Manuela A Östereicher; Lore Becker; Petra Dirscherl; Luise Ernst; Helmut Fuchs; Valerie Gailus-Durner; Lillian Garrett; Florian Giesert; Lisa Glasl; Angelika Hummel; Jan Rozman; Martin Hrabê de Angelis; Daniela M Vogt Weisenhorn; Wolfgang Wurst; Sabine M Hölter

Abstract: Background
Generation and phenotyping of mutant mouse models continues to increase along with the search for the most efficient phenotyping tests. Here we asked if a combination of different locomotor tests is necessary for comprehensive locomotor phenotyping, or if a large data set from an automated gait analysis with the CatWalk system would suffice.

New Method
First we endeavored to meaningfully reduce the large CatWalk data set by Principal Component Analysis (PCA) to decide on the most relevant parameters. We analyzed the influence of sex, body weight, genetic background and age. Then a combination of different locomotor tests was analyzed to investigate the possibility of redundancy between tests.

Result
The extracted 10 components describe 80% of the total variance in the CatWalk, characterizing different aspects of gait. With these, effects of CatWalk version, sex, body weight, age and genetic background were detected. In addition, the PCA on a combination of locomotor tests suggests that these are independent without significant redundancy in their locomotor measures.

Comparison with existing methods
The PCA has permitted the refinement of the highly dimensional CatWalk (and other tests) data set for the extraction of individual component scores and subsequent analysis.

Conclusion
The outcome of the PCA suggests the possibility to focus on measures of the front and hind paws, and one measure of coordination in future

experiments to detect phenotypic differences. Furthermore, although the CatWalk is sensitive for detecting locomotor phenotypes pertaining to gait, it is necessary to include other tests for comprehensive locomotor phenotyping.

Ethical Standards Agreement

Click here to download Ethical Standards Agreement: ethical_standards_agreement.doc

All submissions to the Journal of Neuroscience Methods must contain experiments that conform to the ethical standards printed below.

To confirm your agreement with this, you are required to include the following statement in your cover letter indicating your agreement with these standards: "I have read and have abided by the statement of ethical standards for manuscripts submitted to the Journal of Neuroscience Methods

ETHICAL STANDARDS:

- The authors declare that all experiments on human subjects were conducted in accordance with the Declaration of Helsinki http://www.wma.net and that all procedures were carried out with the adequate understanding and written consent of the subjects.

- The authors also certify that formal approval to conduct the experiments described has been obtained from the human subjects review board of their institution and could be provided upon request.

- If the studies deal with animal experiments, the authors certify that they were carried out in accordance with the National Institute of Health Guide for the Care and Use of Laboratory Animals (NIH Publications No. 80-23) revised 1996 or the UK Animals (Scientific Procedures) Act 1986 and associated guidelines, or the European Communities Council Directive of 24 November 1986 (86/609/EEC).

- The authors also certify that formal approval to conduct the experiments described has been obtained from the animal subjects review board of their institution and could be provided upon request.

- The authors further attest that all efforts were made to minimize the number of animals used and their suffering.

- If the ethical standard governing the reported research is different from those guidelines indicated above, the authors must provide information in the submission cover letter about which guidelines and oversight procedures were followed.

- The Editors reserve the right to return manuscripts in which there is any question as to the appropriate and ethical use of human or animal subjects.

*Revision Notes

Point-by-point response to reviewer´s suggestions

Ref.: Ms. No. JNEUMETH-D-16-00606
Analysis of Locomotor Behavior in the German Mouse Clinic
Journal of Neuroscience Methods

Dear Dr. Crunelli,

Dear Reviewers,

Thank you very much for the very helpful critical review of our submitted manuscript. As we received encouraging responses we are happy to address the remaining concerns by restructuring/rewriting different parts of the manuscript and including more information and details. Please find attached the revised version of our manuscript and below the point-by-point response to the reviewer's questions.

Having made these changes we now hope that the manuscript can be accepted for publication in the Special Issue: Measuring Behavior 2016.

With kind regards,

Annemarie Zimprich

Reviewers' comments:

Reviewer #1:

The present manuscript submitted by Zimprich and collaborators reports a statistical model dedicated to analyse and reduce by Principal Component Analysis a large set of data coming from a multitude of parameters recorded by currents video tracing software such as Catwalk into component which could describe mice motor phenotype.

1)

I found the idea to focus on large set of data relatively interesting and challenging, and the manuscript submitted by Zimprich and collaborators is perfectly in line with this. However, the manuscript is quite difficult to follow in those terms. The way the article is written and it corpus is exclusively linked to statistical correlations and descriptions which make the reading quite difficult to follow. Moreover, the absence of behavioural raw data in parallel to the statistical analysis make difficult as well to attribute a judgement about the quality of the data analysis and the model itself. It would have been nice to have side by side, the Raw Data showing differences between groups in terms of motor differences and the analysis obtained via PCA.

We improved the traceability by rearranging parts and included original data of parameters to illustrate the major points (please see also answer to question 10)).

2)

Moreover, I would advise the author to make the manuscript read by a native English speaker, the quality of the writing is for me not good enough for Journal of Neuroscience Methods.

We now have had a native English speaker critically reading and correcting the manuscript.

Comments:

INTRODUCTION

3)

Regarding the introduction section, my first comment is in regard of the term locomotor phenotype. It might be good to have a clear definition in this section of "locomotor phenotype" which wants to be detected with the Catwalk. Indeed, this term is a bit confusing because this task has been designed mainly to assess motor performance but not proper locomotion or activity which is something quite different.

We inserted a definition of “locomotor phenotype”: “In this paper we focus on locomotor phenotypes, whereby the definition of “locomotor phenotype” is used very broadly. We include gait phenotypes (as analyzed by the CW), as well as activity (as in the Open Field (OF), home cage (HC) and SHIRPA) and motor ability phenotypes (as in the Grip Strength (GS), Rotarod (RR) and Vertical Pole (VP) test) into this term.” (page 3, line 29ff)

4)

Because the methods and result section of the manuscript rely mainly on statistics and PCA it would have been good to mentioned literature which use similar sort of analyses for behaviour.

We quoted literature where a PCA was run on behavioral data (compare page 3, line 19ff) but now elaborate on this further in the Introduction and Discussion. We also included the citation from Ohl et al, which Reviewer #2 mentioned (also see question 26).

5)

- The second paragraph of the introduction is a little bit confusing which is probably related to writing issues. Because it is stated that in the manuscript the focus is carried on locomotor phenotype associated with neurological disease such as Parkinson's disease, Amyotrophic Lateral Sclerosis and Attention Deficit Hyperactivity Syndrome, I expected to see the PCA analysis carried on different mouse model used for these pathologies.

To circumvent this confusion we excluded this sentence from the introduction.

METHODS

6)

- The method part for me needs to be written again or at least better organized, it appear really messy. It would be good to have it extended by information about the animals, a summary table with how many animals has been used for each separate analysis, the age, and genotype would help.

We completely restructured the Methods and added a table in the Supporting Information for details of animal numbers per analysis as suggested (see Table S1).

7)

- I understood that the manuscript has been written to demonstrate the power of PCA for analysing a large pool of data, but it would be good to have for each behavioural test used a description of the different behavioural apparatus with the parameters used being explained or at least defined in a supplementary table.

We included brief descriptions of how the different behavioral tests were performed in the Methods and added tables listing the parameters used in the statistical analysis in the Supporting Information (Table S2 and S7).

8)

- It would good to know as well the type of protocol which has been used for the tasks, for example are the 1499 mice used for the statistical analyses of the 113 parameters of the CW been assessed in the same exact protocol?

Yes, all mice were subjected to the same CW testing protocol. The animals were all transferred to the testing room at least 30 minutes prior to testing. All animals were tested in a darkened room. Two to 4 continuous uninterrupted runs per animal were used for analysis. Only the apparatus version changes, i.e. from 7.1 to XT. The main difference between the versions is the time-related resolution of the cameras. In more detail the CW XT version has a lid above the runway, which has a red light integrated and the animals run towards a so called “goal box” underneath which their home cage is situated. The CW 7.1 version does not have a lid nor a “goal box”.

9)

- In the statistical part, it would nice to have a little definition of what a z transformation and an oblimin rotation are. I understood that the z transformation has been chosen to normalise a set of different parameters in order to compare its together as well as oblimin rotation has been chosen to regroup positive correlation in component, but I think a little definition of each in the method part would be a nice plus.

We elucidate more on the Z-transformation and the oblique rotation in the Methods (page 6, line 19ff).

RESULTS

10)

- Globally the results part is really hard to follow. The fact that all paragraphs are a succession of statistical descriptions make the results part difficult to follow and to rely on actual mouse behaviour.

We included original data to illustrate the statements made in the paragraphs (for example Fig. 2, page 10; Fig. 3, page 12, Fig. 4, page 13 Fig. 6 and 7, pages 17 and 20).

11)

- I am very surprised to see such an enormous difference between CW 7.1 and XT in terms of correlation (Figure 1). Is the difference the same tendency when raw data for parameters are analysed directly? Are the performances of Males of Females mice analysed without z transformation for the XT or 7.1 versions comparable to the data after normalisation?

Yes, when comparing the original data of males and females by the two CW versions we do see the same pattern within the parameters as in the Radar Chart

12)

- The heat map presented in Figure 3 doesn't have a proper figure legend and it is very difficult to understand all the clusters and even to see differences between genotype and groups of mice.

We rearranged the heat map and included a proper legend and explanation (now Fig. 5).

13)

- The authors mentioned in the introduction as a second objective, if a combination of tests measuring locomotor

behaviour is useful for getting a comprehensive phenotyping, or if there are redundancies between tests. This is a very good point to demonstrate indeed. But if I understood right, especially on the table 2 (p13), none of the parameters related to OF or VP could have been clustered with CW parameters. It seems to indicate that each test separately are measuring specific parameters for specific phenotypes which I found interesting and in accordance about what behavioural task has been designed which is to measure specific component of the behaviour. However, how this result can it be interpreted? I am not sure that this point which is for me important is discussed in the discussion part.

We agree with the reviewer. As the single components are always made up by parameters from a single test, we propose that the different tests are not replaceable with each other. As stated above, this confirms the original idea that each test measures distinct aspects of behavior. Nonetheless it is not self-evident that this is actually the case. It would have been possible that activity measures, such as the run duration of the CW and the speed or distance travelled in the OF would cluster together as they measure the same dimension. Also with the IMPC data set we see this separation of tests in the components. Again it would have been entirely possible that the activity measures of SHIRPA, OF and HC would cluster together in one component. Apparently this is not the case, suggesting that other (environmental and/or procedural) factors come into play and alter activity levels according to the test situation. In the discussion we elaborate on this point (page 23, line 23ff).

14)

- Regarding the table 2, why the component presented for the CW data are not similar to the one presented in the table 1 especially when the same number of parameters has been included (113)? Why, the original "spatial dimension" and "temporal dimension" presented in table 1 are respectively named "print dimension" and "step cycle" in table 2? Moreover, why "Stride length" which was originally in Table 1 included in the spatial dimension is in step-cycle dimension in table 2? Does it mean that for different set of data analysed the components can appear to be different? Could this point be critical in terms of reproducibility for an analytic model?

Depending on the input (parameters, tests), the outcome of components can vary. With the first analysis we included 112 parameters only of the CW; in the second analysis we included 147 parameters from the CW (#113), OF (#31) and VP (#3). In both analyses we have the cut off of the number of components used at 80% explained variance. Due to this there are smaller differences of what parameter has its highest loading in what component- and therefor the characterization/naming of the component. Nevertheless, the general dimensions are preserved- when comparing only the CW components- i.e. spatial and temporal dimensions as well as coordination and its variation. The component “Stride Length” in the first analysis is correlated with C 02 “Paw Contact”, a component of the temporal dimension. Therefore it is not too surprising that in the second analysis the” Stride Length” parameters are associated with C 5 “Stand and Stride Length”- again a component of the temporal dimension. Some CW parameters are not independent of each other- i.e. the stride length is influenced by run duration which can also influence stand and swing phase. As such the spatial dimension of the stride length heavily depends on the temporal dimension. We do not think that this is a critical point, as all dimensions of the CW are represented in both analyses. Depending on the input of data (number of parameters and tests) and the processing of this data, the output components can change. Nonetheless we do think that the reproducibility (in terms of taking the same data and re-running the analysis) as well as the replicability (taking a new data set -same tests and parameters assumed) of our data is given.
To strengthen the similarities between the two CW analysis we changed the naming of the dimensions in table 2 (Step Cycle→ Temporal dimension).

DISCUSSION

15)

- Furthermore, in the discussion part, it is status: "We went on to see how the parameters of several different tests would group into principal components. We did this with a subset of the animals used above for the CW

analysis…" Does it mean that the same animal has been assessed twice with the model? And if it is the case, why was it not possible to get the same correlations and then obtain different component between table 1 and 2?

Yes, the second analysis on CW, OF and VP data was done by including a subset of animals from the first analysis (CW only), because not all animals of the first analysis were also tested on OF and VP.

For the question about the difference between table 1 and 2 please refer to the answer above.

16)

- In the first paragraph of the discussion, it is mentioned that the mice age range for the analysis of the 7.1 version was between 3 and 28 months and for the XT version between 3 and 15 month. Why not having considered in the analysis the same age range for the comparison? This point makes the analysis of the difference between Catwalk 7.1 and XT quite confusing then.

As a starting point we wanted as many animals as possible for the analysis, also because we did not expect a difference between the two versions. As we did see differences between the versions in the analysis, we thereupon analyzed only animals in the same age range (and similar genetic backgrounds) and could still pick up the differences between the versions (see page 22, line 5ff). Therefore we decided to stay with the first approach- by including all animals. This also has the benefit, that analysis on age, body weight and background strain can be done with decent animal numbers.

In conclusion, I can see and understand that the manuscript submitted by Zimprich and collaborator represent an astronomical quantity of analyses and work but is not for now convincing enough to attribute a positive point of view on it yet. The idea for me is really interesting and might be followed but the quality of the manuscript needs to be improved. I am convinced that presented differently the model proposed by Zimprich and collaborator can be really nice and useful.

Additional comments:

17)

- Figure legends are missing it is mainly title instead of proper legends.

We included proper figure legends

18)

- "OF and vertical pole testing was already described mouse elsewhere [10]", the sentence is incorrect or some word are missing. (Page 4, Line 10)

We corrected the sentence.

Reviewer #2: To the Authors

19)

Automated systems for quantifying behaviour may produce such a large number of primary and secondary diagnostic parameters that they can become difficult to evaluate. The authors show that principle component analysis is a useful method to reduce parameters to a much smaller set (10 in this case). Despite the reduction the derived parameters can still be allocated to clearly separable general functions. These were the spatial, temporal,

coordination and variation aspects of locomotion. While this is a useful approach it is not directly new. Other studies have used it before and the current study should reference and discuss this more on a general level.

We included more details/references. As we do not want to discuss PCA on a general level, but in combination with CW, we only focused on papers, where a PCA was described in reproducible detail and run on a behavioral data set. (page 3, line 22ff)

20)

The authors found significant differences between two versions of the Cat Walk apparatus. This information is of little interest to the general reader and I recommend reducing it to a few side remarks.

As the two CW versions have a major impact on our analysis of the data, we do need to explain the major differences. Also still CW 7.1 versions are being used and data published and for reasons of comparability across studies the information we give might be useful. Also refer to Chen et al (DOI: 10.1186/1743-0003-11-62), who found differences between the versions and could only detect subtle changes with the XT version but not the CW 7 in a sciatic nerve injury model. Besides, the potential impact of equipment versions, i.e. equipment specifications, on results, particularly on absolute values, unfortunately still keeps being overlooked in the biomedical field, although it might explain parts of the currently debated "reproducibility crisis". For these reasons we would like to keep this part as it is.

21)

While the authors point out that the derived principal components capture essential general parameters it should be better and more critically discussed to what extent the primary measurements are contained in the derived parameters.

We are not sure if we understand the question correctly. For characterizing the components we used the parameters, which have their highest loading (above |0.5|) in this component. The loadings of the individual parameters can be checked by consulting the structure matrix (Table S3, Supporting Information).

22)

I was surprised to see that two different measures of home cage activity (from the home cage and from the open field) apparently were little correlated so that they resulted in different principal components. On one hand this should be better discussed on the other hand it somewhat questioned for me the wider appliccability of this approach.

Yes, we were also surprised to see that the activity measures of the HC and OF (and SHIRPA) were not included in the same component. We now elaborate on this point in the discussion (see page 23, line 37). Please note that the 20 min open field test does not measure home cage activity, but spontaneous locomotion and exploratory activity in a novel environment.

We do not think that this result questions the applicability of the current approach, but it shows that the different test paradigms have a great influence on the parameters measured- even if the parameters are claimed to measure the same quality (in this case activity). Our results illustrate that the circumstances under which activity measures are recorded determine what they reflect. For a further discussion please refer to question 13.

23)

The major result of the study seems to be that results from mouse strains differ. What I find lacking is to what extent the analysis helped in distinguishing for the specific studies performed the experimental animals from the control animals. This is the major question that a researcher wants to answer when performing these measurements.

With our analysis we suggest that for distinguishing between mutant and wild type the measures of FP and HP (for the temporal and spatial dimensions) would be enough and either the couplings or phase dispersion would suffice for a first analysis, thereby reducing the amount of parameters from 113 to 44 (page 21, line 10ff; page 8, line 29ff). The CW can clearly detect differences between control and experimental animals. In the 25 mutant mouse lines used in our analysis 21 of them show a genotype-specific phenotype. This high phenotype- rate in mutant mouse lines subjected to the CW is due to a hypothesis-driven selection of mouse lines (page 23, line 18ff).

To get a comprehensive picture of the "locomotor phenotype" of the mouse line analyzed it is necessary to include a few tests. For instance the CW could be included to analyze the gait, OF, SHRIPA and HC for the activity measures under different circumstances as well as RR, VP and GS for motor performance and muscle strength. All these tests add to the different aspects of locomotor phenotypes (page 23, line 23ff).

In respect to large-scale phenotyping, where the issue of refining and reducing of animals and tests comes into play, we can clearly state that, at least for "locomotor phenotypes", all of the above mentioned tests have their validity and that there are no redundancies between them.

Specific comments

24)

OF open field -> spell out on first use: OF, HC, RH?, others?

We corrected this.

25)

The manuscript would significantly benefit from having a professional English language editor go through it. Since this is available for around 200-400 EUR the authors should not shy that cost.

We now have had a native English speaker critically reading and correcting the manuscript.

26)

This study proposes the use of Principal Component Analysis as a „new method" to reduce dimonsionality and complexity in data sets gathered for behavioural phenotyping by automated instruments. However, PCA is a standard method that has been used in other studies. Thus, I think a general comparison on the use of PCA for behavioural phenotyping by comparing the general merit of this approach with other such studies is warranted. E.g. Vannoni E, Voikar V, Colacicco G, Sánchez MA, Lipp HP, Wolfer DP. 2014. Spontaneous behavior in the social homecage discriminates strains, lesions and mutations in mice. J Neurosci Methods. 30;234:26-37. doi: 10.1016/j.jneumeth.2014.04.026. OR Ohl F, Roedel A, Binder E, Holsboer F. 2003. Impact of high and low anxiety on cognitive performance in a modified hole board test in C57BL/6 and DBA/2 mice. Eur J Neurosci. 17:128-136.

We did not want to indicate that PCA is a new method, but want to show that it is a sound method to reduce dimensions on highly dimensional data sets, such as the CW; thus we edited this part of the Introduction (page3, line 22 ff) and Discussion (page 20, line 23ff). We included the paper from Ohl et al to the mentioned literature (Vannoni et al are already named). We are not aware of any publication refining CW parameters by PCA.

27)

The fact that the two Cat Walk versions give different results is of interest mainly to the authors who know and have used both and maybe to the manufacturer Noldus but is not really of general interest in the context of the method „PCA" to a wider audience. Thus, the sections devoted to the cat walk comparison should be reduced to a

few side remarks. I think it would be sufficient to point out that there is a difference but then concentrate on just one version, e.g. the recent version of the cat walk. Maybe much of the comparative data can be put into a supplement outside the main article.

We do think that the power of this analysis comes from the high amount of animal numbers differing in genotype, sex, body weight and age. Only with this high amount of animals it is possible to analyze age or body weight effects in such detail. Please also see our reply to question 20.

28)

„Running a PCA on data sets to extract individual component scores is a useful tool in interpreting large data sets. „ -> rather „highly dimensional / multidimensional" data sets or similar

We changed this expression.

29)

„it is necessary to include other locomotor tests for a comprehensive phenotyping." -> surely, you do not mean „general behavioural phenotyping" but „locomotor phenotyping". Also: since the only other measure that you really suggest with evidence is „activity" this could maybe mentioned explicitly instead of just „other locomotor tests".

Yes, that is correct, we do mean locomotor phenotyping. We changed the sentence to: “Furthermore, although the CatWalk is sensitive for detecting locomotor phenotypes pertaining to gait, it is necessary to include other tests for comprehensive locomotor phenotyping.” (page 2, line 22f). As we use the term “locomotor phenotype” in a very broad way (Reviewer 1 suggested to include a definition- we did so; page 3, line 29ff), which includes gait, activity as well as motor performance, we would prefer keeping it the way it is and not change it to “activity”.

Journal of Neuroscience Methods

Special Issue „SI: Measuring Behavior 2016“

Analysis of Locomotor Behavior in the German Mouse Clinic

Annemarie Zimprich[1,3,#], Manuela A. Östereicher[2], Lore Becker[2], Petra Dirscherl[3], Luise Ernst[1,3], Helmut Fuchs[2], Valerie Gailus-Durner[2], Lillian Garrett[2,3], Florian Giesert[1,3], Lisa Glasl[3], Angelika Hummel[3], Jan Rozman[2,4], Martin Hrabě de Angelis[2,4,8], Daniela Vogt-Weisenhorn[3], Wolfgang Wurst[3,5,6,7], Sabine M. Hölter[2,3]

[1] Developmental Genetics, Technische Universität München-Weihenstephan, c/o Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

[2] German Mouse Clinic, Institute of Experimental Genetics, Helmholtz Zentrum München, Ingolstaedter Landstr.1, 85764 Neuherberg, Germany

[3] Institute of Developmental Genetics, Helmholtz Zentrum München, Ingolstaedter Landstr.1, 85764 Neuherberg, Germany

[4] German Center for Diabetes Research (DZD), Helmholtz Zentrum München, Ingolstaedter Landstr.1, 85764 Neuherberg, Germany

[5] Chair of Developmental Genetics, Technische Universität München-Weihenstephan, c/o Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

[6] German Center for Neurodegenerative Diseases (DZNE), Site Munich, Feodor-Lynen-Str. 17, 81377 Munich, Germany

[7] Munich Cluster for Systems Neurology (SyNergy), Adolf-Butenandt-Institut, Ludwig-Maximilians-Universität München, Schillerstr. 44, 80336 Munich, Germany

[8] Chair of Experimental Genetics, Technische Universität München-Weihenstephan, c/o Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

[#] Corresponding author:

Dr. Annemarie Zimprich

Helmholtz Zentrum München
Institute of Developmental Genetics
Ingolstaedter Landstr.1
85764 Neuherberg
Germany
Tel.: 089/3187-2371
Fax: 089/3187-3099
Email: annemarie.zimprich@helmholtz-muenchen.de

Highlights

- Assessment of comprehensive locomotor phenotyping strategy
- Refinement of highly dimensional data sets by Principal Component Analysis
- Influence of equipment version, sex, body weight, age and genetic background
- No redundancies detected between different locomotor tests

Journal of Neuroscience Methods

Special Issue „SI: Measuring Behavior 2016“

Analysis of Locomotor Behavior in the German Mouse Clinic

Annemarie Zimprich[1,3,#], Manuela A. Östereicher[2], Lore Becker[2], Petra Dirscherl[3], Luise Ernst[1,3], Helmut Fuchs[2], Valerie Gailus-Durner[2], Lillian Garrett[2,3], Florian Giesert[1,3], Lisa Glasl[3], Angelika Hummel[3], Jan Rozman[2,4], Martin Hrabě de Angelis[2,4,8], Daniela Vogt-Weisenhorn[3], Wolfgang Wurst[3,5,6,7], Sabine M. Hölter[2,3]

[1] Developmental Genetics, Technische Universität München-Weihenstephan, c/o Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

[2] German Mouse Clinic, Institute of Experimental Genetics, Helmholtz Zentrum München, Ingolstaedter Landstr.1, 85764 Neuherberg, Germany

[3] Institute of Developmental Genetics, Helmholtz Zentrum München, Ingolstaedter Landstr.1, 85764 Neuherberg, Germany

[4] German Center for Diabetes Research (DZD), Helmholtz Zentrum München, Ingolstaedter Landstr.1, 85764 Neuherberg, Germany

[5] Chair of Developmental Genetics, Technische Universität München-Weihenstephan, c/o Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

[6] German Center for Neurodegenerative Diseases (DZNE), Site Munich, Feodor-Lynen-Str. 17, 81377 Munich, Germany

[7] Munich Cluster for Systems Neurology (SyNergy), Adolf-Butenandt-Institut, Ludwig-Maximilians-Universität München, Schillerstr. 44, 80336 Munich, Germany

[8] Chair of Experimental Genetics, Technische Universität München-Weihenstephan, c/o Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

[#] Corresponding author:

Dr. Annemarie Zimprich

Helmholtz Zentrum München
Institute of Developmental Genetics
Ingolstaedter Landstr.1
85764 Neuherberg
Germany
Tel.: 089/3187-2371
Fax: 089/3187-3099
Email: annemarie.zimprich@helmholtz-muenchen.de

Abstract

Background

Generation and phenotyping of mutant mouse models continues to increase along with the search for the most efficient phenotyping tests. Here we asked if a combination of different locomotor tests is necessary for comprehensive locomotor phenotyping, or if a large data set from an automated gait analysis with the CatWalk system would suffice.

New Method

First we endeavored to meaningfully reduce the large CatWalk data set by Principal Component Analysis (PCA) to decide on the most relevant parameters. We analyzed the influence of sex, body weight, genetic background and age. Then a combination of different locomotor tests was analyzed to investigate the possibility of redundancy between tests.

Result

The extracted 10 components describe 80% of the total variance in the CatWalk, characterizing different aspects of gait. With these, effects of CatWalk version, sex, body weight, age and genetic background were detected. In addition, the PCA on a combination of locomotor tests suggests that these are independent without significant redundancy in their locomotor measures.

Comparison with existing methods

The PCA has permitted the refinement of the highly dimensional CatWalk (and other tests) data set for the extraction of individual component scores and subsequent analysis.

Conclusion

The outcome of the PCA suggests the possibility to focus on measures of the front and hind paws, and one measure of coordination in future experiments to detect phenotypic differences. Furthermore, although the CatWalk is sensitive for detecting locomotor phenotypes pertaining to gait, it is necessary to include other tests for comprehensive locomotor phenotyping.

Key words: CatWalk; Gait; Principal Component Analysis; Mouse; Locomotion; Activity; Phenotyping

Introduction

Recent advances in genome editing technology have revolutionized and accelerated the generation of mutant mouse lines for modeling human disease. This technological cataclysm has brought with it an upsurge in demand for comprehensive behavioral phenotyping that is reflected in the rising number of large-scale phenotyping centers and consortia [1,2]. To meet the demand, new sophisticated phenotyping devices have been developed to provide an automatic and detailed characterization of mice. The goal of automation is not only to speed up data acquisition and to make it more objective and replicable, but also to reduce the burden on the animal through less handling. Ideally, a few automated

systems would be sufficient to comprehensively phenotype a mouse model. The questions repeatedly arising in this context are how valid are individual tests; which test is the most informative and how many tests are really needed. It would also be beneficial to know if parameters of one test could predict the utility of another test for more detailed phenotypic characterization. Such knowledge would enhance the efficacy of phenotyping strategies, which have to be outlined in license applications for experiments on animals, thus saving time, animals and money.

Here we explore these questions with respect to locomotor phenotyping. Would it be sufficient, for example, to use the Catwalk (CW) from Noldus; an automated gait analysis system with video based paw tracking [3,4]? Previously the assessment of gait parameters involved footprint analysis, where the paws were dipped into ink and the animal walked across a paper. Approximately 20 parameters were measured from these prints, however with the automated CW system, around 230 parameters are measured. With this dramatic increase in parameters/data points data analysis has become more complex. Traditionally only single parameters or a few selected ones were analyzed. But it is hard to formulate a clear picture of the phenotype when dealing with a large amount of parameters for one test. Thus it is necessary to decide which of the parameters give most information and is the best for detecting disease-relevant phenotypes. Alternatively, a compression of the data may be more appropriate to garner a comprehensive overview. As such reducing the number of parameters under consideration (aka dimensions) to a few components (i.e. artificial variables) by Principal Component Analysis (PCA) improves data handling without significant information loss [5-9]. The subsequent analysis then runs on the extracted components that substitute for the more numerous original parameters.

Running a PCA on behavioral data is not new; several groups have used it for analyzing data [5-9]. Yet running a PCA on a multidimensional data set (a high amount of animals and parameters, i.e. over 1000 cases and over 100 parameters, respectively) is not frequently done. With a few exceptions animal numbers are moderate (below 100 animals) and parameters analyzed small (below 50). To answer the above mentioned questions we apply a PCA to a multidimensional data set as collected by the CW, which, to our knowledge, has not been done so far. By including a fairly large sample size (over 1000 cases) we expect to produce reliable results, as factor instability is less likely to occur in a larger sample.

In this paper we focus on locomotor phenotypes, whereby the definition of a “locomotor phenotype” is used very broadly. We include gait phenotypes (as analyzed by the CW), as well as activity (as in the Open Field (OF), home cage (HC) and SHIRPA) and motor ability phenotypes (as in the Grip Strength (GS), Rotarod (RR) and Vertical Pole (VP) test) into this term. Gait phenotypes describe *how* the animal moves, activity phenotypes describe *how frequently* the animal moves, and motor ability phenotypes describe in this case muscle strength, balance and coordination. To test for a disease-relevant locomotor phenotype generally one to several different tests are applied. Here we assess if the CW alone, with its large amount of parameters, can be sufficient to detect locomotor deficits and if a reduction in dimensions is appropriate to detect meaningful components. With historical data from our lab we ran a PCA to reduce the dimensions of the CW data. Furthermore, we characterized the resulting components and analyzed them with different statistical methods to evaluate the impact of different versions of the CW system (we have used two different systems in our lab), sex, body weight, age and genetics.

Likewise we ask if a combination of tests measuring locomotor behavior is necessary for comprehensive locomotor phenotyping, or if there are redundancies between tests. To answer these questions we

included two data sets. In the first set we took historical data from our lab including parameters from the CW, OF and VP test. The second data set comprises publicly available data from the International Mouse Phenotyping Consortium (IMPC; [10]) website (https://www.mousephenotype.org) selecting different measures of motor performance (horizontal as well as vertical activity, muscle strength and coordination), which included parameters measured in the OF, SHIRPA, RR, GS and in the HC during indirect calorimetry at the German Mouse Clinic. With both data sets a PCA was run to see how the different parameters are distributed among the extracted components, thereby assessing possible redundancies between different tests.

Methods:

Behavior:

For all tests mice are transferred to the testing room at a minimum of 30 minutes prior to testing. Mutants are always concurrently tested with respective wild types and males prior to females. After each mouse the test apparatus is cleaned and disinfected. Different mutant lines were used in the behavioral assessments. For sample sizes by sex and genotype please refer to Table S1 in the Supporting Information. All experiments were approved by the government of Upper Bavaria, Germany.

CatWalk:

Mice were tested on two versions of the CW system (Noldus Information Technology, Wageningen, Netherlands): the CW 7.1 version and the CW XT 10.5 version (hereafter referred to as CW XT). On both systems the mouse traverses an elevated glass walkway that is bordered by Plexiglas walls in a darkened room. A camera situated underneath the middle of the walkway tracks the illuminated footprints, which are then analyzed with the CW software. For each animal the mean of 2 to 4 uninterrupted and continuous runs (each included approx. 4-6 step cycles) were calculated and used for the following analysis. The major difference between the two versions is the camera: for the older 7.1 version the camera captures the footprints at a rate of up to 50Hz and for the XT version a high speed camera is used, capturing at a rate of up to 100Hz. The CW system measures various aspects of paw contacts with the floor in a dynamic way and automatically calculates a broad number of spatial and temporal gait parameters in several categories. These include i) parameters related to individual paw prints, such as the width and length, together with a calculation for the front and hind paws (FP and HP, respectively); ii) parameters related to the position of paw prints with respect to each other, for example the stride length and the base of support (BOS), which measures the width between the two FPs and the two HPs respectively; iii) parameters related to time-based relationships between paw pairs (couplings and phase dispersion) and their variation, as well as step patterns. For a more detailed description of the CatWalk method see [3]. Table S2 lists the selected parameters for the statistical analysis.

Open Field and Vertical Pole:

OF and VP testing was already described elsewhere [11]. In brief, mice are placed in a square OF arena surrounded by infra-red beams (ActiMot, TSE, Bad Homburg, Germany) and behavior is recorded for 20 minutes. Several parameters are analyzed for the whole 20 minutes and four of them (distance travelled, number of rearings, time in center [%] and distance in center [%]) are also analyzed in 5 minute bins (Table S2). The VP is a 50cm high, taped pole (diameter 1cm) where the mouse is placed

head upwards at the top. Time to turn and for complete descending is recorded. The time from turning till complete descent is calculated ("TimeDown", Table S2). Animals receive 2-3 training trials and 3-5 test trials with 5-10 minutes inter-trial intervals.

IMPC phenotyping:

Mice from the IMPC consortium undergo a standardized phenotypic screening that involves several different biological and clinical aspects- including behavior and neurological endpoints. In this study we included parameters from the OF, SHIRPA, RR, GS as well as the indirect calorimetry (see Table S7) from the German Mouse Clinic only. For the generation of behavioral data for IMPC lines please see https://www.mousephenotype.org/impress/procedures/14 on the IMPC webpage. In brief, OF testing is performed as mentioned above.

Rotarod:

Motor Performance is tested with a commercially available Rotarod apparatus (Bioseb, Chaville, France). The test phase consists of three trials separated by 15 minute inter-trial intervals. The mice are placed on the rotating rod and the apparatus accelerates from 4 to 40rpm in 300 seconds. The latency to fall is recorded as well as the number of passive rotations.

SHIRPA:

The primary observation screen is a modification of the Irwin procedure [12]. The mouse is placed in a clear cylinder to observe tremors and body positioning. Then the mouse is transferred into an arena (420 x 260 x 180mm) in which a Perspex sheet on the floor is marked with 15 squares. Here several other parameters are taken; one being locomotor activity which is recorded in the first 30 seconds by counting the number of squares being crossed.

Grip Strength:

A grip strength meter (Bioseb, Chaville, France) is used. The mouse is lowered towards the grid, so that the FP cling onto it. By slowly pulling back the mouse the maximum strength is recorded. Then the mouse is lowered so that both FP and HP attach; again the maximum grip strength is recorded. For both measurements the mean of three measurements are calculated. Body weight is recorded. For the PCA the grip/body weight ratio is used.

Home Cage activity:

While the animals are tested for indirect calorimetry (PhenoMaster System, TSE Systems, Germany), activity data (horizontal and vertical) and speed are also recorded. Here animals are singly housed for at least 21h.

Data sets:

Animals were only included in the analysis, when all parameters of all tests were present.

CW data:

Data included 1499 cases and 112 parameters. We included data from two CW versions (i.e. the 7.1 and the XT version) integrating only parameters which both CW versions collect (see Table S2).

CW-OF-VP data:

For the investigation of redundancies between tests we took a data set consisting of animals of the "CW data", which were also tested in OF and VP. In total 1057 cases and 147 parameters from the CW, OF and the VP test were collected (Table S1 and S2).

IMPC data:

We also wanted to investigate an independent data set for redundancies between tests assessing locomotor phenotypes: For this we took a data set, which included 1327 animals and 20 parameters from the OF, SHIRPA, RR, GS and HC activity during indirect calorimetry (Table S1 and S7). Data was taken from our internal database but it is also accessible for the public via http://www.mousephenotype.org/. As we could not include all animals measured (due to excessive animal numbers) we decided to pick mouse lines upon extremes in activity based on 5 parameters (OF-distance, OF-rearings, SHIRPA-locomotion, HC-distance and HC-rearings). The mouse lines were selected based on the effect size (Cohen´s D ($d = (mean_{(mutants)} - mean_{(controls)}) / \sqrt{((SD^2_{(mutants)} + SD^2_{(controls)})/2)}$) – all lines below the 5^{th} and above the 95^{th} percentile were chosen.

Statistics:

All statistical analysis was performed with PASW Statistics 18 (Version 18.0.0; SPSS Inc., Chicago, USA), if not mentioned otherwise, and a p-value ≤0.05 was considered as statistically significant.

Principal Component Analysis:

We conducted three PCAs: a) with CW data, b) with CW, OF and VP and c) with IMPC data.

Generally the handling of data for PCA was as follows:

All parameters of the individual data sets were standardized to a mean of 0 and a standard deviation of 1 via Z-transformation, to be able to compare between parameters. Generally, if the scales between parameters differ greatly and no transformation is being done, then parameters with a large scale can dominate the outcome. Thus, in an exploratory analysis it is recommended to standardize the data set. The sampling adequacy was confirmed by the Kaiser-Meyer-Olkin (KMO) measure and the Bartlett- Test. As we do not know if our components are dependent or independent of each other, we chose the oblique rotation (via the oblimin method) thereby allowing correlation between components. For the extraction of components two criterions were set; firstly, only those with eigenvalues greater than 1 were chosen and secondly, the scree plot was consulted or a cut off at 80% of explained variance was used. As the scree plot becomes difficult to interpret with a high amount of parameters (i.e. for the first two analyses; CW data: 112 parameters and CW-OF-VP data: 147 parameters), we used for the first two data sets the cut off at 80% of explained variance. For the third data set we could confirm the extracted components by consulting the scree plot, as here fewer parameters were measured (20 parameters; for the scree plot see Fig. S1). Individual component scores were calculated. Components were characterized according to the highest loadings (>|0.5|) of individual parameters.

CW data:

With this data set we conducted further analysis to investigate effects of sex, CW version, body weight, age and genotype. For analyzing sex and CW version effects a two-way ANOVA was applied. For the correlation with body weight the Pearson correlation coefficient was used, including only animals where body weight was measured within a week of CW testing. Age effects were evaluated by the Pearson correlation applied to all animals experiencing the CW for the first time and including only wild type animals. A linear regression model was used for analyzing age, sex and body weight influences on components. The medians of the individual component scores of each group were clustered in a hierarchical agglomerative cluster analysis using R (Version 3.0.2, R Foundation for Statistical Computing, Vienna, Austria) to investigate similarities between the lines, the corresponding genotypes and sex (e.g. Group 1: male wild types of mouse line A; Group 2: male mutants of mouse line A). The Ward method was used as a linkage criterion and the Euclidean distance as metric. This method allows the classification of groups on the basis of trait similarities.

Results

PCA on historical CW data

Characterization of the ten principal components revealed the possibility to focus on a reduced parameter-set in future analysis

We firstly ran a PCA only on the CW data set to reduce dimensions and to analyze the impact of sex, body weight, age and genotype. Historical data of 1499 cases was used for the PCA analyzing CW data. As there was a change of the CW version (from 7.1 to XT) in our institute we took data from both, analyzing only parameters that were measured by the two. After the omission of one parameter due to low correlations in the correlation matrix ("StepSequence_Rb"), we analyzed 112 parameters. The sampling adequacy of the data was done via the KMO criterion and Bartlett's test and found to be appropriate (KMO= 0.810; Bartlett- Test: Chi^2= 522783.471, df= 6216, p<0.001). Ten components were extracted via oblique rotation describing 79.5% of the total variance. In other words, we were looking for the components that accounted for the most variation and hence the most likely to reveal differences between the cases and these were the ten revealed. The loadings, which can be described as the correlation coefficient between the parameter and the components, are depicted in the structure matrix table (see Table S3). Components were characterized according to the highest loadings (above |0.5|) of the individual parameters (see Table 1).

Table 1: CW PCA: Component characterization according to the major parameters. The components are arranged due to their representation of different gait dimensions. The explained variance is given for each component and for the dimensions.

		Component name	**Major parameters** (with loadings >|0.5|)	**Explained variance**	
Spatial dimension	C 01	Print Dimensions	Print Length, Print Width, Print Area, Maximum Contact Area	23.2%	26%
	C 08	Stride Length	Stride Length	2.7%	
Temporal dimension	C 02	Paw Contact	Stand Duration, Duty Cycle, Run Duration, Support Diagonal, Support Single	14.7%	27%
	C 04	Swing Phase	Swing Duration, Swing Speed	8.7%	
	C 07	Turning point	Max Contact At	3.4%	

Interlimb coordination	C 05	Coordination of diagonal and ipsilateral pairs	Coupling and Phase Dispersion of the diagonal and ipsilateral pairs, Alternate step patterns	6.1%	11%
	C 06	Coordination of girdle pairs	Couplings and Phase Dispersion of the girdle pairs, Cruciate step patterns	4.5%	
Variation in interlimb coordination	C 03	Variation incl RH	Variation in coordination of paw pairs including the RH, Regularity Index	12.2%	16%
	C 09	Variation excl RH for diagonal and ipsilateral pairs	Variation in coordination of diagonal and ipsilateral pairs not including the RH	2.1%	
	C 10	Variation for FP girdle pairs and Print Position	Variation in coordination of FP girdle pairs, Print Position	1.9%	

RH- right hind paw; FP- front paws

The extracted components can be attributed to different dimensions of gait performance, such as spatial and temporal aspects as well as interlimb coordination and its variation.

Component C 01 *Print Dimensions* (describing 23% of the total variance) contains nearly all spatial print dimensions, which includes the print length and width as well as the area. Component C 08 *Stride Length* (describing 3% of the total variance) accounts for the rest of the spatial measures. Together these components explain 26% of the total variance.

The temporal dimensions are represented by the components C 02 *Paw Contact* (which includes the stand duration, duty cycle as well as the support and run duration, describing 15% of the total variance), C 04 *Swing Phase* (which includes the swing duration and swing speed, describing 8% of the total variance) and C 07 *Turning Point* (which contains the turning point measured by the "Maximum contact at %", describing 3% of the total variance), which explain together 27% of the total variance.

The interlimb coordination is represented by two components, C 05 *Coordination of diagonal and ipsilateral pairs* (including the alternate step patterns (see also Table S2), describing 6% of the total variance) and C 06 *Coordination of girdle pairs* (including the cruciate step patterns, describing 5% of the total variance) explaining together 11% of the total variance.

Another three components, C 03 *Variation including RH* (right hind paw) (contains also the regularity index, describing 12% of the total variance), C 09 *Variation excluding RH for diagonal and ipsilateral pairs* (describing 2% of the total variance), and C 10 *Variation for FP (front paw) girdle pairs and Print Position* (describing 2% of the total variance) represent the variation in the interlimb coordination explaining together 16% of the total variance.

The consequence of allowing for correlations between components by the oblique rotation can be seen in the component correlation matrix (see Table S4). We see slight correlations between some components. The strongest correlation occurs between two components describing the variation; C 03 *Variation including RH* and C 09 *Variation excluding RH for diagonal and ipsilateral pairs* (r= -0.490). There is also a correlation between C 01 *Print Dimensions* and C 10 *Variation FP girdle pairs and Print Position* (r=-0.306), and a weaker one between C 02 *Paw Contact* and C 08 *Stride Length* (r=-0.252).

For the spatial and temporal parameters we have measures from all four paws as well as the combinations for the FPs and HPs (hind paws) respectively (see Table S2 and Fig. 2). All six measures of the same parameter always coincide within the same component (see Table S3); suggesting that analysis of only FP and HP could suffice in future analysis (in contrast to analyzing all single paws). For two

parameters we have only measurements of the FP and HP, and left and right side, respectively; these are the base of support (BOS) (FP and HP) and the print position (left and right side). The print position for left and right side integrated in the same component (C 10). Interestingly, the BOS for FP and HP are within different components, i.e. the BOS FP has its major loading to C 05 *Coordination of the diagonal and ipsilateral pairs* (-0.469) but also loads onto C 02 *Paw Contact* (0.420) and C 08 *Stride Length* (-0.374) with a similar weight; the BOS HP has its highest loading on C 04 *Swing Phase* (0.339). Another noteworthy point is that the BOS of FP and HP do not correlate ($r_{(1498)}$=0.046) and as such have to be regarded as independent parameters, which is also illustrated by the different integration into a component and the differential pattern of loadings onto other components.

Highly significant differences between two CW versions

Over the course of time newer versions of a test apparatus might be installed in a lab and the question arises of how comparable these versions are. This is especially important when comparing mouse lines and experiments before and after such a change has occurred. As we switched from the CW 7.1 to the XT version, we employed the data set to analyze CW version effects and sex effects by a two-way ANOVA. The results showed interactions between sex and CW version for C 01 *Print Dimensions*, 02 *Paw Contact*, 05 *Coordination of diagonal and ipsilateral pairs* and 07 *Turning Point* ($F_{(1,1495)}$=9.109, p=0.003; $F_{(1,1495)}$=11.119, p=0.001; $F_{(1,1495)}$=25.248, p<0.001; $F_{(1,1495)}$=28.247, p<0.001, respectively). We saw highly significant differences between the CW versions in all components except C 03 *Variation incl RH*, which is not significant (C 01 $F_{(1,1495)}$=1579.028, p<0.001; C 02 $F_{(1,1495)}$=443.631, p<0.001; C 04 $F_{(1,1495)}$=259.273, p<0.001; C 05 $F_{(1,1495)}$=43.622, p<0.001; C 06 $F_{(1,1495)}$=11.564, p=0.001; C 07 $F_{(1,1495)}$=20.636, p<0.001; C 08 $F_{(1,1495)}$=186.445, p<0.001; C 09 $F_{(1,1495)}$=3.952, p=0.047; C 10 $F_{(1,1495)}$=128.866, p<0.001). Fig. 1 depicts these differences in a Radar Chart. Thus we continued by separately analyzing the data of the two versions. To further illustrate this difference between the two versions we plotted the original data of the C 01 *Print Dimensions* parameter “Print Area” and for C 08 *Stride Length* “Stride Length” (Fig. 2). We see a larger print area and a longer stride length in the CW XT version compared to the 7.1 version. These results show that a highly significant difference between the two CW versions exists and that caution has to be taken when comparing data of the two.

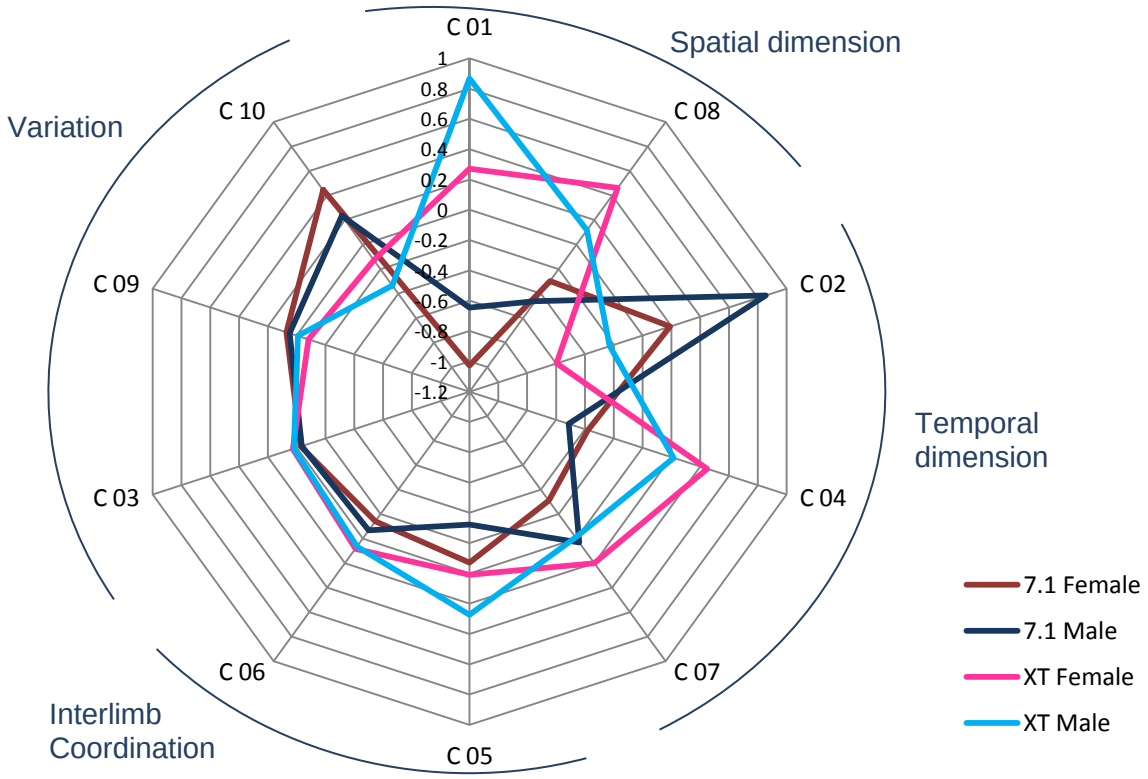


Fig. 1: Radar Chart depicting the differences between CW version and sex (shown are the means per group; $n_{(7.1,M)}$=344; $n_{(7.1,F)}$=283; $n_{(XT,M)}$=464; $n_{(XT,F)}$=408). Notice that components are grouped according to their representation of different gait dimensions.

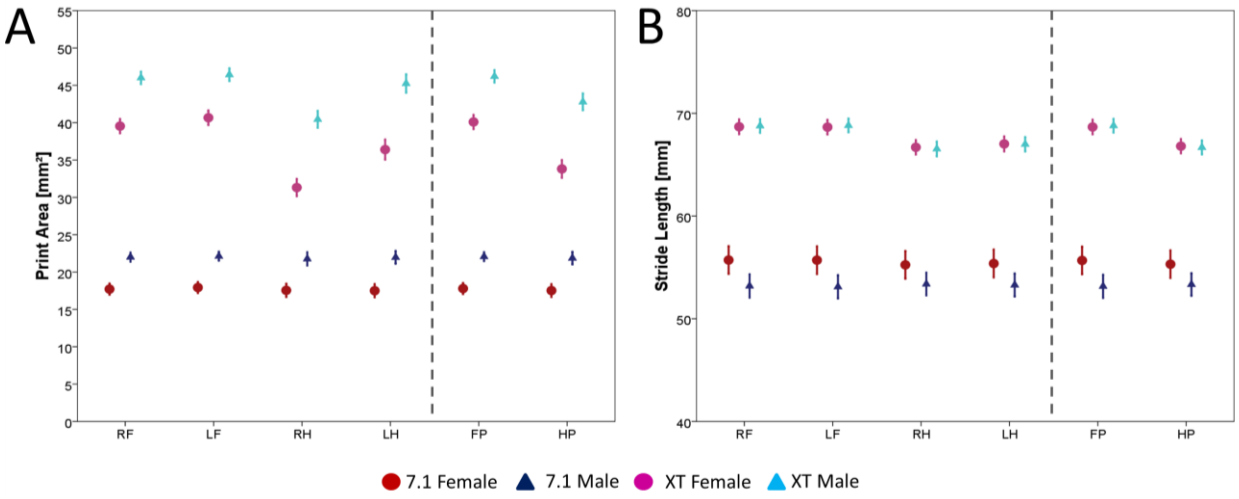


Fig. 2: Original data for "Print Area" (A) and "Stride Length" (B) of all paws, inlcuding front paws and hind paws, to illustrate the differences between the two CW versions. Both "Print Area" and "Stride Length" have higher values in the XT version. Shown are the means and the 95% confidence interval. RF- right front paw, LF- left front paw, RH- right hind paw, LH- left hind paw, FP- front paws, HP- hind paws

Sex and body weight effects on print dimensions in both CW versions

In biomedical research sex effects are often ignored or not thoroughly described, although it is clear that sex differences exist in several different behaviors and the prevalence, severity and etiology of (human) diseases [13-18]. Also body weight can have an effect on gait parameters. To investigate sex and body weight effects we conducted an ANOVA and calculated the Pearson correlation, respectively. As we saw effects of both factors in some components we evaluated their influence with a multiple linear regression model. The analyses occurred separately for both CW versions.

Sex differences appeared in the 7.1 version for C 01 *Print Dimensions*, C 02 *Paw Contact*, C 05 *Coordination of diagonal and ipsilateral pairs*, C 07 *Turning Point* and C 10 *Variation FP girdle pairs and Print Position* (C 01 $F_{(1,625)}$=70.479, p<0.001; C 02 $F_{(1,625)}$=54.920, p<0.001; C 05 $F_{(1,625)}$=8.010, p=0.005; C 07 $F_{(1,625)}$=13.407, p<0.001; C 10 $F_{(1,625)}$=4.864, p=0.028). For the XT version all components of the spatial and temporal dimensions, as well as in the coordination of the diagonal and ipsilateral pairs show a significant difference between the sexes (C 01 $F_{(1,870)}$=140.393, p<0.001; C 02 $F_{(1,870)}$=89.750, p<0.001; C 04 $F_{(1,870)}$=21.799, p<0.001; C 05 $F_{(1,870)}$=19.993, p<0.001; C 07 $F_{(1,870)}$=13.599, p<0.001; C 08 $F_{(1,870)}$=47.880, p<0.001; C 10 $F_{(1,870)}$=17.327, p<0.001).

We went on to analyze the data for correlations between the components scores and the body weight in the two CW versions separately. Only data from animals where body weight was measured within one week of CW testing was included. As the number of animals was still very high (7.1 version: n= 364; XT version: n=872) we got significant correlations for nearly all parameters. Thus we paid little attention to the p-values and assessed the correlation only via the correlation coefficient. The only correlations that are moderately strong in both CW versions are between body weight and C 01 *Print Dimensions* (7.1: $r_{(363)}$= 0.293; XT: $r_{(871)}$= 0.417) and C 02 *Paw Contact* (7.1: $r_{(363)}$= 0.399; XT: $r_{(871)}$= 0.503) respectively, suggesting that only those components are really influenced by the body weight. For the 7.1 version we also see a mild correlation with C 07 *Turning Point* ($r_{(363)}$= 0.312). All other correlation coefficients are below |0.253| and therefore do not suggest a noteworthy influence of body weight on the component.

As noted above there is an influence of sex on C 01 *Print Dimensions*, C 02 *Paw Contact* and C 07 *Turning Point*. To evaluate whether in fact the sex or rather the body weight as a confounder variable influences the component we ran a multiple linear regression analysis on a subset of animals ($n_{(F\ 7.1)}$=170; $n_{(M\ 7.1)}$=194; $n_{(F\ XT)}$=408; $n_{(M\ XT)}$=464), where animals were measured for their body weight within one week of CW testing and evaluated scatter plots. For the 7.1 version 11% of the variance in C 01 *Print Dimensions* can be explained by the linear regression model and here both sex and body weight have a significant influence (Fig. 3A; $F_{(2,361)}$=21.616, p<0.001, R^2=0.107, body weight: β=0.223, p<0.001, sex: β=-0.161, p=0.004). For C 02 *Paw Contact* and C 07 *Turning Point* only 17% and 10 % of the variance, respectively, can be explained by the model. For both components body weight has a significant influence, whereas sex has no impact (C 02: $F_{(2,361)}$=35.927, p<0.001, R^2=0.166, body weight: β=0.359, p<0.001, sex: β=-0.093, p=0.082; C 07: $F_{(2,361)}$=19.622, p<0.001, R^2=0.098, body weight: β=0.300, p<0.001, sex: β=-0.029, p=0.606). For the XT version we saw that C 01 *Print Dimensions* is influenced by both body weight and sex (Fig. 3B; $F_{(2,869)}$=111.077, p<0.001, R^2=0.204, body weight: β=0.304, p<0.001, sex: β=-0.207, p<0.001) explaining about 20% of the variation. In C 02 *Paw Contact* 25% of the variation can be explained by the model and only body weight has a significant influence ($F_{(2,869)}$=147.977, p<0.001, R^2=0.254, body weight: β=0.479, p<0.001, sex: β=-0.044, p=0.209). We correlated body weight to C 01 *Print Dimensions*

separated by sex and CW version and saw positive correlations for the females in both versions (7.1: $r_{(170)}$=0.331; XT: $r_{(408)}$=0.328 respecitvely) and in the males of the XT version ($r_{(464)}$=0.220), but no correlation for the males of the 7.1 version ($r_{(194)}$=0.104) (see Fig. 3A and B). A similar picture appears when analyzing the original data (in Fig. 3C and D "FP Print Area", a parameter of C 01 is depicted).

In summary we detect sex effects in several components, especially in the CW XT versions. In C 01 Print Dimensions we see an effect of sex as well as body weight. This is interesting to note, as it is intuitive that the size of paw prints would correlate with body weight, but less intuitive that there is a sex effect. For C 02 Paw Contact the analysis suggests an influence of body weight and not of sex. As body weight increases with age, we explored the influence of age on the components in a next step.

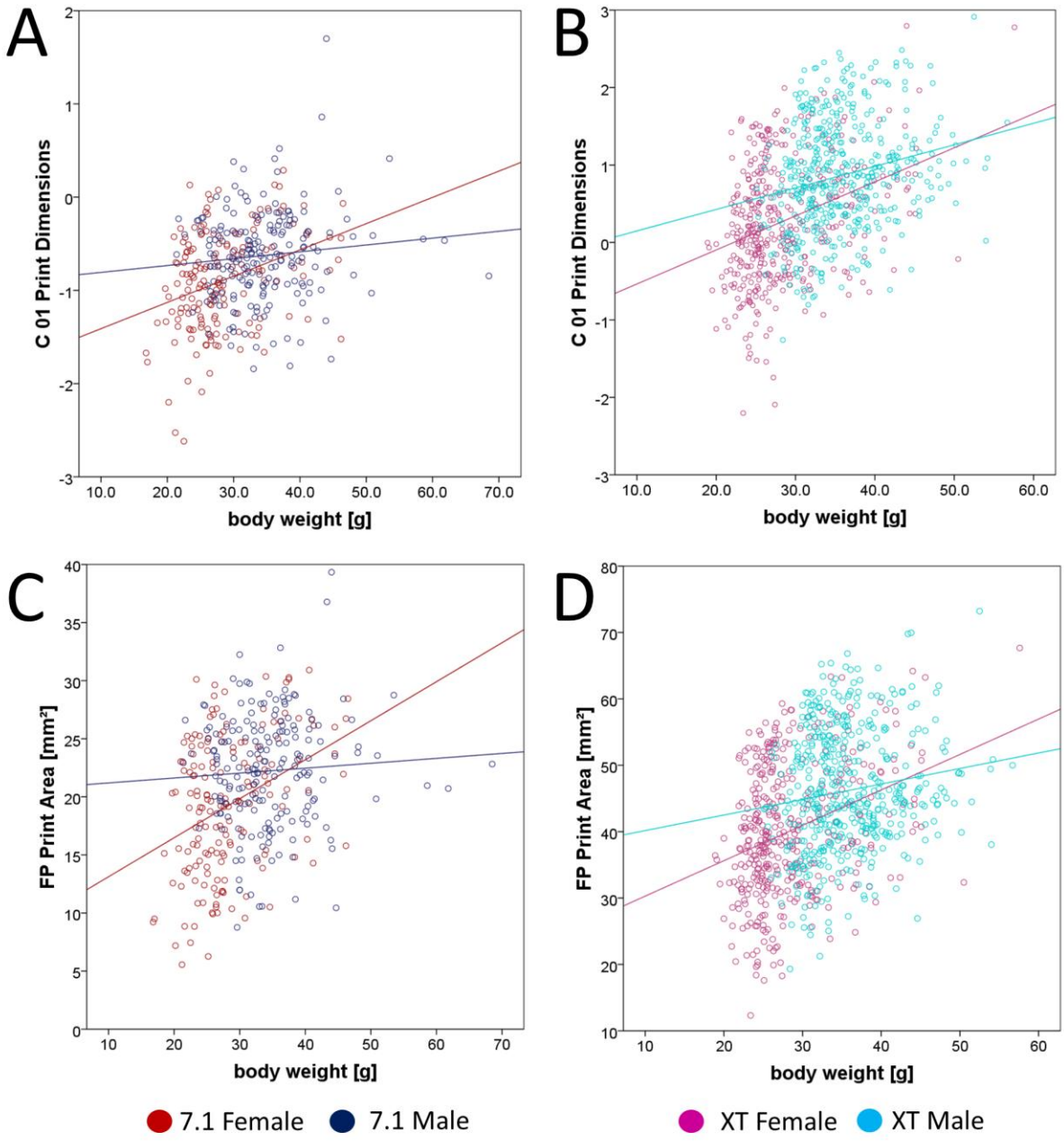


Fig. 3: Scatter plots for correlation between C 01 and Print Area with body weight for the 7.1 version and XT version. In the 7.1 version (A) there is an influence of sex and body weight on the component. When correlating body weight with C 01 per sex, a positive correlation is seen in the females but not in the males. In the XT version (B) we see a positive correlation both for males and females in C 01. In Print Area of the FP, a parameter of C 01, we can see the correlation with body weight also in the original data for both the CW 7.1 (C) and XT (D). FP- front paws

<u>*Paw contact increases with age in both CW versions*</u>

Age can play an important role when analyzing gait parameters. Not only does the body weight change with age (and thus some gait-related parameter change- see above), but also a decline in muscle mass and neuromuscular dysfunction, amongst other changes, can occur, thus leading to an altered gait performance. To investigate possible age effects we calculated the Pearson correlation coefficient and specifically checked the influence of age and body weight on C 02 *Paw Contact* in both versions.

We included in the analysis only wild type animals that were subjected to the CW for the first time to circumvent possible age-mutation interactions. Still we separately analyzed the CW versions and ran a correlation analysis for age and the different components. Again we interpreted only the correlation coefficient and paid less attention to the p-values, as they were nearly all significant due to the high animal number (7.1 n= 209; XT: n= 198). C 02 *Paw Contact* showed a positive correlation to the age of the animal in both versions (7.1: $r_{(208)}$= 0.367, Fig. 4A; XT: $r_{(197)}$= 0.500). The question again arises if the age effects we see in C 02 *Paw Contact* are due to age and not to body weight differences. For this we ran a multiple linear regression analysis for C 02 *Paw Contact* based on age and body weight. In the 7.1 version 16% of the variation can be explained by the model ($F_{(2,95)}$=10.461, p<0.001; adjusted R^2= 0.163), but here only age was found to be significant (β=0.348, p=0.001; body weight: β=0.150, p=0.133). This is illustrated by "HP Stand" (Fig. 4B); an increase in stand duration by age can be seen, but it is not a very dramatic one. For the XT version both predictor variables explain 34% of the variation ($F_{(2,195)}$=52.056, p<0.001; adjusted R^2= 0.341). Both body weight (β=0.362, p<0.001) and age (β =0.318, p<0.001) were significant predictors of C 02.

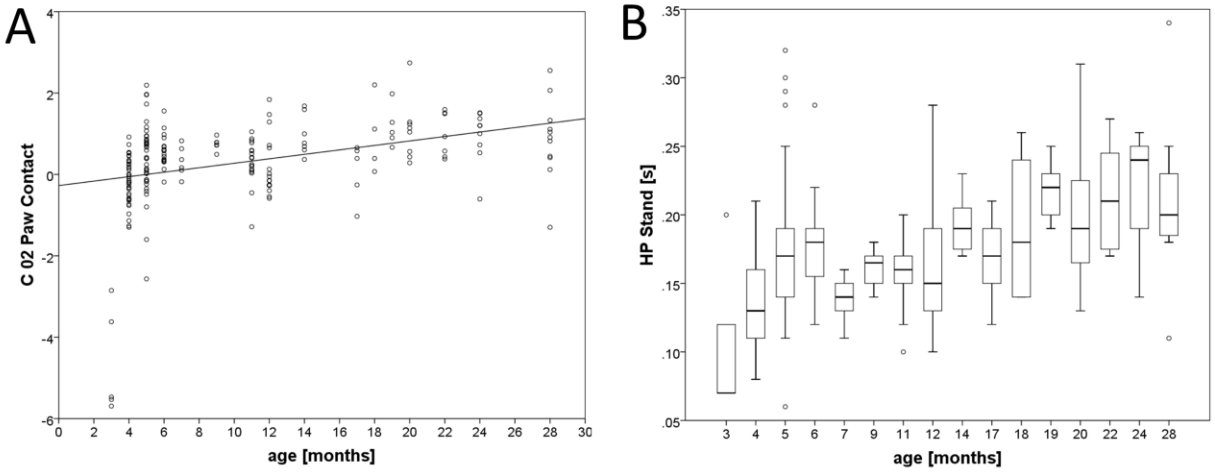


Fig. 4: Plots for depicting correlations between age and C 02 Paw Contact *(A) and original data for "HP Stand" by age (B) of animals of the 7.1 CW version. There is an increase in C 02 due to age, which can be observed also in the original data of, for example, "HP Stand". The stand duration is increasing with age. HP- hind paw*

Two other components showed a correlation coefficient above |0.250| with age. These are C 04 *Swing Phase* and C 08 *Stride Length* for the XT version only ($r_{(197)}$= 0.497 and $r_{(197)}$= -0.255, respectively). These correlations are not very strong. As we saw sex effects in C 04 *Swing Phase* and C 08 *Stride Length* in the analysis of all animals in the XT version we took a closer look at the influence of sex in these components with this reduced animal number. We did a regression analysis for sex and age and found that 29% of the variation in C 04 can be explained by this model ($F_{(2,195)}$=41.217, p<0.001, adjusted R^2= 0.290). Here both sex (β=0.220, p<0.001) and age (β=0.493, p<0.001) were significant predictors, but the beta values suggest a stronger influence of age than sex. The regression analysis for C 08 *Stride Length* revealed an influence of both age (β=-0.258, p<0.001) and sex (β=0.172, p=0.013), but describing thereby only 9% of the variation ($F_{(2,195)}$=10.16, p<0.001; adjusted R^2= 0.085). Again the beta values suggest a stronger influence of age than of sex on C 08 *Stride Length*.

Taken all together we can see rather small effects of sex, body weight and age on the components. We see sex and body weight effects in C 01 *Print Dimensions* in both CW versions. In C 08 *Stride Length,* the second spatial component, we see an influence of age and sex only in the XT version. For C 02 *Paw Contact* we find an age effect in both versions. In the XT version there is also an influence of body weight on this component. For the other temporal dimensions there are different patterns in the two versions: we see a sex effect in both C 04 *Swing Phase* and C 07 *Turning Point* in the XT version, together with an age effect in C 04 *Swing Phase.* In the 7.1 version we only see a body weight effect for C 07 *Turning Point*. In coordination both versions detect sex effects in C 05 *Coordination of diagonal and ipsilateral pairs.* In C 10 *Variation FP girdle pairs and Print Position* we also see an influence of sex in both versions. All other components related to coordination or variation do not show any effects of sex, age or body weight.

<u>*Strong influence of genetic background on the gait profile*</u>

Using hierarchical clustering similarities in gait profile between groups of mice were investigated. We conducted the hierarchical clustering by Wards method using the medians of the individual component scores per mouse line, genotype and sex (one column in Fig. 5 refers to one group, for example: male wild types of a specific mouse line). The dendrogram in Fig. 5 is used to represent the distance between groups and clusters by its branch lengths. The shorter the branch length, the shorter the distance and the greater the similarity between two pairs (e.g. groups or clusters). The heat map depicts the color-coded medians of the components for the single groups. The first separation in the dendrogram was defined by the CW version (see dendrogram in Fig. 5). The heat map suggests that C 01 *Print Dimensions*, C 02 *Paw Contact* and C 04 *Swing Phase* play a crucial role in separating the CW versions. Consecutive separations occur on the basis of genetic background. First a group consisting essentially of

lines on a C3H background separates and then a partitioning between two C57BL6 backgrounds occurs for animals tested on the 7.1 version. Here one group, the "C57BL6-mix", was less often backcrossed to a C57BL6 background compared to the "C57BL6" group. The components C 04 *Swing Phase*, C 05 *Coordination of the diagonal and ipsilateral pairs* and C07 *Turning Point* seem to influence the division between background strains. In Fig. 6 graphs for selected parameters of the original data are presented. For component C 04 the parameters "FP Swing Speed" and "BOS HP", for C 05 "BOS FP" and "PhD RF-RH" (Phase Dispersion RF->RH) and for C 07 "HP Max Contact At" are presented. The C3H background differs from the C57BL6 background in that they have smaller HP BOS but a larger FP BOS, a slightly higher swing speed and a different timing in PhD RF-RH. Also they have an earlier Maximum Contact. The difference between the C57BL6 and the C57BL6 mix becomes obvious in the BOS HP and FP (higher values of the C57BL6 mix compared to the C57BL6 in both parameters) as well as the FP swing phase (a reduced swing speed) ("Run duration": sex-background effect: $F_{(2,334)}$=7.364, p=0.001; Males: $F_{(2,171)}$=9.655, p<0.001; post hoc Bonferroni, mean difference (MD): C57BL6 vs C57BL6 mix: -0.8695, p<0.001; C57BL6 vs C3H: ns; C57BL6 mix vs C3H: 0.7187, p=0.011; Females: ns; "BOS FP": background effect: $F_{(2,334)}$=146.354, p<0.001; post hoc Bonferroni: MD: C57BL6 vs C57BL6 mix: -2.2612, p<0.001; C57BL6 vs C3H: -3.7361, p<0.001; C57BL6 mix vs C3H: -1.4749, p<0.001; "BOS HP": sex-background effect: $F_{(2,334)}$=14.860, p<0.001; Males: $F_{(2,171)}$=37.207, p<0.001; post hoc Bonferroni: MD: C57BL6 vs C57BL6 mix: -1.3186, p=0.01; C57BL6 vs C3H: 3.2298, p<0.001; C57BL6 mix vs C3H: 4.5484, p<0.001; Females: $F_{(2,162)}$=64.546, p<0.001; post hoc Bonferroni: MD: C57BL6 vs C57BL6 mix: -5.0638, p<0.001, C57BL6 vs C3H: 3.4776, p<0.001; C57BL6 mix vs C3H: 8.5414, p<0.001; "FP Swing Speed": background effect: $F_{(2,334)}$=26.425, p<0.001; post hoc Bonferroni: MD: C57BL6 vs C57BL6 mix: 0.1334, p<0.001; C57BL6 vs C3H: -0.1010, p=0.002; C57BL6 mix vs C3H: -0.2343, p<0.001; "PhD RF-RH": background effect: $F_{(2,334)}$=81.338, p<0.001; post hoc Bonferroni: MD: C57BL6 vs C57BL6 mix: ns; C57BL6 vs C3H: -10.774, p<0.001; C57BL6 mix vs C3H: -9.843, p<0.001; "HP Max Contact At": sex-background interaction: $F_{(2,334)}$=9.087, p<0.001; Males : $F_{(2,171)}$=76.711, p<0.001; post hoc Bonferroni, MD: C57BL6 vs C57BL6 mix: -8.1375, p<0.001, C57BL6 vs C3H: 6.3515, p<0.001; C57BL6 mix vs C3H: 14.4890, p<0.001; Females: $F_{(2,162)}$=7.075, p=0.001; post hoc Bonferroni: MD: C57BL6 vs C57BL6 mix: ns; C57BL6 vs C3H: 3.9581, p=0.017; C57BL6 mix vs C3H: 6.3927, p=0.001)

When taking a closer look at the distribution of the mouse lines it becomes clear that there is no strong separation between sexes or genotypes (mutant vs. wild type) within one mouse line. Generally the genotypes of one mouse line cluster together, thus emphasizing again the strong influence of the genetic background on gait parameters.

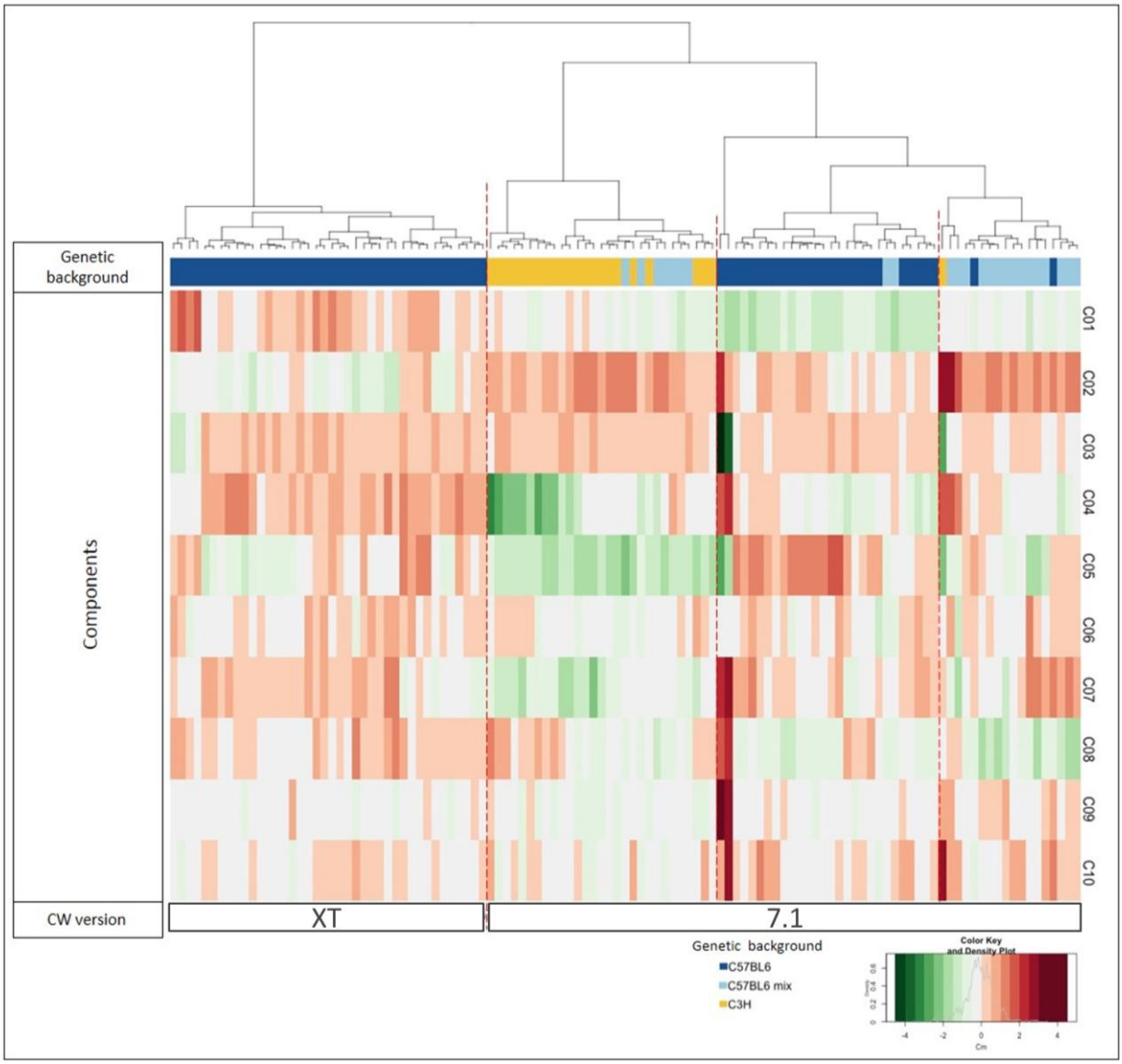
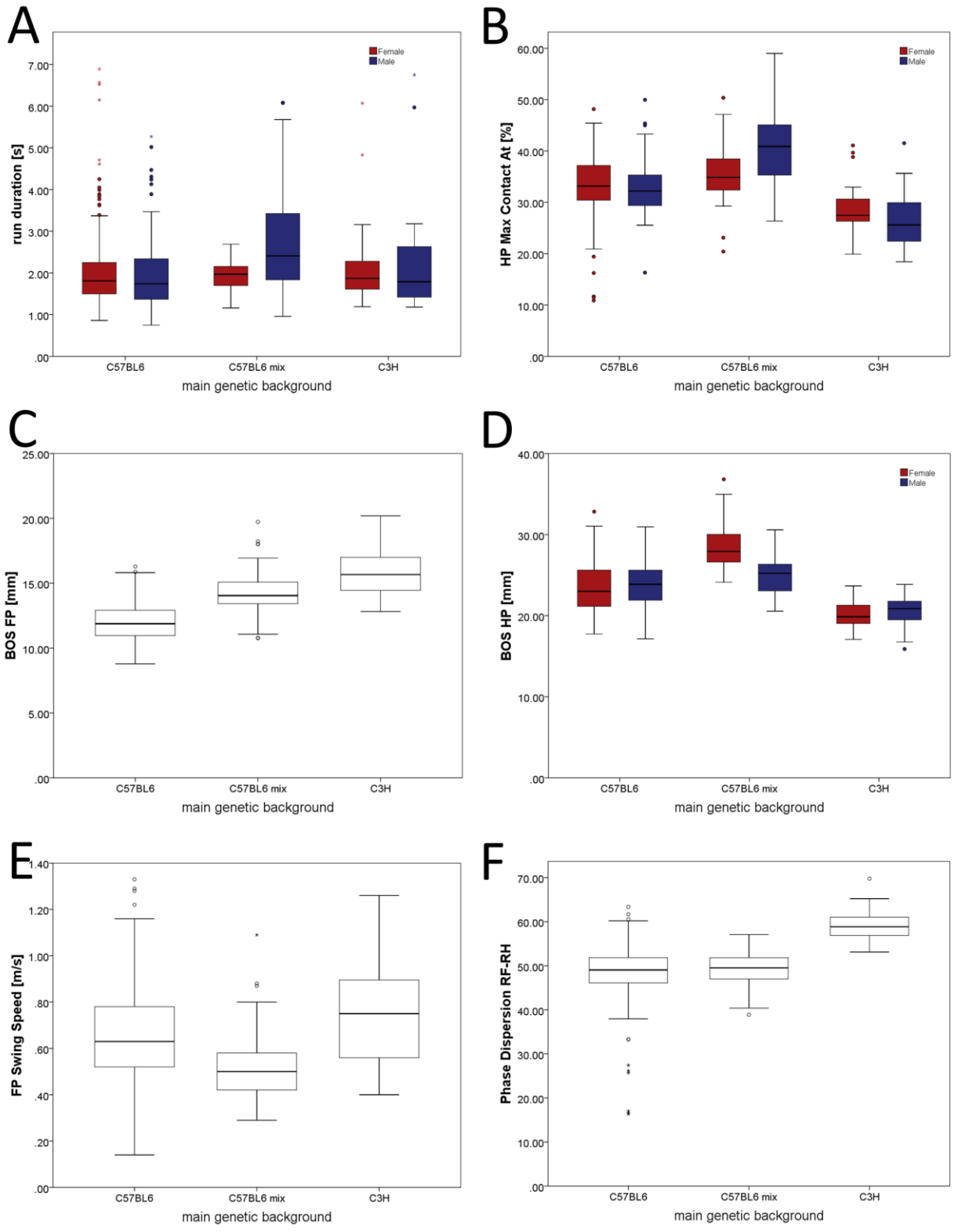


Fig. 5: Dendrogram and heat map from the hierarchical cluster analysis of groups

One vertical line represents one group, for example male wildtypes of a specific mouse line. The horizontal lines represent the color-coded medians per group of the different components (see right hand side; color-coding presented in the key underneath the plot). The first separation in the dendrogram clusters the two different versions. The left branch includes all groups analyzed on the CW XT version and the right branch includes exclusively those tested on the 7.1 version (see bottom of graph). The second division is mainly defined by the background strain (compare bar on the top). With a few exceptions all groups with a C3H background (yellow) cluster in the left branch, whereas most of the groups with a C57BL6 (dark blue) or C57BL6 mix (light blue) background are on the right branch within the 7.1 cluster. Groups of a C57BL6 and C57BL6 mix then diverge with the next branching.

1


A
run duration [s]
7.00
6.00
5.00
4.00
3.00
2.00
1.00
.00
Female
Male
C57BL6
C57BL6 mix
C3H
main genetic background
B
HP Max Contact At [%]
60.00
50.00
40.00
30.00
20.00
10.00
.00
Female
Male
C57BL6
C57BL6 mix
C3H
main genetic background
C
BOS FP [mm]
25.00
20.00
15.00
10.00
5.00
.00
C57BL6
C57BL6 mix
C3H
main genetic background
D
BOS HP [mm]
40.00
30.00
20.00
10.00
.00
Female
Male
C57BL6
C57BL6 mix
C3H
main genetic background
E
FP Swing Speed [m/s]
1.40
1.20
1.00
.80
.60
.40
.20
.00
C57BL6
C57BL6 mix
C3H
main genetic background
F
Phase Dispersion RF-RH
70.00
60.00
50.00
40.00
30.00
20.00
10.00
.00
C57BL6
C57BL6 mix
C3H
main genetic background

2

Fig. 6: Original data from parameters representative for the components that seem to differentiate the genetic background in the 7.1 version. In the run duration (A) we find a sex-background interaction; in the males the C57BL6 mix is different to the two other groups; in females there is no difference between groups. (B) In HP Max Con At (parameter of C 07) we detect a sex-background effect; all male groups differ from each other. The female C3H are different from the other two groups. (C) BOS FP (a parameter of C 05) a background effect is observed- all groups differ. (D) In the BOS HP (a parameter of C 04) there is a sex-background interaction: in males and females all groups differ respectively. (E) FP Swing Speed (a parameter of C 04) there is a background effect detectable. Here the C57BL6 mix group has a lower swing speed compared to the other two groups. (F) In PhD RF-RH (a parameter of C 05) a background effect is detectable. The C3H group differs from the other two groups. HP- hind paws, FP- front paws, BOS- base of support, PhD- Phase Dispersion, RF- right front paw, RH- right hind paw

No redundancy in CW, OF and VP measures

We wanted to know if different tests for locomotor phenotypes indeed assess different aspects of locomotion or if there are redundancies between them. To this end we analyzed two different data sets with PCA: One set included data from CW, OF and VP test and the other set consisted of data from the IMPC including data from OF, SHIRPA, RR, GS and HC.

The PCA on CW, OF and VP data included historical data from our lab. A total of 1057 cases were analyzed including 147 parameters of the three tests (113 from the CW, 31 from the OF and 3 from the VP test; see also Table S2). The sampling adequacy of this parameter set was confirmed by the KMO measure of 0.831 and the significant Bartlett´s test (Chi^2= 512264.116, df= 10731, p<0.001). Extraction of 13 components explaining 80% of the total variance was done using the oblique rotation. The loadings of parameters are depicted in the structure matrix table (see Table S5). The components were characterized according to the highest loadings of the individual parameters (see Table 2).

Table 2: PCA for CW, OF and VP: Component Characterization

			Component name	**Explained variance [%]**	
Open Field		C 1	OF_Activity	19.0 %	23%
		C 7	OF_Emotionality	4.0 %	
CatWalk	Spatial Dimension	C 2	CW_Print Dimensions	17.0 %	17%
	Temporal Dimension	C 3	CW_Swing Phase	11.0 %	19%
		C 5	CW_Stand and Stride Length	5.3 %	
		C 8	CW_Turning Point FPs	2.6 %	
	Interlimb Coordination	C 4	CW_Coordination of diagonal and ipsilateral pairs	7.3 %	13%
		C 6	CW_Coordination of HP girdle pairs	4.2 %	
		C 11	CW_Coordination of FP girdle pairs	1.8 %	
	Variation	C 9	CW_Variation of diagonal and ipsilateral incl LH	2.4 %	6%
		C 10	CW_Variation incl RH	2.3 %	

		C 13	CW_Variation of FP girdle pairs and Turning Point HPs	1.5 %	
Vertical Pole		C 12	Pole	1.6 %	2%

OF- Open Field; CW- CatWalk; FPs- front paws; HPs- hind paws; LH- left hind paw

The parameters clearly cluster according to their respective tests. Component C 12 contains all parameters measured by the VP test; two components contain OF parameters (C 1 and 7) and the rest of the components contain the CW parameters. For the CW parameters it is noticeable that the components have a similar make up of parameters as those in the first PCA and can be allocated to the four dimensions, i.e. spatial and temporal dimensions as well as interlimb coordination and its variation. In the component correlation matrix (see Table S6) it becomes clear that a few components correlate, although not very strongly. The highest correlation is between the two OF components (r=-0.477). As already suggested by the analysis there is no strong correlation between parameters of different tests. From this analysis it becomes very clear that the three tests are not redundant but measure different aspects of locomotion.

IMPC data set shows that OF and HC measure two dissociable aspects of locomotor activity

Also in the IMPC data set the parameters are organized according to their respective tests, suggesting no redundancies between tests. The PCA was done including 1327 animals and 20 parameters comprising of measures from the OF, GS, SHIRPA, RR and HC (Table S7). The sampling adequacy was confirmed by KMO of 0.706 and a significant Bartlett´s test (Chi^2= 64427.87, df= 190, p<0.001). We extracted 6 components explaining 84.1% of the total variance using the oblique rotation (Table 3; for the structure matrix see Table S8).

Table 3: PCA for IMPC data: Component Characterization

	Component name	**Major parameters**	**Explained variance [%]**
C 1	OF Locomotion	Distance and Speed	31.3 %
C 2	OF Emotionality	Time spent in center, Distance travelled in center, Resting time in center, Number of entries in center	18.2 %
C 3	OF Exploration	Resting time in arena and periphery, Number of rearings	11.0 %
C 4	Home cage activity	Distance, Rearings and Speed	9.1 %
C 5	Grip Strength	Grip strength of FP and all paws	8.0 %
C 6	Rotarod	Mean latency to fall, Number of passive rotations	6.5 %

FP- front paws, OF- Open Field

After characterizing the components according to their main parameters we again see an allocation of parameters related to their respective test. An exception is the activity parameter from SHIRPA, which has its major loading in C 5 *Grip Strength* (0.386), which is not very strong, and its second and third highest loadings into C 4 *Home cage activity* (0.369) and C 1 *OF Locomotion* (0.364), respectively, suggesting that all components do not explain a high amount of the parameter´s variation. Interestingly both C 1 and C 4 contain the activity measures from OF and the HC respectively and there is a mild

component correlation between them (r=0.327) (see Table S9). Taking a closer look at the original data of all the horizontal activity parameters from OF, HC and SHIRPA show that there is only a mild correlation between them (OF Distance- HC Distance: $r_{(1326)}$= 0.326, OF Distance- SHIRPA: $r_{(1326)}$=0.370, HC Distance- SHIRPA: $r_{(1326)}$=0.240, see also Fig. 7). This suggests that dissociable aspects of activity are measured by the different tests, again strengthening the interpretation of no redundancy between tests.

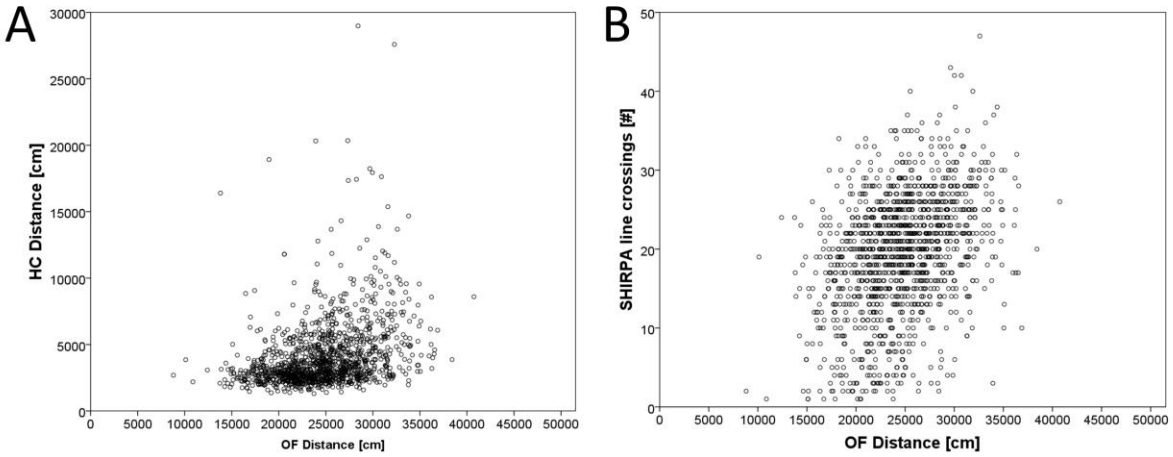


Fig. 7: Scatter plots for correlations between measures of activity. Graph A depicts "OF Distance" and "HC Distance" and graph B "OF Distance" and "SHIRPA Line crossings". There are only weak correlations between the parameters. OF- Open Field, HC- Home Cage

Discussion

In our study we performed PCAs on different data sets related to locomotor phenotypes to refine future data analysis and to investigate possible redundancies between different tests. Knowledge of the latter better informs future decisions on phenotyping strategies: if tests were redundant, one could focus on the simplest one. Our results suggest that for future CW analysis a primary focus can be on FP and HP measures, and not all individual paws, as well as one of the interlimb coordination measures (e.g. only Phase Dispersions). Also we detected minor effects of sex, body weight and age on the components, but a large effect of genetic background and CW version. The analysis of different tests revealed no significant redundancies between them, thereby emphasizing their originality.

A PCA is a useful tool to reduce dimensions of large data sets and consequently analyze the components. This approach has already been successfully used in several papers to identify linear combinations of variables in a high-dimension space best representing the variance that is present in the data [7,19,20]. Nonetheless it is clear that in some publications important specifications for corroborative purposes are missing. This is mainly missing information on sampling adequacy, such as the KMO and Bartlett's test, and/or details on data processing and component extraction methods. All these are important for the reader to judge the validity of the results [5].

Locomotor phenotypes are composed of several different dimensions including gait performance. Gait performance can be analyzed by the CW system measuring over 200 parameters. We used the PCA to examine the relative contribution of 112 parameters to the variance in the data obtained from CW measurements. Our results show that the large amount of parameters measured by the CW system can be compacted via PCA into 10 components explaining 80% of the total variance. These 10 components nicely describe several aspects of gait performance, including spatial (C 01 *Print Dimensions* and 08 *Stride Length*) and temporal (C02 *Paw Contact*, 04 *Swing Phase* and C07 *Turning Point*) dimensions as well as interlimb coordination (C 05 and 06) and its variation (C 03, 09 and 10). Interestingly the spatial and temporal dimensions as well as coordination including its variation, all contribute equally to the explained variance in the data. With respect to future studies, the data suggests that focusing on the parameters of FP and HP is sufficient (instead of considering parameters of each paw individually), as these always cluster together with their respective single paws in one component. Obviously this is only applicable if the experimental design does not dictate a different approach, for example for lateral lesion models, where the sides/single paws have to be compared. Also for the coordination parameters, i.e. couplings and phase dispersions, it is adequate to decide on one. When analyzing individual mouse lines for genotype effects separately it becomes apparent that the FP and HP reliably reflect effects in single paws, if these are not too small. This applies also for phase dispersion representing in essence the couplings. BOS FP and HP should be analyzed individually, as they have quite opposing patterns of loadings onto the different components. This is reflected in their correlation with other parameters. The heavier the mouse the larger the BOS HP, but this effect is far less strong in the BOS FP. When looking at other correlations within the original data and BOS, we have to check the two CW versions separately (see below). Generally, when the BOS gets wider the run duration increases. Furthermore, the stance phase gets longer and the stride length shorter all culminating in a slowing in movement. In the 7.1 version the widening of the BOS FP is correlated with changes in the interlimb coordination, the step pattern used and a reduction in support of diagonal paw pairs. In the XT version we also see a reduction in diagonal support but with a wider BOS HP. Here a positive correlation with “Print position” indicates that the hind paw is placed behind the front paw, again suggesting a slowing of movement. For the BOS FP in the XT version we see a different picture; there is no correlation with run duration or stride length. We do see an increased stance phase, indicated by an increased duty cycle, and a change in lateral support. Also the turning point in the FP is delayed.

Further analysis of the extracted components indicated that the two CW versions show clear differences, as we see in 9 out of 10 components significant version effects. When looking at the parameters we see strong differences between the two versions appearing in spatial and temporal dimensions. As there are less strong or no differences at all in the coordination and variation components between the two versions this again suggests a discrepancy in the accuracy of measuring dimensions between the versions. Possibly the precision of the system has changed and as such dimensions differ. In the XT version a high-speed camera is recording the images at a much higher rate and a red background illumination has been integrated. Thus it is possible that the paw print can be delineated in more accuracy. This is also in accordance with Chen at al [21], who found that the XT version can detect more subtle alterations, which cannot be detected by the earlier version in a model of sciatic nerve injury. Theoretically another issue in our study could be the heterogeneity of animals measured by the two versions. The age range, for example, of the 7.1 version is between 3 to 28 months and for the XT version it is between 3 and 15 months of age (see Fig. S2). Another factor differing between the two versions is the variability in genetic background. The genetic background for the

animals tested on the 7.1 version is broader (including C57BL6, C3H, and mixed backgrounds), for the XT version the animals were of a more similar genetic background. Both factors – age and genetic background- do have an influence on CW performance and this has to be taken into account. It stresses the importance of comparing animals with their corresponding littermates tested at the same time and same test version. To take a closer look at these possible confounding factors in our study we also compared the two versions by including only animals on a C57BL6 (mix) background and only up to an age of 15 months, so that a vast amount of heterogeneity is circumvented. Nonetheless we still observe significant differences between the versions thus confirming a profound difference between the CW versions. Therefore we decided to stay with the approach of including all animals. Also as this has the benefit to analyze effects of age, body weight and genetic background strains with decent animal numbers.

We continued our analysis by checking the two versions separately and looking for the influence of sex, body weight and age on the components. We could see that body weight has an influence on C 01 *Print Dimensions* in both CW versions. This is not too surprising as parameters of print dimension, i.e. print length, width and area, all load highly onto this component and are expected to change with body weight [22]. Interestingly, we see that in females the changes due to body weight seem to be stronger than in males, where especially in the 7.1 version males show no correlation between C 01 and body weight. Body weight also has an influence on C 07 *Turning Point* in the 7.1 version. Here we see an increase in all parameters of "Maximum contact at %", indicating a delayed turning point within the stand phase with increasing body weight.

Age has an influence on C 02 *Paw Contact* in both versions, as well as C 04 *Swing Phase* and C 08 *Stride Length* in the XT version. For C 02 *Paw Contact* we find also a body weight influence in the XT. Generally, in C 02, parameters such as the duty cycle and stand duration increase with age suggesting a slowing of movement, which is also supported by the increase of run duration and a reduction in diagonal support. An age-related increase in swing duration and a reduction in swing speed can be observed, evident mainly in the XT version, where also the corresponding component, C 04 *Swing Phase*, shows a significant age effect. In the XT version C 08 *Stride Length* is influenced by age as well. We see a negative correlation between the parameters of "Stride length" and age, again evidence of a slowing of movement with age. This slowing of movement we see in mice is also true in humans [23,24]. Summing up, both body weight and age do have an effect on the components, but it is rather small.

Two components showing sex dependent effects are C 10 *Variation excl RH for girdle pairs* and C 05 *Coordination of the diagonal and ipsilateral pairs* for both versions. In C 10 *Variation FP girdle pairs and Print Position* the parameter of "Print position", which measures the placement of the HP with respect to the FP, is larger in males than in females. This is partly due to body weight differences. There is no obvious sex difference in the parameters of variation in this component. For C 05 *Coordination of the diagonal and ipsilateral pairs* we see opposing effects in the two CW versions: in the 7.1 version females have a higher score than males and in the XT version it is the opposite. The parameter "Step sequence AB"- the most frequently used step pattern by all animals- has a higher percentage in the males of the 7.1 version compared to the respective females and a higher value for the females of the XT version compared to the respective males. The couplings and phase dispersions of C 05 show the opposite in the two versions between the sexes.

Clustering the medians of the individual component scores of the different mouse lines according to their line, genotype and sex showed that the CW version defines the first separation in the dendrogram, again depicting discrepancies between the two. As mouse lines, independent of their mutation and according to their genetic background, cluster together, a strong influence of the genetic background (and not the mutation) is depicted. Seemingly the genetic background has a profound influence on gait parameters, which can be detected by the CW system. The heat map suggests that C 04 *Swing Phase*, C 05 *Coordination of the diagonal and ipsilateral pairs* and C 07 *Turning Point*, contribute strongly to the separation between a C3H and a C57BL6 background. This can be illustrated by several parameters: although all animals need approximately the same time to cross the runway, we see reduced swing speed in the C57BL6 mix animals and a slightly higher swing speed in the C3H animals compared to C57BL6. Also the inter-paw relations differ between the background strains. The same holds true for BOS; in both the FP and HP C57BL6 mix animals have a broader BOS compared to C57BL6, whereas the C3H have a broader FP BOS but narrower HP BOS compared to C57BL6. For the parameter HP Max Con At [%] C3H animals show an earlier Max contact. It is not too surprising that also here different strains show different patterns. Strain differences have been shown for several tests and this stresses the importance of comparing littermates and not similar genetic backgrounds with each other, when looking for effects of a mutation [25,26].

Still the CW is a valuable tool for investigating genotype specific (in terms of a mutation) gait phenotypes. In the 25 lines we included in our study 21 of them showed a genotype specific phenotype. These numbers seem very high and have to be considered in respect to the selection of lines for CW testing; we subjected lines to the CW for which we hypothesized a gait phenotype, due to the mutation or results from other locomotor tests.

Next we addressed the question of redundancies between tests. We asked how parameters of several different tests would group into principal components. We did this with a subset of the animals used above for the CW analysis, which all performed the CW, OF and VP test. We extracted 13 components explaining 80% of the total variance. The parameters from different tests group into different components, so that each component can be assigned to one test. This indicates that the tests do not predict the outcome of a different one and are therefore useful expansions and complementations of each other and not redundant. This is also suggested by a totally different data set. Here data from IMPC animals of the German Mouse Clinic was taken from different tests measuring different aspects of locomotor performance. This included parameters of the OF, GS, RR, HC activity as well as activity measured by SHIRPA. Again we see a separation of the parameters to their tests. An exception is the activity measure of the SHIRPA protocol which is attributed to C 5 *Grip Strength*. Nonetheless it should be mentioned that the loading onto this component is not very high (0.386) and it loads onto C 4 *Home cage activity* and C 1 *OF Locomotion* with a similar loading (0.369 and 0.364, respectively), suggesting that none of the components describe a high variation of this parameter. The allocation of parameters of a single test to one component stresses the validity of each single test. This is in accordance with the specificity the tests were designed for, for example GS for measuring muscle force, RR for motor coordination and balance and OF for spontaneous locomotion as well as emotional phenotypes. Generally one would not expect that these highly specialized tests can replace another, but it is possible that some of the measured parameters can correlate with parameters of the other test. For example one would not expect that HC observations can replace OF, but it would be possible that the measured activity in HC would hint towards effects on activity parameters of the OF, or vice versa. Two

components, C 1 and C 4, contain the activity measures from OF and the HC, respectively, which moderately correlate. When going back to the original data we can see that the correlation between the activity measures of the different tests are not that impressive ("OF Distance" and "SHIRPA Activity" r=0.370; "OF Distance" with "HC Distance" r=0.326; "SHIRPA Activity" with "HC Distance" r=0.240; "OF ArenaAverageSpeed" with "HC Speed_mean" r=0.336; "OF Rearings" with "HC Rearings" r=0.123), suggesting that, to some extent, there is a small relationship between activity measures between tests but they do not really have predictive value. This can be explained by the different designs for measuring activity. In the SHIRPA protocol used here the activity is measured in the first 30 seconds after placing the animal into a new environment. It is measured by counting line crossings. In the OF and HC, activity is measured via infra-red beam breaks. In OF the activity is recorded for 20 minutes whereas in the HC activity is measured for at least 21 h. In both SHIRPA and OF, activity is measured in a new unfamiliar arena, whereas in the HC the activity is measured in a familiar environment. Thus different aspects come into play, such as stress-reactions and anxiety-related behaviors in a new environment. It is known that animals can differ in these distinct tests and there are underlying genetic factors contributing to these discrepancies as shown by studies with collaborative crosses [27-30]. Again this strengthens the value of each single test applied, as they do measure different aspects of locomotor performance and thus are not redundant.

In conclusion, we have shown that a large amount of data points generated by the CW system can be meaningfully compacted via PCA to a few components. Here we extracted 10 components explaining 80% of the total variance in the data. The loadings of parameters onto components suggest the possibility to focus on FP and HP parameters neglecting the parameters for every single paw, as well as analyzing only one measure of coordination, i.e. couplings or phase dispersion, as long as the experimental question does not require specific analysis. It is necessary to look at the BOS FP and HP separately, as they seem to be independent of each other. The version of the CW system seems to have a great influence on the data, especially in the paw dimensions. Sex, body weight, age and genetic background effects influence the different components, suggesting to carefully consider these in experimental planning- especially when comparing data over time. We could show that the use of a PCA can reduce a big data set to a few meaningful components, which can easily be used to unsheathe underlying patterns. The PCAs, which included different tests, reveal the validity of each single test, as they measure different dimensions of locomotor performance [31].

References

1. Infrafrontier Consortium (2015) INFRAFRONTIER—providing mutant mouse resources as research tools for the international scientific community. Nucleic Acids Research 43: D1171-D1175.
2. de Angelis MH, Nicholson G, Selloum M, White J, Morgan H, et al. (2015) Analysis of mammalian gene function through broad based phenotypic screens across a consortium of mouse clinics. Nature genetics 47: 969-978.
3. Hamers FPT, Lankhorst AJ, van Laar TJ, Veldhuis WB, Gispen WH (2001) Automated Quantitative Gait Analysis During Overground Locomotion in the Rat: Its Application to Spinal Cord Contusion and Transection Injuries. Journal of Neurotrauma 18: 187-201.

4. Batka RJ, Brown TJ, McMillan KP, Meadows RM, Jones KJ, et al. (2014) The Need for Speed in Rodent Locomotion Analyses. Anatomical record (Hoboken, NJ : 2007) 297: 1839-1864.
5. Budaev SV (2010) Using Principal Components and Factor Analysis in Animal Behaviour Research: Caveats and Guidelines. Ethology 116: 472-480.
6. Gonik M, Frank E, Keßler MS, Czamara D, Bunck M, et al. (2012) The endocrine stress response is linked to one specific locus on chromosome 3 in a mouse model based on extremes in trait anxiety. BMC Genomics 13: 579.
7. Vannoni E, Voikar V, Colacicco G, Sánchez MA, Lipp H-P, et al. (2014) Spontaneous behavior in the social homecage discriminates strains, lesions and mutations in mice. Journal of Neuroscience Methods 234: 26-37.
8. Ferguson AR, Irvine K-A, Gensel JC, Nielson JL, Lin A, et al. (2013) Derivation of Multivariate Syndromic Outcome Metrics for Consistent Testing across Multiple Models of Cervical Spinal Cord Injury in Rats. PLOS ONE 8: e59712.
9. Ohl F, Roedel A, Binder E, Holsboer F (2003) Impact of high and low anxiety on cognitive performance in a modified hole board test in C57BL/6 and DBA/2 mice. Eur J Neurosci 17: 128-136.
10. Brown SDM, Moore MW (2012) Towards an encyclopaedia of mammalian gene function: the International Mouse Phenotyping Consortium. Disease Models & Mechanisms 5: 289-292.
11. Hölter SM, Glasl L (2012) High-Throughput Mouse Phenotyping. In: Lane EL, Dunnett SB, editors. Animal Models of Movement Disorders: Volume I. Totowa, NJ: Humana Press. pp. 109-133.
12. Irwin S (1968) Comprehensive observational assessment: Ia. A systematic, quantitative procedure for assessing the behavioral and physiologic state of the mouse. Psychopharmacologia 13: 222-257.
13. Karp NA, Mason J, Beaudet AL, Benjamini Y, Bower L, et al. (2017) Prevalence of sexual dimorphism in mammalian phenotypic traits. Nature Communications accepted
14. Donner NC, Lowry CA (2013) Sex differences in anxiety and emotional behavior. Pflugers Archiv : European journal of physiology 465: 601-626.
15. Gillies GE, Pienaar IS, Vohra S, Qamhawi Z (2014) Sex differences in Parkinson's disease. Frontiers in Neuroendocrinology 35: 370-384.
16. Krizo JA, Mintz EM (2014) Sex Differences in Behavioral Circadian Rhythms in Laboratory Rodents. Frontiers in Endocrinology 5: 234.
17. Manoli DS, Fan P, Fraser EJ, Shah NM (2013) Neural control of sexually dimorphic behaviors. Current opinion in neurobiology 23: 330-338.
18. Beery AK, Zucker I (2011) Sex Bias in Neuroscience and Biomedical Research. Neuroscience and biobehavioral reviews 35: 565-572.
19. Irvine K-A, Ferguson AR, Mitchell KD, Beattie SB, Lin A, et al. (2014) The Irvine, Beatties, and Bresnahan (IBB) Forelimb Recovery Scale: An Assessment of Reliability and Validity. Frontiers in Neurology 5: 116.
20. Catuara-Solarz S, Espinosa-Carrasco J, Erb I, Langohr K, Notredame C, et al. (2015) Principal Component Analysis of the Effects of Environmental Enrichment and (-)-epigallocatechin-3-gallate on Age-Associated Learning Deficits in a Mouse Model of Down Syndrome. Frontiers in Behavioral Neuroscience 9: 330.
21. Chen Y-J, Cheng F-C, Sheu M-L, Su H-L, Chen C-J, et al. (2014) Detection of subtle neurological alterations by the Catwalk XT gait analysis system. Journal of NeuroEngineering and Rehabilitation 11: 62.
22. Parkkinen S, Ortega FJ, Kuptsova K, Huttunen J, Tarkka I, et al. (2013) Gait Impairment in a Rat Model of Focal Cerebral Ischemia. Stroke Research and Treatment 2013: 12.
23. Prince F, Corriveau H, Hébert R, Winter DA Gait in the elderly. Gait & Posture 5: 128-135.

24. Callisaya ML, Blizzard L, Schmidt MD, McGinley JL, Srikanth VK (2008) Sex Modifies the Relationship Between Age and Gait: A Population-Based Study of Older Adults. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences 63: 165-170.
25. Crawley JN, Belknap JK, Collins A, Crabbe JC, Frankel W, et al. (1997) Behavioral phenotypes of inbred mouse strains: implications and recommendations for molecular studies. Psychopharmacology 132: 107-124.
26. Logue SF, Owen EH, Rasmussen DL, Wehner JM (1997) Assessment of locomotor activity, acoustic and tactile startle, and prepulse inhibition of startle in inbred mouse strains and F1 hybrids: Implications of genetic background for single gene and quantitative trait loci analyses. Neuroscience 80: 1075-1086.
27. Wang X-D, Labermaier C, Holsboer F, Wurst W, Deussing JM, et al. (2012) Early-life stress-induced anxiety-related behavior in adult mice partially requires forebrain corticotropin-releasing hormone receptor 1. European Journal of Neuroscience 36: 2360-2367.
28. Bronikowski AM, Carter PA, Swallow JG, Girard IA, Rhodes JS, et al. (2001) Open-field behavior of house mice selectively bred for high voluntary wheel-running. Behav Genet 31: 309-316.
29. DeFries JC, Wilson JR, McClearn GE (1970) Open-field behavior in mice: selection response and situational generality. Behav Genet 1: 195-211.
30. Majdak P, Bucko PJ, Holloway AL, Bhattacharya TK, DeYoung EK, et al. (2014) Behavioral and pharmacological evaluation of a selectively bred mouse model of home cage hyperactivity. Behav Genet 44: 516-534.
31. Brooks SP, Dunnett SB (2009) Tests to assess motor phenotype in mice: a user's guide. Nat Rev Neurosci 10: 519-529.

Acknowledgements

This work was supported by the German Science Foundation Collaborative Research Centre (CRC) 870 to L.E., by funds from the Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst within Bavarian Research Network “Human Induced Pluripotent Stem Cells” (ForIPS) to F.G., the German Federal Ministry of Education and Research (BMBF) through the Integrated Network MitoPD (Mitochondrial endophenotypes of *Morbus Parkinson*), under the auspices of the e:Med Programme (grant 031A430E) to P.D. and the DFG grant 'DJ-1 Linked Neurodegeneration Pathways in New Mouse Models of Parkinson's Disease' (WU 164/5-1) to A.Z.. The scientists of the GMC were supported by the German Federal Ministry of Education and Research (Infrafrontier grant 01KX1012).

We want to thank Bettina Sperling, Jan Einicke and Albert Langer for expert technical assistance.

Supporting Information

Table S1: Number of cases used for PCA

Note that the cases from the CW OF VP Analysis are part of the CW Analysis cases. CW- CatWalk; OF- Open Field; VP- Vertical Pole

Number of cases	**CW Analysis**			**CW OF VP Analysis**			**IMPC Analysis**
	Total	**7.1**	**XT**	***Total***	**7.1**	**XT**	***Total***
Total cases	***1499***	*627*	*872*	***1057***	*358*	*399*	***1327***
female cases	*691*	283	408	*482*	153	329	*618*
male cases	*808*	344	464	*575*	205	370	*709*
Wild type cases	***660***	*276*	*384*	***436***	*157*	*279*	***665***
female cases	*316*	136	180	*197*	71	126	*312*
male cases	*344*	140	204	*239*	86	153	*353*
Mutant cases	***838***	*350*	*488*	***620***	*200*	*420*	***662***
female cases	*375*	147	228	*285*	82	203	*306*
male cases	*463*	203	260	*335*	118	217	*356*
Age range at CW (for IMPC at OF) testing (months)	*3-26*	3-26	3-15	*3-18*	4-18	3-12	*2*

Table S2: Parameters from CatWalk, Open Field and Vertical Pole

CatWalk Parameters			Open Field Parameters	Vertical Pole Parameters
Print Length	for RF, RH, LF, LH, FP and HP	Run_Duration	OF_Dist05	Pole_TimeTurn
Print Width		StepSequence_CA	OF_dist10	Pole_TimeComplete
Print Area		StepSequence_CB	OF_dist15	Pole_TimeDown
Max Contact Area		StepSequence_AA	OF_dist20	
Stand duration		StepSequence_AB	OF_distTot	
Swing duration		StepSequence_RA	OF_Rear05	
Swin Speed		StepSequence_RB	OF_Rear10	
Stride Length		StepSequence_RegularityIndex	OF_Rear15	
Duty Cycle		Support_Zero	OF_Rear20	
Max Conatct At		Support_Single	OF_RearTot	
BOS_FP		Support_Diagonal	OF_%DistCen05	
BOS_HP		Support_Girdle	OF_%DistCen10	
PrintPosition_RP		Support_Lateral	OF_%DistCen15	
PrintPosition_LP			OF_%DistCen20	
PhDig_RFLH_CStat		PhDig_RFLH_CStat_R	OF_%DistCenTot	
PhDig_LFRH_CStat		PhDig_LFRH_CStat_R	OF_%TiCen05	
PhGird_LHRH_CStat		PhGird_LHRH_CStat_R	OF_%TiCen10	
PhGird_LFRF_CStat		PhGird_LFRF_CStat_R	OF_%TiCen15	
PhIps_RFRH_CStat		PhIps_RFRH_CStat_R	OF_%TiCen20	
PhIps_LFLH_CStat		PhIps_LFLH_CStat_R	OF_%TiCenTot	
CoupDig_RFLH_CStat		CoupDig_RFLH_CStat_R	OF_WholeRest	
CoupDig_LFRH_CStat		CoupDig_LFRH_CStat_R	OF_WholeAvSpeed	
CoupDig_LHRF_CStat		CoupDig_LHRF_CStat_R	OF_PerDist	
CoupDig_RHLF_CStat		CoupDig_RHLF_CStat_R	OF_PerRest	
CoupGir_LHRH_CStat		CoupGir_LHRH_CStat_R	OF_PerPerm	
CoupGir_LFRF_CStat		CoupGir_LFRF_CStat_R	OF_PerAvSpeed	
CoupGir_RHLH_CStat		CoupGir_RHLH_CStat_R	OF_CenDist	
CoupGir_RFLF_CStat		CoupGir_RFLF_CStat_R	OF_CenRest	
CoupIpsi_RFRH_CStat		CoupIpsi_RFRH_CStat_R	OF_CenPerm	
CoupIpsi_LFLH_CStat		CoupIpsi_LFLH_CStat_R	OF_CenAvSpeed	
CoupIpsi_RHRF_CStat		CoupIpsi_RHRF_CStat_R	OF_NbEntriesCen	
CoupIpsi_LHLF_CStat		CoupIpsi_LHLF_CStat_R		

RF - right front
RH - right hind
LF - left front
LH - left hind
BOS - base of support

FP - front paws
HP - hind paws
RP - right paws
LP - left paws

Ph - Phase dispersion
Coup - Couplings
Dig - diagonal pairs
Ipsi - ipsilateral pairs
Gir - girdle pairs

CA - Cruciate (RF-LF-RH-LH)
CB - Cruciate (LF-RF-LH-RH)
AA -Alternate (RF-RH-LF-LH)
AB - Alternate (LF-RH-RF-RH)
RA - Rotate (RF-LF-LH-RH)
RB - Rotate (LF-RF-RH-LH)

Cstat - Circular Statistics
CStat_R -Strength of directedness and therefore a measure of Variation

dist- distance
rear- rearings
%DistCen- Distance travelled in the centre (%)
%TiCen- Time in Centre (%)
WholeRest- Resting time in whole arena
WholeAvSpeed- average speed in whole arena
PerDist- Distance Travelled in Periphery
PerRest- Resting time in Periphery
PerPerm- Permanence time in Periphery
PerAvSpeed- Average Speed in periphery
CenDist- Distance Travelled in Center
CenRest- Resting time in Center
CenPerm- Permanence time in Center
CenAvSpeed- Average Speed in Center
NbEntriesCen-Number of Entries into Center

TimeTurn-time to turn
TimeComplete- total time
TimeDown- time needed from turning to fully descend

Table S3: Structure matrix for CatWalk PCA

For abbreviations used see Table S2

Structure matrix

	Component									
	1	2	3	4	5	6	7	8	9	10
HP_PrintArea	**.965**	.005	-.006	.008	.170	.031	-.244	.083	-.011	-.306
FP_MaxContactArea	**.963**	-.164	-.017	.094	.106	.053	-.048	.253	-.015	-.290
FP_PrintArea	**.960**	-.179	-.019	.141	.086	.056	-.037	.296	-.020	-.283
HP_MaxContactArea	**.958**	-.061	-.004	.008	.189	.039	-.233	.104	-.017	-.305
RF_PrintArea	**.958**	-.167	-.013	.135	.089	.036	-.044	.294	-.017	-.284
RF_MaxContactArea	**.957**	-.143	-.016	.083	.104	.024	-.057	.246	-.010	-.293
LF_MaxContactArea	**.954**	-.182	-.018	.105	.108	.080	-.038	.256	-.018	-.284
LF_PrintArea	**.952**	-.189	-.025	.147	.083	.076	-.028	.294	-.022	-.280
LH_PrintArea	**.950**	-.040	-.050	.023	.162	.095	-.214	.091	-.049	-.290
LH_MaxContactArea	**.941**	-.095	-.047	.016	.181	.084	-.207	.108	-.053	-.288
RH_MaxContactArea	**.931**	-.019	.036	.000	.190	-.014	-.251	.094	.025	-.312
RH_PrintArea	**.931**	.056	.033	-.010	.171	-.043	-.266	.070	.033	-.312
FP_PrintLength	**.899**	-.115	.012	.149	.027	.065	-.001	.438	-.080	-.270
RF_PrintLength	**.879**	-.110	.020	.145	.036	.041	-.012	.429	-.071	-.259
HP_PrintLength	**.879**	.247	.038	-.049	.066	.001	-.155	.040	-.046	-.303
LF_PrintLength	**.878**	-.116	.003	.148	.018	.088	.013	.427	-.085	-.270
FP_PrintWidth	**.874**	.102	.048	-.021	.015	.044	-.082	.365	-.105	-.324
LH_PrintLength	**.858**	.197	-.014	-.014	.051	.085	-.135	.057	-.085	-.277
LF_PrintWidth	**.845**	.056	.047	.002	-.002	.104	-.056	.367	-.127	-.307
RF_PrintWidth	**.841**	.139	.044	-.041	.032	-.018	-.099	.337	-.074	-.321
RH_PrintLength	**.814**	.270	.076	-.079	.075	-.080	-.160	.022	.000	-.303
LH_PrintWidth	**.787**	.329	-.003	-.116	.084	.053	-.290	-.072	-.107	-.267
HP_PrintWidth	**.765**	.432	.058	-.169	.078	.002	-.335	-.098	-.048	-.280
RH_PrintWidth	**.653**	.489	.103	-.206	.064	-.053	-.345	-.112	.023	-.266
HP_Stand	-.049	**.894**	-.188	.287	-.304	-.061	.038	-.308	.229	.127
RH_Stand	-.062	**.891**	-.160	.261	-.294	-.112	.027	-.299	.256	.112
HP_DutyCycle	-.124	**.888**	.009	-.312	-.141	-.063	-.050	-.476	.007	-.114
LH_Stand	-.034	**.874**	-.217	.306	-.305	-.008	.047	-.302	.198	.137
RH_DutyCycle	-.142	**.854**	.014	-.320	-.121	-.193	-.069	-.444	.044	-.108
LH_DutyCycle	-.090	**.851**	-.004	-.277	-.151	.088	-.024	-.471	-.034	-.116
FP_DutyCycle	.012	**.815**	.021	-.113	-.595	-.009	.103	-.192	-.042	-.137
RF_DutyCycle	.024	**.789**	.033	-.105	-.575	-.098	.097	-.185	-.020	-.141
FP_Stand	.008	**.789**	-.242	.428	-.418	-.033	.090	-.148	.263	.102
LF_DutyCycle	-.002	**.784**	.008	-.112	-.573	.081	.103	-.187	-.060	-.126
LF_Stand	.009	**.783**	-.244	.428	-.417	-.009	.095	-.153	.250	.098
RF_Stand	.012	**.781**	-.236	.428	-.409	-.056	.085	-.141	.271	.099
Support_Diagonal	.309	**-.751**	.208	.110	.438	.095	.027	.448	-.245	-.219
Run_Duration	-.195	**.613**	-.254	.322	-.265	-.037	.089	-.363	.259	.206
Support_Single	-.064	**-.527**	-.008	.303	.060	.072	.214	.416	-.040	.071
Support_Zero	-.160	**-.358**	-.022	.186	-.035	.045	.064	.250	-.045	.073
Support_Girdle	-.112	**-.340**	-.037	.208	-.057	.025	.212	.201	.066	.085

CoupIpsi_RFRH_CStat_R	.000	-.052	**.945**	-.079	-.048	.050	.092	.145	-.502	-.188
PhIps_RFRH_CStat_R	.045	-.058	**.932**	-.057	-.047	.056	.098	.172	-.501	-.196
PhDig_LFRH_CStat_R	-.040	.034	**.918**	-.053	-.165	.034	.071	.105	-.432	-.096
CoupDig_LFRH_CStat_R	-.035	.059	**.918**	-.049	-.135	.036	.079	.109	-.444	-.130
CoupIpsi_RHRF_CStat_R	-.003	-.070	**.915**	-.017	-.037	.045	.049	.176	-.437	-.173
PhGird_LHRH_CStat_R	-.053	-.021	**.863**	-.035	-.095	.057	.054	.030	-.712	-.001
CoupGir_RHLH_CStat_R	-.107	-.021	**.862**	-.030	-.106	.033	.034	.021	-.685	.047
CoupGir_LHRH_CStat_R	-.086	-.031	**.844**	-.046	-.101	.039	.037	.019	-.734	.027
StepSequence_RegularityIndex	-.180	-.053	**.615**	-.075	-.118	.041	-.050	-.107	-.482	.054
CoupDig_RHLF_CStat_R	.191	-.122	**.600**	-.097	.016	.106	.009	.304	-.556	-.472
StepSequence_RA	.006	.100	**-.451**	.071	-.008	-.012	.090	.141	.174	.073
HP_SwingSpeed	.077	.073	.195	**-.908**	.151	-.028	-.237	.224	-.230	-.089
RH_SwingSpeed	.042	.103	.198	**-.880**	.142	-.150	-.241	.205	-.183	-.085
LH_SwingSpeed	.111	.034	.168	**-.875**	.151	.103	-.219	.230	-.262	-.092
FP_Swing	.166	.197	-.030	**.839**	.180	.058	.076	.108	.050	-.199
RF_Swing	.145	.196	-.051	**.827**	.172	.125	.085	.105	.044	-.177
HP_Swing	.199	-.279	.100	**.817**	-.100	.115	.185	.440	-.100	-.063
LF_Swing	.177	.184	-.012	**.813**	.167	-.019	.071	.097	.056	-.217
LH_Swing	.195	-.242	.056	**.801**	-.070	-.026	.182	.444	-.101	-.078
RH_Swing	.189	-.276	.114	**.749**	-.108	.213	.174	.375	-.077	-.049
FP_SwingSpeed	.242	-.318	.194	**-.714**	.023	.022	-.182	.593	-.276	-.097
LF_SwingSpeed	.231	-.315	.181	**-.706**	.026	.068	-.173	.584	-.280	-.093
RF_SwingSpeed	.246	-.313	.204	**-.704**	.023	-.023	-.184	.584	-.263	-.103
BOS_HP	.299	.077	-.102	**.339**	.209	.031	.259	-.108	.156	-.242
StepSequence_AB	-.136	.290	.074	.056	**-.864**	-.088	-.092	-.164	-.026	.083
PhIps_RFRH_CStat	-.038	.404	.158	.016	**-.827**	.102	.113	.121	-.142	-.066
CoupIpsi_RFRH_CStat	-.038	.411	.152	.021	**-.823**	.093	.117	.126	-.137	-.046
StepSequence_AA	.024	-.132	-.088	-.082	**.816**	.025	-.051	.113	.028	-.069
CoupIpsi_LFLH_CStat	-.059	.420	.121	-.013	**-.798**	-.083	.052	.100	-.252	-.037
PhIps_LFLH_CStat	-.058	.414	.101	-.013	**-.798**	-.092	.071	.092	-.247	-.049
CoupIpsi_RHRF_CStat	.171	-.185	.101	.011	**.794**	.001	.112	.137	-.218	-.369
CoupIpsi_LHLF_CStat	.214	-.229	.173	.007	**.784**	.203	.148	.150	-.152	-.418
CoupDig_LFRH_CStat	.141	-.058	-.117	.013	**.708**	-.417	.033	.137	.015	-.175
CoupDig_RHLF_CStat	-.075	.091	.105	.047	**-.704**	.425	.008	-.115	-.061	.084
PhDig_LFRH_CStat	.111	-.030	-.113	.011	**.703**	-.404	.027	.101	.021	-.170
CoupDig_LHRF_CStat	-.093	.155	.121	.003	**-.671**	-.512	-.014	-.050	-.019	-.047
CoupDig_RFLH_CStat	.138	-.121	-.121	.054	**.663**	.544	.049	.085	.035	-.044
PhDig_RFLH_CStat	.101	-.100	-.104	.054	**.657**	.538	.044	.062	.018	-.050
BOS_FP	-.185	.420	-.108	-.125	**-.469**	-.044	.047	-.374	.125	.067
CoupGir_LHRH_CStat	.129	.068	.264	.034	-.038	**.785**	.183	.145	-.200	-.329
PhGird_LHRH_CStat	.128	.048	.270	.028	-.030	**.783**	.162	.143	-.231	-.339
StepSequence_CA	.140	-.240	.008	.038	.191	**.750**	.113	.043	-.037	.024
StepSequence_CB	.056	-.085	.078	-.054	.209	**-.709**	.091	.064	-.008	-.114
CoupGir_LFRF_CStat	.083	.145	.156	.066	-.003	**.616**	.212	.200	-.389	-.367
PhGird_LFRF_CStat	.089	.150	.160	.062	.005	**.615**	.199	.197	-.391	-.369
CoupGir_RHLH_CStat	.104	.250	.038	.008	.041	**-.605**	.171	.233	-.251	-.313
CoupGir_RFLF_CStat	.161	.209	.272	.011	.004	**-.438**	.191	.291	-.285	-.429
HP_MaxContactAt	-.304	.010	.121	.059	.091	.062	**.812**	.024	-.144	.128

LH_MaxContactAt	-.309	.021	.095	.047	.066	.095	**.723**	.030	-.109	.133
FP_MaxContactAt	.073	.021	-.039	.129	-.079	.042	**.719**	-.065	.054	-.394
RH_MaxContactAt	-.226	-.005	.122	.057	.097	.010	**.718**	.012	-.145	.090
RF_MaxContactAt	.201	-.083	.014	.191	-.022	-.007	**.620**	.036	-.005	-.330
LF_MaxContactAt	-.080	.120	-.082	.022	-.112	.078	**.583**	-.144	.097	-.330
RF_StrideLength	.552	-.466	.179	-.072	.284	.078	-.159	**.788**	-.263	-.276
FP_StrideLength	.556	-.469	.181	-.076	.282	.077	-.159	**.787**	-.266	-.279
HP_StrideLength	.516	-.462	.323	-.115	.240	.075	-.152	**.783**	-.373	-.188
LF_StrideLength	.555	-.471	.182	-.080	.282	.077	-.155	**.781**	-.267	-.283
RH_StrideLength	.504	-.453	.380	-.110	.231	.068	-.146	**.780**	-.313	-.198
LH_StrideLength	.519	-.460	.243	-.117	.246	.081	-.155	**.768**	-.414	-.182
Support_Lateral	.106	-.306	-.100	.324	.173	.116	.215	**.379**	.038	.067
CoupIpsi_LFLH_CStat_R	.044	-.047	.540	-.100	-.020	.073	.062	.146	**-.926**	-.117
PhIps_LFLH_CStat_R	.088	-.089	.533	-.091	-.017	.080	.069	.174	**-.921**	-.145
CoupDig_RFLH_CStat_R	.015	-.022	.501	-.079	-.070	.071	.019	.104	**-.919**	-.114
PhDig_RFLH_CStat_R	.020	-.016	.483	-.074	-.115	.040	.013	.105	**-.914**	-.077
CoupIpsi_LHLF_CStat_R	.074	-.136	.545	-.087	.017	.080	.034	.166	**-.889**	-.177
CoupDig_LHRF_CStat_R	.152	-.164	.497	-.087	.056	.073	-.039	.266	**-.718**	-.481
CoupGir_LFRF_CStat_R	.273	-.219	.429	-.072	.150	.136	.054	.375	-.564	**-.677**
PhGird_LFRF_CStat_R	.276	-.224	.406	-.048	.120	.159	.095	.385	-.549	**-.659**
CoupGir_RFLF_CStat_R	.260	-.229	.449	-.114	.146	.138	.075	.344	-.596	**-.638**
PrintPos_LP	.325	-.029	-.002	.250	.122	.084	.130	-.083	.219	**-.613**
PrintPos_RP	.341	-.033	-.102	.219	.143	.034	.106	-.108	.073	**-.586**

Table S4: Component correlation matrix for CatWalk PCA

Component correlation matrix

Component	1	2	3	4	5	6	7	8	9	10
1	1.000	-.009	.012	.033	.111	.040	-.133	.198	-.034	-.306
2	-.009	1.000	-.060	-.016	-.206	-.053	.000	-.252	.047	-.050
3	.012	-.060	1.000	-.092	-.046	.036	.001	.088	-.490	-.138
4	.033	-.016	-.092	1.000	-.027	.043	.183	.029	.118	.000
5	.111	-.206	-.046	-.027	1.000	.010	-.023	.050	.012	-.094
6	.040	-.053	.036	.043	.010	1.000	.053	.028	-.070	-.029
7	-.133	.000	.001	.183	-.023	.053	1.000	.040	-.026	-.120
8	.198	-.252	.088	.029	.050	.028	.040	1.000	-.210	-.090
9	-.034	.047	-.490	.118	.012	-.070	-.026	-.210	1.000	.153
10	-.306	-.050	-.138	.000	-.094	-.029	-.120	-.090	.153	1.000

1 *Table S5: Structure matrix for CatWalk- Open Field- Vertical Pole PCA*

2 *For abbreviations see Table S2*

Structure matrix

	Component												
	1	2	3	4	5	6	7	8	9	10	11	12	13
OF_WholeAvSpeed	.950	-.162	.163	-.037	-.197	.046	-.631	.105	-.236	.263	.085	-.076	-.097
OF_PerAvSpeed	.944	-.154	.177	-.029	-.187	.042	-.627	.108	-.230	.265	.082	-.069	-.094
OF_PerDist	.934	-.173	.160	-.054	-.172	.044	-.444	.114	-.228	.253	.068	-.044	-.129
OF_distTot	.932	-.169	.161	-.044	-.181	.045	-.627	.112	-.232	.262	.074	-.058	-.117
OF_RearTot	.926	-.094	.169	-.038	-.198	.053	-.597	.064	-.240	.245	.089	-.120	.040
OF_dist10	.903	-.143	.172	-.046	-.169	.048	-.604	.101	-.206	.256	.076	-.055	-.130
OF_Rear10	.884	-.087	.169	-.033	-.187	.057	-.575	.052	-.220	.235	.091	-.114	.041
OF_dist15	.883	-.128	.171	-.062	-.164	.032	-.585	.127	-.193	.221	.078	-.033	-.111
OF_dist20	.870	-.110	.149	-.083	-.150	.035	-.557	.104	-.183	.202	.051	-.043	-.083
OF_Rear15	.864	-.073	.145	-.041	-.162	.054	-.571	.060	-.214	.214	.090	-.110	.057
OF_Dist05	.858	-.232	.122	.008	-.193	.045	-.608	.097	-.270	.291	.063	-.069	-.106
OF_Rear20	.852	-.035	.113	-.063	-.165	.025	-.551	.043	-.198	.211	.071	-.113	.047
OF_Rear05	.842	-.148	.197	-.007	-.217	.063	-.523	.084	-.253	.245	.077	-.100	.012
OF_WholeRest	.752	-.119	.060	.009	-.230	.044	-.480	-.029	-.214	.176	.143	-.200	.150
OF_PerRest	.743	-.145	.074	.003	-.206	.053	-.360	-.029	-.218	.161	.144	-.184	.150
OF_CenAvSpeed	.452	-.018	.007	-.102	-.237	.074	.131	.069	-.071	.112	-.001	-.097	-.211
HP_PrintArea	-.153	.962	.024	-.180	-.012	-.010	.118	-.061	.102	-.127	-.071	-.046	-.376
FP_MaxContactArea	-.074	.960	.016	-.105	-.201	.029	.066	.050	.103	-.153	-.061	-.089	-.360
FP_PrintArea	-.040	.957	.052	-.079	-.252	.039	.048	.050	.099	-.153	-.055	-.099	-.365
RF_PrintArea	-.042	.955	.053	-.079	-.244	.033	.040	.039	.104	-.149	-.087	-.090	-.357
HP_MaxContactArea	-.150	.954	-.015	-.188	-.056	-.010	.118	-.050	.086	-.123	-.060	-.069	-.394
RF_MaxContactArea	-.085	.953	.014	-.101	-.186	.011	.063	.039	.108	-.150	-.093	-.077	-.352
LF_MaxContactArea	-.064	.951	.019	-.107	-.211	.047	.069	.060	.097	-.154	-.029	-.098	-.361
LF_PrintArea	-.038	.949	.051	-.078	-.256	.046	.056	.061	.094	-.157	-.023	-.107	-.368
LH_PrintArea	-.145	.944	.021	-.173	-.045	.045	.124	-.037	.047	-.189	-.026	-.064	-.366
LH_MaxContactArea	-.150	.934	-.020	-.182	-.080	.028	.131	-.031	.035	-.183	-.026	-.083	-.381
RH_PrintArea	-.157	.930	.025	-.180	.026	-.072	.107	-.084	.161	-.063	-.118	-.024	-.368
RH_MaxContactArea	-.144	.928	-.009	-.187	-.026	-.053	.099	-.070	.141	-.061	-.095	-.050	-.390
FP_PrintLength	.002	.910	.083	-.036	-.321	.054	.032	-.008	.091	-.131	-.044	-.094	-.342
HP_PrintLength	-.183	.905	.074	-.110	.136	-.009	.127	-.054	.135	-.095	-.092	.044	-.229
LF_PrintLength	.008	.896	.065	-.032	-.310	.062	.038	.006	.078	-.127	-.010	-.099	-.352
FP_PrintWidth	-.208	.891	-.035	-.014	-.090	.048	.116	-.143	.117	-.117	-.093	.000	-.287
RF_PrintLength	-.003	.888	.098	-.038	-.318	.045	.026	-.022	.100	-.130	-.075	-.086	-.319
LH_PrintLength	-.167	.881	.090	-.095	.073	.073	.118	-.043	.070	-.165	-.032	.026	-.222
LF_PrintWidth	-.172	.862	-.036	.003	-.125	.074	.106	-.118	.075	-.102	.010	-.038	-.297
RF_PrintWidth	-.229	.849	-.030	-.029	-.047	.020	.118	-.156	.151	-.124	-.187	.038	-.254
RH_PrintLength	-.183	.842	.051	-.116	.183	-.089	.124	-.059	.191	-.031	-.142	.055	-.217
LH_PrintWidth	-.276	.802	.044	-.107	.284	.014	.197	-.199	.049	-.169	-.064	.093	-.219
HP_PrintWidth	-.327	.796	.032	-.107	.376	-.028	.206	-.260	.143	-.077	-.102	.136	-.212
RH_PrintWidth	-.342	.682	.013	-.093	.422	-.071	.190	-.291	.230	.014	-.129	.162	-.182
FP_Swing	.183	.150	.883	-.016	-.077	.042	-.078	-.047	.218	-.131	-.026	.009	-.018
RF_Swing	.182	.138	.864	-.008	-.074	.027	-.078	-.049	.220	-.130	.135	.012	-.022
FP_SwingSpeed	-.179	.190	-.863	-.114	-.360	-.003	.038	-.343	-.162	.093	-.071	-.042	-.227

LF_SwingSpeed	-.178	.184	-.858	-.114	-.351	-.017	.034	-.339	-.161	.090	.032	-.046	-.225
LF_Swing	.186	.158	.857	-.022	-.074	.054	-.070	-.031	.204	-.124	-.178	.002	-.027
HP_SwingSpeed	-.326	.046	-.856	-.253	.123	-.062	.140	-.383	-.118	.116	-.100	.098	-.036
RF_SwingSpeed	-.177	.192	-.846	-.112	-.359	.012	.039	-.336	-.159	.094	-.173	-.038	-.222
LH_SwingSpeed	-.295	.078	-.840	-.260	.093	.090	.127	-.365	-.168	.057	-.049	.072	-.045
RH_SwingSpeed	-.336	.017	-.816	-.229	.141	-.200	.145	-.372	-.060	.151	-.140	.113	-.032
LF_Stand	-.191	.057	.740	.441	.478	.087	.126	-.120	.359	-.142	-.014	.328	.300
FP_Stand	-.195	.051	.733	.442	.469	.098	.129	-.134	.370	-.153	-.083	.322	.314
RF_Stand	-.196	.047	.721	.436	.453	.103	.127	-.140	.373	-.161	-.139	.317	.318
LH_Stand	-.230	-.006	.660	.325	.638	.108	.148	-.166	.272	-.151	-.072	.345	.394
Run_Duration	-.111	-.137	.655	.338	.553	.025	.103	-.002	.275	-.133	-.036	.280	.342
LH_Swing	.363	.139	.652	.210	-.534	-.040	-.218	.050	-.072	.045	.003	-.121	-.146
HP_Swing	.405	.149	.651	.219	-.555	.138	-.260	.077	-.071	.103	.055	-.128	-.151
RH_Swing	.397	.151	.594	.200	-.509	.260	-.259	.108	-.056	.129	.080	-.127	-.141
BOS_HP	.038	.274	.365	-.128	.092	.067	-.003	.338	.218	-.176	-.113	.074	-.117
CoupIpsi_LFLH_CStat	-.142	-.105	.209	.901	.096	-.134	.037	-.089	-.018	.096	.118	.261	.147
CoupIpsi_LHLF_CStat	.156	.145	-.180	-.883	-.099	.162	-.046	.144	-.017	-.059	-.098	-.267	-.191
PhIps_LFLH_CStat	-.127	-.093	.168	.880	.082	-.159	.034	-.043	-.083	.136	.132	.235	.094
StepSequence_AB	.003	-.126	.291	.864	.142	.067	-.020	.032	-.036	.119	-.021	.174	.065
CoupIpsi_RFRH_CStat	-.097	-.080	.225	.855	.092	.401	.039	-.043	.006	.108	-.133	.214	.162
PhIps_RFRH_CStat	-.098	-.075	.222	.851	.092	.404	.043	-.036	-.008	.114	-.126	.210	.155
CoupIpsi_RHRF_CStat	.101	.131	-.183	-.824	-.060	-.379	-.066	.094	-.020	-.032	.136	-.201	-.179
StepSequence_AA	-.061	.019	-.195	-.816	-.066	-.080	.039	-.102	.031	-.090	-.038	-.103	-.098
CoupDig_RFLH_CStat	.029	.093	-.016	-.706	-.066	.348	-.003	-.022	.060	-.175	.359	-.133	-.073
PhDig_RFLH_CStat	.049	.065	-.011	-.705	-.061	.331	-.002	-.033	.032	-.150	.349	-.139	-.070
CoupDig_LHRF_CStat	-.029	-.057	.086	.705	.066	-.324	.019	.051	.032	.135	-.355	.127	.035
CoupDig_LFRH_CStat	-.037	.166	-.050	-.668	-.070	-.459	.028	.017	.082	-.137	-.338	-.049	-.104
CoupDig_RHLF_CStat	.050	-.101	.117	.668	.072	.429	-.028	.034	-.073	.086	.355	.048	.062
PhDig_LFRH_CStat	-.050	.131	-.042	-.665	-.046	-.447	.040	.033	.078	-.128	-.334	-.045	-.077
HP_DutyCycle	-.391	-.109	.077	.071	.931	-.044	.234	-.129	.166	-.015	-.087	.358	.321
LH_DutyCycle	-.335	-.080	.103	.066	.893	.137	.200	-.100	.100	-.058	-.021	.334	.314
RH_DutyCycle	-.408	-.122	.048	.069	.880	-.200	.243	-.143	.211	.015	-.137	.346	.295
Support_Diagonal	.320	.236	-.290	-.354	-.774	.002	-.228	.081	-.299	.117	.063	-.388	-.418
FP_DutyCycle	-.323	.019	.257	.585	.719	.092	.178	-.009	.197	.001	-.069	.388	.286
RF_StrideLength	-.018	.508	-.339	-.267	-.699	.028	-.041	-.334	-.096	.006	-.087	-.175	-.471
FP_StrideLength	-.014	.511	-.350	-.270	-.699	.025	-.042	-.329	-.101	.010	-.081	-.176	-.478
LF_StrideLength	-.011	.511	-.358	-.271	-.695	.022	-.044	-.322	-.105	.014	-.074	-.177	-.482
HP_StrideLength	.018	.449	-.395	-.218	-.685	.022	-.075	-.384	-.270	.211	-.071	-.150	-.393
RF_DutyCycle	-.306	.029	.255	.552	.684	.108	.171	-.007	.191	-.004	-.292	.367	.279
LF_DutyCycle	-.308	.009	.234	.559	.682	.067	.167	-.010	.185	.005	.165	.371	.264
LH_StrideLength	.007	.450	-.390	-.224	-.675	.016	-.069	-.385	-.337	.087	-.051	-.153	-.389
RH_StrideLength	.022	.437	-.385	-.208	-.671	.028	-.073	-.364	-.178	.303	-.090	-.142	-.392
HP_Stand	-.251	-.018	.647	.329	.652	.046	.162	-.174	.321	-.112	-.091	.355	.397
RH_Stand	-.265	-.027	.618	.321	.648	-.013	.172	-.182	.362	-.078	-.110	.357	.387
Support_Single	.340	-.135	.093	-.038	-.598	.023	-.133	.151	-.012	-.067	.082	-.242	-.068
Support_Lateral	.269	.141	.146	-.238	-.532	.067	-.153	.079	.003	-.083	.102	-.091	-.046
BOS_FP	-.225	-.116	.124	.390	.434	.026	.173	.186	.122	-.046	-.016	.264	.197
Support_Zero	.166	-.187	.055	.049	-.382	.038	-.075	.051	-.060	-.048	.069	-.155	-.054
Support_Girdle	.222	-.161	.062	.077	-.330	-.028	-.136	.187	.016	-.006	.111	-.186	-.033
CoupGir_LHRH_CStat	.087	.039	.045	-.014	-.012	.937	-.025	.090	.019	.056	.312	-.035	-.019

PhGird_LHRH_CStat	.084	.040	.041	-.017	-.025	.937	-.024	.083	-.011	.038	.323	-.047	-.033
CoupGir_RHLH_CStat	-.079	.080	-.019	-.032	.023	-.879	.010	.008	-.024	.001	-.284	.023	-.058
StepSequence_CB	-.042	.075	-.127	-.089	-.013	-.631	.026	.070	.038	.081	-.571	-.023	.011
OF_CenPerm	.525	-.106	.075	-.003	-.105	.000	-.988	.064	-.170	.199	.067	-.080	-.011
OF_PerPerm	-.525	.106	-.075	.003	.105	.000	.988	-.064	.170	-.199	-.067	.080	.011
OF_PerTiCenTot	.525	-.107	.075	-.003	-.106	.000	-.988	.064	-.170	.199	.067	-.080	-.011
OF_PerDistCenTot	.512	-.119	.005	-.070	-.146	.017	-.966	.063	-.183	.176	.068	-.118	-.042
OF_PerDistCen15	.480	-.106	.038	-.090	-.117	.012	-.885	.045	-.178	.129	.075	-.102	-.035
OF_PerTiCen15	.472	-.092	.085	-.030	-.090	.003	-.883	.047	-.160	.144	.072	-.076	-.008
OF_PerTiCen10	.399	-.118	.025	.003	-.091	.006	-.877	.031	-.118	.174	.066	-.103	-.022
OF_PerDistCen10	.425	-.151	-.037	-.066	-.117	.013	-.876	.041	-.149	.168	.067	-.120	-.043
OF_CenDist	.770	-.133	.136	-.019	-.167	.039	-.869	.089	-.200	.236	.074	-.075	-.075
OF_PerDistCen20	.510	-.065	.026	-.099	-.095	-.015	-.846	.093	-.128	.160	.034	-.043	.009
OF_PerTiCen20	.493	-.051	.081	-.039	-.065	-.036	-.837	.089	-.112	.160	.043	-.013	.029
OF_NbEntriesCen	.802	-.137	.119	-.010	-.178	.037	-.824	.094	-.214	.244	.078	-.076	-.082
OF_PerDistCen05	.429	-.142	-.002	.038	-.172	.031	-.815	.045	-.219	.197	.053	-.130	-.047
OF_PerTiCen05	.439	-.121	.064	.074	-.131	.039	-.811	.051	-.204	.216	.047	-.086	-.041
OF_CenRest	.542	-.008	-.002	.023	-.230	.004	-.683	-.021	-.135	.164	.092	-.185	.104
FP_MaxContactAt	-.074	.057	.088	.059	.040	.018	.018	.715	.146	-.070	-.022	.140	.095
RF_MaxContactAt	.011	.163	.134	.019	-.076	-.077	-.022	.612	.086	-.026	-.015	.074	.017
LF_MaxContactAt	-.134	-.072	.009	.079	.145	.110	.052	.561	.158	-.092	-.024	.156	.141
PrintPositions_LP	.189	.318	.276	-.108	-.010	.124	-.133	.448	.306	-.050	-.048	-.065	-.438
PhIps_LFLH_CStat_R	.216	-.070	-.163	.019	-.077	.001	-.194	-.057	-.940	.385	.075	-.059	-.044
CoupIpsi_LFLH_CStat_R	.197	-.101	-.157	.032	-.030	-.001	-.184	-.082	-.936	.388	.072	-.014	-.011
CoupDig_RFLH_CStat_R	.185	-.130	-.110	.094	-.024	-.007	-.165	-.103	-.928	.340	.074	-.017	.038
PhDig_RFLH_CStat_R	.185	-.126	-.099	.128	-.032	-.017	-.165	-.103	-.925	.337	.051	-.031	.029
CoupIpsi_LHLF_CStat_R	.197	-.075	-.145	-.008	-.100	-.009	-.194	-.073	-.920	.377	.086	-.088	-.091
CoupDig_LHRF_CStat_R	.121	.024	-.209	-.051	-.266	-.058	-.096	-.029	-.632	.285	.051	-.222	-.366
StepSequence_RB	-.141	.031	.142	.090	.025	.016	.022	-.127	.367	-.288	-.030	.033	.101
CoupIpsi_RFRH_CStat_R	.246	-.155	-.100	.091	-.018	.027	-.206	-.006	-.368	.960	-.007	.002	.002
CoupDig_LFRH_CStat_R	.212	-.164	-.048	.144	.020	.034	-.173	-.021	-.325	.949	.000	.033	.036
PhIps_RFRH_CStat_R	.254	-.115	-.087	.090	-.040	.041	-.198	-.003	-.369	.947	-.008	-.002	-.015
PhDig_LFRH_CStat_R	.217	-.172	-.059	.166	.019	.039	-.175	-.017	-.338	.945	.013	.034	.025
CoupIpsi_RHRF_CStat_R	.228	-.132	-.036	.087	-.065	.037	-.194	-.052	-.306	.916	-.013	-.031	-.026
CoupGir_RHLH_CStat_R	.277	-.220	-.024	.129	.063	.013	-.225	-.084	-.674	.837	.037	.019	.114
PhGird_LHRH_CStat_R	.289	-.194	-.021	.123	.059	.035	-.225	-.057	-.702	.826	.048	.008	.075
CoupGir_LHRH_CStat_R	.284	-.223	-.036	.131	.053	.011	-.232	-.080	-.735	.799	.052	.001	.097
StepSequence_RegularityIndex	.150	-.213	-.040	.103	.045	.027	-.157	-.035	-.550	.619	.075	.022	.041
StepSequence_RA	-.108	.079	.028	-.056	-.061	.050	.104	-.040	.131	-.532	.059	-.009	.034
CoupDig_RHLF_CStat_R	.081	.084	-.226	-.056	-.270	.030	-.117	.025	-.398	.480	.070	-.166	-.341
CoupGir_LFRF_CStat	.070	-.002	.030	-.004	-.029	.287	-.078	.030	-.098	.011	.936	-.003	-.051
PhGird_LFRF_CStat	.056	.005	.041	-.002	-.018	.295	-.057	.010	-.085	.000	.924	.015	-.032
CoupGir_RFLF_CStat	-.008	.090	.039	-.024	-.008	-.243	.010	.066	.023	.044	-.913	-.057	-.055
StepSequence_CA	.137	.088	-.078	-.154	-.109	.582	-.061	.035	-.074	-.050	.604	-.107	-.003
Pole_TimeComplete	.063	-.105	.010	.113	.041	-.001	-.005	.130	.000	.038	.040	.909	.061
Pole_TimeTurn	.126	-.066	.053	.105	.070	.011	-.065	.171	.005	.055	.037	.829	.037
Pole_TimeDown	-.055	-.126	-.052	.112	-.016	-.019	.087	.033	-.006	.002	.040	.798	.080
HP_MaxContactAt	.153	-.461	-.014	-.059	-.027	.039	-.101	.421	-.101	.097	.040	.022	.678
CoupGir_LFRF_CStat_R	.140	.224	-.307	-.219	-.415	.046	-.133	.136	-.350	.208	.044	-.218	-.648
PhGird_LFRF_CStat_R	.164	.220	-.265	-.190	-.421	.061	-.156	.171	-.341	.201	.060	-.227	-.634

RH_MaxContactAt	.119	-.369	-.008	-.050	-.026	-.004	-.052	.386	-.086	.078	.000	.022	.616
CoupGir_RFLF_CStat_R	.138	.178	-.353	-.241	-.367	.008	-.161	.178	-.410	.234	.050	-.212	-.589
LH_MaxContactAt	.150	-.445	-.017	-.056	-.023	.070	-.122	.362	-.089	.093	.065	.017	.587
PrintPositions_RP	.209	.349	.249	-.132	-.011	.000	-.132	.431	.126	-.190	.019	-.084	-.437

Table S6: Component correlation matrix for CatWalk- Open Field- Vertical Pole PCA

Component correlation matrix

Component	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1.000	-.129	.155	-.065	-.235	.048	-.477	.124	-.186	.190	.084	-.128	-.057
2	-.129	1.000	.023	-.106	-.076	.021	.075	-.076	.094	-.101	-.076	-.019	-.323
3	.155	.023	1.000	.183	.115	.049	-.050	.105	.190	-.102	.010	.059	.104
4	-.065	-.106	.183	1.000	.102	.053	.016	-.013	.001	.087	-.010	.157	.129
5	-.235	-.076	.115	.102	1.000	-.017	.117	-.030	.126	-.006	-.041	.232	.223
6	.048	.021	.049	.053	-.017	1.000	-.004	.007	-.001	.000	.256	-.006	.008
7	-.477	.075	-.050	.016	.117	-.004	1.000	-.043	.152	-.166	-.059	.080	.013
8	.124	-.076	.105	-.013	-.030	.007	-.043	1.000	.048	-.034	.023	-.010	.004
9	-.186	.094	.190	.001	.126	-.001	.152	.048	1.000	-.378	-.086	.085	.071
10	.190	-.101	-.102	.087	-.006	.000	-.166	-.034	-.378	1.000	-.027	.011	-.037
11	.084	-.076	.010	-.010	-.041	.256	-.059	.023	-.086	-.027	1.000	-.043	.007
12	-.128	-.019	.059	.157	.232	-.006	.080	-.010	.085	.011	-.043	1.000	.135
13	-.057	-.323	.104	.129	.223	.008	.013	.004	.071	-.037	.007	.135	1.000

Table S7: Parameters of the IMPC data set

IMPC parameters
OF_Distance
OF_Rearings
OF_ArenaAverageSpeed
OF_ArenaRestingTime
OF_PeripheryDistance
OF_PeripheryAverageSpeed
OF_PeripheryRestingTime
OF_CenterAverageSpeed
OF_%CenterTime
OF_%CenterDistance
OF_CenterRestingTime
OF_CenterEntries
HC_Speed_mean

HC_Distance_mean
HC_Rearings_mean
GS_FP_bw_ratio
GS_4paws_bw_ratio
SHIRPA_activity
RR_PassiveRotations
RR_mean_Latency

OF-Open Field
HC- home cage
GS- Grip Strength
FP- front paws
bw- body weight
RR- Rotarod

Fig. S1: Scree plot for the IMPC data set

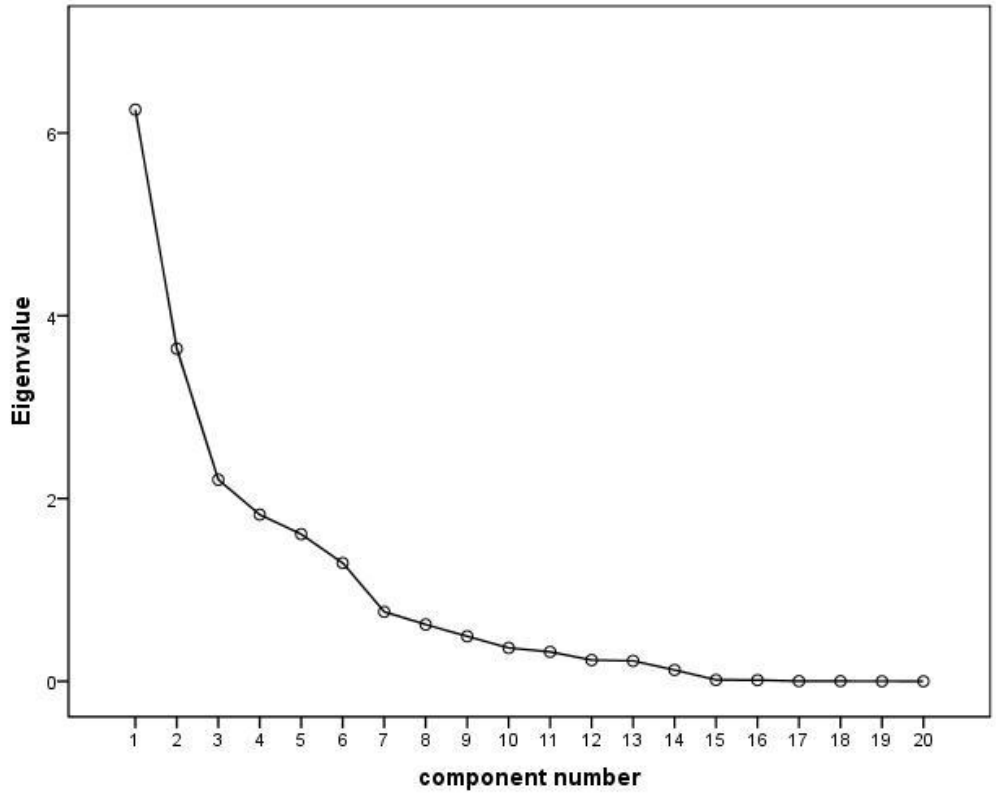


Table S8: Structure matrix for the IMPC data set

For abbreviations see Table S7

Structure matrix

	Component					
	1	2	3	4	5	6
OF_Distance	.982	.150	.181	.393	.214	.125
OF_ArenaAverageSpeed	.945	.247	.446	.418	.282	.158
OF_PeripheryAverageSpeed	.895	.328	.449	.416	.316	.162
OF_CenterAverageSpeed	.830	-.303	.293	.271	.099	.068
OF_PeripheryDistance	.776	-.574	.073	.212	.067	.050
OF_%CenterTime	.004	.986	.126	.128	.189	.081
OF_%CenterDistance	.058	.977	.117	.151	.166	.080
OF_CenterRestingTime	.029	.757	.612	.095	.258	.108
OF_CenterEntries	.586	.702	.133	.311	.203	.112
OF_ArenaRestingTime	.179	.368	.950	.191	.300	.156
OF_PeripheryRestingTime	.245	-.116	.894	.205	.224	.141
OF_Rearings	.380	.082	.832	.095	.135	.086
HC_Speed_mean	.257	.059	.087	.963	.154	.024
HC_Distance_mean	.257	.060	.088	.963	.154	.025
HC_Rearings_mean	.291	.131	.162	.671	.238	.123
GS_FP_bw_ratio	.116	.155	.181	.170	.933	.182
GS_4paws_bw_ratio	.132	.050	.157	.182	.924	.145
SHIRPA_activity	.364	.277	.271	.369	.386	.188
RR_PassiveRotations	-.009	-.012	-.004	-.064	.068	.858
RR_mean_Latency	.147	.096	.189	.172	.232	.847

Table S9: Component correlation matrix for the IMPC data set

Component correlation matrix

Component	1	2	3	4	5	6
1	1.000	.031	.238	.327	.171	.100
2	.031	1.000	.142	.116	.159	.068
3	.238	.142	1.000	.148	.221	.119
4	.327	.116	.148	1.000	.232	.080
5	.171	.159	.221	.232	1.000	.192
6	.100	.068	.119	.080	.192	1.000

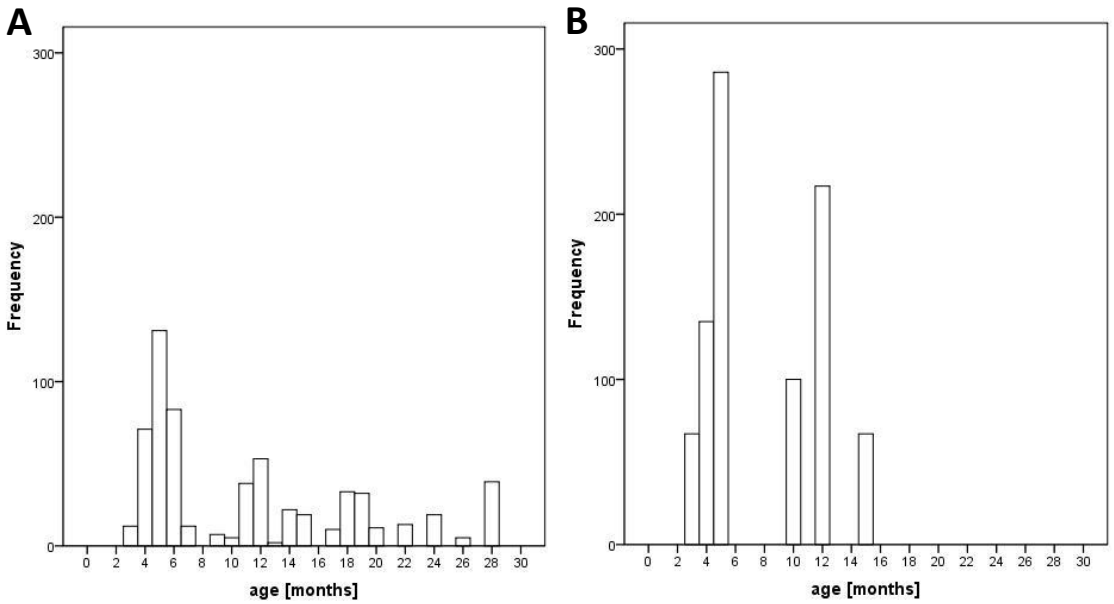


Fig. S2: Histogram for age in the 7.1 (A) and XT CW version (B)