OTKA Zárójelentés

A pikkelymintázat és egyéb kapcsolt tulajdonságok öröklődésének vizsgálata halakon

(**77681**) Projektvezető: Bercsényi Miklós

Háttér

A projekt célkitűzése az volt, hogy egy korábban évtizedeken át használt öröklődési modell hiányosságait feltárja, állítson fel egy - a valóságot jobban leíró - modellt, és próbálja meg a történéseket molekuláris oldalról alátámasztani. A kezdeményezők úgy látták, hogy egy hazai ponty tájfajtában a bőrponty változat nem mutatja azt a letalitást, amit a korábbi modell jósolt, ugyanakkor az u.n. oldalvonal-soros változatokból is más arányok hasadtak ki, mint az a modell szerint várható volt. Felmerült az a kérdés is, hogy milyen a kapcsolat egyéb fenotípusos tulajdonságok és a pikkelymintázat között. Azt is láttuk, hogy előfordul olyan fenotípus is, amilyet korábban nem írtak le, és amit a korábbi modell nem értelmez. Részben a projekt előkészítése idején, részben annak folyamán kiderült, hogy három másik fajon is előfordul a vad (teljesen pikkelyezett) fenotípustól eltérő változat is.

A kérdéskör tisztázására két irányban indítottuk el a vizsgálatainkat. Az egyik irány a

- különféle fenotípusok összegyűjtése, és ezekkel történő célzott keresztezések elvégzése után morfológiai és növekedési vizsgálatok, valamint hasadási arányok megállapítása;

a másik irány pedig

- a pikkelyek kialakulásáért felelős gének keresése és azok esetleges alléljainak az elkülönítése volt.

Mintázati fenotípusok gyűjtése, új altípusok, célzott keresztezések, hasadási arányok

A különböző pikkelymintázatú halakat tógazdaságokból, természetes vizekből és díszhal-kereskedésekből szereztük be. Ezek között voltak vadnak nevezhető fajták, mint pl. pikkelyes amúri vadponty, valamint nemesített változatok, mint a tatai pikkelyes, vagy a legtöbb tükrös változat. A teljesen pikkelytelen bőrpontyot egy kimondottan európai nemes formát mutató hazai változatból (hajdúböszörményi bőrponty), ill. egy megnyúlt, ázsiai koi-változatból (szingapuri bőrkoi) választottuk ki. Az oldalvonal-soros anyahalakat részben hazai (Köröstarcsa), részben ázsiai változatokból használtuk fel. A projekt során találtunk egy oldalvonal-soros amúrt is. Együttműködő partnereink pedig a zebradánió fajból is izoláltak tükrös változatot.

A rendelkezésre álló fenotípusok felhasználásával 30 keresztezést végeztünk el. A keresztezéseket először Magyarországon, majd részben magyar bőrponty, ill. annak keresztezett utódjai felhasználásával, Szingapurban hoztuk létre és neveltük fel. A szaporítás és nevelés – ellentétben a kirpichnikovi, nyílt tavakban végzett kísérletekkel – végig kontrollált labor körülmények között történt. Keresztezéseket és visszakeresztezéseket 3 éven át végeztünk. Az ikrák termékenyülését és a kelési százalékot minden esetben megállapítottuk és összevetettük a korábbi modell szerinti várható értékekkel. Az egy-egy szaporító-pártól származó utódcsoportokat elkülönítetten neveltük. Abban az esetben, ha egy családon belül az ivadékok extrém mértékű szétnövést mutattak, a halakat nagyság szerint szétválogattuk és úgy neveltük addig a méretig, amíg a pikkelyezettség egyértelműen megállapíthatóvá vált. Ekkor a halakat egyenként mindkét oldalról lefényképeztük, és belőlük úszómintát vettünk későbbi molekuláris vizsgálatok elvégzéséhez. Az egyes fenotípusokból garatfogaik vizsgálatára is tartósítottunk mintákat.

Ahogyan az várható volt, a homozigóta pikkelyes változatok önmagukkal és a tükrös változatokkal egyöntetű pikkelyes utódokat eredményeztek. A harmadik és negyedik évben már csak az. u.n. érzékeny változatokra koncentráltunk. Azok keresztezési kombinációit komplettáltuk, ill. többeket megismételtünk, európai és ázsiai eredetű szülőkkel egyaránt. Munkánk során találtunk egy olyan amúr változatot, amelyik oldalvonal-soros mintázatot mutatott.

Felelős gének keresése, allélok elkülönítése

Megállapítottuk, hogy a ponty részleges pikkelyvesztését az *fgfr1a1* gén (ez a Kirpichnikov által korábban feltételezett 's' gén) mutációja okozza, mely homozigóta állapotban részleges funkcióvesztéshez vezet. A második gént, azaz a Kirpichnikov által 'N'-nek nevezett lókuszt, mind a mai napig nem sikerült azonosítani.
Hipotézisünk szerint a második gén valószínűleg a fibroblaszt növekedési faktor kaszkádon keresztül fejti ki hatását. Ezt kétféleképpen teheti meg:
a) a kaszkádban szerepet vállalva, vagy
b) a kaszkádba torkolló folyamatokban kifejtve hatását.

Első lépésként az Fgf kaszkád két célgénjét teszteltük, hogy bizonyítsuk: a második mutáció valóban ezen az úton, nem pedig egy független kaszkádon keresztül fejti ki hatását. Az eredmények igazolták feltételezésünket: az Fgf kaszkád célgénjeinek aktivitása lépésenként csökkent a pikkelyes- től a tükrösön át a bőrpontyig. A fibroblaszt növekedési faktor kaszkád mindkét célgénje, *dusp6* és *sef,* tendenciózusan csökkenő szintet mutatott a pikkelyes, tükrös és bőrponty egyedekben. A különböző fenotípusokból származó minták összehasonlítása igazolja a korábbi eredményt az *fgfr1a1* génben bekövetkező mutáció hatásáról, valamint arra utalnak, hogy az N gén szintén ezen a kaszkádon keresztül fejti ki (közvetlenül vagy közvetve) a hatását.

Ezután hasonló eljárással teszteltük négy olyan jelátvivő folyamat, ill. transzkripciós faktor, egyenként két-két célgénjét, melyekről korábban kimutatták, hogy az Fgf kaszkád működését szabályozzák. A három folyamat a következő volt: 'ectodysplasin' (Eda/Edar), kanonikus Wnt és retinol sav (RA), míg a transzkripciós faktor a T-box 5 (Tbx5). Mind a négy esetben ugyanazt az eredményt kaptuk: nem volt változás a három fenotípusból izolált minták expressziója között.

Elemeztük az úszók alakváltozásait valamint a garatfogak számának csökkenését. Megállapítottuk, hogy ezek a fenotípusos változások összekapcsolhatók a pikkelymintázat alakulásával, illetve az Fgf allélok expressziójával. Úgy látjuk, hogy az általunk megnevezett új szórt altípusok a hagyományos tükröshöz képest megnövekedett Fgf jelek következményei, legyenek bár azok akár az Fgf út egy további mutációjának, akár egy fentről, ebbe az útba torkolló funkcionális kaszkád egyik génje megváltozásának eredményei. Megjegyezzük, hogy ebben a tekintetben jelentős különbségeket találtunk a magyar és az ázsiai eredetű bőrpontyok tekintetében. Az ázsiai bőrpontyok egészen extrém úszó- és garatfog degradációt mutattak. Ezen halak mozgásában komoly problémák léptek fel, ami a növekedési-erély jelentős csökkenésével is párosult. Ugyanez a magyar bőrpontyoknál garatfogak tekintetében nem jelentkezett, és az úszók is csak kis mértékben deformálódtak.

Izoláltunk egy teljes hosszúságú új *fgfr1* paralogot is pontyból. Ez a paralog azonban a tükrös és a bőr változatok között nem mutatott különbséget.

Oldalvonal-soros amúr - bemetszés után regenerálódó - farokúszójából RNS-t izoláltunk, ami alkalmas lesz – akár közvetlen szekvenálása, akár cDNS-ének felhasználása révén, a pontyban talált szekvenciákkal való azonosság, vagy különbözőség kimutatására.

Modell-készítés

Készítettünk egy modellt, ami a korábban elfogadottnál alkalmasabb a kísérleti eredmények magyarázatára. Ennek képi megjelenítése a mellékelt kézirat 7.sz. ábráján található. Ez a modell lehetővé teszi a fokozatos pikkelyvesztés okának indoklását. Nem tételezi fel a keléskori letalitást, de az egyes pikkelyezettségi fenotípusok relatív fitneszeinek jelentős különbségét figyelembe tudja venni. A korábban „S" génnek nevezett gén szerepét leírja, és lehetőséget nyújt arra, hogy akár „N" gén nélkül, akár a jövőben pontosítandó, ennek szerepét betöltő génnel a fenotípusok kialakulását magyarázhassa.

A projekten dolgozó személyzet

Pannon Egyetem, Állattudományi és Állattenyésztéstani Tanszék
Bercsényi Miklós, a szaporítások és fenoípus-értékelések felelőse, projektvezető (bértámogatás nélkül),
Németh Szabolcs biológus, (OTKA bértámogatással
Szűcs Réka PhD hallgató, fő kutatási témája alapján, (bértámogatás nélkül)
valamint: Meiszter Máté BSc hallgató,
Havasi Máté, Németh Sándor, és Felföldi Zoltán PhD hallgatók (Németh Szabolcsnak a projektből való kiválását követően időszakos megbízással, részfeladatok elvégzésére, (OTKA bértámogatással)

Külföldi együttműködő partnerként, OTKA támogatás nélkül
Temasek Life Sciences Laboratories, Singapore
Orbán László, laborvezető, senior kutató, a molekuláris vizsgálatok legfőbb vezetője/mozgatórugója
Shubha Vij, Laura Casas, Natascha May Thevasagayam,
Chin Heng Goh és Purushothaman Kathiresan, beosztott kutatók

University of Tübingen, Institute of Developmental Biology
Matthew Harris csoportvezető
Nicholas Rohner, PhD hallgató
Christiane Nüsslein-Volhard intézetigazgató

Publikációk

Publikációink feltöltve a honlapra. Ezeken túl feltöltöttünk egy kéziratot is, amit egy viszonylag magas impaktú folyóiratba szánunk, és ami részletezi a projekt eredményeit.
Mivel ez a kézirat olyan információkat tartalmaz, amelyeket publikálás előtt nem szeretnénk illetéktelenek számára hozzáférhetővé tenni, szeretnénk kérni az OTKA által biztosított azon opciót, hogy a megjelenésig – de legkésőbb 1 év leteltéig – ezt ne hozza nyilvánosságra. Mihelyt a kézirat publikálásra elfogadásra kerül, azt azonnal jelezzük és örömmel járulunk majd hozzá a közzétételhez.

Jövőbeni tervek

Szeretnénk a jövőben azt megvizsgálni, hogy a természetes módon, vad állapotban is pikkely nélküli (csupasz) halak, pl. több harcsafaj minek következtében veszítette el a pikkelyeit. Ezen túl szeretnénk majd ennek a jelenségnek a humán vonatkozásaihoz is közelebb kerülni.

Disappearing scales in carps:
Re-visiting Kirpichnikov's model on the genetics of scale pattern formation

Réka Szűcs[1*], Shubha Vij[2*], Laura Casas[2*&], Natascha May Thevasagayam[2], Chin Heng Goh[3], Sándor Németh[1], Purushothaman Kathiresan[2], Zsigmond Jeney[4], Matthew Harris[5], Miklós Bercsényi[1+] and László Orbán[2,1,6+]

1) Department of Animal Sciences and Animal Husbandry, Georgikon Faculty, University of Pannonia, Keszthely, Hungary
2) Reproductive Genomics Group, Strategic Research Program, Temasek Life Sciences Laboratory, Singapore
3) Fish Facility, Temasek Life Sciences Laboratory, Singapore
4) Research Institute for Fisheries, Aquaculture and Irrigation, Szarvas, Hungary
5) Harvard Medical School, Cambridge, MA, USA
6) Department of Biological Sciences, National University of Singapore, Singapore

* These authors have contributed equally to the production of data presented in the paper.

& Current address: XXX

+ Corresponding authors: Miklos Bercsényi - Tel: +36-20-971-6055; Fax: +36-83-545-143; E-mail: bm@georgikon.hu; László Orbán - Tel: +65-6872-7413; Fax: +65-6872-7007; Email: laszlo@tll.org.sg

Keywords: cyprinid, mirror carp, linear carp, scaleless carp, nude carp, Fgfr1

Abstract

The body of most fishes is fully covered by scales that typically form tight, partially overlapping rows. While the molecular processes leading to the formation and growth of fish scales have been investigated, very little is known about the genetic mechanisms regulating scale pattern formation. Although the existence of two genes (s and N) regulating scale coverage in cyprinids have been predicted nearly eighty years ago (Kirpichnikov and Balkashina, 1935&1936), their identity was unknown until recently, when one of the was found to be a paralog of fibroblast growth factor receptor 1, *fgfr1a1*. The current study describes the first steps of our continuing search towards the identification of the second gene, called N.

We re-visited the original model of Kirpichnikov that proposed four major scale pattern types through the analysis of offspring generated by a large number of crosses involving loss-of-scale mutants of European and Asian origin. We showed that varieties of the so-called scattered phenotype with a larger number of non-overlapping scales often appear in offsprings of mirrrors and nudes. Therefore, we divided the scattered type into three sub-types: irregular, incomplete scaled and classical mirror. We also analyzed the survival rates of offspring groups potentially inheriting two N alleles and found distinct differences between Asian and European crosses, indicative of the presence of a strong N allele with homozygous lethality in the former one and a weaker, non-lethal one in the latter.

We analyzed the inheritance patterns, deformations of fins and losses of pharyngeal teeth and found that phenotypic changes show gradations in crosses as opposed to a few distinct groups. We propose that the new sub-types of scattered were formed due to increased levels of Fgf signals compared to mirrors and especially nudes, either due to an additional mutation in one of the FGF signaling pathway genes or that in an upstream pathway functionally connected the Fgf signaling.

We isolated the full-length transcript of a new *fgfr1* paralog, *fgfr1b* from common carp. When the sequence of *fgfr1b* was compared between mirror and nude individuals was compared, no difference was found.

Finally, we describe ongoing and potential future approaches for the isolation of the N gene, the mutation of which leads to complete scale loss in individuals carrying homozygous mutations in the *fgfr1a1* gene.

Introduction

Cyprinid teleosts account for over 30% of worldwide aquaculture production and according to the FAO, common carp *(Cyprinus carpio L.)* is the species with the third highest production today (http://www.fao.org/fi/default.asp). Common carp was probably the earliest domesticated fish species for alimentary purposes, with records of ancient Chinese documents showing that cultivation of common carp in China began in the twelfth century BC (1-3). In Europe, common carp was first domesticated by the Romans before the sixth century (1-4).

Today, common carp is divided into at least two subspecies: the separation of Central-Asian/European *(C. carpio carpio)* and East-Asian subspecies (*C. carpio haematopterus)* is well supported by microsatellite and mitochondrial genetic data (5-8). In addition, the existence of a potential third subspecies *(C. c. rubrofuscus or C. c. viridiviolaceus)* is possible, but not confirmed based on the genotypes (6). Earlier, a Central-Asian subspecies *(C. c. aralensis)* was proposed by Kirpichnikov (9). However, recent studies (5, 6, 10) have demonstrated that the European and Central-Asian forms of common carp are actually quite closely related, with the latter comprising a subset of the genetic diversity of the former. The authors subsequently classified both European and Central-Asian carp as subspecies *carpio*. Based on the analysis of mtDNA sequences, Froufe and colleagues (11) concluded that the European common carps were likely introduced from Asia.

The domestication of common carp led to the emergence of different varieties, among them various scalation patterns. The wild phenotype was a fully scaled torpedo-shaped fish, but through artificial selection a number of scalation variants have been developed over the centuries. These variants, characterized by the reduction of the scale coverage, have been favoured as they were easier to de-scale for cooking (12). According to Kirpichnikov (9, 13, 14), the main scalation types of common carp are: scaled, linear, scattered and nude (Sulpplementary File S1A-D). In addition to the above phenotypes, several additional varieties, including irregular and incomplete scaled have also been reported (13, 15), but they have mostly been regarded as deviations and therefore, have not been included in the genetic model (see below).

The distribution of scales over the body of cyprinids is genetically determined. Rudzinsky (16, 17) was the first to point out that scaled variety of common carp is dominant over the mirror one. Based on data obtained by remarkably simple tools, such as survival rates and

phenotypic analysis of individuals grown in ponds, Kirpichnikov and colleagues (18, 19) proposed a 'two genes – four alleles' type model for the inheritance of scale pattern in common carp. According to their model, scaled fish are of SS/nn or Ss/nn genotype, scattered carps are ss/nn or ss/Nn, linears (or 'linear mirrors' with a line of scale running along the lateral line) are SS/Nn or Ss/Nn, while nudes (or 'leathers' without scales) are ssNn [for review see: (13, 14); Supplementary File 1]. Based on their observations, NN results in lethality in any combination with ss, SS or Ss [for review see: (13, 14)].

Over the next decades, the majority of textbooks took over the model and it became the most well-known example for two-genic inheritance in fish genetics (see e.g. (20, 21)). Although some of the crosses were repeated subsequently and yielded data similar to the original ones (see e.g. (15, 22, 23)), according to our knowledge, nobody has re-visited the issue by performing a systematic analysis with a larger set of crosses. Recently, two findings motivated us to reconsider the model. The first result was that nude x nude common carp crosses performed at one of the Hungarian fish farms repeatedly failed to show either the 25% lethality, or the 25% of scattered phenotypes (15) expected on the basis of the Kirpichnikov model (13). The second was the discovery of a "mirror" variant in zebrafish and the identification of the mutant gene responsible for this phenotype: one of the paralogs of fibroblast growth hormone receptor 1, *fgfr1a* in zebrafish and *fgfr1a1* in common carp (24). In other words, this is the 's' gene predicted earlier based on data from common carp by Kirpichnikov and his team (13, 14, 18, 19). This discovery has paved the way for a more informed search for the second member of this interesting gene pair, the so-called 'N' gene.

In this manuscript, we describe the ratio of scale pattern phenotypes in offspring groups originating from crosses involving brooders with partial or full loss of scale sets. We also isolate and structurally characterize a hitherto missing member of the Fgfr1 receptor family, *fgfr1b*, and show that its sequence has not been mutated in nude individuals in comparison to mirrors. Finally, we propose a model that could explain the 'deviating phenotypes' observed in some of the crosses described above.

Material and Methods

Brooders

For the crosses performed in Hungary, common carp brooders (males and females) have been selected from the following sources: scaled carp - Amur wild type carp, and Tata common carp from the live cyprinid collection of HAKI (Szarvas, Hungary); mirror carp: Line No2 from HAKI; linear carp - from Tiszaker fish farm (Kőröstarcsa, Hungary); nude carp - from Béke fish farm (Hajdúböszörmény, Hungary).

For the crosses performed in Singapore, a European nude male carp was shipped from Hungary to Singapore and used as a father for a large number of crosses. In addition to that, koi carps of the four major and some minor scale pattern types were purchased from XXX, and used as brooders.

Artificial propagation

The breeders were prepared for the artificial propagation by hypophysation according to (4). Small batches of eggs (ca. 50g) from each female were fertilized by 2 ml of fresh milt collected earlier from the chosen male(s). For the crosses performed in Hungary, two minutes after fertilization, the eggs were stacked onto a tulle netting that was stretched onto a metal frame. This provided easy and accurate tracking of embryonic development, as fertilization rate and hatching percent were calculated by counting the live or dead eggs using digital photos on the eggs stacked to the net. For the crosses performed in Singapore, the stickiness of fertilized eggs was first removed through a treatment with Woynarowich solution (25) and later they were placed into traditional Zuger jars and they were hatched there. Survival rates were calculated by removing a random sample of eggs and counting live vs. dead individuals under a stereo microscope.

Hatching, larviculture and phenotyping

Fry were hatched out in separate tanks in order to avoid potential mixing of different families. Feeding of fry started on the 3rd day after hatching by live brine shrimp nauplii. From the end of the second week, live food was gradually replaced with dry pelleted feed over a week

transition. In Singapore, mutants were separated from the rest and grown in smaller tanks. As the rest of fish grew in the aquaria, their number was reduced systematically by random removal to keep the density acceptable. In Hungary, fish were transferred to earthen ponds at XXX age and fed with XXX. The families were reared for four months when the scale pattern could be clearly identified. At this timepoint, for the first two crosses performed in Singapore (NN1&2; Supplementary File S2) classification was performed directly through visual observation of the fish, whereas for the remaining Singaporean crosses and all crosses performed in Hungary, fingerlings were individually photographed from both sides and scalation was assessed based on the photos. Phenotypic analysis was performed by assessing the scale patterns based on a classification (see Supplementary File S3) that has been a modified version of Kirpichnikov's (13), as our classification contained a total of six categories instead of the four used earlier. We have retained three of the four major scale patterns, namely, scaled, linear, and nude (Supplementary File S1). In addition, we have divided Kirpichnikov's 'scattered (or mirror)' category into three sub-categories: irregular, incomplete scaled and classical mirror (Fig. 1; for descriptions see Supplementary File S3).

In few cases, the scale pattern on the two sides of the fish were different. In these instances, the fish were classified based on the overall phenotype, e.g., if an individual had 10-20% scales on one side and 70-80% scales on the other side then it was classified as an irregular and not an incompletely scaled individual. Phenotype frequencies within the families as percentage were compared to the expected values calculated from the Kirpichnikov model.

Isolation of pharyngeal teeth

For isolating pharyngeal teeth, individuals were culled by placing them into 2% ethyl 3-aminobenzoate methanesulfonate salt (MS222; Sigma-Aldrich, St. Louis, MO, USA) for 15 minutes. Then, their head portion was cut off at the distal end of the operculum and submerged in 4% potassium hydroxide to dissolve the soft parts. After 2-3 days, the pharyngeal teeth were picked from the remaining mass of tissue and thoroughly washed in water and dried. The number of teeth was counted under a Leica M125 stereomicroscope and the photographs were taken with a Leica MZ 10F stereomicroscope fitted with a Nikon DXM 1200F camera.

Sample collection and isolation of nucleic acids for genotyping and sequencing

Individuals showing different scale coverage were tranquilized in 2% MS222. Their fin clips were collected, placed into 95% ethanol and stored at 4ºC until use for DNA isolation. Genomic DNA was isolated using the standard phenol-chloroform method (26).

For RNA isolation, the ends of caudal fins were cut using a sharp scalpel, the fins were then allowed to regenerate for three to five days. Following this period, the regenerated part of the fin was collected, immediately immersed in Trizol and stored at -80ºC until further analysis. Total RNA was extracted from regenerating fins samples collected on the 3rd to 5th day following the cut by using the Trizol Reagent (Invitrogen, Carlsbad, CA, USA) according to manufacturer´s protocol.

The quality and concentration of nucleic acids was tested by spectrophotometry using a Nanodrop Spectrophotometer ND-1000 UV/Vis (Nanodrop Technologies, Wilmington, DE, USA), followed by agarose gel electrophoresis.

Isolation, sequencing and comparative analysis of cDNA sequences from candidate genes

For the isolation of the additional two presumed copies of common carp *fgfr1b*, specific primers targeting the differential regions between teleost paralogues *fgfr1a* and *fgfr1b* were designed. The design was based on the alignment of the two already described *fgfr1a* paralogues from common carp with *fgfr1a* and *fgfr1b* from zebrafish (*Danio rerio*; *fgfr1a* LG8, Ensembl ID: ENSDARG00000011027, *fgfr1b* LG10: ENSDARG00000011190), three-spined stickleback (*Gasterosteus aculeatus*; ENSGACG00000012410, ENSGACG00000015518), green spotted pufferfish (*Tetraodon nigroviridis*; ENSTNIG00000018850, ENSTNIG00000013597), Japanese fugu (Takifugu rubripes; ENSTRUG00000016527, ENSTRUG00000018627) and Japanese medaka (*Oryzias latipes;* ENSORLG00000014206, ENSORLG00000000321). (For the alignment of these sequences please see Supplementary File S4) cDNAs from the regenerating fin samples of two common carp individuals showing the mirror phenotype were PCR-amplified with primers fgfr1b_1F and fgfr1b_2R under the following conditions: Reactions were carried out in a total volume of 25µl using the AmpliTaq DNA Polymerase package (Applied Biosystems, Foster City, CA, USA) containing 1X PCR buffer, 2µM primer, 100 µM dNTP mix, 2 mM MgCl2, 20 ng

cDNA template, and 0.5 U Taq polymerase. The PCR reaction was performed in a PTC-100 thermal cycler (MJ Research, Watertown, MA, USA) by using the following program: an initial denaturation at 94ºC for 2 minutes followed by 30 cycles at 94ºC for 15 seconds, 62ºC for 45 seconds and 2 minutes at 72ºC for extension. A final step was performed at 72ºC for 5 minutes for final elongation.

PCR products (20 µl) were separated on a 2% agarose gel (Bio-Rad, Hercules, CA, USA) in 1X TBE buffer containing either 0.5µg/ml ethidium bromide or 10nl/ml Gelstar (FMC BioProducts, Rockland, ME, USA). The gel was placed onto a UV-lamp to excise the band using a scalpel and the DNA content was isolated using the GFX^{TM} PCR DNA and Gel Band Purification Kit (Amersham Biosciences, Piscataway, NJ, USA). The fragment was then ligated into the pGEM T-easy Vector System (Promega, Madison, WI, USA) and fifty clones containing an insert of the expected size were sequenced a minimum of five times on both strands using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturer´s recommendations in an ABI Prism 3100 sequencer (Perkin Elmer, Foster City, CA, USA). Sequencing reactions were carried out in a total volume of 20µl, containing 2µl 5X BigDye sequencing buffer, 4 µl 2.5X Terminator Ready Reaction Mix, 3.2 pmol universal primer T7 or SP6, and 1 µl of purified DNA. Produced sequences were edited and assembled using SequencherTM v4.0.5 analysis software (Gene Codes Corporation, Ann Arbor, MI, USA),

RACE Procedure

The full-length cDNAs were obtained by using the RACE Technique or Rapid amplification of cDNA ends. Reverse transcription and rapid amplification of cDNA ends was carried out using the FirstChoice RLM-RACE Kit (Ambion) following manufacturer´s protocol. The gene-specific primers provided by the user were designed based on RACE requirements using the Primer 3 Program (27) and their sequences are described in Supplementary File S4.

Cloning of the RACE products was done using the pGEM T-easy Vector System (Promega, Madison, WI, USA). Twenty independent clones for each of the 5′- and 3′-RACE products were sequenced a minimum of five times on both strands using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturer´s recommendations in an ABI Prism 3100 sequencer (Perkin Elmer, Foster

City, CA, USA). Sequencing reactions were carried out as previously described. The assembled cDNA sequences were aligned using ClustalX 2.0 software (12) and BLAST searched (28) against GenBank (http://www.ncbi.nlm.nih.gov/blast).

Searching CarpBase XXX

Comparative analysis was performed by sequencing PCR-amplified cDNA of the selected genes from three mirror and three nude common carp siblings and comparing the sequences. XXX

Bioinformatic analysis

Proteins were represented using the DOG 1.0 software (29). The phylogeny of fish *fgfr1s* was reconstructed using the Maximum Parsimony Method implemented in MEGA4.0.2 software (30). Confidence in the resulting unrooted tree was assessed by bootstrapping (1000 replicates). The tree was represented using the Archeopteryx software (29).

Results

Lack of expected lethality among the offspring of European common carps with nude and linear scale pattern types

Altogether, we have performed XXX crosses at two different locations (Supplementary File S2) and estimated the survival rates of their offspring either by i) counting fertilized eggs with eye spots (viable embryos) and those without (dead eggs) from nets; or ii) by sorting a few hundred embryos randomly removed from the Zuger jar under a dissecting scope. Analysis of the survival rates showed the expected 25% lethality in all nude x nude (N x N) crosses performed in Singapore (data not shown), but not among the offspring of European linear x linear (L x L), linear x nude (L x N) or N x N crosses done in Hungary. The mean survival rates for these latter three offspring groups were XXX+/-XXX%, XXX+/-XXX% and XXX+/- XXX%, respectively, not significantly different from the mean of the other types of European crosses tested (XXX+/XXX%; p>XXX; XXX).

The scale pattern phenotypes of the offspring originating from seventeen different crosses involving XXX brooders (see Supplementary File S2 for details) were analyzed in detail. All of these crosses involved brooders with reduced scale pattern types: 14 were between the classical scalation types (i.e. linear, mirror and nude), whereas in the remaining three one of the parents showed the irregular scale pattern (see Supplementary File S3 for detailed description). When classified according to the origin of the parents, ten crosses involved partners originating from the same subspecies (European x European or koi x koi), three of them were between the two subspecies and the remaining four involved one or two F1 hybrids from a cross between the two subspecies.

In several cases, we have found substantial deviation from the ratios predicted based on Kirpitchnikov's model (XXXref). Two of the three 'all European' L x L crosses yielded a majority (95% and 65%) of linear offspring with the rest showing irregular (I), incomplete scaled (Isc), and classical mirror (M) phenotypes (see Fig. 2 for representative examples and Supplementary File S3 for detailed phenotype description). The offspring phenotypes from the third L X L cross showed a very similar proportion of Ls (32%) and Ms (31%), while the rest was divided between Isc (21%) and I (14%; Supplementary File S2). No classical scaled offspring was found in any of the three crosses.

In the two crosses involving an irregular and a classical mirror type parent, the combined proportion the two new sub-categories (I and Isc) dominated the phenotype list (82% and 64%, respectively). In addition to classical mirrors, a few nudes (4% in both crosses) also showed up (Fig. 3A). No classical scaled or linears were found among the offspring.

In the three European L X N crosses, instead of the expected high proportio of Ls (33-67%) very few of them (1-9%) showed up. Most offsprings were classical Ms in all three crosses (88-99%) with a small proportion of unexpected nudes (2-9%) in two crosses.

One of the two N x M crosses (MN36) yielded 96% Ms, 3% Ns and 1% Ls, a substantial deviation from the expected equal proprtion of Ms and Ns. When an irregular female was crossed with a nude male, both I and Isc phenotypes appeared among the offspring, resulting in the combined proportion of 63% scattered together with classical Ms (I+Isc+M; expected: 50%).

In the four N x N crosses involving at least one koi parent, 33% Ms and 67% Ns were expected after the initial loss one quarter of the offspring. Interestingly, one or both new sub-categories of scattered appeared in all crosses, their combined proportion ranging from 15% to 53%. The ratio of nudes was lower than the expected 67% in all four crosses (range: 40-59%). In the only cross between two European nudes (NN26), the proportion of Ns has increased to 87% (expected: 75%) due to the lack of lethality, but the remaining 13% of the offspring were all classical Ms (Fig. 3B).

The deformity/disappearance of fins and gradual decrease in pharyngeal teeth count could be observed in all three subtypes of scattered, not just the nudes

XXX Fig. 1J-L

We tested potential associations between various levels of scale loss and fin deformity and/or loss in irregular, incomplete scaled, mirror and nude individuals from four families originating from crosses involving European and Asian grandparents (XXX, XXX, XXX and XXX). Fin defects showed a progressive increase with the decrease in the number of scales such that the irregular individuals had the least of these abnormalities in terms of fins being distorted (reduced/stunted) or absent while the nude group had the maximum number of such

defects. In fact, amongst the irregular, incomplete scaled and mirror groups, the dorsal fin was the most affected and barring it, the observed defects were <10% for the remaining fins. Conversely for the nudes, ~95% of the individuals had at least one fin defect with the dorsal fin absent from ~80% and the pectoral and pelvic pair fins missing from ~60% in the group (Figure 4A). Fin defects were also quanitified on a per-fish basis using an arbitrary scale by assigning one point for distortion of a fin and two points for each fin loss. Only ~1% of the irregular, incomplete scaled and mirror, but at least 50% of the nude fish had >10 points. Likewise, >75% of the irregular, incomplete scaled and mirror fish had >2 points (Figure 4B).

The association between the scale pattern and the number of pharyngeal teeth was also tested. There was a progressive loss of pharyngeal teeth in parallel with decreasing scale coverage. Almost 70% of the nudes entirely lacked teeth, while the rest of them had between 1-4 teeth only. The teeth numbers for the other three groups were: incomplete scaled – 4-8, mirror – 5-8 and irregular – 5-9, with almost ~70% of the individuals in each of these three groups showing the presence of at least 5 teeth (Figure 5). At the other end of the scale, most scaled individuals (XXX%) had a complete set of pharyngeal teeth, whereas the rest were missing just one (XXX%) or two of them (XXX%), thus the range for those was 8-10.

We have also compared the averaged relative size of the biggest scales in the four different phenotypes with scale loss (I, Isc, M and N) and found that they decreased in the following order: Isc>I>M>N (Supplementary File S5).

Isolation and structural characterization of the fgfr1b *paralog from common carp*

Earlier, two Fgfr1 paralogs, *fgfr1a1* and *fgfr1a2*, have been described from common carp (24). As common carp is a tetraploid species (31, 32), its genome could potentially contain additional two paralogs that might play a role in scale pattern formation. We have performed PCR-amplification of cDNAs from the regenerating fin samples of mirror carp individuals with primers binding to those regions of *fgfr1b* that were most different from *fgfr1a* in other teleost species. Analysis of the products has revealed two overlapping contings (1,562 and 682 bp) which presumably correspond to new *fgfr1b* paralogue(s) in common carp. As the two sequences have shown a 96% nucleotide identity with 655 of 682 nucleotides being identical, and three different efforts of sequencing of the common carp transcriptome from several organs yielded only a single contig (XXXcarpbase; XXXSpaink; our unpublished

data), we concluded that the two contigs represented transcript variants expressed from the same locus. When the 1,562 bp long consensus sequence was Blasted against GenBank, it produced a highly significant alignment with the zebrafish *fibroblast growth factor receptor 1b* (*fgfr1b*), mRNA (EU919571), showing a maximum identity of 88%, coverage 99% and e-value of 0.0%. When blasted against the two existing common carp *fgfr1* paralogs, *fgfr1a1* and *fgfr1a2*, XXX.

Subsequent 5'- and 3'-RACE reactions have successfully revealed the complete coding sequence of a new *fgfr1b* paralogue in common carp. XXX The sequence information has been deposited in GenBank under accession number XXXX. XXX

The deduced protein contained only two extracellular Immunoglobulin c2 type domains (IGc2a&b) compared to three in both *fgfr1a* paralogues (IGc2a-c) described earlier (Fig. 6A). In order to investigate whether the domain in question was lost or gained during the evolution of bony fishes, we compared the primary structure of the above proteins from several teleost species with their orthologs from cartilaginous fish (Uniprot accession numbers: picked dogfish, *Squalus acanthias* - D5FGJ8, D5FGF2; little skate, *Leucoraja erinacea* - D5FGF3). According to current estimates, the ancestor of the latter was separated from the common ancestor of bony fishes about 420 million years ago (33). The reconstructed phylogeny revealed that the Fgfr1 protein had originally three IGc2 domains in the ancestral fishes. One domain was likely lost from one of the two paralogs following the fish-specific duplication event (3R: (34, 35)) of the genome of the common teleost ancestor, but before the speciation of the bony fish species included in the analysis (Fig. 6B). Nonetheless, despite of the absence of one extracellular IGc2 domain, the *fgfr1b* paralog of common carp is likely to be functional, since it has been demonstrated in zebrafish that there is a functional redundancy of both forms during early embryonic development (24). After duplication, paralogous genes that are not silenced may acquire new functions through a process called neofunctionalization (36, 37). Others may subfunctionalize, or partition old functions as a strategy to escape disabling mutations that would lead to their eradication, or they can function redundantly (36, 37).

Comparative sequence analysis of full-length fgfr1b *cDNAs found no difference between mirror and nude common carp siblings*

The *fgfr1b* transcript was amplified and sequenced both from mirror and nude individuals in order to determine whether the latter contained mutation(s) associated with complete scale loss. Comparison of the 2,208 bp cDNA fragment from three mirror and three nude siblings did not identify any consistent difference between the two groups (data not shown), therefore we excluded the new paralog from among the potantial candidates of the N gene.

Quantifying the expression level of downstream target genes of the Fgf pathway in irregular, incomplete scaled, mirror and nude common carp individuals

Quantifying the expression level of two downstream target genes of the Fgf pathway by qRT-PCR to find out whether we can detect differences among the Fgf signal intensities in the two new sub-types compared to mirror (and nudes)...

Discussion

Our proposed extension of the Kirpichnikov model contains three sub-types of scattered: Irregular, incomplete scaled and classical mirror

Nearly a century ago, Rudzinsky (16, 17) described the first set of data on the genetic regulation of scale pattern formation in common carp. Later, Kirpichnikov and Balkashina (18, 19) added more details that eventually led to a complete model (13) that proposed existence of two loci and four alleles, the combination of which resulted in four major phenotypes (listed in the order of decreasing scale cover): fully scaled (wild type), linear, scattered and nude. In addition to the four major phenotypes, several sub-types were also described (13) as potential deviations from linear or mirror with extra number of scales, but their exact relationship to the main phenotypes was not determined.

The experiments described in this manuscript were initiated by two observations. The first one was the frequent appearance of sub-types among the offspring from crosses involving linears, mirrors and nudes that were clearly different from the four major phenotypes. (The second was the lack of lethality suggested by the present model, when two European nude individuals crossed that will be discussed in the next section.)

Here, we propose that a model where the completeness of scale pattern is dependent on the overall level of Fgf signal at the locations where scales are formed. According to our model, although the two genes proposed by Kirpitchnikov (S and n; (13)) would be located on two different chromosomes, functionally they would not be fully independent, as they would act along the same pathway(s) regulating the overall level of Fgfs signaling and thereby the activity of their downstream targents (Fig. 7). This is supported by our preliminary experiments that detected lower transcript levels of target genes of Fgf signaling in nudes than in scattered (data not shown). The combination of the variable effects from the two genes would result in a rheostat-like system, where intermediate phenotypes could appear among the major ones. We argue that instead of removing the sub-types from the system and labeling them as aberrations, they should be included, as their analysis will help us to gain bettwer understand of this complex system. Accordingly, we have sub-divided Kirpichnikov's scattered phenotype into three sub-types, and followed their inheritance in several crosses.

Based on the results, we propose that the increased number of scales in the irregular and incomplete scaled sub-types are the result of an elevated level of Fgf signaling compared to

classical mirrors and nudes. This level is higher than that in the classical mirrors, resulting in the formation of scales at many locations over the body surface, but lower than those that are required for the formation of the wild type pattern. Similar phenotypes with large non-overlapping scales were observed in carps with SssNnn genotype generated by triploidization of the eggs from a scaled and nude brooder, presumably due to incomplete dominance of the 'N' allele over two wild type 'n' alleles (38). Moreover, triploid nude carps with sssNnn genotype showed less severe phenotypic effects (reduced scale cover and number of anal fin rays) than to their diploid counterparts (ssNn; (39)).

We do not know the reason why these scales in the irregular and incomplete scaled sub-types are often bigger and why they aren't arranged in the tigth, partially overlapping order as those on the fully scaled wild types are. There might be a temporal increase in one of the signals in these individuals during scale formation that results in the fusion of their precursors. Additional research would be needed to find a reason for these phenotypes.

Lack of lethality in the offspring of European nudes indicates the presence of a new 'N' allele with milder phenotypic effect

When two European brooders carrying the proposed 'N' allele were crossed, no lethality was observed among the offspring (Fig. 2). Also, the distortions and losses of fins (Fig. 4) as well as severely reduced pharyngeal teeth counts (Fig. 5) often observed in Asian nudes, were not observed in most of their European counterparts. These observations seem to indicate that the European and Asian populations contain two different mutant 'N' alleles: a stronger one in the former and a weaker in the latter. The European allele causes the loss of scales, but it has limited, if any, effect on teeth and fin formation, whereas the strong Asian allele exerts strong, lasting effects on the formation of all three structures. In fact, the cummulative effects of the strong 'N' allele are so strong that those nude individuals that survive the early development are often not able to swim properly and exhibit a distorted body shape either due to skeletal deformations or as a consequence of the lack of fins. When such mutants are grown together with their unaffected (i.e. scattered, linear or fully scaled) siblings in larger tanks, most of them disappear during the first two months as they loose out in competition for food and get cannibalized by their stronger kins. From their observations it seems likely that the lines

Kirpichnikov and his colleagues worked with carried the stronger ’N’ allele, not the weaker one.

The effects of the ’N’ allele are dependent on location and developmental timing

Fgf signaling is essential for several important developmental processes throughout the animal kingdom (40-42). For instance, in humans they have a role in bone formation, smelling and reproduction (review: (43)), whereas they are essential for limb formation in mammals and birds (XXXRef). In fish, various Fgf ligands and receptors were shown to be involved with the formation of i) scales (24, 44); ii) median fin fold, the precursor of dorsal fin (45); iii) paired fins (46) and iv) lateral line in the zebrafish model (reviews: (47-49)), as well as fin regeneration (reviews: (50, 51)). From the above processes, the mutant ’N’ allele exerts the most severe negative effects on scale and dorsal fin development.

Interestingly, loss or reduction of dorsal fin has been documented from a number of other fish species (see e.g. (52-55)), especially those under intensive culture. The phenotype is called ’saddleback’, it is characterized by entirely missing or severely distorted dorsal fins, often together with fusion of some of the vertebrae. It was first described in blue tilapia as a genetically inherited trait, caused by a dominant, lethal mutation (56). Although this mutation does not usually result in scale-loss, its additional phenotypes, including decreased stress resistance and increased sensitivity to infections, make it likely that it affects similar processes in tilapia, as in ’N’ does in nude carps.

One of the advantages of scale-loss phenotypes is that they reveal preferential locations of scale formation that are not detectable on wild type individuals. The two locations, where scales tend to appear even in the case of severe scale loss are the area above the lateral line (in linears) and that below the dorsal fin (in scattered and some nudes). In case of the former, it seems likely that the increased Fgf levels are maintained during the period of scale formation, resulting in the formation of a line even when the general levels are reduced below the threshold necessary for scale fomation at most locations of the body surface. Such phenotypes have been observed in other cyprinids, including the goldfish (http://mirrorscalegoldfish.blogspot.com/) and grass carp (see Fig. 3 of (57)) and even in a more distantly related Patagonian species, the naked characin (*Gymnocharacinus bergi*, Steindacher, 1903). In this threatened species, the scales first develop over the whole body

surface, later they are re-absorbed with the exception of the area covering the lateral line resulting in a linear phenotype (58). The situation with the other region is more complicated, as there are individuals with a missing dorsal fin and a line of scale below. There are two potential explanation for such phenomena: a) the threshold of Fgf levels required for fin initiation is higher than that needed for scale formation; or b) the early effect of mutation is stronger that the late one.

Future outlook

It took more than eighty years after the first publication on the involvement of genetic mechanisms in scale-loss phenotype (16, 17) to figure out the identity of the 's' gene (24). We are currently working on the isolation of the second member of this gene pair by following three parallel routes.

Firstly, we have isolated several key members of the Fgf signaling cascade and genes from those upstream pathways that were shown earlier to regulate this process (see e.g. (40, 59)). Comparative sequence analysis of these cDNAs from nude and mirror sibling groups might allow for the identification of the N gene.

Secondly, we have generated several F2 mapping families by crossing European and Asian representatives of the species with partial or full scale-loss phenotype. Genetic linkage mapping that is becoming a routine exercise in common carp (see e.g. (60-62)) will reveal the chromosomal location that harbors the gene in question. Comparative bioinformatic analysis of the genes contained in syntenic regions of the sequenced teleost models, especially zebrafish might allow for narrowing down the list of potential candidates. Should that approach fail to identify the mutant gene, a map-based positional cloning can be performed for its identification.

Thirdly, rapidly increasing sequence information from traditional (63) and NGS-based sequencing efforts (64, 65) have already yielded benefits for isolation and characterization full-length cDNA sequences. One of the short-term benefits of these activities will be a publicly available high quality transcriptome (65) allowing for RNAseq-based transcriptomics, a substantial improvement of the from the current method of choice, the cDNA microarray (66).

According to our hope, parallel application of these three approaches will soon lead to the identification of the ’N’ gene and more complete understanding of the complex process of scale pattern formation in cyprinids and possibly other teleosts.

Acknowledgements

The authors thank Hans Komen for useful discussions; Herman Spaink for access to transcriptome assembly; Nicolas Roehner for *fgfr1a* sequences and primers; Alex Chang, Akos Horvath, Gyula Kovacs, Janos Merth, Dora Posfai and XXX for technical help with fish breeding and larval rearing; Sarah Lah Bifeng for help with the optimization of qRT-PCRs. During the course of this work, LC was supported by a travel fellowship from the Galician Government (XXX, Spain), LO by internal funds from the Strategic Research Program of Temaek Life Sciences Laboratory, whereas MB by the National Research Foundation of Hungary **OTKA-77681**

LIMBO:

Scaleless carp (several papers)

Figure legends

Figure 1: Kirpichnikov's "scattered" scale pattern can be further divided into three phenotypes: A-D) Irregular; E-H) Incomplete scaled; and I-L) Classical mirror. For detailed description of phenotyping criteria, please see Supplementary File S3.

Figure 2: Lack of lethality in a cross involving two nude brooders. Common carp eggs were stuck to a nylon mesh by taking advantage of their natural stickiness immediately after fertilization. The meshes were immersed into separate Zuger jars and kept there for for 48 hours. Survival rates were estimated by counting surviving embryos with eye spots versus the opaque ones (empty egg shells). A) Mirror x nude (MN) cross; B) Nude X nude (NN) cross.

Figure 3: The two new sub-types of scattered are inherited to the offspring from irregular or even nude parents and reduce the proportion of mirrors within the scattered group. Panel A) In „Mirror x Irregular" type crosses (MI33 & MI37) the irregular scale pattern was inherited from the parent to the offspring substantially reducing the proportion of mirrors from the expected 100%. The combined proportion of irregular, incomplete scaled and mirror phenotypes are very close to 100%. Panel B) In the first three „Nude X Nude" crosses (NN1, NN2 & NN41) the irregular and incomplete scaled phenotypes appeared among the offspring, resulting in a deviation of the proportion of phenotypes from the expected ratio. In the case of NN26, two European nude individuals were crossed and no lethality was observed. As expected, the proportion of nudes increased in comparison to the other crosses with lethality.

Figure 4: Association between the level of scale loss and the type and number of distorted or missing fins in irregular, incomplete scaled, mirror and nude phenotypes in four families. Panel A) The percentage of distorted/absent fins is shown across the four major phenotypes. Panel B) Fin defects were quantified on a per fish basis (distortion of one fin: 1 point; loss of one fin: 2 points) and the percentage of individuals belonging to each of the four phenotypic categories is shown in relation to the number of defects observed.

Figure 5: The number of pharyngeal teeth gradually decreases with the reduction of scale coverage of the body surface from completely scaled to nudes. The percentage of individuals representing the five phenotypes (from the right: completely scaled, irregular, incomplete scaled, mirror and nude) is plotted against the total number of pharyngeal teeth identified per individual. The lower panel shows a representative picture of the different number of teeth observed (from 10 to 0).

Figure 6: Comparative analysis of the Fgfr1 paralogs of common carp and zebrafish. A) Domain organization of the three Fgfr1 paralogs in common carp in comparison to their two orthologs in zebrafish. Green circles: Immunoglobulin C-2 typedomains; blue rectangle: transmembrane domain and pink hexagon: tyrosine kinase domain. B) Phylogeny and domain architecture of *fgfr1* homologs in cartilaginous and bony fishes reconstructed using a Maximum Parsimony approach. Confidence in the resulting unrooted tree was assessed by

bootstrapping (1000 replicates). Posterior probability values are shown for each branch. Circle denotes the teleost-specific whole-genome duplication event (3R).

Figure 7: Our working hypothesis showing the rheostat-like action of mutations to the signals from the Fgf pathway. XXX

Supplementary Files

Supplementary File S1: Typical representatives of the four major scale pattern phenotypes in common carp, as classified by Kirpichnikov. A) Fully scaled; B) Scattered; C) Linear) and D) Nude individuals.

Supplementary File S2: Distribution of scale phenotypes from XXX different crosses. XXX

Supplementary File S3: Description of our revised scale pattern classification

Supplementary File S4: List of PCR primers used

Supplementary File S5: The relative scale size in nude individuals is significantly smaller than that of the other three phenotypic groups (mirror, incomplete scaled and irregular). The height and width of three largest scales from twenty individuals representing each of the four phenotypes were measured and normalized by taking the standard length of the fish into account. The error bars represent standard deviation of the mean. Columns labeled with different letters indicate statistically significant values (p-value <0.01; Student's t-test).

References

1. Balon EK: Origin and domestication of the wild carp, *Cyprinus carpio* - from Roman gourmets to the swimming flowers. Aquaculture 129: 3-48, 1995.
2. Balon EK: About the oldest domesticates among fishes. Journal of Fish Biology 65: 1-27, 2004.
3. Orban L and Wu QJ: Cyprinids. In: Kole CR and Kocher TD (eds). Genome Mapping and Genomics in Animals: Fishes and Aquatic Animals. Springer Verlag, Berlin, Germany, 2008, pp 45-83.
4. Horvath L, Tamas G and Tolg I: Special methods in pond fish husbandry. Akademiai Kiado and Halver Corporation, Budapest (Hungary) and Seattle (WA, USA), 1984, pp. 1-148.
5. Kohlmann K, Gross R, Murakaeva A and Kersten P: Genetic variability and structure of common carp (*Cyprinus carpio*) populations throughout the distribution range inferred from allozyme, microsatellite and mitochondrial DNA markers. Aquatic Living Resources 16: 421, 2003.
6. Kohlmann K, Kersten P and Flajshans M: Microsatellite-based genetic variability and differentiation of domesticated, wild and feral common carp (*Cyprinus carpio L.*) populations. Aquaculture 247: 253-266, 2005.
7. Zhou J, Wu Q, Wang Z and Ye Y: Molecular phylogeny of three subspecies of common carp Cyprinus carpio, based on sequence analysis of cytochrome b and control region of mtDNA. Journal of Zoological Systematics and Evolutionary Research 42: 266, 2004.
8. Zhou JF, Wang ZW, Ye YZ and Wu QJ: PCR-RFLP analysis of mitochondrial DNA ND5/6 region among 3 subspecies of common carp (Cyprinus carpio L.) and its application to genetic discrimination of subspecies. Chinese Science Bulletin 48: 465, 2003.
9. Kirpichnikov VS: Homologous hereditary variation and evolution of wild common carp *Cyprinus carpio* (in Russian). Genetika (Moscow) 8: 65-72, 1967.
10. Memis D and Kohlmann K: Genetic characterization of wild common carp (Cyprinus carpio L.) from Turkey. 258: 257-262, 2006.
11. Froufe E, Magyary I, Lehoczky I and Weiss S: mtDNA sequence data supports an Asian ancestry and single introduction of common carp into the Danube basin. Journal of Fish Biology 61: 301-304, 2002.
12. Michaels VK: Carp Farming. Fishing News Book, Farnham, Surrey, UK, 1998.
13. Kirpichnikov VS: Genetic basis of fish selection. Springer Verlag, Berlin, Germany, 1981.
14. Kirpichnikov VS: Genetics and breeding of common carp. Hydrobiologie et Aquaculture: 1-97, 1999.
15. Czuczka P: Studies on the inheritance of scales in common carp (*Cyprinus carpio* L.) (in Hungarian. University Diploma; Veszprem University, Georgikon Agricultural School, Keszthely, Hungary. 1-41, 2002.

16. Rudzinsky E: Uber Kreuzungsversuche bei Karpfen I. Fischerei Zeitung 30: 593-597, 1928.

17. Rudzinsky E: Uber Kreuzungsversuche bei Karpfen II. Fischerei Zeitung 31: 613-618, 1928.

18. Kirpichnikov VS and Balkashina EI: Materials on genetics and selection of common carp I. (in Russian). Zoologichesky Journal (Moscow) 14: 45-78, 1935.

19. Kirpichnikov VS and Balkashina EI: Materials on genetics and selection of common carp II. (in Russian). Biologichesky Journal (Moscow) 5: 327-376, 1936.

20. Purdom CE: Genetics and fish breeding. Chapman and Hall, London, UK, 1993.

21. Tave D: Genetics for fish hatchery managers. Springer Verlag, 1993, pp. 1-432.

22. Komen J: Clones of common carp, *Cyprinus carpio*: New perspectives in fish research. In: Wageningen University, Wageningen, The Netherlands, 1990, pp 1-169.

23. Nicolescu C: The phenotypical expression of the common carp's scale cover pattern in accordance with the genotypic structure and the gene dosage. Buletin USAMV-CN 60: 78-83, 2004.

24. Rohner N, Bercsenyi M, Orban L, Kolanczyk ME, Linke D, Brand M, Nusslein-Volhard C and Harris MP: Duplication of *fgfr1* permits Fgf signaling to serve as a target for selection during domestication. Current Biology 19: 1642-7, 2009.

25. Woynarovich E: Hatching carp eggs in Zuger glasses and breeding of carp larvae until the age of 10 days. Israeli Journal of Aquaculture - Bamidgeh 4: 38-46, 1962.

26. Sambrook J and Russell D: Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2001, pp. 1-2344.

27. Rozen S and Skaletsky H: Primer3 on the WWW for general users and for biologist programmers. Methods in molecular biology 132: 365-86, 2000.

28. Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ: Basic local alignment search tool. J Mol Biol 215: 403-10, 1990.

29. Han MV and Zmasek CM: phyloXML: XML for evolutionary biology and comparative genomics. BMC Bioinformatics 10: 356, 2009.

30. Tamura K, Dudley J, Nei M and Kumar S: MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596-9, 2007.

31. David L, Blum S, Feldman MW, Lavi U and Hillel J: Recent duplication of the common carp (*Cyprinus carpio* L.) genome as revealed by analyses of microsatellite loci. Molecular Biology and Evolution 20: 1425, 2003.

32. Larhammar D and Risinger C: Molecular genetic aspects of tetraploidy in the common carp *Cyprinus carpio*. Molecular Phylogenetics and Evolution 3: 59-68, 1994.

33. Benton MJ and Donoghue PC: Paleontological evidence to date the tree of life. Molecular Biology and Evolution 24: 26-53, 2007.

34. Christoffels A, Koh EG, Chia JM, Brenner S, Aparicio S and Venkatesh B: Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. Molecular Biology and Evolution 21: 1146-51, 2004.

35. Meyer A and Van de Peer Y: From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). BioEssays 27: 937-45, 2005.

36. Force A, Lynch M, Pickett FB, Amores A, Yan YL and Postlethwait J: Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151: 1531-1545, 1999.

37. Loh YH, Christoffels A, Brenner S, Hunziker W and Venkatesh B: Extensive expansion of the claudin gene family in the teleost fish, Fugu rubripes. Genome Research 14: 1248-57, 2004.

38. Gomelsky B, Schneider KJ, Glennon RP and Plouffe DA: Effect of ploidy on scale-cover pattern in linear ornamental (koi) common carp *Cyprinus carpio*. Journal of Fish Biology 81: 1201-9, 2012.

39. Gomelsky BI, Emelyanova OV and Recoubratsky AV: Application of the scale cover gene (N) to identification of type of gynogenesis and determination of ploidy in common carp. Aquaculture 106: 233-237, 1992.

40. Gibert Y, Bernard L, Debiais-Thibaud M, Bourrat F, Joly JS, Pottin K, Meyer A, Retaux S, Stock DW, Jackman WR, Seritrakul P, Begemann G and Laudet V: Formation of oral and pharyngeal dentition in teleosts depends on differential recruitment of retinoic acid signaling. FASEB Journal 24: 3298-309, 2010.

41. Goldfarb M: Fibroblast growth factor homologous factors: evolution, structure, and function. Cytokine & Growth Factor Reviews 16: 215-20, 2005.

42. Coulier F, Pontarotti P, Roubin R, Hartung H, Goldfarb M and Birnbaum D: Of worms and men: an evolutionary perspective on the fibroblast growth factor (FGF) and FGF receptor families. Journal of Molecular Evolution 44: 43-56, 1997.

43. Wilkie AO: Bad bones, absent smell, selfish testes: the pleiotropic consequences of human FGF receptor mutations. Cytokine & Growth Factor Reviews 16: 187-203, 2005.

44. Kondo S, Kuwahara Y, Kondo M, Naruse K, Mitani H, Wakamatsu Y, Ozato K, Asakawa S, Shimizu N and Shima A: The medaka rs-3 locus required for scale development encodes ectodysplasin-A receptor. Current Biology 11: 1202-6, 2001.

45. Abe G, Ide H and Tamura K: Function of FGF signaling in the developmental process of the median fin fold in zebrafish. Developmental Biology 304: 355-66, 2007.

46. Prykhozhij SV and Neumann CJ: Distinct roles of Shh and Fgf signaling in regulating cell proliferation during zebrafish pectoral fin development. BMC developmental biology 8: 91, 2008.

47. Ma EY and Raible DW: Signaling pathways regulating zebrafish lateral line development. Current Biology 19: R381-6, 2009.

48. Gallardo VE, Liang J, Behra M, Elkahloun A, Villablanca EJ, Russo V, Allende ML and Burgess SM: Molecular dissection of the migrating posterior lateral line primordium during early development in zebrafish. BMC Developmental Biology 10: 120, 2010.

49. Sarrazin AF, Nunez VA, Sapede D, Tassin V, Dambly-Chaudiere C and Ghysen A: Origin and early development of the posterior lateral line system of zebrafish. The Journal of Neuroscience 30: 8234-44, 2010.

50. Poss KD, Shen J, Nechiporuk A, McMahon G, Thisse B, Thisse C and Keating MT: Roles for Fgf signaling during zebrafish fin regeneration. Developmental Biology 222: 347-58, 2000.

51. Stoick-Cooper CL, Moon RT and Weidinger G: Advances in signaling in vertebrate regeneration as a prelude to regenerative medicine. Genes & Development 21: 1292-315, 2007.

52. Diggles BK: Saddleback deformities in yellowfin bream, Acanthopagrus australis (Gunther), from South East Queensland. Journal of Fish Diseases: 2012.

53. Morrison C, Cornick J, Shum J and Zwicker B: Microbiology and histopathology of 'saddleback' disease of underyearling Atlantic salmon, *Salmo salar* L. Jourbnal of Fish Disease 4: 243-258, 1981.

54. Koumoundouros G: First record of saddleback syndrome in wild parrotfish *Sparisoma cretense* (L., 1758) (Perciformes, Scaridae). Journal of Fish Biology 72: 737-741, 2008.

55. Setiadi E, Tsumura S, Kassam D and Yamaoka K: Effect of saddleback syndrome and vertebral deformity on the body shape and size in hatchery-reared juvenile red spotted grouper, *Epinephelus akaara* (Perciformes: Serranidae): a geometric morphometric approach. Journal of Applied Ichthyology 22: 49-53, 2006.

56. Tave D, Bartels JE and Smitherman RO: Saddleback: a dominant, lethal gene in *Sarotherodon aureus* (Steindachner) (= *Tilapia aurea*). Journal of Fish Diseases 6: 59-73, 1983.

57. Bercsényi M, Szűcs R, Orbán L, Lehoczky I and Jeney Z: Research on fish genetics for the purposes of strain improvement and population protection (in Hungarian). Állattenyésztés és Takarmányozás 60: 267-279, 2011.

58. Ortubay S and Cussac V: Threatened fishes of the world: *Gymnocharacinus bergi*, Steindacher, 1903 (Characidae). Environmental Biology of Fishes 58: 144, 2000.

59. Haworth KE, Wilson JM, Grevellec A, Cobourne MT, Healy C, Helms JA, Sharpe PT and Tucker AS: Sonic hedgehog in the pharyngeal endoderm controls arch pattern via regulation of Fgf8 in head ectoderm. Developmental Biology 303: 244-58, 2007.

60. Zhang X, Zhang Y, Zheng X, Kuang Y, Zhao Z, Zhao L, Li C, Jiang L, Cao D, Lu C, Xu P and Sun X: A consensus linkage map provides insights on genome character and evolution in common carp (*Cyprinus carpio* L.). Marine Biotechnology, in press: 2012.

61. Zheng X, Kuang Y, Zhang X, Lu C, Cao D, Li C and Sun X: A genetic linkage map and comparative genome analysis of common carp (*Cyprinus carpio* L.) using microsatellites and SNPs. Molecular Genetics and Genomics 286: 261-77, 2011.

62. Cheng L, Liu L, Yu X, Wang D and Tong J: A linkage map of common carp (*Cyprinus carpio*) based on AFLP and microsatellite markers. Animal Genetics 41: 191-8, 2010.

63. Christoffels A, Bartfai R, Srinivasan H, Komen H and Orban L: Comparative genomics in cyprinids: common carp ESTs help the annotation of the zebrafish genome. BMC Bioinformatics 7 Suppl 5: S2, 2006.

64. Ji P, Liu G, Xu J, Wang X, Li J, Zhao Z, Zhang X, Zhang Y, Xu P and Sun X: Characterization of common carp transcriptome: sequencing, de novo assembly, annotation and comparative genomics. PloS one 7: e35152, 2012.

65. Henkel CV, Dirks RP, Jansen HJ, Forlenza M, Wiegertjes GF, Howe K, van den Thillart GE and Spaink HP: Comparison of the exomes of common carp *(Cyprinus carpio)* and zebrafish *(Danio rerio)*. Zebrafish 9: 59-67, 2012.

66. Williams DR, Li W, Hughes MA, Gonzalez SF, Vernon C, Vidal MC, Jeney Z, Jeney G, Dixon P, McAndrew B, Bartfai R, Orban L, Trudeau V, Rogers J, Matthews L, Fraser EJ, Gracey AY and Cossins AR: Genomic resources and microarrays for the common carp *Cyprinus carpio* L. Journal of Fish Biology 72: 2095-2117, 2008.

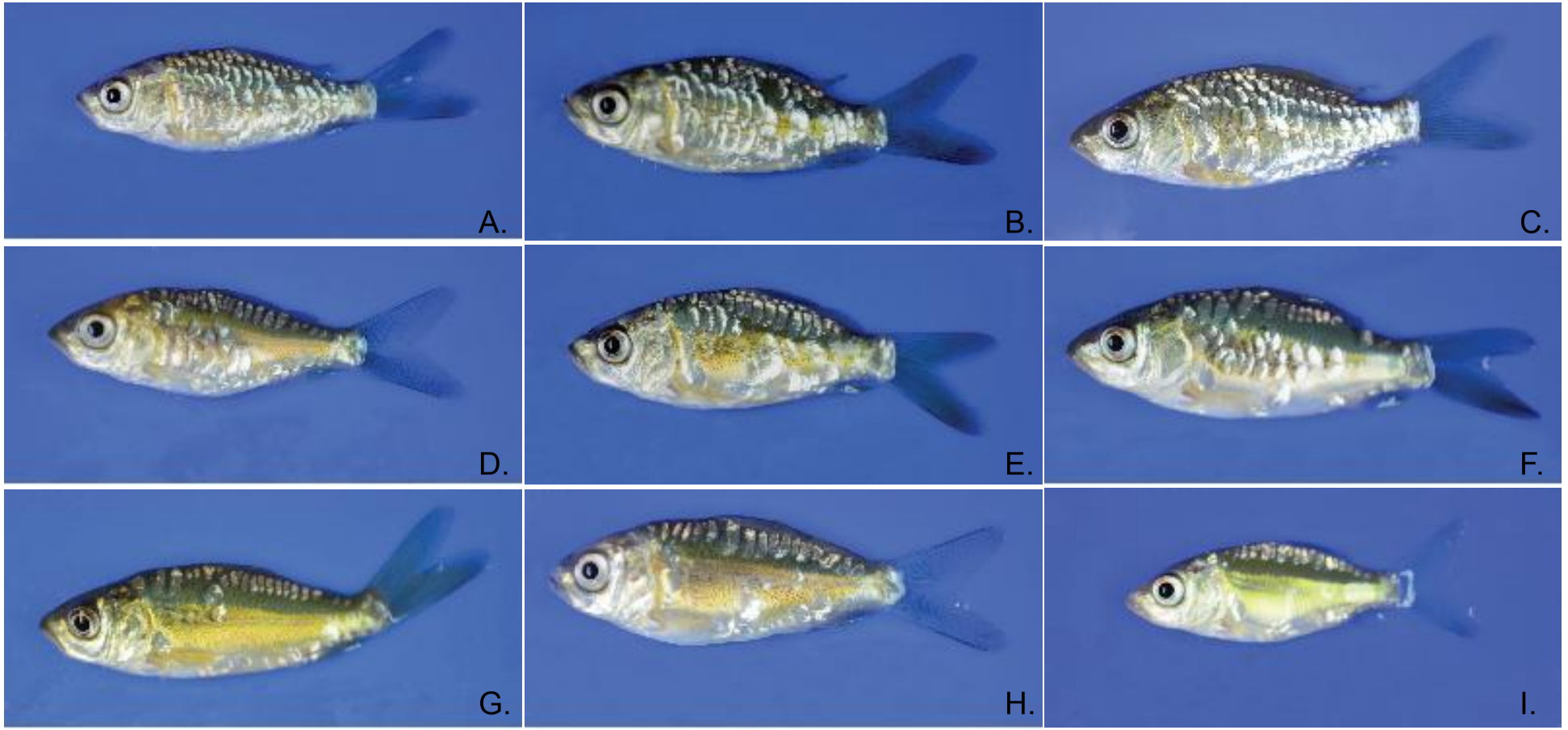


Figure 1. Kirpichnikov's "scattered" scale pattern can be further divided into three phenotypes: A-C: Irregular; D-F: Incomplete scaled and G-I: Classical mirror.

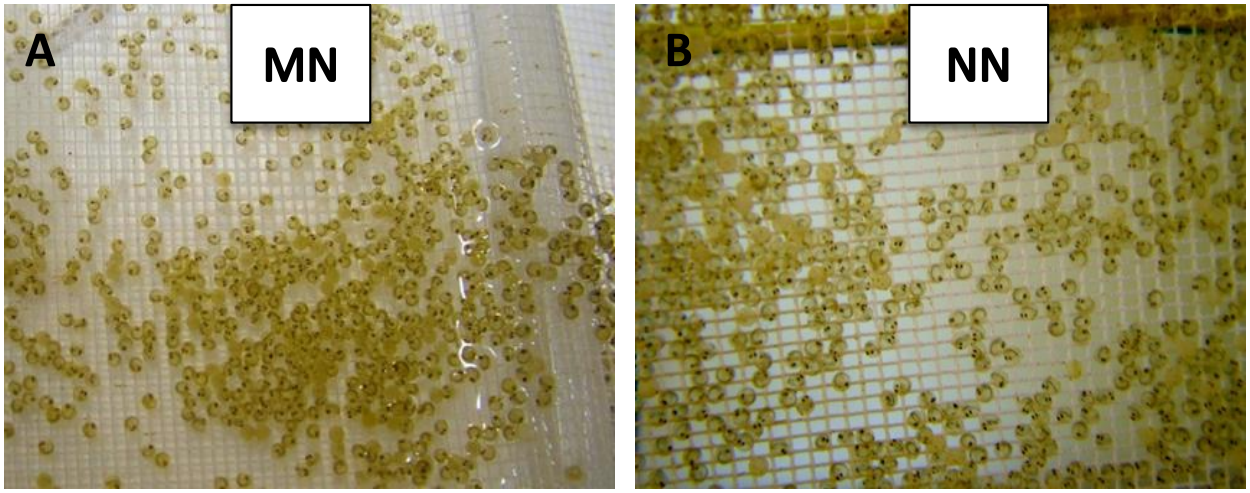


Figure 2: Lack of lethality in a cross involving two nude brooders. Common carp eggs were stuck to a nylon mesh by taking advantage of their natural stickiness immediately after fertilization. The meshes were immersed into separate Zuger jars and kept there for for 48 hours. Survival rates were estimated by counting surviving embryos with eye spots versus the opaque ones (empty egg shells). A) Mirror x nude (MN) cross; B) Nude X nude (NN) cross.

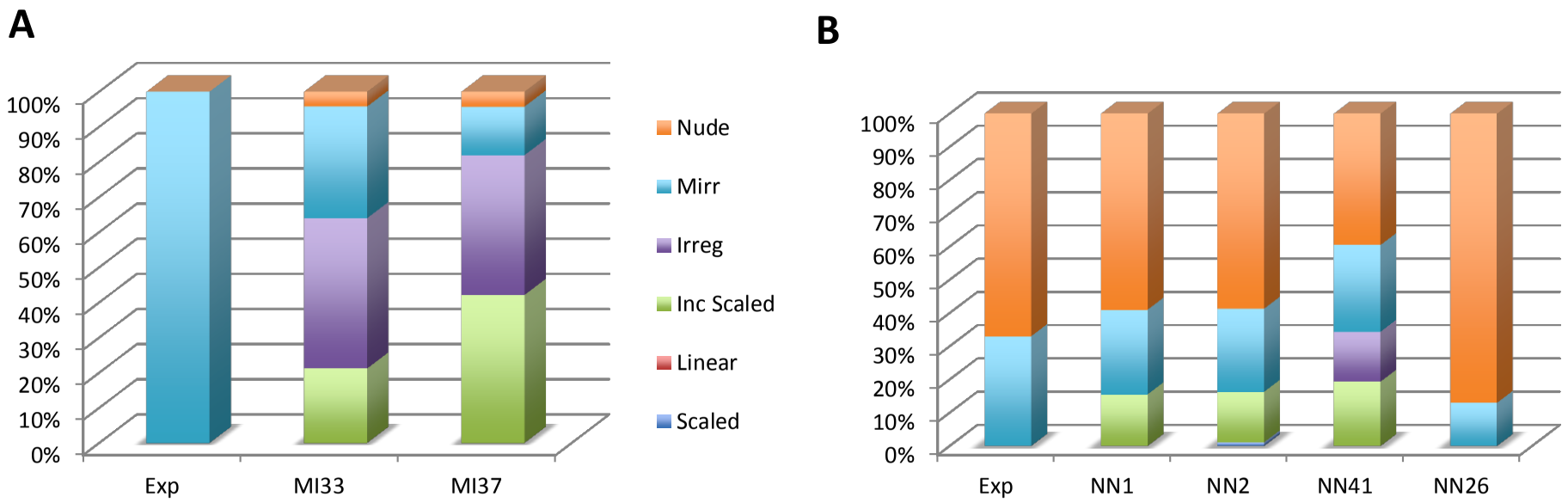
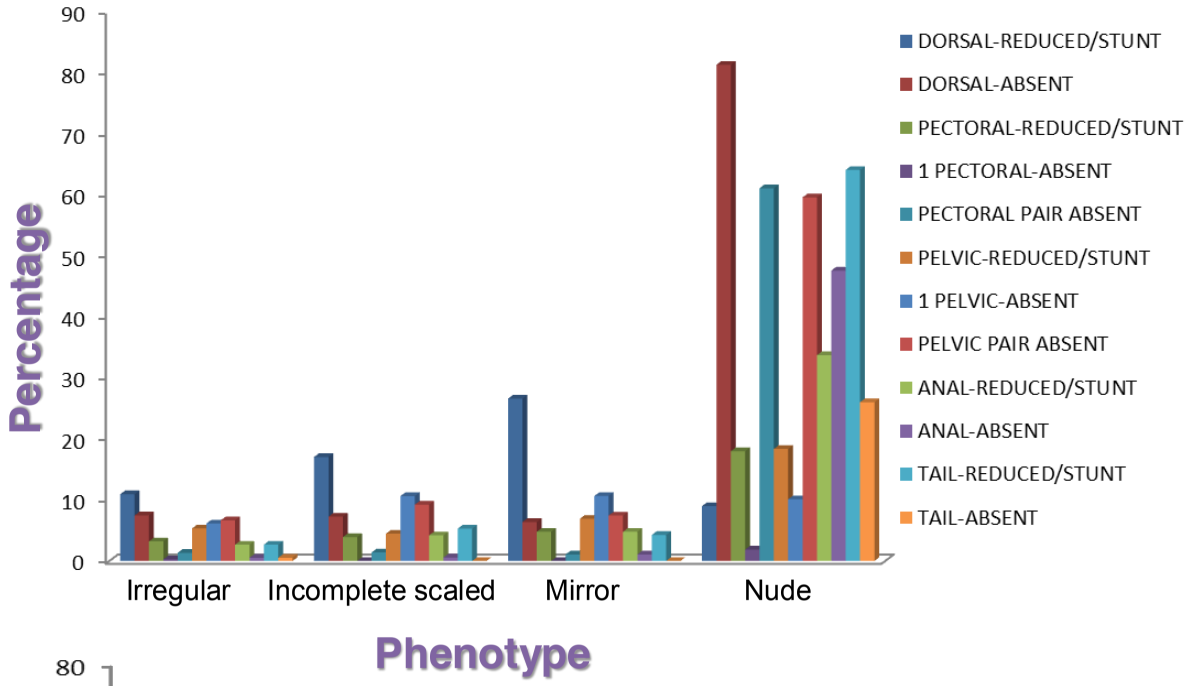


Figure 3: The two new sub-types of scattered are inherited to the offspring from irregular or even nude parents and reduce the proportion of mirrors within the scattered group. Panel A) In „Mirror x Irregular" type crosses (MI33 & MI37) the irregular scale pattern was inherited from the parent to the offspring substantially reducing the proportion of mirrors from the expected 100%. The combined proportion of irregular, incomplete scaled and mirror phenotypes are very close to 100%. Panel B) In the first three „Nude X Nude" crosses (NN1, NN2 & NN41) the irregular and incomplete scaled phenotypes appeared among the offspring, resulting in a deviation of the proportion of phenotypes from the expected ratio. In the case of NN26, two European nude individuals were crossed and no lethality was observed. As expected, the proportion of nudes increased in comparison to the other crosses with lethality.

A



B

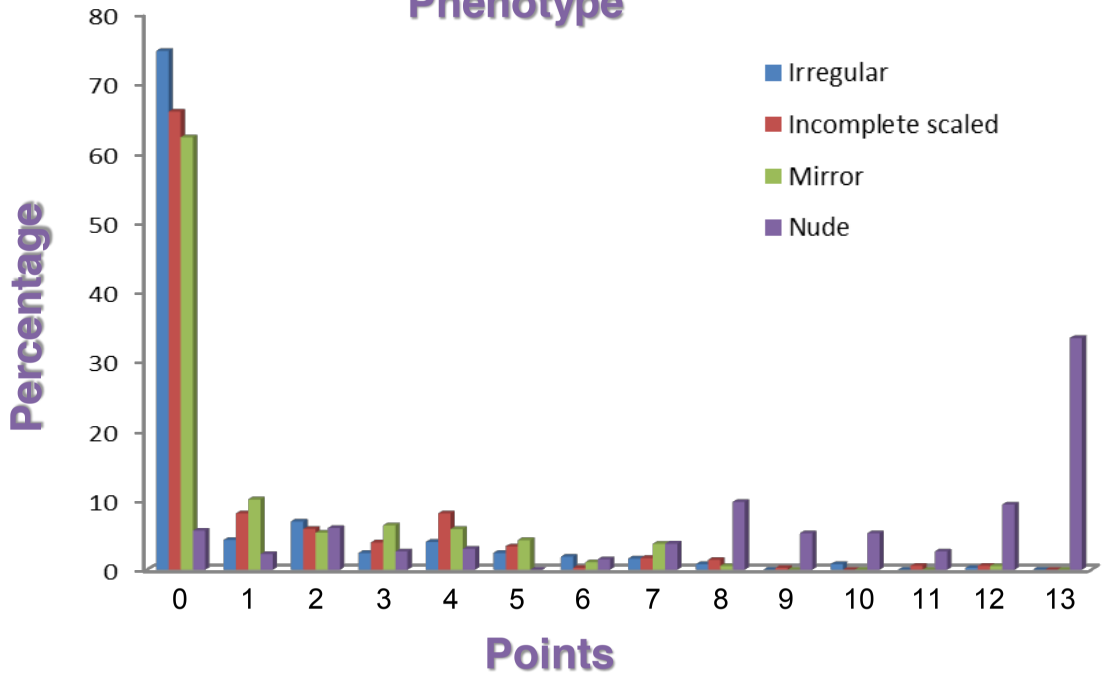


Figure 4. Association between the level of scale loss and the type and number of distorted or missing fins in irregular, incomplete scaled, mirror and nude phenotypes in four families. Panel A) The percentage of distorted/absent fins is shown across the four major phenotypes. Panel B) Fin defects were quantified on a per fish basis (distortion of one fin: 1 point; loss of one fin: 2 points) and the percentage of individuals belonging to each of the four phenotypic categories is shown in relation to the number of defects observed.

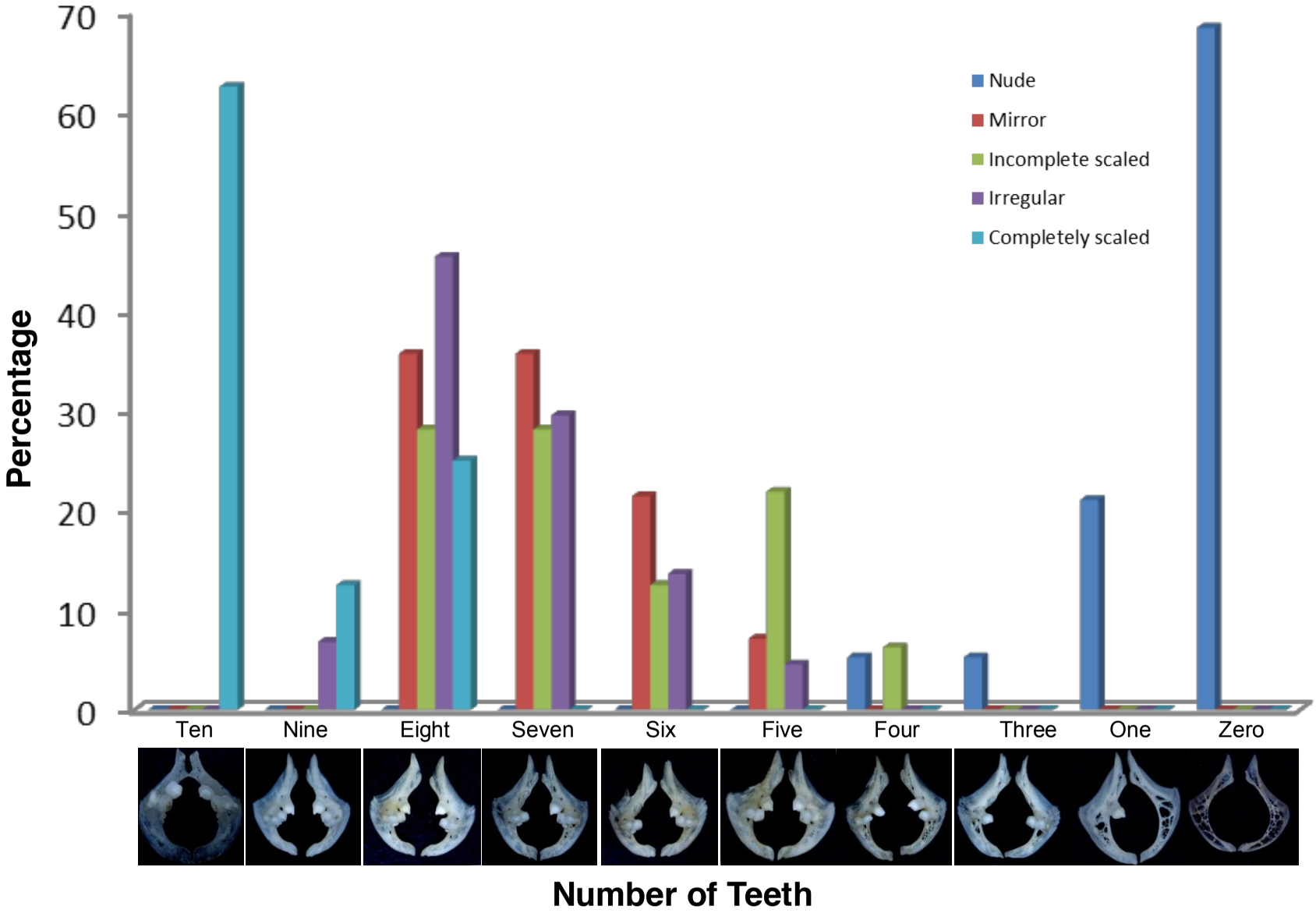
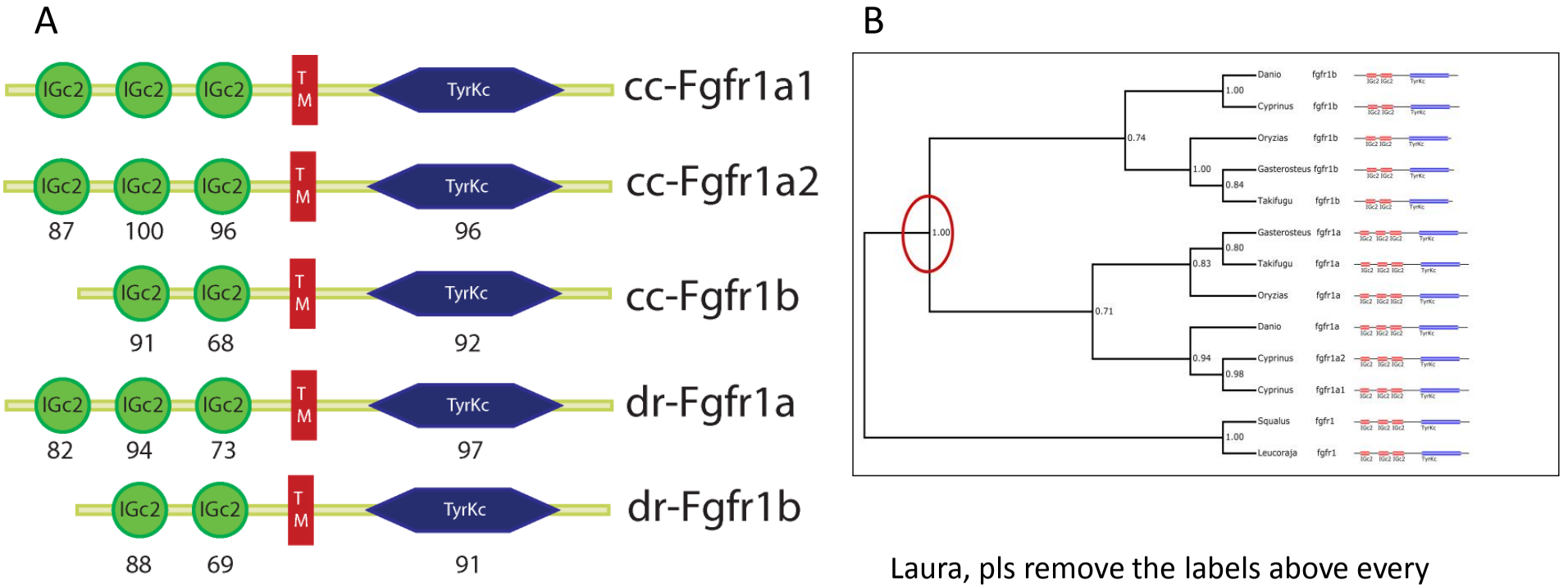


Figure 5: The number of pharyngeal teeth gradually decreases with the reduction of scale coverage of the body surface from completely scaled to nudes. The percentage of individuals representing the five phenotypes (from the right: completely scaled, irregular, incomplete scaled, mirror and nude) is plotted against the total number of pharyngeal teeth identified per individual. The lower panel shows a representative picture of the different number of teeth observed (from 10 to 0).



Laura, pls remove the labels above every domain, as I have already indicated them in the legend.

Figure 6: Comparative analysis of the Fgfr1 paralogs of common carp and zebrafish. A) Domain organization of the three Fgfr1 paralogs in common carp in comparison to their two orthologs in zebrafish. Green circles: Immunoglobulin C-2 typedomains; blue rectangle: transmembrane domain and pink hexagon: tyrosine kinase domain. B) Phylogeny and domain architecture of *fgfr1* homologs in cartilaginous and bony fishes reconstructed using a Maximum Parsimony approach. Confidence in the resulting unrooted tree was assessed by bootstrapping (1000 replicates). Posterior probability values are shown for each branch. Circle denotes the teleost-specific whole-genome duplication event (3R).

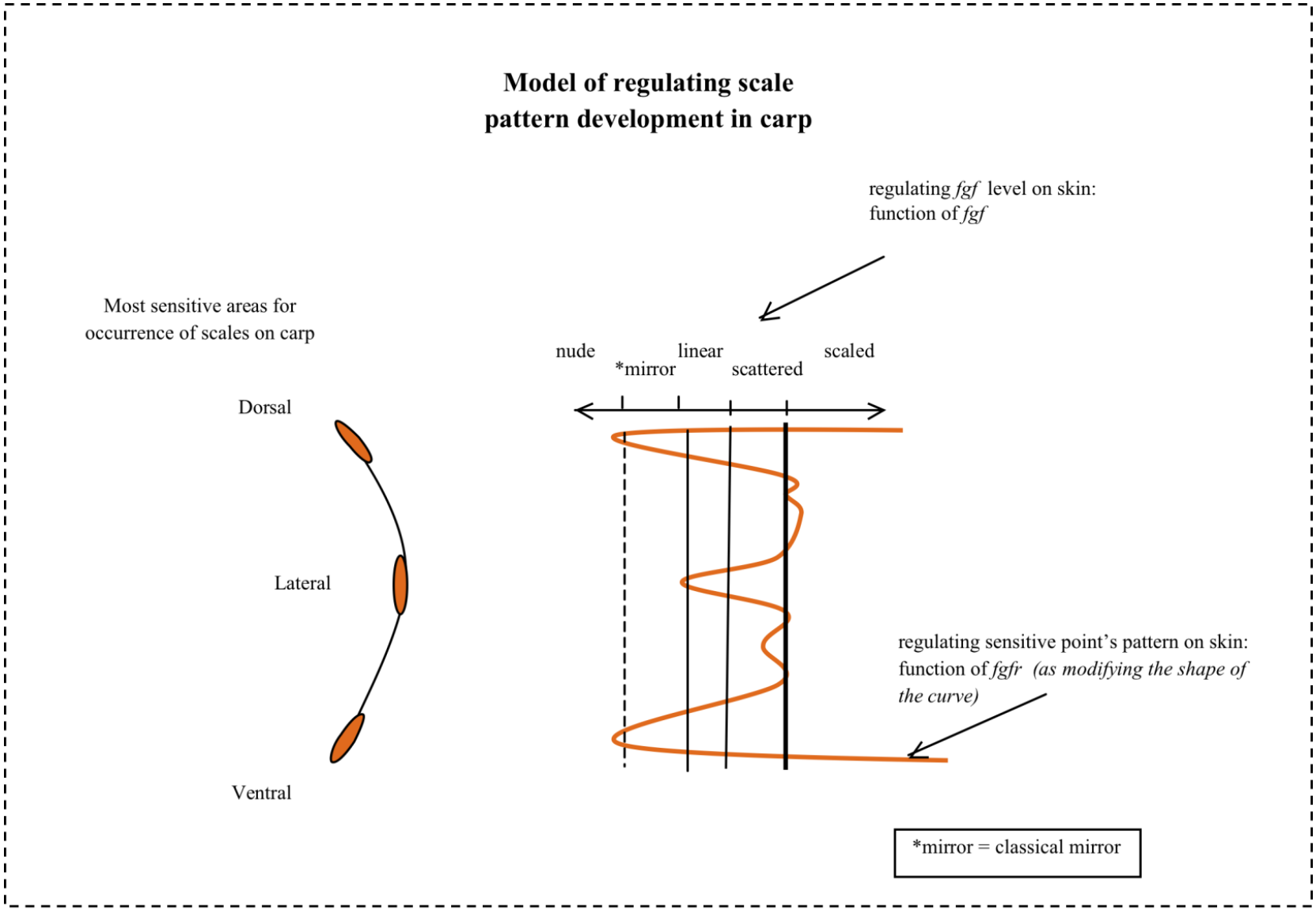
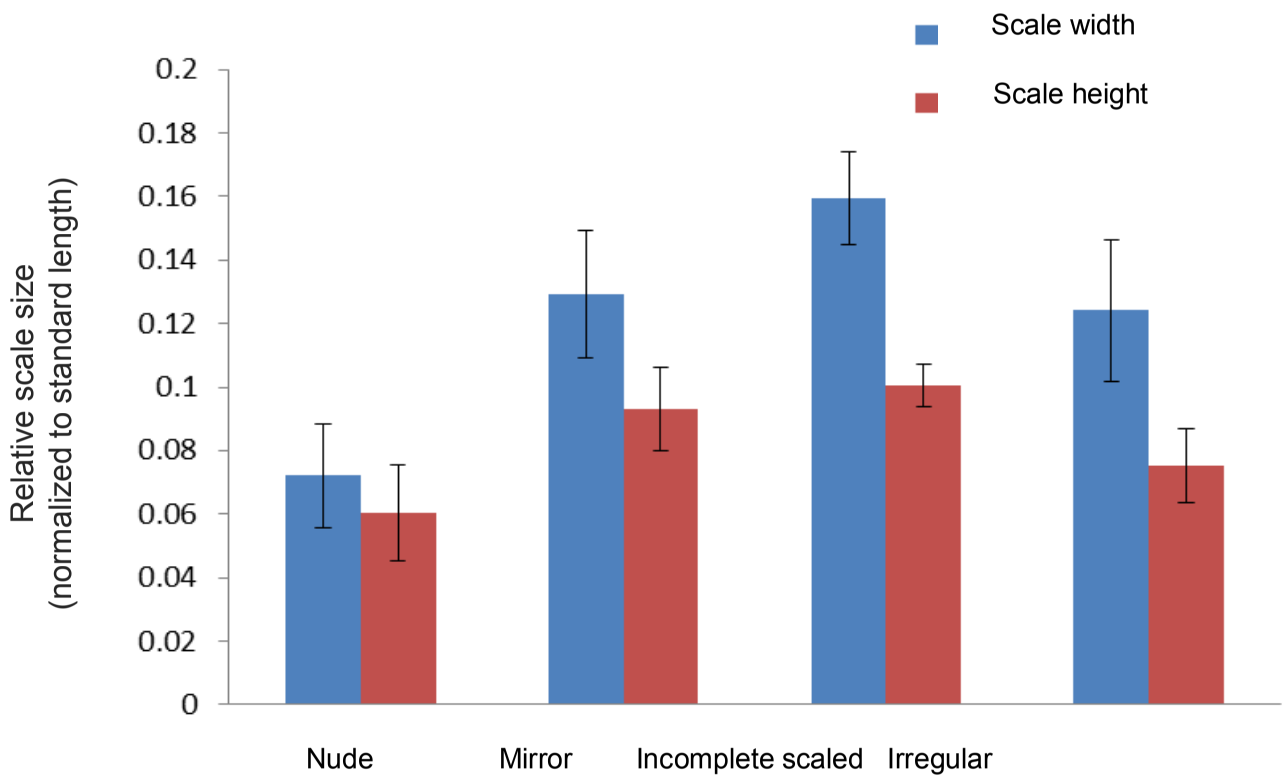


Figure 7: Our working hypothesis showing the rheostat-like action of mutations to the signals from the Fgf pathway. XXX (Both graph and legend need improvements)

SUPPLEMENTARY FIGS

Supplementary File S1. Typical representatives of the four major scale pattern phenotypes in common carp, as classified by Kirpichnikov. A) Fully scaled; B) Scattered; C) Linear) and D) Nude individuals.



Supplementary File S4: The relative scale size in nude individuals is significantly smaller than that of the other three phenotypic groups (mirror, incomplete scaled and irregular). The height and width of three largest scales from twenty individuals representing each of the four phenotypes were measured and normalized by taking the standard length of the fish into account. The error bars represent standard deviation of the mean. Columns labeled with different letters indicate statistically significant values (p-value <0.01; Student's t-test).

LIMBO

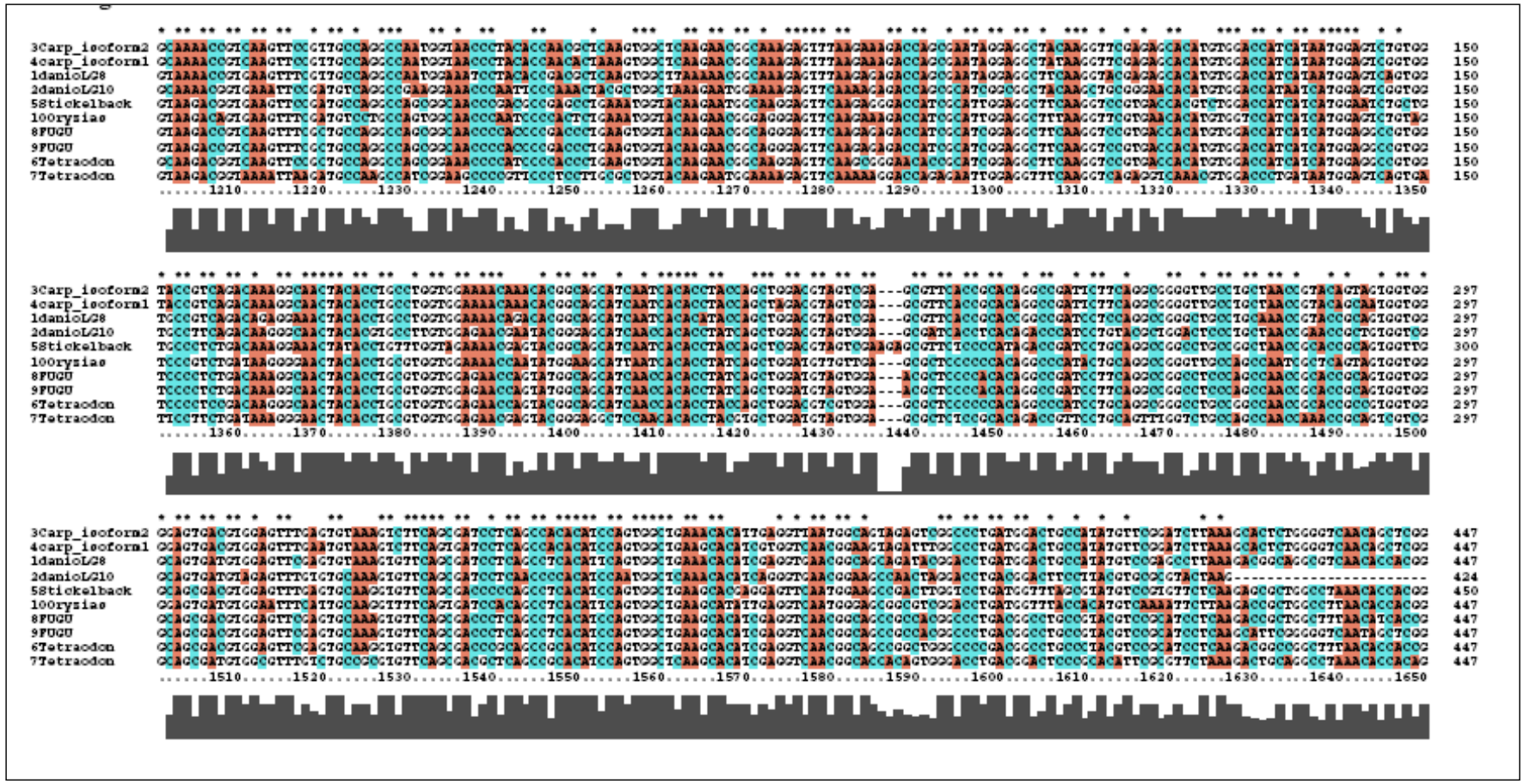
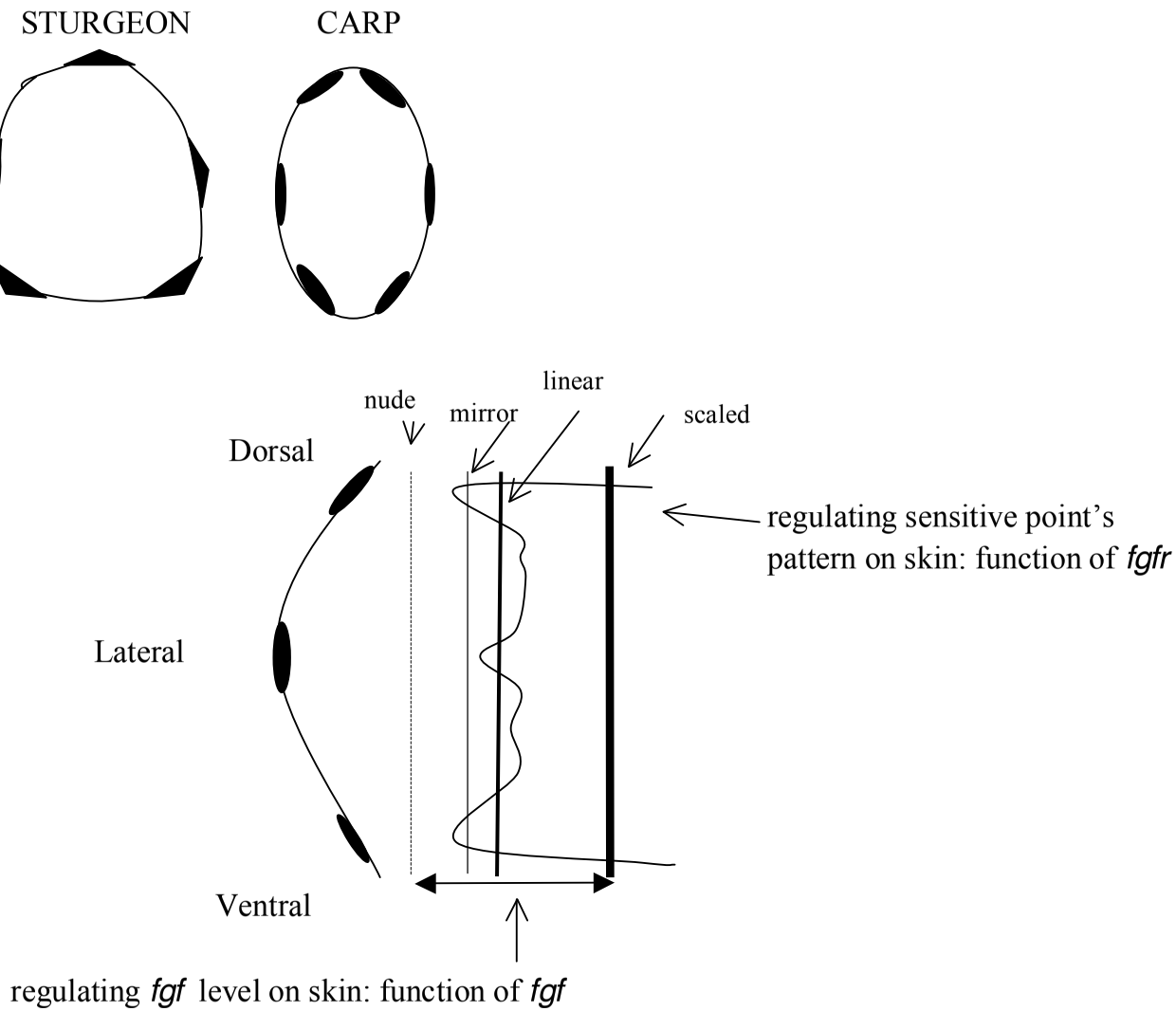


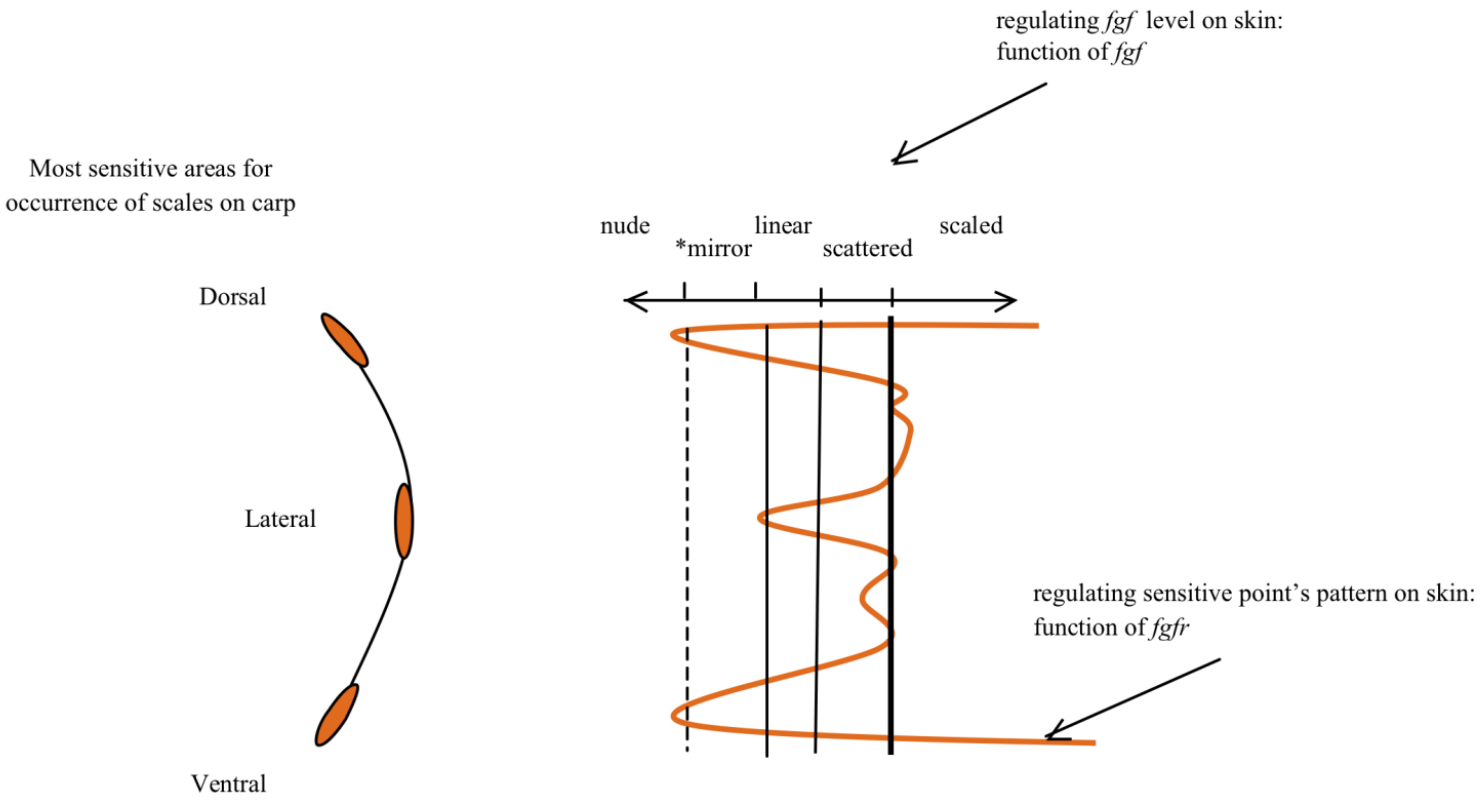
Figure XXX: ClustalX alignment of *fgfr1* sequences in carp-*fgfr1a2*, carp-*fgfr1a1*, zebrafish-LG8, zebrafish LG10, stickleback, Oryzias, Fugu-iso 1, fugu iso-2, Tetraodon-iso 1, Tetraodon iso-2 from top to bottom.

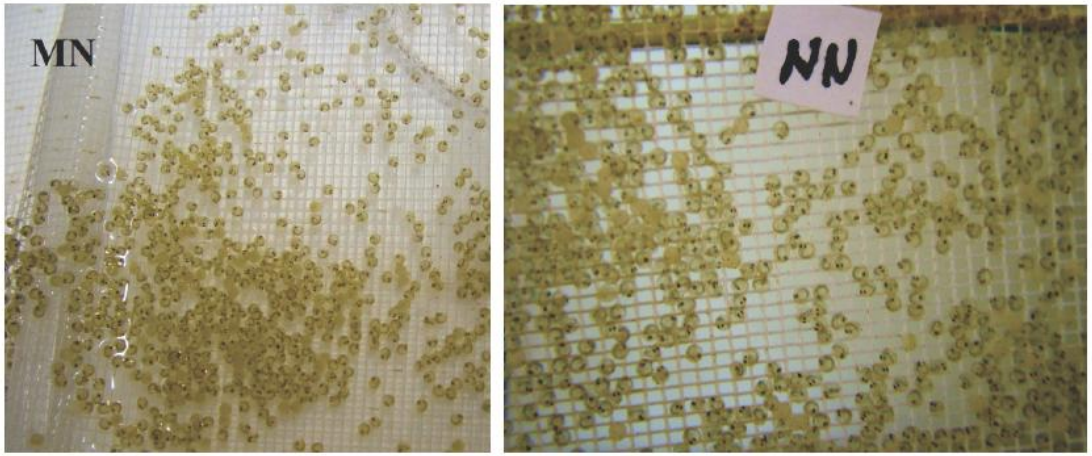


Prof Bercsenyi will re-draw this figure

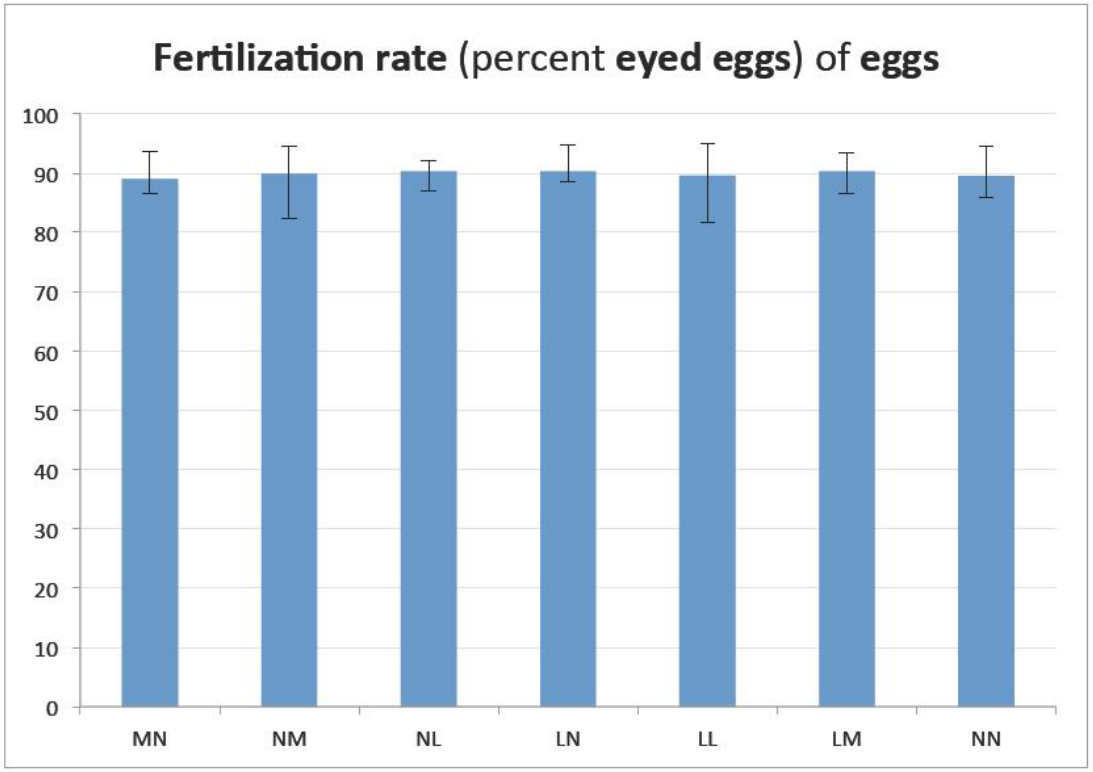
Figure 8: Our working hypothesis showing the rheostat-like action of mutations to the signals from the Fgf pathway. XXX (Both graph and legend need improvements)

Model of regulating scale
pattern development in carp

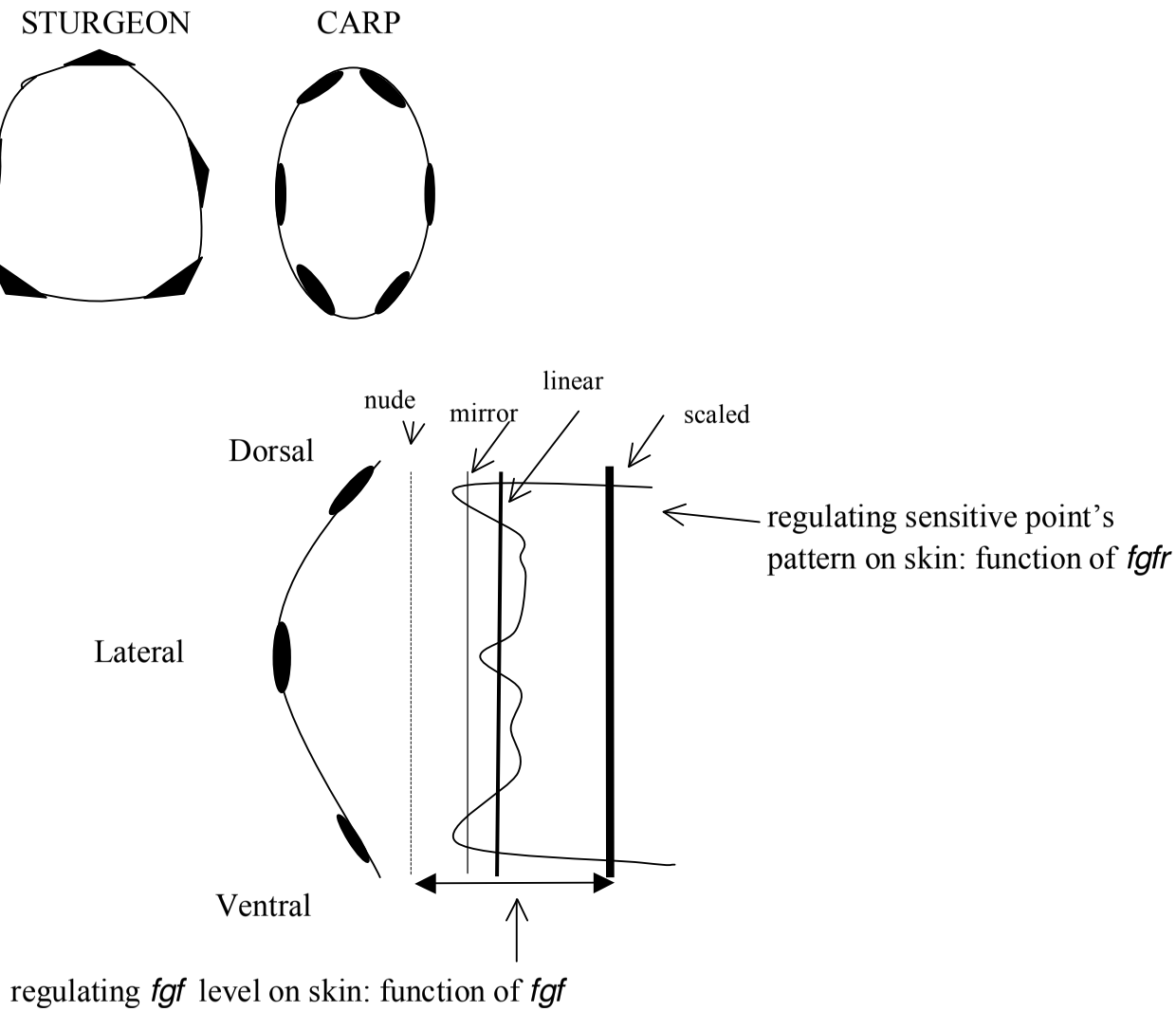




XXX Pic. .... Mirror X Nude



XXX . Figure ...



Prof Bercsenyi will re-draw this figure

Figure 8: Our working hypothesis showing the rheostat-like action of mutations to the signals from the Fgf pathway. XXX (Both graph and legend need improvements)

Cross	Location	M_Pheno	M_Orig	Fe_Pheno	Fe_Orig	F1_No	Lethality
LL7	HUN	Linear	Eur	Linear	Eur	97	N/A
LL8	HUN	Linear	Eur	Linear	Eur	21	N/A
LL24	HUN	Linear	Eur	Linear	Eur	234	N/A
LM23	HUN	Mirror	Eur	Linear	Eur	204	N/A
MI33	SIN	Mirror	F1hyb	Irregular	F1hyb	236	N/A
MI37	SIN	Mirror	F1hyb	Irregular	F1hyb	304	N/A
NL27	HUN	Linear	Eur	Nude	Eur	289	25%
LN10	HUN	Nude	Eur	Linear	Eur	47	25%
LN22	HUN	Nude	Eur	Linear	Eur	118	25%
IN39	SIN	Irregular	F1hyb	Nude	F1hyb	54	N/A
MN36	SIN	Mirror	F1hyb	Nude	Koi	177	N/A
MN21	HUN	Nude	Eur	Mirror	Eur	186	N/A
NN1	SIN	Nude	Eur	Nude	Koi	161	25%
NN2	SIN	Nude	Koi	Nude	Koi	92	25%
NN26	HUN	Nude	Eur	Nude	Eur	208	N/A
NN35	SIN	Nude(gp)	Eur	Nude	Koi	218	25%
NN41	SIN	Nude(gp)	Eur	Nude	Koi	253	25%

Expected%				Observed %					
Scaled %	Linear%	Mirror%	Nude%	Scaled	Linear	Inc Scaled	Irreg	Mirr	Nude
19-33%	56-67%	0-6%	0-19%	0%	69%	15%	5%	10%	0%
19-33%	56-67%	0-6%	0-19%	0%	95%	5%	0%	0%	0%
19-33%	56-67%	0-6%	0-19%	0%	32%	21%	14%	31%	2%
0-50%	50%	0-50%	0%	0%	33%	0%	0%	67%	0%
		100%		0%	0%	21%	43%	32%	4%
		100%		0%	0%	42%	40%	14%	4%
0-33%	33-67%	0-33%	0-33%	0%	1%	0%	0%	99%	0%
0-33%	33-67%	0-33%	0-33%	0%	9%	0%	0%	89%	2%
0-33%	33-67%	0-33%	0-33%	0%	3%	0%	0%	88%	9%
		50%	50%	0%	0%	15%	30%	19%	37%
		50%	50%	0%	0%	51%	22%	2%	25%
		50%	50%	0%	1%	0%	0%	96%	3%
		33%	67%	0%	0%	16%	0%	25%	59%
		33%	67%	1%	0%	15%	0%	25%	59%
		25%	75%	0%	0%	0%	0%	13%	87%
		33%	67%	0%	0%	26%	27%	1%	46%
		33%	67%	0%	0%	19%	15%	26%	40%

Classifying common carps based on their scale pattern

(an extended version of Kirpichnikov's system)

Scaled (Sc): The whole body is covered with regularly arranged scales. Every scale partially covers the one located behind it.

Iregular (Ir): Large portion of the body surface is covered with large (presumably fused) scales. The scales do not overlap, often they do not even reacgh each other, leaving the skin exposed among them. Occasionally, a more or less complete line of scales can be found over the lateral line.

Incomplete scaled (earlier 2/3 mirror or M+; ISc): All individuals lacking scales over at least 33% of their body surface should be placed into this group. Occasionally, a more or less complete line of scales can be found over the lateral line.

Linear (Li): The line of scales is clearly defined, consisting uniform scales of normal size. The line might be incomplete. In addition, a lesser number (<10) scales can be found scattered over the body surface.

Mirror (Mi): All the fins are intact. The anal fin has five rays. There is a row of scales (sometimes incomplete) below the dorsal find, and occasionally another row above the belly (could also be incomplete). In addition to these, there might be other scales scattered over the body, especially the in the tail region. There is no uniform line of scales over the lateral line and the majority of the body surface (>90%) is scaleless.

Nude (Nu): The individuals must be classified as a nude, if the phenotype is similar to that of the mirror, but one of the following criteria is fulfilled:

1) There is no scale on the body surface;
2) The scale line below the dorsal fin is missing and there are less than three scales on the body surface;
3) There are less then five rays on the anal fin;
4) The isolated pharyngeal arches have less than three teeth in total;
5) At least three fins are severely degraded or missing.

Primer name	**Purpose**
fgfr1b_1F	Amplification of carp *fgfr1b* cDNA fragment
fgfr1b_2R	Amplification of carp *fgfr1b* cDNA fragment
	5'RACE gene-specific *fgfr1b*
	5'RACE gene-specific inner primer *fgfr1b*
	5'RACE gene-specific Outer primer *fgfr1b*
	3’RACE gene-specific Outer Primer *fgfr1b*
	3’RACE gene-specific Inner Primer *fgfr1b*
fgfr1b_3F	Amplification of carp *fgfr1b* cDNA (coding, full)
fgfr1b_4R	Amplification of carp *fgfr1b* cDNA (coding, full)

Sequence (5'-3')

GGAGCATCAATCACACCTATCA
AAGTTTGCTTCCATTCACCAGT
AGCATCCTCAAAGGACACATTC
GATGGCACCTGAGGCTTTGTTT
ATCCAGGAGTGCCWGTGGAAGA