Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Drosophila connectome prior enables adaptable embodied control

Siyu Yu
Tsinghua University
Kanglei Zhou
Tsinghua University
Hongwei Yan
Tsinghua University
Xinyi Wang
Tsinghua University
Sen Song
Tsinghua University
Bo Hong
Tsinghua University
Liyuan Wang
liyuanwang@tsinghua.edu.cn

Tsinghua University

Research Article

Keywords: embodied intelligence, connectome, locomotion control, adaptability

Posted Date: November 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11557476/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

A *Drosophila* connectome prior enables adaptable embodied control

Siyu Yu[1,2,3], Kanglei Zhou[1,3], Hongwei Yan[3,4], Xinyi Wang[1], Sen Song[2,3], Bo Hong[2,3] and Liyuan Wang[1,3*]

[1]Department of Psychological and Cognitive Sciences, Tsinghua University, Beijing, China.
[2]School of Biomedical Engineering, Tsinghua University, Beijing, China.
[3]IDG/McGovern Institute for Brain Research, Tsinghua University, Beijing, China.
[4]School of Life Sciences, Tsinghua University, Beijing, China.

*Corresponding author(s). E-mail(s): liyuanwang@tsinghua.edu.cn;
Contributing authors: yusy23@mails.tsinghua.edu.cn; zhoukanglei@tsinghua.edu.cn; yanhw22@mails.tsinghua.edu.cn; wangxinyi@aibrainsong.com; songsen@tsinghua.edu.cn; hongbo@tsinghua.edu.cn;

Abstract

Embodied intelligence requires effective control as environments, movement demands and body morphologies change. Biological nervous systems support such adaptability through highly structured connectivity, yet whether evolved neural wiring itself confers a computational advantage remains unclear. Here we use the adult *Drosophila* brain connectome as a fixed recurrent structural prior for closed-loop embodied control. In a neuromechanical digital fly, connectome-based and multilayer-perceptron controllers perform comparably during nominal locomotion but diverge as control demands change. Across vertical terrain negotiation, lateral direction turning and longitudinal speed modulation, the connectome exhibits a consistent advantage, improving obstacle negotiation and speed tracking while supporting smoother directional reconfiguration. Mechanistic analyses reveal spontaneous recruitment of behaviourally relevant sensorimotor pathways and distinct local policy geometry associated with robust adaptation. The same connectome prior improves locomotion after transfer to a quadrupedal body and outperforms representative artificial controllers in maze navigation,

with its advantage widening as maze complexity increases, while remaining effective across two-, three-, four- and six-legged embodiments. These results identify adaptability as a functional advantage of connectome structure and establish evolved neural wiring as a transferable prior for adaptive embodied control.*

Keywords: embodied intelligence, connectome, locomotion control, adaptability

1 Introduction

Intelligence is expressed through a body. Animals continuously transform sensory and internal states into actions through closed-loop interactions among neural circuits, biomechanics and changing environments [1–3]. Locomotion makes this coupling especially explicit, requiring an agent to coordinate many degrees of control while continually adjusting where to move, how fast to move and how to interact with the surrounding physical world. Biological nervous systems are naturally evolved controllers whose defining strength is *adaptability*: effective behaviour is maintained as these demands change. Artificial agents can now acquire impressive motor skills, including agile locomotion over challenging terrain [4–7], but such capabilities commonly rely on controllers optimized for particular bodies and training distributions, often supplemented by extensive domain randomization, explicit adaptation mechanisms or morphology-specific designs. Understanding how evolved neural wiring supports adaptable embodied control across changing environments, movement demands and body morphologies is therefore a shared challenge for neuroscience and embodied intelligence.

Connectomics provides an unusually direct route to the neural organization underlying embodied control. Synapse-resolution reconstructions have progressed from the complete larval *Drosophila* brain [8] to the adult brain [9], the adult ventral nerve cord (VNC) [10], and, most recently, an integrated brain-cord connectome spanning sensory input to motor output [11]. These datasets now support executable models in which measured biological wiring explicitly constrains computation. Connectome-constrained networks can reproduce neural responses of the fly visual system [12]. Whole-brain models based on synaptic connectivity and neurotransmitter identity can predict sensorimotor transformations underlying feeding and grooming [13]. Also, increasingly realistic neuromechanical platforms enable closed-loop simulation of sensory processing, navigation and whole-body locomotion in digital flies [14, 15], and recent work has instantiated the adult fly connectome itself as a trainable locomotor controller [16]. These advances establish that connectomic organization can support functional computation and embodied behaviour, but leave unresolved *what advantage its evolved neural wiring confers*.

We hypothesize that this advantage lies in *adaptability*. The fly nervous system embeds local sensorimotor circuits within recurrent ascending and descending pathways and distributed behaviour-related circuitry [10, 11], an organization well suited

*The project webpage: https://thu-neuroml.github.io/FlyControl/

to coordinating behaviour as sensorimotor demands change. We test this hypothesis by using the adult *Drosophila* brain connectome as a fixed, evolved recurrent structural prior for closed-loop control of a neuromechanical digital fly (Fig. 1). During nominal straight, constant-speed locomotion on level ground, connectome-based and multilayer-perceptron (MLP) controllers perform comparably, suggesting little benefit when control demands remain unchanged (Fig. 2a–d). The advantage emerges across three dimensions of adaptation: vertical terrain negotiation, lateral direction turning and longitudinal speed modulation. Across these settings, the connectome improves obstacle negotiation and speed tracking while supporting smoother directional reconfiguration. Mechanistic analyses reveal spontaneous recruitment of mechanosensory, ascending and descending pathways, together with distinct local policy geometry under connectome-constrained control.

We ask whether this advantage extends beyond the body in which the connectome evolved. Coupled to newly learned interfaces on a quadrupedal Ant, the same fly connectome improves locomotor speed and path efficiency even on flat ground, indicating that embodiment change alone is sufficient for a benefit to emerge. During spatial navigation, the connectome outperforms diverse artificial controllers, with its success advantage widening from U-Maze to Medium and Large mazes. Extending the same prior across two-, three-, four- and six-legged embodiments further demonstrates transfer across distinct body plans. The hexapod achieves the highest absolute performance, consistent with its closer morphological correspondence to the native fly body, whereas the largest relative gain emerges in the biped, the most morphologically divergent embodiment, with clear benefits also retained in other embodiments. These results reveal adaptability as a core computational advantage of evolved neural wiring and point to its broader potential for embodied intelligence.

2 Results

2.1 Connectome-derived recurrent prior for adaptable embodied control

We use the adult *Drosophila* brain connectome as the principal substrate for embodied control [9]. Its directed synaptic organization defines a fixed recurrent core, while trainable sensory and motor interfaces map task observations into the network and recurrent activity into actions. Recurrent connectivity is held fixed throughout learning, allowing us to isolate how biological connectomic organization shapes motor outcomes. We later examine the larval brain [8] and the integrated adult brain-VNC connectome [10, 11] as complementary biological comparisons.

We instantiate this controller in a closed-loop neuromechanical fly built on recent whole-body simulation platforms [14, 15] (Fig. 1a, Methods). Body state, contact information, central pattern generator (CPG) signals, and task-specific variables enter through the learned sensory interface, while recurrent activity is mapped through the motor interface and combined with CPG-generated stepping patterns to control the fly's 42 leg joints and adhesion. Body-environment interactions update proprioceptive and contact feedback, closing the sensorimotor loop. This places measured connectomic wiring within an executable whole-body control system, extending

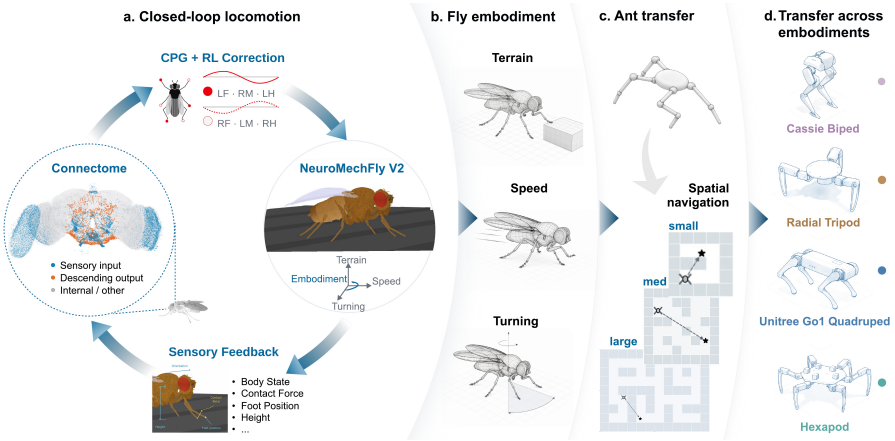


Fig. 1 Connectome-derived recurrent prior across locomotion demands and embodiments. a, Closed-loop control of a neuromechanical digital fly, with the adult *Drosophila* brain connectome providing a fixed recurrent controller that generates reinforcement learning (RL)-based corrections to central pattern generator (CPG)-driven locomotion in NeuroMechFly v2. Body-state, contact-force, foot-position and height feedback close the sensorimotor loop. LF, left front. LM, left middle. LH, left hind. RF, right front. RM, right middle. RH, right hind. **b,** Adaptation within the native fly embodiment. Terrain, speed and turning vary vertical, longitudinal and lateral control demands, respectively. **c,** Transfer of the fly-connectome prior to the Ant quadruped, followed by spatial navigation across mazes of increasing complexity. **d,** Transfer of the same recurrent prior across substantially different body plans, including a Cassie biped, a radial tripod, a Unitree Go1 quadruped and a hexapod, with body-specific sensory and motor interfaces learned for each embodiment.

recent efforts to implement connectomes as functional models of neural computation and behaviour [12, 13, 16].

The key comparison is between nominal control and adaptation. During straight, constant-speed locomotion on level ground, Adult-connectome and matched MLP controllers perform comparably (Fig. 2c, flat condition), providing a baseline against which adaptation can be assessed. We then vary the environment and movement demand within the same body, the embodiment itself, and combinations of both to test how the connectomic prior supports control as sensorimotor demands change (Fig. 1b–d).

2.2 Connectome structure supports adaptability within the fly embodiment

We first examine adaptability while keeping the neuromechanical fly embodiment fixed. Movement demands are varied along three complementary dimensions, i.e., terrain, direction and speed, which probe vertical, lateral and longitudinal adaptation, respectively.

Adaptation to vertical terrain variation.

We perturb the physical environment while preserving the fly body and its locomotor objective, using raised blocks of increasing height to vary terrain difficulty (Fig. 2a). On flat ground, Adult-connectome and matched MLP controllers show comparable

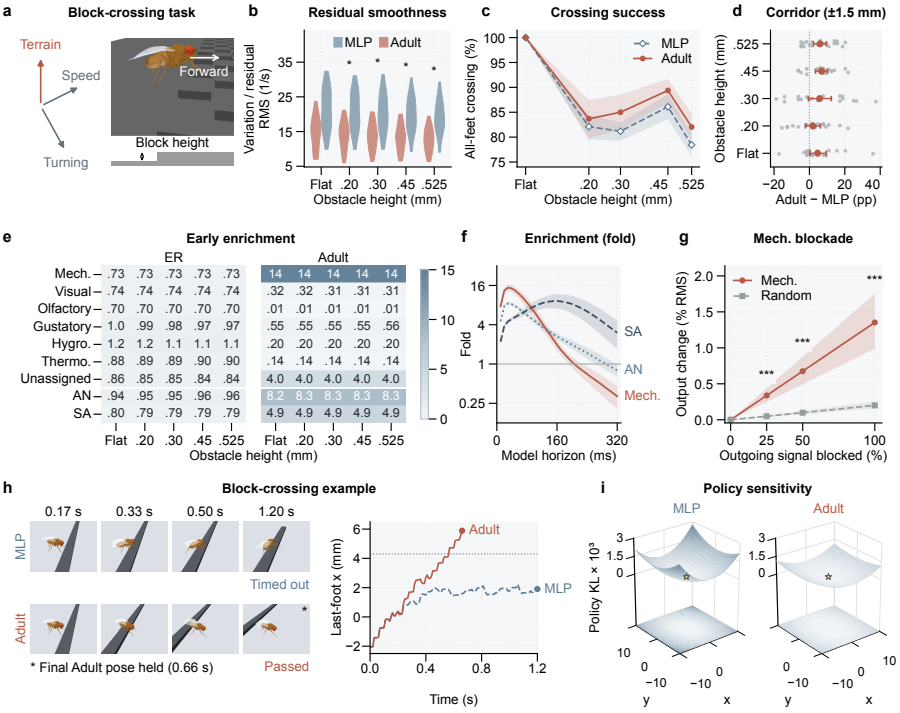


Fig. 2 Connectome control supports terrain adaptation and mechanosensory routing. a, Coordinate system and block-terrain task. **b,** Residual variation rate normalized by residual root-mean-square (RMS) amplitude. Violins show seed distributions. **c,** All-feet crossing success of Adult-connectome and matched multilayer perceptron (MLP) controllers across obstacle heights after 32,768 training decisions and a 1-s run-in (20 training seeds, 64 shared test phases per seed and height). Success requires all six feet to cross while maintaining an upright posture. **d,** Paired Adult-MLP differences in crossing success with lateral deviation additionally restricted to ± 1.5 mm. Grey points denote seed pairs. Flat denotes 0 mm. **e,** Cell-class output-sensitivity enrichment at 40 ms in Adult and randomized Erdős-Rényi (ER) controllers trained under matched conditions (20 paired seeds). Squared output derivatives with respect to initial neural states are normalized per cell to the whole-network mean. **f,** Adult sensitivity enrichment across model horizons sampled every 10 ms. Inputs beyond 80 ms repeat the final observed window. Mechanosensory neurons (Mech.), ascending neurons (ANs) and sensory-ascending neurons (SAs) exhibit distinct temporal profiles. **g,** RMS change in policy output after graded blockade of mechanosensory outgoing signals, compared with count-, partition- and degree-matched random controls on fixed input histories. **h,** Outcome-selected paired example at 0.525 mm obstacle height. Adult crosses the obstacle, whereas MLP times out. The asterisk marks the final Adult pose. **i,** Local policy-sensitivity landscapes quantified by policy Kullback-Leibler (KL) divergence under two-dimensional perturbations around the learned policy. The two axes vary body-velocity and joint-velocity input gains. Lines or points show 20-seed means, and shading or bars denote descriptive 95% seed-bootstrap intervals. Asterisks denote Holm-adjusted P values from two-sided exact paired sign-flip tests: **b**, $^{*}P < 0.05$ (130 comparisons); **g**, $^{***}P < 0.001$ (6 comparisons).

control and crossing performance. As obstacle height increases, however, the Adult controller produces smoother residual dynamics, with lower amplitude-normalized residual variation (Fig. 2b), while maintaining higher all-feet crossing success across increasingly challenging blocks (Fig. 2c). The same advantage persists under the stricter requirement that successful crossing remain within ± 1.5 mm laterally: paired Adult-MLP differences remain near zero on flat ground but become positive across

non-zero obstacle heights (Fig. 2d). A representative 0.525-mm trial illustrates this behavioural separation, with the Adult controller clearing the obstacle while the MLP advances only part-way before timing out (Fig. 2h, Supplementary Video 1).

We next examine how connectomic organization routes information during terrain adaptation. Compared with randomized Erdős-Rényi (ER) controllers trained under matched conditions, the Adult connectome shows substantially stronger early output-sensitivity enrichment in mechanosensory neurons (Fig. 2e). Within the Adult controller, this influence evolves through recurrent processing, with mechanosensory neurons, ascending neurons (ANs) and sensory-ascending neurons (SAs) exhibiting distinct temporal profiles across model horizons (Fig. 2f). Functional intervention further supports a privileged role for mechanosensory pathways. Graded blockade of mechanosensory outgoing signals produces markedly larger policy-output changes than count-, partition- and degree-matched random blockades, with the separation increasing as more signal is removed (Fig. 2g). Complementary analyses using unit-state and activity-scaled perturbations further resolve this temporal organization, revealing distinct enrichment and absolute-sensitivity profiles across mechanosensory neurons, ANs and SAs over recurrent propagation (Supplementary Fig. S1).

Adaptation to lateral turning variation.

We vary the desired movement direction while keeping the body and terrain fixed, with signed target headings of $\pm 30^\circ$, $\pm 60^\circ$ and $\pm 90^\circ$ (Fig. 3a). Terminal accuracy remains direction-dependent. Adult-connectome and matched MLP controllers show similar errors at positive targets, whereas the MLP exhibits larger errors and greater between-seed variability for negative targets (Fig. 3d). A clearer separation emerges in how the turn is executed. The Adult controller reduces amplitude-normalized command roughness by 29.9% (1.099 versus 1.568, Fig. 3b) and total command variation by 44.6% (11.264 versus 20.342, Fig. 3c), indicating smoother directional reconfiguration with substantially less command fluctuation. A matched $+60^\circ$ example illustrates this behaviour, with the Adult controller approaching and stabilizing around the target while the MLP overshoots and continues to rotate (Fig. 3i, Supplementary Videos 2,3).

We next examine how this directional reconfiguration is organized within the connectome. Future-output sensitivity follows a temporally structured progression across neuronal populations: mechanosensory influence is strongest at short propagation horizons, ANs remain enriched throughout, and SAs become progressively more prominent at longer horizons (Fig. 3e). Input learning further reorganizes this pathway. With the trained readout fixed, the learned input mapping increases relative AN enrichment by 25.1% while reducing absolute AN sensitivity by 93.7% at 100 ms (Fig. 3f), concentrating AN influence in relative terms while suppressing its absolute sensitivity. Whole-turn activity provides complementary evidence, with ANs reaching 1.46-fold enrichment and exceeding the input-population average in all seeds (Fig. 3g). Functionally, cutting the connections from ANs to descending neurons (DNs) increases mean heading error by 3.52°, compared with 1.16° after matched

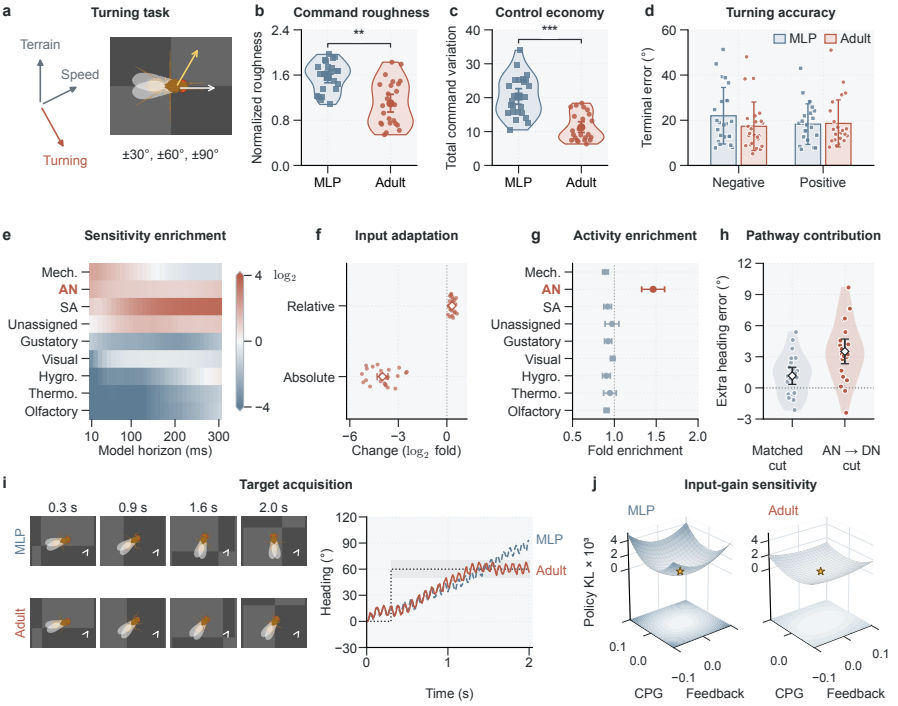


Fig. 3 Connectome control supports turning adaptation and ascending pathway recruitment. a, Turning task at target headings of $\pm 30°$, $\pm 60°$ and $\pm 90°$. $0°$ denotes straight walking. All controllers were trained with 24 paired seeds. **b,c,** Amplitude-normalized command roughness (**b**) and total command variation (**c**). Adult produces smoother commands (1.099 versus 1.568) with 44.6% less total variation (11.264 versus 20.342). Violins show seed distributions. Large symbols and bars denote means and 95% bootstrap confidence intervals. Asterisks denote Holm-adjusted P values from two-sided exact paired sign-flip tests: $^{**}P < 0.01$ and $^{***}P < 0.001$. **d,** Terminal heading error after averaging the three target magnitudes within each signed direction. Bars show mean $\pm$ s.d. and points denote seed-level averages. **e,** Future-output sensitivity across neuronal classes and propagation horizons. Mechanosensory influence dominates at short horizons, ascending neurons (ANs) remain enriched throughout and sensory-ascending neurons (SAs) become more prominent at longer horizons. **f,** Effect of input learning on AN sensitivity at 100 ms. With the trained readout fixed, learned input mappings increase relative AN enrichment by 25.1% while reducing absolute squared sensitivity by 93.7% (Holm-adjusted $P < 0.001$, 18 comparisons). Diamonds and bars show means and 95% t confidence intervals. **g,** Whole-turn activity enrichment relative to matched $0°$ trajectories. AN activity reaches 1.46-fold enrichment and exceeds the population average in all 24 seeds. **h,** Functional contribution of connections from AN to descending neurons (DNs). AN→DN cuts increase heading error more than matched control cuts ($3.52°$ versus $1.16°$, Holm-adjusted $P < 0.001$). **i,** Matched $+60°$ example showing target acquisition and stabilization by Adult and overshoot by MLP. Shading denotes the $\pm 10°$ target band. **j,** Local policy-sensitivity landscapes quantified by policy KL divergence under two-dimensional perturbations around the learned policy. The two axes vary turn-feedback and central pattern generator (CPG) state input gains. Stars mark the unperturbed policies.

control cuts (Fig. 3h). These results link smoother directional reconfiguration to a temporally organized ascending pathway.

Adaptation to longitudinal speed demands.

The third axis varies how fast the fly must move while preserving its body, movement direction and nominal terrain. We impose ramp commands that require continuous

speed adjustment and sinusoidal commands that require repeated acceleration and deceleration (Fig. 4a). Across 20 paired training seeds, the Adult-connectome controller tracks both demands more accurately than the matched MLP. Ramp-rate error and sinusoidal amplitude-ratio error are both significantly lower (Fig. 4b). Around ramp onset, the Adult controller also follows the commanded speed increments more closely (Fig. 4c). A matched ramp episode further illustrates this dynamic advantage, with the Adult controller repeatedly adjusting its speed in closer correspondence with the changing target (Fig. 4h, Supplementary Videos 4,5). Speed modulation therefore extends the connectome advantage to continuous reconfiguration of an internally specified movement demand.

We then examine the neural organization supporting this adaptation. Future-output sensitivity is enriched across mechanosensory neuron, AN and SA populations with distinct temporal profiles (Fig. 4d), a pattern that extends across both acceleration and deceleration (Supplementary Fig. S2). Holding target-related input fixed reveals class-specific activity responses during acceleration (Fig. 4e), while withholding the changing target signal from individual populations shows the largest loss of drive response for ANs, with substantial contributions from mechanosensory neurons (Fig. 4f). Across the 20 paired models, stronger Adult advantages in target-driven neural response are associated with larger reductions in ramp-rate error (Fig. 4g). These results suggest that mechanosensory and ascending circuitry may serve as a key neural substrate for continuous speed adaptation.

Convergent computational organization across tasks and connectomes.

Across all three adaptation settings, perturbation analyses reveal a convergent pattern of reduced local policy sensitivity under connectome-constrained control. During terrain control, joint perturbations to body- and joint-velocity input gains produce smaller policy shifts in the Adult controller than in the matched MLP (Fig. 2i). For speed control, matched parameter perturbations reveal a broader low-sensitivity neighbourhood around the Adult solution (Fig. 4i), and the same tendency emerges in turning control with the clearest behavioural improvements (Fig. 3j). Reduced local sensitivity therefore recurs across distinct changing demands, providing a common computational signature of connectome-constrained adaptation. This pattern conceptually echoes the flat-minima perspective in machine learning, where broader solution neighbourhoods are associated with improved generalization and robustness [17, 18].

We further examine the larval brain and integrated adult brain-VNC connectomes. The larval connectome supports directional turning and dynamic speed control in the adult fly body, while the Adult connectome shows greater mean command-aligned rotation with lower between-model variability during turning, together with faster acceleration and deceleration responses and better preservation of response amplitude at the higher command frequency during speed control (Supplementary Figs. S3 and S4). Extending the framework to the brain-VNC connectome reveals a complementary role for readout location. DN readouts develop a more structured concentration of sensitivity across input populations and depend more strongly on brain-derived signals, whereas motor-neuron readouts depend more

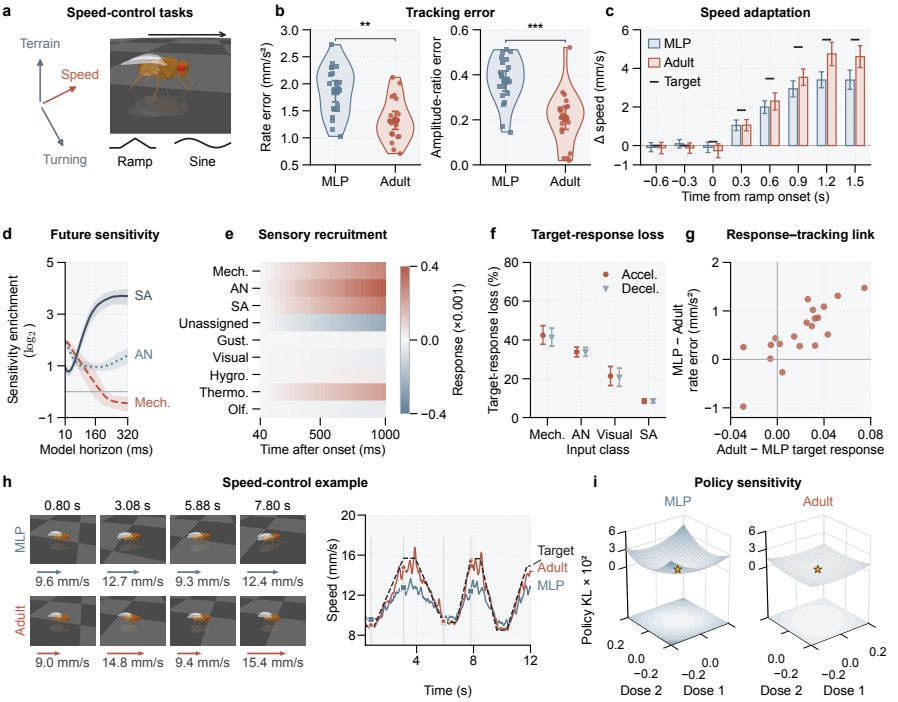


Fig. 4 Connectome control supports speed adaptation and sensorimotor pathway recruitment. a, Ramp and sinusoidal speed-control tasks. **b,** Ramp-rate error and absolute sine amplitude-ratio error for Adult-connectome and matched MLP controllers across 20 paired training seeds after 15,872 decisions. Both errors are lower for Adult. Violins show seed distributions. Symbols and bars denote means and 95% bootstrap confidence intervals. Asterisks denote Holm-adjusted P values from two-sided exact paired sign-flip tests: $^{**}P < 0.01$ and $^{***}P < 0.001$. **c,** Speed increments around ramp onset for the first predefined case, measured in 100-ms windows and referenced to each model's preceding 100-ms baseline. Black markers denote target increments. **d,** Cell-count-normalized future-output sensitivity enrichment across mechanosensory neurons (Mech.), ascending neurons (ANs) and sensory-ascending neurons (SAs). **e,** Class-specific activity responses during acceleration relative to interventions holding target-related signals fixed. **f,** Loss of drive response when the changing target signal is withheld from individual neuronal classes. **g,** Relationship between paired target-response and ramp-rate-error advantages across 20 model pairs (Spearman $\rho = 0.792$, Holm-adjusted $P < 0.01$, 45 tested associations). **h,** Matched frames and speed traces from a completed ramp episode. **i,** Local policy-sensitivity landscapes quantified by policy KL divergence under two-dimensional perturbations around the learned policy. The two axes follow two filter-normalized parameter directions. In **c–f**, curves, bars or points show means with 95% seed-bootstrap intervals, and holds denote policy-internal interventions.

strongly on VNC-derived signals, with corresponding losses in closed-loop recovery under regional message substitutions (Supplementary Figs. S5 and S6). These comparisons show that developmental wiring and readout location systematically shape connectome-constrained control, while reinforcing adaptability as a functional advantage of connectomic organization.

2.3 Cross-embodiment transfer of the fly-connectome prior

We next ask whether the adaptability conferred by connectome structure extends across bodies. The adult fly connectome is retained as a fixed recurrent prior, while embodiment-specific sensory and motor interfaces are learned anew for each morphology, isolating the transfer of connectomic organization across changing sensorimotor systems.

Quadruped locomotion and spatial navigation.

We first transfer the fixed fly-connectome prior to the simulated Ant quadruped, whose four-legged morphology, actuation and body dynamics differ substantially from those of the six-legged digital fly (Fig. 5a,b). We begin with flat-ground locomotion to isolate embodiment change from additional environmental or spatial demands. Even under this nominal setting, the connectome controller moves faster and more directly than the matched MLP (Fig. 5b–d). Across 20 paired training seeds, mean forward velocity reaches 1.94 m/s versus 0.68 m/s, while signed path efficiency increases from 0.701 to 0.885 (Fig. 5c,d). This contrasts with the native fly, where the two controllers perform comparably during nominal locomotion, indicating that embodiment change alone is sufficient for a connectomic advantage to emerge. This stronger forward drive is accompanied by a higher failure fraction (38.6% versus 7.0%), revealing a trade-off with episode stability. Across altered floor friction, the connectome retains its velocity advantage while also showing higher failure fractions, reinforcing this trade-off (Supplementary Fig. S7, Supplementary Video 6).

The advantage broadens when embodiment change is combined with spatial navigation. Across seven representative controllers, including the Adult connectome, Transformer, MLP, gated recurrent unit (GRU), long short-term memory (LSTM), Reservoir and Reservoir with readout-only adaptation (Reservoir+RO), performance separates as maze complexity increases, with the Adult-connectome controller achieving the highest success in both Medium and Large mazes (Fig. 5e, Supplementary Videos 7-9). Large-maze analyses further resolve these differences across trajectories, speed profiles and route-complexity-stratified navigation outcomes (Supplementary Fig. S8).

Focusing on the Adult-MLP comparison, success remains similar in the simpler U-Maze (84.8% versus 81.5%) but increases from 44.1% to 64.2% in the more demanding Medium maze (Supplementary Videos 10-12). The Adult controller also maintains higher success across evaluation time budgets (Fig. 5g) and increasingly stringent geodesic-efficiency thresholds (Fig. 5h). The separation is largest for the most complex routes: the success-rate advantage is 14.2 percentage points for routes requiring 3–4 turns, 12.9 points for 5 turns and 30.3 points for 6–7 turns, reaching 50.6% versus 20.4% in the most complex group (Fig. 5i). Under the primary Medium-maze evaluation, physical failures are also reduced from 15.1% to 9.2% (Fig. 5j), while matched successful trajectories illustrate shorter and faster routes under connectome control (Fig. 5f). The strongest gains thus emerge as navigation requires greater directional reconfiguration.

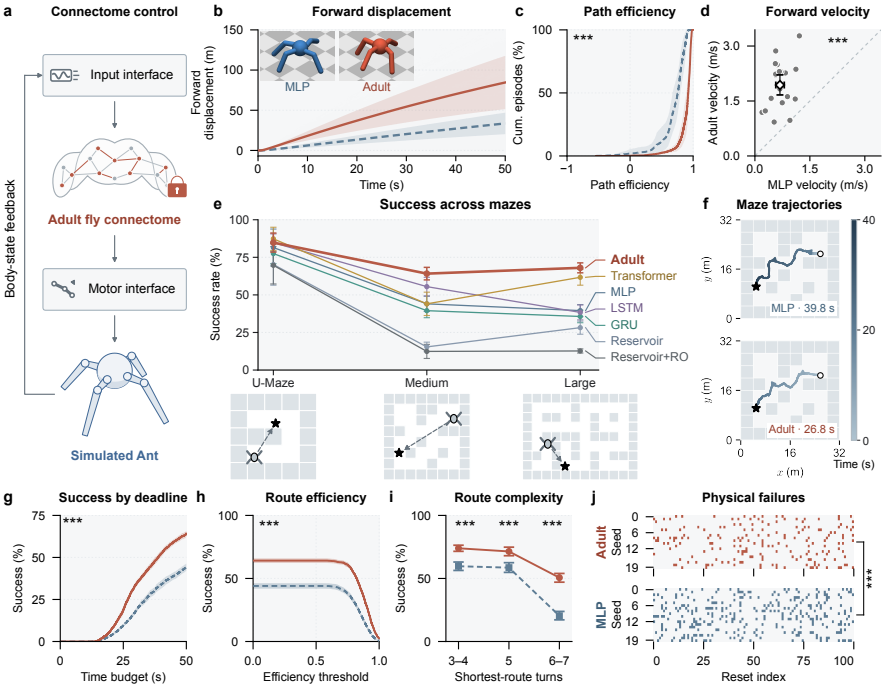


Fig. 5 Fly-connectome control transfers to quadrupedal locomotion and navigation. a, Fixed adult *Drosophila* brain connectome coupled to learned sensory and motor interfaces for the Ant quadruped. Additional maze inputs are omitted for clarity. **b–d,** Flat-ground locomotion: forward displacement (**b**), cumulative signed path-efficiency distributions (**c**) and paired seed-mean forward velocity (**d**). **e,** Success rates across U-Maze, Medium and Large mazes for seven controller architectures. Schematics below show the corresponding maze layouts. **f,** Matched successful trajectories illustrating a shorter and faster route under connectome control. Circles mark starts, stars mark goals and colour denotes elapsed time. **g,** Medium-maze success as a function of time budget, with all evaluation resets retained in the denominator. **h,i,** Medium-maze success above increasing geodesic-efficiency thresholds (**h**) and stratified by shortest-route turn count (**i**). **j,** Physical failures across Medium-maze seed-reset combinations. Comparisons use 20 paired training seeds in **b–d** and **g–j**, and five paired training seeds in **e**, with 100 matched evaluation resets per seed. Curves in **b** average 2,000 episodes per method. Shading denotes ± 1 s.d. across seed means. Error bars and bands in **c,d,g–i** denote 95% paired-bootstrap intervals. Asterisks denote Holm-adjusted P values from two-sided exact paired tests: $^{***}P < 0.001$ (26 comparisons).

Control across diverse embodiments.

We next ask whether cross-embodiment transfer extends beyond Ant to substantially different body plans. The same fixed fly-connectome core is coupled to learned body-specific sensory and motor interfaces for four simulated embodiments spanning two to six legs: a Cassie biped, a radial tripod, a Unitree Go1 quadruped and a hexapod (Fig. 6a,b, Supplementary Videos 13-20). Within each body, connectome and MLP controllers receive the same interaction budget and have matched trainable parameter counts. The tasks span bipedal postural balance and multi-legged forward locomotion under nominal and perturbed conditions, testing whether the structural prior remains useful across distinct morphologies and control regimes (Fig. 6c,d).

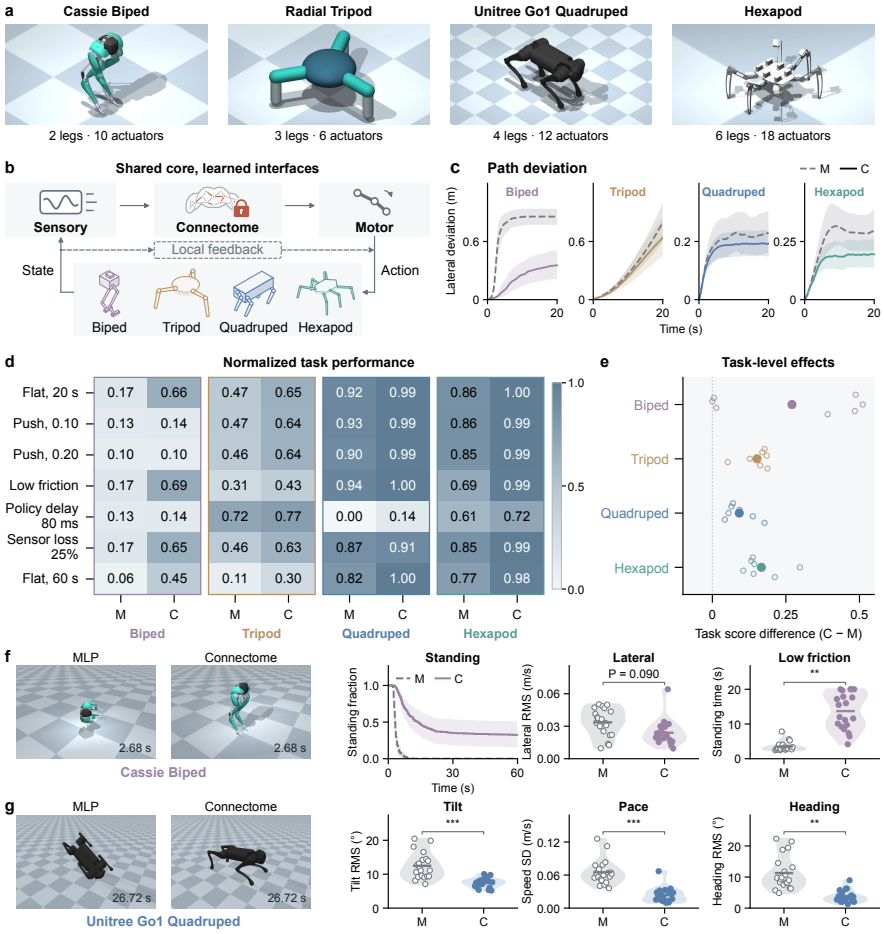


Fig. 6 A fly-connectome prior supports robust control across diverse embodiments. a, Simulated embodiments spanning two to six legs. **b,** Shared frozen fly-connectome core with learned body-specific sensory and motor interfaces. **c,** Mean absolute lateral displacement during nominal trials for Adult-connectome (C, solid) and MLP (M, dashed) controllers. **d,** Mean normalized performance across nominal control, lateral pushes, reduced friction, policy delay, partial sensor loss and extended-duration evaluation. **e,** Condition-level C−M effects corresponding to **d**. Open symbols denote the seven condition effects, and filled symbols denote their equal-weight mean. **f,** Matched reduced-friction Cassie snapshots with cohort-level flat-ground standing fraction over 60 s, early lateral-velocity root-mean-square (RMS) over 0–1 s and reduced-friction standing duration over 20 s. **g,** Matched Go1 snapshots from a 60-s walking trial with flat-ground torso-tilt RMS, variability of successive 1-s forward speeds and heading-error RMS over 2–8 s. All analyses use 20 paired training seeds and 20 matched evaluation states per seed and condition. Violins and points show seed means, short bars show group means and shading denotes 95% seed-bootstrap confidence intervals. Asterisks denote Holm-adjusted P values from two-sided exact paired-seed sign-flip tests: $^{**}P < 0.01$ and $^{***}P < 0.001$. Corrections are applied over the prespecified 10-endpoint family for lateral motion, tilt and pace, and over 3,150 exploratory comparisons for duration and heading. Controllers are matched within each embodiment for interaction budget and trainable parameter count.

Performance benefits are evident across all four embodiments, although their magnitude varies with morphology and control demand (Fig. 6d,e). Averaged across seven evaluation conditions spanning nominal control, physical perturbations, policy delay, partial sensor loss and extended-duration evaluation, connectome control improves normalized performance over the matched MLP by 0.271 in Cassie, 0.151 in the radial tripod, 0.091 in Go1 and 0.167 in the hexapod. The six-legged hexapod achieves the highest absolute performance under connectome control, averaging approximately 0.95 across conditions, consistent with its closer morphological correspondence to the native fly facilitating transfer. In comparison, the Cassie biped shows the largest mean improvement over its matched MLP despite representing the greatest departure from the native six-legged body plan. The radial tripod and Go1 quadruped also show consistent gains across conditions, with the smaller margin in Go1 reflecting an already strong MLP baseline. These complementary patterns indicate that connectome structure supports adaptability across embodiment change, enabling both high absolute performance in morphologically aligned bodies and substantial relative gains in markedly different ones.

Beyond aggregate task performance, connectome control also improves stability across embodiments. In Cassie, it increases flat-ground standing fraction, reduces early lateral motion and prolongs standing under reduced friction. In a matched reduced-friction trial, the connectome-controlled robot remains upright for 14.18 s, compared with 2.68 s for the MLP (Fig. 6f). In the radial tripod, connectome control reduces accumulated lateral displacement, heading error and late-stage lateral drift during extended walking, with corresponding improvements in outcome composition and task success across evaluation conditions (Supplementary Fig. S9). In Go1, it reduces torso tilt, forward-speed variability and heading error, and completes a matched 60-s trial in which the MLP fails at 26.72 s (Fig. 6g). The hexapod likewise shows reduced lateral displacement and late drift and retains the prescribed lateral corridor for longer across multiple tolerance thresholds, despite both controllers sharing the same analytic CPG (Supplementary Fig. S10). These effects extend cross-embodiment transfer from task-level performance to sustained postural and trajectory control.

3 Discussion

Our study identifies *adaptability* as a functional advantage of biological connectome structure for embodied control. In the native fly body, connectome and MLP controllers perform similarly during nominal locomotion but diverge when control must be reconfigured across terrain, direction and speed. These behavioural advantages are accompanied by spontaneous recruitment of task-relevant sensorimotor pathways and a convergent reduction in local policy sensitivity across adaptation settings. The same connectome prior remains effective after substantial changes in embodiment, supporting locomotion and postural control across two-, three-, four- and six-legged bodies. In spatial navigation, its advantage over representative artificial controllers becomes more pronounced as maze and route complexity increase. Comparisons using the larval brain and integrated adult brain-VNC connectomes further show that

developmental wiring and circuit readout location systematically shape the resulting control dynamics. Across these settings, connectomic organization confers its clearest advantage when embodied control must adapt to change.

The context-dependent advantage of connectome control helps clarify the computational benefit of evolved neural wiring. Previous work has shown that measured connectivity can constrain visual computation, reproduce whole-brain sensorimotor transformations and support embodied behaviour [11–13]. Our results extend this view by showing that biological wiring can bias learning towards adaptable control, beyond simply providing a functional network architecture. During locomotion, optimization spontaneously recruits behaviourally relevant mechanosensory and ascending pathways without prescribing which neuronal populations should dominate the solution. The stronger dynamic responses of the adult connectome relative to its larval counterpart are consistent with the extensive circuit remodelling accompanying *Drosophila* metamorphosis and the emergence of adult-specific behaviours [19]. The recurring reduction in local policy sensitivity across terrain, turning and speed further suggests that the recurrent prior shapes the geometry of learned control as well as task-specific performance.

The transfer results extend this perspective to embodied intelligence. Generalist robot policies increasingly aggregate experience across tasks and platforms [20, 21], yet changes in morphology, sensing and dynamics can substantially alter the observation-action relationship on which control depends. Our results suggest an architectural complement to scaling data and adaptation procedures in which a reusable recurrent prior is preserved while body-specific sensory and motor interfaces are relearned. In spatial navigation, the connectome outperforms diverse artificial controllers, with its advantage widening as maze complexity increases, showing that the benefit extends beyond a single architectural baseline and becomes more pronounced as control demands grow. The consistently high absolute performance of the six-legged hexapod suggests that morphological compatibility can facilitate transfer, while benefits remain evident in substantially different bodies, including the Cassie biped, radial tripod and Go1 quadruped. Connectome structure may therefore offer a biologically grounded route towards embodied controllers that remain effective as bodies and control demands change.

The framework also provides a way to study biological intelligence through behaviour. Neuromechanical simulation places measured neural structure back into a closed sensorimotor loop, allowing anatomical organization to be perturbed and evaluated through whole-body consequences [3, 14]. The fly connectome is particularly suited to this approach because local sensorimotor circuits are embedded within recurrent ascending and descending pathways and broader interconnected populations that support integration and distributed control [11, 22–24]. In our models, spontaneous recruitment of locomotor-related populations, pathway-specific interventions and readout-dependent brain-VNC dependencies illustrate how connectome-constrained embodiment can generate circuit-level hypotheses for subsequent biological testing.

Several limitations define the scope of these findings. Our principal connectome is a static anatomical prior, whereas biological nervous systems adapt through synaptic plasticity, neuromodulation, intrinsic dynamics and structural remodelling [25–27]. Preserving recurrent wiring while learning sensory and motor interfaces therefore captures only part of biological adaptability and abstracts away the developmental processes that couple nervous systems to bodies. Although our analyses implicate sensorimotor routing, recurrent dynamics and pathway organization, the specific structural properties responsible for the observed advantage remain unresolved. Recurrence, sparsity, modularity and ascending-descending organization warrant systematic dissection. Finally, all embodied experiments are conducted in simulation. Physical systems introduce sensing noise, latency, actuator uncertainty and unmodelled contact dynamics, providing a more stringent test of connectome-derived control.

The broader opportunity extends beyond *Drosophila*. The fly offers a rare combination of synapse-resolution connectivity at large scale and sufficiently rich behaviour for embodied computational studies. Connectomic resources are now emerging at complementary scales beyond the fly, from brain-wide mesoscale connectivity in mouse [28] to synapse-resolution reconstructions of mouse visual cortex linked to functional measurements [29] and millimetre-scale human cortical tissue [30]. Comparative analyses across species, developmental stages and functional systems may help distinguish organism-specific solutions from more general principles of adaptable computation. Translating such principles into scalable architectures could turn connectomics from a description of biological circuitry into a source of design principles for artificial systems that remain effective as environments, tasks and bodies change.

4 Methods

Experimental framework. We use the *Drosophila* connectome as a fixed recurrent prior and learn interfaces that map observations to neural activity and recurrent activity to motor commands. Simulated controllers receive body-state feedback after each action, forming a closed loop. We vary terrain, turning and speed within the fly embodiment, then train independent body-specific interfaces for five additional embodiments. Gait generators, navigation modules and local feedback branches form part of the deployed controller where specified.

Connectome construction and recurrent dynamics. We construct the principal recurrent substrate from the adult brain connectome, FlyWire materialization 783 [9, 31], which yields 138,584 neurons and 3,732,460 directed connections after preprocessing. We retain connections supported by at least one synapse and aggregate records by ordered neuron pair. We assign connection signs from predicted neurotransmitter identity, with γ-aminobutyric acid (GABA) assigned negative weights and acetylcholine, glutamate, dopamine, serotonin and octopamine assigned positive weights. Connections without a supported neurotransmitter assignment receive zero weight. For signed synapse count A_{ij} from neuron i to neuron j, we normalize outgoing weights as $W_{ij} = A_{ij} / \sum_{\ell=1}^{N} |A_{i\ell}|$, where N is the number of neurons. Rows

corresponding to neurons without outgoing connections remain zero, and the resulting sparse weight matrix W remains fixed throughout training. Recurrent dynamics follow

$$\begin{aligned} \mathbf{r}_k &= \tanh(\mathbf{h}_k), \\ \mathbf{h}_{k+1} &= (1-\alpha)\mathbf{h}_k + \alpha \tanh\big(W^{\mathsf{T}}\mathbf{r}_k + P_{\mathrm{in}}(E_\theta \widetilde{\mathbf{o}}_k + \mathbf{b}_\theta)\big), \end{aligned} \tag{1}$$

where k indexes recurrent updates, $\mathbf{h}_k$ is the recurrent state, $\mathbf{r}_k$ is its bounded activation and $\widetilde{\mathbf{o}}_k$ is the observation supplied at update k. The learned sensory projection matrix E_θ and bias $\mathbf{b}_\theta$ map the observation to the neuronal population selected by the fixed input operator P_{in}. The recurrent update coefficient α is set to 0.5, corresponding to a nominal 10-ms model step and a 20-ms time constant. Turning and speed policies repeat the current observation over four recurrent updates per 40-ms control decision. Terrain policies instead process four successive 10-ms sensory frames, while controllers for the additional embodiments use one recurrent update per decision. Recurrent states are reset between episodes and initialized to zero for deterministic evaluation.

The fixed output selector P_{out} supplies recurrent activations to a learned motor readout. Fly policies use 1,290 descending neurons as outputs, with 2,279 ascending and sensory-ascending neurons as inputs for terrain and 18,756 afferent neurons as inputs for turning and speed. The 18,756 afferents comprise 16,477 sensory neurons, 1,716 ascending neurons and 563 sensory-ascending neurons. Controllers for the additional embodiments use the same 18,756 afferents together with 109 neurons annotated as motor outputs [31].

Policy learning. Policies use diagonal Gaussian distributions with state-dependent means and learned state-independent standard deviations. For the main fly and additional-embodiment policies, Gaussian means are bounded as $3\tanh(\mathbf{m}/3)$, where $\mathbf{m}$ denotes the actor output. Ant policies use unsaturated Gaussian means. Graph readouts apply layer normalization followed by a linear output layer. Sampled actions are transformed by $\tanh$ before entering the actuator interface, and deterministic evaluation uses the transformed policy mean.

All tasks except maze adaptation use proximal policy optimization (PPO) [32], generalized advantage estimation [33] and a separate two-layer critic with 256 units per layer. For the main fly and additional embodiment comparisons, multilayer perceptron (MLP) actors match the trainable capacity of the corresponding connectome controllers within 0.5%. Paired architectures share critic initialization, task definitions, actuator mappings, rewards and interaction budgets. Fly rewards include potential-based shaping [34]. Maze students instead minimize squared error between policy means and the inverse-$\tanh$ transform of clipped teacher actions. The teacher is absent during evaluation. Optimization and adaptation settings are provided in Supplementary Tables S1–S3.

Neuromechanical fly tasks. NeuroMechFly v2, simulated through FlyGym and MuJoCo [14, 35], has 42 actuated leg joints, adhesion and a six-oscillator central pattern generator (CPG). Adult-connectome and MLP policies modulate a reference gait using proprioceptive, contact, body-state and CPG feedback. All fly comparisons use

final-budget checkpoints and matched evaluation initializations. Task interfaces and secondary metrics are specified in Supplementary Information.

Terrain. The policy outputs six leg-specific residual corrections for block crossing. 20 model pairs receive 32,768 training decisions per model, with block heights of 0.20, 0.30, 0.45 and 0.525 mm sampled equally with ± 0.025-mm jitter. Evaluation uses the same four heights and flat ground, with 64 shared phases per condition. Each 1.2-s evaluation trial follows 1 s of zero-residual walking while maintaining continuous physical state, whereas training omits this run-in. Success requires all six feet to pass the far edge while the body-up component remains at least 0.8 for 0.1 s.

Turning. Two policy outputs modulate forward drive and differential left-right CPG amplitude. Each trial begins with a 0° target, which switches at 0.3 s to $\pm 30^\circ$, $\pm 60^\circ$ or $\pm 90^\circ$. 24 model pairs receive 8,192 training decisions per model and are evaluated from 20 matched initializations per target over 2 s, together with a 0° reference condition. Terminal accuracy is measured as absolute wrapped heading error. Stable holding requires heading error within 10°, a body-up component of at least 0.8 and mean forward speed of at least 3 mm/s for 0.2 s.

Speed. Two policy outputs modulate locomotor drive and turning. 12-s training episodes sample ramp and sinusoidal target-speed profiles with equal probability over 8–16 mm/s. 20 model pairs receive 15,872 training decisions per model and are evaluated on 4 fixed ramps and 4 fixed sinusoids. Outcomes include full-horizon speed error, ramp-rate error and sinusoidal amplitude-ratio error. Missing samples after early termination are assigned an error of 20 mm/s. Ramp-rate error is the absolute difference between the commanded slope and a linear fit to the measured speed, excluding 0.21 s at each ramp boundary. Sinusoidal fits begin at 4 s and include an intercept, a linear trend, and sine and cosine terms at the commanded frequency. Amplitude-ratio error is the absolute deviation of the fitted-to-commanded amplitude ratio from one.

Ant locomotion and maze navigation. We use Gymnasium Ant-v5 and the Gymnasium-Robotics AntMaze U-Maze, Medium Diverse and Large Diverse environments [36, 37], all based on the same eight-actuator Ant embodiment. Locomotor pretraining uses approximately 3.0 million PPO interactions followed by 0.2 million stability-refinement interactions. Checkpoints are selected according to training return, stability and velocity. Flat-ground evaluation uses 20 paired Adult and MLP models, with 100 matched resets per model over 50 s.

Maze comparisons include Adult, MLP, gated recurrent unit (GRU), long short-term memory (LSTM), Transformer, Reservoir and Reservoir with readout-only adaptation (Reservoir+RO). The additional artificial architectures approximately match the trainable capacity of the MLP. Each controller family has 20 optimization replicates that share the specified teacher data, routing modules and source policies. Final adaptation checkpoints are evaluated on 100 matched resets per replicate. Evaluation lasts 35 s in U-Maze and 50 s in Medium and Large, with success defined as reaching within 0.45 m of the goal. Source-policy progression and routing procedures are described in Supplementary Information.

Signed forward efficiency is defined as net forward displacement divided by planar path length. Success-weighted geodesic efficiency is $L_\star / \max(L_\star, L)$ for successful episodes and zero for failures, where $L_\star$ denotes the shortest grid-route length and L denotes the travelled path length to first goal attainment.

Additional embodiments. Cassie and Go1 use public MuJoCo models, whereas the radial tripod and hexapod use custom models [35, 38]. Cassie is trained for postural balance, while the other embodiments track forward-speed commands through the body-specific interfaces specified in Supplementary Table S4. Rewards combine forward progress or postural support with stability, path-following and control-effort terms. Each architecture has 20 training replicates per body. Cassie and Go1 receive 524,288 decisions per model, whereas the tripod and hexapod receive 131,072.

Final checkpoints are evaluated from 20 matched initial states under seven conditions comprising nominal control, two lateral pushes, reduced friction, action delay, observation loss and endurance. Standard trials last 20 s, while endurance trials last 60 s. For the walking embodiments, success requires survival over the full evaluation horizon, mean net forward velocity of at least 70% of the commanded speed and maximum absolute lateral displacement no greater than 0.75 m. Cassie's primary score is the fraction of the evaluation horizon for which balance is maintained. Full-interval Cassie success additionally requires absolute mean forward velocity no greater than 0.05 m/s together with the same lateral bound. Body-level paired scores assign equal weight to the seven evaluation conditions.

Each lateral push lasts 0.4 s and delivers impulse $\mathcal{I} = \eta m \sqrt{g h_0}$, where $\eta \in \{0.10, 0.20\}$, m denotes body mass, g gravitational acceleration and h_0 the initial reference height. Friction is reduced to 60% of its nominal value, and action delay is fixed at 80 ms. Observation loss removes 25% of physical-input packets over a 4-s interval. Missing physical inputs retain their most recent values, while command and controller-state inputs remain available.

Cell-class sensitivity and interventions. Neuronal groups are defined using FlyWire identifiers and annotations [31]. We differentiate future bounded policy outputs with respect to the current recurrent state while holding subsequent observations fixed. For neuronal class G, reference population $\mathcal{R}$ and propagation horizon ℓ, sensitivity enrichment is

$$\mathcal{E}_G(\ell) = \frac{|G|^{-1} \sum_{i \in G} \|\partial \mathbf{y}_{k+\ell} / \partial h_{k,i}\|_2^2}{|\mathcal{R}|^{-1} \sum_{i \in \mathcal{R}} \|\partial \mathbf{y}_{k+\ell} / \partial h_{k,i}\|_2^2}, \tag{2}$$

where $h_{k,i}$ is the ith component of the recurrent state at update k and $\mathbf{y}$ denotes the turn command, drive command or six-component terrain residual. Ratios are computed within individual windows before averaging at the model level. The reference population comprises afferent neurons for turning and speed and all neurons for the main terrain analyses.

Turning analyses compare the initial and trained sensory projections while keeping the trained readout fixed, with each projection replayed under its corresponding operating states. Sensitivity is evaluated over 36 shared windows per model and propagation horizons of 10–300 ms, with input-learning contrasts assessed at 100 ms. Pathway interventions remove 14,840 connections from ascending to descending

neurons beginning at 0.32 s. Matched control cuts preserve edge count and connection sign while approximately matching postsynaptic targets and weight magnitude. Two states per direction are used to quantify the increase in mean absolute heading error over 0.32–2 s.

Speed replays hold target-related inputs either globally or within individual neuronal classes while updating tracking-error inputs consistently. Terrain analyses compare 20 Adult-connectome controllers with 20 matched Erdős-Rényi (ER) controllers that preserve node and edge counts, interface partitions and the signed-weight multiset. Graded blockade scales the outgoing activation of selected neurons by $1-d$, where $d \in \{0, 0.25, 0.5, 1\}$, and compares the resulting effects with matched random masks. Replay windows, control construction and connectome comparisons are specified in Supplementary Information.

Local policy sensitivity. Policy Kullback-Leibler (KL) divergence quantifies local sensitivity on fixed observation histories. Reference and perturbed policies share the original action covariance Σ and are initialized from the corresponding recurrent state of the reference policy. Let $\boldsymbol{\mu}_\theta$ and $\boldsymbol{\mu}_{\theta'}$ denote the Gaussian action means under the reference and perturbed policy parameters θ and θ', respectively. We compute

$$\overline{D}_{\mathrm{KL}} = \frac{1}{2} \left\langle (\boldsymbol{\mu}_{\theta'} - \boldsymbol{\mu}_\theta)^{\mathsf{T}} \Sigma^{-1} (\boldsymbol{\mu}_{\theta'} - \boldsymbol{\mu}_\theta) \right\rangle , \tag{3}$$

where Σ is the shared action covariance and the average is taken over evaluation windows and control decisions. Terrain surfaces vary body-velocity and joint-velocity input gains, while turning surfaces vary turn-feedback and encoded CPG-state gains. Gains span $[0.9, 1.1]$ on a 21×21 grid, with all policy parameters held fixed. Terrain axes are selected using 10 model pairs and assessed on the remaining 10, while the displayed surfaces average all 20 pairs. Turning feedback comprises averaged yaw rate and lateral displacement, and CPG inputs are selected by screening for comparable response amplitudes. Speed surfaces instead perturb trainable actor parameters along two random filter-normalized directions [39], with coefficients spanning $[-0.2, 0.2]$ on the same grid. Axis selection, replay sampling and directional checks are specified in Supplementary Information.

Statistical analysis. Training or optimization replicates are the units of inference. Episodes, conditions, observation windows, neurons and perturbation directions are aggregated within each model before statistical testing, while paired analyses preserve matched model identities and evaluation conditions. Unless otherwise specified, comparisons use two-sided exact paired sign-flip tests with the absolute mean paired difference as the test statistic. Turning pathway-specificity analyses use paired t tests, developmental comparisons use exact unpaired permutation tests, and speed associations use permutation-tested Spearman correlations. Holm correction [40] is applied within the analysis families defined in Supplementary Information.

Intervals are pointwise 95% unless otherwise stated and are generally obtained by resampling models while preserving paired identities and associated conditions. Resampling counts and interval-specific exceptions are provided in Supplementary

Information. Behavioural failures remain in the relevant denominators, and incomplete kinematic windows are excluded according to matched-pair coverage without imputation. All prespecified models are retained in the complete analysis cohorts.

Data Availability

The public data and benchmark resources used in this study are available from their original repositories. The *Drosophila* connectome (FlyWire FAFB v783) and associated neuronal annotations are available through Codex (https://codex.flywire.ai/?dataset=fafb) [9, 31]. NeuroMechFly v2 resources are available from the Harvard Dataverse (https://doi.org/10.7910/DVN/3MCEYR) [14]. The Ant-v5 environment is available through Gymnasium (https://gymnasium.farama.org/environments/mujoco/ant/) [36], and the AntMaze environments are available through Gymnasium-Robotics (https://robotics.farama.org/main/envs/maze/ant_maze/). The Medium Diverse GR configuration is additionally archived in Minari (https://minari.farama.org/datasets/D4RL/antmaze/medium-diverse-v2/) [37].

Code Availability

The implementation code, configuration files, and evaluation protocols are available at https://github.com/THU-NeuroML/FlyControl. Supplementary Videos can be downloaded at https://raw.githubusercontent.com/THU-NeuroML/FlyControl/v1.0.0/Supplementary-Videos.zip.

Acknowledgments

This work was supported by the NSFC Project (No. T2622023, No. 62406160 and No. 62595773), the Beijing Nova Program (No. 202604841279), and the Beijing Major Science and Technology Project (No. Z251100008425003).

Author Contributions Statement

S.Y. and L.W. conceived the project. S.Y. designed the computational framework and performed main experiments. S.Y. and L.W. analyzed the results. S.Y. and L.W. wrote the manuscript. All authors discussed the results and revised the manuscript. L.W. supervised the project.

Competing Interests Statement

The authors declare no competing interests.

References

[1] Merel, J., Botvinick, M., Wayne, G.: Hierarchical motor control in mammals and machines. Nature Communications **10**(1), 5489 (2019)

[2] Gupta, A., Savarese, S., Ganguli, S., Fei-Fei, L.: Embodied intelligence via learning and evolution. Nature Communications **12**, 5721 (2021)

[3] Zador, A., Escola, S., Richards, B., *et al.*: Catalyzing next-generation artificial intelligence through neuroai. Nature Communications **14**, 1597 (2023)

[4] Hwangbo, J., Lee, J., Dosovitskiy, A., Bellicoso, D., Tsounis, V., Koltun, V., Hutter, M.: Learning agile and dynamic motor skills for legged robots. Science Robotics **4**(26), 5872 (2019)

[5] Lee, J., Hwangbo, J., Wellhausen, L., Koltun, V., Hutter, M.: Learning quadrupedal locomotion over challenging terrain. Science Robotics **5**(47), 5986 (2020)

[6] Miki, T., Lee, J., Hwangbo, J., Wellhausen, L., Koltun, V., Hutter, M.: Learning robust perceptive locomotion for quadrupedal robots in the wild. Science Robotics **7**(62), 2822 (2022)

[7] Kumar, A., Fu, Z., Pathak, D., Malik, J.: RMA: Rapid motor adaptation for legged robots. In: Proceedings of Robotics: Science and Systems (2021)

[8] Winding, M., Pedigo, B.D., Barnes, C.L., Patsolic, H.G., Park, Y., Kazimiers, T., Fushiki, A., *et al.*: The connectome of an insect brain. Science **379**(6636), 9330 (2023)

[9] Dorkenwald, S., Matsliah, A., Sterling, A.R., Schlegel, P., Yu, S.-c., McKellar, C.E., *et al.*: Neuronal wiring diagram of an adult brain. Nature **634**, 124–138 (2024)

[10] Azevedo, A., Lesser, E., Phelps, J.S., Mark, B., Elabbady, L., Kuroda, S., Sustar, A., Moussa, A., Khandelwal, A., Dallmann, C.J., Agrawal, S., Lee, S.-Y.J., Pratt, B., Cook, A., Skutt-Kakaria, K., Gerhard, S., Lu, R., Kemnitz, N., Lee, K., Halageri, A., Castro, M., Ih, D., Gager, J., Tammam, M., Dorkenwald, S., Collman, F., Schneider-Mizell, C., Brittain, D., Jordan, C.S., Dickinson, M., Pacureanu, A., Seung, H.S., Macrina, T., Lee, W.-C.A., Tuthill, J.C.: Connectomic reconstruction of a female drosophila ventral nerve cord. Nature **631**(8020), 360–368 (2024)

[11] Bates, A.S., Phelps, J.S., Kim, M., Yang, H.H., Matsliah, A., *et al.*: Distributed control circuits across a brain-and-cord connectome. Nature **656**(8129), 957–970 (2026)

[12] Lappalainen, J.K., Tschopp, F.D., Prakhya, S., McGill, M., Nern, A., Shinomiya, K., Takemura, S.-y., Gruntman, E., Macke, J.H., Turaga, S.C.: Connectome-constrained networks predict neural activity across the fly visual system. Nature **634**, 1132–1140 (2024)

[13] Shiu, P.K., Sterne, G.R., Spiller, N., Franconville, R., Sandoval, A., *et al.*: A drosophila computational brain model reveals sensorimotor processing. Nature **634**, 210–219 (2024)

[14] Wang-Chen, S., Stimpfling, V.A., Lam, T.K.C., Özdil, P.G., Genoud, L., Hurtak, F., Ramdya, P.: NeuroMechFly v2: simulating embodied sensorimotor control in adult drosophila. Nature Methods **21**, 2353–2362 (2024)

[15] Vaxenburg, R., Siwanowicz, I., Merel, J., Robie, A.A., Morrow, C., Novati, G., Stefanidi, Z., Both, G.-J., Card, G.M., Reiser, M.B., Botvinick, M.M., Branson, K.M., Tassa, Y., Turaga, S.C.: Whole-body physics simulation of fruit fly locomotion. Nature **643**, 1312–1320 (2025)

[16] Jin, Z., Zhu, Y., Zhang, C., Sui, Y.: Whole-Brain Connectomic Graph Model Enables Whole-Body Locomotion Control in Fruit Fly (2026)

[17] Hochreiter, S., Schmidhuber, J.: Flat minima. Neural Computation **9**(1), 1–42 (1997)

[18] Foret, P., Kleiner, A., Mobahi, H., Neyshabur, B.: Sharpness-aware minimization for efficiently improving generalization. In: International Conference on Learning Representations (2021)

[19] Yaniv, S.P., Schuldiner, O.: A fly's view of neuronal remodeling. WIREs Developmental Biology **5**(5), 618–635 (2016)

[20] Open X-Embodiment Collaboration, O'Neill, A., Rehman, A., *et al.*: Open x-embodiment: Robotic learning datasets and rt-x models. In: IEEE International Conference on Robotics and Automation, pp. 6892–6903 (2024)

[21] Ghosh, D., Walke, H.R., Pertsch, K., Black, K., Mees, O., Dasari, S., Hejna, J., Kreiman, T., Xu, C., Luo, J., Tan, Y.L., Chen, L.Y., Vuong, Q., Xiao, T., Sanketi, P.R., Sadigh, D., Finn, C., Levine, S.: Octo: An open-source generalist robot policy. In: Proceedings of Robotics: Science and Systems, Delft, Netherlands (2024)

[22] Chen, C.-L., Aymanns, F., Minegishi, R., Matsuda, V.D.V., Talabot, N., Günel, S., Dickson, B.J., Ramdya, P.: Ascending neurons convey behavioral state to integrative sensory and action selection brain regions. Nature Neuroscience **26**(4), 682–695 (2023)

[23] Braun, J., Hurtak, F., Wang-Chen, S., Ramdya, P.: Descending networks

transform command signals into population motor control. Nature **630**(8017), 686–694 (2024)

[24] Lin, A., Yang, R., Dorkenwald, S., Matsliah, A., Sterling, A.R., Schlegel, P., Yu, S.-c., McKellar, C.E., Costa, M., Eichler, K., Bates, A.S., *et al.*: Network statistics of the whole-brain connectome of drosophila. Nature **634**, 153–165 (2024)

[25] Destexhe, A., Marder, E.: Plasticity in single neuron and circuit computations. Nature **431**(7010), 789–795 (2004)

[26] Caroni, P., Donato, F., Muller, D.: Structural plasticity upon learning: regulation and functions. Nature Reviews Neuroscience **13**(7), 478–490 (2012)

[27] Roelfsema, P.R., Holtmaat, A.: Control of synaptic plasticity in deep cortical networks. Nature Reviews Neuroscience **19**(3), 166–180 (2018)

[28] Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., Mortrud, M.T., Ouellette, B., Nguyen, T.N., Sorensen, S.A., Slaughterbeck, C.R., Wakeman, W., Li, Y., Feng, D., Ho, A., Nicholas, E., Hirokawa, K.E., Bohn, P., Joines, K.M., Peng, H., Hawrylycz, M.J., Phillips, J.W., Hohmann, J.G., Wohnoutka, P., Gerfen, C.R., Koch, C., Bernard, A., Dang, C., Jones, A.R., Zeng, H.: A mesoscale connectome of the mouse brain. Nature **508**, 207–214 (2014)

[29] MICrONS Consortium: Functional connectomics spanning multiple areas of mouse visual cortex. Nature **640**, 435–447 (2025)

[30] Shapson-Coe, A., Januszewski, M., Berger, D.R., Pope, A., Wu, Y., Blakely, T., Schalek, R.L., Li, P.H., Wang, S., Maitin-Shepard, J., Karlupia, N., Dorkenwald, S., Sjostedt, E., Leavitt, L., Lee, D., Troidl, J., Collman, F., Bailey, L., Fitzmaurice, A., Kar, R., Field, B., Wu, H., Wagner-Carena, J., Aley, D., Lau, J., Lin, Z., Wei, D., Pfister, H., Peleg, A., Jain, V., Lichtman, J.W.: A petavoxel fragment of human cerebral cortex reconstructed at nanoscale resolution. Science **384**(6696), 4858 (2024)

[31] Schlegel, P., Yin, Y., Bates, A.S., Dorkenwald, S., Eichler, K., Brooks, P., Han, D.S., Gkantia, M., dos Santos, M., Munnelly, E.J., Sterling, A.R., Yu, S.-c., McKellar, C.E., Consortium, F., Costa, M., Seung, H.S., Murthy, M., Hartenstein, V., Bock, D.D., Jefferis, G.S.X.E.: Whole-brain annotation and multi-connectome cell typing of drosophila. Nature **634**, 139–152 (2024)

[32] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O.: Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347 (2017)

[33] Schulman, J., Moritz, P., Levine, S., Jordan, M.I., Abbeel, P.: High-dimensional continuous control using generalized advantage estimation. In: International

Conference on Learning Representations (2016)

[34] Ng, A.Y., Harada, D., Russell, S.: Policy invariance under reward transformations: Theory and application to reward shaping. In: International Conference on Machine Learning, pp. 278–287 (1999)

[35] Todorov, E., Erez, T., Tassa, Y.: MuJoCo: A physics engine for model-based control. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 5026–5033. IEEE, Vilamoura, Algarve, Portugal (2012)

[36] Towers, M., Kwiatkowski, A., Balis, J., De Cola, G., Deleu, T., Goulão, M., Kallinteris, A., Krimmel, M., KG, A., Perez-Vicente, R., Terry, J.K., Pierré, A., Schulhoff, S., Tai, J.J., Tan, H., Younis, O.G.: Gymnasium: A standard interface for reinforcement learning environments. In: Advances in Neural Information Processing Systems, vol. 38 (2025)

[37] Fu, J., Kumar, A., Nachum, O., Tucker, G., Levine, S.: D4rl: Datasets for deep data-driven reinforcement learning. In: International Conference on Learning Representations (2021)

[38] Zakka, K., Tassa, Y., MuJoCo Menagerie Contributors: MuJoCo Menagerie: A Collection of High-Quality Simulation Models for MuJoCo

[39] Li, H., Xu, Z., Taylor, G., Studer, C., Goldstein, T.: Visualizing the loss landscape of neural nets. In: Advances in Neural Information Processing Systems, vol. 31, pp. 6389–6399 (2018)

[40] Holm, S.: A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics **6**(2), 65–70 (1979)

A Implementation Details

Training settings. Supplementary Table S1 summarizes PPO settings. Policy and value clipping thresholds are both 0.2, the value-loss weight is 0.5 and the gradient-norm limit is 0.5, with reward normalization applied throughout. Main fly and additional-embodiment policies use fixed observation scaling, zero-initialized outputs, separate gradient clipping for actor and critic, and action likelihoods evaluated at the stored pre-squashing coordinates. Their MLP actors contain two hidden layers with exponential-linear activations. Actor and critic learning rates start at 10^{-4}. The adaptive actor learning rate targets KL divergence of 0.015, stops updates at 0.03 and is constrained to 10^{-5}–10^{-3}, with a 10^{-6} floor during speed continuation. Log-standard-deviation learning rates are 10^{-3} for fly tasks and 10^{-4} for additional embodiments.

The discount factor γ and advantage-decay parameter λ are parameterized as $\gamma = e^{-\Delta t/T_\gamma}$ and $\lambda = e^{-\Delta t/T_\lambda}$, where Δt is the decision interval and T_γ and T_λ are the corresponding timescales. We use $(T_\gamma, T_\lambda) = (5, 0.8)$ s for terrain and additional embodiments and $(2, 0.2)$ s for turning and speed. Potential-based shaping subtracts the current potential from the discounted next-state potential, with terminal potential set to zero. Microstep rewards are discounted and summed within each control decision, and fly-task deadlines are treated as terminal. Ant observation normalization is frozen during evaluation. Stability refinement caps rewarded speed at 2.2 m/s and penalizes instability. Flat-ground Ant evaluation uses the native environment dynamics and termination rules with an action scale of 0.9.

Table S1. PPO configurations.

Setting	Main fly tasks	Other embodiments	Ant pretraining
Parallel environments	8	4	4
Rollout length	32	256	512
Sequence length	8	32	64
Optimization epochs	4	4	10
Minibatches	8	8	4 for Adult & MLP, 32 for others
Decision interval	40 ms	20 ms for Cassie, 40 ms for others	Native step
Discounting	Task-specific timescales	Task-specific timescales	$\gamma = 0.99$, $\lambda = 0.95$
Learning rates	10^{-4} critic, adaptive actor	10^{-4} critic, adaptive actor	3×10^{-4} base, 10^{-4} refinement

Fly interfaces and metrics. Terrain uses 50-μs physics steps and 172-feature observation frames that include block height. The MLP concatenates four successive frames. Initial tripod phase is sampled across eight bins with within-bin randomization. The block spans forward positions of 3.0–4.3 mm, with its placement and task coordinates updated after the run-in. Residual corrections modulate a 12-Hz gait,

with the norm of each leg-specific correction bounded by $0.15\sqrt{3}$ rad and correction authority increasing from zero to full over forward positions of 1.8–2.6 mm. The reward combines minimum-foot progress and crossing success with penalties for failure, posture deviation, lateral deviation and correction magnitude. Lateral retention requires maximum absolute lateral displacement within 1.5 mm. Residual variation is computed by summing absolute successive changes across outputs, averaging across output dimensions and normalizing by duration, with optional normalization by root-mean-square (RMS) amplitude.

Turning and speed use 100-μs physics steps. Turning uses 64 observation features, including target error, timing and preceding commands. Its reward combines heading alignment, forward progress and uprightness with penalties for rotation and control effort. Command roughness is the RMS successive command change divided by the centred RMS command amplitude over 0.4–2 s. Total variation sums the Euclidean changes in policy output, including the initial change from zero. Speed uses 51 observation features, including target speed, measured speed, tracking error and heading. Its gait interface adjusts amplitude, duty factor and inter-leg phase, with speed and yaw rate causally averaged over 100 ms. The reward combines speed tracking and uprightness with penalties for heading error, yaw motion and failure. Ramp commands connect speed targets in the ranges 8–10 and 14–16 mm/s with slopes of 2–8 mm/s^2 and durations of at least 1 s. Sinusoidal commands use amplitudes of 2–4 mm/s and frequencies of 0.15–0.30 Hz. Training comprises 12,800 initial and 3,072 continuation decisions, with the total budget and continuation learning-rate floor selected in an earlier eight-pair development study. Ramp-onset increments compare each 100-ms window with the mean speed over the preceding 100 ms.

Maze adaptation. Supplementary Tables S2 and S3 specify controller architectures and adaptation settings. Reservoir controllers use a leak rate of 0.5, a fixed orthogonal recurrent matrix with spectral radius 0.9 and nonlinear readouts terminating in linear output layers. The additional comparator families reserve four goal-coordinate inputs during locomotor pretraining. The Adult-connectome controller retains its pretrained locomotor mapping and trains a residual module receiving recurrent motor activity, body state and goal features. Other controllers update their trainable action networks, while routing modules and fixed reservoir components remain frozen. The Adult and MLP adaptation modules contain 89,992 and 96,008 trainable parameters, respectively.

U-Maze adapts locomotion checkpoints using teacher trajectories followed by teacher-labelled student rollouts, with a grid-waypoint planner supplying local goals. Medium uses one fixed source policy per architecture, whereas Large starts from the corresponding Medium checkpoint. Routers for Medium and Large receive the final goal as input. The additional comparator families in Medium use U-Maze source policies and the shared MLP router together with spatial anchors and fixed routing rules. Large uses a single router shared across all controller families, trained with 120,000 teacher steps and 60,000 updates using walkable-cell anchors and a 0.45-m latch radius. Burn-in steps update recurrent state before the supervised loss is

evaluated. Applied actions follow

$$\mathbf{a}_t = \beta s \widehat{\mathbf{a}}_t + (1-\beta)\mathbf{a}_{t-1},$$

where t indexes control decisions, $\mathbf{a}_t$ is the action applied to the environment, $\widehat{\mathbf{a}}_t$ is the bounded student-policy output, β is the smoothing weight and s is the action scale.

Deadline and efficiency-threshold curves retain all evaluation episodes in the denominator. The Medium-maze deadline endpoint is the area under the success-by-deadline curve over 0–50 s, normalized by the 50-s evaluation horizon. Route complexity is defined as the number of turns along the shortest grid route. Friction tests use ten paired Adult-connectome and MLP models at floor-friction coefficients of 0.50, 0.75, 1.25 and 1.50.

Table S2. Maze comparator architectures.

Family	Architecture	Trainable components
MLP	Two 256-unit rectified-linear hidden layers	Action network
GRU	One 128-unit recurrent layer	Recurrent action network
LSTM	One 108-unit recurrent layer	Recurrent action network
Transformer	One causal block with context length 16, width 64, 4 attention heads and feedforward width 512	Action network with learned positional embeddings and no dropout
Reservoir	512 leaky $\tanh$ units with a 72-unit $\tanh$ readout	Input projection and readout
Reservoir+RO	Same reservoir with a 176-unit $\tanh$ readout	Readout only

Table S3. Supervised maze adaptation.

Setting	U-Maze	Medium	Large
Teacher steps	40,000	60,000 (shared)	60,000 (shared)
Teacher-labelled student steps	30,000	–	–
Adaptation updates	10,000 per stage	5,000	5,000
Batch size	32	64	64
Learning rate	3×10^{-4}	10^{-4}	10^{-4}
Adult-connectome burn-in	8	8	8
MLP burn-in	0	8	8
Other-family burn-in	16	16	16
Observation clip bound	10	30	60
Smoothing weight β	0.8	0.95	0.95
Connectome action scale s	1	0.8	0.8
Other-family action scale s	1	1	1

Table S4. Body-specific motor interfaces.

Body	Motor coordination	Actuators	Target speed
Cassie	Bilateral balance coordinates with learned local feedback	10	0 m/s
Go1	Phase-conditioned joint control with learned local feedback	12	0.8 m/s
Tripod	Phase-conditioned joint control with learned local feedback	6	0.1 m/s
Hexapod	Alternating-tripod CPG with learned joint corrections and a normalized graph readout	18	0.5 m/s

Additional-embodiment perturbations. Controllers are trained from scratch for each body using the interfaces in Supplementary Table S4. Inputs encode scaled body state, joint state, preceding control and task commands. Push, action-delay and observation-loss perturbations begin at 1 s for Cassie and 8 s for the walking embodiments. Kinematic analyses use windows spanning 0–1 s, pre-perturbation intervals of 0.5–1 s for Cassie or 2–8 s for the walking embodiments, 0.4–2.4 s after perturbation onset, 10–20 s, and 40–60 s during endurance evaluation. Tilt and heading are summarized using RMS values, while speed variability is quantified by the standard deviation of successive 1-s forward-speed estimates. Path-retention duration ends at lateral-boundary crossing or episode termination and is capped by the evaluation horizon. The displayed Go1 pre-perturbation metrics retain 393 of 400 matched episode pairs.

Mechanistic assays. Fixed-history probes retain the recorded preceding-action inputs. Relative-enrichment analyses exclude cases with zero-output initial readouts. Turning activity is quantified as the squared per-cell activation difference between $\pm 60^\circ$ turning trials and matched 0° trials, averaged across three states per direction over the 1.7 s following cue onset.

Speed-sensitivity analyses use early and middle windows from four accelerating and four decelerating ramps, with matched Adult and MLP histories and propagation horizons up to 320 ms. Target inputs are held at their ramp-onset values. Signed activity is defined as the class-mean difference between original and target-held activations. Drive loss is measured relative to the all-class target hold over 0.8–1.0 s after ramp onset. Target-response probes apply ± 1 mm/s perturbations while updating tracking-error inputs consistently. The five predictors are target response, measured-speed response, parameter-perturbation KL, bounded-action change and speed-input projection gain. Associations are evaluated against three tracking outcomes within Adult-connectome controllers, within MLP controllers and across paired architecture differences. Response advantage is defined as Adult minus MLP response, whereas tracking advantage is defined as MLP minus Adult error.

Randomized terrain controllers use the same training budget as the Adult terrain controllers. Distinct non-self edges are sampled uniformly, and signed weights are permuted without renormalization. Eight phases, five obstacle heights and both Adult and ER controllers yield 80 shared histories per model. Sensitivity is evaluated over

propagation horizons of 40–320 ms at 10-ms intervals, with the final sensory window repeated beyond 80 ms. Blockade effects are quantified as the 40-ms output change normalized by the RMS output of the intact controller. Mechanosensory and olfactory controls use five masks per target class that are disjoint from the target population and matched for cell count, interface partition and total-degree decile. Total degree is the sum of incoming and outgoing edge counts, and mask effects are averaged within each model.

Six supplementary Adult models replay 80 histories beginning at 0.12 s over propagation horizons of 10–320 ms. Unit-state sensitivity sums squared output derivatives across the six bounded, ungated policy outputs. Activity-scaled sensitivity weights each neuron's squared sensitivity by the squared magnitude of its initial recurrent state. Absolute sensitivity is defined as the per-neuron RMS output derivative, and enrichment is normalized to the 18,756 afferent neurons. Obstacle-evoked activity is evaluated across five obstacle heights and 64 phases, with block and flat-ground trials matched by model, phase and obstacle height. Height-cue exchanges isolate the contribution of the cue itself. Activation differences are aligned to the time at which the first foot reaches the front edge of the block, retaining only shared frames and complete windows from -100 to $+200$ ms.

KL-surface sampling. Terrain surfaces use two-decision windows drawn from the 80 shared histories described above. Turning surfaces use 48 shared four-decision windows per model and are averaged across all 24 model pairs. Speed surfaces perturb trainable actor parameters along directions normalized by the corresponding trained matrix-row norms or, for vector parameters, by tensor norms, while keeping graph weights, critic parameters and action variance fixed. Each model uses 24 four-decision windows drawn from four matched ramp and sinusoidal histories, and the displayed surfaces are averaged across all 20 model pairs. An additional eight-direction analysis evaluates both perturbation signs at radii of 0.05, 0.10 and 0.20. Auxiliary analyses report unit-variance KL and changes in bounded policy outputs.

Developmental comparison. Six models per connectome control the adult fly body through a 51-feature sensory interface and a two-command motor interface. The processed Larva connectome contains 2,952 neurons and 63,545 connections weighted by non-negative synapse counts. Inputs are sensory neurons, while outputs comprise descending neurons to the subesophageal zone (DN-SEZ) and ring gland neurons (RGNs). Larval weights use a global scaling factor of 0.04 for turning and outgoing row-ℓ_1 normalization for speed, whereas the Adult connectome uses outgoing row normalization. Final-budget policies use one recurrent update per 10-ms control decision, linear readouts, a two-layer critic with 128 units per layer and 100-μs physics steps. PPO uses 16 environments, 160-step rollouts, 32-step sequences, 2 optimization epochs, 5 minibatches, a learning rate of 10^{-4}, $\gamma = 0.99$ and $\lambda = 0.95$, for a total of 499,200 decisions per model.

Directional trials last 1.5 s and use 20 initializations per direction for each model. Outcomes are rotation sign and net rotation aligned with the commanded direction. Speed training samples fixed targets uniformly over 8–16 mm/s for 1.5 s. Dynamic reevaluation uses five initializations per condition for 8-to-16-mm/s and 16-to-8-mm/s step commands and sinusoidal commands at 0.25 and 0.75 Hz. Velocity

projected onto the initial heading is exponentially filtered with a coefficient of 0.25. Step-response latency is defined as the time to first attain 90% of the commanded speed change. Sinusoidal amplitude ratio and phase lag are fitted over 1–9 s.

BANC readouts and regional interventions. The Brain And Nerve Cord (BANC) connectome covers the adult *Drosophila* brain and VNC [11]. The processed graph contains 112,885 neurons and 2,672,709 directed connections, with weights normalized by the sum of absolute outgoing weights. The full graph remains recurrently active in both readout conditions. A learned projection maps 51 observations to 12,250 afferent neurons, while the readout selects either 1,312 descending neurons or 831 motor neurons to drive a two-command CPG interface. Six models per readout follow the supplementary fixed-target speed protocol. Initial and trained policies replay 32 shared step and ramp histories drawn from four controller families. Post-switch squared drive sensitivity is evaluated at a ten-update horizon, normalized per cell to the BANC afferent population and aggregated over 16 neuronal classes matched by root identifiers and annotations.

Three recovery checkpoints per readout are trained for approximately 0.5 million decisions with a 40-ms push of magnitude 10 applied at 0.75 s in simulator force units. Nine matched states per checkpoint, comprising three initializations and three CPG phases, receive either no push or 40-ms pushes of ± 10, ± 15 and ± 20. Regional message substitution modifies each connection contribution as

$$W_{ij}\left[(1-f_{ij})r_i + f_{ij}r_i^0\right],$$

where f_{ij} is the fraction of synapses on edge $i \to j$ assigned to the substituted anatomical region, r_i is the current presynaptic activation and r_i^0 is the matched activation from the unpushed reference branch. Intervention branches share physical, CPG and recurrent states at the point of branching. Synapse-location fractions retain neck contributions without renormalization.

Fixed-history probes replace messages originating from the brain, VNC, VNC interneurons, VNC interneurons projecting to descending neurons, or ascending neurons projecting to the brain. Each checkpoint uses one reference state and six nonzero-force branches. Action-change norms are normalized by the intact pushed-versus-unpushed response. Closed-loop tests replace brain-derived, VNC-derived or combined brain-and-VNC messages, or replace all observations, with their matched unpushed references. Recovery requires survival for 0.5 s and terminal heading error no greater than 0.25 rad. Error integrals assign π rad after termination, with intervention effects averaged across states, push directions and push magnitudes within each checkpoint.

Statistical details. Holm correction families include both displayed and undisplayed contrasts. Terrain analyses comprise five success conditions, 26 secondary endpoints across five obstacle heights, blockade tests across two neuronal classes and three nonzero blockade levels, and 45 Adult-connectome versus ER contrasts. Held-out gain analyses comprise four axis-by-metric contrasts and two surface-metric contrasts. Turning analyses comprise four primary and 32 secondary behavioural contrasts, input-learning analyses across nine neuronal classes and two endpoints,

four pathway-specificity contrasts and ten intervention-versus-intact contrasts. Input-learning contrasts use base-2 log ratios. Speed analyses comprise three behavioural outcomes and 45 neural-behavioural associations. The associations are tested with 19,999 random permutations together with the observed assignment.

Ant analyses comprise 26 main contrasts, including the exploratory deadline integral, and 25 friction contrasts. Additional-embodiment scores comprise four body-level comparisons. Kinematic analyses treat the ten endpoints prespecified before cohort expansion separately from 3,150 exploratory comparisons across episodes, analysis windows and nominal-control differences, including duration and heading measures. Developmental analyses comprise 22 terrain and turning contrasts and 20 speed contrasts.

Bootstrap analyses use 50,000 resamples for terrain behaviour and Ant, 20,000 for fly mechanistic analyses, turning command metrics, speed and BANC sensitivity, and 10,000 for the additional embodiments. Turning activity, pathway effects and input-learning log ratios use t-based intervals, with geometric ratios obtained by back-transformation. Ant displacement, seven-family maze comparisons and developmental speed summaries report between-model standard deviations. BANC sensitivity and regional-intervention effects are reported descriptively. For visualization, Ant displacement and multibody lateral trajectories retain the final position after episode termination.

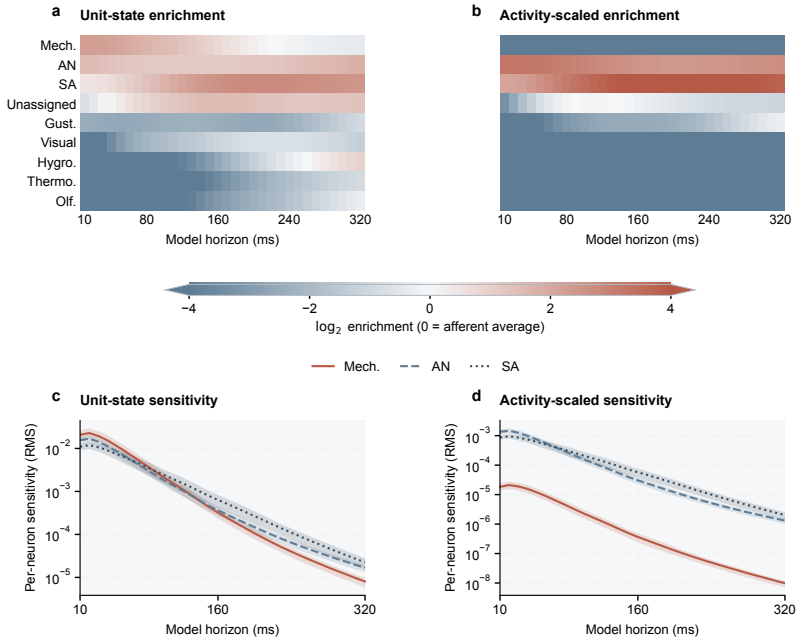


Fig. S1 Temporal organization of neuronal output sensitivity during terrain adaptation. a,b, Per-neuron output-sensitivity enrichment across connectome-defined neuronal populations during terrain control under unit-state perturbations (**a**) and activity-scaled perturbations (**b**), normalized to the mean sensitivity across afferent neurons. Colours show mean $\log_2$ enrichment across six independently trained Adult-connectome controllers over propagation horizons of 10–320 ms. **c,d,** Absolute per-neuron sensitivity of mechanosensory neurons, ascending neurons (ANs) and sensory-ascending neurons (SAs) under the corresponding unit-state (**c**) and activity-scaled (**d**) perturbations. Lines and shading denote means and 95% bootstrap intervals across controllers.

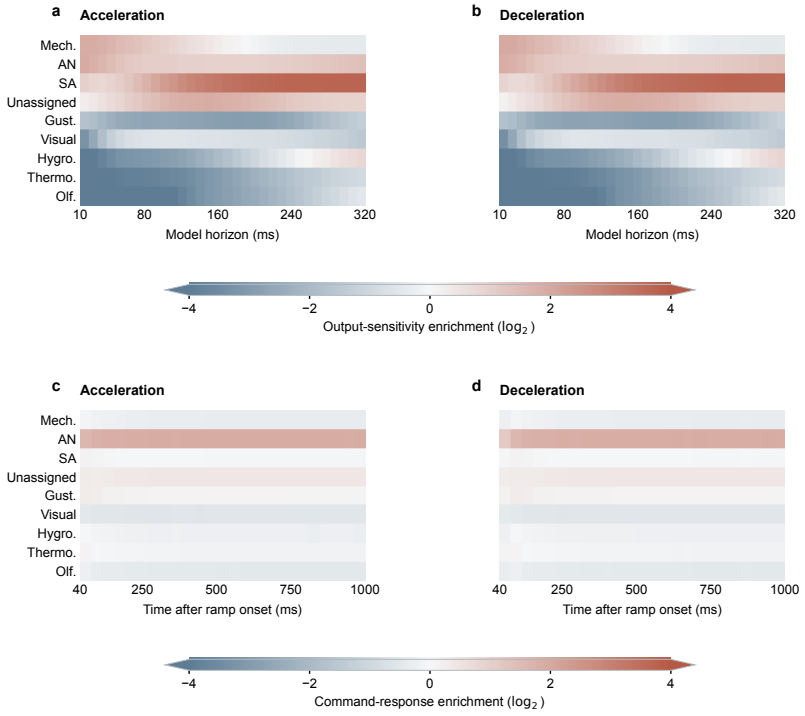


Fig. S2 Neural dynamics during speed adaptation. a,b, Cell-count-normalized enrichment of future-drive sensitivity across connectome-defined neuronal populations during acceleration (**a**) and deceleration (**b**), evaluated over propagation horizons of 10–320 ms. **c,d,** Neural-response enrichment evoked by changing speed commands during acceleration (**c**) and deceleration (**d**), measured relative to replay with target speed held at its onset value. Heatmaps show mean $\log_2$ enrichment across 20 Adult-connectome controllers, with zero denoting the population-average response per afferent neuron.

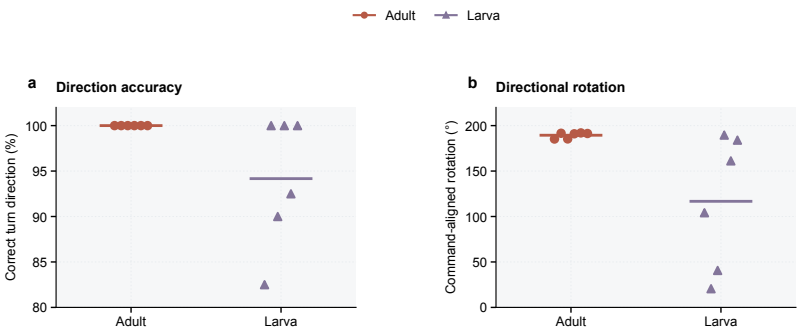


Fig. S3 Developmental connectome comparison in directional turning control. a, Correct-direction fraction for Adult- and Larva-connectome controllers under left- and right-turn commands. **b,** Net body rotation aligned with the requested direction. Each dot denotes one of six independently trained models per group, averaged over 20 evaluation initializations for each direction (480 episodes in total). Horizontal segments denote group means. Both connectomes support reliable directional control in the adult fly body, with the Adult connectome showing greater mean command-aligned rotation and lower between-model variability.

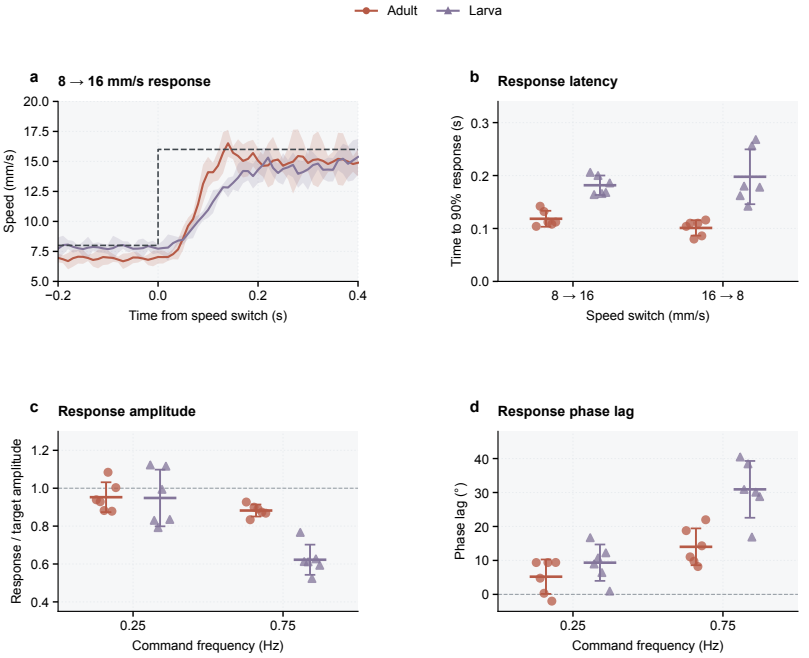


Fig. S4 Developmental connectome comparison in dynamic speed control. a, Response to an $8 \rightarrow 16$ mm/s speed switch for Adult- and Larva-connectome controllers. The dashed line denotes the target command. Lines and bands show mean $\pm$ s.d. across six trained models. **b,** Response latency, measured as the time to first attain 90% of the commanded speed change, during acceleration and deceleration. **c,d,** Response amplitude relative to the target amplitude (**c**) and phase lag (**d**) under sinusoidal commands at 0.25 and 0.75 Hz. Dots denote individual model means. Horizontal bars and capped vertical bars denote group means and s.d. Adult-connectome controllers respond faster during both acceleration and deceleration and retain response amplitudes closer to the target at the higher command frequency.

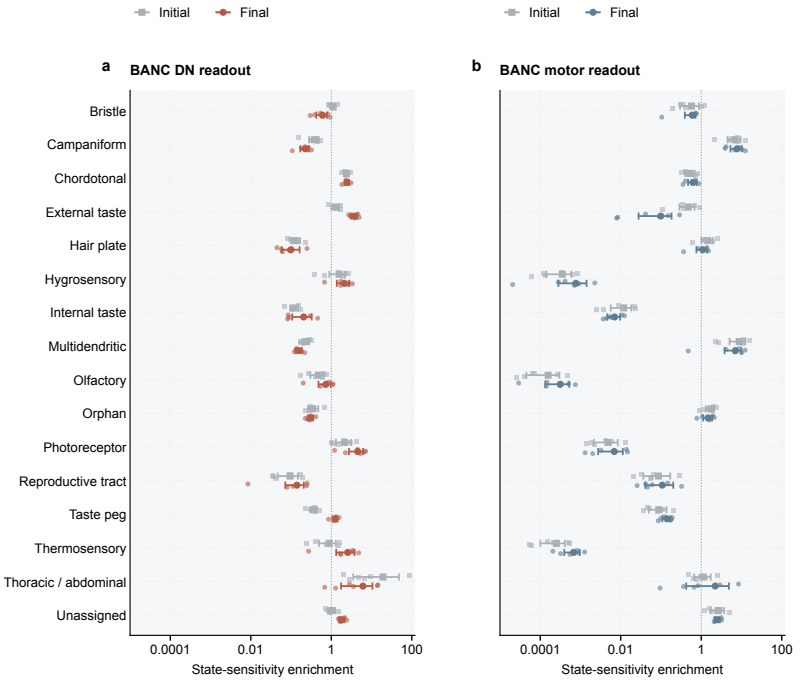


Fig. S5 Readout-dependent sensitivity organization in BANC speed controllers. a,b, State-sensitivity enrichment across 16 input classes for BANC controllers using descending-neuron (DN) or motor-neuron readouts, shown before and after training. Enrichment is defined as each class's share of squared drive-output sensitivity divided by its share of input neurons, with one denoting the population-average sensitivity per neuron. Points denote 6 independently trained models per readout, and intervals show 95% model-bootstrap intervals. Both readouts develop class-selective sensitivity profiles after training, with the DN readout showing a more structured concentration of sensitivity across input populations.

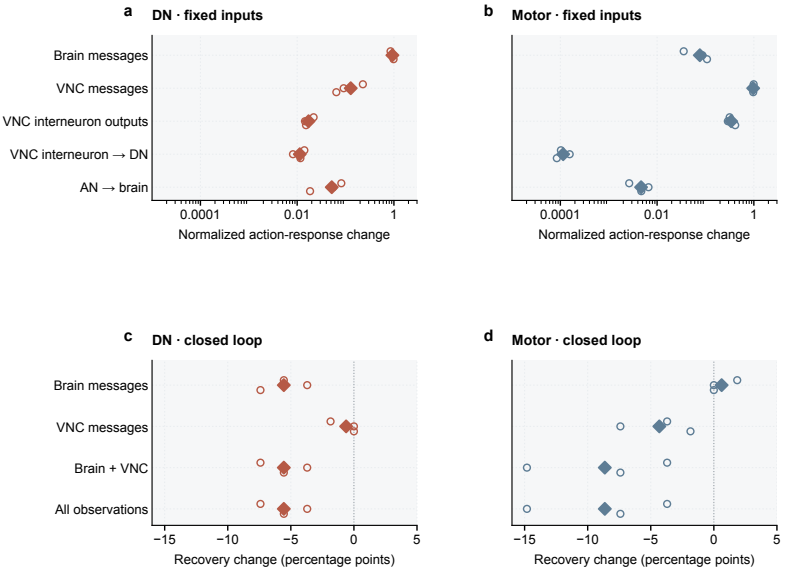


Fig. S6 Readout-dependent regional organization of feedback in BANC controllers. a,b, Normalized changes in action response after replacing messages from the indicated brain and ventral nerve cord (VNC) regions or pathways with matched reference messages, evaluated on fixed recorded observations for BANC controllers using descending-neuron (DN) or motor-neuron readouts. **c,d,** Changes in joint survival-and-heading recovery under the corresponding regional substitutions in closed-loop evaluations, averaged across push magnitudes and directions. Circles denote three independently trained checkpoints per readout, and diamonds denote means. Brain-message substitution produces a larger normalized action-response change for the DN readout (0.930 versus 0.076), whereas VNC-message substitution has the larger effect for the motor-neuron readout (0.969 versus 0.128). The corresponding closed-loop substitutions reduce mean recovery by 5.56 and 4.32 percentage points for DN-readout brain interventions and motor-readout VNC interventions, respectively.

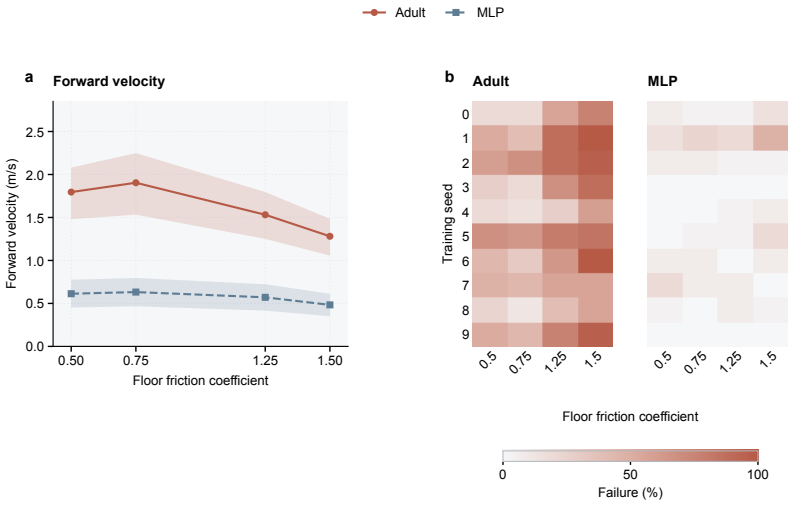


Fig. S7 Forward velocity and stability across floor friction in Ant. a, Forward velocity of Adult-connectome and matched MLP controllers across four floor-friction coefficients, summarized over 10 paired training seeds. **b,** Physical failure fraction for each training seed, controller and friction condition. Connectome control maintains higher forward velocity across friction conditions while also exhibiting higher failure fractions, complementing the flat-ground Ant results.

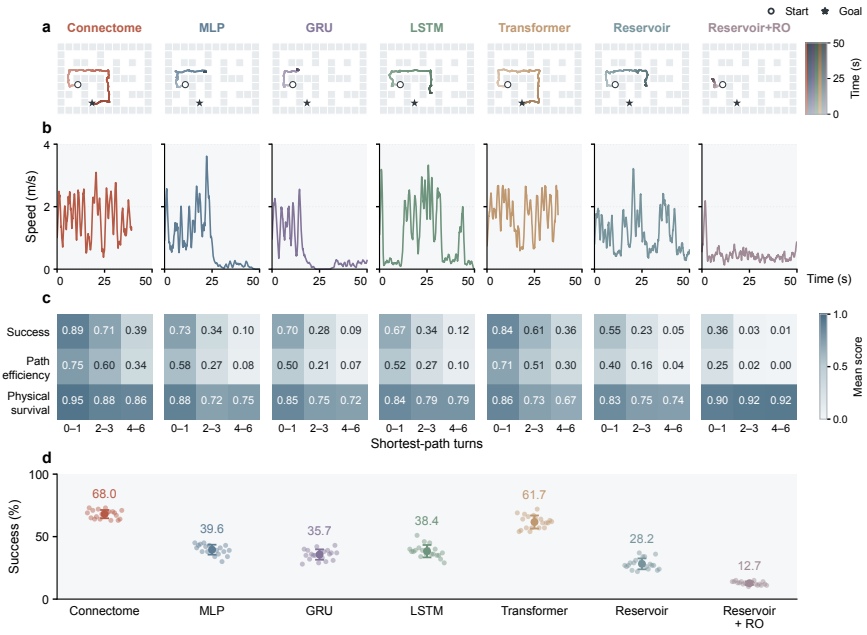


Fig. S8 Navigation and performance of seven controller families in the Large maze. a, Navigation trajectories from a matched initial state to the same goal in a 48×36-m maze. Model-specific colours darken with elapsed time (0–50 s), with the seven adjacent gradients following the column order. Circles mark starts and stars mark goals. Each column shows one illustrative episode selected by registered index (optimization seed 0, reset 9400000), without selection by outcome. **b,** Planar speed derived from consecutive position samples for the episodes in **a**, shown with a centred ± 0.5-s moving average (21 samples). Boundary windows use only observed samples and are renormalized. **c,** Success fraction, success-weighted path efficiency (SPL) and physical survival, stratified by shortest-path turn count. The 0–1, 2–3 and 4–6 turn groups contain 31, 42 and 27 paired resets, respectively. Colours and numbers denote means across 20 optimization seeds on a shared 0–1 scale. Physical survival does not imply navigation success. **d,** Overall success across 100 paired resets per optimization seed. Large dots and error bars show mean $\pm$ s.d. across 20 seeds per controller family. Numbers denote mean success rates in percent. All 14,000 episodes are retained.

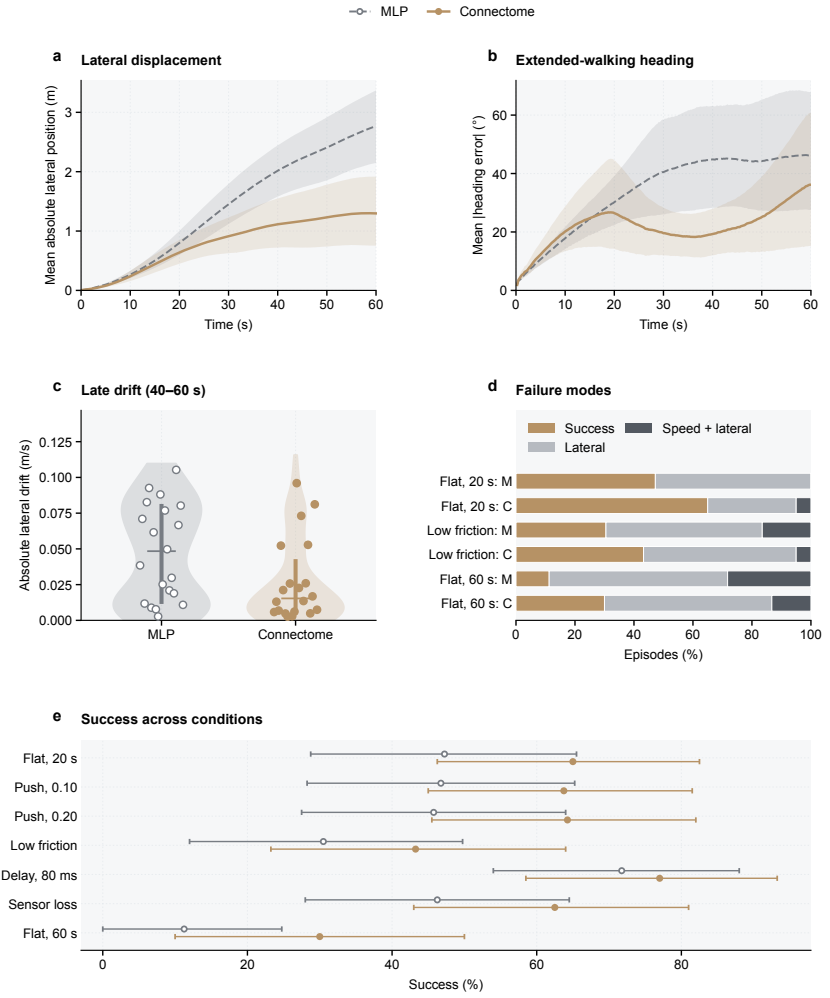


Fig. S9 Directional and path stability in the radial tripod. a,b, Mean absolute lateral displacement (**a**) and heading error (**b**) during 60-s flat-ground walking, relative to the first archived state. Heading error is wrapped to $\pm 180°$. **c,** Absolute net lateral drift rate over the final 20 s. **d,** Outcome composition in 20-s flat-ground, low-friction and 60-s flat-ground trials, separating successful episodes from lateral and combined speed-lateral failures. M and C denote MLP and Connectome controllers, respectively. **e,** Task success across seven evaluation conditions. Each condition includes 20 paired training seeds and 20 matched evaluation initial states per seed (400 episode pairs). All episodes remain physically stable and contribute throughout the corresponding evaluation horizon. Violins show episode distributions with interquartile ranges and medians. Open MLP and filled Connectome markers denote seed means. Curves and condition summaries show equally weighted seed means with 95% seed-bootstrap intervals (10,000 resamples).

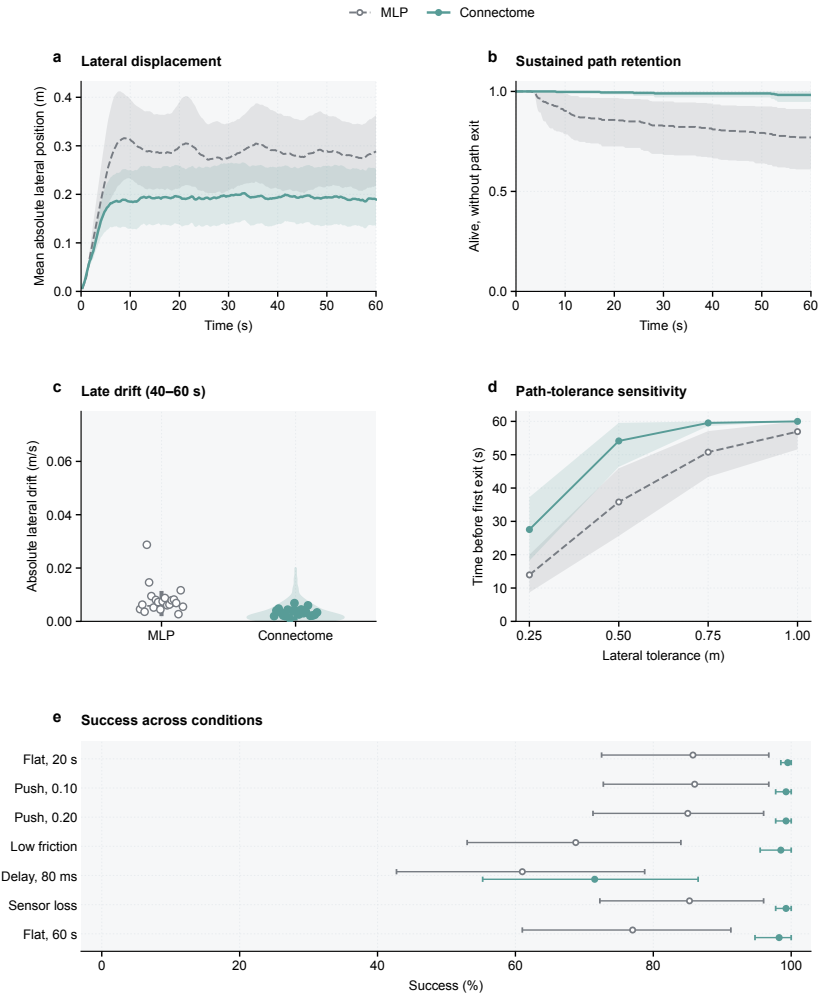


Fig. S10 Path stability in the hexapod with a shared analytic gait interface. a, Mean absolute lateral displacement during 60-s flat-ground walking, relative to the first archived state. **b,** Fraction of trials remaining physically stable without exiting the ± 0.75-m lateral corridor. **c,** Absolute net lateral drift rate over the final 20 s. **d,** Mean time to first corridor exit across four lateral tolerances, bounded by the 60-s evaluation horizon. **e,** Task success across seven evaluation conditions. Both controllers use the same analytic CPG and differ in the learned control architecture. Each condition includes 20 paired training seeds and 20 matched evaluation initial states per seed (400 episode pairs). All episodes remain physically stable and contribute throughout the corresponding evaluation horizon. Violins show episode distributions with interquartile ranges and medians. Open MLP and filled Connectome markers denote seed means. Curves and condition summaries show equally weighted seed means with 95% seed-bootstrap intervals (10,000 resamples).