

Cooperation by non-kin during birth underpins sperm whale social complexity

Alaa Maalouf,^{1,2,3,*} Joseph DelPreto,^{1,2,*} Maxime Lucas,^{1,4,5,6,†} Simone Poetto,^{1,7,6,†} Jacob Andreas,^{1,2} Antonio Torralba,^{1,2} Shane Gero,^{1,8,9,‡} Giovanni Petri,^{1,10,6,‡} Daniela Rus,^{1,2,‡} and David F. Gruber^{1,11,‡}

¹*Project CETI, New York, NY, USA and Roseau, Dominica, W.I.*

²*CSAIL, Massachusetts Institute of Technology, Cambridge, MA, USA*

³*Department of Computer Science, University of Haifa, Haifa, Israel*

⁴*Department of Mathematics and Namur Institute for Complex Systems (naXys), Université de Namur, Namur, Belgium*

⁵*Mycology Laboratory, Earth and Life Institute, Université Catholique de Louvain, Louvain-la-Neuve, Belgium*

⁶*CENTAI Institute, Turin, Italy*

⁷*Center for Modern Interdisciplinary Technologies, Nicolaus Copernicus University, Toruń, Poland*

⁸*Department of Biology, Carleton University, Ottawa, ON, Canada*

⁹*The Dominica Sperm Whale Project, Roseau, Dominica, W.I.*

¹⁰*Network Science Institute, Northeastern University London, London, UK*

¹¹*Department of Natural Sciences, Baruch College and The Graduate Center, City University of New York, New York, NY, USA*

Cooperation poses an evolutionary challenge that has been central to the study of many large-brained mammals [1, 2]. Cooperative care and protection of offspring is often central to these questions [3, 4], yet the details and dynamics driving mammalian social structures are largely unknown, particularly for marine mammals. Sperm whales offer an excellent opportunity to probe these questions, since they have a modular social structure that is thought to have evolved due to collective care of offspring [5–10]. Despite extensive behavioral studies, observations of births are exceedingly rare [11, 12] and thus there is little direct quantitative evidence linking cooperative behavior to calf survival. Here, we show that the interactions during a sperm whale birth event in the Eastern Caribbean were marked by a collective response across kinship lines. Using high-resolution drone footage, a computer vision pipeline, new interactive segmentation tools, and multiscale network analyses, we quantify interactions among all members of the social unit actively participating in the birth, including both kin and non-kin. We show that birth attendance and assistance to the mother during labor was led by a related multiparous female. After delivery, all participating whales shifted their orientation towards the newborn and collectively engaged in lifting behavior, which is a key caregiving behavior required due to newborns being negatively buoyant [6, 13]. While the closest kin were responsible for the majority of lifting, all members of the unit contributed, taking turns and coordinating in smaller groups. Our findings provide evidence that collective cooperation in allocare behaviors drive the social bonds underpinning the complex, modular, social organization which has evolved in sperm whales.

I Introduction

The evolution of cooperative groups has been a focus of animal studies for decades [1, 2, 14–18]. Many toothed whales live in complex social networks, from fluid fission-fusion communities to modular multi-leveled social organizations [19–21]. Complex social interactions driven by socio-ecological principles underlie their societies. Yet, as a result of their marine habitat and their deep-diving lifestyles, it has historically been challenging to document the detailed behavioral interactions among individuals when cooperation occurs. As such, our understanding of the mechanisms that drive cooperation and group living in these species is limited [22].

Communal care and defense of calves are particularly central to cetacean societies for deep-diving, k -selected species [19], because every calf is a significant investment for female cetaceans and calf survival is central to the female’s fitness. From this perspective, analyses of singular events such as births can provide gateways to significant biological insights

* Core contributor; Machine learning and dataset curation lead, network analysis co-lead

† Core contributor; Network analysis lead, data curation co-lead

‡ Co-senior author

into the mutualistic benefits of group living in such contexts. Yet, despite centuries of study, direct observations of cetacean births and the behavioral intricacies involved in their social interactions remain largely undocumented in a rigorous, quantitative fashion. This leaves critical gaps in our understanding of group dynamics and specific behaviors during these crucial life events [11, 12, 23–25], and in our understanding of how social structures and coordination are shaped at larger scales.

Against this background, sperm whales (*Physeter macrocephalus*) offer a particularly compelling case study. These deep-diving cetaceans exhibit a modular, multi-level matrilineal society that is thought to have evolved to enhance collective defense and communal care of calves [5–7, 10, 26]. Within their stable, multi-matrilineal social units, individuals work together in highly coordinated ways, such as taking turns to socialize, forage, and provide care for the social unit’s young [5, 27, 28]. As a result, calves play a critical role as social hubs in the interactions among unit members [29]. Furthermore, evidence suggests that newborn sperm whales are born negatively buoyant [6, 13] and need to be supported as neonates in order to breathe; such early vulnerability could underscore the evolutionary importance of collective behavior for the survival of the unit’s calves [7, 30]. Birth events thus have the potential to provide biological insights into the emergence of cooperation in sperm whale society. Unfortunately, due to the inherent difficulties of observing births at sea and to sperm whales’ lengthy gestation [25], there have only been four reported observations in the last 60 years, which are either largely qualitative or connected to the whaling industry [11, 31–33].

Here, we present the first high-quality observation of the birth of a sperm whale, and provide quantitative evidence of cooperative behavior among immediate kin and non-kin during the event. We recorded a unique dataset of high-resolution aerial drone videos, which allows for detailed analysis of the social interactions and coordination of sperm whales during one of the most critical moments of their life cycle. We adapt advanced machine learning and computer vision techniques to automatically detect and track the movements and orientations of whales in the videos. A custom segmentation and annotation tool then enables curating the results and augmenting them with manually annotated whale identities and historical kinship data.

We then employ network theory tools to quantify whale interactions and coordination across multiple scales: the family unit, the matriline within, and roles of individuals. This quantitative investigation provides new insights that substantially exceed what is obtainable from previous observational reports. At the unit level, by tracking the unit’s network cohesion and spatial distribution, we show that the sperm whales organized throughout the event to first provide constant support to the mother and then to the newborn. At the matriline level, we find that familial ties influenced proximity to the newborn and the roles that individual whales assumed in collective post-birth care. Despite this, sperm whales freely mixed across kinship lines in the interest of post-birth care; this is in stark contrast with previous observations of strict kin segregation during ordinary foraging behavior. Finally, we highlight the specific roles of individual key whales in relation to the newborn and its mother, pointing to the emergence of socio-behavioral profiles driven by individual experiences. Our findings shed light on long-standing evolutionary questions about cooperation, caregiving, and social organization in one of the ocean’s most enigmatic species.

II Results

II.A The birth of a sperm whale calf

On the morning of 8 July 2023, off the west coast of Dominica, a team aboard a Project CETI research vessel encountered a unit of sperm whales gathered at the surface for an extended time. This unit, “Unit A” [30], is composed of two matriline. Despite sharing a mitochondrial DNA haplotype, these two matriline are not closely related [26, 34], and indeed are rarely observed associating physically at the surface when foraging [8, 10, 26, 30]. This is confirmed by the kinship network and the association network being virtually identical (Pearson correlation between corresponding adjacency matrices $r = 0.88$, $p \ll 10^{-5}$, Fig. 1A), amounting to a social divide between the two matriline within the unit.

Motivated by the anomalous sighting of the two matriline together, two CETI aerial drones were launched to record footage of the interactions among the whales (drone model DJI Mavic 3 recording at a resolution of $3,840 \times 2,160$ pixels). Fig. 1B highlights representative images captured throughout the event. The 11 sperm whales of Unit A were first observed in a tight formation, vocalizing and making shallow dives before becoming more active and heading south at 10:48 local time. At 11:12, we started observing the birth of a sperm whale calf. The process began with the calf’s flukes emerging from the mother (*Rounder*), while other females oriented themselves towards the birth, diving and positioning closely around the mother. Between 11:42 and 11:45, brown-red plumes of blood and feces appeared, and at 11:45:45 the birth was completed. The calf surfaced beside its mother shortly after, marking the end of a 34-minute

delivery process. The unit’s behavior changed dramatically post-birth, becoming highly active and cohesive; adults and juveniles physically engaged with the newborn, lifting it, pushing it, and keeping it near the surface. This intense social behavior continued as the unit approached within a meter of the passive research vessel, and persisted as they encountered short-finned pilot whales and Fraser’s dolphins. Despite no physical contact with the dolphins, multiple close interactions with pilot whales were recorded. By 12:30, the cohesive unit began to break into smaller clusters, and by 13:39, the whales resumed foraging. Unit A was not seen again until over a year later, when the calf was observed off Dominica alongside its siblings (*Accra* and *Aurora*); reaching over one year old is a milestone linked to increased survival [35]. Additional descriptive details of the observed event can be found in [12].

While the entire dataset spans over six hours, here we focus on those relevant to our understanding of the dynamics of the social interactions and coordination before, during, and after the birth.

II.B Data pipeline: dataset curation, computer vision, and network analysis

To quantify the unit’s interactions during this event, we developed a data processing and analysis pipeline able to extract whale positions and behaviors from the drone footage. This pipeline is summarized in Fig. 1C. The first stage of the pipeline uses computer vision techniques to segment individual whales in each video, track whales across frames, and determine the orientation of each whale. Challenges included interference due to occlusions caused by whale interactions, sun glare, waves, and whales that frequently dived and then resurfaced elsewhere. We addressed these challenges by leveraging state-of-the-art machine learning methods and expert human annotations. We developed pipelines for segmenting whales in individual frames and for associating whale masks across frames, and we provided manual annotations to improve the robustness of these models particularly for challenging portions of the footage. In parallel, we developed zero-shot deep learning and heuristic methods to estimate whale orientations. See Section M.1 for additional details.

The next stage of the pipeline focuses on verifying these computer vision results and linking them with longitudinal genetic and behavioral data. To this aim, we developed an interactive online segmentation and annotation tool named *Whale Tales* to adjust and annotate model outputs in real time. This tool eases the process of manually correcting segmentations, inter-frame tracking, and whale orientations to ensure accurate results. In addition, *Whale Tales* enables annotations that augment the machine learning outputs with biological data; in particular, it was used by an experienced cetologist (SG) to identify whales based on features such as the trailing edge of flukes, with additional aid of body marks and scars. See Section M.2 for additional details.

This pipeline combined automation and human expertise to yield a comprehensive dataset featuring (1) whale positions and orientations with fine temporal and spatial resolution (including per-frame whale masks, orientations, centroids, and bounding boxes), and (2) whale identities linked to kinship information.

Throughout the following discussions of analyses performed on this dataset, the event is divided into four main phases: *pre-birth*, *during-birth*, *post-birth*, and *end-of-event*. *Pre-birth* and *during-birth* are separated by when the newborn flukes emerged. *During-birth* and *post-birth* are separated by when delivery was completed. *End-of-event* constitutes one video selected approximately 2.3 hours later during a return to foraging, to analyze a return to steady-state behavior. Additional details about the data timelines and segmentations are provided in Fig. M1.

II.C Collective group dynamics

To analyze the processed dataset, we first focused on unit-level dynamics. In contrast to typical foraging behavior, in which sperm whales are seen at the surface in clusters of 1-3 whales [30], we observed all 11 individuals during the birth event, corresponding to the entire Unit A. In particular, sperm whales qualitatively displayed both tight formations with many animals in physical contact, as well as active surface social behavior [36].

To quantify multiple aspects of the unit’s behaviors with fine temporal resolution, we used three metrics. Two capture the compactness of the unit in space, and a third describes unit members’ orientations: (i) *spatial cohesion* (Fig. 2A-B), (ii) the *connected group size* (Fig. 2C-D), and (iii) *orientation synchrony* (Fig. 2E-F). These metrics are summarized below, and additional details are provided in Section M.3.2.

Spatial cohesion measures how tightly grouped the whales are over time. For each frame, we compute the convex hull encompassing all whale segmentation masks, and then compute the ratio of the area occupied by whale masks to that of the hull. This ratio tends towards 1 for very tight formations with little empty space between whales, and towards 0 for progressively sparser formations. Results are illustrated in Fig. 2A. We find that the spatial cohesion increases leading

up to the birth, peaks immediately following the birth, plateaus in a cohesive group after the birth, and then decreases as the unit begins to disperse and resume foraging. As seen in Fig. 2B and Fig. S1, results indicate significantly different distributions of cohesion between the phases of the event (Kruskal-Wallis, $p < 10^{-4}$, and Mann-Whitney U, $p < 10^{-4}$, for all pairs of distributions). In particular, cohesion is much larger in the *post-birth* phase, as indicated by large effect sizes (Cliff’s $|\delta| > 0.80$ for each pairwise test).

The *connected group size* measures the number of whales in the largest sub-network of whales that are in physical contact. For each frame, we build the network of whale physical contacts based on distances between segmentation masks, find the largest connected component, and normalize its size by the total number of whales present in that frame. As seen in Fig. 2C, this normalized connected group size evolves similarly to the spatial cohesion: it increases leading up to the birth, stays at its highest after the newborn is delivered, then drops hours later as the whales disperse. As seen in Fig. 2D and Fig. S1, the distributions are significantly different across the four phases of the event. In particular, the normalized connected group size is much larger in the *post-birth* phase, as indicated by large effect sizes (Cliff’s $|\delta| \geq 79$ for each pairwise test), consistent with the cohesion metric.

Finally, we quantified the *orientation synchronization* among the whales. In particular, we compare the directions in which whales are facing using the standard Kuramoto order parameter for oscillators [37], which ranges from 1 when whales are all pointed in the same direction, to close to 0 when their orientations are isotropically distributed. As seen in Fig. 2E, the unit maintained high synchrony throughout the whole event and synchronization reached its highest value during the birth itself, suggesting a highly coherent formation to assist the mother during delivery. As shown in Fig. 2E and Fig. S1, the reported distributions are significantly different across phases. Pairwise effect sizes indicate that the effect is small (Cliff’s $|\delta| \leq 30$ for all pairs).

These findings suggest a crucial role of spatially coordinated physical interactions in supporting the newborn and in maintaining the social cohesion of the unit during the event. They are consistent with previous observations [26, 30], underscore the highly social nature of the event, and further support the conclusion that sperm whales come together to provide assistance during birth and postnatal care.

II.D Dynamics within the unit

To build on the results of unit-level dynamics, we developed metrics to shed light on how support for the newborn is explicitly realized within the unit itself. In particular, we focus on two measures of social organization: (i) kinship homophily, which reflects how much the two distinct matriline mix within the unit, and (ii) collective attention on specific individuals.

II.D.1 Matriline mixing and kinship homophily

We define *kinship homophily* as the tendency of individuals to be in close proximity to individuals within their own matriline rather than to those outside of their matriline, and measure it as the network assortativity coefficient [38, 39]; see Section M.3.3 for details. Homophily ranges from 1, when animals are exactly segregated in their matriline, to 0, when individuals are fully mixed regardless of kinship.

As shown in Fig. 3A, we find that homophily is near 0 throughout all active portions of the birth event, implying an extreme level of mixing between the two matriline. Importantly, as the unit begins to disperse at the end of the event, homophily returns to levels observed during typical foraging behavior (horizontal gray dashed line). This corresponds to the re-establishment of the typical matriline segregation as illustrated in the rightmost representative inset of Fig. 3A. Statistical tests applied to associated distributions of homophily confirm significant differences between the *end-of-event* phase and the others (see Fig. 3B, Fig. S2, and Section M.3.3). In particular, homophily is much larger in the *post-birth* phase, as indicated by large effect sizes (Cliff’s $|\delta| = 1$ for each associated pairwise test).

These findings suggest a flexibility of sperm whale social interactions during this critical event, in which active mixing between matriline is adopted to facilitate coordinated support for the newborn.

II.D.2 Shifts in collective attention

In addition to analyzing mixing across matriline, we investigated how patterns of *attention* within the unit changed throughout the event. We quantify these dynamics at a fine-grained level using a directionality metric based on the

cosine of the angle between an observer’s orientation vector and the vector pointing from the observer to a target’s spatial position (see Fig. 3D). This attention metric is 1 when the observer is oriented perfectly towards the target, -1 when the attending whale is oriented away from the target, and 0 when they are perpendicularly oriented (additional details in Section M.3.3).

As shown in Fig. 3C, the unit focused strongly on the mother until the delivery was completed, at which point the collective attention markedly shifted to the newborn. This highlights the unit’s prioritization of the newborn’s immediate safety. The heightened attention to the newborn persisted in the *post-birth* phase.

Interestingly, this analysis also yielded predominantly negative attention values for the adolescent male present (*Allan*). This quantitatively supports the observation that he was consistently kept to the outside of the unit with the other individuals turned away from him. Indeed, the adolescent male was rapidly turned away despite his repeated attempts to interact with the rest of the unit.

These results suggest how collective attention is dynamically allocated within the unit to reflect the changing needs of its vulnerable members throughout the event. They also showcase the remarkable adaptability of sperm whale social behaviors, driven by a balance of kinship, cooperation, and attention to individual needs.

II.E Individual roles in cooperative birth dynamics

To further investigate social dynamics after the delivery, we analyzed the fine-scale interactions and roles among individual whales. In particular, we focused on interactions with the newborn and defined metrics that capture proximity, physical contacts, and different types of support coordination. These metrics and results are summarized below, and additional details can be found in Section M.3.4.

II.E.1 Collective and continuous newborn lifting by kin and non-kin

The first key observation is that, in the aftermath of the delivery, multiple whales were typically in contact with the newborn to lift it out of the water. To quantify this, we measured the number of whales that touch the newborn in each frame, where touching is defined as their segmentation masks intersecting. As seen in Fig. 4A,B, we found that (i) two or more whales were typically in contact with the newborn, (ii) only less than 3% of the time contained no whales touching the newborn, and (iii) there were no prolonged gaps without contact.

The high frequency of multiple whales touching the newborn, combined with the rarity and brevity of moments without it, reflects continuous, collective care and high attentiveness from the unit. This suggests that cooperation among a larger number of whales may be essential for the survival of newborns during the first hours of life. Adults consistently lifted the newborn out of the water, reinforcing the importance of this behavior and supporting the hypothesis that newborns are negatively buoyant at birth [6, 13].

II.E.2 Turn-taking of newborn support by all unit members

We introduce a *support rank* metric, which captures how prominently each individual is supporting the newborn in a given frame. It first considers whales whose masks overlap with the newborn, and sorts them based on the extent of the overlap. Next, whales whose masks do not overlap with the newborn are sorted by their distance to the newborn. Fig. 4C demonstrates that every unit member served as the primary supporter for at least 1% (7 seconds) of the time when the newborn was identified (followed by 4% or 22 seconds). This highlights the significant role each individual played in supporting the newborn during the event, creating a robust and distributed support network where all unit members contribute to the care of the newborn.

Similarly, as presented along the diagonal of the matrix in Fig. 4D, every whale touched the newborn for at least 5% of the total time in which the newborn was identified. This is further expanded in Fig. S8, which shows that every whale touched the newborn for at least 8.9% of the time during which both the whale and the newborn were simultaneously identified – representing over 4% (32 seconds) of the total analyzed period.

II.E.3 Coordinated groups supporting the newborn

To provide further insight into the continuity and distribution of support, we characterized the coordination in lifting by measuring the percent of time that individual whales, pairs of whales, or groups of whales touched the newborn. As seen in Fig. 4D and Fig. S11, a *core* group of whales emerged that spend a large fraction of time supporting the newborn, either individually or together; this group includes the mother (*Rounder*), her half-sister (*Aurora*), a non-kin juvenile female (*Ariel*), and a grandmother (*Atwood*). The remaining whales, mostly non-kin individuals, spend a smaller fraction of time supporting the newborn and mostly do so with a member of the core group.

Analysis of the core support group reveals that the mother and her half-sister maintained contact with the newborn for 86% of the time the newborn was identified. Including the non-kin juvenile female increased this coverage to 93%, suggesting that meaningful contributions also came from outside the kin network. Finally, adding the fourth member of the core group increases coverage to 96%. Interestingly, adding a non-kin individual (*Snow*) further increased the coverage to 98%, despite this whale touching the newborn only 8% of the time; this means that 2% out of these 8% (25% of this whale’s touching duration) occurred precisely when no core group member was in contact – this suggests a critical complementary role. Beyond this point, adding the remaining whales had minimal impact on the coverage, indicating that most of the continuous support was provided by this core set of individuals.

Furthermore, Fig. 4D confirms a degree of turn-taking within this core group of individuals. The mother’s half-sister was found to be touching the newborn for 60% of the time during which the mother was not touching the newborn. During the times when neither whale from this pair was touching the newborn (approximately 14% of the time the newborn is identified), the juvenile female (*Ariel*) touched the newborn 50% of that time; adding the final whale in the core group (an elder female in the mother’s matriline), they together touched the newborn 72% of that time.

Such patterns are also evident when considering the distance to the newborn rather than only contact with the newborn (Fig. 4E). Interestingly, when none of the core group are with the newborn, younger females from the other matriline provide support (specifically, *Snow* and *Soursop*).

Crucially, we note that while individuals outside the core group spend less time lifting the newborn, they do spend a significant fraction of the time in this behavior (approximately 30% by a single pair in some cases); this highlights once again the cross-kin nature of the observed coordination patterns.

II.E.4 The central role of immediate kin and juvenile females

Focusing on the core support group points to a pivotal role played by immediate kin and by juvenile, nulliparous females who exhibited higher levels of interaction with the newborn. These results can be seen in Fig. 4C,D,E and in Fig. S8.

Roles of immediate kin: The mother (*Rounder*) and her half-sister (*Aurora*) formed a core support group, spending the most time in physical proximity to the newborn and providing the majority of lifting assistance. Within frames where the newborn was identified, the mother was the primary supporter approximately 31% of the time (Fig. 4C) and touched the newborn approximately 65% of the time (Fig. 4D); this corresponds to approximately 73.9% of the analyzed period where the newborn and the mother were both identified (Fig. S8). She was never observed to be greater than 1.5 newborn body lengths away from the newborn (Fig. 4E). Meanwhile, her sister *Aurora* touched the newborn for 42% of the same period. Considering *Rounder* and *Aurora* together, Fig. 4C and Fig. 4D show that at least one of them was the primary supporter for approximately 50% of the time and touched the newborn for 86% of the time. Similarly, *Aurora* spent most of the time in very close proximity to the newborn.

Learning-to-Mother: A juvenile female *Ariel* (born 2018), was among the most engaged non-kin supporters; her support rank varied between the first and fifth positions. *Ariel* frequently made physical contact with the newborn and exhibited behaviors that might best be interpreted under the “learning-to-mother” hypothesis [40] in which young nulliparous females learn caregiving skills through active interaction with newborns. Together with *Aurora* (born 2016) and *Rounder*, at least one of these three whales touched the newborn during 93% of the time. This finding underscores dual roles that events such as the birth can serve: ensuring immediate survival and facilitating social learning for younger members.

II.F Coordinated touching of the newborn

Beyond individual behaviors, we analyzed coordinated interactions with the newborn among pairs and triplets of whales. To quantify this, we defined a coordination metric that compares the observed probability of whales touching the newborn simultaneously to the probability expected if each whale acted independently (see Section M.3.4 for additional details).

The results for pairs, as shown in Fig. 4F, reveal that many whale pairs were more likely to touch the newborn together than would be expected by chance (i.e., with a metric greater than 1) and thus suggest possible coordination. Interestingly, most of these pairs do not involve close relatives of the newborn and were not part of the core support group, suggesting that coordinated caregiving extended beyond kinship ties.

Note that the mother (*Rounder*) and her half-sister (*Aurora*), while central in previous analyses, appear less dominant here due to their high baseline probabilities of independently touching the newborn; this high baseline makes coordinated touching less distinguishable under this metric.

In addition, we also extended the metric formulation to triplets of whales. We again consider the observed likelihood of all whales in the group touching the newborn together as opposed to an assumption of them touching the newborn independently. In this case, three notable groups emerged: (1) *Rounder*, *Ariel*, and *Accra*, (2) *Aurora*, *Ariel*, and *Soursop*, and (3) *Aurora*, *Ariel*, and the *Young-of-Year*. Interestingly, the juvenile female *Ariel* appears in all of these groups, suggesting a central role potentially due to her age and the general centrality of calves.

To further support this observation, we analyzed the percent of time pairs of whales were simultaneously touching the newborn versus when only one was touching. We also analyzed patterns of spatial proximity to the newborn, revealing additional coordination around the newborn and offering deeper insight into the support structure and turn-taking dynamics. See Section S3.2 for details.

III Discussion

This study provides the first quantitative analysis of a cetacean birth, offering unprecedented insights into the social and evolutionary dynamics surrounding parturition in one of the ocean’s most socially complex mammals. Through the integration of aerial footage, machine learning, and long-term behavioral and kinship data, we document cooperative behaviors during the birth of a sperm whale calf. It has been hypothesized that group living in this species was driven primarily by the need for communal defense and care for calves [5–7, 10, 29, 41, 42], and this study provides quantitative evidence that adult females collectively cooperated in direct allocare behaviors to ensure a newborn’s survival, and that these drive the social bonds underpinning the complex, modular, social organization which has evolved in sperm whales. These findings challenge prevailing assumptions about cetacean sociality and place sperm whale birth behavior in comparative context with terrestrial mammals, including primates and humans.

Birth attendance as a derived social trait

The behaviors observed during the period leading up to and including parturition reveal a striking pattern of social coordination. Analysis results indicate synchronized orientations, larger groups than those seen during typical foraging behaviors, and pronounced attention towards the parturient. Several females positioned themselves in close proximity to the parturient, orienting their bodies toward her and often engaging in synchronous movement. Multiparous individuals, including the grandmother of the newborn, performed rolls across the mother’s abdomen, suggesting a deliberate involvement in the birthing process [12]. These observations provide the first empirical evidence of birth attendance or assistance in a non-primate species – a behavior long considered characteristic of humans and their close relatives [43, 44]. The potential convergence between human midwifery and cetacean birth assistance raises compelling questions about the evolutionary pressures and cognitive architectures that support such behaviors.

Unit-level coordination and cooperation

At the unit level, results indicate that the birth event was marked by a dramatic shift in spatial cohesion, physical contact, and orientation synchrony. As the birth progressed, the unit compacted into a tight formation, maximizing contact and presumably enabling rapid support of the newborn. These dynamics are distinct from typical daily

association and interactions for these sperm whales [6, 30, 42], and suggest the presence of an emergent collective behavior specific to birth events.

Kinship and cross-kin cooperation

Despite the strongly matrilineal social structure of sperm whales, caregiving behaviors during the birth event transcended kin lines. Individuals from both matriline not only remained in close proximity but also actively engaged in supporting the newborn. Analyses reveal that each adult in the unit served as the primary physical supporter of the newborn at least once, suggest coordinated newborn support, highlight a core support group with individuals spanning matriline, and reveal mixing across matriline as measured by kinship homophily. This inter-matrilineal cooperation contrasts sharply with the usual pattern of kin-preferential associations observed in this unit and across this community [7, 26, 27], and suggests that birth events elicit a form of transient social integration potentially grounded in mutualistic or reciprocal dynamics.

Individual roles and caregiving profiles

Beneath the collective coordination, distinct individual-level caregiving profiles emerged. A core group of three whales – the mother, her half-sister, and a caregiving-experienced female – accounted for the majority of physical contact and proximity to the newborn. These individuals appeared to engage in coordinated turn-taking, maintaining nearly continuous support coverage. Simultaneously, juvenile females, lacking reproductive experience, displayed behaviors suggestive of exploratory caregiving, remaining close to the newborn and central to the unit. These patterns may reflect a caregiving gradient, wherein younger individuals gain experience through “learning to mother” phenomena documented in other mammals [40, 43].

Roles of elders and the adolescent male

The most senior females in the unit – the grandmothers – were less involved in direct care than expected. Although present during the delivery and briefly involved, they often shifted to peripheral positions or attended to other calves in the unit. This may reflect a strategic reallocation of roles, in which elder individuals facilitate learning opportunities for younger females while providing broader unit-level surveillance or coordination [7, 29]. This aligns with the “grandmother hypothesis” [45], which suggests that elder females enhance offspring survival indirectly through social scaffolding and knowledge transfer.

The role of the adolescent male (*Allan*) adds further complexity. Though socially peripheral and seldom in contact with the newborn, he remained with the unit throughout much of the birth and postnatal period. Whether this reflects protective behavior or residual attachment to his natal unit remains unclear. His surprising presence and ambiguous position offer a window into the transitional dynamics of juvenile males who are ostracized from their highly social natal units and face a dramatic change in their place in sperm whale societies [29].

Cognitive and evolutionary implications

The behaviors documented here suggest a level of social cognition and coordination across long time periods that is rare among non-human animals. Delayed reciprocation [15, 46, 47], across years between births, of prosocial behaviors relating to parturient assistance and neonatal survival across their social networks could underpin the evolution of their modular society. This interpretation aligns with existing neuroanatomical evidence; sperm whales possess large gyrified neocortices rich in spindle neurons, and extensive limbic systems, supporting the necessary capacity and advanced social cognition to track relationships, monitor reciprocal interactions, and manage cooperative behaviors over extensive periods [48–53]. Vocal communication during the event, e.g., by shared identity coda production [54, 55], vocal style learning [56], overlapping and matching of coda types [57], or matching exchange-level rubato [58], may serve a role analogous to vocal grooming in primates, reinforcing social bonds and enabling coordination [59, 60]; this might additionally carry a function to further promote expected partners’ future, prosocial, reciprocal behavior [61, 62]. The possibility that complex, multi-modal social cognition has evolved independently in large-brained marine mammals underscores the broader evolutionary significance of our findings.

Methodological innovations and future directions

This study presents biological insights enabled by aerial drone recordings, state-of-the-art machine learning pipelines, interactive annotation tools, and kinship-informed network analysis techniques. The tools developed for this study were the result of close multidisciplinary collaborations among computer scientists, network scientists, and biologists. This pipeline allowed for high-resolution, quantitative tracking of behavioral and social dynamics that would otherwise remain inaccessible in marine environments. These building blocks can help lay the foundation for providing researchers with platforms and tools to better document, annotate, quantify, and understand animal behaviors.

While this study documents a single birth event, the level of detail supports the generation of hypotheses about the structure and function of social care during birth. Future work should aim to collect additional observations across diverse social units and species, with multimodal datasets incorporating acoustic, visual, and physiological channels [63]. Comparative analysis across taxa – particularly with species exhibiting extended parental care and social learning – may help illuminate the convergent and divergent paths through which social birth assistance evolves.

Data and code availability

The code for the machine learning and analysis pipelines is available on Project CETI’s GitHub repository at <https://github.com/Project-CETI/whale-birth-data-and-analysis-suite>.

The videos, segmentations, and annotations used for this study can be explored and visualized on a read-only version of *Whale Tales* accessible at <https://whale-tales.csail.mit.edu>. Reviewers can access the platform anonymously by using the username `anonymous0` and password `SpermWhaleBirth123`. If it says that the account is currently in use, increment the username number and try again with the same password; this can continue up to the username `anonymous9`.

We will make the dataset publicly available on an archival platform upon acceptance/publication of the manuscript.

Acknowledgments

This study was funded by Project CETI via grants from Dalio Philanthropies and Ocean X; Sea Grape Foundation; Virgin Unite, and Rosamund Zander/Hansjorg Wyss through The Audacious Project: a collaborative funding initiative housed at TED.

Association and kinship data was derived from longitudinal work of The Dominica Sperm Whale Project which was funded by a FNU fellowship for the Danish Council for Independent Research supplemented by a Sapere Aude Research Talent Award, a Carlsberg Foundation expedition grant, a grant from Focused on Nature, two Explorer Grants from the National Geographic Society, and supplementary grants from the Arizona Center for Nature Conservation, Quarters For Conservation to S.G., with supplemental funding from the Dansk Akustisks Selskab, Oticon Foundation, and the Dansk Tennis Fond to Pernille Tønnesen. Further funding was provided by Discovery and Equipment grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to Hal Whitehead of Dalhousie University and a FNU large frame grant and a Villum Foundation Grant to Peter Madsen of the Marine Bioacoustics Lab at Aarhus University.

M.L. is a Postdoctoral Researcher of the Fonds de la Recherche Scientifique-FNRS.

Project CETI’s research is conducted under scientific research permit LS 27-200-21 to Project CETI Dominica granted by Dominica Fisheries Division of the Ministry of Agriculture, Fisheries, Blue and Green Economy. Drone research was approved under Harvard University’s IACUC Protocol ID 21-02-379-1. Regarding the long-term whale datasets, the field protocols for approaching, photographing, tagging, and recording sperm whales were approved by either the University Committee on Laboratory Animals of Dalhousie University, Canada; the Animal Welfare and Ethics Committee of the University of St Andrews, Scotland or Aarhus University, Denmark; or the Institutional Animal Care and Use Committees at Harvard University; and sometimes several or all of these across years.

We thank Lakshyana KC for integrating and running the segmentation and tracking code on a cluster, and for her valuable discussions. We also thank Krishna Murthy Jatavallabhula for his early-stage insights and constructive brainstorming sessions. Finally, we thank all members of the CETI team for their helpful comments.

-
- [1] T. Clutton-Brock, *Social evolution in mammals*, Science **373**, eabc9699 (2021).
 - [2] R. Axelrod and W. D. Hamilton, *The evolution of cooperation*, science **211**, 1390–1396 (1981).
 - [3] M. L. Riedman, *The evolution of alloparental care and adoption in mammals and birds*, The Quarterly review of biology **57**, 405–435 (1982).
 - [4] M. D. Jennions and D. W. Macdonald, *Cooperative breeding in mammals*, Trends in Ecology & Evolution **9**, 89–93 (1994).
 - [5] H. Whitehead, *Babysitting, dive synchrony, and indications of alloparental care in sperm whales*, Behavioral Ecology and Sociobiology **38**, 237–244 (1996).
 - [6] H. Whitehead, *Sperm whales: social evolution in the ocean* (University of Chicago press, Chicago, Illinois 60637, USA, 2003).
 - [7] S. Gero, D. Engelhaupt, L. Rendell, and H. Whitehead, *Who cares? between-group variation in alloparental caregiving in sperm whales*, Behavioral Ecology **20**, 838–843 (2009).
 - [8] S. Gero, J. Gordon, and H. Whitehead, *Individualized social preferences and long-term social fidelity between social units of sperm whales*, Animal Behaviour **102**, 15–23 (2015).
 - [9] S. Gero, A. Böttcher, H. Whitehead, and P. T. Madsen, *Socially segregated, sympatric sperm whale clans in the atlantic ocean*, Royal Society Open Science **3**, 160061 (2016).

- [10] C. M. Konrad, T. R. Frasier, H. Whitehead, and S. Gero, *Kin selection and allocate in sperm whales*, Behavioral Ecology **30**, 194–201 (2019).
- [11] L. S. Weilgart and H. Whitehead, *Observations of a sperm whale (physeter catodon) birth*, Journal of Mammalogy **67**, 399–401 (1986).
- [12] Y. Aluma, Z. Baron, R. Barrett, C. Baumgartner, G. Beguš, S. Bhattacharya, M. M. Bronstein, S. Dahan, O. Davis, S. de Haas, J. Defoe, J. DelPreto, R. Dessi, R. Diamant, J. Gatesy, K. George, S. Gero, D. Gibbons, D. Gibbons, S. Gil, D. F. Gruber, S. Goldwasser, O. Harve, A. Hernandez, M. Ishay, N. Jadhav, L. KC, A. Kenny, B. Kim, J. Lee, A. Leitao, M. Lucas, A. Maalouf, P. Malkin, Y. Mevorach, S. Pagani, O. Paradise, G. Petri, S. Poetto, E. Rossi, D. Rus, M. Salino-Hugg, A. Santoro, P. Sharma, D. Tchernov, A. Torralba, P. Tønnesen, D. M. Vogt, and R. J. Wood, *Detailed documentation of cooperative lifting during a sperm whale birth and evidence of this behavior since the late eocene* (2025), manuscript in preparation.
- [13] P. Tønnesen, S. Gero, M. Ladegaard, M. Johnson, and P. T. Madsen, *First-year sperm whale calves echolocate and perform long, deep dives*, Behavioral Ecology and Sociobiology **72**, 1–15 (2018).
- [14] F. B. De Waal and S. F. Brosnan, in *Cooperation in primates and humans: Mechanisms and evolution* (Springer, Berlin, Heidelberg, 2006) pp. 85–105.
- [15] R. C. Connor, *Cooperation beyond the dyad: on simple models and a complex society*, Philosophical Transactions of the Royal Society B: Biological Sciences **365**, 2687–2697 (2010).
- [16] L. A. Dugatkin, *Cooperation in animals: an evolutionary overview*, Biology and Philosophy **17**, 459–476 (2002).
- [17] R. Boyd and P. J. Richerson, *The evolution of indirect reciprocity*, Social networks **11**, 213–236 (1989).
- [18] R. Boyd and P. J. Richerson, *Culture and the evolution of human cooperation*, Philosophical Transactions of the Royal Society B: Biological Sciences **364**, 3281–3288 (2009).
- [19] L. Rendell, M. Cantor, S. Gero, H. Whitehead, and J. Mann, *Causes and consequences of female centrality in cetacean societies*, Philosophical Transactions of the Royal Society B **374**, 20180066 (2019).
- [20] M. N. Weiss, S. Ellis, and D. P. Croft, *Diversity and consequences of social network structure in toothed whales*, Frontiers in Marine Science **8**, 688842 (2021).
- [21] S. Gowans, B. Würsig, and L. Karczmarski, *The social structure and strategies of delphinids: predictions based on an ecological framework*, Advances in marine biology **53**, 195–294 (2007).
- [22] R. C. Connor, in *Cetacean societies: Field studies of dolphins and whales*, edited by J. Mann, R. Connor, P. Tyack, and H. Whitehead (University of Chicago Press, Chicago, USA, 2000) pp. 199–218.
- [23] T. A. Jefferson, S. Leatherwood, and M. A. Webber, *Marine Mammals of the World: FAO Species Identification Guide* (Food and Agriculture Organization of the United Nations, Rome, 1993).
- [24] G. N. di Sciara and F. Serena, *Observations of a fin whale Balaenoptera physalus giving birth in the mediterranean sea*, Marine Mammal Science **3**, 356–357 (1987).
- [25] P. B. Best, P. A. S. Canham, and N. MacLeod, *Patterns of reproduction in sperm whales, Physeter macrocephalus*, Reports of the International Whaling Commission **Special Issue 6**, 51–79 (1984).
- [26] C. M. Konrad, S. Gero, T. Frasier, and H. Whitehead, *Kinship influences sperm whale social organization within, but generally not among, social units*, Royal Society Open Science **5**, 180914 (2018).
- [27] S. Gero, D. Engelhaupt, and H. Whitehead, *Heterogeneous social associations within a sperm whale, Physeter macrocephalus, unit reflect pairwise relatedness*, Behavioral Ecology and Sociobiology **63**, 143–151 (2009).
- [28] R. L. Pitman, L. T. Ballance, S. I. Mesnick, and S. J. Chivers, *Killer whale predation on sperm whales: observations and implications*, Marine mammal science **17**, 494–507 (2001).
- [29] S. Gero, J. Gordon, and H. Whitehead, *Calves as social hubs: dynamics of the social network within sperm whale units*, Proceedings of the Royal Society B: Biological Sciences **280**, 20131113 (2013).
- [30] S. Gero, M. Milligan, C. Rinaldi, P. Francis, J. Gordon, C. Carlson, A. Steffen, P. Tyack, P. Evans, and H. Whitehead, *Behavior and social structure of the sperm whales of dominica, west indies*, Marine Mammal Science **30**, 905–922 (2014).
- [31] R. Gambell, C. Lockyer, and G. Ross, *Observations on the birth of a sperm whale calf*, South. African J. Sci. , 147–148 (1973).
- [32] K. Pinkerton and R. Gambell, *Aerial observations of sperm whale behaviour*, Norsk hvalfangst-tidende **57**, 126–138 (1968).
- [33] A. Pervushin, *Nabliudeniia za rodami u kasholotov. (observations of delivery of sperm-whales)*, Zoologicheskii Zhurnal **45**, 1892–1893 (1966).
- [34] A. Alexander, D. Steel, B. Slikas, K. Hoekzema, C. Carraher, M. Parks, R. Cronn, and C. S. Baker, *Low diversity in the mitogenome of sperm whales revealed by next-generation sequencing*, Genome biology and evolution **5**, 113–129 (2013).
- [35] S. Gero and H. Whitehead, *Critical decline of the eastern caribbean sperm whale population*, PLoS One **11**, e0162019 (2016).
- [36] H. Whitehead and L. Weilgart, *Patterns of visually observable behaviour and vocalizations in groups of female sperm whales*, Behaviour , 275–296 (1991).
- [37] Y. Zhang, P. S. Skardal, F. Battiston, G. Petri, and M. Lucas, *Deeper but smaller: Higher-order interactions increase linear stability but shrink basins*, Science Advances **10**, eado8049 (2024).
- [38] M. E. Newman, *Assortative mixing in networks*, Physical review letters **89**, 208701 (2002).
- [39] F. Fu, M. A. Nowak, N. A. Christakis, and J. H. Fowler, *The evolution of homophily*, Scientific reports **2**, 845 (2012).
- [40] J. B. Lancaster, *Play-mothering: the relations between juvenile females and young infants among free-ranging vervet monkeys (cevcopithecus aethiops)*, Folia primatologica **15**, 161–182 (1971).
- [41] P. B. Best, in *Behavior of marine animals: Current perspectives in research* (Springer USA, Boston, USA, 1979) pp. 227–289.
- [42] J. C. Gordon, *Sperm whale groups and social behaviour observed off sri lanka*, Report of the International Whaling

- Commission **37**, 205–217 (1987).
- [43] E. Demuru, P. F. Ferrari, and E. Palagi, *Is birth attendance a uniquely human feature? new evidence suggests that bonobo females protect and support the parturient*, *Evolution and human behavior* **39**, 502–510 (2018).
 - [44] P. H. Douglas, *Female sociality during the daytime birth of a wild bonobo at luikotale, democratic republic of the congo*, *Primates* **55**, 533–542 (2014).
 - [45] K. Hawkes, J. F. O’Connell, N. G. Blurton Jones, H. Alvarez, and E. L. Charnov, *Grandmothering, menopause, and the evolution of human life histories*, *Proceedings of the National Academy of Sciences* **95**, 1336–1339 (1998).
 - [46] D. J. Rankin and M. Taborsky, *Assortment and the evolution of generalized reciprocity*, *Evolution* **63**, 1913–1922 (2009).
 - [47] F. C. Santos and J. M. Pacheco, *Scale-free networks provide a unifying framework for the emergence of cooperation*, *Physical review letters* **95**, 098104 (2005).
 - [48] L. Marino, R. C. Connor, R. E. Fordyce, L. M. Herman, P. R. Hof, L. Lefebvre, D. Lusseau, B. McCowan, E. A. Nimchinsky, A. A. Pack, *et al.*, *Cetaceans have complex brains for complex cognition*, *PLoS Biology* **5**, e139 (2007).
 - [49] P. R. Hof and E. Van der Gucht, *Cortical complexity in cetacean brains*, *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology* **287**, 1142–1152 (2005).
 - [50] C. Butti, C. C. Sherwood, A. Y. Hakeem, J. M. Allman, and P. R. Hof, *Total number and volume of von economo neurons in the cerebral cortex of cetaceans*, *The Journal of Comparative Neurology* **515**, 243–259 (2009).
 - [51] J. N. Bruck, *Decades-long social memory in bottlenose dolphins*, *Proceedings of the Royal Society B: Biological Sciences* **280**, 20131726 (2013).
 - [52] J. R. Davies, E. Garcia-Pelegrin, L. Baciadonna, C. Pilenga, L. Favaro, and N. S. Clayton, *Episodic-like memory in common bottlenose dolphins*, *Current Biology* **32**, 3436–3442 (2022).
 - [53] M. J. Silk and D. J. Hodgson, *Differentiated social relationships and the pace-of-life-history*, *Trends in Ecology & Evolution* **36**, 498–506 (2021).
 - [54] T. A. Hersh, S. Gero, L. Rendell, M. Cantor, L. Weilgart, M. Amano, S. M. Dawson, E. Slooten, C. M. Johnson, I. Kerr, *et al.*, *Evidence from sperm whale clans of symbolic marking in non-human cultures*, *Proceedings of the National Academy of Sciences* **119**, e2201692119 (2022).
 - [55] S. Gero, H. Whitehead, and L. Rendell, *Individual, unit and vocal clan level identity cues in sperm whale codas*, *Royal Society Open Science* **3**, 150372 (2016).
 - [56] A. Leitao, M. Lucas, S. Poetto, T. A. Hersh, S. Gero, D. Gruber, M. Bronstein, and G. Petri, *Evidence of social learning across symbolic cultural barriers in sperm whales*, *arXiv preprint arXiv:2307.05304* (2023).
 - [57] T. M. Schulz, H. Whitehead, S. Gero, and L. Rendell, *Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function*, *Animal Behaviour* **76**, 1977–1988 (2008).
 - [58] P. Sharma, S. Gero, R. Payne, D. F. Gruber, D. Rus, A. Torralba, and J. Andreas, *Contextual and combinatorial structure in sperm whale vocalisations*, *Nature Communications* **15**, 3617 (2024).
 - [59] K. McComb and S. Semple, *Coevolution of vocal communication and sociality in primates*, *Biology letters* **1**, 381–385 (2005).
 - [60] D. L. Cheney and R. M. Seyfarth, *Flexible usage and social function in primate vocalizations*, *Proceedings of the National Academy of Sciences* **115**, 1974–1979 (2018).
 - [61] R. I. M. Dunbar, *Grooming, gossip, and the evolution of language* (Harvard University Press, Boston, USA, 1996).
 - [62] E. Chereskin, R. C. Connor, W. R. Friedman, F. H. Jensen, S. J. Allen, P. M. Sørensen, M. Krützen, and S. L. King, *Allied male dolphins use vocal exchanges to “bond at a distance”*, *Current Biology* **32**, 1657–1663 (2022).
 - [63] S. L. King and F. H. Jensen, *Rise of the machines: Integrating technology with playback experiments to study cetacean social cognition in the wild*, *Methods in Ecology and Evolution* **14**, 1873–1886 (2023).

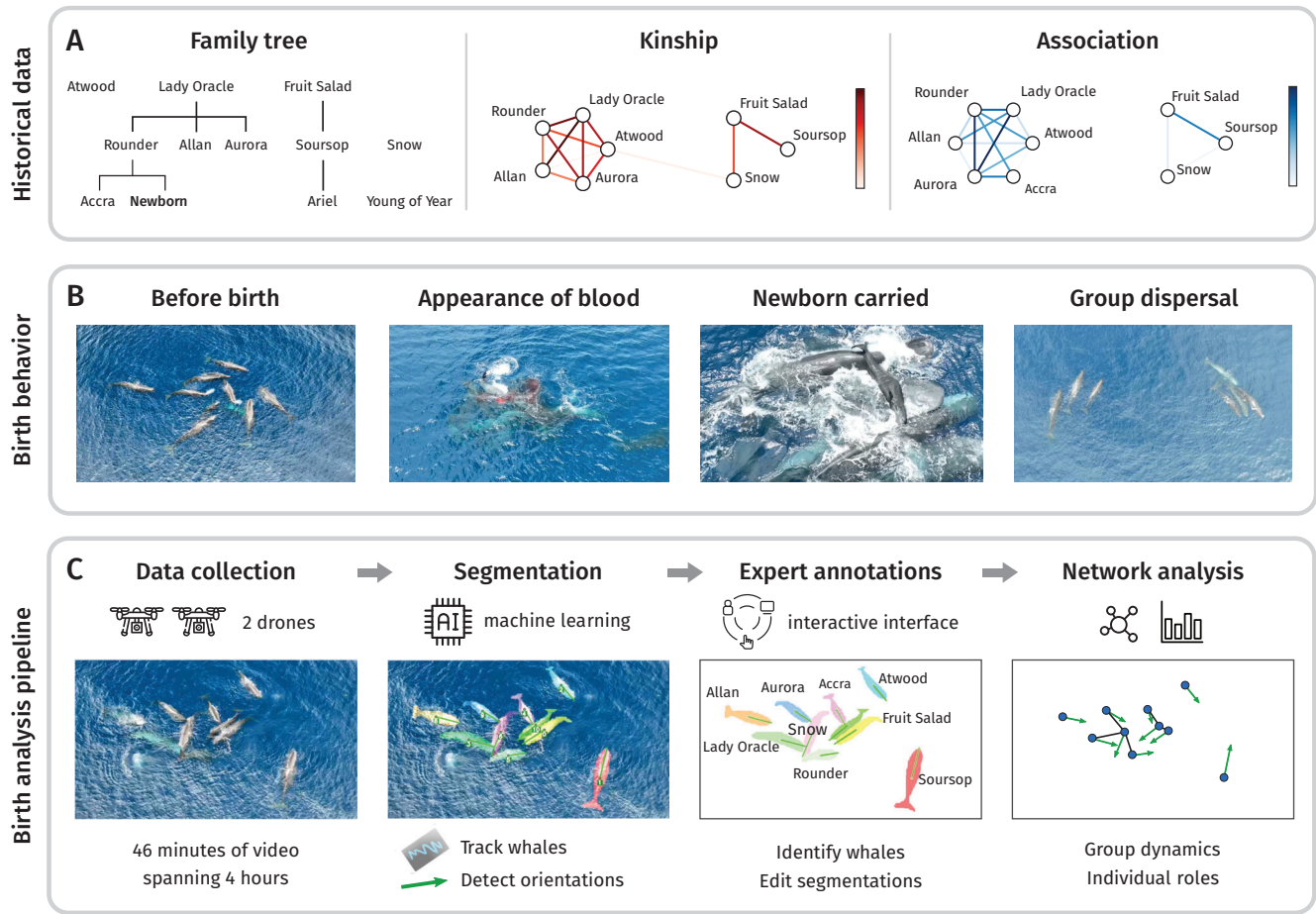


FIG. 1. **The birth of a sperm whale through the lens of a drone.** (A) Kin structure of the members of Unit A, showing the two distinct genetic branches and the respective kin relationships. (B) Examples of drone footage from the birth event. These include the plume of blood immediately after delivery completion, the density of whales from both branches of the family, and the adults apparently supporting the newborn to prevent it from sinking. (C) Overview of the pipeline developed to analyze the drone footage, including data collection and processing, computer vision, interactive manual annotation, and behavioral analysis.

IV Figures

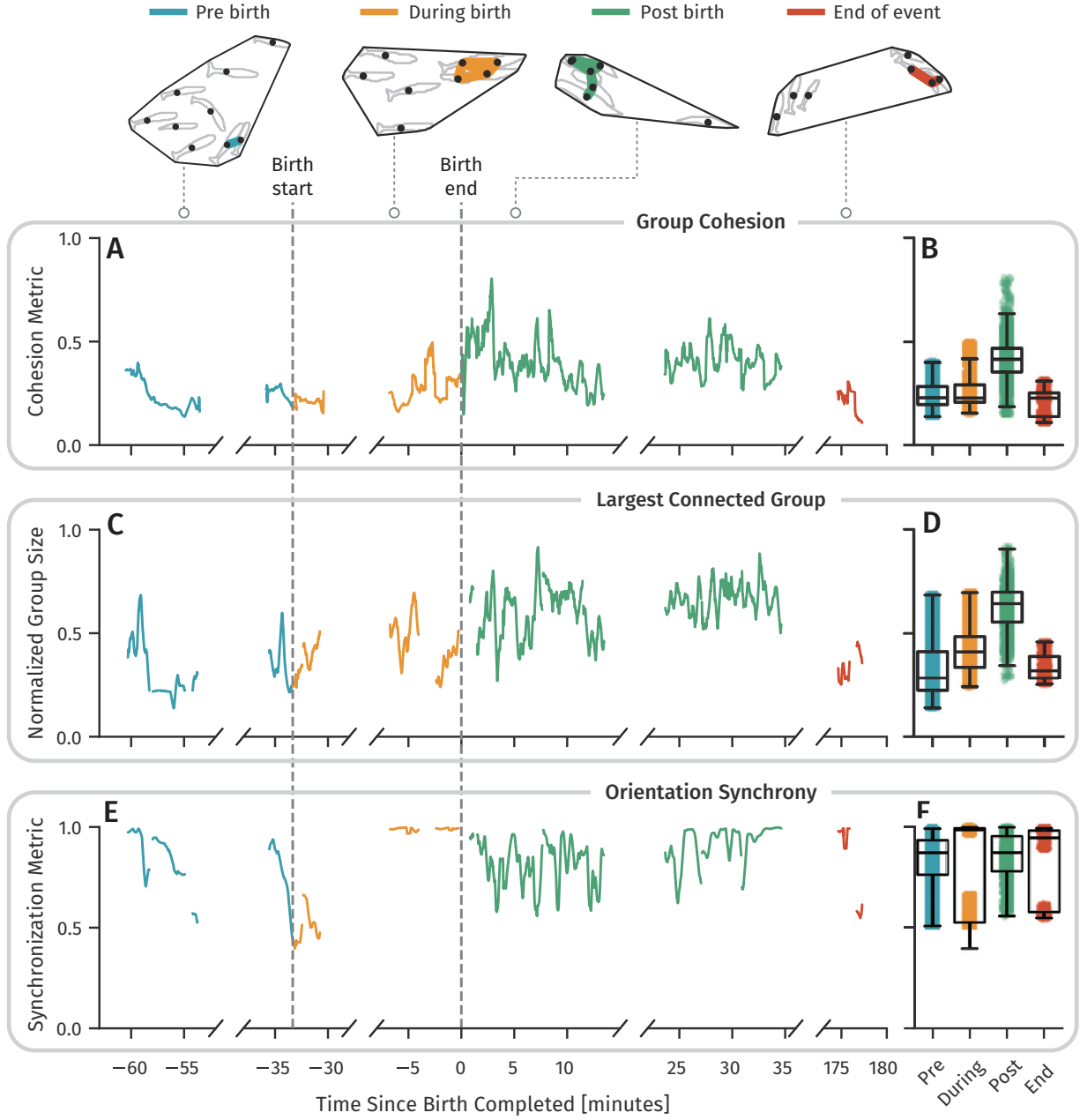


FIG. 2. Compactification and relaxation of the unit's structure throughout the birth event. We quantified the evolution of the unit-level structure of Unit A during the birth event using three metrics: (A, B) spatial cohesion, (C, D) the normalized connected group size, and (E, F) synchrony of orientations. For each metric, we show (A, C, E) its evolution over time, and (B, D, F) its distributions across the four phases: *pre-birth* (blue), *during-birth* (orange), *post-birth* (green), and *end-of-event* (red). The spatial cohesion and the normalized connected group size display similar evolutions: they increase leading up to the birth, plateau while the newborn needs assistance, and drop when returning to foraging. Orientation synchrony is also high approaching and following the birth then drops hours later when the whales disperse. These metrics indicate high social engagement and tight coordinated formations. We compared all pairs of distributions in (B, D, F), and obtained significant differences in all cases using Kruskal-Wallis ($p < 10^{-4}$) and Mann-Whitney ($p < 10^{-4}$) tests. See Fig. S1 for additional details.

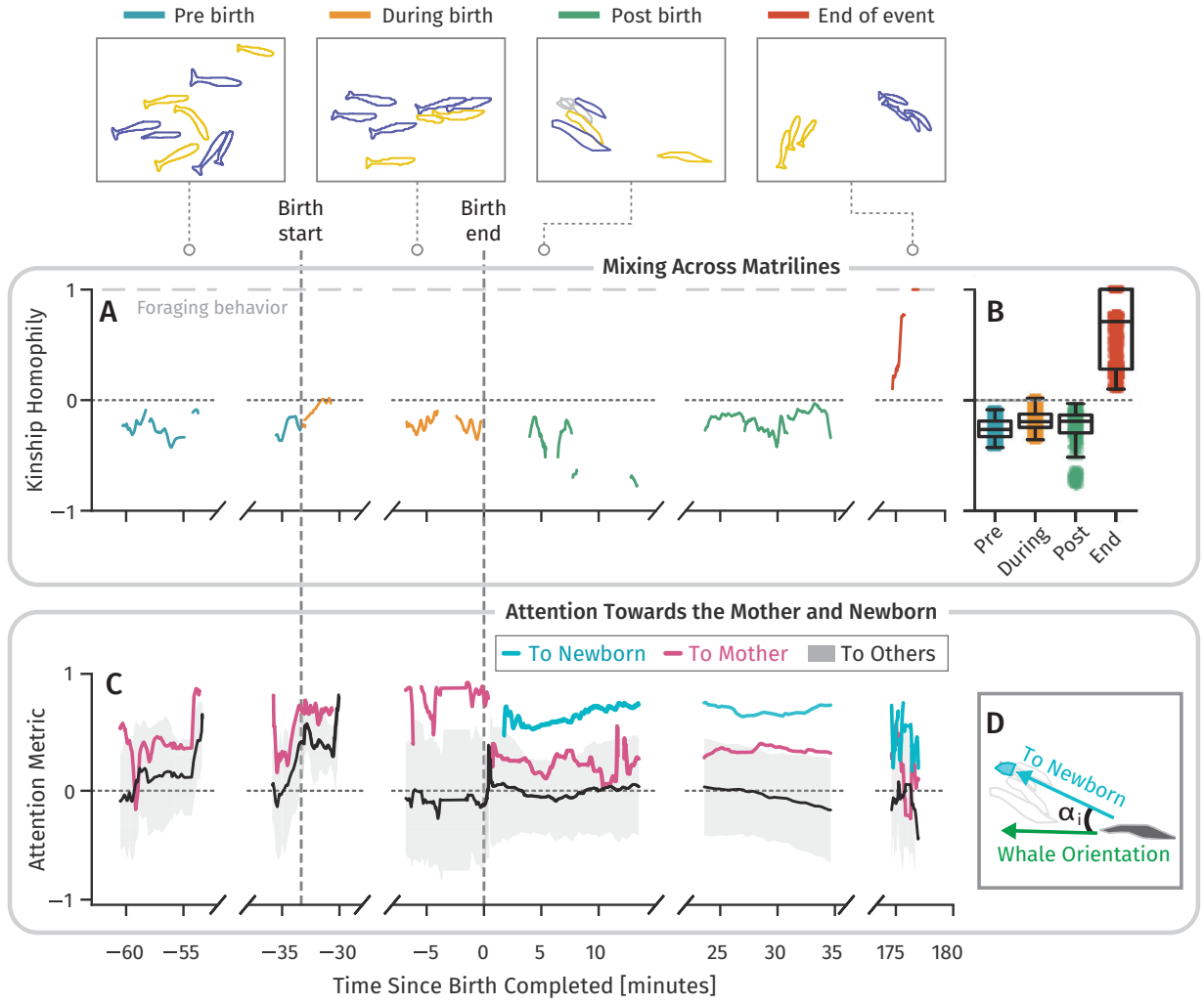


FIG. 3. **Collective dynamics within the unit.** We quantified the evolution of structure within the unit during the birth event using two metrics: (A, B) kinship homophily, and (C, D) attention. In (A, C) we show the evolution of each metric over time. In (B) we show the distributions of kinship homophily across the four phases of *pre-birth* (blue), *during-birth* (orange), *post-birth* (green), and *end-of-event* (red). Insets show representative snapshots of whale configurations, with individuals colored according to family branch: *Fruit Salad's* matriline (purple), *Lady Oracle's* matriline (yellow), and non-identified (gray). (A, B): Homophily is near zero throughout all phases except *end-of-event*, indicating that the two matriline of Unit A (see Fig. 1A) are mixed together – this is in contrast to behavior historically observed during typical foraging behavior. During the *end-of-event* phase, homophily peaks to typical foraging levels (gray dotted line) as individuals segregate according to the matriline. (C): The attention metric for an individual is close to 1 when all whales are oriented towards it, but close to -1 if they are oriented away (see diagram in (D)). The mother (pink) receives the most attention until the end of the *during-birth* phase, when the collective attention shifts to the newborn (turquoise). For reference, we show the average (black) and standard deviation (shaded region) of attention towards all other whales.

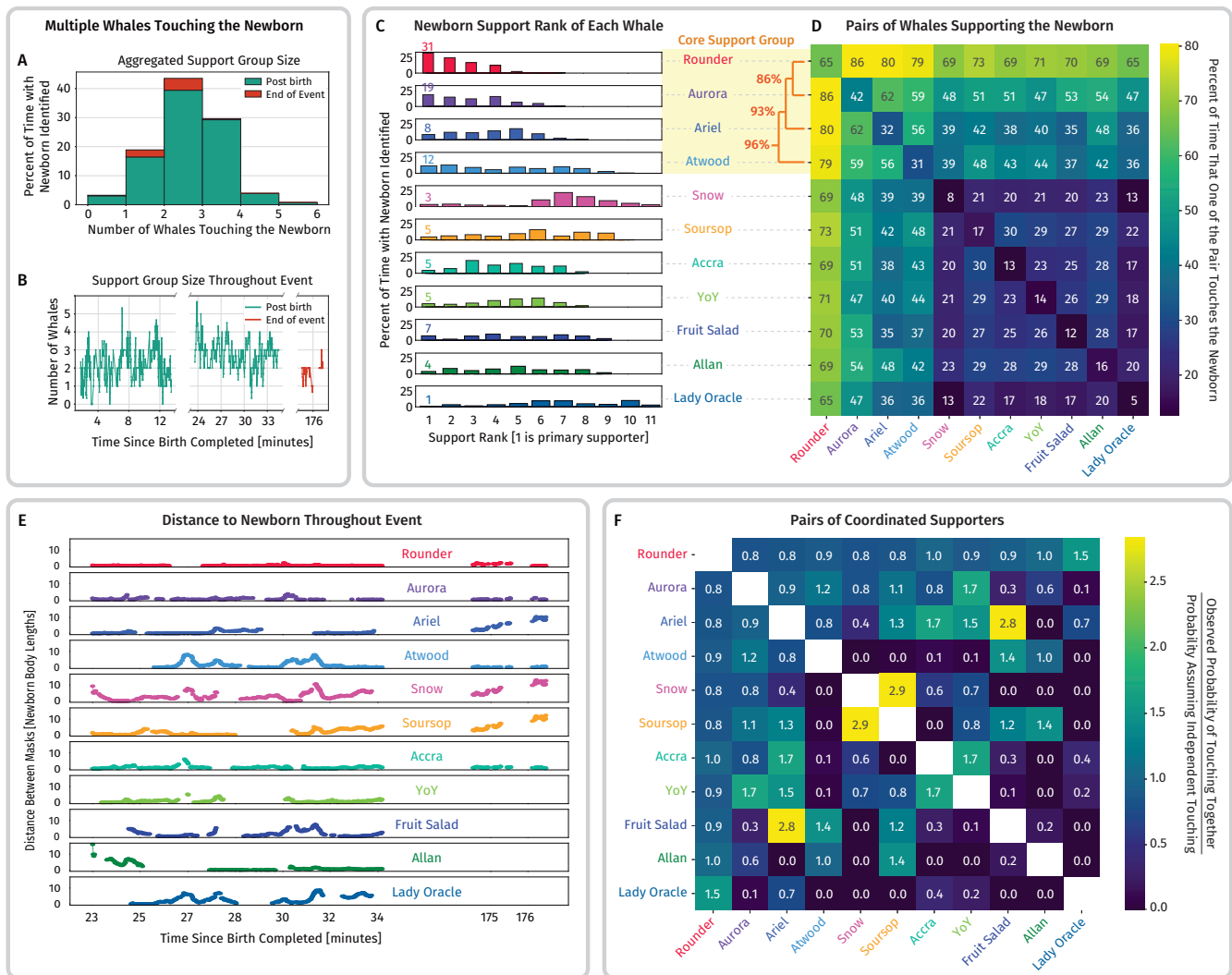


FIG. 4. Individual roles and interactions with the newborn. (A, B) A histogram and timeline of the number of whales touching the newborn. Two or more whales typically made physical contact with the newborn, with no significant gaps where no whales are touching the newborn. Results are after applying a moving average filter with duration 2 seconds. (C) Every whale served as the primary supporter of the newborn for at least 1% of the time in which the newborn was identified. (D) A core group of four whales helped support the newborn for 96% of the time, and other whales interacted with the newborn mostly in conjunction with a core group member. (E) We visualize the distance of each whale to the newborn as a function of time. (F) We identify patterns of coordinated contact by analyzing the observed probability of pairs of whales touching the newborn as compared to an assumption that their touching of the newborn is independent.

M Methods

During the birth event, two aerial drones were used to record 70 videos over the course of approximately 5 hours. As illustrated in Fig. M1, 17 of these videos were selected for processing; these total approximately 46 minutes in duration and span approximately 4 hours of the event. Videos are resized to be $2,500 \times 1,407$ pixels, and have 30 frames per second. Machine learning followed by interactive manual annotation was used to segment whales in individual frames, track them over time, and label whale identities. To achieve this, tailored computer vision pipelines and interactive dataset curation tools were developed. Following this dataset curation phase, network analysis tools were developed and applied to extract biological insights.

M.1 Video segmentation

To enable our analysis, we first segment individual whales in each frame, track the whales across frames, and determine the orientation of each whale. This yields a whale mask, bounding box, centroid, head location, and fluke location for each whale in each frame.

The video segmentation pipeline differs according to how challenging the video is to segment. In particular, we split the videos into two categories: *clean* and *noisy*. Noisy videos are characterized by significant interference that complicates the segmentation process, such as extensive whale interactions, diving whales, sun glare, and water disturbances.

M.1.1 Segmentation and tracking of clean videos

We trained an instance segmentation YOLO-v8-X [1] model, initialized with FastSAM [2] weights, on 682 frames of the recorded videos that were manually annotated with 5,119 sperm whale masks. Subsequently, we fine-tuned a specific model on the annotated frames of each video individually to get segmentation models tailored to each video. In some videos, we employed median filtering with a window of approximately 30 frames to remove water noise.

We applied the trained segmentation model to obtain masks of the whales in each frame of each video. Notably, each frame is considered independently, so mask indexes are not persistently associated with individual whales at this stage. As seen in the examples of Fig. M2A, some whales that are present in a certain frame may not be segmented, and indexes may swap between frames. The next phase of the pipeline then aims to match or track segmentations across consecutive frames to persistently associate masks with whales over time. Our approach is summarized in Fig. M2B. Given a frame and its segmentation masks, we apply a DeAoT tracker [3] to propagate the masks to the next frame. We then compare these propagated masks to the masks that were generated by the segmentation model for that frame, using Intersection over Union (IoU) for each pair of masks. If a mask is matched, i.e., if a mask was generated for a whale by both the tracker and the segmentation model, then we choose the one from the segmentation model. Otherwise, if one of the networks generated a mask that is not matched, then we use that mask directly in the new frame; this case represents either the tracker having a low confidence or the segmentation model missing a whale in a frame. Notably, when no mask is generated for a whale by either model, we retain its last known mask for up to 100 frames (approximately 3 seconds). Any new segmented unassigned mask is then compared to this stored mask to help

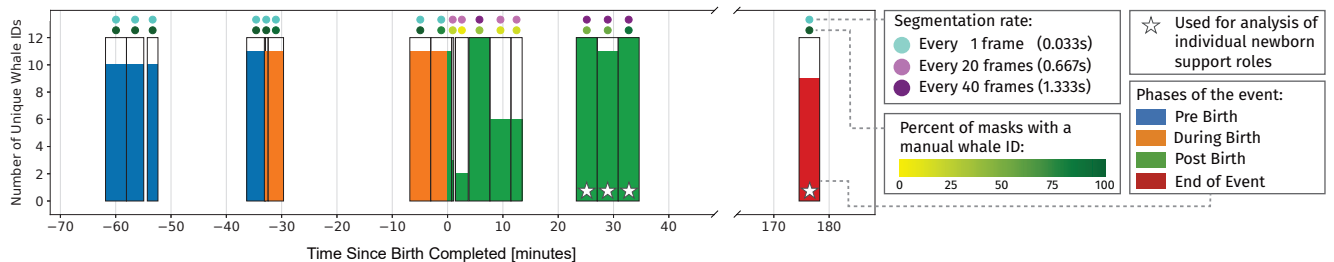


FIG. M1. **The timeline of data and annotations made throughout the birth event.** We recorded detailed drone videos throughout the birth event, and processed and annotated them using a tailored computer vision pipeline and new interactive tools. In the above plot, each drone video is represented as a box with a black outline. Each box is colored according to the event phases, and the height of each coloring represents the number of whales that were manually identified in the video. The top row of dots above each box indicates the segmentation density achieved, with sparser segmentations indicating more challenging noisier videos to process. The next row of dots indicates how many of the segmentation masks were manually assigned a whale identification. The stars denote videos that were used for analyzing individual interactions with the newborn (the results in Fig. 4C-F).

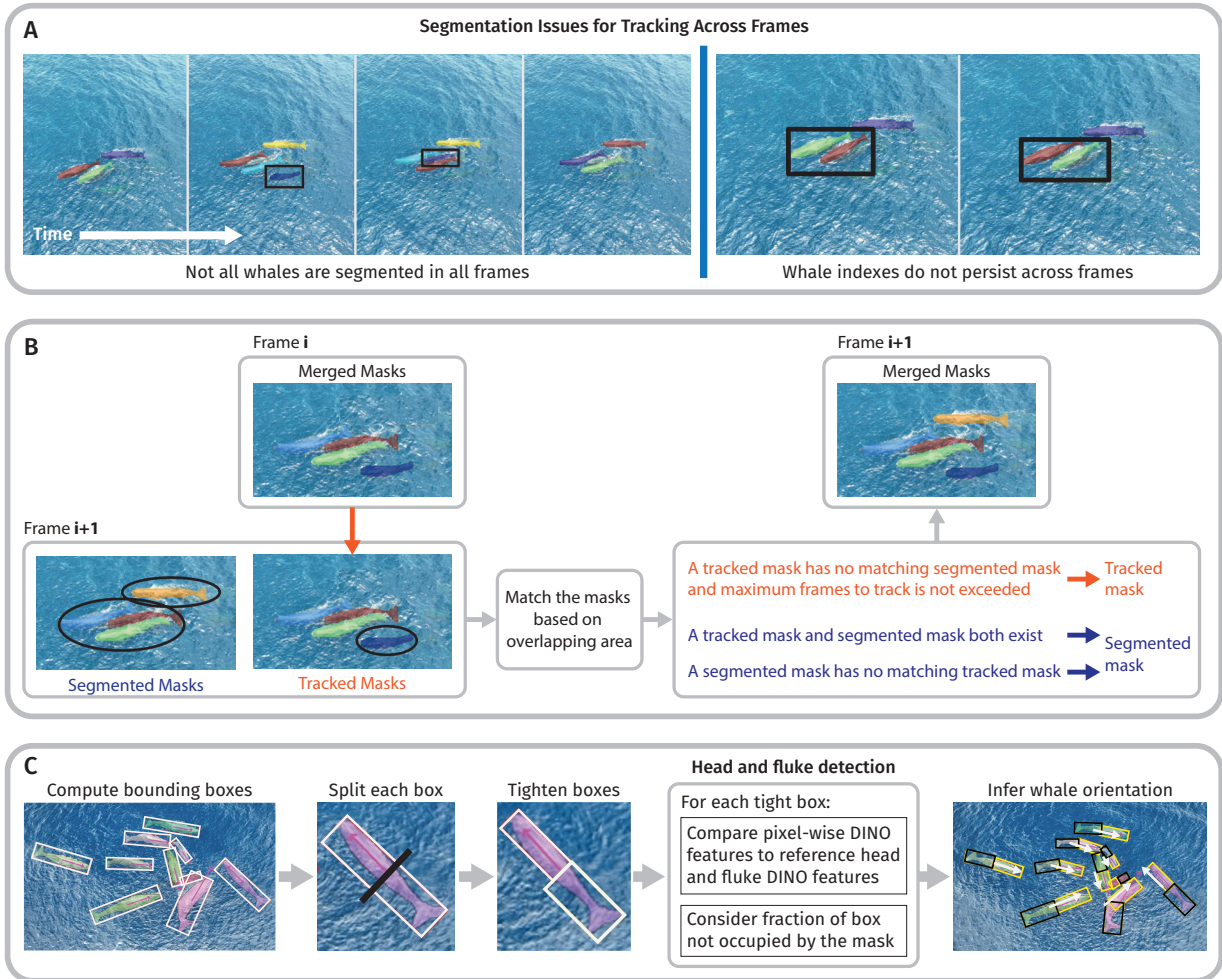


FIG. M2. **Segmentation, tracking, and orientation detection.** (A) Applying the trained segmentation model to generate masks in each frame independently is not sufficient since some visible whales may be missed and since there is no persistent notion of each whale across frames. (B) To address this, we track whales across frames by merging both the segmentation model and the tracker. (C) For each segmentation mask, we infer the orientation of the whale by combining two zero-shot methods for detecting the head and fluke: a heuristic based on shape analysis and a comparison of DINO features. Note that while the goal is to compute the orientation, it is shown in the illustration from the outset to support a clearer understanding of the image.

maintain identity continuity. This approach ensures consistent tracking, even when a whale is briefly obscured due to water disturbances. If a whale has not been detected by the segmentation model in 100 frames, then we stop applying the tracker to that mask; this ensures that noise or water is not tracked after a whale has dived. Once this process is complete for a video, we exclude all whales with less than 150 masks across all frames.

M.1.2 Segmentation and tracking of noisy videos

To handle the noisy videos, we manually annotated whales every 20 or 40 frames as indicated in Fig. M1. To associate masks of the same whale across frames, we applied a tracker when possible using a similar pipeline to that described for the clean videos. However, in most cases, water disturbance and sun glare precluded the use of a tracker. To address these cases, we first divided the videos into segments and noted whether the drone camera was substantially moving during each segment. For segments without substantial camera movement, we directly used IoU to match segmentation masks across subsequent frames. For segments with substantial camera movement, we first find the best rotation and translation to align the manually segmented whale mask centroids between two subsequent annotated frames. We do this by solving the 2D registration problem [4] by applying the iterative closest point (ICP) algorithm [5] on a provably approximated initialization for the global minimum as suggested in [6]. The computed transformation is then applied to the mask coordinates. Finally, we track whales by matching segmentations in subsequent frames based on their IoU after applying the mapping transformation.

M.1.3 Zero-shot orientation estimation

To determine the orientation of a whale, we first consider the 2D mask contour points of its mask and fit an ellipse in a least-squares sense [7]. The major axis of this ellipse indicates the axis of the mask. However, the direction of this axis still needs to be determined by identifying the whale’s head or fluke.

To determine this direction, we introduce and apply two zero-shot methods for detecting the head and fluke as summarized in Fig. M2C. For each whale mask in each frame, we first split it into two sub-masks. We divide it along the midline that splits its bounding box into two halves, where this midline is the minor axis of the fit ellipse. We then independently apply two detection methods as described below.

Foundation model feature-based approach: The first method is based on DINO features [8]. We utilize the method from [9] to extract pixel-level semantic feature descriptors from vision transformer models. We manually selected the head and fluke in representative frames at the start of the process and extracted their features. Then, for every pixel inside each sub-mask, we measure its descriptor cosine similarity to these previously extracted head or fluke features. We average these similarity scores within each sub-mask, and assign each sub-mask a head or fluke label based on the maximum average similarities computed. The certainty level is determined by the maximum cosine similarity that guided the labeling decision.

Shape analysis: We first calculate a tight bounding box around each sub-mask, then determine the percent of pixels that lie within the bounding box but outside the sub-mask. Typically, the sub-mask associated with the fluke has a higher percentage; the head tends to fill its tight bounding box, while the fluke tends to have more empty space within its tight bounding box. The certainty level is also represented by this computed percentage.

After applying both detection methods to a whale mask, we choose the result with the highest certainty estimate. Note that the shape analysis approach was observed to be chosen in most cases (provided higher certainty) and was much more robust (provided correct head/fluke detection), even when its certainty was slightly low. For videos in the *pre-birth* and *during-birth* phases, only the shape analysis approach was used.

Finally, after computing the orientations for all appearing whales across an entire video, for every whale, we identify the frames with the most certain detections and align the other frames around these keyframes.

M.2 Interactive segmentation corrections and fusion with biological data

Given the outputs of the machine learning pipelines, it is important to keep the human in the loop to verify their accuracy and curate a high-quality dataset. Towards this end, an interactive platform was developed to facilitate close collaboration between machine learning experts and biologists. The system, dubbed *Whale Tales*, provides an interactive web-based interface for exploring, visualizing, editing, and creating segmentations and annotations. Underlying these capabilities are algorithms for real-time segmentation and annotation processing, algorithms to alleviate manual effort with semi-automated propagation of edits, and efficient data storage. It also supports a variety of features to streamline the user experience including supporting multiple users, dataset versioning, timeline-based data exploration, and multiple video resolutions.

We made extensive use of *Whale Tales* to adjust model outputs. This included tasks such as merging segmentation instances over time to mark them as the same whale across dives, deleting spurious segmentations, adjusting whale orientations, and drawing new whale masks.

In parallel with these adjustments, *Whale Tales* was also used to identify whales based on features such as fluke profiles. An expert biologist (SG) interactively annotated whale masks with identities of the whales based on historical data, and these identities could be propagated throughout the videos based on the segmentation tracking.

This process yielded a comprehensive dataset with fine temporal resolution facilitated by a combination of manual annotation, automated machine learning, and human expertise. The dataset includes per-frame whale masks, orientations, centroids, bounding boxes, and identities.

M.3 Network analysis

M.3.1 Defining networks of sperm whales

Networks are a powerful abstraction to model the structure of interactions between units in a complex system. Formally, a network is defined by a set of N nodes and a set of M edges, where each edge represents the pairwise interaction between two nodes [10, 11]. Here, we naturally define the nodes to represent individual sperm whales. We then define the “interactions” in three different ways, building three different networks, to answer different biological questions. In one case, we define a *contact network* by saying that two sperm whales are interacting, and thus are connected by an edge, if they physically touch as determined by their segmentation masks. This approach can also be extended to create a *distance network* with weighted edges. Finally, we also create an *attention ego-network* to reflect whether the group is oriented towards individual whales.

Once a network is defined, network theory offers a plethora of metrics to analyze its structure from the local to global scales. An example of a local metric is the *degree*: the degree of a node (a whale) is the number of edges it has, i.e., the number of its neighbors in the network. An example of a global metric is the *density* of the network, defined as the number of edges divided by the total number of possible edges if all nodes were connected: $\rho = \frac{M}{N(N-1)}$ where M is the number of edges in the network and N is the number of nodes.

Note that individual sperm whales may appear and disappear from a video and that the interaction between two individuals may vary over time. For example, two individuals may be physically touching at some times but not others. This means that we are considering *temporal* networks, where edges can change over time [12, 13]. Here, a temporal network is a sequence of static networks, one for each video frame. Metrics such as degree or density can then be calculated over time.

A network at a given time can be represented by an $N \times N$ adjacency matrix A . If there is an edge between nodes i and j then $A_{ij} \neq 0$, and if there is no edge then $A_{ij} = 0$. A network is said to be *undirected* if its adjacency matrix is symmetric, with $A_{ij} = A_{ji}$, and *directed* otherwise. A network is also said to be *unweighted* if its edges have no weight, such that $A_{ij} \in \{0, 1\} \forall i, j$, and *weighted* otherwise. In a temporal network, the adjacency matrix is a function of time given by $A(t)$. Note that $A(t)$ is an $N \times N$ matrix for all t , even if sperm whales are in general not present at all times. To deal with the absence of a sperm whale at a time t , we set appropriate default values to its associated edges (details are below for each particular network).

M.3.2 Metrics of collective unit dynamics

Here, we define the global metrics shown in Fig. 2. These summarize properties of the whole unit of sperm whales using a single number at each timestep.

Connected group size using a contact network We define a contact network such that nodes represent individual whales, and there is an edge between them at a given time if they are in physical contact. Contact networks have been used in humans and non-human animals, for example, to study disease propagation [14, 15]. In each video frame, two sperm whales can be said to be in physical contact if their segmentation masks computed by the pipeline in Section M.1 overlap. However, due to the segmentation challenges and expected resolution, masks for two sperm whales that are touching may be close but not overlapping. We thus put an edge between two individuals when the smallest distance between any two points of their masks is smaller than a threshold of 15 pixels; formally, $A_{ij} = 1$ if $\text{dist}(\text{mask}_i, \text{mask}_j) < 15$ pixels and 0 otherwise. We also set the diagonal to zero, $A_{ii} = 0$, to avoid self-loops in the network. Finally, we set $A_{ij} = 0$ if either of the sperm whales is not present at that time. This network is undirected and unweighted.

In the contact network, the degree of a sperm whale at a given time represents the number of other sperm whales with which it is in contact. Similarly, the number of edges in the network represents the number of pairs of whales that are in contact. A complementary quantity in this case is the size of largest connected group, so-called *giant connected component* (GCC) in network science literature [10, 11]. A connected component of a network is a subset of nodes for which there is a path of edges between any two nodes. The GCC is the number of nodes in the largest subset that is thusly connected – it is the size of the largest group of sperm whales that are in contact with at least one sperm whale from this group. This GCC metric can distinguish between different ways in which a network could have a large number of edges – for example whether the network has many disconnected groups of two sperm whales in contact, or whether there is a single large group of whales in contact with each other.

Note that to compute the contact network and its GCC in a video frame, we only need whale segmentations in that frame – no whale identifications are needed. We can thus compute the GCC for every frame with segmentation masks to analyze its evolution over time. To smooth over spurious changes of the metric due to small changes of segmentation masks, a moving average filter with a duration of approximately 20s was applied. The results can be seen in Fig. 2C,D.

Group cohesion score: Group cohesion can be an important biological metric related to coordination and social interaction. To measure this quantitatively using the drone videos, we developed a metric to capture spatial cohesion without relying on knowing absolute distances that can be hard to extract from videos. The approach is based on the notion of a convex hull, which is defined as the minimum polygon that completely encompasses a given set of points. For each frame, we compute the convex hull of all segmented whale masks. We then compute the total area of all whale masks, and divide by the total area of the convex hull. This thus considers the amount of empty space within the bounds of the unit’s formation; if the whales are tightly packed, then the metric will be close to 1 since almost all of the hull area is occupied by whales. Computing this metric for every frame provides a continuous metric of the unit’s spatial density throughout the duration of the recordings.

This metric does not rely on whale identifications, and can thus be computed for every frame with whale segmentations. However, only frames with at least three whales are selected for analysis. To account for transient changes in segmentations between frames, the resulting series of cohesion scores is smoothed over time. In particular, a rolling mean is applied to the data with a window duration of 20 seconds. The results are plotted over time in Fig. 2A and aggregated by phase of the event in Fig. 2B.

Orientation synchrony: Synchrony is a standard metric in dynamical systems and network science [16–18] which has also been used to study animal behavior [19, 20]. Here, we use synchrony to measure the degree to which whale orientations are spatially aligned. Given the orientation $\theta_i \in [0, 2\pi]$ for each whale in a group of N whales, the Kuramoto order parameter [21] is defined as:

$$Z = Re^{i\Phi} = \frac{1}{N} \sum_{j=1}^N e^{i\theta_j} \quad (\text{M1})$$

which is a complex number. The level of synchrony is given by its magnitude: $R = |Z|$. This yields $R = 1$ when all whales are oriented with the same angle θ , and tends to 0 when their orientations are isotropically distributed.

This metric does not rely on whale identifications, and can thus be computed for every frame in which orientations were extracted from the segmentation masks. To smooth over spurious changes of the metric due to small changes of segmentation masks, a moving average filter with a duration of approximately 20s was applied. Results can be seen in Fig. 2E,F. Note that small uncertainties in the extraction of those orientations for one or a small number of whales do not impact R strongly because it is defined as an average over all whales.

M.3.3 Metrics of dynamics within the unit

Kinship homophily using a distance network: A distance network can first be constructed by creating a node for each individual sperm whale and then adding an edge between any pair of two individuals (it is a complete network). Each edge is then weighted with a weight equal to the distance between the segmentation masks of those two sperm whales. Formally, $A_{ij} = \text{dist}(\text{mask}_i, \text{mask}_j)$. We also set $A_{ij} = \text{NaN}$ (Not a Number) when either of the sperm whales is not present at that time. This is an undirected, but weighted network. This distance network can also be binarized by considering an edge between two individuals only if they are closer than a given threshold $d_{\max}$, such that $A_{ij} = 1$ if $\text{dist}(\text{mask}_i, \text{mask}_j) < d_{\max}$, and 0 otherwise. Similar to the contact network, computing the distance network in a video frame only requires whale segmentations – no identifications are needed. Note, however, that distances are not consistent between video frames due to changing zoom levels, altitudes, and orientations of the drone cameras that makes determining absolute distances a hard task.

This distance network can be used to investigate *kinship homophily*. Kinship homophily reflects the tendency of individuals to associate with individuals from the same matriline rather than a different matriline. To measure this, we use a standard metric from network science: the *assortativity coefficient* [22]. Given a categorical attribute (here the matriline) which can take values “ n ” or “ m ”, the assortativity coefficient is defined as

$$r = \frac{\sum_m e_{mm} - \sum_m a_m b_m}{1 - \sum_m a_m b_m}, \quad (\text{M2})$$

where e_{mn} is the fraction of edges connecting nodes of categories m and n , and $a_m = \sum_n e_{mn}$ and $b_n = \sum_m e_{mn}$ represent the fraction of edges incident to nodes of categories m and n , respectively. The coefficient r ranges from -1 (disassortative mixing) to $+1$ (assortative mixing), with $r = 0$ indicating no preference for within-group connections.

Defined this way, assortativity works well for large networks with balanced categories. In fact, large networks make the quantities used – fractions of edges – more stable by providing larger statistics. In the present study, however, we often deal with small networks (at most 11 uniquely identified whales) with few edges (depending on the chosen network constructions). To alleviate challenges that arise due to these small network sizes, we considered the distance network with a threshold set as half the length of the largest whale (determined in each frame) so as to have more edges than in the contact network. Moreover, the two matriline are not represented in balanced proportions in all frames – in some frames, one of the matriline is represented by no individual, or by a single individual. To deal with these imbalances, we used the adjusted assortativity coefficient proposed in [23] and designed for that purpose:

$$r_{\text{adjusted}} = \frac{\sum_m e_{mm}^* - \sum_m a_m^* b_m^*}{1 - \sum_m a_m^* b_m^*}, \quad (\text{M3})$$

where $e_{mn}^* = \frac{e_{mn}}{f_m f_n}$ are the elements of the adjusted mixing matrix, with f_m being the proportional size of group m . Similarly, the following quantities are also adjusted: $a_m^* = \sum_n e_{mn}^*$ and $b_n^* = \sum_m e_{mn}^*$.

Attention ego-network: The idea of the *attention ego-network* is to capture how much attention a whale is receiving from all other whales. As a measure of attention, we compute the degree to which an individual is oriented towards another target individual. Consider an individual i with its mask centroid located at (x_i, y_i) and its orientation directed at angle θ_i , and a target individual j with mask centroid located at (x_j, y_j) . The orientation angle that i would need to assume to be aiming directly at j can be computed as $\alpha_{ij} = \arctan2(y_j - y_i, x_j - x_i)$. We then define *attention* from i to j as

$$a(i, j) = \cos(\Delta\theta) \quad (\text{M4})$$

where $\Delta\theta = \alpha_{ij} - \theta_i$ is the angle between i 's current orientation and the orientation it should have to be directed towards j . By definition, $\Delta\theta = 0$ if i is directed towards j , $\Delta\theta = \pi$ if it is looking away from j , and takes values in between for all situations in between. By taking the cosine of this angle, the attention $a(i, j) = +1$ for whales directed perfectly towards the target, $a(i, j) = -1$ for whales looking perfectly away from it, and $a(i, j) = 0$ for whales looking perpendicularly to it.

We can use this measure to define an ego-network, which is a standard tool in social network studies [24, 25]. Here, the ego-network is defined for each sperm whale; its nodes represent individual sperm whales, and it contains edges only from all sperm whales i to the target sperm whale j (the *ego* whale), with a weight equal to the associated attention $a(i, j)$. Formally, $A_{ij} = a(i, j)$ if j is the target, and 0 otherwise. We also set the diagonal entries A_{jj} to NaN (Not a Number). Finally, we set $A_{ij} = \text{NaN}$ when either of the sperm whales i or j is not present in that frame. This is an ego-network that is directed, because $a(i, j) \neq a(j, i)$, and weighted.

We can build the ego-network for different target whales and compare them. In particular, we can compare aggregated attention on the mother or the newborn to attention on all other whales and how these metrics evolve over time. Results can be seen in Fig. 3C.

M.3.4 Metrics of individual interactions and cooperative birth dynamics

Of particular interest when considering the roles of individuals is how they behave relative to the newborn. We hypothesize that the newborn requires active support following birth to compensate for its negative buoyancy, that providing this support is a highly social and cooperative endeavor, and that the amount of support individual whales provide will be related to their kinship or to their caregiving experience. Results are summarized in Figure 4, and leverage four key metrics described below.

Distance to the newborn: We compute the distance between a pair of whales in a frame by finding the shortest distance between their segmentation masks. In particular, the analyses focus on the distance from the newborn to each other whale. By definition the newborn is thus present in each frame considered, so it can also be used as a reference length scale; we measure distances within each frame in units of the newborn body length in that frame.

Physical contact with the newborn: If the distance between the newborn and another whale is below a threshold, the pair of whales are considered *touching*. This is a binary measure that is either true or false, and the current analyses use a threshold of 20 pixels.

Support rank: The distance and touching metrics described above are also combined to *rank whales in order of their level of support* within each frame. First, whales that touch the newborn are sorted according to the amount of mask overlap, with larger overlaps indicating a greater supporting role as they carry the newborn. Then whales that do not touch the newborn are sorted by distance to the newborn, with closer distances indicating a higher likelihood of future interactions. In the final ranking, lower numbers indicate greater levels of support – a whale with the support rank of 1 is the primary supporter in that frame.

Coordinated touching of the newborn: To investigate coordination of support among pairs of whales, we define a metric that compares the observed probability of the pair providing simultaneous support to the probabilities of each of them providing support if their support was independent of each other. For a whale X_i , we compute $\Pr(X_i)$ as the proportion of seconds in which whale X_i is touching the newborn relative to the total seconds where both the newborn and whale X_i are identified. Then for a pair of whales X_i and X_j , we compute $\Pr(X_i, X_j)$ as the proportion of seconds where both whales are simultaneously touching the newborn, relative to the total number of seconds where the three individuals (X_i , X_j , and the newborn) are identified. To assess the likelihood of coordinated touching, with respect to the baselines of X_i and X_j , we report the ratio $\frac{\Pr(X_i, X_j)}{\Pr(X_i) \cdot \Pr(X_j)}$. This ratio compares the observed probability of whales touching together with the probability expected under the assumption of independent touching. Results are presented in Fig. 4. In Section S3.2, we also report the results of extending the analysis from pairs of whales to triplets of whales.

M.4 Statistical analysis

Throughout some of the above analyses, distributions of a metric were aggregated with respect to phases of the event (*pre-birth*, *during-birth*, *post-birth*, and *end-of-event*). In particular, see the distributions in Fig. 2B,C,D and Fig. 3B. The distributions violated normality assumptions and exhibited unequal sample sizes. In each case, the effect was tested using Kruskal-Wallis tests with a significance level of 0.05. When a significant effect was detected, post hoc pairwise comparisons were performed using Mann-Whitney U tests with Bonferroni correction to adjust for multiple comparisons. We also reported pairwise effect sizes with Cliff’s δ which measures directed overlap between two distributions and takes values within $[-1, 1]$ (0 indicates no overlap) [26, 27]. See Fig. S1 and Fig. S2 for results.

-
- [1] R. Varghese and S. M., *Yolov8: A novel object detection algorithm with enhanced performance and robustness*, in *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)* (2024) pp. 1–6.
 - [2] X. Zhao, W. Ding, Y. An, Y. Du, T. Yu, M. Li, M. Tang, and J. Wang, Fast segment anything (2023), [arXiv:2306.12156 \[cs.CV\]](https://arxiv.org/abs/2306.12156).
 - [3] Z. Yang and Y. Yang, *Decoupling features in hierarchical propagation for video object segmentation*, in *Advances in Neural Information Processing Systems (NeurIPS)* (2022).
 - [4] F. Pomerleau, F. Colas, R. Siegwart, et al., *A review of point cloud registration algorithms for mobile robotics*, *Foundations and Trends® in Robotics* **4**, 1–104 (2015).
 - [5] P. J. Besl and N. D. McKay, *Method for registration of 3-d shapes*, in *Sensor fusion IV: control paradigms and data structures*, Vol. 1611 (Spie, 1992) pp. 586–606.
 - [6] I. Jubran, A. Maalouf, R. Kimmel, and D. Feldman, *Provably approximated point cloud registration*, in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2021) pp. 13269–13278.
 - [7] A. Fitzgibbon and R. Fisher, *A buyer’s guide to conic fitting.*, in *Proceedings of the British Machine Vision Conference*, edited by D. Pycock (1995) pp. 511 – 5110, doi:10.5244/C.9.51.
 - [8] M. Caron, H. Touvron, I. Misra, H. Jégou, J. Mairal, P. Bojanowski, and A. Joulin, *Emerging properties in self-supervised vision transformers*, in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2021).
 - [9] S. Amir, Y. Gandelsman, S. Bagon, and T. Dekel, *On the effectiveness of vit features as local semantic descriptors*, in *Computer Vision – ECCV 2022 Workshops*, edited by L. Karlinsky, T. Michaeli, and K. Nishino (Springer Nature Switzerland, Cham, 2023) pp. 39–55.
 - [10] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford; New York, 2010).
 - [11] A.-L. Barabási and M. Pósfai, *Network Science* (Cambridge University Press, Cambridge, 2016).
 - [12] P. Holme and J. Saramäki, eds., *Temporal Networks*, Understanding Complex Systems (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013).
 - [13] N. Masuda and R. Lambiotte, *A Guidance to Temporal Networks* (World Scientific, Singapore, 2016).
 - [14] M. Salathé, M. Kazandjieva, J. W. Lee, P. Levis, M. W. Feldman, and J. H. Jones, *A high-resolution human contact network for infectious disease transmission*, *Proceedings of the National Academy of Sciences* **107**, 22020–22025 (2010).

- [15] L. A. White, J. D. Forester, and M. E. Craft, *Using contact networks to explore mechanisms of parasite transmission in wildlife*, [Biological Reviews](#) **92**, 389–409 (2017).
- [16] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, UK).
- [17] S. H. Strogatz, *Sync: The Emerging Science of Spontaneous Order* (Penguin UK, London, 2004).
- [18] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, *Synchronization in complex networks*, [Physics Reports](#) **469**, 93–153 (2008).
- [19] I. D. Couzin, *Synchronization: the key to effective communication in animal collectives*, *Trends in cognitive sciences* (2018).
- [20] N. W. Bode, J. J. Faria, D. W. Franks, J. Krause, and A. J. Wood, *How perceived threat increases synchronization in collectively moving animal groups*, *Proceedings of the Royal Society B: Biological Sciences* **277**, 3065–3070 (2010).
- [21] Y. Zhang, P. S. Skardal, F. Battiston, G. Petri, and M. Lucas, *Deeper but smaller: Higher-order interactions increase linear stability but shrink basins*, *Science Advances* **10**, eado8049 (2024).
- [22] R. Noldus and P. Van Mieghem, *Assortativity in complex networks*, [Journal of Complex Networks](#) **3**, 507–542 (2015).
- [23] F. Karimi and M. Oliveira, *On the inadequacy of nominal assortativity for assessing homophily in networks*, [Scientific Reports](#) **13**, 21053 (2023).
- [24] J. Leskovec and J. Mcauley, *Learning to discover social circles in ego networks*, in *Advances in Neural Information Processing Systems*, Vol. 25, edited by F. Pereira, C. Burges, L. Bottou, and K. Weinberger (2012).
- [25] V. Arnaboldi, M. Conti, M. La Gala, A. Passarella, and F. Pezzoni, *Ego network structure in online social networks and its impact on information diffusion*, [Computer Communications](#) **76**, 26–41 (2016).
- [26] N. Cliff, *Dominance statistics: Ordinal analyses to answer ordinal questions.*, *Psychological bulletin* **114**, 494 (1993).
- [27] J. Romano, J. D. Kromrey, J. Coraggio, J. Skowronek, and L. Devine, *Exploring methods for evaluating group differences on the nsse and other surveys: Are the t-test and cohen'sd indices the most appropriate choices*, in *annual meeting of the Southern Association for Institutional Research*, Vol. 14 (Citeseer, 2006).

Supplementary Material

S1 Collective unit dynamics: statistical results

In Fig. S1, we show the distributions of the three metrics related to unit-level dynamics. These metrics are previously illustrated in Fig. 2B,D,F. Here, the distribution plots are augmented with results of statistical comparisons. For each metric, a Kruskal-Wallis test indicated a significant effect of the event phase ($p < 10^{-4}$), and pairwise Mann-Whitney U tests (with Bonferroni correction to adjust for multiple comparisons) indicated significant differences for each pair of phases (all $p < 10^{-4}$). In Fig. S1, we display the pairwise effect sizes measured with Cliff's δ . Absolute values of $|\delta| \simeq 0$ indicate a small effect and $|\delta| \simeq 1$ indicate a large effect.

S2 Dynamics within the unit

S2.1 Kinship homophily: statistical results

In Fig. S2, we show the distributions of kinship homophily in the four phases; this is the same data as in Fig. 3B, but here augmented with results of statistical comparisons. A Kruskal-Wallis test indicated a significant effect of the phase ($p < 10^{-4}$) on homophily, and pairwise Mann-Whitney U tests (with Bonferroni correction to adjust for multiple comparisons) indicated significant differences for each pair (all $p < 10^{-4}$). In Fig. S2, we also display the pairwise effect sizes measured with Cliff's δ . Absolute values of $|\delta| \simeq 0$ indicate a small effect and $|\delta| \simeq 1$ indicate a large effect.

S2.2 Matriline association during the birth event compared to legacy matriline association

First, we verified that in typical foraging behavior outside of the birth event, association data strongly correlates with kinship data (see Fig. 1A), indicating that whales are typically seen at the surface with whales from their own matriline but not from the other one. This is indicated by a strong Pearson correlation ($r = 0.88$, $p \ll 10^{-5}$) between the weighted adjacency matrix of the kinship network and that of the association network (see Fig. S3).

Second, in addition to the kinship homophily metric, we compared the association network from typical foraging behavior (described above) with the association network at a given moment during the birth event. Specifically, we used the network of inverse distances between whales, averaged over a given birth video recording. For each

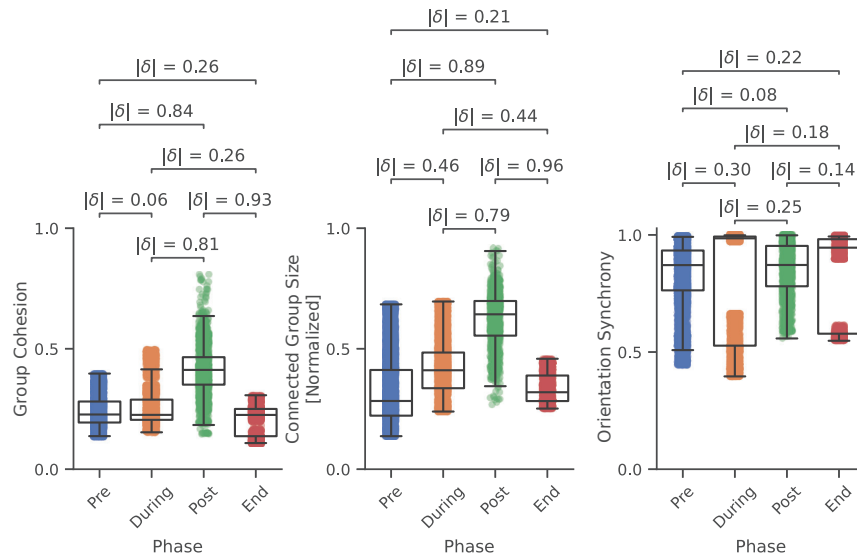


FIG. S1. Distributions of the three unit-level metrics (same data as in Fig. 2B,D,F) with Cliff's δ effect size values corresponding to the Mann-Whitney U pairwise tests, Bonferroni-adjusted.

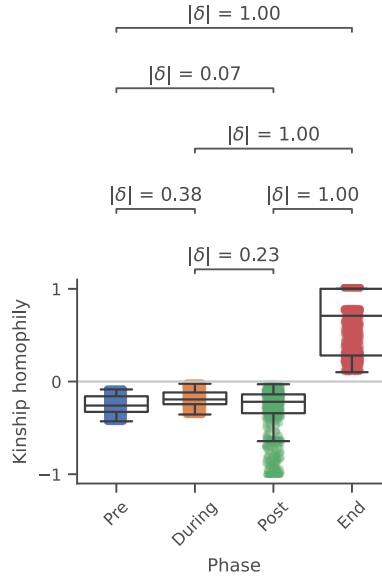


FIG. S2. Distributions of kinship homophily (same data as in Fig. 3B) with Cliff's δ effect size values corresponding to the Mann-Whitney U pairwise tests, Bonferroni-adjusted.

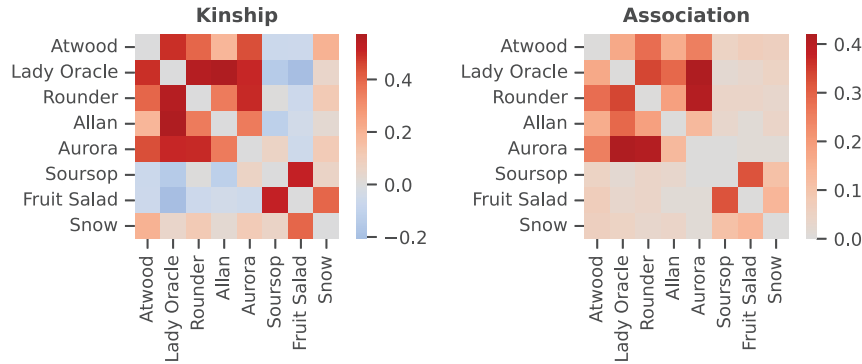


FIG. S3. In typical foraging behavior outside of the birth event, the kinship graph strongly correlates with the association graph, indicating that whales are typically seen at the surface with whales from their own matriline but not from the other one.

video, we computed the Pearson correlation between the adjacency matrix of the network and that of the non-birth network. Results are shown in Fig. S4 where we show the p -value and the correlation coefficient r . Note that, for each comparison, we only kept the graph of the whales that were identified in the video and that were also present in the typical foraging data. The resulting number of whales was sometimes very low (mostly because of the difficulty of identifying whales in videos close to the birth), and we decided to discard videos with fewer than four whales in the final network – those videos have no data points in the figure. Results indicate that before and around the birth, correlation is mostly non-significant, but it becomes significant much later when they segregate into their two matriline (during the *end-of-event* phase), with correlation coefficients larger than 0.5. This is consistent with the dynamics of the kinship homophily metric shown in Fig. 3A,B.

S2.3 Attention dynamics

In Fig. S5, we illustrate the attention data previously illustrated in Fig. 3C, but here aggregated by video (there is a box for each video). The boxes summarize the attention level given to all whales excluding the newborn and the mother (*Rounder*). The attention values for the mother are shown as a pink whale symbol, and values for the newborn are shown as a smaller cyan whale symbol. For the newborn data points, the z values represent the number of standard deviations by which the newborn's attention level exceeds the average attention level of the other whales.

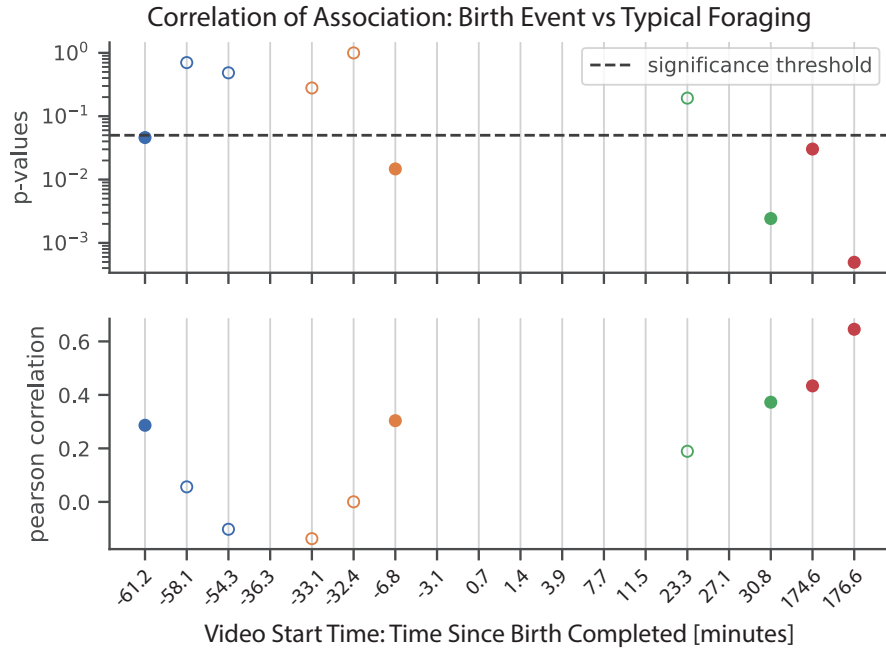


FIG. S4. Pearson correlation between the non-birth association network and the inverse distance network in each birth video recording: (top) p -value, and (bottom) correlation coefficient r . The horizontal dashed line indicates the significance threshold of 0.05. Videos with significant p -values are shown with solid circles, and non-significant videos with empty circles. Colors indicate the phase of the video: *pre-birth* (blue), *during-birth* (orange), *post-birth* (green), and *end-of-event* (red).

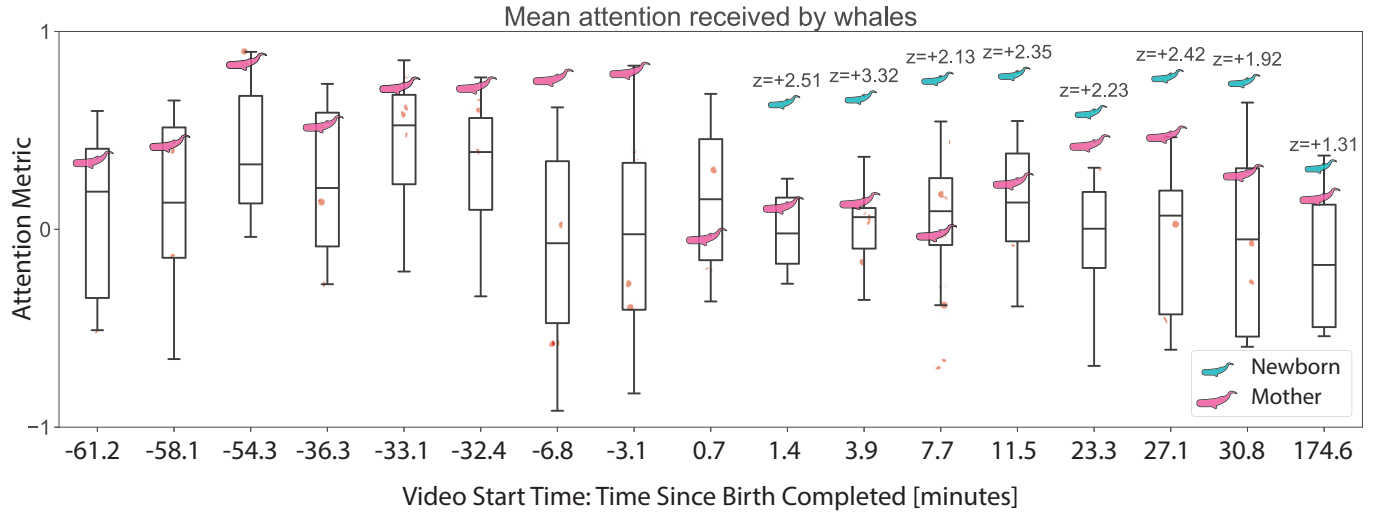


FIG. S5. The attention received by each whale in each video recording, highlighting the increased attention towards the mother and the newborn. Note that this graphs reflects the same underlying data as shown in Fig. 3C

This visualization makes it clear that the mother is the center of attention of the unit during the *pre-birth* phase, and that attention shifts to the newborn after delivery is completed (with the mother still maintaining a high attention level as well).

S2.4 Centrality of individual whales

An important metric for social interactions and group dynamics is whether whales are in the center or the edge of the unit. This can be closely related to attention and to behavioral roles within the unit. Towards this end, a metric was developed to indicate whether individual whales are towards the center or the edge of the unit's formation.

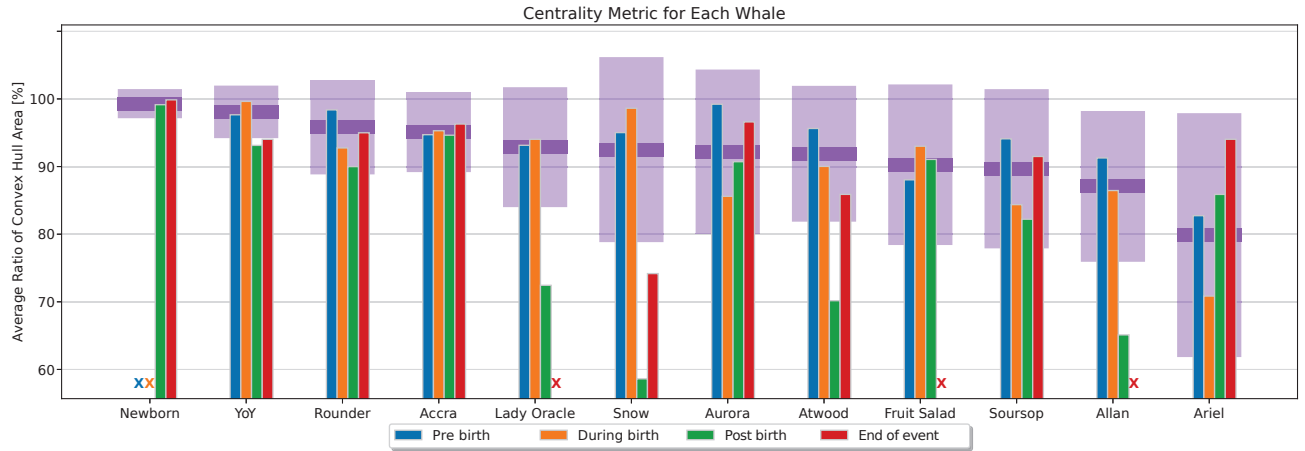


FIG. S6. A metric based on the convex hull was developed to investigate whether individual whales tend to be in the center of the unit or at the edge. The purple boxes indicate the aggregated centrality of each whale throughout all videos in which they were identified, with the dark purple indicating the mean and the light purple spanning a standard deviation in each direction. The thinner bars represent the centrality metric for each whale across all frames in which they were identified, aggregated by the birth event phases. An “X” indicates an analysis region in which the whale was not identified.

Firstly, a convex hull encompassing all whale segmentation masks is computed for each frame. Each whale segmentation mask is then iteratively ignored, and a new convex hull is computed; if a whale is completely in the interior of the unit then the convex hull will not change when it is removed, but a whale at the extreme of the formation will have a substantial effect convex hull. We thus compute the ratio between the hull area of the group without a given whale and the hull area of all whales. This will tend towards 1 if the whale is central to the group, and towards 0 if the whale is very far from the group. This provides a continuous metric of centrality that does not rely on absolute distances.

This metric can be computed for each frame and each identified whale. Results are then aggregated over all videos, to indicate which whales tend to be in the center of the unit throughout the event. Results are also aggregated by analysis region, to see how centrality changes over time relative to the key birth event milestones.

Results are presented in Fig. S6. They indicate that the newborn and the youngest calf (*Young-of-Year*) are predominantly in the interior of the unit. They also suggest trends of individuals relative to key phases of the birth event; for example, *Rounder* is typically in the interior of the unit but spends more time towards the edge during and after the birth, perhaps as the other whales follow her through the process. *Lady Oracle* is typically near the unit’s interior before and during birth, but then shifts to the edge after birth. Similar trends are seen for *Snow* and *Atwood*. The opposite trend is suggested for *Aurora* and *Ariel*, who tend to move towards the interior over time starting just after the birth as the younger animals press towards the newborn and displace the older females. *Ariel* in particular has a large variation over the course of the event, in contrast to whales such as *Accra* (mostly interior) or *Fruit Salad* that maintain a consistent relative position over time. *Allan* on average is near the edge of the unit, but is more interior before and during the birth when the entire unit is close together.

This metric thus begins to explore trends in the centrality of individual whales over time, which is closely related to analyses of caregiving and newborn interactions. Future work can also further investigate this centrality relative to the unit’s orientation and motion; for example, being on the front edge of the unit can have different behavioral implications than being on the rear edge of the unit.

S3 Interactions with the newborn

We provide additional analyses to further investigate how whales interacted with the newborn during the event. First, Fig. S7A presents histograms of each whale’s distance from the newborn, demonstrating that the core support group (*Rounder*, *Aurora*, *Ariel*, and *Atwood*) identified in Figure 4 consistently maintained close proximity to the newborn. Specifically, the mother (*Rounder*) was never observed with a distance exceeding 1.3 newborn body lengths, and she was typically less than 0.5 newborn body lengths from the newborn. Second, Fig. S7B shows the support rank of each whale over time, smoothed using a rolling median filter with a window size of 7 consecutive segmented frames. This highlights that the mother and aunt (*Aurora*) had consistently dominant support rankings (between 1 to 3) throughout the observed period, with no substantial gaps in their roles as primary supporters.

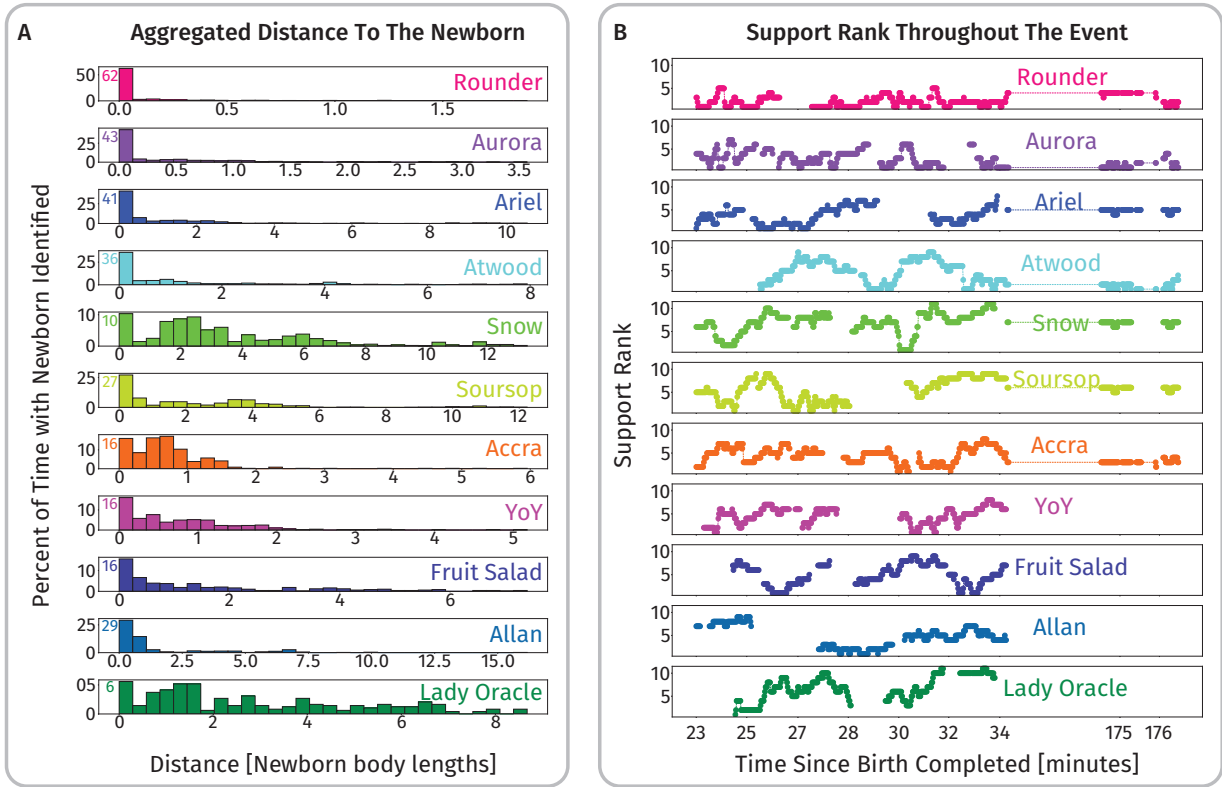


FIG. S7. (A) A histogram of the distance from each whale to the newborn during the *post-birth* and *end-of-event* phases. (B) The support rank of each whale as a function of time.

Fig. S8A displays the percent of time, relative to the newborn's identified presence, during which each whale was also identified (see the parenthetical blue values). It further shows the proportion of that time that each whale was touching the newborn; it highlights that all whales touched the newborn for at least 8% of their co-identified time (at least 32 seconds). Notably, the core support group showed contact ranging from approximately 40% to 74% of their co-identified time. Finally, Fig. S8B visualizes each whale's contact with the newborn across time, confirming that members of the core group touched the newborn throughout the observation period without major interruptions.

S3.1 Additional behavioral patterns of individual whales

Young whales: Interestingly, the youngest members of the unit, including the newborn's sister *Accra* and the *Young-of-Year* from the previous season, also remain relatively close to the newborn but usually without direct interaction.

Elder whales: Surprisingly, the elders of both matriline (*Fruit Salad* and the newborn's grandmother *Lady Oracle*) demonstrated inconsistent support of the newborn; they often maintained proximity but contributed minimally to physical support of the newborn. This may be partly explained by the observation that *Fruit Salad* was often escorting the *Young-of-Year* at the edge of the cluster of whales.

Male whale: While the adolescent male (*Allan*) remained closer to the newborn for more time than the grandmother and touched the newborn more often, he was also often at the unit's periphery. This may reflect the ongoing process of separating from his natal unit through social marginalization [1].

S3.2 Turn-taking, coordination, and coverage for supporting the newborn

To further investigate the dynamics of cooperative caregiving among individuals, we expand on the analyses depicted in Figure 4 by presenting additional data on pairwise and unit-level cooperation. Specifically, we examine (i) the similarity in proximity to the newborn over time between whale pairs or groups, and (ii) the duration of simultaneous and separate contact with the newborn.

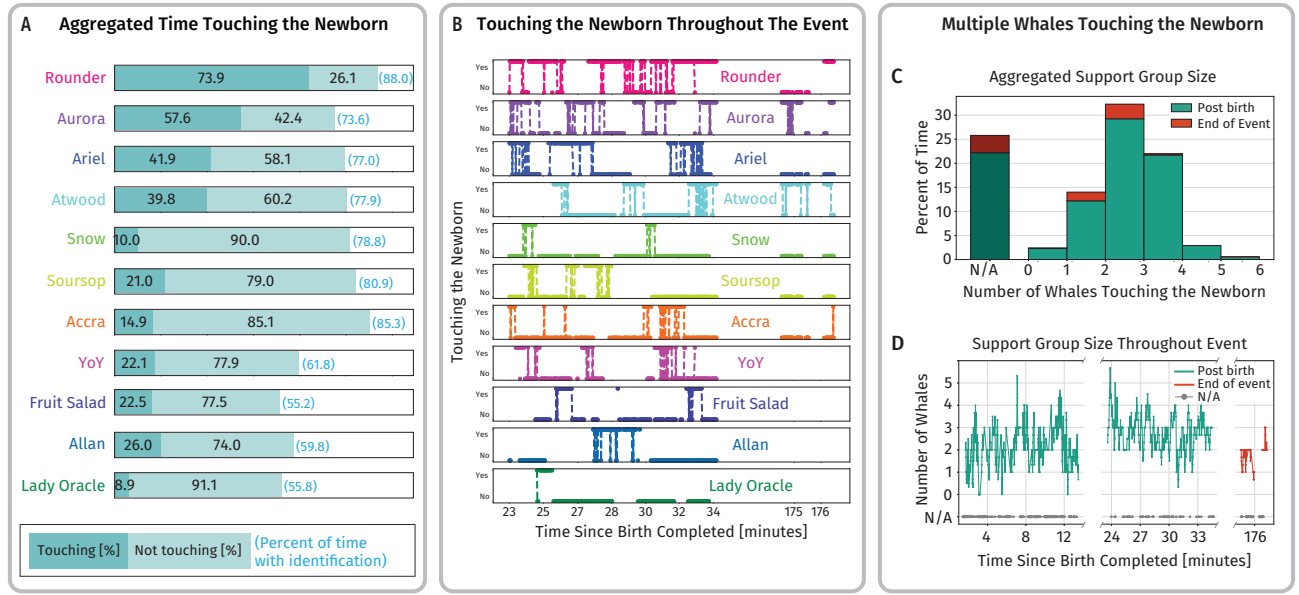


FIG. S8. (A) How often each whale touched the newborn during the *post-birth* and *end-of-event* phases. (B) The presence or absence of physical contact with the newborn over time. (C) A histogram of the number of whales simultaneously touching the newborn, after smoothing with a window duration of 2 s. The N/A category represents when the newborn was not identified. (D) A timeline of the number of whales simultaneously touching the newborn, after smoothing with a window duration of 2 s. The N/A line denotes when the newborn was not identified.

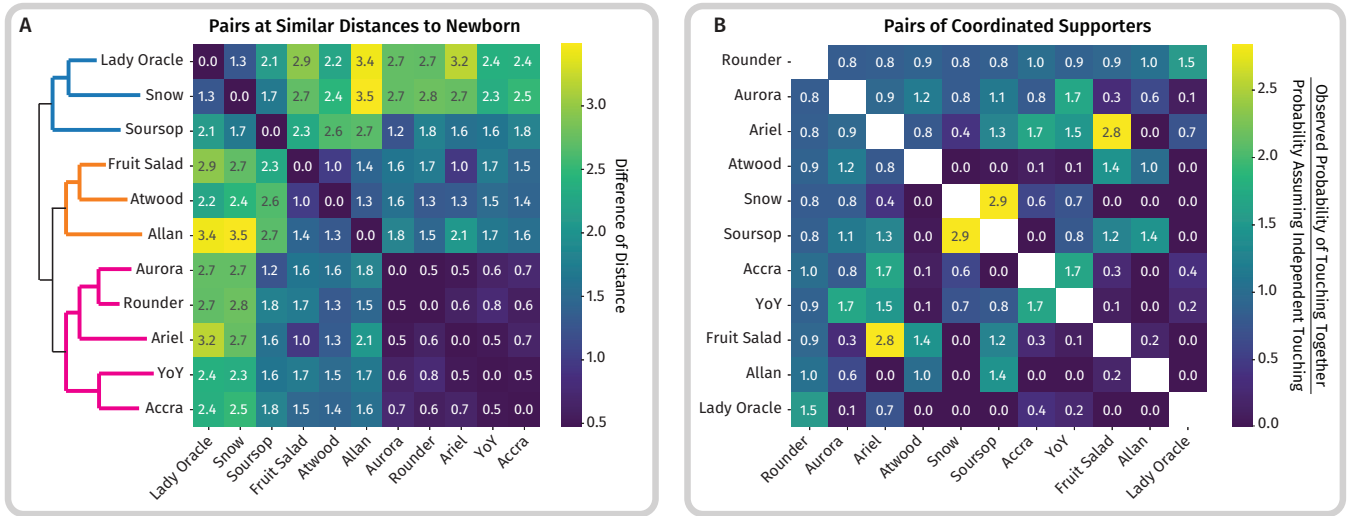


FIG. S9. (A) Spatial proximity patterns: For each pair of whales, we computed the absolute difference in their distances to the newborn in each frame and then averaged these values across all frames. (B) Coordinated touching probability: We compared the observed probability that both whales in a pair were touching the newborn simultaneously to the probability under the assumption of each whale touching the newborn independently.

Patterns of spatial proximity

To compare patterns of proximity with the newborn, we considered all pairs of whales. For each pair, we computed the difference between their distances to the newborn in each frame and then averaged these differences across time. This indicates whether two whales were typically the same distance from the newborn. The results are visualized in Fig. S9A, where three distinct clusters emerge by performing hierarchical clustering on the generated matrix of distance differences. Each cluster exhibited similar spatial behavior and maintained a consistent distance from the newborn within the unit. The first cluster is of particular interest and includes the mother (*Rounder*), the aunt (*Aurora*), the sister (*Accra*), the youngest calf (the *Young-of-Year*), and a juvenile (*Ariel*). The whales within this cluster maintained a similar distance to the newborn, with the differences varying from 0.47 to 0.73 newborn body lengths. This clustering may reflect a cooperative strategy for protecting calves within the unit during interactions

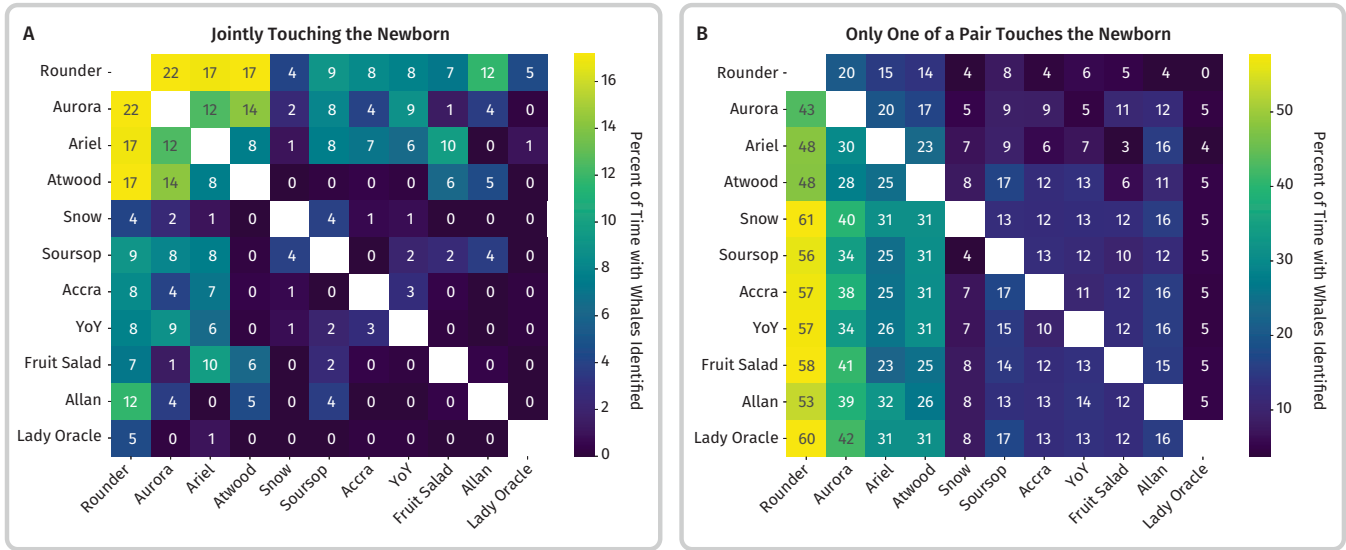


FIG. S10. (A) Percent of time both whales in a pair were simultaneously touching the newborn, calculated over the period when the newborn was identified. (B) Percent of time exactly one whale from the pair was touching the newborn during the same period.

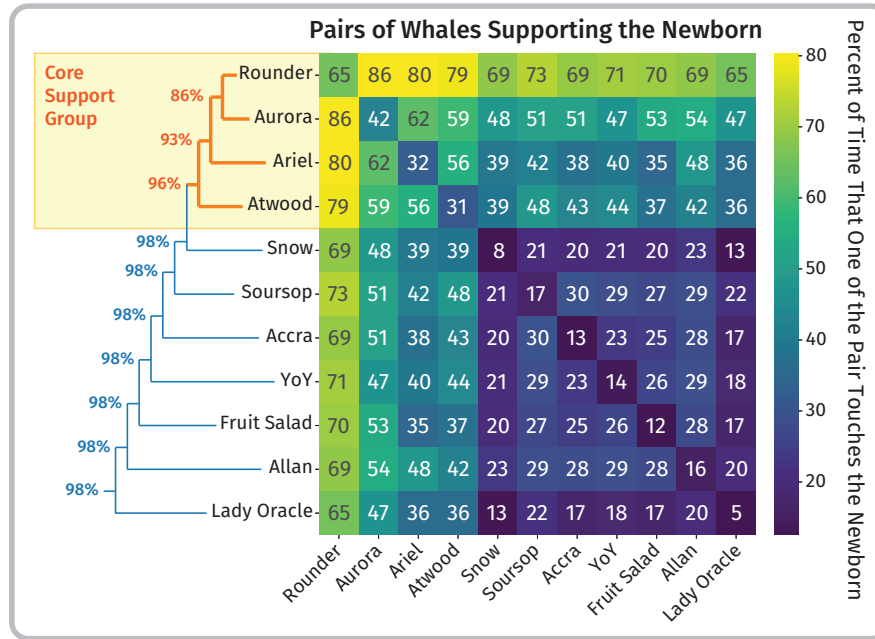


FIG. S11. Percent of time during which at least one whale in a given pair was touching the newborn, calculated over the period when the newborn was identified. The dendrogram on the left of the heatmap illustrates how iteratively considering additional whales increases the percent of time that the group is in contact with the newborn; a core support group of four whales emerges from this analysis, after which there are diminishing returns of increasing support coverage.

with pilot whales, consistent with reports that sperm whales position their calves centrally in defensive formations [2–4]. The second cluster includes *Atwood*, *Fruit Salad*, and *Allan*, while the third cluster consists of *Soursop*, *Snow*, and *Lady Oracle*.

Coordinated versus independent touching of the newborn

For each whale X , we compute $\Pr(X)$ as the proportion of seconds in which whale X is touching the newborn relative to the total seconds where both the newborn and whale X are identified. For every pair of whales X_i and X_j , we compute $\Pr(X_i, X_j)$ as the proportion of seconds where both whales are simultaneously touching the newborn, relative

to the total number of seconds where the three individuals (X_i , X_j , and the newborn) are identified. To assess the likelihood of coordinated touching, we report the ratio $\frac{\Pr(X_i, X_j)}{\Pr(X_i) \cdot \Pr(X_j)}$ for each pair of whales X_i and X_j . This ratio compares the observed probability of whales touching together with the probability expected under the assumption of independent touching.

Results are shown in Fig. S9B, and indicate that many whale pairs have a higher likelihood of touching the newborn together than expected by chance; this suggests potential coordination. Notable pairs include the *Young-of-Year* and *Aurora* (both young animals but from different matriline), *Snow* and *Soursop* (both reproductive females from the other matriline), *Soursop* and *Fruit Salad* (an adult mother-calf pair from the other matriline), *Ariel* and *Fruit Salad* (a granddaughter and grandmother pair from the other matriline), and *Lady Oracle* with *Rounder* (the newborn’s grandmother and mother). Interestingly, these pairs are mostly not direct relatives of the newborn. Note that the mother (*Rounder*) and the aunt (*Aurora*) appear less dominant in this matrix due to their already high probabilities of individual touching, $\Pr(\text{Rounder})$ and $\Pr(\text{Aurora})$.

We also extended this analysis to triplets of whales. For each triplet X_i , X_j , and X_k , we compute $\Pr(X_i, X_j, X_k)$ as the proportion of frames where all three whales are simultaneously touching the newborn, relative to frames where all three individuals are identified. We report the metric $\phi(X_i, X_j, X_k) := \frac{\Pr(X_i, X_j, X_k)}{\Pr(X_i) \Pr(X_j) \Pr(X_k)}$, restricting the analysis to triplets observed together for more than a specified time threshold. Notably, only three groups exhibited a clear signal (with $\phi(X_i, X_j, X_k) > 1.7$): $\phi(\text{Rounder}, \text{Ariel}, \text{Accra}) = 1.79$, $\phi(\text{Aurora}, \text{Ariel}, \text{Soursop}) = 2.35$, and $\phi(\text{Aurora}, \text{Ariel}, \text{Young-of-Year}) = 3.43$. Interestingly, *Ariel* appears in all of these groups, suggesting a central role in coordinated touching behavior, potentially due to her age.

Duration of touching the newborn together, separately, or as part of a group

Fig. S10A presents the percent of time that the newborn was identified during which each pair of whales touched the newborn together. *Aurora* and *Rounder* spend 22% of the time touching the newborn together, *Ariel* and *Rounder* touch together for 17% of the time, and *Rounder* and *Atwood* touch together for 17% as well. Surprisingly, *Rounder* and *Allan* touch the newborn together for 13% of the time – this represents most of the time that *Allan* is touching the newborn.

Fig. S10B shows the number of seconds each whale touched the newborn without the other whale doing the same. *Rounder*, followed by *Aurora*, are the whales who most frequently touched the newborn independently of other whales.

Finally, Fig. S11 shows that a core support group is formed around the newborn. It consists mainly of the mother and her sister, with at least one of them touching the newborn 86% of the time that the newborn was identified. Adding *Ariel* to this group increases the coverage to 93%, and adding *Atwood* increases the coverage to 96%. Notably, results also indicate that each whale has touched the newborn for at least 5% of the time.

-
- [1] S. Gero, J. Gordon, and H. Whitehead, *Calves as social hubs: dynamics of the social network within sperm whale units*, Proceedings of the Royal Society B: Biological Sciences **280**, 20131113 (2013).
 - [2] R. L. Pitman, L. T. Ballance, S. I. Mesnick, and S. J. Chivers, *Killer whale predation on sperm whales: observations and implications*, Marine mammal science **17**, 494–507 (2001).
 - [3] G. L. Gemmell, J. D. McInnes, S. J. Heinrichs, and G. de Silva Wijeyeratne, *Killer whale (orcinus orca) predation on whales in sri lankan waters.*, Aquatic Mammals **41** (2015).
 - [4] H. Whitehead, *Sperm whales: social evolution in the ocean* (University of Chicago press, Chicago, Illinois 60637, USA, 2003).