



ORIGINAL ARTICLE

Phenotypic architecture of sociality and its associated genetic polymorphisms in zebrafish

Claúdia Gonçalves¹ | Kyriacos Kareklas¹  | Magda C. Teles¹ |
Susana A. M. Varela¹ | João Costa¹ | Ricardo B. Leite¹ | Tiago Paixão¹ |
Rui F. Oliveira^{1,2,3} 

¹Integrative Behavioural Biology Laboratory, Gulbenkian Institute of Science, Oeiras, Portugal

²Department of Biosciences, ISPA-Instituto Universitário, Lisbon, Portugal

³Champalimaud Neuroscience Program, Champalimaud Foundation, Lisbon, Portugal

Correspondence

Rui F. Oliveira, Integrative Behavioural Biology Laboratory, Gulbenkian Institute of Science, Oeiras, Portugal.

Email: ruiol@ispa.pt

Funding information

Fundação para a Ciência e a Tecnologia, Grant/Award Numbers: PTDC/BIA-ANM/0810/2014, PTDC/BIA-COM/30627/2017, SFRH/BD/132562/2017, UIDB/04555/2020

Abstract

Sociality relies on motivational and cognitive components that may have evolved independently, or may have been linked by phenotypic correlations driven by a shared selective pressure for increased social competence. Furthermore, these components may be domain-specific or of general-domain across social and non-social contexts. Here, we used zebrafish to test if the motivational and cognitive components of social behavior are phenotypically linked and if they are domain specific or of general domain. The behavioral phenotyping of zebrafish in social and equivalent non-social tests shows that the motivational (preference) and cognitive (memory) components of sociality: (1) are independent from each other, hence not supporting the occurrence of a sociality syndrome; and (2) are phenotypically linked to non-social traits, forming two general behavioral modules, suggesting that sociality traits have been co-opted from general-domain motivational and cognitive traits. Moreover, the study of the association between single nucleotide polymorphisms (SNPs) and each behavioral module further supports this view, since several SNPs from a list of candidate “social” genes, are statistically associated with the motivational, but not with the cognitive, behavioral module. Together, these results support the occurrence of general-domain motivational and cognitive behavioral modules in zebrafish, which have been co-opted for the social domain.

KEYWORDS

phenotypic correlations, SNP, social cognition, social recognition, social tendency, zebrafish

1 | INTRODUCTION

Sociality is ubiquitous among animals, with animal aggregations and the formation of social groups occurring across most animal taxa.¹ Group living has benefits, such as enhanced foraging efficiency and

predator protection, but it also carries costs, such as intra-group competition for resources and increased risk of pathogens' transmission.¹ Therefore, the evolution of the preference to join social groups in animals, which reflects the balance between conflicting approach/withdrawal social motivation, must result from a positive balance of this cost/benefit trade-off. As a result, the tendency to form social groups is expected to vary across species, and within species with

Claúdia Gonçalves and Kyriacos Kareklas share the first authorship of the article.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. Genes, Brain and Behavior published by International Behavioural and Neural Genetics Society and John Wiley & Sons Ltd.

varying ecological conditions, depending on the relative weight of these costs and benefits. Selection for increased sociality (i.e., higher social tendency) should also increase the selective pressure for the evolution of social abilities that optimize this trade-off, namely by increasing the social competence of animals, hence decreasing their costs of group living. Consequently, social living is predicted to drive the evolution of cognitive abilities that enhance social competence (aka social brain hypothesis,^{2,3}), and increased social preference is expected to be correlated with enhanced socially-related cognitive abilities. The social brain hypothesis has been extensively studied using comparative studies of phylogenetically related species with different degrees of sociality with conflicting results.^{2,4} However, this hypothesis can also be tested within species. At this level, three predictions can be generated: (1) there should be phenotypic correlations between measures of social preference and measures of social cognition; if these phenotypic correlations result from a common genetic or physiological mechanism for social preference and social cognition that evolved in response to selection for sociality, then: (2) they should be maintained across different environments; and (3) they should share, at least partially, their genetic basis.

For testing the occurrence of phenotypic correlations between the motivational drive to form social groups and cognitive abilities that enhance fitness in a social environment two elementary components of social behavior can be considered: (1) a measure of approach response towards conspecifics (social tendency) that leads to the formation of social groups; and (2) the cognitive ability to recognize different conspecifics (social recognition) that allows individuals to selectively adjust the expression of their behavior to different individuals they encounter. It should be mentioned that social recognition can range from being more coarse or categorical, as the ability to recognize categories of conspecifics (e.g., male/female, familiar/stranger), to being fine grained, as the ability to recognize specific individuals within the group (e.g., pair mate). As mentioned above, given the role of these two behaviors for sociality, one can predict them to be selected together (i.e., co-evolve) during social evolution, leading to a phenotypic correlation between them. However, these traits can have evolved from other similar traits that have initially evolved in a non-social domain and were co-opted for the social domain when selection for group living increased. For example, social recognition may reflect a general domain cognitive ability, that evolved to allow animals to discriminate different entities, social or not (e.g., edible vs. non-edible food), in the environment, rather than a domain-specific trait selected by sociality.^{5,6} In this case a phenotypic correlation would be expected between social recognition and non-social (e.g., object) recognition. Similarly, social tendency may reflect a general domain response to threat perception in the environment, since cohesiveness in animal aggregations is known to increase with perceived danger (i.e., aka defensive aggregation; e.g., rats: Reference 7; zebrafish: Reference 8). In this case a phenotypic correlation would be expected between social tendency and behavioral measures of anxiety/stress. Thus, the phenotypic architecture of sociality can be characterized by the pattern of phenotypic correlations among these behavioral traits.

The evolution of correlated traits can be explained by two alternative hypotheses, which are not necessarily mutually exclusive: (1) the constraint hypothesis, that postulates the occurrence of shared proximal mechanisms such as a pleiotropic effect of a gene, or a hormone with multiple target tissues; or (2) the adaptive hypothesis, that proposes that positive correlations between traits only occur in environments that favor them, such that selection can break apart maladaptive combinations of traits.⁹ These two hypotheses generate different predictions that can be tested by comparing the patterns of correlated characters across different populations of the same species. The constraint hypothesis predicts traits to be correlated across populations irrespective of ecological conditions, whereas the adaptive hypothesis predicts correlations between traits to vary between populations depending on local conditions. Thus, these two scenarios also have different evolutionary consequences, with the correlated traits acting as evolutionary constraint in the first case and, the correlation being itself an adaptation in the latter. Although, this rationale has been used to study the evolution of behavioral syndromes (aka personality),⁹ to the best of our knowledge, it has not been applied yet to analyze the evolution of correlated social behavior traits.

Finally, it can also be tested if the genetic architecture of correlated traits is shared or not. Given the complexity of social behavior traits, they are expected to be under the influence of multiple genes, with small effects of each of them. In fact, several genes involved in neurotransmission (e.g., dopamine, serotonin¹⁰⁻¹²), neuromodulation (e.g., oxytocin¹³⁻¹⁵) and synaptic plasticity mechanisms (e.g., neuroligins/neurexins¹⁶⁻¹⁹) have been reported to influence social behavior in multiple ecological domains across a wide range of vertebrate taxa. Moreover, these “social” genes are expressed in brain regions that together form an evolutionary conserved social decision-making network in vertebrates.^{20,21} Therefore, the question is to what extent these candidate genes show specific or shared patterns of association with the motivational and cognitive components of sociality discussed above.

Enough variation in both social tendency and social recognition occurs across species and between individuals of the same species, which should allow to test the abovementioned hypotheses. The tendency to associate with conspecifics varies considerably among species, ranging from weakly social species, in which social interactions only occur at specific times (e.g., breeding), to highly social species, in which individuals stay all their lives in close proximity and interacting with others. Similarly, variation in social recognition ability also occurs across species, from basic levels of recognition (e.g., conspecific vs. heterospecific), to increasingly more elaborate ones with high degree of specificity (e.g., kin vs. non-kin; particular individuals).²² Moreover, variation in both social tendency and social recognition also occur within species, both intra- (e.g., with age and life-history stage) and inter-individually.

In this study we aim to characterize the phenotypic architecture of sociality in zebrafish (*Danio rerio*) by characterizing social tendency, social recognition and object recognition across multiple laboratory zebrafish populations that have evolved separately in captivity for multiple generations and by characterizing the genetic polymorphisms of candidate “social” genes associated with these behavioral traits. In

zebrafish, isogenic lines are not viable due to inbreeding depression.²³ Hence, laboratory zebrafish populations differ from those of other model organisms in that they are recurrently outcrossed to maintain diversity.²⁴ As a result, laboratory zebrafish populations contain significant but varying levels of genetic diversity.^{25,26} In parallel, zebrafish lines (e.g., AB, TU and WIK) have already been shown to vary in many behaviors, some of them interlinked, including locomotor activity, anxiety traits, stress reactivity, learning abilities and shoaling.^{27–38} The paralleled variation in genetic diversity^{25,26} and several behavioral phenotypes, provides the rationale that constitutive genetic variation may contribute to the observed behavioral variability.

Here we specifically aim to test: (1) if there is an association between social tendency and social recognition; (2) if social and non-social cognitive abilities (i.e., social vs. object recognition) are independent from each other, or if they co-vary supporting a general domain factor; (3) if there is an association between social tendency and anxiety trait; (4) if the phenotypic correlations found are fixed or vary across lines (populations), in order to test the constraint versus adaptive hypothesis; (5) to what extent the genetic architecture of each of these behavioral traits is shared or not, which would provide evidence for genetic pleiotropic effects underlying a putative sociality syndrome. For the latter, we have assessed the association between known single nucleotide polymorphisms (SNPs) in zebrafish for a set of candidate “social” genes (see Section 2 for details) and each behavioral trait.

2 | METHODS

2.1 | Zebrafish lines and housing conditions

Zebrafish were raised in the Fish Facility of the Gulbenkian Institute of Science under laboratory conditions. A total of 164 experimentally naive adult zebrafish from different wild type lines of both sexes, aged 6–8 months, were used in this study as focal subjects (AB: M = 8, F = 14; TU: M = 9, F = 12; WIK: M = 12, F = 4; TL: M = 13, F = 10; LEO: M = 7, F = 10; 5D M = 32, F = 33). Focal fish were raised and housed separately from fish used as stimuli to prevent effects of prior familiarity. Fish used as stimuli were of the same line as the focal fish. Housing was in groups of 35 fish kept in 3.5 L aquaria of a recirculating system (ZebraTec, 93 Tecniplast), with water parameters set at 27–28°C, 7.5 ± 0.2 pH, ~900 µSm and <0.2 ppm nitrites, <50 ppm nitrates and 0.01–0.1 ppm ammonia. Daily photoperiods were alternated between 14 h light and 10 h dark and feeding occurred twice-daily and included a combination of live (*Paramecium caudatum*; *Artemia salina*) and processed dry food (GEMMA Micro).

2.2 | Experimental setup and procedures

The behavior of each experimental fish was assessed in four different tests: (1) a shoal preference test to measure social tendency; two one-trial recognition tests using either objects (2) or conspecifics (3) as stimuli to measure non-social and social recognition/exploration,

respectively; and (4) an open-field test to measure the anxiety trait (Figure 1A–E; see Supplementary material for details). All tests occurred during the light period between 09:00 and 19:00, before which fish were kept overnight in an aquarium with individual compartments for identification purposes.

Behavior during tests was recorded using black and white mini surveillance cameras (Henelec 300B) suspended above the experimental tank. Videos were analyzed using a commercial video-tracking software (EthoVision XT, Version 11.5, Noldus Information Technology) and behavioral measures were extracted from each test. Regions of interest (ROI) marked were kept at an average body length distance from the target location (gray regions in Figure 1A,B). Social tendency during the shoal preference test was quantified by the proportion of time in ROIs spent near the shoal (Figure 1C), social and non-social discrimination during the conspecific and object recognition tests was measured by the proportion of time in ROIs spent near the preferred stimulus (familiar or novel; Figure 1D,E, respectively), while the overall time spent in ROIs near both stimuli was used as a measure of exploration. Anxiety in the open field test is typically exhibited by thigmotaxis (i.e., the propensity to avoid exposed areas), which was measured as the proportion of time spent within the ROI near the periphery following first entry (to control for any initial freezing in the center), while the average distance (in cm) from the wall was used to quantify the edge or wall orienting tendency associated with fear-induced thigmotaxis.³⁹

2.3 | Genetic polymorphisms analysis

At the end of the behavioral phenotyping, animals were anesthetized by immersion into an ethyl 3-aminobenzoate methanesulfonate salt solution (MS222) 100–200 mg/L, a fin clip collected from the caudal fin of each experimental fish, and preserved in a digestion mix (PK, 10 mg/ml, Lysis solution [Fermentas #K0512], TE buffer) until further processing. Subsequently, DNA was extracted from preserved fin clips using DNA Extraction kit (Fermentas #K0512) with some adjustments to the protocol provided by the manufacturer (see Supplementary material for details).

We built a list of candidate genes to test their association with the behavior traits, based on evidence from the literature for their involvement in the regulation of social behavior. This gene list included genes for: neurotransmitter systems (e.g., dopamine, serotonin), neuro-modulators (e.g., oxytocin, AVT and NPY), neuroplasticity (e.g., bdnf, neurexins and neuroligins), and genes linked to autism (e.g., shank3a). A total of 139 SNPs in the genes of interest were successfully sequenced (see Supplementary material and Table S1 for details), but we had to remove 7 for lack of variation between the 164 tested zebrafish.

2.4 | Statistical analysis

One-sample *t*-tests ($\mu \neq 0.5$ vs. >0.5) were used to test if the scores of social tendency, object discrimination and social discrimination were significantly different from chance levels for each sex and for each line.

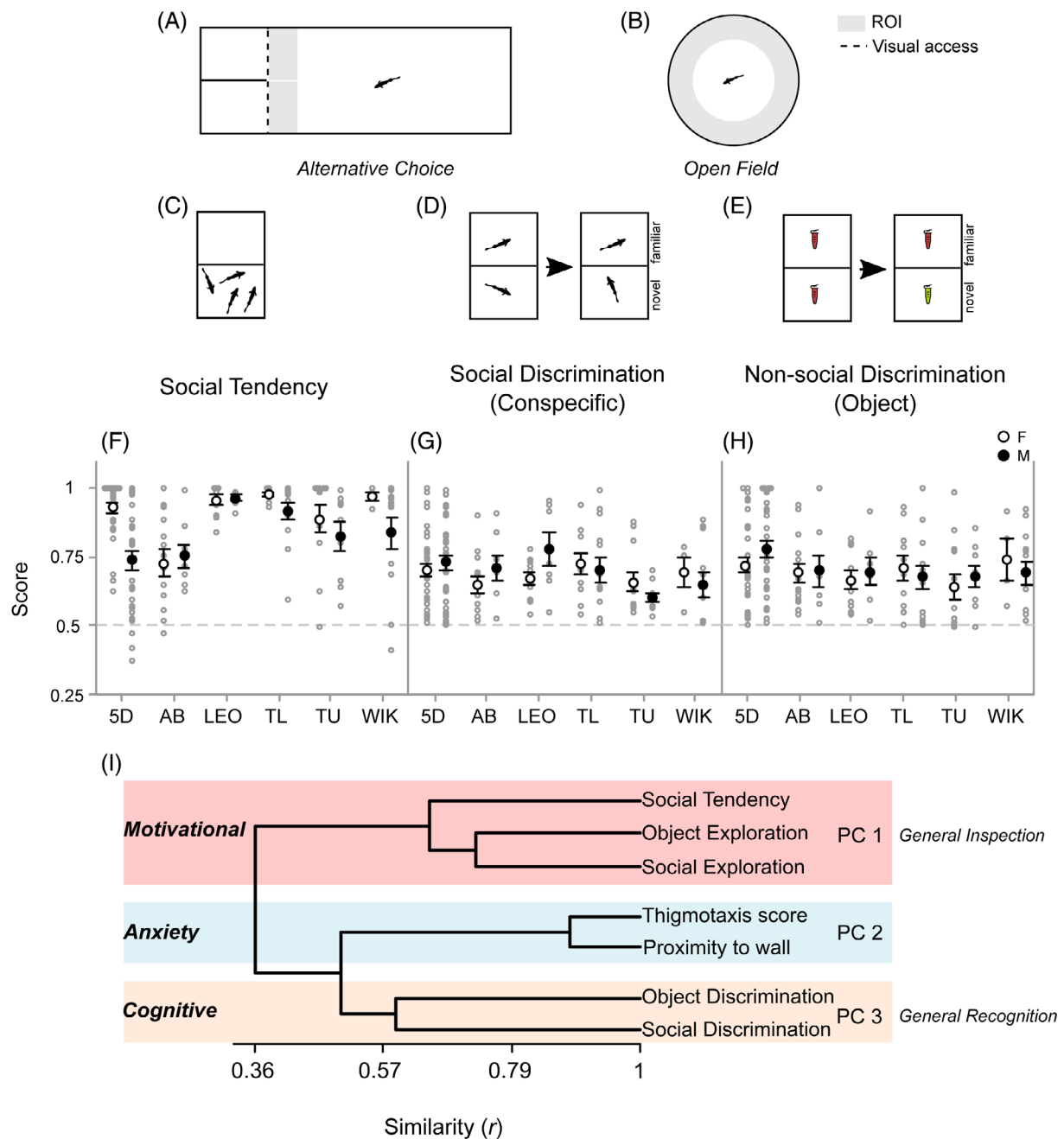


FIGURE 1 Social and associated behaviors in zebrafish. (A) Across lines, a two alternative-choice set-up was used to measure social preference and recognition abilities and (B) an open field test for measuring anxiety-driven thigmotaxis towards the periphery and edge-orienting. Regions of interest (ROI) were set within 1 standard body-length from target locations or stimuli. (C) Social tendency was measured by interaction preferences towards a shoal. Social (D) and non-social (E) discrimination tests were comprised of two phases: an acquisition phase, in which the focal fish was exposed to two unfamiliar items (two fish or two objects, respectively) followed (as indicated by arrow in D and E) by a probe-test phase, in which the focal fish had to discriminate between one of the previously seen items (fish or object) and a novel one; recognition in both the social (D) and non-social (E) context were measured by the ability to discriminate between a familiar and a novel stimulus. Males (full circles) and females (open circles) of all lines (5D, AB, LEO, TL, TU, Wik) exhibited above chance (dashed line) preference for shoal over an empty tank (social tendency, F) and discrimination between a novel and familiar stimulus in both a social (conspecific; G) and non-social (object; H) context (bars indicate 95% CI). Behavioral measures exhibited different degrees of correlation (r), illustrated in the cladogram as degrees of association (I), based on which factor analysis revealed three principal components (PC): PC1 aggregates social tendency and social and object exploration corresponding to a motivational component of sociality; PC2 aggregates thigmotaxis and (i.e., proportion time in periphery) and edge-orienting (distance to wall) measured in the open field test, corresponding to an anxiety component; PC3 aggregates object and social discrimination, corresponding to a general-domain cognitive component

Next, we extracted behavioral modules that aggregate correlated behaviors by carrying out a principal component analysis (PCA) followed by varimax rotation, based on the correlation matrix of all behavioral measures (social tendency, social discrimination, social exploration, object discrimination, object exploration, thigmotaxis and edge-orienting). These analyses were carried out in the statistical software Minitab® version 17 (Minitab Inc., State Collage, PA). The remaining analyses described below were carried out in the statistical software R, version 4.0.4⁴⁰ (see Supplementary material for a list of packages used).

To test if the behavioral modules are differently related with each other in each zebrafish line, we used the quadratic assignment procedure (QAP) correlation test with 5000 permutations,⁴¹ to assess the association between any two correlation matrices between different zebrafish lines on UCINET 6.⁴² Given that the null hypothesis of the QAP test is that there is no association between matrices, a significant *p*-value indicates that the correlation matrices are similar.

To check whether the genetic distances between subjects are structured by line or represent a uniform population, we computed a genetic distance (i.e., jaccard distance) matrix between all subjects (using their genetic data from the list of 132 SNPs), based on which we performed a hierarchical clustering with complete-linkage.

To assess the associations between genetic polymorphisms and behavior, we tested each of the 132 SNPs independently against each behavioral phenotype (the 7 behaviors and 3 PC scores). We did not include three zebrafish subjects in this analysis because their sample call rate was below 5%, meaning they lack genetic information for most SNPs. For the behaviors that followed a linear distribution (general inspection, general recognition, anxiety and edge-orienting) we used linear models (LM). For the behaviors that were proportions (social tendency, social discrimination, social exploration, object discrimination, object exploration and thigmotaxis), we used generalized linear models (GLM) with beta regression implemented. In all models, the behaviors were the response variables, SNP was the explanatory variable and line was a co-variate. Because we run 132 independent tests for each SNP, we corrected the *p*-values with the false discovery rate (FDR) adjustment method.

2.5 | Ethics

All experimental procedures were reviewed by the institutional internal Ethics Committee at the Gulbenkian Institute of Science and approved by the National Veterinary Authority (Direção Geral de Alimentação e Veterinária, Portugal; permit number 0421/000/000/2017).

3 | RESULTS

3.1 | Phenotypic architecture of sociality in zebrafish

Scores of social tendency (i.e., preference for shoal over empty tank), as well as object and conspecific discrimination scores (i.e., preference

between a novel and a familiar stimulus) were all significantly different than chance for individuals of both sexes and for all lines tested (one-sample *t*-test: $\mu \neq 0.5$, $p < 0.001$; see Table S2; Figure 1F–H), indicating that social affiliation and social and object recognition abilities are present in males and females across zebrafish lines.

The PCA used to assess the phenotypic architecture of sociality, based on the correlation matrix between measures extracted from the four separate tests of social and associated behaviors (sampling adequacy: KMO > 0.5; sphericity: Bartlett's $\chi^2_{21} = 253.76$, $p < 0.001$; determinacy of multicollinearity: $\rho = 0.754$), identified three principal components (PC) with eigenvalues ≥ 1 (Figure 1I and Table 1). PC1 shows a strong loading of social tendency measured in the social preference test and of social and object exploration measured in the social and object discrimination tests, respectively, suggesting the occurrence of a general inspection behavioral module that is expressed both in social and non-social contexts. PC2 shows a strong loading of thigmotaxis and edge-orienting measured in the open-field test, corresponding to an anxiety behavioral module. Finally, PC3 shows a strong loading of object and social discrimination, measured in the object and social discrimination tests, respectively, suggesting the occurrence of a general recognition behavioral module that is expressed both in social and non-social contexts.

To test if the behavioral modules described above (i.e., General inspection, General recognition and Anxiety) can evolve differently from each other in each zebrafish line—which represent different laboratory populations established by different wild type founders and that have evolved independently from each other in somewhat similar lab conditions—we computed correlation matrices between individual scores for each module (varimax rotated PC scores) for each of the different zebrafish lines. We then used the QAP correlation test to compare the correlation matrices of the different lines. The results identified a single significant negative correlation ($r = -0.9988$, $p = 0.0002$) between 5D and WIK correlation matrices. Thus, none of the correlation matrices were similar between each other (Figure 2), rejecting the constraints hypothesis, that predicts similar patterns of phenotypic correlations across different populations.

Although it was not the central question of this study, the occurrence of sex and line differences in the expression of the behavioral modules identified above can be informative when choosing lines to run specific behavioral tests in zebrafish, and we report them as Supplementary material (Supplementary results and Figure S1).

3.2 | Genetic polymorphisms associated with behavioral modules

To assess if the different behavioral modules identified above were linked by a shared genetic architecture, we have investigated the association between a set of genetic polymorphisms (SNPs) in a list of candidate “social” genes and each of the measured behaviors and PC behavioral modules. Given the fact that we have phenotyped individuals from six different wild type lines, we checked for structured genetic variation by computing the genetic distance between the

TABLE 1 Loadings extracted by the varimax rotation of principal components from the correlation matrix of behaviors across tests, for zebrafish of all lines

Test	Behavior	Parameters	Principal component loadings ^a			
			PC 1	PC 2	PC 3	Comm. ^b
			General Inspection	Anxiety	General Recognition	
SP	Social tendency	Proportion total ROI time spent with shoal	0.725	−0.033	0.180	0.559
SR	Social discrimination	Proportion total ROI time spent with preferred conspecific	0.113	−0.031	0.852	0.739
	Social exploration	Proportion test time spent in ROI of both conspecifics	0.762	0.085	−0.224	0.639
OR	Non-social discrimination	Proportion total ROI time spent with preferred object	−0.363	0.139	0.619	0.535
	Non-social exploration	Proportion test time spent in ROI of both objects	0.754	−0.042	−0.111	0.583
OF	Thigmotaxis	Proportion time spent within ROI of periphery, after first entry (control for initial freezing in center)	0.107	0.932	0.084	0.888
	Edge-orienting	Average distance from wall (cm)	0.112	− 0.941	0.009	0.897
Eigenvalue ^c			1.845	1.784	1.211	4.840
% Variance explained			0.264	0.255	0.173	0.691

Note: Bold type indicates the strongest contributors (coefficient >0.5) to each principal component (PC).

Abbreviation: ROI, regions of interest.

^aCorrelation between components and variable values.

^bCommunalities: Proportion of variable variance explained by all principal components.

^cVariance of transformed data used for each principal component.

phenotyped individuals for the SNPs under study. We found that genetic variation for the SNPs of interest is highly structured with individuals from the same wild type lines clustering together (Figure 3A). Therefore, we have used the line as a covariate in the model that assessed the association between each SNP and each of the behavioral modules.

Out of the 132 SNPs that showed variation in our sampled individuals, 53 (which mapped to 28 genes) were significantly associated with General Inspection, none with General Recognition and 8 (which mapped to 6 genes) with Anxiety (Table 2). Regarding the 3 behaviors that loaded to the General Inspection behavioral module, 6 SNPs (mapping to 6 genes) were associated with social tendency, 11 (mapping to 10 genes) with social exploration, and 3 (mapping to 3 genes) with object exploration. Of these 20 SNPs associated with these behaviors that load to General Inspection, only one (mapping to the serotonin receptor gene 5HTR 2c12) is not also associated with General Inspection (Figure 3B; Table 2). Moreover, of the 29 SNPs associated with General inspection, 16 are also associated at least with one of the behaviors that constitutes these behavioral module (Figure 3B; Table 2). However, there is a reduced overlap between the SNPs associated with these different behaviors: only one SNP affects both social tendency and social exploration (matching the gene 5HTR-1aa), and only another SNP affects both social exploration and object exploration (matching the gene 5HTR-2c11) (Figure 3B; Table 2).

The SNPs associated with the General Inspection behavioral module are widely distributed across the zebrafish genome being absent

only from chromosomes 11, 12, 19, 21 and 23 (Figure 3C). However, one can find SNPs associated with behaviors that load to General Inspection module in some of these chromosomes; SNPs associated with social exploration in chromosome 11, 19 and 21; and SNPs associated with social tendency and with object exploration in chromosome 21 (Figure 3C).

The list of SNPs associated with the General Inspection module include genes involved in neurotransmission (e.g., serotonin and dopamine receptors), neuromodulation (e.g., NPY, oxytocin), synaptic plasticity (e.g., neuroligins, neurexins) and epigenetic marking (e.g., methyl CpG binding protein 2) (see Figure 4 for arbitrarily selected illustrative examples).

4 | DISCUSSION

In this study we have characterized the phenotypic architecture of sociality in zebrafish. We have behaviorally phenotyped males and females of six different wild type laboratory lines in four behavioral tests (social tendency, social and object discrimination and open-field) and showed that social tendency (i.e., preference to associate with conspecifics) and the ability to discriminate between conspecifics (social recognition) is present in both sexes of all lines tested. A factor analysis identified three main behavioral modules: (1) general inspection, which includes social tendency measured in the social preference test and social and object exploration, measured in the social and

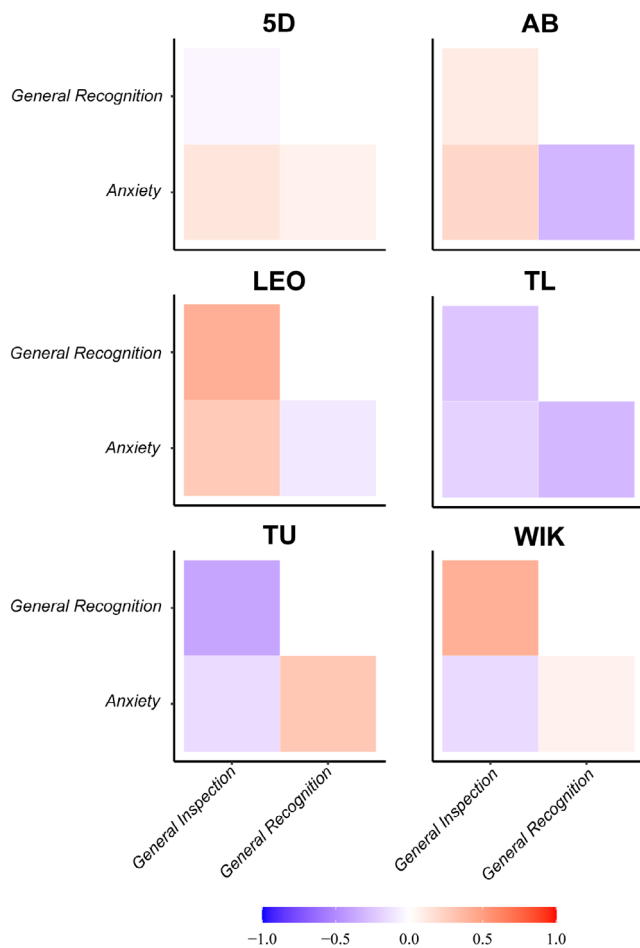


FIGURE 2 Phenotypic correlation matrices. Phenotypic correlation matrices for the behavioral components of General Inspection, General Recognition and Anxiety across six different zebrafish laboratory lines (i.e., populations). Color code represents correlation (r) values

object discrimination tests, respectively; (2) general recognition, which includes social and object discrimination, measured in the social and object discrimination tests, respectively; and (3) anxiety, which include the behavioral measures of thigmotaxis and edge-orienting taken in the open field test. Therefore, the motivational (social tendency) and cognitive (social recognition) aspects of sociality are not phenotypically correlated, a result that does not support the occurrence of a sociality syndrome, which could be predicted by shared selective pressures on these two traits for the evolution of sociality. Moreover, the fact that both social tendency and social recognition are phenotypically correlated with similar non-social behaviors (i.e., object exploration and object recognition, respectively), integrating two general-domain behavioral modules (general inspection and general recognition), supports the hypothesis that these behaviors are not domain specific and have been evolutionarily co-opted from general-domain motivational and cognitive traits. These results agree with a recent study showing that in zebrafish both social recognition and object recognition, but not social tendency, are oxytocin-dependent. Thus, both studies suggest a common proximate mechanism indicative

of a general-domain cognitive trait (i.e., social and social memory).⁴³ Finally, it is worth noting that even though sociality has been proposed to be promoted by predator pressure as a defensive mechanism,⁴⁴ anxiety forms an independent behavioral module from those where social traits are included.

Even with the motivational and the cognitive components of sociality being part of two different behavioral modules, a shared selective pressure on both for the enhancement of social competence could result in a physiological linkage between the two behavioral modules; for example, due to the evolution of a common neuromodulator that phenotypically integrates the independent neural mechanisms underlying general inspection and general recognition. In fact, even though that social affiliation and social memory have been shown to rely on separate neural circuitry, some neuromodulators, such as oxytocin have been shown to regulate both mechanisms,^{45,46} opening the possibility for the evolution of physiological constraints that phenotypically link these two domains. We tested the constraint hypothesis, which predicts traits to be correlated across populations irrespective of ecological conditions,^{9,47} in our data set by comparing the matrices of phenotypic correlations among the three behavioral modules extracted from the factor analysis across the six wild type lines used in this study. Given that these wild type laboratory lines have been established independently from different founders collected in the wild and have been evolving independently from each other in similar stochastic lab environments, they can be seen as independent representative populations of this species (despite living in artificial environments). Contrary to the prediction of the constraint hypothesis, the phenotypic correlation matrices were not similar between any pair of zebrafish laboratory lines studied. In fact, there was only one significant QAP correlation between the 5D and Wik matrices, but it was a negative correlation suggesting an asymmetric structure of the matrices. It should be noted that despite the fact that our data rejects the occurrence of constraints, it does not provide evidence for the alternative adaptive hypothesis, that proposes that positive correlations between traits are the result of historical selection favoring particular trait combinations (i.e., selection-induced linkage disequilibrium,^{48,49} such that the evolution of different combinations between the different behavioral modules is not physiologically or genetically linked), given the similarities in lab environments across all six lines tested. Thus, the different combinations of positive phenotypic correlations across lines, which evolved in similar lab environments, must represent stochastic variation.

The study of the association between a set of genetic polymorphisms (SNPs), in candidate genes that have been implicated in social behavior in vertebrates (“social genes”), and the behavioral modules that emerged from our factor analysis indicates that only the general inspection (motivational) module is associated with SNPs in the “social genes,” further supporting the lack of genetic linkage between this module and the general recognition (cognitive) module. Thus, the “social genes” studied here seem to be associated with a general domain motivational component of social behavior, rather than with a general domain cognitive component, which probably relies on memory related genes not included in our “social genes” list. Moreover,

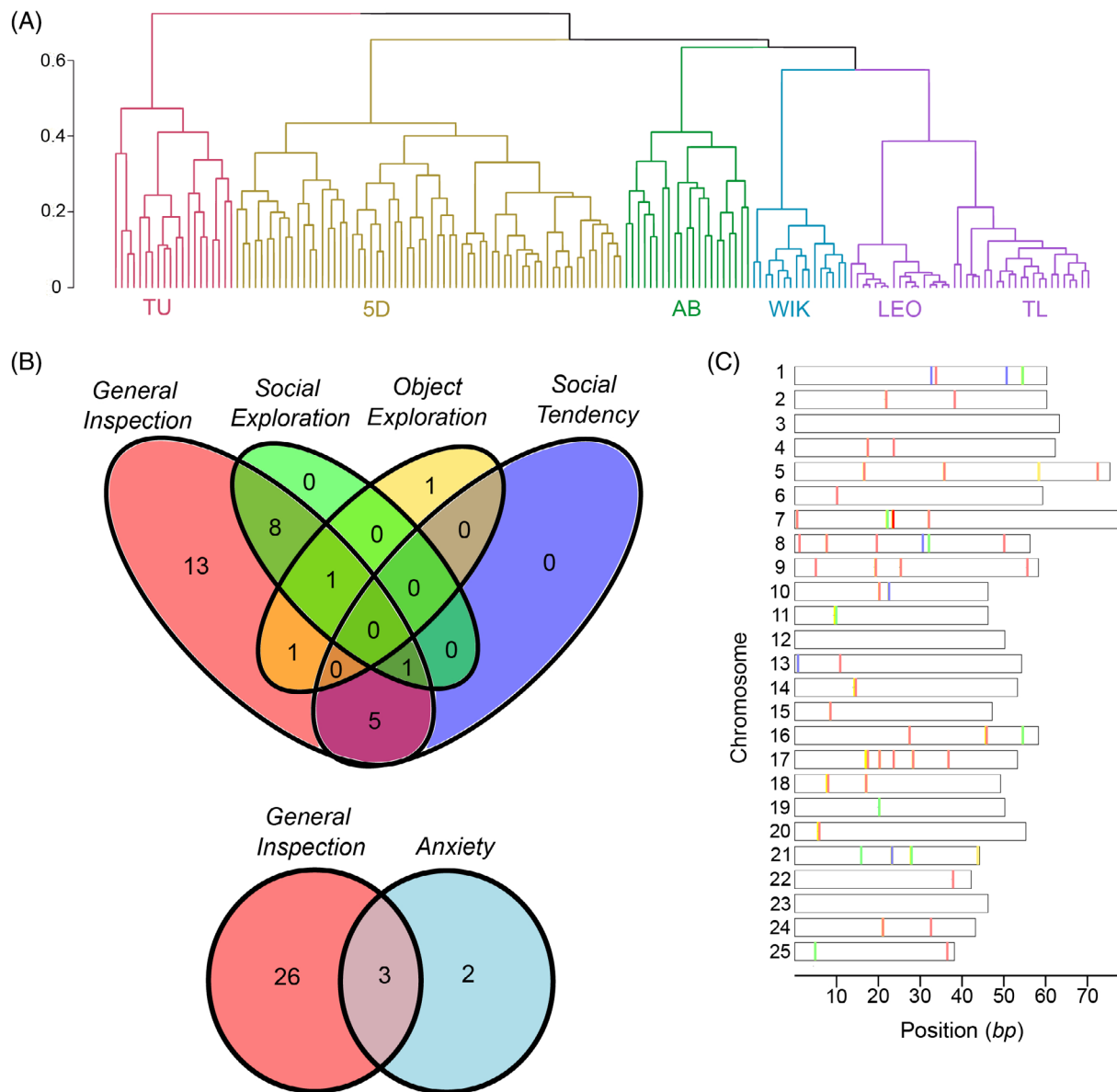


FIGURE 3 Genetic clustering and behavioral associations. (A) Hierarchical clustering of genetic distances (Jaccard distance) between the sampled individuals indicates the occurrence of five major clusters that overall match the six wild type lines used (pink cluster: TU; gold cluster: 5D; green cluster: AB; blue cluster: Wik), with LEO and TL included in the purple cluster but subsequently segregated from each other in two lower order clusters. (B) Venn diagrams representing the number of SNPs the General Inspection component shares with its constitutive behaviors (social tendency, social exploration and object exploration) and the Anxiety component. (C) Chromosome mapping of the SNPs that are significantly associated with the General Inspection component and its constitutive behaviors, following the color code used by the Venn diagrams, and with the position of each SNP on each chromosome is given in bp. SNPs, single nucleotide polymorphisms

our results also indicate a low overlap in the genetic polymorphisms association (3 out of 29 SNPs) between the general inspection and the anxiety modules, which suggests that despite these two behavioral modules relying on motivational mechanisms they have significantly different genetic architectures.

Interestingly, all except one of the genetic polymorphisms (5HTR 2c12) associated with the three behaviors that loaded to the General Inspection behavioral module, are also associated with this behavioral module indicating an agreement between phenotypic (i.e., behavioral

correlations) and the genetic (i.e., genetic polymorphisms) data supporting the occurrence of this behavioral module. The genetic polymorphisms associated with these behaviors include neurotransmitter and neuromodulator systems known to modulate motivational states, such as serotonergic (social tendency is associated with, 5HTR1aa, 5HTR3a and social exploration with 5HTR-1aa, 5HTR-2c11) and neuropeptidergic pathways (social exploration is associated with GnRH2 and NPY), as well as genes involved in synaptic plasticity, such as the neuroligin/neurexin system (social tendency is associated with

TABLE 2 Lists of genes with SNPs associated with the behavioral modules General Inspection (and its contributing behaviors) and anxiety

Gene name	General Inspection	Social tendency	Social exploration	Object exploration	Anxiety
5HTR-1aa	rs180146258, rs180146259	rs180146258	rs180146259		
5HTR-2cl1	rs180151790		rs180151790	rs180151790	
5HTR-2cl2				rs180074553	
5HTR-3a	rs180073160, rs180073162, rs180073164, rs180168240	rs180168240			
5HTR-3b	rs180168236, rs180168238				
5HTR-7b	rs180131627, rs180131628				
5HTR-7c	rs180162109, rs40616624, rs40650859				
Chd7					rs180043369
Cyp19a1b	rs180038734		rs180038734		
D2b	rs180032799, rs180107813, rs180173350			rs180173350	
D3	rs180060870, rs180060872				
Dkk2	rs180052655	rs180052655			
GnRH-2	rs40618151		rs40618151		
itprid1	rs180062152				
MECP2	rs180034118, rs180034123		rs180034118		
Nlgn1	rs180124055, rs180124079		rs180124079		
Nlgn2a	rs180151551, rs180151563		rs180151551		rs180151551
Nlgn2b	rs180131390, rs180109431	rs180131390			
Nlgn4xa	rs180050066	rs180050066			
Npas1	rs180107067				
NPY	rs180080888		rs180080888		
Nr4a2a					rs180134986
Nr4a3	rs180101713				
Nrxn1b	rs180110916, rs180110942	rs180110942			
Nrxn2a	rs180168558		rs180168558		
Nrxn2b	rs180174009				
Nrxn3b					rs180149774
oxytocin	rs180034306, rs180034305				rs180034306, rs180034305
Shank3a	rs179558694, rs180084393, rs180084400, rs180084434				
Syngap1b	rs180104498				
TryptophanH2	rs180086458, rs180086462				
Tsc2	rs180053194, rs180053196, rs180053204, rs180053446, rs180055121		rs180053204, rs180053446		
TyrosineH2					rs180036401, rs180036402

Note: SNP names are provided for each cell. Gene name abbreviations: 5HTR = serotonin receptor; D = dopamine receptor; Cyp19a1b = cytochrome P450, family 19, subfamily A, polypeptide 1b; Nrxn = neurexin; Nlgn = neuroligin; Npas1 = Neuronal PAS Domain Protein 1; NPY = neuropeptide Y; Nr4a3 = nuclear receptor subfamily 4, group A, member 3; Nr4a2a = nuclear receptor subfamily 4, group A, member 2a; Dkk2 = dickkopf WNT signaling pathway inhibitor 2; itprid1 = ITPR interacting domain containing 1; MECP2 = methyl CpG binding protein 2; Syngap1b = synaptic Ras GTPase activating protein 1b; Tsc2 = TSC complex subunit 2; Chd7 = chromodomain helicase DNA binding protein 7.

Abbreviation: SNPs, single nucleotide polymorphisms.

Nrxn1b, Nlgn2b and Nlgn4xa, and social exploration with Nlgn1, Nlgn2a and Nrxn2a) and epigenetic marking (social exploration is associated with the methyl CpG binding protein 2).

On the other hand, the genetic polymorphisms associated with object exploration include less “social genes” (only 3), which are restricted to the serotonergic and dopaminergic neurotransmitter

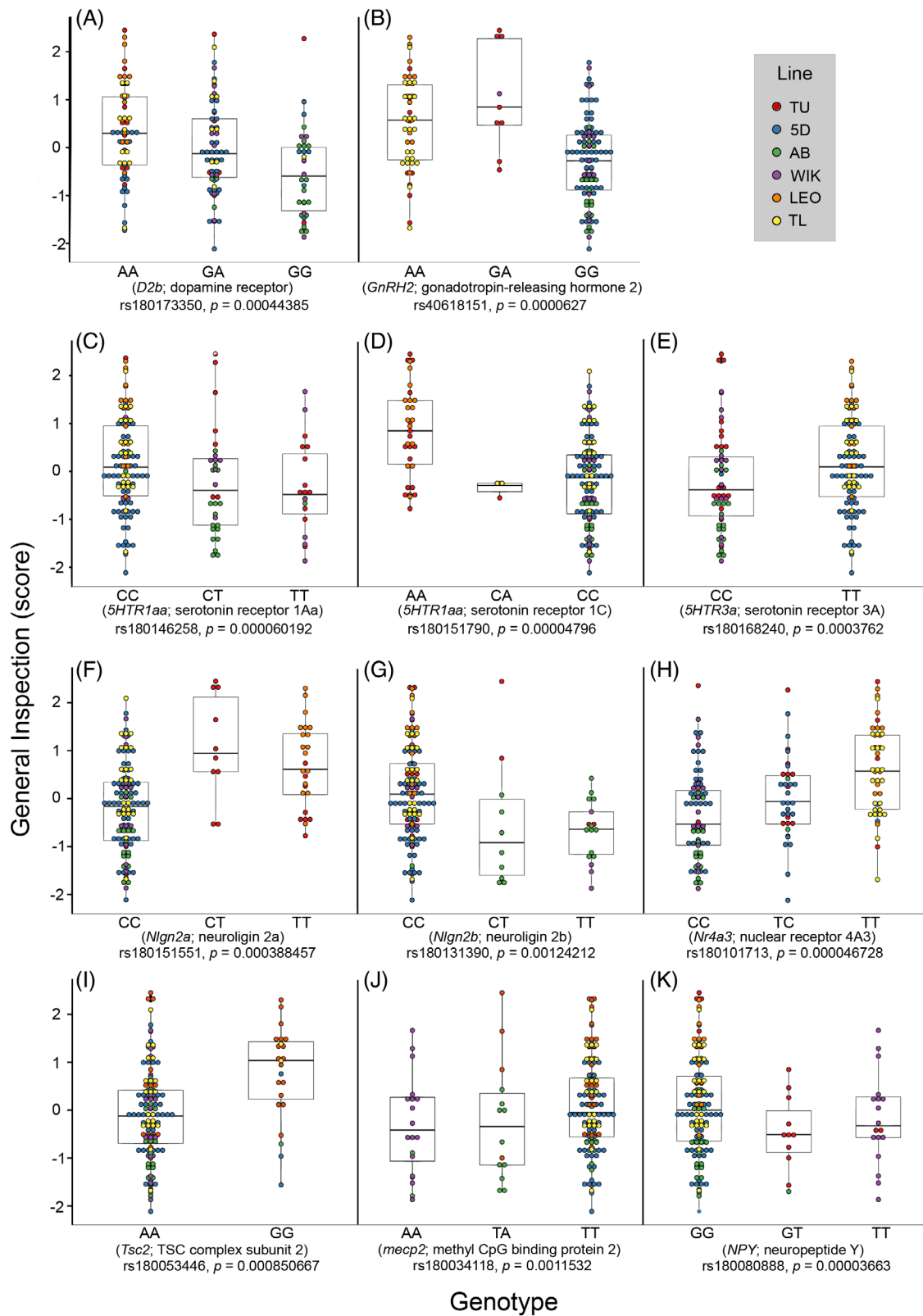


FIGURE 4 Illustrative examples of SNPs associated with the General Inspection component. (A) *D2b*; (B) *GnRH2*; (C) *5HTR1aa*; (D) *5HTR2d1*; (E) *5HTR3a*; (F) *Nlgn2a*; (G) *Nlgn2b*; (H) *Nr4a3*; (I) *Tsc2*; (J) *MECP2*; (K) *NPY*. Individuals of the different lines are represented by different colors according to color code indicated in the figure legend. SNPs, single nucleotide polymorphisms

pathways (5HTR-2c1, 5HTR-2c2 and D2b). Thus, even within a behavioral module it is possible to observe a significant partitioning of the genetic associations with the different component traits of that module. This conclusion is further supported by the fact that there are only two SNPs, in the same gene (5HTR-1aa), that are associated both with social tendency and social exploration, and only another SNP in one gene (5HTR-2c1) associated both with social and object exploration. The current availability of CRISPR/Cas9 mutagenesis in zebrafish⁵⁰ will allow in the future to test the functional role of the SNPs found in these study to be associated with specific components of sociality, on each of the behavioral traits.

The SNPs associated with the General Inspection behavioral module are distributed across 20 of the 25 chromosomes that constitute the zebrafish genome, being absent only from chromosomes 11, 12, 19, 21 and 23. However, one can find SNPs associated with behaviors that load to the general inspection module in chromosomes that do not contain SNPs associated with the behavioral module itself (e.g., SNPs associated with social exploration in chromosome 11, 19 and 21, and the SNPs associated with social tendency and with object exploration in chromosome 21). In a previous study that aimed to identify quantitative trait loci (QTL) in zebrafish for behavioral and morphological traits, QTLs for social tendency have been identified when using one of the two statistical methods used (genetic algorithm mapping vs. interval mapping) in chromosomes 18 and 24.⁵¹ In our study, variation in social tendency is associated with SNPs located in chromosomes 1(#2), 8, 10, 13 and 21. However, the General Inspection module, where social tendency is included, has associated SNPs on chromosomes 18 and 24. Hence, this mismatch between the QTL results and our results presented here can be due either to a false detection of these QTLs by the genetic algorithm mapping method, given the lack of support from the interval mapping method in the previous study, which led the authors not to claim these QTLs themselves⁵¹; or to an indirect association through the link between social tendency and the general inspection module. Either way, our results show that the SNPs associated with both the general inspection module and the behaviors that constitute this module are widespread across the genome, supporting a many gene (each with small effects) genetic architecture for these traits.

ACKNOWLEDGMENTS

The authors thank Ibukun Akinrinade for extracting the DNA samples for the SNP analysis. This study was funded by Fundação para a Ciência e a Tecnologia (FCT, Portugal) through grants PTDC/BIA-ANM/0810/2014 and PTDC/BIA-COM/30627/2017 awarded to RFO and a PhD fellowship awarded to C.G. (SFRH/BD/132562/2017). All the genetic work was run at the Genomics Unit of Instituto Gulbenkian de Ciência which is funded by a FCT R&D unit grant (UIDB/04555/2020).

CONFLICT OF INTERESTS

The authors declare they have no conflicts of interests.

DATA AVAILABILITY STATEMENT

Data has been deposited at the public repository Dryad (<https://doi.org/10.5061/dryad.v15dv41wh>).⁵²

ORCID

Kyriacos Kareklas  <https://orcid.org/0000-0003-2453-9556>

Rui F. Oliveira  <https://orcid.org/0000-0003-1528-618X>

REFERENCES

1. Ward A, Webster M. *Sociality: the Behaviour of Group-Living Animals*. Springer-Verlag; 2016.
2. Dunbar RIM, Shultz S. Evolution in the social brain. *Science*. 2007; 317:1344-1347.
3. Taborsky B, Oliveira RF. Social competence: an evolutionary approach. *Trends Ecol Evol*. 2012;27:679-688.
4. Farris SM. Insect societies and the social brain. *Curr Opin Insect Sci*. 2016;15:1-8.
5. Heyes C, Pearce JM. Not-so-social learning strategies. *Proc R Soc B*. 2015;282:20141709.
6. Varela SAM, Teles M, Oliveira RF. The correlated evolution of social competence and social cognition. *Funct Ecol*. 2019;34(2):332-343.
7. Bowen MT, Keats K, Kendig MD, Cacic V, Callaghan PD, McGregor IS. Aggregation in quads but not pairs of rats exposed to cat odor or bright light. *Behav Process*. 2012;90(3):331-336.
8. Kleinhappel TK, Pike TW, Burman OHP. Stress-induced changes in group behaviour. *Sci Rep*. 2019;9:17200.
9. Bell A. Behavioural differences between individuals and two populations of stickleback (*Gasterosteus aculeatus*). *J Evol Biol*. 2005; 18(2):464-473.
10. Sören K, Frieder P, Maren B, Tilo K. The rewarding nature of social interactions. *Front Behav Neurosci*. 2010;4:22.
11. Gunaydin L, Deisseroth K. Dopaminergic dynamics contributing to social behavior. *Cold Spring Harb Sym*. 2014;79:221-227.
12. Walsh JJ, Christoffel DJ, Heifets BD, et al. 5-HT release in nucleus accumbens rescues social deficits in mouse autism model. *Nature*. 2018;560(7720):589-594.
13. Donaldson Z, Young L. Oxytocin, vasopressin, and the neurogenetics of sociality. *Science*. 2008;322:900-904.
14. Goodson J, Thompson R. Nonapeptide mechanisms of social cognition, behavior and species-specific social systems. *Curr Opin Neurobiol*. 2010;20(6):784-794.
15. Goodson J. Deconstructing sociality, social evolution and relevant nonapeptide functions. *Psychoneuroendocrinology*. 2013;38(4): 465-478.
16. Südhof T. Neuroligins and neuroligins link synaptic function to cognitive disease. *Nature*. 2008;455:903-911.
17. Grayton M, Missler M, Collier D, Fernandes C. Altered social behaviours in neuroligin 1α knockout mice resemble core symptoms in neurodevelopmental disorders. *Plos One*. 2013;8(6):e67114.
18. Rabaneda L, Robles-Lanuza E, Nieto-González J, Scholl F. Neuroligin dysfunction in adult neurons results in autistic-like behavior in mice. *Cell Rep*. 2014;8(2):338-346.
19. Hörnberg H, Pérez-Garci E, Schreiner D, et al. Rescue of oxytocin response and social behaviour in a mouse model of autism. *Nature*. 2020;584:252-256.
20. O'Connell LA, Hofmann HA. The vertebrate mesolimbic reward system and social behavior network: a comparative synthesis. *J Comp Neurol*. 2011;519:3599-3639.
21. O'Connell LA, Hofmann HA. Evolution of a vertebrate social decision-making network. *Science*. 2012;336(6085):1154-1157.
22. Tibbetts E, Dale J. Individual recognition: it is good to be different. *Trends Ecol Evol*. 2007;22(10):529-537.
23. Mrakovcic M, Haley L. Inbreeding depression in the zebra fish *Brachydanio rerio* (Hamilton Buchanan). *J Fish Biol*. 1979;15(3):323-327.
24. Nasiadka A, Clark M. Zebrafish breeding in the laboratory environment. *ILAR J*. 2012;53(2):161-168.
25. Brown KH, Dobrinski KP, Lee AS, et al. Extensive genetic diversity and substructuring among zebrafish strains revealed through copy

- number variant analysis. *Proc Natl Acad Sci USA*. 2012;109(2):529-534.
26. Balik-Meisner M, Truong L, Scholl EH, Tanguay RL, Reif DM. Population genetic diversity in zebrafish lines. *Mamm Genome*. 2018;29:90-100.
 27. Oswald M, Robison BD. Strain-specific alteration of zebrafish feeding behavior in response to aversive stimuli. *Can J Zool*. 2008;86(10):1085-1094.
 28. Egan RJ, Bergner CL, Hart PC, et al. Understanding behavioral and physiological phenotypes of stress and anxiety in zebrafish. *Behav Brain Res*. 2009;205(1):38-44.
 29. Sackerman J, Donegan JJ, Cunningham CS, et al. Zebrafish behavior in novel environments: effects of acute exposure to anxiolytic compounds and choice of *Danio rerio* line. *J Comp Psychol*. 2010;23(1):43-61.
 30. Barba-Escobedo P, Gould G. Visual social preferences of lone zebrafish in a novel environment: strain and anxiolytic effects. *Genes Brain Behav*. 2012;11(3):366-373.
 31. Lange M, Neuzeret F, Fabreges B, et al. Inter-individual and inter-strain variations in zebrafish locomotor ontogeny. *PLoS One*. 2013;8(8):e70172.
 32. Mahabir S, Chatterjee D, Buske C, Gerlai R. Maturation of shoaling in two zebrafish strains: a behavioral and neurochemical analysis. *Behav Brain Res*. 2013;247:1-8.
 33. Maximino C, Puty B, Benzecry R, et al. Role of serotonin in zebrafish (*Danio rerio*) anxiety: relationship with serotonin levels and effect of buspirone, WAY 100635, SB 224289, fluoxetine and parachlorophenylalanine (pCPA) in two behavioral models. *Neuropharmacology*. 2013;71:83-97.
 34. Vignet C, Bégout M-L, Péan S, Lyphout L, Leguay D, Cousin X. Systematic screening of behavioral responses in two zebrafish strains. *Zebrafish*. 2013;10(3):365-375.
 35. Vital C, Martins EP. Strain differences in zebrafish (*Danio rerio*) social roles and their impact on group task performance. *J Comp Psychol*. 2011;125(3):278-285.
 36. Liu X, Guo N, Lin J, et al. Strain-dependent differential behavioral responses of zebrafish larvae to acute MK-801 treatment. *Pharmacol Biochem Behav*. 2014;127:82-89.
 37. Gorissen M, Manuel R, Pelgrim TNM, et al. Differences in inhibitory avoidance, cortisol and brain gene expression in TL and AB zebrafish. *Genes Brain Behav*. 2015;14(5):428-438.
 38. Séguret A, Collignon B, Halloy J. Strain differences in the collective behaviour of zebrafish (*Danio rerio*) in heterogeneous environment. *Roy Soc Open Sci*. 2016;3(10):160451.
 39. Kalueff AV, Gebhardt M, Stewart AM, et al. And the zebrafish neuroscience research consortium (ZNRC). Towards a comprehensive catalog of zebrafish behavior 1.0 and beyond. *Zebrafish*. 2013;10(1):70-86.
 40. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing 2021. <https://www.R-project.org/>
 41. Borgatti SP, Everett MG, Johnson JC. *Analyzing Social Networks*. Sage Publications; 2013.
 42. Borgatti SP, Everett MG, Freeman LC. *Ucinet for Windows: Software for Social Network Analysis*. Analytic Technologies; 2002.
 43. Ribeiro D, Nunes AR, Glikberg M, Anbalagan S, Levkowitz G, Oliveira RF. Oxytocin receptor signalling modulates novelty recognition but not social preference in zebrafish. *J Neuroendocrinol*. 2020;32:e12834.
 44. Groenewoud F, Frommen J, Josi D, Tanaka H, Jungwirth A, Taborsky M. Predation risk drives social complexity in cooperative breeders. *Proc Natl Acad Sci USA*. 2016;113(15):4104-4109.
 45. Ferguson JN, Aldag JM, Insel TR, Young LJ. Oxytocin in the medial amygdala is essential for social recognition in the mouse. *J Neurosci*. 2001;21(20):8278-8285.
 46. Resendez SL, Namboodiri VMK, Otis JM, et al. Social stimuli induce activation of oxytocin neurons within the paraventricular nucleus of the hypothalamus to promote social behavior in male mice. *J Neurosci*. 2020;40(11):2282-2295.
 47. Dochtermann N, Dingemans N. Behavioral syndromes as evolutionary constraints. *Behav Ecol*. 2013;24(4):806-811.
 48. Saltz J, Hessel F, Kelly M. Trait correlations in the genomics era. *Trends Ecol Evol*. 2017;32(4):279-290.
 49. Royauté R, Hedrick A, Dochtermann NA. Behavioural syndromes shape evolutionary trajectories via conserved genetic architecture. *Proc R Soc B*. 2020;287:20200183.
 50. Varshney GK, Pei W, LaFave MC, et al. High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9. *Genome Res*. 2015;25:1030-1042.
 51. Wright D, Nakamichi R, Krause J, Butlin RK. QTL analysis of behavioral and morphological differentiation between wild and laboratory zebrafish (*Danio rerio*). *Behav Genet*. 2006;36:271-284.
 52. Oliveira R, Gonçalves C, Kareklas K, et al. Phenotypic architecture of sociality and its associated genetic polymorphisms in zebrafish. *Dryad, Dataset*; 2021. [10.5061/dryad.v15dv41wh](https://doi.org/10.5061/dryad.v15dv41wh)

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Gonçalves C, Kareklas K, Teles MC, et al. Phenotypic architecture of sociality and its associated genetic polymorphisms in zebrafish. *Genes, Brain and Behavior*. 2022;21(5):e12809. doi:[10.1111/gbb.12809](https://doi.org/10.1111/gbb.12809)