Supplementary Material

Experimental Procedures

Study design

After initial acclimatization (at least two weeks) to the animal facility of the Biozentrum (Basel, Switzerland), subgroups of mice were implanted with a battery free transponder for the later application of a telemetric system to track spontaneous locomotor activity and core body temperature (Fig. 1A). In the weeks prior to the exercise and/or rIL-6 interventions, mice were subjected to a baseline characterization of behavior (spontaneous activity), gait (CatWalk analysis), motor coordination (Rotarod test), physical performance [running capacity and/or peak oxygen uptake ($\dot{V}O_{2peak}$)], metabolic parameters [body mass, body composition, rate of oxygen uptake ($\dot{V}O_2$) and carbon dioxide production ($\dot{V}CO_2$), heat production, food intake, whole body glucose disposal] and systemic markers for metabolic and inflammatory status.

At 22 months of age, mice were divided into four experimental groups. Two groups of animals were kept under sedentary conditions, of which one was subcutaneously injected with low doses (10 µg/kg) of murine rIL-6 (406-ML-200/CF, R&D Systems) three times per week, to mimic the transient increase of IL-6 plasma levels observed in response to single endurance exercise bouts (Sed+IL-6) and the other was sham injected with saline solution (Sed+Saline). Two additional groups of animals were engaged in a moderate-intensity, low-volume treadmill-training regimen, in one group paired with sham injections (Ex+Saline), and in the other group with rIL-6 administration (Ex+IL-6). All interventions lasted for a period of 12 weeks and until mice reached an age of 25 months in order to investigate the long-term effect of recurrent transient exposure to rIL-6 and/or exercise training. Experimental groups were generated in a semi-randomized manner by first randomly assigning mice to one of four groups and then equating for baseline body mass and maximum distance achieved in the maximum running capacity test (long duration incremental step protocol). Of note, group equalization was performed with the initial number of animals before starting the rIL-6 and/or exercise treatment and therefore prior to any dropouts or deaths occurring during the intervention.

Habitual activity and body temperature were monitored regularly throughout the study whereas gait, motor coordination, endurance performance and metabolic parameters were retested during the last third of the study. Of note, retests were scheduled as close as possible towards the end of the study but had to be distributed over the last 3 weeks due to feasibility reasons and in order to minimize the animal's stress level. After 12 weeks, the function of the *M. tibialis anterior* was evaluated *in situ* and

subsequently, mice were sacrificed to obtain terminal blood and to harvest limb muscles and other tissues. This was performed 48 h after the last injection and/or exercise session in order to minimize acute effects, as the goal of the current study was to assess long-term adaptation to training and/or rIL-6 treatment. Of note, data presented in this study resulted from three independent cohorts and most of the tests and measurements were only performed in subgroups of animals. Young sedentary mice were tested as a separate cohort at the age of 6 months.

Plasma sampling and analyses

Blood was drawn from the tail vein before (baseline) and after 12 weeks of the intervention in heparin coated tubes (Sarstedt). Tubes were centrifuged at 2000 g for 5 min at 4°C, supernatant transferred to a fresh tube, snap-frozen in liquid nitrogen and stored at –80°C until use.

Plasma cytokines were measured with a V-PLEX Proinflammatory Panel 1 Mouse Kit (Meso Scale Discovery) customized for five analytes (INF-γ, IL-1β, IL-6, IL-10, TNF-α) using a 2-fold dilution and following the manufacturer's instructions.

Terminal plasma levels of alanine transaminase (ALT), aspartate transaminase (AST), albumin, lactate dehydrogenase (LDH), total cholesterol, low-density lipoprotein (LDL), high-density lipoprotein (HDL), triglycerides and lipase were measured using the cobas c 111 analyzer (cobas®, Roche Diagnostics AG) using a 3-fold dilution in water and the manufacturer's test kits for the individual analytes.

Body composition

Body composition was assessed with an EchoMRI-100 analyzer (EchoMRI Medical Systems) in mice that did not carry a transponder. Mice were placed into a transparent measuring tube containing small holes for air exchange in the wall at the end of the tube. To minimize movement while the measurement is taking place, mice were minimally restrained by introducing another tube with a smaller diameter. Measurements were taken at baseline and the day before euthanasia and dissection.

Energy homeostasis measurements with the comprehensive laboratory animal monitoring system (CLAMS)

In brief, food intake, the rate of oxygen consumption ($\dot{V}O_2$) and carbon dioxide production ($\dot{V}CO_2$) and spontaneous movements were measured using the Comprehensive Laboratory Animal Monitoring System (CLAMS, Columbus Instruments) and Oxymax System (Columbus Instruments) for indirect calorimetry. After measuring body mass, each mouse was placed into a sealed chamber (21.3 cm x 11.6 cm x 11.5 cm) individually. Measurements were continuously taken for 94 h in an environment-

controlled cabinet chamber set at 23°C with a 12 h light to dark cycle (lights on at 06:00 am *i.e.*, at zeitgeber time 0; ZT0). During this time, food and water were provided to the animals through the feeding and drinking devices provided by the CLAMS system. The amount of food consumed by each animal was monitored through a precision balance installed under the chamber. A specific gas blend (Primary Standard Grade; 20.5% O_2 and 0.5% CO_2 in N_2) was used to calibrate the span or gain of both the O_2 and CO_2 sensors of the Oxymax system before each experiment and flow was set to 0.6 l/m. $\dot{V}O_2$ and $\dot{V}CO_2$ samplings were done sequentially in a 15 min interval. Respiratory exchange ratio (RER) was calculated as $\dot{V}CO_2/\dot{V}O_2$.

Glucose homeostasis

Mice were acclimatized to restraining a few days before the baseline measurement of glucose homeostasis in order to reduce the stress level during the experiment. Mice were fasted 12 h overnight before intraperitoneal injection of a glucose solution diluted in 0.9% NaCl (2 g glucose / kg body mass). Blood glucose levels were recorded from the tail vein with a glucose meter (Accu-Chek, Roche Diagnostics) at 0, 15, 30, 45, 60, 90 and 120 min after glucose injection. Area under the curve (AUC) was determined as incremental area above zero. In old mice, the same procedure was repeated in the 9^{th} week after the start of the intervention.

Body temperature and locomotor activity recordings

General locomotor activity and core body temperature data were acquired with the E-Mitter Telemetry System (STARR, Life Sciences Corp.) from animals placed with their home-cages in an environment-controlled cabinet (UniProtect Air Flow Cabinet, Bioscape). In short, small transponders (G2 E-Mitter, 15.5 mm x 6.5 mm, 1.1 g; STARR, Life Sciences Corp.) were implanted into the abdominal cavity of the mice under isoflurane anesthesia (2% isoflurane + O_2). Mice were treated with Meloxicam (1 mg/kg) pre- and post-operatively and allowed to recover for at least three weeks before the first measurement. Horizontal activity and core body temperature was recorded with a PC-based acquisition system connected to ER4000 Receivers (VitalView, STARR Life Sciences Corp.) for 3 to 4 day periods.

Gait analysis

The Noldus CatWalk XT system was used according to the CatWalk XT 10.6 Reference Manual and as previously described [1]. Before the baseline measurement, mice were habituated to the darkened experimental room and trained to cross the illuminated walkway for three consecutive days (with three trials per day). On the fourth day, each animal was tested by being placed at one extremity of the glass plate and being allowed to move freely back and forth the walkway. Run duration was

between 0.5 and 5 s with a maximum allowed speed variation of 60% to be considered as successful. The minimum number of compliant runs to acquire was set to three. The camera gain was adjusted to 20 dB and the detection threshold to 0.1 with a red ceiling and a green walkway light of 17.7 V and 16.5 V, respectively. In old mice, the same test procedure was applied during week 12 of the intervention, but without acclimatization. Runs recorded on the testing days were manually controlled for the following compliance criteria: runs were considered compliant if the animal walked continuously and straight *i.e.*, without turning the head to the side. One to three trials were used for the final analyses. Runs were analyzed by an independent investigator using the CatWalk XT software which measures diverse gait parameters such as those reported in this study *i.e.*, body speed (cm/s), cadence (number of steps/s), stride length (distance between successive placements of the same paw in cm), stand time (s) and swing speed (speed of the paw during swings, which is the duration in seconds of no contact of a paw with the glass plate), step cycle (s), base of support (distance between paws of fore or hind limbs in cm).

Evaluation of maximum running capacity (treadmill exhaustion test)

The running capacity test was performed on a six-lane open treadmill system (Columbus Instruments). Before the baseline test, mice were subjected to a two-day familiarization procedure. Day 1: mice were placed on the static treadmill band for 5 min followed by running at 5, 8 and 10 m/min for 5 min each at 5° inclination. Day 2: running at 5, 8, 10 and 12 m/min for 5 min each at 5° inclination. On the testing day (at least one day after acclimatization), mice were weighed and basal values for blood lactate were measured. Maximum running capacity was then evaluated by letting the mice run to exhaustion on an open treadmill with 5° inclination using a long duration incremental step protocol: 5 min at 5 m/min, 5 min at 8 m/min, 15 min at 10 m/min, followed by an increase of 2 m/min every 15 min until exhaustion (Fig. S5G). As older muscle shows a shift towards a higher proportion of fibers expressing slower myosin heavy chain isoforms compared to young muscle, a low intensity and long lasting protocol might be most suited to assess endurance performance in old mice. Exhaustion was determined by the animal failing to remain on the treadmill belt despite a mild electric stimulus (0.5 mA, 200 ms pulse, 1 Hz) and prodding. At this point, the mouse was taken off the treadmill to determine blood lactate levels immediately. Maximum running capacity was estimated using four parameters: the distance ran (m), the duration of the run (min), peak power [P_{peak} (W)] reached and total work [W_{tot} (J)] performed. Power was calculated as P_{peak} = body mass (kg) · gravity (9.81 m/s^2) · vertical speed (m/s · angle), and total work as W_{tot} (J) = body mass (kg) · gravity (9.81 m/s^2) · vertical speed (m/s · angle) · time (s). Mice of old cohorts were retested with the same protocol during week 12 of the intervention. Tests were performed in the early morning (lights on) and at least 48 h after the last training session and/or rIL-6 injection to minimize acute effects of the training and/or

injections. For the test, mice were randomly assigned to groups of 5 to 6 animals and the testers were blinded for their intervention group assignment.

Ramp-sprint test for $\dot{V}O_{2peak}$ measurements

Mice were familiarized with the Metabolic Modular Treadmill (Columbus Instruments) set to 5° slope on three consecutive days (in the early morning, lights on) prior to the baseline measurement. The familiarization period consisted of: Day 1, placing the mice in the treadmill for 10 min without speed followed by 5 min at 5 m/min; Day 2, running at 5 m/min, 7 m/min and 10 m/min for 5 min each; Day 3, running for 8 m/min, 10 m/min and 12 m/min for 5 min each. After one resting day, a short and high-intensity ramp-sprint test was performed in the early morning using a protocol optimized to measure $\dot{V}O_{2peak}$ [2]: After an 8 min resting-measurement followed by a 3 min warm-up at 8 m/min, speed was continuously increased by 0.03 m/min/s (*i.e.*, slowly ramped up) (Fig. S5A). Slope was set to 5°. Stainless steel grids at the end of each lane provided a mild electrical stimulus (0.5 mA, 200 ms pulse, 1 Hz) to keep the mice running. Mice ran until $\dot{V}O_2$ values plateaued and/or if the animal remained on the electrical grid for more than 5 s without any attempt to go back on the treadmill. Before being placed on the treadmill, mice were weighed and tail blood lactate (Lactate Plus meter, Nova Biomedical) values were determined. Lactate values were measured again immediately after mice met the above described abortion criteria. In old mice, a retest using the same protocol was performed during week 10 of the intervention. For the tests, mice were randomly assigned to groups of two animals. A specific gas blend (Primary Standard Grade; 20.5% O_2 and 0.5% CO_2 in N_2) was used to calibrate the span or gain of both the O_2 and CO_2 sensors of the Oxymax system before each test session. Flow was set to 0.6 l/min. $\dot{V}O_2$ and $\dot{V}CO_2$ were measured every 5 s.

Treadmill training and saline or rIL-6 injections

Mice were trained 3 times per week on an open treadmill (Columbus Instruments) in the early morning (between 06:00 am and 09:30 am, lights on). Training intensity was determined based on the average peak speed (v_{peak}) reached during the maximum running capacity test (long duration incremental step protocol). Maximum speed was 50% v_{peak} in the first training session and was gradually increased every following session until 80% v_{peak}, which was reached after 5 weeks. In each training session, the speed was slowly ramped up during the first 5-10 min up to the scheduled maximum speed of the given session. Mice were encouraged to run continuously on the belt without resting by gentle prodding. Training sessions lasted 45 to 50 min until all mice covered the same total distance. Mice were injected with a recombinant mouse IL-6 protein (406-ML-200/CF, R&D Systems) diluted in PBS (Dulbecco's Phosphate Buffered Saline, Sigma) to a dose of 10 µg/kg per injection (injection volume: 250-350 µl/mouse). Sham treated animals were injected with PBS (Dulbecco's Phosphate Buffered Saline,

Sigma) at volumes that would correspond to rIL-6 dilutions. All mice were injected three times per week in the early morning (lights on) and exercising mice received their injection 30 min prior to each training session to expose them to elevated IL-6 levels during and following the exercise bouts.

Evaluation of muscle function in situ

Function of the *M. tibialis anterior* was evaluated *in situ* using the 1300A Whole-Animal System (Aurora Scientific) as previously described [1]. In brief, mice were anesthetized using pentobarbital (Esconarkon; 90-150 mg/kg) and kept on a heating pad during the whole procedure. The hind limb skin was carefully removed under a microscope to expose the *M. tibialis anterior*. The distal tendon was freed from connective tissue, tied with square knots, then released and attached to the mechanical force transducer. The knee was immobilized by passing a syringe needle behind the patellar tendon. Stimulation electrodes were placed underneath the sciatic nerve, which was previously exposed at the hip. Exposed tendon, nerve and muscle were constantly kept moist with pre-warmed mineral oil (Sigma). In all experiments, 0.2 s pulses at 15 V were used and the muscle was allowed to rest after each stimulation. Prior to twitch and tetanic force measurements, two 100 Hz tetanic stimulations were applied to tighten the knots. Muscle optimal length (L_O) was then determined by applying successive single-twitch stimulations and adjusting basal muscle tension until maximal isometric twitch force (P_t). The muscle was then stimulated at different frequencies to obtain the force-frequency relationship and to determine absolute maximal tetanic force. To test fatigue resistance, the *M. tibialis anterior* was subjected to a series of successive 100 Hz tetanic stimulations delivered every 2 s for 4 min. Recovery from fatigue was assessed after 1, 2 and 3 min with a single 100 Hz tetanic stimulation. Muscle mass and Lo were used to determine cross-sectional area to obtain specific isometric twitch force (sP_t; kN/mm^2) and tetanic force (sP_O; kN/mm^2). Animals were euthanized at the end of the experiment by thoracic opening and by drawing terminal blood *via* heart puncture. AUC was determined as incremental area above zero for the fatigue curve and area of the curve (AOC) was determined for the following recovery period as incremental area above the terminal value of the fatigue protocol.

Grip strength in vivo

Muscle force was estimated *in vivo* by measuring peak force of whole limb grip (*i.e.*, fore and hind limbs together) using a Grip Strength Meter (Chatillon; Columbus Instruments) in the 11^{th} week. In brief, a mouse was held by the base of its tail between the thumb and the forefinger and placed with all four limbs on the angled mesh pull bar assembly connected to the force transducer. When the mouse was grasping properly with all four limbs, it was pulled along a straight line leading away from the sensor until the grip was released. A total of four trials were performed with at least 10 min break

between trials. Each trial consisted of a series of four pulls with a short latency between each pull, where the highest force attained during the pulls was recorded in kilogram-force (kgf; one kgf is equal to 9.806650 N). The median of all four trials was normalized to body mass.

Assessment of motor coordination (Rotarod test)

Mice were familiarized with the Rotarod instrument (Ugo basile, Model 47600) and taught to walk straight ahead without turning around on three consecutive days, by performing three 3 min trials in the fix mode with a constant speed of 4, 8 and 12 rpm on each day. On the test day, the Rotarod was set to acceleration mode (4 to 40 rpm within 5 min). Each mouse performed four consecutive test trials (with at least 10 min break between trials) and the time (s) as well as the speed (rpm) at which a mouse fell was recorded. Animals showing noncompliant behavior *i.e.*, animals clinging to the rod and completing passive rotations were not included into the analysis. The median of trials longer than 100 s was taken as readout. Old mice were tested at baseline (before the intervention) and during week 11 of the intervention.

Total muscle RNA extraction and quantitative real-time PCR

Total RNA was extracted from the *M. quadriceps femoris* of the unstimulated (*in situ*) left leg using a hybrid method as previously described [1]. In brief, TRI-Reagent (Sigma) was combined with on-column RNA purification with DNase treatment (RNease Mini Kit, Qiagen). RNA quantity and purity was evaluated using the NanoDrop OneC spectrophotometer (Thermo Scientific; A260/A280 and A260/A230) values were >1.9) and 1 µg of total RNA was reverse transcribed using the High Capacity cDNA RT Kit (Applied Biosystems). Quantitative real-time PCR was performed with Fast SYBR Green (Applied Biosystems) using the QuantStudio Real-Time PCR System (Applied Biosystems). PCR reactions were done in duplicate with the addition of negative controls (*i.e.*, no reverse transcription and no template controls). Relative gene expression levels were determined using the comparative ΔΔCT method to normalize target gene mRNA to *Tbp* and to the average of the control group (Sed+Saline). Primer sequences are summarized in Table S2.

Protein extraction and immunoblotting

Equal amounts (30 mg) of frozen powdered tissue of the unstimulated (*in situ*) *M. quadriceps femoris* was homogenized in 300 µl of ice cold RIPA buffer [150 mM NaCl, 1% v/v Nonidet-P40 substitute, 0.2% v/v Na-deoxycholate, 0.1% v/v SDS, 50 mM TRIS-HCl (pH 7.5), 1 mM EDTA, 0.5 mM EGTA, 1 mM DTT, 10 mM Nicotinamide] containing freshly added protease inhibitors (cOmplete mini EDTA free, Roche) and phosphatase inhibitors (PhosStop, Roche) using a bead homogenizer and metal beads [Qiagen, stainless steel 5 mm (200 pc)] and incubated for 1 h at 4°C. Next, tubes were centrifuged at 16’000 g

for 10 min at 4°C and the supernatant was transferred to a fresh tube. Protein concentration was measured with the Bradford protein assay (Quick Start Bradford 1x Dye reagent, BioRad) and BSA standards (Pierce). Samples were diluted in RIPA buffer and 4X Laemmli Sample Buffer (BioRad) with 2-Mercaptoethanol (Sigma) and boiled for 5 min at 95°C or at 50°C for OXPHOS blots. Equal amounts of protein were separated on 4-20% Mini-Protean TGX Precast Protein Gels (BioRad) and transferred on Nitrocellulose membranes (Amersham Protran 0.45 NC, 10600007, 0-45 µm). Membranes were blocked for 1 h in 5% milk in Tris-buffered saline + 0.1% Tween 20 (TBS-T) before overnight incubation at 4°C and shaking with the following antibodies: LDH-H (1:5000 in 5% BSA, TBS-T, NB110-57160, NOVUS Biologicals), Sarcomeric α-Actinin (1:5000 in 5% BSA, TBS-T, A7732, Sigma-Aldrich), Total OXPHOS Rodent WB Antibody Cocktail (1:1000 in 1% milk, PBS, ab110413, Abcam), eEF2 (1:1000 in 5% BSA, TBS-T #2332, Cell Signaling), NDUFB5 (1:500 in 5% BSA, TBS-T, #PA5-98000, Invitrogen Thermo Fisher Scientific), PDK4 P-24 (1:500 in 5% BSA, TBS-T, sc-120841, Santa Cruz), STAT3 (1:1000 in 5% milk, TBS-T, #9139, Cell Signaling). After washing, membranes were incubated with peroxidase-conjugated secondary antibodies (1:10'000 in TBS-T, Dako) for 1 h at room temperature, followed by antibody binding detection using chemiluminescence horseradish peroxidase substrate detection kits (ECL, Supersignal West Dura or Femto; Pierce) and a Fusion FX imager (Viliber). Quantification was done with the Fusion FX software. Relative protein levels were determined by normalizing the band intensity of the target to the loading control (α-Actinin, eEF2 or Ponceau S stain) and to the average of the control group (Sed+Saline) within one gel.

Statistical analyses

Data are presented as mean ± standard error of the mean (SEM) including individual values where indicated, mean and individual values, or as box and whisker plots. In the latter, boxes depict the 25th and 75th percentiles (upper and lower perimeters), the median (midline) and the mean (cross). Whiskers are plotted based on the interquartile range (IQR) *i.e.*, the difference between the 25th and 75th percentiles. The upper whisker is drawn to the largest value in the data set that is smaller than (or equal to) the 75th percentile plus 1.5 times IQR (upper fence) and the lower whisker is drawn to the smallest value in the data set that is greater than (or equal to) the 25th percentile minus 1.5 times IQR (lower fence). Any values greater than the upper or smaller than the lower fence are plotted as individual points. Data were represented and analyzed using the GraphPad Prism 8.0 software (GraphPad Software, San Diego, CA, USA). The *n* number used for each experiment is indicated in the figure legends. Within group comparisons of baseline test and retest measurements were performed with a paired Student's *t*-test. In this case, paired values are indicated by connecting individual values with a black line. For between group comparisons, either one-way or two-way ANOVA was performed followed by Sidak's multiple comparisons as post-hoc test. Dropout curves of the distance ran (m)

during the maximum running capacity test were plotted as Kaplan-Meier plots and were compared using a logrank test for trend. This test is using chi-square to compute a *P* value by testing the null hypothesis that there is no linear trend between column order (*i.e.*, the order of plotted curves from left to right) and survival (*i.e.*, keeping on running). The level of significance was set at $P < 0.05$ for all statistical tests.

References

1. Delezie J, Weihrauch M, Maier G, Tejero R, Ham DJ, Gill JF, et al. BDNF is a mediator of glycolytic fiber-type specification in mouse skeletal muscle. Proc Natl Acad Sci U S A. 2019;116:16111-20.
2. Ayachi M, Niel R, Momken I, Billat VL, Mille-Hamard L. Validation of a Ramp Running Protocol for Determination of the True VO2max in Mice. Front Physiol. 2016;7:372.

Supplementary figure legends

Figure S1 Equal metabolic parameters under non-exercise conditions

(A) Oxygen uptake ($\dot{V}O_2$), (B) carbon dioxide uptake ($\dot{V}CO_2$), (C) respiratory exchange ratio (RER; $\dot{V}CO_2/\dot{V}O_2$) and (D) heat production over 48 h (left panels; ZT: zeitgeber time) and averaged for light and dark phase of this period (right panels) during week 11 of the intervention (Sed+Saline, *n*=13; Sed+IL-6, *n*=13; Ex+Saline, *n*=10; Ex+IL-6, *n*=15). Data are presented as mean ± SEM (left panels A-D) including individual values (right panels in A-D).

Figure S2 Normal blood biochemistry in rIL-6 treated mice

(A) Blood plasma levels of alanine aminotransferase (ALT), aspartate aminotransferase (AST), albumin, lactate dehydrogenase (LDH), total cholesterol, low-density lipoprotein (LDL), high-density lipoprotein (HDL), triglycerides and lipase for all four groups after 12 weeks of treatment. (B) Plasma levels of interferon-gamma (INF-γ) and interleukin-beta (IL-β) before (Baseline) and at the end (Terminal) of the study and of an independent cohort of young sedentary mice (Sed+Saline, *n*=15; Sed+IL-6, *n*=15; Ex+Saline, *n*=13; Ex+IL-6, *n*=16; Young Sed, *n*=7). (B) Y-axis has a logarithmic scale (log 10). Data are presented as box and whiskers (A) or individual paired values connected with a black line (B). One-way ANOVA followed by Sidak's multiple comparisons (A), paired Student's *t*-test and one-way ANOVA followed by Sidak's multiple comparisons of Terminal measures of old groups and the Young Sed group (B). $^{*}P < 0.05$. In (J and K): $^{\#\#}P < 0.01$, $^{\#\#\#}P < 0.001$ old groups *vs*. Young Sed group.

Figure S3 Moderate-intensity low-volume endurance training preserves glucose tolerance with aging

Relative change from baseline area under the curve (Sed+Saline, *n*=10; Sed+IL-6, *n*=10; Ex+Saline, *n*=8; Ex+IL-6, *n*=11; Young Sed, *n*=12).

Figure S4 Effects of moderate-intensity low-volume endurance training and/or rIL-6 treatment on body composition

(A) Absolute fat mass at baseline and after 12 weeks of the intervention (left) and relative fat mass change from baseline to 12 weeks in percent (right). (B) Absolute lean mass at baseline and after 12 weeks (left) and relative lean mass change from baseline to 12 weeks in percent (right). (C) Percent body fat at baseline and after 12 weeks. (A-C) Sed+Saline, *n*=6; Sed+IL-6, *n*=7; Ex+Saline, *n*=6; Ex+IL-6, *n*=6. (D) Fat depot masses upon dissection of epididymal white adipose tissue (eWAT) and anterior subcutaneous white adipose tissue (sWAT). (E) Individual muscle masses of *M. quadriceps femoris* (QUAD), *M. gastrocnemius* (GAS), *M. triceps brachii* (TRI), *M. extensor digitorum longus* (EDL), *M. soleus* (SOL) and *M. plantaris* (PLAN) upon dissection (Sed+Saline, *n*=15; Sed+IL-6, *n*=16; Ex+Saline,

n=13; Ex+IL-6, *n*=16). Data are presented as mean and individual paired values connected with a black line (left panel in A and B and in C), box and whiskers (right panel in A and B), or individual values and mean ± SEM (F and G). Paired Student's *t*-test (left panel in A and B and in C) and one-way ANOVA followed by Sidak's multiple comparisons (individual tissues in D). *P < 0.05, **P < 0.01, ***P < 0.001 as indicated.

Figure S5 Treadmill protocols and additional corresponding measurements

(A) Graphical representation of the ramp-sprint protocol used to assess peak oxygen uptake ($\dot{V}O_{2peak}$). (B) Blood lactate levels before (Basal) and after (Exhausted) the ramp-sprint test at baseline and after 10 weeks (retest) in the old groups. (C) Peak speed reached and (D) distance covered until exhaustion during the retest (10 weeks) in old groups and in an independent young sedentary group. (E) Relative $\dot{V}O_{2peak}$ change from baseline (–3 weeks) to 10 weeks in percent. (F) Increase in blood lactate expressed as the difference exhausted-basal (Δ) during the retest (10 weeks) and in young sedentary mice. (B-F) Sed+Saline, *n*=6; Sed+IL-6, *n*=7; Ex+Saline, *n*=6; Ex+IL-6, *n*=5; Young Sed, *n*=14. (G) Graphical representation of the long duration incremental step protocol used to evaluate running capacity. (H) Peak power (P_{peak}) achieved at baseline and after 12 weeks in old groups and in young sedentary mice. (I) Blood lactate levels in basal and exhausted state. (H and I) Sed+Saline, *n*=15; Sed+IL-6, *n*=16; Ex+Saline, *n*=13; Ex+IL-6, *n*=15; Young Sed, *n*=14. Data are presented as mean and paired individual values connected with a black line (A and B), mean ± SEM (E) including individual values (H), and box and whiskers (C, D, E, F). One-way ANOVA followed by Sidak's multiple comparisons (B-F, H and I) and paired Student's *t*-test (H). In (B and I) ***P < 0.001 basal *vs.* exhausted within group; $^{\#}P$ < 0.05, $^{\#\#\#}P$ < 0.001 Baseline exhausted *vs.* 12 weeks exhausted within group; In (C, D, F and H) $^{\#}P$ < 0.05 $^{\#\#}P$ < 0.01 $^{\#\#\#}P$ < 0.001, Young Sed *vs*. older groups. In (E) * or $^{\#}P$ < 0.05, Ex-IL-6 *vs*. Sed+Saline or Sed+IL-6; ** or $^{\#\#}P$ < 0.01 Ex+Saline *vs*. Sed+Saline or Sed+IL-6. In (E) $^{\delta}$P < 0.05, $^{\delta\delta\delta}$P < 0.001 Sed+Saline 12 weeks exhausted *vs.* Ex+Saline or Ex+IL-6 12 weeks exhausted.

Figure S6 No effect of moderate-intensity low-volume endurance training and/or rIL-6 treatment on contractile properties in situ and grip strength in vivo

(A) Area of the curves (AOCs) of the recovery period following the fatigue protocol. (B) Normalized *in vivo* muscle force estimated by measuring peak force of whole limb grip (kgf, kilogram-force; one kgf is equal to 9.806650 N). Electrical sciatic nerve stimulation evoked (C) absolute muscle force-frequency relationship, (D) specific twitch force and (E) specific tetanic force. (F) Masses and lengths of the *M. tibialis anterior* used to calculate specific forces. (G) Single twitch time-to-peak tension and (H) single twitch half-relaxation time. Sed+Saline, *n*=14; Sed+IL-6, *n*=15; Ex+Saline, *n*=12; Ex+IL-6, *n*=14, Young Sed, *n*=9. Data are presented as box and whiskers (A, B, D, E, G and H), mean ± SEM (C) including

individual values (F). One-way ANOVA followed by Sidak's multiple comparisons (C). $^{\#}P < 0.05$ Young Sed *vs*. old groups.

Figure S7 Additional parameters obtained with the CatWalk voluntary gait analysis system and Rotarod test

(A-H) Additional parameters of quantitative voluntary gait analysis at baseline and after 12 weeks of the intervention and an independent cohort of young sedentary mice: (A-F) Swing speed (cm/s), time (s) of swing phase and time (s) of stand phase of hind limbs (A-C) and fore limbs (D-F). (G) Average time (s) of a step cycle and (H) base of support *i.e.*, distance (cm) between paws of hind limbs. (A-H) Sed+Saline, *n*=9; Sed+IL-6, *n*=9; Ex+Saline, *n*=7; Ex+IL-6, *n*=11; Young Sed, *n*=11. (I) Motor coordination assessed by challenging mice with an accelerated Rotarod test at baseline and in the 11th week of the intervention (Sed+Saline, *n*=8; Sed+IL-6, *n*=9; Ex+Saline, *n*=7; Ex+IL-6, *n*=11; Young Sed, *n*=14). Data are presented as mean and individual paired values connected with a black line. Paired Student's *t*-test and one-way ANOVA followed by Sidak's multiple comparisons of retest of old groups and young sedentary group. $^{*}P < 0.05$, $^{**}P < 0.01$; $^{\#}P < 0.05$, $^{\#\#}P < 0.01$, $^{\#\#\#}P < 0.001$ Young Sed *vs*. old groups.

Figure S8 Expression of fatty acid and glucose metabolism related genes

Relative *M. quadriceps femoris* mRNA levels of genes involved in (A) fatty acid transport, synthesis and oxidation and (B) glucose metabolism (Sed+Saline, *n*=9; Sed+IL-6, *n*=9; Ex+Saline, *n*=7; Ex+IL-6, *n*=11). Expression values were determined by qPCR and normalized to *Tbp*. Data are presented as the mean fold-change ± SEM relative to the expression of the control (Sed+Saline) set to 1. (C) Relative *M. quadriceps femoris* total STAT3 protein levels assed by western blot (*n*=6 per group). Target band intensities were normalized to the loading control (Ponceau S stain) and data are shown as the mean fold-change ± SEM including individual values relative to the control (Sed+Saline) set to 1. One-way ANOVA followed by Sidak's multiple comparisons (C). $^{**}P < 0.01$, $^{***}P < 0.001$.

Table S1. Dying and excluded animals

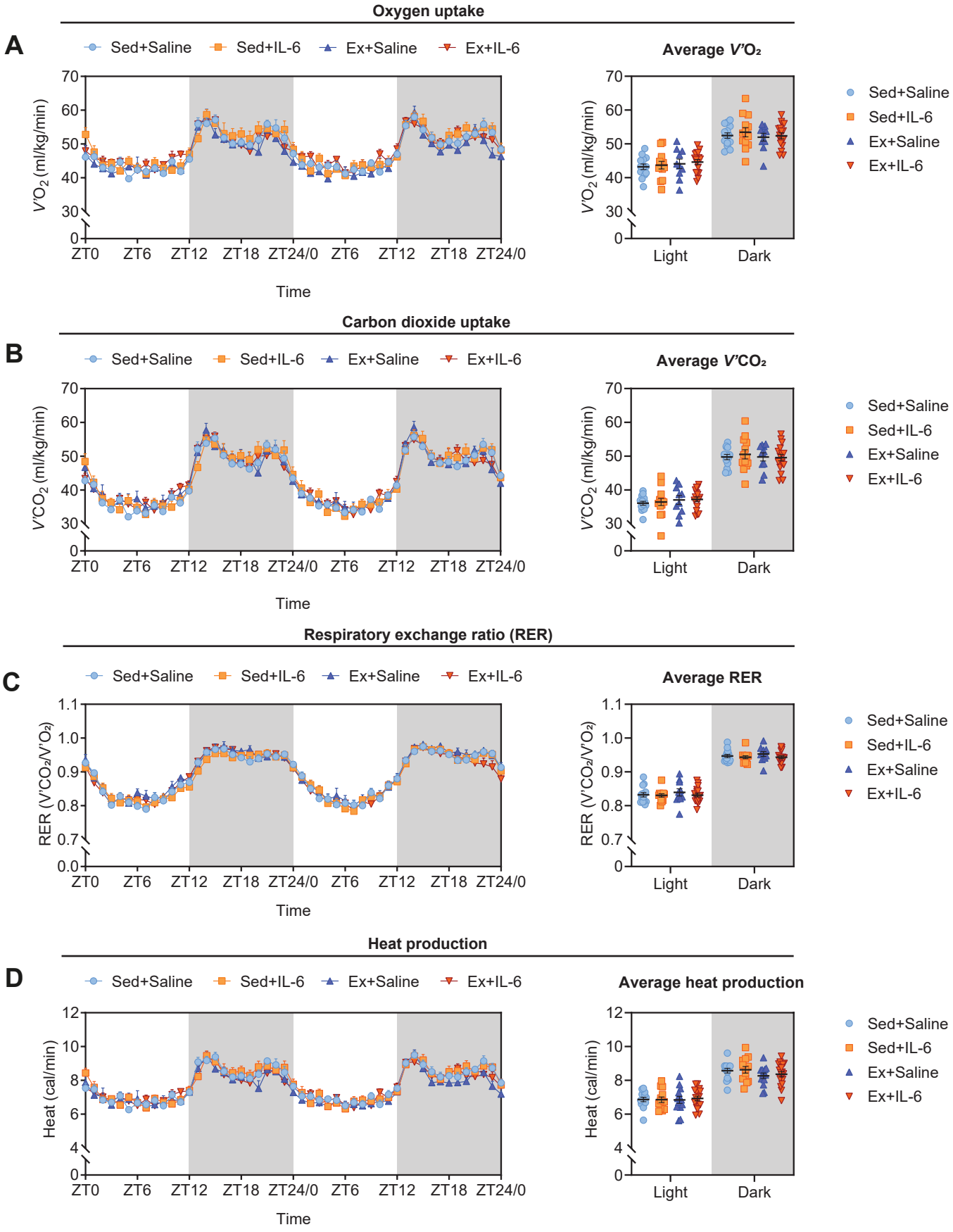
	Sed+Saline	Sed+IL-6	Ex+Saline	Ex+IL-6
Week 0 (*n*)	19	17	18	19
Week 12 (*n*)	15	16	13	16
Number of deaths	2	0	5	1
Deaths (%)	13	0	38	6
Limb joint problem	1			1
Tumor	1			1
Dermatitis		1		
Total exclusions	2	1	0	2

Table S2. Primer sequences used for qPCR

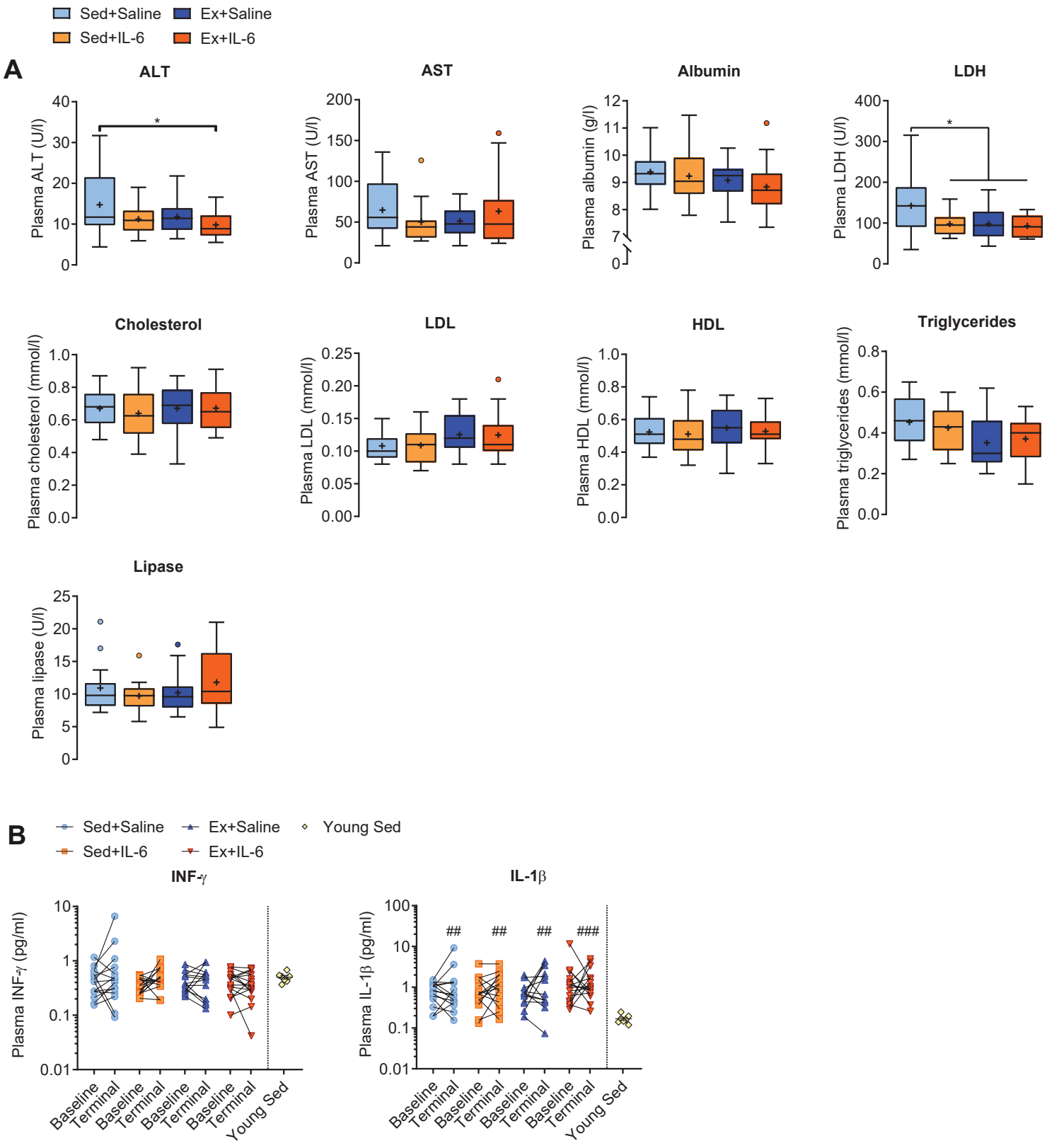
Gene name	Forward primer	Reverse primer
Acadm	AACACTTACTATGCCTCGATTGCA	CCATAGCCTCCGAAAATCTGAA
Acc1	AAGGCTATGTGAAGGATG	CTGTCTGAAGAGGTTAGG
Acc2	GGGCTCGGGCATGATTG	CAGGTAAGCCCCGATTCCA
Atp5	TCTCCATGCCTCTAACACTCG	CCAGGTCAACAGACGTGTCAG
Atp6	ACTATGAGCTGGAGCCGTAATTACA	TGGAAGGAAGTGGGCAAGTG
Cd36	AAGAGGTCCTTACACATACAGAGTTC	AGCTGCTACAGCCAGATTCA
Cpt1b	GCACACCAGGCAGTAGCTTT	CAGGAGTTGATTCCAGACAGGA
Cox1	TGCTAGCCGCAGGCATTACT	GCGGGATCAAAGAAAGTTGTG
Cs	CCCAGGATACGGTCATGCA	GCAAACTCTCGCTGACAGGAA
Cyc	ACAAGAAGACTCAAATGTGTTTCAGTTT	TGCACTGTCAAGAATAGACAGTTGC
Fh1	TGCTCTCAGTGCAAAATCCAA	CGTGTGAGTTCGCCCAATT
Glut4	CATGGCTGTCGCTGGTTTC	AAACCCATGCCGACAATGA
Gyg1	CGGCCACACTATGACAGATCA	TGTGGGCTGGTGAGTACAAC
Gys1	TACACTGTGCTGCAGACGAA	CGCCCAAAATACACCTTACAAC
Hadha	TGCATTTGCCGCAGCTTTAC	GTTGGCCCAGATTTCGTTCA
Hadhb	TGCTGTCAGGCACTTCGTATAAA	AAACCCGAAAGTGCAGCTCTAG
Hk2	CCCTGCCACCAGACGAAA	GACTTGAACCCCTTAGTCCATGA
Mlycd	ACTCCATCAGCCTGACCCAG	ACCCCTTGAGGCTCTCGTGA
Ndufa5	ACATGCAGCCTATAGAAAATACACAGA	TCCGCCTTGACCATATCCA
Ndufb5	TTTTCTCACGCGGAGCTTTC	TGCCATGGTCCCCACTGT
Ndufb8	CAAGAAGTATAACATGCGAGTGGAA	CCATACCCCATGCCATCATC
Pdha	GAAGATGCTTGCCGCTGTATC	CCGATGAAGGTCACATTTCTTAAT
Pdk4	AAAATTTCCAGGCCAACCAA	CGAAGAGCATGTGGTGAAGGT
Pdp1	CGGGCACTGCTACCTATCCTT	ACAATTTGGACGCCTCCTTACT

Pfkm	TGTGGTCCGAGTTGGTATCTT	GCACTTCCAATCACTGTGCC
Pygm	GGTTTATGGTGCCGAGGACT	GGCGGCGGGAATAACTTTCT
Pgm1	AAAAATCAGGGCTTGCGGC	TGGCAACGTCCTTCTCGTAG
Sdha	GCTGGTGTGGATGTCACTAAGG	CCCACCCATGTTGTAATGCA
Tbc1d1	CATAAAGAACACACTCCCCAACCT	TGCTTGGCGATGTCCATCT
Tbc1d4	GTACCGACCGGATATGATGTCA	CGGTGGTAGTCATGAAGGAGTCT
Tbp	TGCTGTTGGTGATTGTTGGT	CTGGCTTGTGTGGGAAAGAT
Uqcrc2	CCCATCTTGCTTTGCTGTCTG	AATAAAATCTCGAGAAGGACCCG

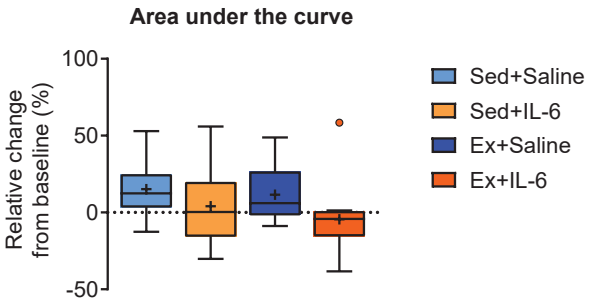
Supplementary figure 1



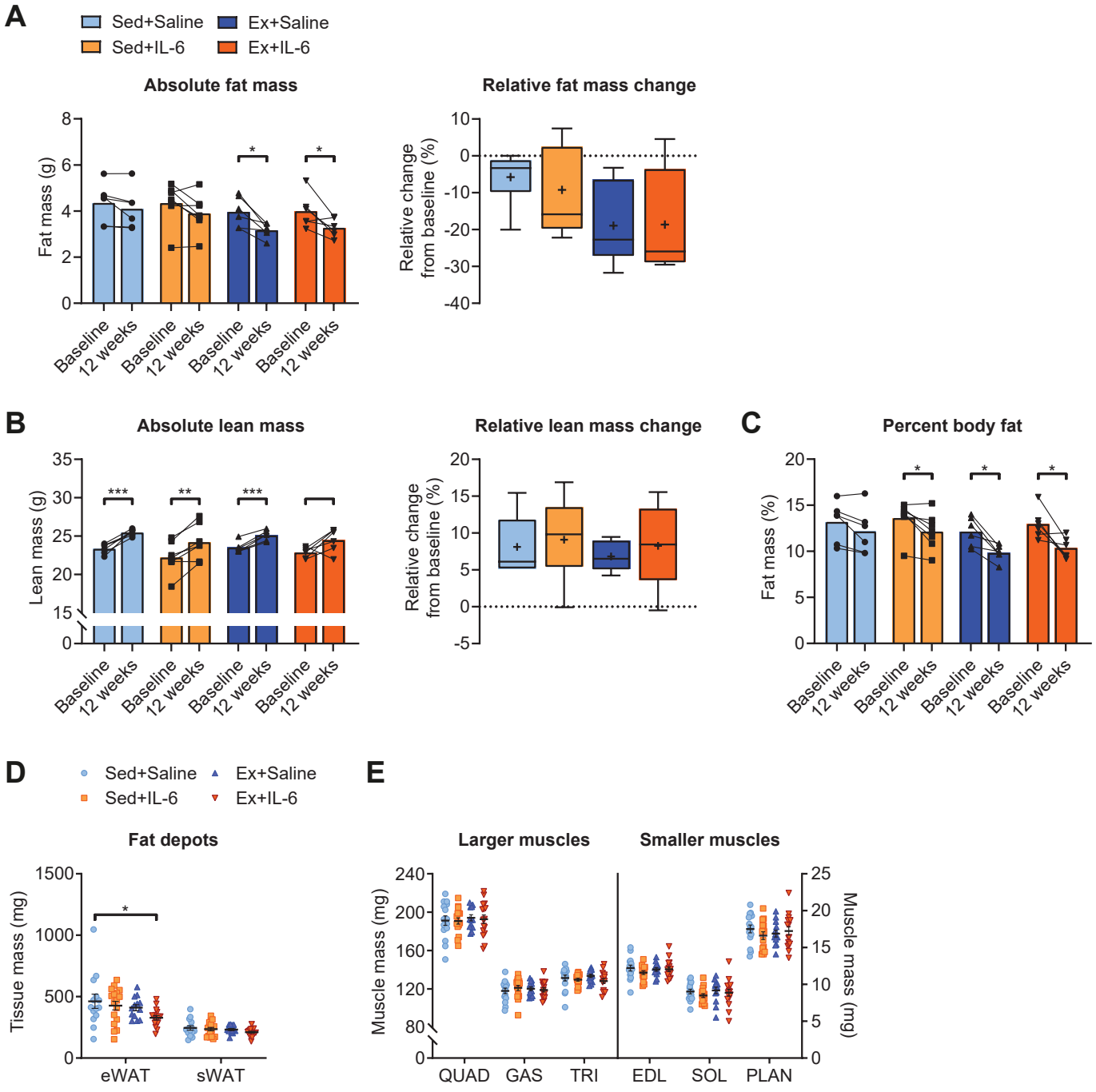
Supplementary figure 2



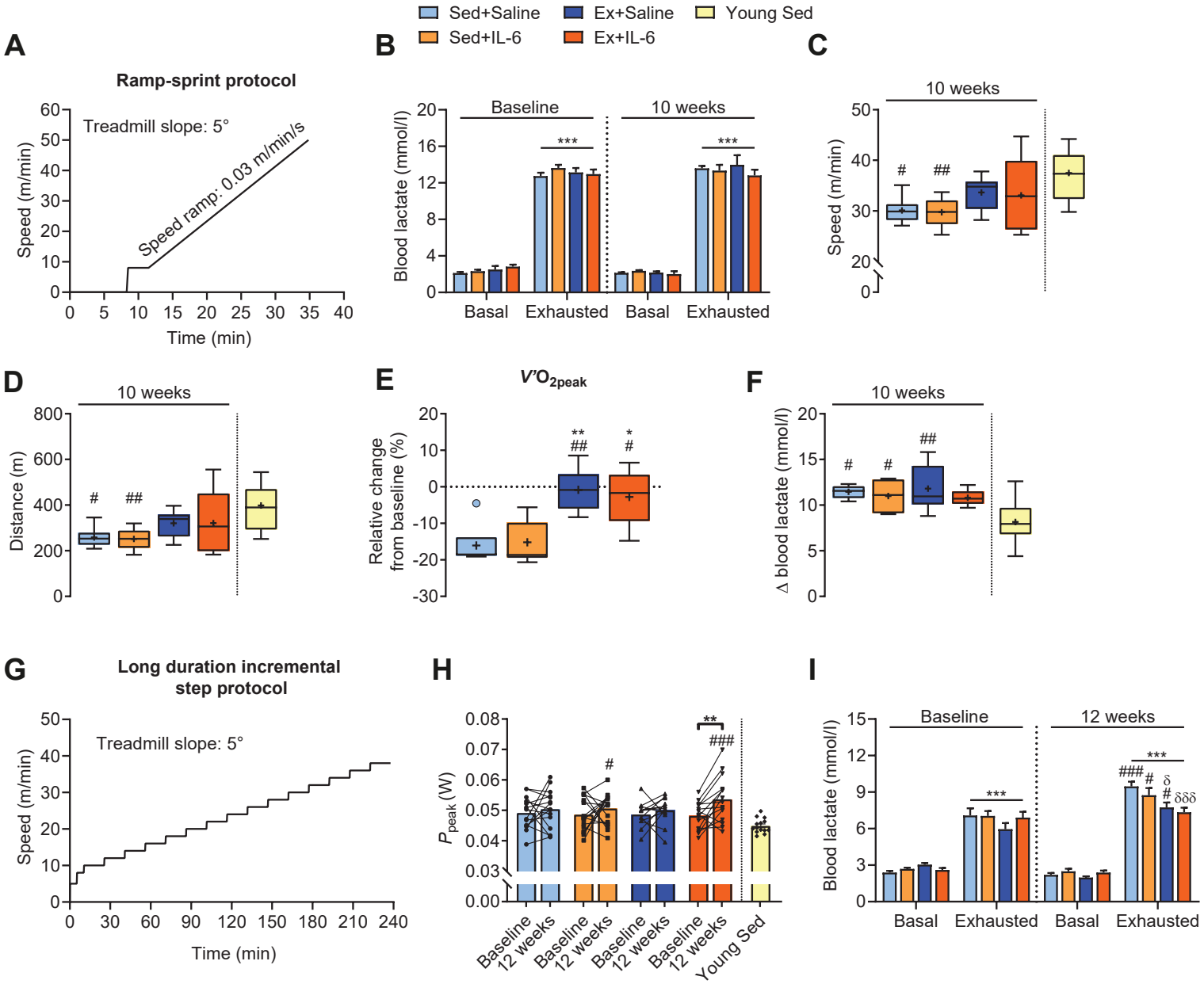
Supplementary figure 3



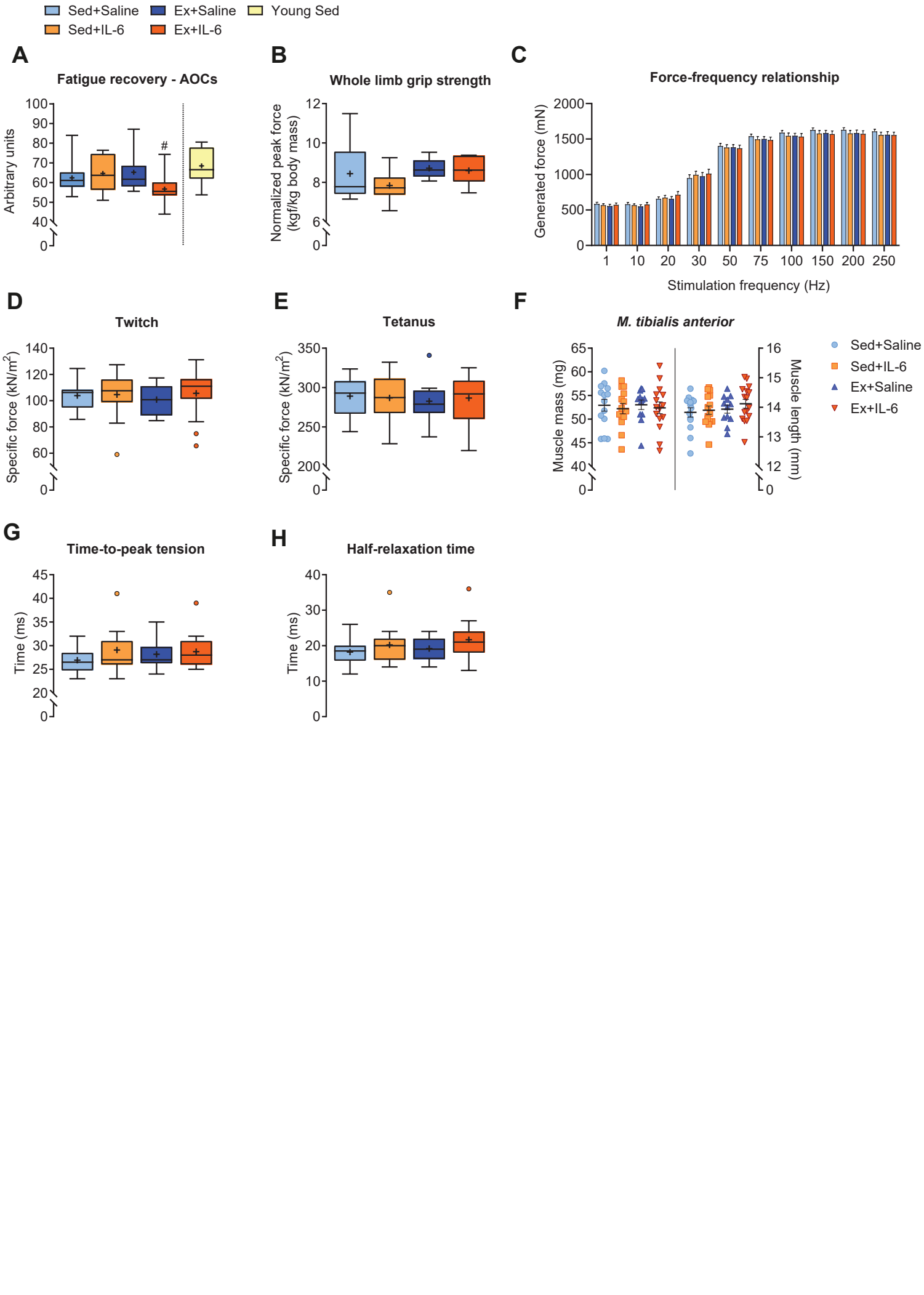
Supplementary figure 4



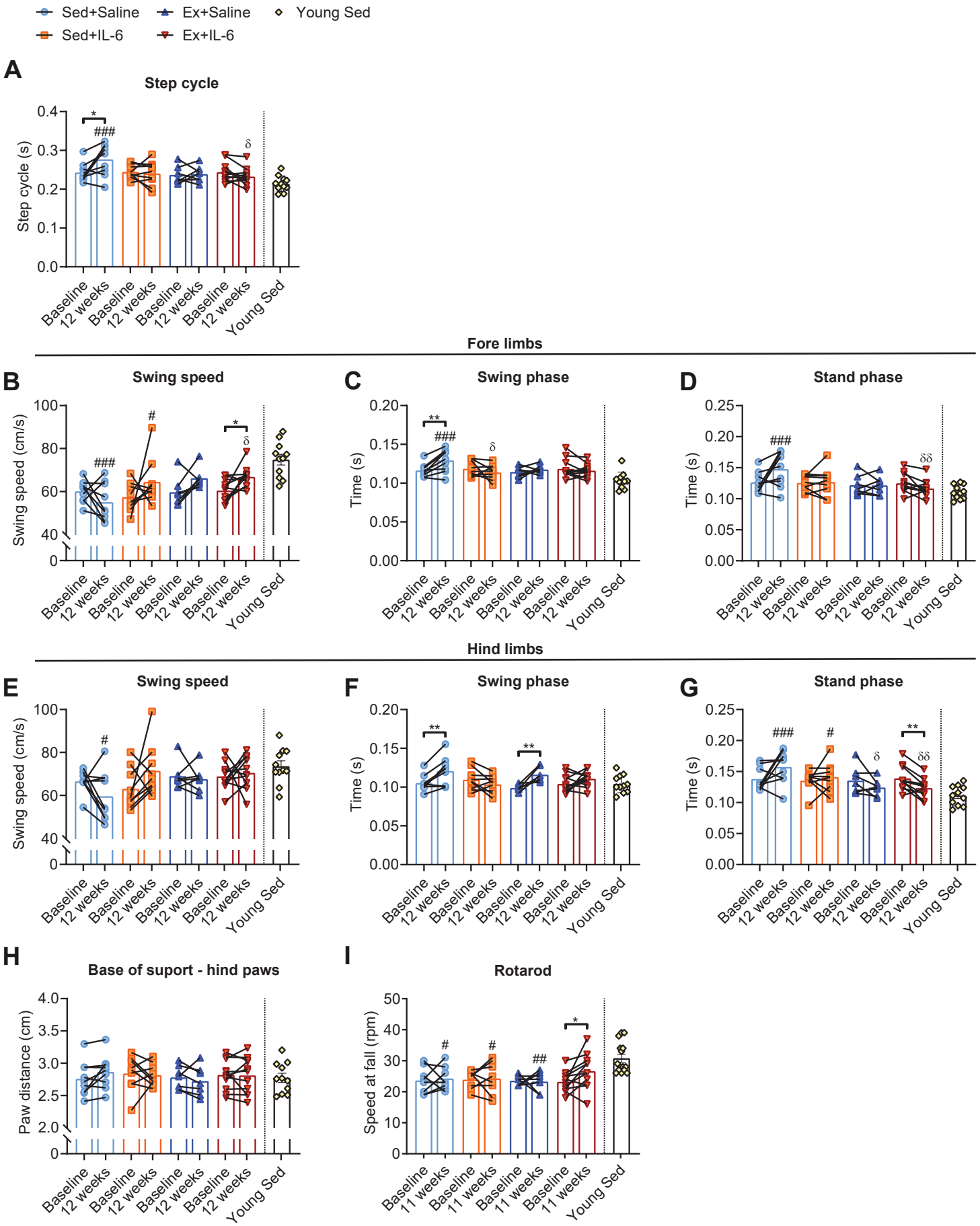
Supplementary figure 5



Supplementary figure 6



Supplementary figure 7



Supplementary figure 8