



**HAL**  
open science

## Individuals have unique muscle activation signatures as revealed during gait and pedaling

François Hug, Clément Vogel, Kylie Tucker, Sylvain Dorel, Thibault Deschamps, Éric Le Carpentier, Lilian Lacourpaille

### ► To cite this version:

François Hug, Clément Vogel, Kylie Tucker, Sylvain Dorel, Thibault Deschamps, et al.. Individuals have unique muscle activation signatures as revealed during gait and pedaling. *Journal of Applied Physiology*, 2019, 127 (4), pp.1165-1174. 10.1152/japphysiol.01101.2018 . hal-03563380

**HAL Id: hal-03563380**

**<https://hal.science/hal-03563380v1>**

Submitted on 9 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## RESEARCH ARTICLE

# Individuals have unique muscle activation signatures as revealed during gait and pedaling

 François Hug,<sup>1,2,5</sup> Clément Vogel,<sup>1</sup> Kylie Tucker,<sup>2,3</sup> Sylvain Dorel,<sup>1</sup> Thibault Deschamps,<sup>1</sup> Éric Le Carpentier,<sup>4</sup> and Lilian Lacourpaille<sup>1</sup>

<sup>1</sup>Movement, Interactions, Performance, Nantes Université, EA 4334, Nantes, France; <sup>2</sup>National Health and Medical Research Council Centre of Clinical Research Excellence in Spinal Pain, Injury and Health, School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Queensland, Australia; <sup>3</sup>School of Biomedical Sciences, The University of Queensland, Brisbane, Queensland, Australia; <sup>4</sup>Laboratory LS2N, École Centrale de Nantes, Nantes, France; and <sup>5</sup>Institut Universitaire de France, Paris, France

Submitted 13 December 2018; accepted in final form 29 July 2019

**Hug F, Vogel C, Tucker K, Dorel S, Deschamps T, Le Carpentier É, Lacourpaille L.** Individuals have unique muscle activation signatures as revealed during gait and pedaling. *J Appl Physiol* 127: 1165–1174, 2019. First published August 8, 2019; doi:10.1152/jappphysiol.01101.2018.—Although it is known that the muscle activation patterns used to produce even simple movements can vary between individuals, these differences have not been considered to prove the existence of individual muscle activation strategies (or signatures). We used a machine learning approach (support vector machine) to test the hypothesis that each individual has unique muscle activation signatures. Eighty participants performed a series of pedaling and gait tasks, and 53 of these participants performed a second experimental session on a subsequent day. Myoelectrical activity was measured from eight muscles: vastus lateralis and medialis, rectus femoris, gastrocnemius lateralis and medialis, soleus, tibialis anterior, and biceps femoris-long head. The classification task involved separating data into training and testing sets. For the within-day classification, each pedaling/gait cycle was tested using the classifier, which had been trained on the remaining cycles. For the between-day classification, each cycle from *day 2* was tested using the classifier, which had been trained on the cycles from *day 1*. When considering all eight muscles, the activation profiles were assigned to the corresponding individuals with a classification rate of up to 99.28% (2,353/2,370 cycles) and 91.22% (1,341/1,470 cycles) for the within-day and between-day classification, respectively. When considering the within-day classification, a combination of two muscles was sufficient to obtain a classification rate >80% for both pedaling and gait. When considering between-day classification, a combination of four to five muscles was sufficient to obtain a classification rate >80% for pedaling and gait. These results demonstrate that strategies not only vary between individuals, as is often assumed, but are unique to each individual.

**NEW & NOTEWORTHY** We used a machine learning approach to test the uniqueness and robustness of muscle activation patterns. We considered that, if an algorithm can accurately identify participants, one can conclude that these participants exhibit discernible differences and thus have unique muscle activation signatures. Our results show that activation patterns not only vary between individuals, but are unique to each individual. Individual differences should, therefore, be considered relevant information for addressing fundamental questions about the control of movement.

Address for reprint requests and other correspondence: Prof. F. Hug, Laboratory Movement, Interactions, Performance (EA 4334), University of Nantes, 25 bis Blvd. Guy Mollet, 44300 Nantes, France (e-mail: francois.hug@univ-nantes.fr).

electromyography; gait; muscle coordination; pedaling; support vector machines

## INTRODUCTION

Each individual is unique, with distinctive patterns or characteristics by which he/she can be identified, leading to the notion of a personal signature. Using relatively simple biometric identifiers, algorithms have been developed that identify individuals based on their face, iris, or fingerprints. However, differences between individuals go well beyond differences in physical characteristics: individuals also differ in their ways of interacting with their environment. Our ability to identify a friend by his or her walk (9), for example, suggests the existence of identifiable movement signatures. Recently, a number of studies have supported the existence of individual movement signatures identified from kinematic or kinetic features (17, 27, 31). For instance, using support vector machines for pattern recognition, Horst et al. (17) showed that the ground reaction force pattern measured in 128 healthy individuals during gait could be assigned to the corresponding individuals with a classification rate >99%. In other words, participants could be accurately identified based on their ground reaction force pattern. Considering muscle coordination as the distribution of force among muscles to produce a given motor task (20), the existence of these movement signatures could be explained by the following: 1) individual anatomical/mechanical differences that result in different movement kinematics/kinetics, despite similar muscle coordination strategies (32); 2) the existence of muscle coordination strategies unique to each individual (20); or 3) both. In other words, it is unclear whether movement signatures result from the existence of individual muscle coordination signatures.

The most comprehensively studied contributor to muscle coordination is muscle activation, which is classically estimated using surface electromyography (EMG). Human studies that focus on individual data report differences in muscle activation patterns between individuals during multijoint tasks, such as gait (1, 2, 28), pedaling (22), and giant swing on the high bar (14). Large individual differences in the distribution of activation among synergist muscles have also been observed during well-controlled tasks, such as isometric plantar flexion

(26) and isometric knee extension (19). It is, therefore, well known that the distribution of muscle activation required to produce even simple movements can vary between individuals. However, the existence of such an interindividual variability does not prove the existence of individual muscle activation signatures. Indeed, if we consider a signature as distinctive patterns or characteristics by which someone can be identified, there are two important factors that need to be confirmed to support the existence of individual signatures. First, individual differences in muscle activation should persist over time. To the best of our knowledge, the vast majority of studies reporting interindividual variability of activation strategies did not test the robustness of these strategies across days. Second, these strategies should be unique in the sense that they should not be exactly the same for any two individuals performing the same task. As classical approaches to assess interindividual variability are often limited to descriptive statistics on time-discrete EMG variables; they cannot test the uniqueness of the time-varying EMG patterns. To test the uniqueness and robustness of muscle activation strategies, pattern recognition tools can be used; if an algorithm can accurately identify participants, one can conclude that these participants exhibit discernible differences and thus have unique signatures.

Bearing these considerations in mind, this study aimed to test the hypothesis that each individual has unique muscle activation signatures through a machine learning approach (support vector machine). Specifically, we hypothesized that activation patterns estimated using surface EMG over lower limb muscles during pedaling and gait can be assigned to the corresponding individual with a low classification error. If this hypothesis is supported, it would demonstrate that strategies not only vary between individuals, as is often assumed, but are also unique to each individual. This would open new research perspectives in which individual differences are considered relevant information for addressing fundamental questions about the control of movement in health, aging, and disease.

## METHODS

### Participants

Eighty healthy volunteers (62 men and 18 women; Table 1) participated in this study. Participants had no history of lower leg pain limiting function within the previous 2 mo. The ethics committee, CPP Ile de France XI, approved this study (no. 2018-A00110-55/18020), and all procedures adhered to the Declaration of Helsinki. Participants provided informed, written consent.

Table 1. Demographic and anthropometric data for the tested population

	Men	Women
<i>n</i>	62	18
Age, yr	24.1 ± 5.6 (18–46)	22.1 ± 4.7 (18–38)
Height, cm	179.3 ± 6.1 (158–193)	167.8 ± 6.0 (150–177)
Body mass, kg	74.8 ± 8.2 (51–90)	59.1 ± 4.5 (51–68)
Body mass index, kg/m <sup>2</sup>	23.2 ± 1.9 (18.7–28.1)	21.0 ± 1.7 (18.8–25.2)
Maximal power output, W	937 ± 144 (600–1,291)	637 ± 82 (464–778)
Left footed, <i>n</i>	9 (14.5%)	0 (0%)

Values are means ± SD (with range in parentheses); *n*, no. of subjects. Maximal power output was assessed during isokinetic (80 rpm) pedaling sprints.

### Experimental Design

The experimental session consisted of a series of locomotor tasks: two all-out isokinetic pedaling sprints used to standardize the intensity of the submaximal pedaling tasks, pedaling at four submaximal intensities, and walking on a treadmill at 1.11 m/s. The order of both of the tasks (pedaling and walking) and the intensities of the pedaling tasks were randomized. Of these 80 participants, 53 (12 women and 41 men) performed a second experimental session ~13 days later (range: 1–41 days; standard deviation: 10 days). This second session included all of the submaximal tasks.

### Myoelectrical Activity

Myoelectrical activity data were collected via surface EMG from eight muscles of the right leg: vastus lateralis (VL), rectus femoris (RF), vastus medialis (VM), gastrocnemius lateralis (GL), gastrocnemius medialis (GM), soleus (SOL), tibialis anterior (TA), and biceps femoris-long head (BF). For each muscle, a wireless surface electrode (Trigno Flex, Delsys, Boston) was attached to the skin at the site recommended by SENIAM (15). We intentionally did not mark the electrode location, such that day-to-day variability of the electrode placement could have occurred. It was important not to exclude this source of between-day variability because electrode placement might explain, at least in part, interindividual variability in estimated activation strategies. Before electrode application, the skin was shaved and cleaned with an abrasive pad and alcohol. Electrodes were well secured to the skin with double-sided tape and a tubular elastic bandage (tg fix, Lohmann & Rauscher International, Germany). EMG signals were band-pass filtered (10–850 Hz) and digitized at a sampling rate of 2,000 Hz using an EMG acquisition system (Trigno, Delsys, Boston).

### Experimental Protocol

**Pedaling.** The pedaling task was performed on an electronically braked cycloergometer (Excalibur Sport; Lode, Groningen, the Netherlands) equipped with standard cranks (170 mm) and clipless pedals. The saddle height was standardized such that the lower limb was straight when the heel was positioned in the middle of the pedal axle with the crank at the bottom of the cycle and aligned with the seat tube. To standardize the saddle setback, the knee cap was aligned with the ball of the foot, while the pedal was positioned at 90° from the top dead center (highest position of the pedal). The handlebar height was matched to that of the saddle. Participants were instructed to maintain a seated position throughout the tasks and to keep their hands on the dropped portion of the handlebars. This overall standardization procedure was adopted to limit the influence of pedaling positions on the observed variability of activation strategies between participants and across the two testing sessions.

After familiarization with the cycloergometer and a standardized warm-up, participants were asked to perform two 5-s all-out isokinetic pedaling sprints at 80 rpm, separated by 2 min of rest. The torque exerted on the cranks was measured by strain gauges in the crank arms of the cycloergometer (Excalibur Sport; Lode, Groningen, the Netherlands). The average of the two cycles with the highest power output was considered as the maximal power output ( $P_{max}$ ). Because we cannot discount the possibility that the individual differences, and thus the excellent classification rate, stems from different relative exercise intensities across participants, both absolute and relative pedaling intensities were tested. Specifically, participants were asked to pedal at four different intensities (80 W, 150 W, 10% of  $P_{max}$ , and 15% of  $P_{max}$ ; randomized order) each at 80 rpm for 90 s, with 30 s of rest in between. A transistor-transistor logic pulse indicated the top dead center of the right pedal and was recorded on the EMG acquisition system such that the crank position and the EMG data were synchronized.

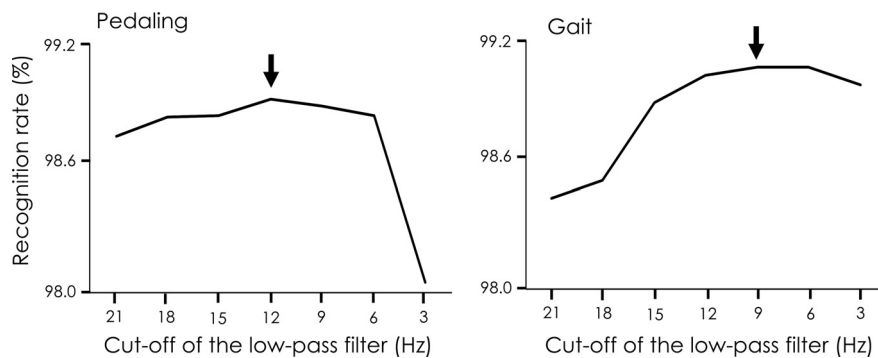


Fig. 1. Recognition rate as a function of the low-pass cut-off frequency used to compute the electromyography envelope. Results are presented for the leave-one-out method applied to data from *day 1* when considering all of the eight muscles ( $n = 78-79$  participants). Data for the four pedaling conditions are averaged. Arrows show the cut-off of the low-pass filter selected for each of the two conditions.

**Gait.** The gait experiments were conducted on a treadmill (Power 795i, Pro-form, France) to minimize perturbations induced by the external environment and to ensure that all participants adopted the same walking speed. Participants walked barefoot and were familiarized with the treadmill before starting the experimental task, which consisted of walking at 1.11 m/s for 90–120 s. A force-sensitive resistor (Delsys) was taped under the heel of the right foot to detect the onset of foot contact (i.e., the onset of the stance phase). These signals were recorded by the acquisition system used for EMG such that the foot pressure and the EMG data were synchronized.

#### Data Analysis and Statistics

All EMG data were processed using MATLAB (The Mathworks, Naticks). Raw EMG signals were first band-pass filtered (20–700 Hz) with a second-order Butterworth filter. Then EMG signals were inspected for noise or artifacts. At this stage, some data were discarded due to movement artifacts, leaving data for analysis from 78–79 participants for *day 1* and 49–50 participants for the follow-up (*days 1* and 2). The number of participants is indicated for each task in Tables 2 and 3.

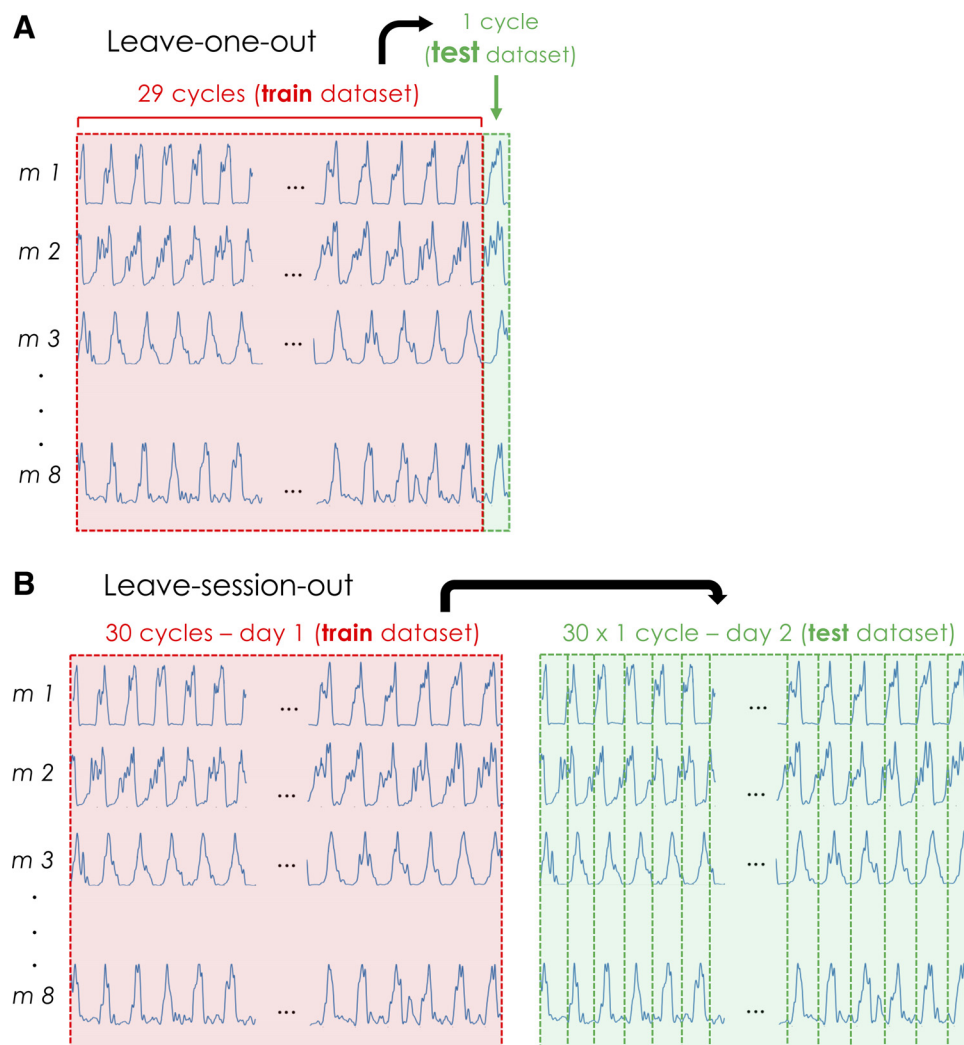


Fig. 2. Data analysis. The classification task typically involved separating data into training and testing sets. *A*: for the leave-one-out method applied to data from *day 1*, each pedaling/gait cycle was tested using the classifier, which had been trained on the remaining 29 cycles. *B*: for the leave-session-out method applied to data from *days 1* and 2, each pedaling/gait cycle from *day 2* was tested using the classifier, which had been trained on the 30 cycles from *day 1*. Each instance was predicted once, and the average accuracy provided the percentage of correctly classified data over the total number of cycles (i.e., 30 cycles  $\times$  number of participants). *m1–m8*, muscles 1–8.



Table 2. Participant classification using support vector machines for data from day 1

No. of Muscles	Pedaling				Gait
	15% P <sub>max</sub> (n = 79)	150 W (n = 79)	10% P <sub>max</sub> (n = 78)	80 W (n = 79)	1.11 m/s (n = 79)
1	60.34 GM	61.35 GM	57.78 SOL	58.35 SOL	58.69 BF
2	82.49 RF, GM	82.95 RF, GM	79.62 VM, SOL	78.0% RF, SOL	81.43 TA, BF
3	91.94 RF, GL, SOL	93.67 RF, SOL, BF	92.69 RF, GL, SOL	90.25 RF, GL, SOL	91.52 GM, TA, BF
4	95.74 RF, GL, TA, BF	96.84 RF, GM, SOL, BF	95.85 RF, GL, SOL, BF	94.22 RF, GL, SOL, TA	95.06 GM, SOL, TA, BF
5	97.59 RF, GL, GM, SOL, TA	98.02 RF, VM, GM, SOL, BF	97.69 RF, GL, SOL, TA, BF	96.54 RF, GL, GM, SOL, BF	97.09 VL, GL, SOL, TA, BF
6	98.52 RF, GL, GM, SOL, TA, BF	98.65 RF, VM, GL, GM, SOL, BF	98.63 RF, VM, GL, SOL, TA, BF	97.55 RF, VM, GL, GM, SOL, TA	98.02 RF, VM, GL, SOL, TA, BF
7	99.20 VL, RF, GL, GM, SOL, TA, BF	98.99 RF, VM, GL, GM, SOL, TA, BF	98.89 RF, VM, GL, GM, SOL, TA, BF	98.10 VL, RF, GL, GM, SOL, TA, BF	98.82 RF, VM, GL, GM, SOL, TA, BF
8	99.16 All	99.28 All	99.02 All	98.22 All	98.86 All

Values are the highest classification rate in percentage, along with the combination of muscles that led to this rate (n = 78 or 79 subjects). P<sub>max</sub>, maximal power output; BF, biceps femoris; GL, gastrocnemius lateralis; GM, gastrocnemius medialis; RF, rectus femoris; SOL, soleus; TA, tibialis anterior; VL, vastus lateralis; VM, vastus medialis.

To quantify interindividual differences in activation strategies, we considered the EMG time-varying profiles measured during pedaling and gait. The classification of EMG patterns was performed using support vector machines, which consist of a supervised machine learning algorithm for pattern recognition (5). The L2-regularized L2-loss support vector classification (primal) of the Liblinear Toolbox 2.11 (13) was used with a linear kernel function. *Option C* was

selected such that the best *C* was first determined by cross-validation. *C* is a parameter that controls the trade-off between smooth decision boundary and classifying the training data correctly. For both the pedaling and the gait task, the first 20 cycles were excluded from analysis. Then the first 30 consecutive cycles that were free of any artifacts were selected. The EMG signals were rectified and low-pass filtered at 12 and 9 Hz for pedaling and gait, respectively. This

Table 3. Participant classification using support vector machines for data from day 1 and 2

No. of Muscles	Pedaling				Gait
	15% P <sub>max</sub> (n = 49)	150 W (n = 49)	10% P <sub>max</sub> (n = 49)	80 W (n = 49)	1.11 m/s (n = 50)
1	39.18 GM	39.80 RF	35.44 GL	32.24 RF	35.60 TA
2	64.42 RF, GL	60.68 RF, GL	56.19 RF, GL	56.94 RF, GM	55.20 GL, SOL
3	76.05 RF, GL, GM	72.65 RF, GM, GL	69.93 RF, GL, GM	70.48 RF, GL, GM	67.40 SOL, TA, BF
4	82.17 RF, GL, GM, TA	80.82 RF, GL, GM, BF	75.85 RF, GL, GM, SOL	76.46 RF, GL, GM, TA	77.67 GM, SOL, TA, BF
5	85.58 RF, GL, GM, SOL, TA	85.85 RF, GL, GM, SOL, BF	79.18 RF, GL, GM, TA, BF	82.45 RF, GL, GM, TA, BF	83.67 GL, GM, SOL, TA, BF
6	85.84 VL, RF, GL, GM, SOL, TA	88.71 RF, GL, GM, SOL, TA, BF	83.33 RF, GL, GM, SOL, TA, BF	85.78 RF, VM, GL, GM, TA, BF	84.87 VM, GL, GM, SOL, TA, BF
7	89.79 VL, RF, VM, GL, GM, SOL, TA	89.80 VL, RF, GL, GM, SOL, TA, BF	85.85 RF, VM, GL, GM, SOL, TA, BF	86.87 VL, RF, VM, GL, GM, TA, BF	86.20 VL, VM, GL, GM, SOL, TA, BF
8	90.81 All	91.22 All	86.94 All	88.30 All	86.73 All

Values are the highest classification rate in percentage, along with the combination of muscles that led to this rate; n = 49 or 50 subjects. P<sub>max</sub>, maximal power output; BF, biceps femoris; GL, gastrocnemius lateralis; GM, gastrocnemius medialis; RF, rectus femoris; SOL, soleus; TA, tibialis anterior; VL, vastus lateralis; VM, vastus medialis.

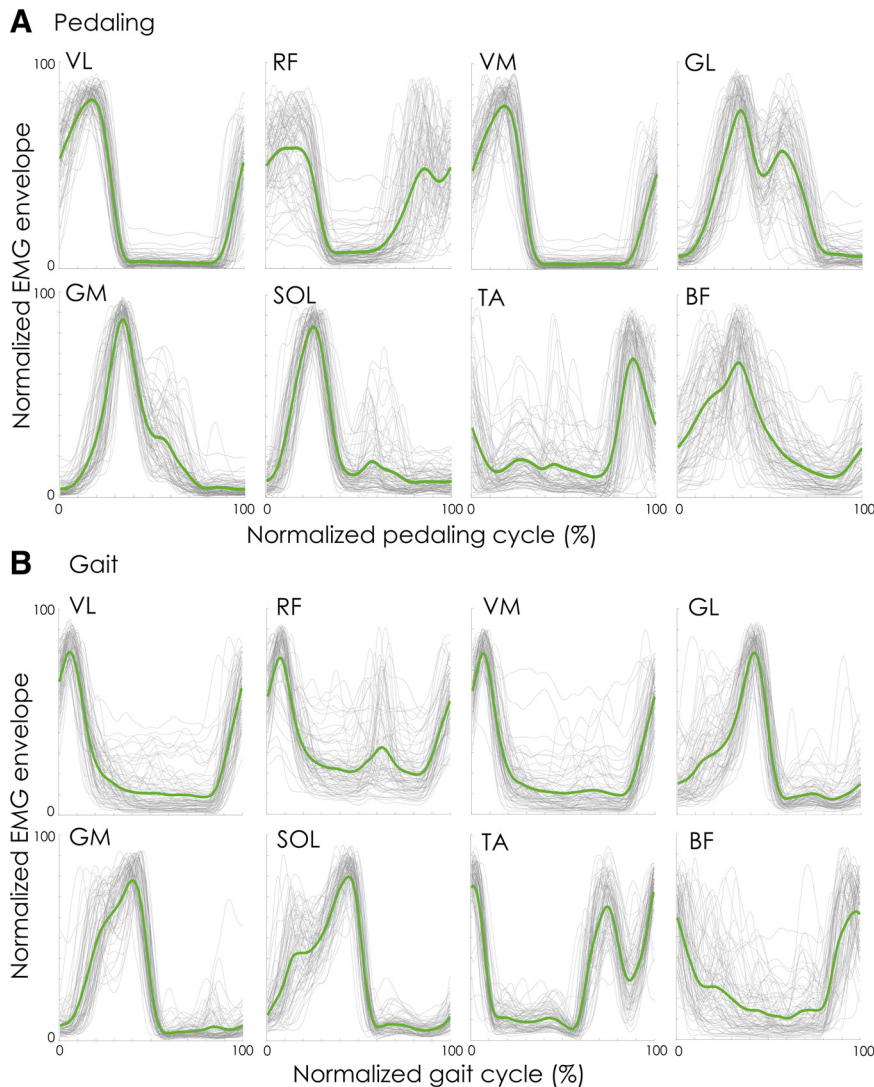


Fig. 3. Interindividual variability of electromyography (EMG) patterns during pedaling at 150 W (A) and gait at 1.11 m/s (B). For each participant, the mean profile over the 30 cycles is shown in gray. These profiles were normalized to their maximal value. The mean profile over the 79 participants is shown in green. Note that activation of the quadriceps muscles was low during gait, which explains the relatively low signal-to-noise ratio for some participants. BF, biceps femoris; GL, gastrocnemius lateralis; GM, gastrocnemius medialis; RF, rectus femoris; SOL, soleus; TA, tibialis anterior; VL, vastus lateralis; VM, vastus medialis.

low-pass filter was chosen because it provided the highest recognition rate out of those tested (from 3 to 21 Hz; Fig. 1) and falls well within the 4- to 15-Hz range classically used for smoothing of EMG patterns measured during locomotor tasks under similar velocity/frequency conditions (21). For every muscle, data for each cycle were interpolated to 200 time points and normalized to its maximal EMG amplitude (within that cycle), leading to a signal amplitude between 0 and 1. This normalization procedure ensured that all muscles and cycles contributed equally to the classification.

The classification of EMG patterns was performed by considering all possible combinations of  $n$  muscles from  $n = 1-8$ . This resulted in 8, 28, 56, 70, 56, 28, 8, and 1 possible combinations when considering 1, 2, 3, 4, 5, 6, 7, and 8 muscles, respectively.

Data from the first session were used to assess individual differences in activation strategies. The data set consisted of a  $c \times m$  matrix [ $c = 30$  cycles  $\times$  number of participants (78 or 79);  $m = 200$  time points  $\times$  number of muscles (1-8)]. The classification task typically involved separating data into training and testing sets. To this end, we used the leave-one-out method. Each pedaling/gait cycle was tested using the classifier, which had been trained on the remaining 29 cycles (Fig. 2). Each instance was predicted once, and the average accuracy provided the percentage of correctly classified cycles over 2,340 or 2,370 cycles [i.e., 30 cycles  $\times$  number of participants (78 or 79, respectively)].

Data from participants who performed two sessions on separate days were used to assess the robustness of the activation strategies

over time. The data set consisted of two  $c \times m$  matrices, i.e., one for each day [ $c = 30$  cycles  $\times$  number of participants (49 or 50);  $m = 200$  time points  $\times$  number of muscles (1-8)]. To this end, we used the leave-session-out method. Each pedaling/gait cycle from *day 2* was tested using the classifier, which had been trained on the 30 cycles from *day 1* (Fig. 2). Each instance was predicted once, and the average accuracy provided the percentage of correctly classified cycles over 1,470 or 1,500 cycles [i.e., 30 cycles  $\times$  number of participants (49 or 50, respectively)].

The percentage of correctly classified data (the classification rate) was interpreted in regard to the random classification rate that was expected by chance (1/number of participants). In addition, according to previous data on gait kinetics and kinematics (16, 17), we considered a recognition rate  $>80\%$  as strong evidence of discernible individual patterns.

## RESULTS

During both pedaling and gait, interindividual variability of the EMG patterns was substantial, particularly for some biarticular muscles (e.g., RF, GL, BF) and the SOL muscle (Fig. 3). Figure 4 depicts the EMG patterns during pedaling at 150 W and during gait for four participants, highlighting some individual-specific patterns that can be visually identified. For example, when considering the pedaling task, a double burst of

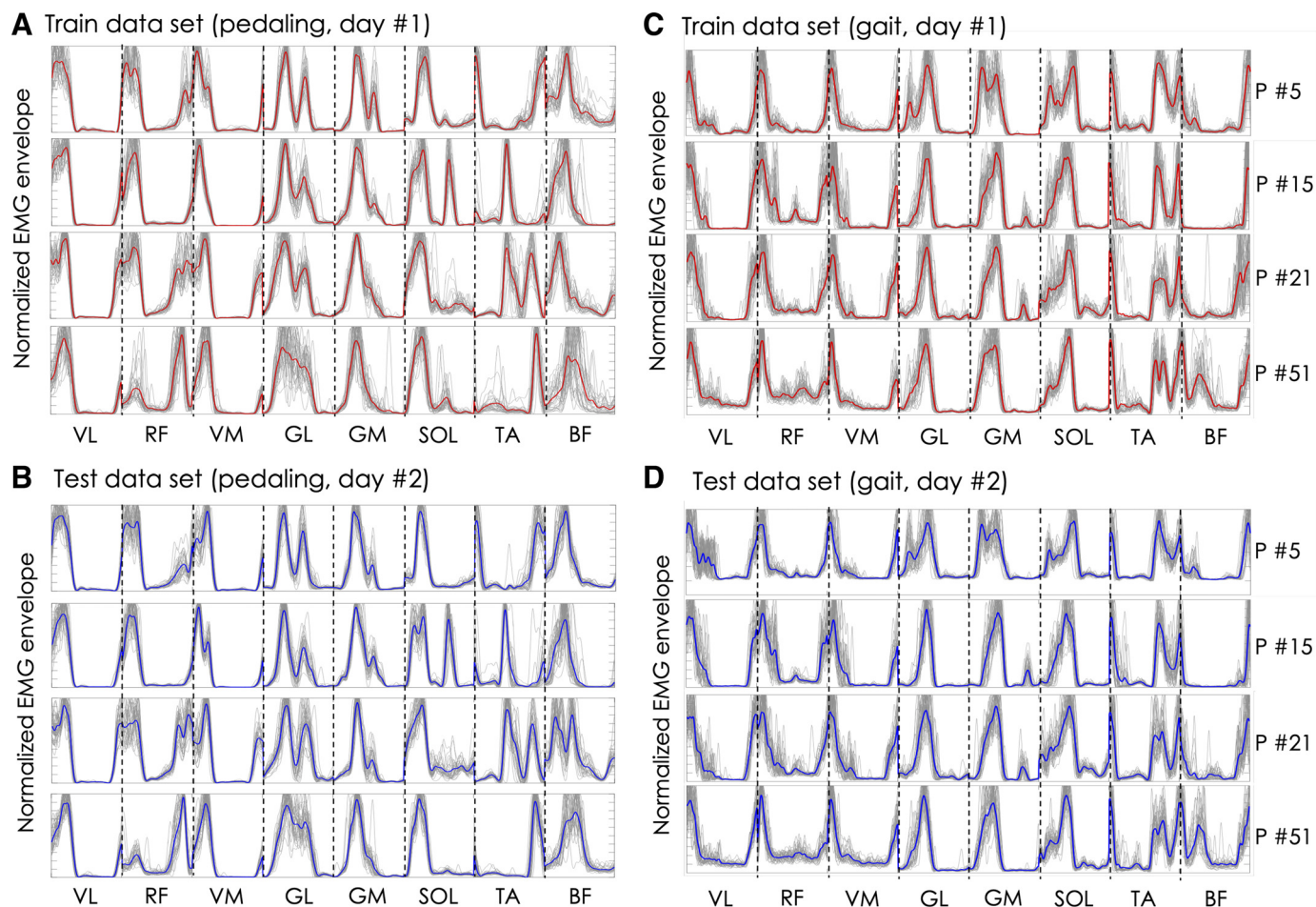


Fig. 4. Examples of individual-specific electromyography (EMG) patterns measured during either pedaling (*A* and *B*) or gait (*C* and *D*). Each time-varying profile in gray corresponds to one of the 30 consecutive cycles used as the training (*day 1*; *A* and *C*) or testing data set (*day 2*; *B* and *D*). These profiles were normalized to their maximal value. The mean profile over the 30 cycles is shown in red or blue but was not considered in the actual data sets. Data for four participants (P #5, P #15, P #21, P #51) are depicted to highlight differences between individuals. BF, biceps femoris; GL, gastrocnemius lateralis; GM, gastrocnemius medialis; RF, rectus femoris; SOL, soleus; TA, tibialis anterior; VL, vastus lateralis; VM, vastus medialis.

GL activation is clearly visible for *participants 5, 15, and 21*, whereas the other participant exhibited only one burst. Also, each of the four participants exhibited a different TA activation pattern. Even though individual variability was less clear for quadriceps muscles, RF muscle exhibited much clearer activation during the upstroke phase of the pedaling cycle for *participants 21 and 51* than for the others. Interestingly, all of these specific patterns were observed on both days. Of note, such individual differences have already been reported (18, 22); the novelty of the present study lies in quantifying the uniqueness of these strategies using a machine learning approach.

The support vector machines analysis was first performed on participants from *day 1*. When considering eight muscles, classification rate reached 99.28% (2,353/2,370 cycles) for pedaling at 150 W and 98.86% (2,343/2,370 cycles) for gait (Fig. 5; Table 2). Even though recognition rate was lower when fewer muscles were considered in the analysis, it remained well beyond the theoretical random classification rate (1/79 participants = 1.27%). Specifically, when considering  $n$  muscles and their various combinations for the pedaling tasks, the highest classification rate was 61.35% for one muscle, 82.95% for two muscles, 93.67% for three muscles, 96.84% for four

muscles, 98.02% for five muscles, 98.65% for six muscles, and 99.20% for seven muscles (Fig. 5A). Muscle combinations that led to these recognition rates are presented in Table 2. RF, SOL, GM, and GL were the muscles the most represented within these combinations, which suggests that their activation profile exhibited the largest variation between participants. Interestingly, the majority of the pedaling cycles (i.e., >15/30) were correctly classified in all of the participants when considering two or more muscles during pedaling at 150 W. This demonstrates that two muscles (here, RF and GM) were sufficient to correctly identify all of the participants from their EMG patterns.

When considering  $n$  muscles and their various combinations for gait, the highest classification rate was 58.69% for one muscle, 81.43% for two muscles, 91.52% for three muscles, 95.06% for four muscles, 97.09% for five muscles, 98.02% for six muscles, and 98.82% for seven muscles (Fig. 5B; Table 2). BF and TA muscles were the muscles most represented within these combinations. Interestingly, most cycles (>15/30) were correctly classified in each participant when considering three muscles and more.



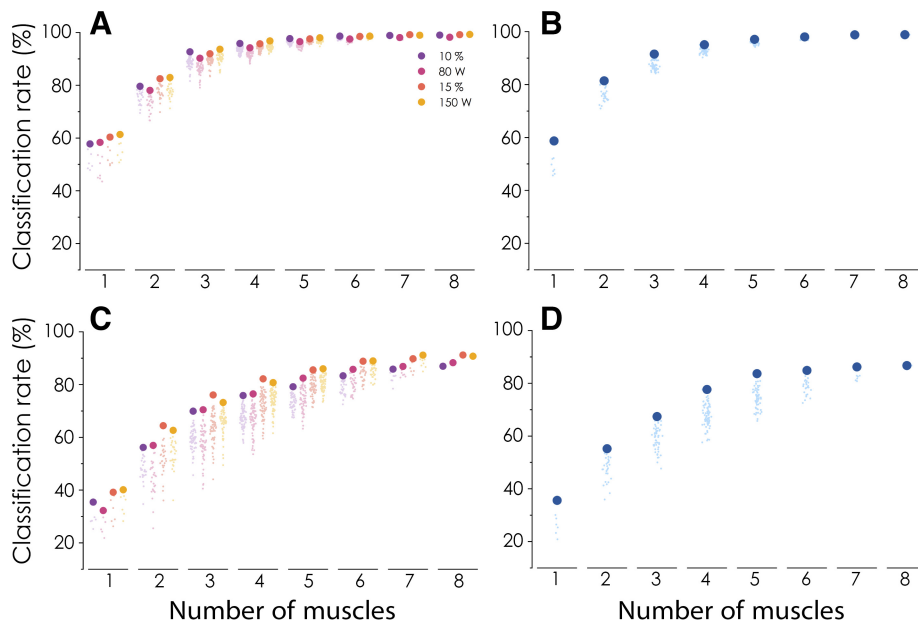


Fig. 5. Classification rate obtained from the leave-one-out [pedaling (A) and gait (B)] and the leave-session-out method [pedaling (C) and gait (D)]. The classification of electromyography patterns was performed by considering all possible combinations of  $n$  muscles from  $n = 1-8$ , which corresponded to 8, 28, 56, 70, 56, 28, 8, and 1 combinations for  $n = 1, 2, 3, 4, 5, 6, 7,$  and 8 muscles, respectively. The result for each of these combinations is shown as a small light dot, whereas the color refers to the condition. The highest classification rate for each condition is depicted as a bigger dark dot. Depending on the pedaling condition,  $n = 78-79$  participants for the leave-one-out method and  $n = 49-50$  participants for the leave-session-out method. For the gait condition,  $n = 79$  and 50 participants for the leave-one-out method and the leave-session-out method, respectively.

To test the robustness of the EMG patterns over time, a subset of participants performed a second session. When considering eight muscles, the classification rate reached 91.22% (1,341/1,470 cycles) for pedaling at 150 W and 86.73% (1,301/1,500 cycles) for gait (Table 3). Again, the recognition rate was lower when fewer muscles were considered in the analysis, but it remained well beyond the theoretical random classification rate (1/49 participants = 2.04%). Specifically, when considering  $n$  muscles and their various combinations for the pedaling tasks, the highest classification rate was 39.80% for one muscle, 64.42% for two muscles, 76.05% for three muscles, 82.17% for four muscles, 85.84% for five muscles, 88.71% for six muscles, and 89.80% for seven muscles (Fig. 5C). RF, GM, and GL were the muscles the most represented within these combinations, which suggests that their activation profile exhibited the largest variation between participants. The majority of the pedaling cycles (i.e., >15/30) were correctly classified in >44/49 participants when considering four or more muscles during pedaling at 15%.

When considering  $n$  muscles and their various combinations for gait, the highest classification rate was 35.60% for one muscle, 55.20% for two muscles, 67.40% for three muscles, 77.67% for four muscles, 83.67% for five muscles, 84.87% for six muscles, and 86.20% for seven muscles (Fig. 5D; Table 3). SOL and TA muscles were the muscles the were most represented in these combinations. Interestingly, the majority of cycles (>15/30) were correctly classified in >45/49 participants when considering four or more muscles.

## DISCUSSION

This study used a machine learning approach to classify EMG profiles measured in up to eight lower limb muscles during pedaling and gait, separately. When considering the eight lower limb muscles, the EMG profiles were assigned to the corresponding individuals with a classification rate of up to 99.28 and 91.22% (pedaling 150 W) for the within-day and between-day classification, respectively. Even though the recognition rate was lower when fewer muscles were considered

in the analysis, it remained >80% for two and four muscles for within-day and between-day classification, respectively. This finding supports the assumption that each individual has unique muscle activation signatures, i.e., activation strategies not only vary between individuals, but also are unique to each individual.

### *Our Approach to Demonstrate the Existence of Individual Activation Signatures*

It is well known that individuals exhibit different motor styles, known as the individuality principle (33). This is evidenced here as different EMG time-varying profiles. Such a large variability of activation patterns has been reported in experiments in animals (23, 24) and humans (2, 22). However, finding different activation patterns does not conclusively prove the existence of an individual muscle activation signature, as the concept of individual signatures implies that the activation patterns are unique features by which individuals can be identified. In addition, the vast majority of previous research did not test the robustness of the EMG time-varying profiles over time, making it impossible to conclude that the interindividual variability of these profiles reflects individual strategies rather than random variation.

We took advantage of support vector machines to test the uniqueness of muscle activation strategies. We considered that, if an algorithm can accurately classify participants, one could conclude that these participants exhibit discernible differences and thus have unique signatures. Using a similar approach on kinematics and/or kinetics data, recent studies reported an almost error-free assignment of gait patterns to the corresponding individual (16, 17, 27, 29). Although this suggests the existence of individual gait pattern signatures, the existence of individual muscle activation signatures cannot be inferred from these data. This is because the force produced by a muscle depends on the complex interplay between its activation and several biomechanical factors (physiological cross-sectional area, specific tension, force-length, and force-velocity relationship). As such, different mechanical outputs can be theoretic-



cally produced by the same activation strategies and vice versa. To the best of our knowledge, our study is the first to use a supervised learning approach to classify participants based on their muscle activation patterns. We considered two locomotor tasks with different degrees of mechanical constraints: pedaling, for which the foot trajectory, movement velocity, and power output can be matched between participants; and gait, which is a less constrained form of locomotion.

#### *Participant Classification from Day 1*

The classification of data from *day 1* was almost error free when considering the eight lower limb muscles all together during both pedaling (range: 98.22–99.28%; Table 2) and gait (98.86%; Table 2). Despite not being measured in the present study, gait was very likely associated with more interindividual variability in movement kinematics than pedaling. Within this context, the similar classification rates observed between gait and pedaling suggest that individual activation strategies were not solely due to differences in movement kinematics between individuals.

Although the classification rate was lower when considering fewer muscles (Table 2), it remained well beyond the theoretical random classification rate. When considering all possible combinations of  $n$  muscles, from  $n = 1$ –7, the classification rate was  $>80\%$  for combinations of two muscles and  $>90\%$  for combinations of three muscles (Table 2). Of note, this rate is based on the classification of the pedaling/gait cycles. If we consider that the participants are accurately identified when the majority of their pedaling or gait cycles (i.e.,  $>15/30$ ) are correctly assigned to them, 100% of the participants were correctly identified from a combination of two muscles for pedaling and from a combination of three muscles for gait.

Within the muscle combinations that gave the best recognition rates, the most represented muscles were RF, GM, GL, and SOL for pedaling and TA and BF for gait. This suggests that the activation profile of these muscles exhibited unique features, which is further confirmed by inspection of Fig. 4. Interestingly, interindividual variability of GM and GL EMG patterns during pedaling was mostly localized during the upstroke phase, which is not the main component of the pedaling cycle in terms of power production (30). Monoarticular muscles, such as VL and VM, were the least represented within the combinations that gave the best recognition rates during pedaling. This might be explained by the functions of these muscles during pedaling, in which quadriceps (mostly the vastii) are the primary power producers and thus are essential for the task and exhibit less variation between participants. Overall, these results are in line with those obtained from behaving California sea hares [*Aplysia californica* (7, 8)], which demonstrated that the degree to which a motor component varied between animals depends on the role of that motor component in the behavior. Specifically, the motor neuronal activity that is essential for biting or swallowing was similar among animals, whereas motor neuronal activity that is less essential for that behavior varied more across individuals (8).

Of note, the recognition rate achieved by the support vector machines analysis might have been affected by two limitations. First, the classification task was performed on the basis of only eight muscles, whereas pedaling or gait are produced by the coordinated activation of many more muscles. Second, the

normalization procedure of the EMG signals did not consider possible between-muscle differences in magnitude of activation. By considering both more muscles and information about the magnitude of activation in respect to maximal activation, it is very likely that the support vector machines analysis would be able to achieve an even higher classification rate, possibly 100%.

#### *Participant Follow-up Classification (Days 1 and 2)*

The excellent classification rate from data obtained within a single session (*day 1*) provides strong evidence that the EMG time-varying profiles are unique to each individual. However, to determine whether these profiles represent individual strategies rather than random variability, it was necessary to demonstrate their robustness over time. When data from *day 2* were tested using the classifier, which had been trained on data from *day 1*, the participant follow-up classification rate remained high (range: 86.94–91.22% for the different pedaling conditions and 86.73% for gait; Table 3). This classification rate was slightly higher than that previously reported from ground reaction force data during gait [83.48%; Horst et al. (17)]. Again, when considering all possible combinations of  $n$  muscles from  $n = 1$ –7, the classification rate logically decreased but was  $>80\%$  for combinations of four muscles during pedaling and for combinations of five muscles during gait (Table 3). Overall, the good participant follow-up classification rate demonstrates that individual characteristics of EMG time-varying profiles persist over time, allowing us to consider that each individual exhibits his/her own muscle activation signature.

Even though the participant follow-up classification rate was good, it was lower than the within-day participant classification rate. This result can be explained by two sources of variability. First, it is possible that the EMG electrode placement differed slightly between the sessions, leading to different EMG time-varying profiles. It was for the purpose of allowing this source of variability that we chose not to mark the electrode location. Indeed, by allowing the electrode location to slightly vary between the sessions, the impact of different electrode placement among participants on the participant-classification rate was allowed to influence the data. Because classification rate decreased only marginally for the participant follow-up classification compared with the participant classification, we are confident that the excellent classification rate obtained for data from *day 1* is not due to different electrode placement among participants. However, a more controlled experiment with consideration of spatial variability of muscle activation is needed to quantify the impact of the electrode placement in the recognition rate. Second, it is possible that the lower participant follow-up classification rate is explained by natural intra-individual variability of the activation strategies between days; however, this variability, if any, remained marginal.

#### *Origin and Consequence of Individual Muscle Activation Signatures*

By showing discernible differences in muscle activation patterns, our results suggest that the control of locomotor tasks is unique to an individual. Although the origin of these individual differences remains unclear, they can be discussed in regard to current motor control theories. The optimal feedback

control theory (34) proposes that motor patterns are selected such that movement costs (e.g., smoothness, activation, jerk, energy) are constantly minimized. Within this framework, one would hypothesize that each individual optimizes a different cost and/or that optimizing the same cost(s) requires different activation strategies across individuals, due to individual characteristics (e.g., morphology, anatomy, neural constraints). In this way, we observed a positive correlation between the distribution of muscle activation measured during an isometric task and the distribution of muscle force-generating capacity between VL and VM (19) and between GM and GL (6). Although this correlation does not imply causation, it suggests that people might exhibit different distribution of activation among synergists, based on their specific distribution of force-generating capacity. This strategy might aim to optimize the energetic or activation cost. Overall, it is possible that each individual optimizes their movement with the muscle activation strategies that are best, given that individual's mechanical and/or neural constraints (e.g., relative force-generating capacity of the synergist muscles, shape of articular surfaces, or specific bone configuration). An alternative motor control theory, the "good-enough" theory, proposes that a hierarchy of sensorimotor networks gradually adapts through trial-and-error learning to produce effective movements that are good enough to achieve the task goal (25). It is, therefore, possible that individuals develop different "good-enough" muscle activation strategies through motor exploration, experience, and training, leading to habitual rather than optimal strategies (10). In this way, there is evidence from animal models that the central pattern generators are capable of learning and adaptation, explaining, at least in part, the interanimal variability (23). Even though we recruited a homogeneous population of active participants in terms of age and body mass index, it is possible that their past and present experience with specific motor skills (i.e., motor history) might have played a role in shaping individual strategies. To address this question, retrospective studies on large cohorts or longitudinal studies performed at different lifespans are needed. Finally, we cannot rule out that part of these strategies is innate. In this way, interindividual variability of muscle synergies has been identified in neonates during the stepping reflex (Fig. 2B in Ref. 12). Overall, these results provide the impetus to determine whether factors, such as development, aging, and disease, might alter a person's individuality in muscle activation.

Even though reporting discernible differences in activation patterns between individuals does not provide direct evidence that these differences are biomechanically or functionally significant, it is known that even small differences in activation level can have a large effect on movement (3, 11). It is, therefore, possible that each individual muscle activation signature has specific mechanical effects on the musculoskeletal system (20). As such, some strategies may place some individuals at greater risk of developing musculoskeletal disorders than others. For example, we observed a quite large variation of GM, GL, and SOL activation patterns between participants (Figs. 3 and 4). Because each of these three muscles are attached to a different fascicle bundle of the Achilles tendon, different activation strategies might induce unique patterns of load distribution within the Achilles tendon, with some strategies being more likely to lead to tendon problems (4). To test this assumption, it is necessary to consider muscle and tendon

biomechanical characteristics, which may either exacerbate or counteract differences in activation.

### Conclusion

Although it is known that muscle activation can vary between individuals, these differences have been very rarely considered as relevant information to expand our knowledge of motor control. By showing the uniqueness and robustness of EMG time-varying profiles, our results strongly suggest that activation strategies not only vary between individuals, but are unique to each individual. Similar to other biometrics, such as fingerprints, speech, and the structure of the iris, muscle activation strategies seem to be a unique feature that may be used to identify an individual. This study highlights the need to consider these individual differences as relevant information to address fundamental questions about the control of movement in health, aging, and disease.

### Data availability

All raw EMG data are available as supplemental material at <https://doi.org/10.6084/m9.figshare.8273774>.

### ACKNOWLEDGMENTS

The authors thank Fabian Horst (Johannes Gutenberg-University Mainz, Germany) for insights on the support vector machines method, and Simon Avrillon (University of Nantes, France) for helping with Fig. 5.

### GRANTS

F. Hug was supported by a fellowship from the Institut Universitaire de France. Support was received from the French National Research Agency (ANR-19-CE17-0002-01; COMMODE project).

### DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

### AUTHOR CONTRIBUTIONS

F.H. and L.L. conceived and designed research; F.H., C.V., and L.L. performed experiments; F.H. and E.L.C. analyzed data; F.H., K.J.T., S.D., T.D., E.L.C., and L.L. interpreted results of experiments; F.H. prepared figures; F.H. drafted manuscript; F.H., K.J.T., S.D., T.D., E.L.C., and L.L. edited and revised manuscript; F.H., C.V., K.J.T., S.D., T.D., E.L.C., and L.L. approved final version of manuscript.

### REFERENCES

1. Ahn AN, Kang JK, Quitt MA, Davidson BC, Nguyen CT. Variability of neural activation during walking in humans: short heels and big calves. *Biol Lett* 7: 539–542, 2011. doi:10.1098/rsbl.2010.1169.
2. Arsenaault AB, Winter DA, Marteniuk RG. Is there a 'normal' profile of EMG activity in gait? *Med Biol Eng Comput* 24: 337–343, 1986. doi:10.1007/BF02442685.
3. Barradas VR, Kutch JJ, Kawase T, Koike Y, Schweighofer N. When 90% of the variance is not enough: residual EMG from muscle synergy extraction influences task performance. *bioRxiv*. In press. doi:10.1101/634758.
4. Bojsen-Møller J, Magnusson SP. Heterogeneous loading of the human achilles tendon in vivo. *Exerc Sport Sci Rev* 43: 190–197, 2015. doi:10.1249/JES.0000000000000062.
5. Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 20: 273–297, 1995. doi:10.1007/BF00994018.
6. Crouzier M, Lacourpaille L, Nordez A, Tucker K, Hug F. Neuromechanical coupling within the human triceps surae and its consequence on individual force-sharing strategies. *J Exp Biol* 221: jeb187260, 2018. doi:10.1242/jeb.187260.
7. Cullins MJ, Gill JP, McManus JM, Lu H, Shaw KM, Chiel HJ. Sensory feedback reduces individuality by increasing variability within subjects. *Curr Biol* 25: 2672–2676, 2015. doi:10.1016/j.cub.2015.08.044.

8. Cullins MJ, Shaw KM, Gill JP, Chiel HJ. Motor neuronal activity varies least among individuals when it matters most for behavior. *J Neurophysiol* 113: 981–1000, 2015. doi:10.1152/jn.00729.2014.
9. Cutting JE, Kozlowski LT. Recognizing friends by their walk: gait perception without familiarity cues. *Bull Psychon Soc* 9: 353–356, 1977. doi:10.3758/BF03337021.
10. de Rugy A, Loeb GE, Carroll TJ. Muscle coordination is habitual rather than optimal. *J Neurosci* 32: 7384–7391, 2012. doi:10.1523/JNEUROSCI.5792-11.2012.
11. Diong J, Gandevia SC, Nguyen D, Foo Y, Kastre C, Andersson K, Butler JE, Héroux ME. Small amounts of involuntary muscle activity reduce passive joint range of motion. *J Appl Physiol* (1985) 127: 229–234, 2019. doi:10.1152/jappphysiol.00168.2019.
12. Dominici N, Ivanenko YP, Cappellini G, d'Avella A, Mondì V, Cicchese M, Fabiano A, Silei T, Di Paolo A, Giannini C, Poppele RE, Lacquaniti F. Locomotor primitives in newborn babies and their development. *Science* 334: 997–999, 2011. doi:10.1126/science.1210617.
13. Fan RE, Chang KW, Hsieh CJ, Wang XR, Lin CJ. LIBLINEAR: a library for large linear classification. *J Mach Learn Res* 9: 1871–1874, 2008.
14. Frère J, Hug F. Between-subject variability of muscle synergies during a complex motor skill. *Front Comput Neurosci* 6: 99, 2012. doi:10.3389/fncom.2012.00099.
15. Hermens HJ, Freriks B, Disselhorst-Klug C, Rau G. Development of recommendations for SEMG sensors and sensor placement procedures. *J Electromyogr Kinesiol* 10: 361–374, 2000. doi:10.1016/S1050-6411(00)00027-4.
16. Horst F, Kramer F, Schäfer B, Eekhoff A, Hegen P, Nigg BM, Schöllhorn WI. Daily changes of individual gait patterns identified by means of support vector machines. *Gait Posture* 49: 309–314, 2016. doi:10.1016/j.gaitpost.2016.07.073.
17. Horst F, Mildner M, Schöllhorn WI. One-year persistence of individual gait patterns identified in a follow-up study—a call for individualised diagnose and therapy. *Gait Posture* 58: 476–480, 2017. doi:10.1016/j.gaitpost.2017.09.003.
18. Hug F, Drouet JM, Champoux Y, Couturier A, Dorel S. Interindividual variability of electromyographic patterns and pedal force profiles in trained cyclists. *Eur J Appl Physiol* 104: 667–678, 2008. doi:10.1007/s00421-008-0810-y.
19. Hug F, Goupille C, Baum D, Raiteri BJ, Hodges PW, Tucker K. Nature of the coupling between neural drive and force-generating capacity in the human quadriceps muscle. *Proc Biol Sci* 282: 20151908, 2015. doi:10.1098/rspb.2015.1908.
20. Hug F, Tucker K. Muscle coordination and the development of musculoskeletal disorders. *Exerc Sport Sci Rev* 45: 201–208, 2017. doi:10.1249/JES.0000000000001122.
21. Hug F, Turpin NA, Dorel S, Guével A. Smoothing of electromyographic signals can influence the number of extracted muscle synergies. *Clin Neurophysiol* 123: 1895–1896, 2012. doi:10.1016/j.clinph.2012.01.015.
22. Hug F, Turpin NA, Guével A, Dorel S. Is interindividual variability of EMG patterns in trained cyclists related to different muscle synergies? *J Appl Physiol* (1985) 108: 1727–1736, 2010. doi:10.1152/jappphysiol.01305.2009.
23. Loeb GE. Asymmetry of hindlimb muscle activity and cutaneous reflexes after tendon transfers in kittens. *J Neurophysiol* 82: 3392–3405, 1999. doi:10.1152/jn.1999.82.6.3392.
24. Loeb GE. The distal hindlimb musculature of the cat: interanimal variability of locomotor activity and cutaneous reflexes. *Exp Brain Res* 96: 125–140, 1993. doi:10.1007/BF00230446.
25. Loeb GE. Optimal isn't good enough. *Biol Cybern* 106: 757–765, 2012. doi:10.1007/s00422-012-0514-6.
26. Masood T, Bojsen-Møller J, Kalliokoski KK, Kirjavainen A, Áärimaa V, Peter Magnusson S, Finni T. Differential contributions of ankle plantarflexors during submaximal isometric muscle action: a PET and EMG study. *J Electromyogr Kinesiol* 24: 367–374, 2014. doi:10.1016/j.jelekin.2014.03.002.
27. Pataky TC, Mu T, Bosch K, Rosenbaum D, Goulermas JY. Gait recognition: highly unique dynamic plantar pressure patterns among 104 individuals. *J R Soc Interface* 9: 790–800, 2012. doi:10.1098/rsif.2011.0430.
28. Pedotti A. A study of motor coordination and neuromuscular activities in human locomotion. *Biol Cybern* 26: 53–62, 1977. doi:10.1007/BF00363992.
29. Phinyomark A, Osis ST, Hettinga BA, Kobsar D, Ferber R. Gender differences in gait kinematics for patients with knee osteoarthritis. *BMC Musculoskelet Disord* 17: 157, 2016. doi:10.1186/s12891-016-1013-z.
30. Prilutsky BI, Zatsiorsky VM. Optimization-based models of muscle coordination. *Exerc Sport Sci Rev* 30: 32–38, 2002. doi:10.1097/00003677-200201000-00007.
31. Słowiński P, Zhai C, Alderisio F, Saless R, Guegnon M, Marin L, Bardy BG, di Bernardo M, Tsaneva-Atanasova K. Dynamic similarity promotes interpersonal coordination in joint action. *J R Soc Interface* 13: 20151093, 2016. doi:10.1098/rsif.2015.1093.
32. Smoger LM, Fitzpatrick CK, Clary CW, Cyr AJ, Maletsky LP, Rullkoetter PJ, Laz PJ. Statistical modeling to characterize relationships between knee anatomy and kinematics. *J Orthop Res* 33: 1620–1630, 2015. doi:10.1002/jor.22948.
33. Ting LH, Chiel HJ, Trumbower RD, Allen JL, McKay JL, Hackney ME, Kesar TM. Neuromechanical principles underlying movement modularity and their implications for rehabilitation. *Neuron* 86: 38–54, 2015. doi:10.1016/j.neuron.2015.02.042.
34. Todorov E. Optimality principles in sensorimotor control. *Nat Neurosci* 7: 907–915, 2004. doi:10.1038/nm1309.