Transcriptomic analysis of the trade-off between endurance and burst-performance in the frog *Xenopus allofraseri*

Valérie Ducret 1*, Adam J. Richards 2, Mathieu Videlier 3, Thibault Scalvenzi 4, Karen A. Moore 5, Konrad Paszkiewicz 5, Camille Bonneaud 2,6, Nicolas Pollet 4 and Anthony Herrel 2,7

**Abstract**

**Background:** Variation in locomotor capacity among animals often reflects adaptations to different environments. Despite evidence that physical performance is heritable, the molecular basis of locomotor performance and performance trade-offs remains poorly understood. In this study we identify the genes, signaling pathways, and regulatory processes possibly responsible for the trade-off between burst performance and endurance observed in *Xenopus allofraseri*, using a transcriptomic approach.

**Results:** We obtained a total of about 121 million paired-end reads from Illumina RNA sequencing and analyzed 218,541 transcripts obtained from a de novo assembly. We identified 109 transcripts with a significant differential expression between endurant and burst performant individuals (FDR \( \leq 0.05 \) and logFC \( \geq 2 \)), and blast searches resulted in 103 protein-coding genes. We found major differences between endurant and burst-performant individuals in the expression of genes involved in the polymerization and ATPase activity of actin filaments, cellular trafficking, proteoglycans and extracellular proteins secreted, lipid metabolism, mitochondrial activity and regulators of signaling cascades. Remarkably, we revealed transcript isoforms of key genes with functions in metabolism, apoptosis, nuclear export and as a transcriptional corepressor, expressed in either burst-performant or endurant individuals. Lastly, we find two up-regulated transcripts in burst-performant individuals that correspond to the expression of myosin-binding protein C fast-type (mybpc2). This suggests the presence of mybpc2 homoeologs and may have been favored by selection to permit fast and powerful locomotion.

**Conclusion:** These results suggest that the differential expression of genes belonging to the pathways of calcium signaling, endoplasmic reticulum stress responses and striated muscle contraction, in addition to the use of alternative splicing and effectors of cellular activity underlie locomotor performance trade-offs. Ultimately, our transcriptomic analysis offers new perspectives for future analyses of the role of single nucleotide variants, homoeology and alternative splicing in the evolution of locomotor performance trade-offs.

**Keywords:** Anura, Limb, Muscle, Myosin, RNA-sequencing, Stamina

* Correspondence: Valerie.ducret@gmail.com
1 UMR 7179 MECADeV, CNRS/MNHN, Département Adaptations du Vivant, 55 Rue Buffon, 75005 Paris, France
Full list of author information is available at the end of the article

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
Background
Locomotor performance has a strong impact on the survival and reproduction of many organisms [1–3]. Burst performance is often most relevant in the context of prey capture and predator escape, whereas endurance is relevant in the context of territory defense, dispersal, or migration. Yet, the evolution of locomotor performance can be constrained if performance traits are involved in traded-offs, as often observed between burst performance and endurance capacity in vertebrates [4–9]. Conflicting demands on muscles to express either fast-twitch glycolytic fibers that facilitate burst performance or slow-twitch oxidative muscle fibers that enhance stamina may explain in part this performance trade-off [10–13]. Although the physiological basis of this performance trade-off has been documented, how it is governed at the gene expression level remains poorly understood. Uncovering the molecular basis and biological pathways underlying performance trade-offs is therefore essential for understanding the adaptive evolution of these traits.

Because locomotor performance is heritable [14–16], efforts have been made to explain differences in physical performance by variation in coding DNA in humans [17–19], racing pigeons [20], mice [21], horses [22] and dogs [23]. While those studies highlighted a remarkable number of genetic variants associated with variation in physical performance, they provide little insight into the potential processes underlying performance trade-offs. Altogether, the myriad of genetic variants with little phenotypic effects has led to the consensus that physical performance is a polygenic trait that is governed by features such as transcriptional regulation. Recently, microRNAs have been found to regulate the expression of target genes in skeletal muscle [24, 25], as well as target genes involved in muscle cell proliferation, differentiation, motility and regeneration [26]. In humans, a transcriptional map established after endurance exercise training highlighted an important regulation of gene expression to increase aerobic capacity [27]. Although a few transcriptomic analyses have been performed in the context of physical performance [20, 27, 28], none have tried to understand the factors underlying performance trade-offs.

In this study, we analyzed the transcriptomes of eight adult *Xenopus allofraseri* males from a single population that show a marked trade-off between endurance and burst-performance capacity. We performed a RNA-seq analysis of genes expressed in limb muscle that allowed us to highlight the genes, signaling pathways, and regulatory processes such as alternative splicing likely underlying this locomotor performance trade-off.

Results and discussion
Raw sequencing data, de novo assembly and quality control
We obtained a total of about 121 million paired-end reads using Illumina RNA sequencing. After trimming and quality filtering, biological replicates produced between 5.2 and 28 million paired-end reads (Table 1). The number of reads in each group was well balanced with 5.5 million in the endurant group and 6.6 million in the burst-performant group. The BUSCO analysis resulted in 65.4% gene identification (54.9% completeness and 10.5% of fragmented genes), which is relatively good as only one muscle tissue was sampled. Next, we evaluated the Trinity de novo assemblies by mapping the trimmed reads. We obtained an overall alignment rate of > 97% percent identity and > 89% of reads aligned as proper pairs (Table 1). The de novo assembly consisted of 218,541 transcripts and 163,981 genes with an E90N50 value (i.e. the N50 for transcripts that represent 90% of the total normalized expression data) of 1462 pb. These different metrics testify that our transcriptome assemblies were of good quality.

Physical performance
Transcript levels were quantified with respect to endurant and burst performant classifications after measuring four physical performance traits: maximum distance jumped before exhaustion (m), maximum time jumped before exhaustion (s), maximum burst velocity (m.s\(^{-1}\)), and maximum burst acceleration (m.s\(^{-2}\)) (Table 2). The

| Sample | Paired-end reads | Total singleton reads | > = Q30 (%) | Mean quality score |
|--------|-----------------|----------------------|-------------|-------------------|
| A      | 8,961,655       | 17,923,310           | 94.02       | 36.42             |
| B      | 16,672,325      | 33,344,650           | 93.97       | 36.41             |
| C      | 28,509,462      | 57,018,924           | 93.71       | 36.31             |
| D      | 14,145,979      | 28,291,958           | 94.07       | 36.44             |
| E      | 20,664,890      | 41,329,780           | 94.19       | 36.49             |
| F      | 15,697,497      | 31,394,994           | 94.06       | 36.44             |
| G      | 11,647,327      | 23,294,654           | 94.37       | 36.54             |
| H      | 5,288,288       | 10,576,576           | 94.00       | 36.41             |

Q30: Phred quality scores when probability of incorrect base call is 1 in 1000
principal component analysis (PCA) followed by the agglomerative hierarchical clustering allowed to clearly segregate individuals into the two groups (burst performant vs. endurant individuals; Fig. 1) confirming the existence of a locomotor trade-off in this species. Maximum distance, maximum time and maximum velocity contributed mainly to the first axis of the PCA (respectively 92.1, 90.5 and 81.3%), whereas maximum acceleration contributed to the second axis (75.3%).

Phylogenetic analysis
Phylogenetic analysis of the mitogenomes indicated that mitochondrial DNA from the eight *Xenopus* males (Sample A to H, Fig. 2) are closely related and correspond to specimens of the species *Xenopus allofraseri*. These mitochondrial sequences are sister to those of *Xenopus pygmaeus* and markedly diverge from other *Xenopus* species such as *Xenopus laevis* and *Xenopus tropicalis*. Noticeably, the eight *Xenopus allofraseri* males were captured in a geographic range that was not previously reported for this species [29].

Differentially expressed transcripts
We identified 109 transcripts with a significant differential expression between endurant and burst performant individuals (Fig. 3). Six of those transcripts yielded no similarities to either the Uniprot or the NCBI databases. The blast searches resulted in 103 protein-coding genes (Table S1) matching either *Xenopus laevis* (*n* = 94) or *Xenopus tropicalis* (*n* = 9) proteins. Due to alternative

### Table 2 Individual measures of locomotor performance of the eight males *Xenopus allofraseri* (named sample A to H)

| Sample | Category       | Velocity (m.s\(^{-1}\)) | Acceleration (m.s\(^{-2}\)) | Time (s) | Distance (m) |
|--------|----------------|------------------------|--------------------------|---------|-------------|
| A      | Endurant       | 1.17                   | 54.41                    | 71      | 1.190       |
| B      | Burst-performant| 1.67                   | 47.75                    | 46      | 0.530       |
| C      | Burst-performant| 1.87                   | 61.83                    | 36      | 0.590       |
| D      | Endurant       | 1.10                   | 45.17                    | 55      | 0.840       |
| E      | Endurant       | 1.20                   | 43.80                    | 96      | 1.310       |
| F      | Burst-performant| 1.87                   | 49.69                    | 54      | 0.575       |
| G      | Endurant       | 1.56                   | 46.05                    | 73      | 1.040       |
| H      | Burst-performant| 1.44                   | 48.42                    | 32      | 0.560       |

*Fig. 1* Principal Component Analysis (PCA) and agglomerative hierarchical clustering of the four locomotor performance traits in eight males *Xenopus allofraseri* (named sample A to H): distance (total distance jumped until exhaustion), time (maximum time spent moving until exhaustion), acceleration (maximal instantaneous acceleration during an escape locomotor burst), velocity (maximal instantaneous speed during an escape locomotor burst).
splicing, some transcripts blasted to the same gene, therefore we identified 90 unique protein-coding genes. Using the human STRING database, we generated nine networks involving 46 differentially expressed protein-coding genes (Fig. 4).

We highlighted differentially expressed protein-coding genes involved in the structural organization and functioning of muscle cells, such as actin cytoskeleton and microtubule composition, conformation, mitochondrial activity, and cellular trafficking. Yet, it appears that many of those transcripts have regulatory properties or are effectors of downstream signaling cascades, starting from stimuli in the extracellular matrix and involving cell surface or transmembrane proteins. Consequently, endurant and burst-performant individuals differ in the functional pathways that are initiated by those up-stream effectors.

**Transmembrane proteins and focal adhesions**

Focal adhesion are macromolecular assemblies that play key roles in linking the extracellular matrix to the cytoskeleton [33] and act as important signal transducer [34]. In loading muscle, previous study highlighted the role of focal adhesion kinase (FAK, PTK2 gene) to act as a mediator, and transmit a stress and strain signal by integrins (transmembrane receptors) that activate multiple anti-apoptotic, cell growth pathways [35] and increase muscle mass [36]. Whereas FAK lead to slow twitch muscle generation and to an up-regulation of genes involved in mitochondrial metabolism [37], FAK-related non-kinase (FRNK) - a protein transcribed from the FAT portion of the FAK gene - acts to inhibit FAK in many cell types, including skeletal muscle [38]. In our study, we find *ptk2* and other protein-coding genes (*gca, lmo7*) involved in focal adhesion and the signal transduction cascade through the activation of Rho-GTPases (e.g. RhoG, rac1, cdc42) to be up-regulated in burst-performant individuals. Also, kinectin 1 (*ktn1*), a receptor for kinesin that accumulates in integrin-based adhesion complexes, is up-regulated in endurant individuals, whereas *mef2a*, a DNA-binding transcription factor of
ktn1, is up-regulated in burst-performant individuals. Remarkably, kinectin interacts with RhoG to activate rac1 and cdc42 through a microtubule-dependent pathway [39]. Indeed, kinesins are major microtubule motor proteins that have different functional properties depending on the ‘cargo’ (i.e. vesicle) they transport. We found several genes involved in microtubule composition and elongation, such as tubg1 and ckap5, to be up-regulated in burst-performant individuals. Those genes interact with a centromere protein R-like (an ortholog of ITGB3BP) that is up-regulated in endurant individuals (Fig. 4).

Furthermore, we found a differential expression of several protein-coding genes related to cellular trafficking and the Golgi apparatus. This central organelle system of the secretory pathway biosynthesizes proteoglycans [40]. It is also an important center for the formation of microtubules for its own functioning, also called ‘MTOC’ [41]. We found RhoGDI-3 (arhgdig) to be up-regulated in burst-performant individuals and it targets RhoG from the Golgi apparatus to be activated locally [42]. Interestingly, arafgap1, which codes for a GTPase-activating protein involved in membrane trafficking and vesicle transport from the Golgi complex, is up-regulated in endurant individuals. Yet, arafgap1 interacts with copz1 (Fig. 4), which codes for a coatamer (i.e., a protein complex that associates with Golgi coated vesicles and mediate transport from the endoplasmic reticulum). This protein-coding gene is up-regulated in burst-performant individuals, as well as rab12 and grasp, which both play a role in intracellular trafficking. We emit the hypothesis that endurant and burst-performant individuals differ in a range of downstream effectors, transcription regulators, molecules involved in cellular trafficking and microtubule activity in order to biosynthesize distinct extracellular matrix molecules and cell surface proteins, such as proteoglycans in the Golgi apparatus.

**Extracellular matrix and proteoglycans**

The extracellular matrix (ECM) is a primary macrostructure composed of several molecules such as collagen, hyaluronan, proteoglycans and glycoproteins that assemble into an organized meshwork [31, 43]. Proteoglycans for instance have diverse and essential roles in matrix remodeling and can act as receptors or co-receptors to affect signaling pathways but also to initiate and modulate signal transduction cascades independently of other receptors [44, 45]. In this study, we highlighted the up-regulation of genes coding members of two large groups of proteoglycans: neurocan (ncan), a chondroitin sulphate proteoglycan that is up-regulated in burst-performant
individuals, and glypican (gpc5), a heparan sulfate proteoglycan that is up-regulated in endurant individuals. In addition, we found an up-regulation of a cartilage oligomeric matrix protein-like (comp) in endurant individuals that has the molecular functions to bind calcium, heparin or proteoglycans. Therefore, it is plausible that endurant and burst-performant individuals differ in the proteoglycans and other extracellular proteins synthesized because their diversity and properties make them advantageous for powerful bursts of speed or long-duration exercise.

Chondroitin sulfates that partly compose aggrecan are able to absorb shocks by binding and releasing water content during compression in cartilageous tissues, tendons, or ligaments [46, 47] which can protect against injury during short and powerful physical performance. In addition, it has been shown that glypican-1 is able to enhance growth factor activity and is therefore used in therapeutic treatment to create new vasculature and restore blood flow in ischemic tissues [48]. Therefore, there could be a link between gpc5 and the positive relationship between endurance training and capillary densities [49], which may be beneficial for transporting oxygen to muscle [50]. Interestingly, we also found ppox, which codes for an essential component of hemoglobin and myoglobin, and spib, a hematopoietic transcription factor, to be up-regulated in endurant individuals. The coupling of increased blood oxygenation and muscle microvasculature is expected to render the aerobic pathway used during prolonged exercise more efficient. Finally, the study of Mao and colleagues [51] suggests that spib could be phosphorylated and activated by mitogen-activated protein kinase 8 (mapk8), which is also up-regulated in endurant individuals, and is part of a vast network comprising numerous genes involved in lipid metabolism, mitochondrial activity, and stress responses.

**Lipid metabolism, mitochondrial activity and stress response**

Almost all differentially expressed transcripts related to lipid metabolism, energy production, mitochondrial activity (mfn1, esrra, atp5b, dgat2, gls2, nfs1) are up-regulated in endurant individuals compared to burst-performant individuals. Yet, one protein-coding gene, a A-kinase anchor protein 1 (akap1), has 2 transcript isoforms, one being up-regulated in burst-performant and
one in endurant individuals. Those splice variants are proteins found in the mitochondria transmembrane, but at different position (position 7–26 and 42–61 in burst-performant and endurant individuals, respectively). The A-kinase anchor protein 1 binds to different regulatory subunits of protein kinase A (PKA) that has regulatory properties in lipid, sugar, and glycogen metabolism. Interestingly, we found an up-regulation of fsd2 in burst-performant individuals, which is an important paralog of CMYA5 that mediates subcellular compartmentation of protein kinase A and may attenuate the ability of calcineurin to induce a slow-fiber gene program in muscle [52]. Thus, we suggest that alternative splicing of akap1, in association with other mitochondrial or cytoplasmic genes, is a mechanism enabling the shift between different types of metabolism in endurant and burst-performant individuals.

Furthermore, our results are consistent with the fact that endurant individuals rely preferentially on lipid metabolism, because oxidative phosphorylation of fatty acids in muscle mitochondria produces a high yield of ATP, necessary for prolonged contraction of muscle fibers [53, 54]. On the contrary, individuals excelling at burst performance may rely mostly on anaerobic glycolysis in the cytosol (fast rate but low yield of ATP) [55]. In this context, we found diacylglycerol acyltransferase 2 (dgat2) to be up-regulated in endurant individuals. This endoplasmic reticulum enzyme catalyzes the final step in triglyceride synthesis and is part of the glycerolipid metabolism [56]. In addition, we found an up-regulation of atp5b, a mitochondrial ATP synthase subunit, by the estrogen-related receptor α (ERRα, coded by esrra) that regulates the transcription of metabolic genes and has a role in oxidative metabolism (Fig. 4) [57, 58]. ERRα has been found to be under control of myocyte enhancer factor 2 (MEF2) [59], a transcription factor that belongs to the MADS-box superfamily and that activates numerous muscle specific, growth factor-induced and stress-induced genes [60, 61]. Yet, we found a transcript that matches the mRNA of myocyte enhancer factor 2A L homoeolog of Xenopus laevis (mef2a) to be up-regulated in burst-performant individuals. This transcript has a non-synonymous mutation in the coding part of the MADS-box protein domain (Arg4Lys) which is responsible for DNA recognition and cofactor interaction. Therefore, it is not clear if the mef2a transcript of our study negatively regulates esrra (and also ktn1) expression or if it activates another gene that has yet to be identified. Intriguingly, we found an up-regulation of an inhibitor of cyclin-dependent kinase (CDK1 xic1) in endurant individuals, while cyclin-dependent kinase (CDK5) has been found to inhibit MEF2 [62].

Several studies have suggested a link between the MEF2 family of transcription factors and calcium-dependent signaling pathways [63, 64]. Calcium signaling is known to be essential for increasing endurance, oxidative capacity, and mitochondrial biogenesis [65, 66]. Likewise, we found an up-regulation in endurant individuals of the calcium/calmodulin-dependent protein kinase (CAMK) 2 A (camk2a) along with filamin B (flnb), an actin-binding protein (Fig. 4). Interestingly, CAMKs have also been found to activate mitogen-activated protein kinase (MAPK) which mediates early gene expression in response to various cell stimuli. Consistently, mapk8, which is up-regulated in endurant individuals, is known to positively regulate the expression of bnip3, an apoptosis-inducing protein located in the outer mitochondrial membrane [67]. On the contrary, bnip3 is negatively controlled by the translation initiation factor 5B (eif5b) [68], the latter having an increased expression in burst-performant individuals, along with the ribosomal protein S4 (rps4x) and the ribosomal protein S6 kinase a3 (rps6ka3). Noticeably, Clarke and colleagues [69] predicted the translation factor Eif6 to be a key regulator of energy metabolism, affecting mitochondrial respiration efficiency, reactive oxygen species (ROS) production, and exercise performance. Also, mapk8 and a transcription factor jun-D-like (jund) interact with ddit3 (Fig. 4) which encodes a member of the C/EBP family of transcription factors implicated in adipogenesis, erythropoiesis or promoting apoptosis, and which has two transcript isoforms up-regulated in endurant individuals and one transcript isoform up-regulated in burst-performant individuals.

We found a notable relationship between the calcium signaling pathway and stress-induced genes that are up-regulated in either endurant or burst-performant individuals. This is consistent with previous reports of a link between endoplasmic reticulum (ER) stress, unfolded protein response, and the contractile activity of muscle [70, 71] and suggests a need to further recycle damaged proteins and organelles that are used during muscle activity [72]. For instance, one of those actively used proteins during contraction and relaxation of the muscle is the calcium cycling protein parvalbumin that reduces the free calcium concentration in the sarcoendoplasmic reticulum and cytoplasm [73, 74]. In our study, ocm4.1, which codes for a protein that belongs to the parvalbumin family, is significantly up-regulated in burst performant frogs compared to endurant individuals. Similarly, the parvalbumin gene (pvallb) was found to be highly expressed in beltfish (Trichiurus lepturus), a fish species with high swimming activity [75] and particularly associated with fast contracting muscle fibers [76].
Rho-GTPases, ARP2/3 and WAVE complexes, and actin cytoskeleton

Many studies interested in physical performance in general and performance trade-offs in particular, have focused on the physiological aspects of muscle and particularly on fiber type differences [12, 77]. Whereas one major factor contributing to the differences in contractile properties between fiber types is the presence of different myosin heavy chain (MHC) isoforms [78, 79], we found the third most up-regulated transcripts in burst-performant individuals (LOC108701289) to be an ortholog of Xenopus tropicalis spec gene, which encodes a protein with similarity to members of the myosin light chain kinase family. In addition, we primarily found genes involved in the polymerization and depolymerization of the actin filament to be differentially expressed between endurant and burst-performant individuals. Indeed, the network of actin and actin binding proteins, along with the microtubules and intermediate filaments, constituting the actin cytoskeleton of skeletal striated muscle, is highly dynamic and allows crucial processes like cell migration and division, signal transduction, organelle transport and coordination of muscle contraction [33]. This dynamic system is made possible through the reversible polymerization of globular actin monomers (G-actin) into filaments (F-actin) [80].

Polymerization and depolymerization of the actin filaments is the culmination of a signaling cascade that begins with extracellular stimulation, adhesion interaction (ECM cell-cell interaction) or mechanical stress, which then acts upon guanine-nucleotide-exchange factors (GEFs) and GTPase-activating proteins (GAPs) to control the activation state of the small GTPases Rho, Rac, and Cdc42. After activation, the GTPases bind to a variety of effectors to stimulate downstream signaling pathways [81]. For instance, the non-receptor protein kinase 2 (ptk2, up-regulated in burst-performant individuals) can modulate the RhoA regulation pathway, but also activates MAP kinase signaling cascade and mediates activation of the Rho GTPase rac1. Rac1 activates the WAVE regulatory complex that drives ARP2/3 complex-mediated actin polymerization [82]. In our study, arpc4, an actin-binding component of the ARP2/3 complex is up-regulated in burst-performant individuals, whereas abi2, coding a component of the WAVE complex, is up-regulated in endurant individuals. Interestingly, we found two isoforms of the nif3l1 gene, one isoform being up-regulated in endurant individuals, the other in burst-performant individuals. This gene may function as a transcriptional corepressor and interacts with arpc4 and abi2 in the network involving also ptk2 and comp (Fig. 4).

Further examination would be necessary to test the hypothesis that the two isoforms of the nif3l1 transcript serve as a switch to activate or inhibit proteins of the WAVE and ARP2/3 complex, and thus actin polymerization.

Additionally, we found two up-regulated genes in endurant individuals that are responsible for the actin thin filament length. An F-actin uncapping protein, lrc16a, generates uncapped barbed ends that enhance actin polymerization, as well as a CapZ-interacting protein, rcsd1, which induces phosphorylation of CapZIP and regulates the ability of the F-actin-capping protein to remodel actin filament assembly. Antagonistically, the tropomodulin-4 gene (tmod4) is up-regulated in burst performant individuals and codes for a type of actin-capping protein that blocks the depolymerization of the actin filaments at the pointed end, thus contributing to the formation of short actin protofilaments. In one network, this gene interacts with a myosin-binding protein C fast-type (mybpcc2) for which we find two different up-regulated transcripts in burst-performant individuals (Figs. 3, 4). These two transcripts are globally dissimilar, which suggests the presence of two mybpcc2 homologs in the allotetraploid Xenopus allofraseri. Homoeologs are homologous genes in the same species that started diverging through speciation but were reunified in the same genome by allopolyploidization [83]. The up-regulation of mybpcc2 transcription in burst performant individuals, along with the over-expression of tmod4, may lead to the creation and renewal of a short fast-type actin-like filament, necessary for fast and powerful locomotion. Therefore, it seems that there is further depolymerization and polymerization of the actin filaments in respectively burst-performant and endurant individuals, and this may be linked to the need to rapidly recycle and rearrange or to stabilize the actin cytoskeleton.

Conclusions

Locomotor performance trade-offs have received considerable attention in the literature over the past three decades [11, 13, 84–87], yet the molecular origins of such trade-offs remain unclear. Because locomotor performance is heritable [14–16], substantial effort has been devoted to uncover its molecular basis. However, contradictory results regarding candidate genes suggest that locomotor performance is a complex polygenic trait and that gene expression regulation could be a non-negligible factor. Accordingly, our study reveals numerous transcription factors (DNA and RNA binding), ribosomal proteins, protein kinase, ubiquitin ligase, methyltransferase, and effectors of the signaling cascade that may help explain the trade-off between burst performance and endurance observed in male Xenopus allofraseri. Specifically, endurant individuals show an over-expression of protein-coding genes related directly or indirectly to lipid metabolism, mitochondrial activity, ATP
production and muscle oxygen supply. Moreover, endurance individuals appear to have increased actin polymerization, whereas burst-performant individuals have increased actin depolymerization along with an up-regulation of two fast-type myosin-binding protein C transcripts. Finally, burst-performant and endurance individuals show several differentially expressed transcripts coding for proteoglycans and extracellular matrix proteins, or proteins involved in intracellular trafficking, apoptosis and the ER stress response. Interestingly, several differentially expressed protein-coding genes are involved in both the calcium signaling and mitogen-activated pathways. How, and if, this relationship could explain the evolution of performance trade-offs remains unclear and would require further investigation.

Previous studies in humans have shown that differences in fiber type can be affected by a stop codon polymorphism (R577X) at actn3 [88, 89] and this same mutation can alter muscle function in mice [90]. Yet, there are contradictory results regarding the effects of actn3 polymorphism to explain performance trade-offs in humans [91]. In our study, none of the transcripts corresponding to actn3 or any member of the α-actin binding protein gene family were differentially expressed. The great variation observed within or between species for muscle fiber type composition can be attributed to the use of alternative splice forms by structural proteins [92]. Indeed, the regulation of, for instance, alternative splicing plays a major role in the production of functional complexity [93] and interestingly, a previous study detected an association between flight performance in dragonflies and alternative splicing in relation to muscle calcium sensitivity [94]. In our study, we found key genes (akap1 for metabolism, ddit3 for apoptosis, ranbp3 for nuclear export, nif3l1 as a transcriptional co-repressor) with transcript isoforms expressed in either burst-performant or endurance individuals. Although our results are biologically meaningful, we want to acknowledge that we were not able to validate these observations by quantitative methods such as RT-QPCR. In addition, RNAseq and qPCR results are known to be closely correlated [95] and the confirmation of these results would rather require testing additional samples. Finally, future studies dedicated to clarify the critical role of alternative splicing and its regulatory mechanism in explaining physical performance trade-offs would be insightful.

**Methods**

**Model species**

Eight *Xenopus allofraseri* males were caught in the wild (December 2009) in a single pond between Manengoteng (N 04.8090, E 09.8011) and Manjo (N 04.8435, E 09.8218) in Cameroon (Fig. 2). Animals were exported to France with authorization from the Cameroonian Ministry of Forestry and Wildlife (MINFOF) and were housed at the National Museum of Natural History in Paris, France. Frogs were placed in aquaria (60 × 30 × 30 cm) at 24°C and fed every week with beef heart, earthworms or mosquito larvae ad libitum. All individuals were pit tagged (Nonatec, Rodange, Luxembourg) for permanent identification.

**Physical performance**

Performance traits were measured for the eight males at a fixed temperature of 24°C. Maximal exertion capacity was measured by chasing each individual down a three-meter long circular terrestrial track until exhaustion, as indicated by the lack of a righting response. The floor of the track was covered with moistened cork to improve traction and prevent dehydration. For each individual, we recorded both the total distance covered and time spent moving until exhaustion. Burst performance capacity was quantified by measuring maximal instantaneous swimming speed and acceleration (see additional information about materials and measurement protocol in [8]). From the four physical performance traits, we performed a Principal Component Analysis (PCA) followed by an agglomerative hierarchical clustering using respectively the ade4 and FactoMineR packages of the R software.

**Illumina transcriptome sequencing and de novo assembly**

The individuals used for performance measurements were euthanized with a lethal injection of sodium pentobarbital (dosage of 150 mg/kg), a chemical compound acting quickly on the central nervous system, rendering the animal unconscious with little distress (a method validated by the European Commission (https://op.europa.eu/s/olzw)). The knee extensor muscles of the right leg were extracted for subsequent RNA sequencing. Tissues were extracted, transferred to labeled tubes containing RNA-later and conserved at −80°C until further processing. The protocol for RNA extraction using Trizol, RNA quantification and quality checking can be found in Dhorne-Pollet et al. [96]. PolyA-RNA was isolated and sequencing libraries prepared using ScriptSeq (Illumina). Pooled libraries were 100 paired-end sequenced using Illumina HiSeq 2500 located at the University of Exeter Sequencing Service facility.

Paired-end sequence reads were pooled together to generate a de novo transcriptome assembly. The raw sequence reads were trimmed and Illumina adapters emoved using Trimmomatic [97] with the following parameters: leading:5 trailing:5 slidingwindow:4:15 minlen:36. Transcriptome assembly was then performed de novo with the program Trinity [98]. We assessed the completeness of our transcriptome assembly by
searching for a tetrapod set of 3950 orthologs using BUSCO version 4.0.2 [99]. To obtain assembly quality statistics, paired-end reads were aligned back to the assembly with Bowtie2 [100]. A high-quality transcriptome assembly is expected to have strong representation of the reads input to the assembler and specifically for a trinity transcriptome assembly at least 80% of reads should mapped back to the assembly and exist as proper pairs. Transcript-level abundance was estimated using Kallisto [101], in addition to a normalized measure of transcript expression (TPM).

Phylogenetic analysis

To evaluate the phylogenetic position of the eight *Xenopus* males, we implemented a phylogenetic analysis with previously published mitogenomes of most known *Xenopus* species downloaded from GenBank [102] (Table S2). We retrieved reads mapping to the mitochondrial genome and performed a reference-based assembly of the complete mitochondrial genome for each sample using Geneious (www.geneious.com) and the *Xenopus allopriaderi* mtDNA sequence [98]. We used MAFFT v.7 to compute a multiple alignment of all *Xenopus* mtDNA [103]. From this multiple alignment, we discarded the control region and all positions with too many gaps or misalignments using Gblocks. The phylogeny was then constructed using PhyML v.3.0 [30] including the Shimodaira-Hasegawa (SH) statistic test. For visualization purposes, the phylogenetic tree was designed using Figtree v.1.4.4 [104] (Fig. 2).

Differential expression analysis

To detect differentially expressed transcripts, we ran edgeR [105], integrated in the R Bioconductor suite [106], using as input the TPM expression values that were cross-sample normalized using the Trimmed Mean of M-values (TMM) method. We chose edgeR rather than DESeq2 as it can detect differentially expressed transcripts between our two conditions even for transcripts between our two conditions even for transcripts with previously published mitogenomes of most known *Xenopus* species downloaded from GenBank [102] (Table S2). We retrieved reads mapping to the mitochondrial genome and performed a reference-based assembly of the complete mitochondrial genome for each sample using Geneious (www.geneious.com) and the *Xenopus allopriaderi* mtDNA sequence [98]. We used MAFFT v.7 to compute a multiple alignment of all *Xenopus* mtDNA [103]. From this multiple alignment, we discarded the control region and all positions with too many gaps or misalignments using Gblocks. The phylogeny was then constructed using PhyML v.3.0 [30] including the Shimodaira-Hasegawa (SH) statistic test. For visualization purposes, the phylogenetic tree was designed using Figtree v.1.4.4 [104] (Fig. 2).

Transcript annotation

We compared the significantly differentially expressed transcript sequences to the combined proteomes of *Xenopus laevis* and *Xenopus tropicalis* extracted from the Uniprot database using BLAST. We manually annotated the transcript sequences lacking detectable protein homologies by comparing them to the nucleotide (nr/nt) database using BLASTN with an e-value threshold of $1 \times 10^{-5}$. Differentially expressed transcripts were analyzed using STRING [107]. The STRING database allows the construction of protein-protein interaction networks, ranging from direct protein-protein interactions to indirect interactions (such as co-expression and text mining). The human database was selected for the network analysis as it contained substantially more information than the one for *Xenopus tropicalis* (the STRING database does not include data for *Xenopus laevis*). Gene symbols for each protein were used to find human orthologues and generate STRING networks using default settings. TSV files with interaction data were then exported and processed using Cytoscape [32] for visual inspection.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-021-07517-1.

Acknowledgements

We would like to thank Michel Baguette and Hervé Philippe for providing valuable resources and discussion. We would like to thank Eric Fokam and Legrand-Nono Gonzalo for their help in the field and without whom this study would not have been possible. We want to thank the Sequencing Service and Computational core facilities at the University of Exeter for producing quality RNA-seq data. We also want to thank the IUCN red list for making the spatial data publicly available.

Authors’ contributions

AJR, CB and AH conceived and designed the study; MV and AH collected the samples; AH and CB measured the locomotor performance; MV and TS performed the RNA extraction; KM and KP performed the RNA sequencing; AJR assembled the de novo transcriptome; NP performed the quality control of the assembled transcriptome, the phylogenetic analysis, and provided support with the analyses; VD conducted the differential expression analysis, annotation, protein network, statistical analysis, wrote and illustrated the manuscript; All authors read and approved the final manuscript.

Funding

Support for this work was provided by the Agence Nationale de la Recherche (ANR; France) MOBIGEN (ANR-09-PEXT-003). RNA-seq data was produced by the Sequencing Service and Computational core facilities at the...
University of Exeter with generous support from Medical Research Council Clinical Infrastructure award (MR/M008924/1), Wellcome Trust Institutional Strategic Support Fund (WT077835/MF), Wellcome Trust Multi User Equipment Award (WT101650MA) and BBSRC LOLA award (BB/K003240/1).

Availability of data and materials

The annotation of the 109 differentially expressed transcripts and the STRING TSV interaction file are available on Zenodo (https://doi.org/10.5281/zenodo.4028544). Processed and raw data files generated by this transcriptomic analysis have been deposited in NCBI’s Gene Expression Omnibus [108] and are accessible through GEO Series accession number GSE157915 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE157915). The geographic ranges of Xenopus species are publicly available at the IUCN red list website (https://www.iucnredlist.org/). Species name with their corresponding accession numbers of mitochondrial sequences downloaded from GenBank and used in the phylogenetic analysis are indicated in Table S2.

Declarations

Ethics approval and consent to participate

Animals were exported to France with the authorization for export from the Cameroonian Ministry of Forestry and Wildlife (MINFOR). All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. All experiments were approved by the Institutional ethics committee at the National Museum of Natural History in Paris (#68–25). In accordance to the directive 2010/63/EU of the European Parliament and French legislation, our study did not require specific authorization because our observations did not cause any pain, suffering, distress, or lasting harm.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

1 UMR 7179 MECADEV, CNRS/M.N.H.L., Département Adaptations du Vivant, 55 Rue Buffon, 75005 Paris, France. 2 Station d’Ecologie Expérimentale du CNRS, USR 2936, 09200 Moulis, France. 3 Functional Ecology Lab, Department of Biology, University of Ottawa, 30 Marie Curie, Ottawa, ON K1N 6N5, Canada. 4 Evolution, Génomes, Comportement & Écologie, Université Paris-Saclay, CNRS, IRD, 91198 Gif-sur-Yvette, France. 5 Exeter Sequencing Service, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QD, UK. 6 Evolutionary Morphology of Vertebrates, Ghent University, B-9000 Gent, Belgium.

Received: 17 September 2020 Accepted: 8 March 2021

References

1. Irshick DJ, Garland T Jr. Integrating function and ecology in studies of adaptation: investigations of locomotor capacity as a model system. Annu Rev Ecol Evol Syst. 2001;32(1):367–96. https://doi.org/10.1146/annurev. eco.32.110199.114048.

2. Le Galliard JF, Clokit J, Ferriere R. Physical performance and Darwinian fitness in lizards. Nature. 2004;425(7165):502–5. https://doi.org/10.1038/nature02057.

3. Husak JF, Fox SF, Lovern MB, Van den Bussche RA. Faster lizards sire more offspring: sexual selection on whole-animal performance. Evolution. 2006;60(10):2122–30. https://doi.org/10.1111/j.1022-7058.2006.tb01849.x.

4. van Damme R, Wilson RS, Vanhooydonck B, Aerts P. Performance constraints in decathletes. Nature. 2002;415(6873):795–6. https://doi.org/10.1038/415758a.

5. Stephens PR, Wiens JJ. Testing for evolutionary trade-offs in a phylogenetic context: ecological diversification and evolution of locomotor performance in emydid turtles. J Evol Biol. 2008;21(1):77–87. https://doi.org/10.1111/j.1420-9101.2007.01467.x.

6. Dlugosz EM, Chappell MA, McGillivray DG, Syme DA, Garland T Jr. Locomotor trade-offs in mice selectively bred for high voluntary wheel running. J Exp Biol. 2009;212(Pt 16):2612–8. https://doi.org/10.1242/jeb.029058.

7. Oufiero CE, Walsh MR, Reznick DN, Garland T Jr. Swimming performance trade-offs across a gradient in community composition in Trinidadian killifish (Rivulus hartii). Ecology. 2011;92(3):791–80.

8. Herrel A, Bonneaud C. Trade-offs between burst performance and maximal exertion capacity in a wild amphibian, Xenopus tropicalis. J Exp Biol. 2012;215(7):1306–11. https://doi.org/10.1242/jeb.072090.

9. Vanhooydonck B, James RS, Tallis J, Aerts P, Tadic Z, Tolley KA, Measey GJ, Herrel A. Is the whole more than the sum of its parts? Evolutionary tradeoffs between burst and sustained locomotion in lacertid lizards. Proc Biol Sci. 2014;281(1777):20132677.

10. Comm M. Physiological and biomechanical correlates of muscle function: effects of muscle structure and stretch-shortening cycle on force and speed. Exerc Sport Sci Rev. 1984;12(1):81–122.

11. Garland T Jr. Genetic basis of activity metabolism. I. Inheritance of speed, stamina, and antipredator displays in the garter snake Thamnophis sirtalis. Evolution. 1988;42(2):335–50. https://doi.org/10.1111/j.1558-5646.1988.tb04137.x.

12. Bonine KE, Gleeson TT, Garland T Jr. Comparative analysis of fiber-type composition in the illicibularus muscle of Phynosomatid lizards (Squamata). J Morphol. 2001;250(3):265–80. https://doi.org/10.1002/jmor.1069.

13. Wilson RS, James RS. Constraints on muscular performance: trade-offs between power output and fatigue resistance. Proc Biol Sci. 2004;271:1225–5.

14. Wilson JS, Gow JL, Taylor EB, Hendry AP. Quantifying the constraining influence of gene flow on adaptive divergence in the lake-stream threespine stickleback system. Evolution. 2007;61(8):2015–26. https://doi.org/10.1111/j.1558-5646.2007.00168.x.

15. Rico-Sanz J, Rankinen T, Joannisse DR, Leon AS, Skinner JS, Wilmore JH, Rao DC, Bouchard C, Study HF. Familial resemblance for muscle phenotypes in the HERITAGE family study. Med Sci Sports Exerc. 2007;39(8):1360–6. https://doi.org/10.1249/01.MSS.0000079031.22755.63.

16. Costa AM, Breitenfeld L, Silva AJ, Pereira A, Izquierdo M, Marques MC. Genetic inheritance effects on endurance and muscle strength. An update. Sports Med. 2012;42(4):449–458. DOI: https://doi.org/10.2165/11650560-0000000000000000.

17. Bray MS, Hagberg JM, Perusse L, Rankinen T, Roth SM, Wolfarth B, Bouchard C. The human gene map for performance and health-related fitness phenotypes. Med Sci Sports Exerc. 2009;41(1):35–73. https://doi.org/10.1249/MSS.0b013e3181844179.

18. Hagberg JM, Rankinen T,Loos RJ, Perusse L, Roth SM, Wolfarth B, Bouchard C. Advances in exercise, fitness, and performance genomics in 2010. Med Sci Sports Exerc. 2011;43(5):743–52. https://doi.org/10.1249/MSS.0b013e318255c221.

19. Sarzynski MA, Loos RJ, Lucia A, Perusse L, Roth SM, Wolfarth B, Rankinen T, Bouchard C. Advances in exercise, fitness, and performance genomics in 2015. Med Sci Sports Exerc. 2016;48(10):1906–16. https://doi.org/10.1249/MSS.0000000000000982.

20. Gazda MA, Andrade P, Afonse S, Dilyte J, Archer JP, Lopes RJ, Faria R, Cameiro M. Signatures of selection on standing genetic variation underlying athletic and navigational performance in racing pigeons. Mol Biol Evol. 2018;35(5):1176–89. https://doi.org/10.1093/molbev/msy030.

21. Kelly SA, Bell TA, Selfsry SR, Buus RJ, Hua K, Weinstock GM, Garland T Jr. Pardo-Manuel de Villena P, Pompe D: a novel intronic single nucleotide polymorphism in the myosin heavy polypeptide 4 gene is responsible for the mini-muscle phenotype characterized by major reduction in hind-limb muscle mass in mice. Genetics. 2013;195(6):1385–95. https://doi.org/10.1534/genetics.113.154476.

22. Ricard A, Robert C, Blouin G, Baste F, Torquet G, Morgenstier M, Savoie L, Mach N, Mata X, Schiller L, et al. Endurance exercise ability in the horse: a trait with a complex polygenic determination. Front Genet. 2017;8:899. https://doi.org/10.3389/fgen.2017.00899.

23. Kim J, Williams FJ, Dreger DL, Plassais J, Davis BW, Parker HG, Ostlund EA. Genetic selection of athletic success in sport-hunting dogs. Proc Natl Acad Sci U S A. 2018;115(30):E7212–21. https://doi.org/10.1073/pnas.1804551115.

24. Ehler T, Simon P, Moser DA. Epigenetics in sports. Sports Med. 2013;43(2):93–110. https://doi.org/10.1007/s40279-012-0012-y.

25. Liu Y, Li M, Ma J, Zhang J, Zhou C, Wang T, Gao X, Li X. Identification of differences in microRNA transcriptions between porcine oxidative and
v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607–13. https://doi.org/10.1093/nar/gky1131.

108. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res. 2012;41(D1):D991–5. https://doi.org/10.1093/nar/gks1193.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.