American Genetic Association

Journal of Heredity, 2019, 194–210
doi:10.1093/jhered/esy071
Original Article
Advance Access publication December 31, 2018

OXFORD

Original Article

Two Proximally Close Priority Candidate Genes for *diplopodia-1*, an Autosomal Inherited Craniofacial-Limb Syndrome in the Chicken: *MRE11* and *GPR83*

Elizabeth A. O'Hare, Parker B. Antin, and Mary E. Delany

From the Department of Animal Science, University of California, Davis, CA 95616 (O'Hare and Delany); and the Department of Molecular and Cellular Medicine, University of Arizona, Tucson, AZ 85724 (Antin). Elizabeth A. O'Hare is now at the Department of Biological Sciences, Towson University, Towson, MD 21252.

Address correpondence to Mary E. Delany at the address above, or e-mail: medelany@ucdavis.edu

Received April 11, 2018; First decision June 13, 2018; Accepted December 29, 2018.

Corresponding Editor: Susan J. Lamont

Abstract

Next-generation sequencing (NGS) and expression technologies were utilized to investigate the genes and sequence elements in a 586 kb region of chicken chromosome 1 associated with the autosomal recessive *diplopodia-1* (*dp-1*) mutation. This mutation shows a syndromic phenotype similar to known human developmental abnormalities (e.g., cleft palate, polydactyly, omphalocele [exposed viscera]). Toward our goal to ascertain the variant responsible, the entire 586 kb region was sequenced following utilization of a specifically designed capture array and to confirm/validate fine-mapping results. Bioinformatic analyses identified a total of 6142 sequence variants, which included SNPs, indels, and gaps. Of these, 778 SNPs, 146 micro-indels, and 581 gaps were unique to the UCD-Dp-1.003 inbred congenic line; those found within exons and splice sites were studied for contribution to the mutant phenotype. Upon further validation with additional mutant samples, a smaller subset (of variants [51]) remains linked to the mutation. Additionally, utilization of specific samples in the NGS technology was advantageous in that fine-mapping methodologies eliminated an additional 326 kb of sequence information on chromosome 1. Predicted and confirmed protein-coding genes within the smaller 260 kb region were assessed for their developmental expression patterns over several stages of early embryogenesis in regions/tissues of interest (e.g., digits, craniofacial region). Based on these results and known function in other vertebrates, 2 genes within 5 kb of each other, *MRE11* and *GPR83*, are proposed as high-priority candidates for the *dp-1* mutation.

Subject area: Genomics and gene mapping
Keywords: capture array, congenital malformations, developmental mutation, next-generation sequencing, vertebrate development

Introduction

Animal models contribute enormously to our understanding of the genetic and molecular basis of human diseases and disorders. The chicken embryo offers a unique advantage to the study of vertebrate developmental biology, due to the many naturally occurring mutants that exhibit developmental defects (Pisenti et al. 1999; Delany 2004; Robb et al. 2011; Robb and Delany 2012a) as well as the benefit of *in ovo* embryogenesis allowing for easy access, assessment, and manipulation. The many benefits of utilizing the chick embryo as a model system for human health-related research has been described in both primary research articles and reviews (Schock et al. 2016; Davey et al. 2018; Stern 2018; The Chick Embryo Model System 2018). One area of biomedical research that makes excellent use of the chick embryo model system is the study of inherited congenital malformations (Burt 2007; Davey and Tickle 2007; Schock et al. 2016; Davey et al. 2018). The inherited mutations in the chicken contribute new knowledge regarding the etiology and pathways underpinning complex syndromes involving craniofacial, limb and organ malformations (Davey et al. 2006; Schock et al. 2016). Diplopodia is a congenital malformation occurring in tetrapods involving duplication of elements of the foot and limb as well as other structural abnormalities. Human and chicken diplopodia (Figure 1) share phenotypic features (Karchinov 1973; Jones et al. 1978; Narang et al. 1982; Hamanishi et al. 1985; Brower et al. 2003; Khan et al. 2008; Kadir et al. 2011) including extreme polydactyly (more than one additional digit), micromelia (truncated limbs), and visceral abnormalities. Although similar to polydactyly in that supernumerary digits are duplicated, diplopodia differs from common polydactyly in that affected individuals have additional metatarsal and tarsal bones in addition to the extra digits. *Diplopodia*, derived from the Greek roots diplo, meaning "double" and pod meaning "foot," refers to the pre-axial polydactyly observed in the mutant embryo (Figure 1), resulting in the doubling of the structures of the foot (i.e., metatarsals) (Taylor and Gunns 1947).

Diplopodia-1 (*dp-1*) is 1 of 5 *diplopodia* mutations (named *dp-1* through *dp-5*) described in chicken and was shown to be non-complementary with the other 4 mutations (Robb et al. 2011 and references therein). The morphological severity of the wing and leg defects was assessed and ordered as such: *dp-3* < *dp-1* < *dp-4* < *dp-2* (Taylor 1972). The *dp-1* mutation originated within a Single Comb White Leghorn commercial stock, was acquired by researchers at the University of California in 1941, and later incorporated into a congenic inbred line (a.k.a. UCD-Dp-1.003) in the late 1980s to early 1990s. This mutation is characterized by dwarfism, craniofacial defects (e.g., cleft palate), exposed visceral organs (omphalocele), micromelia, and polydactyly ranging from the addition of 1–5 pre-axial digits (pre-appendage) affecting the legs and wings (Taylor and Gunns 1947; Landauer 1956; Robb et al. 2011; Robb 2012). Digit pattern conformation analyses in *dp-1* embryos showed that a majority (~80%) of mutants had 6 digits on each foot arranged in 2 sets of 3, in a 2′-1′-1-2-3-4 digit conformation (Landauer 1956). (The primed numbers [e.g., 1′] refer to the extra, duplicated digit; 1-2-3-4 is the foot pattern of a normal embryo.) During the characterization studies conducted by Taylor and Gunns (1947), *dp-1* was found to be an autosomal recessive, embryonic lethal mutation affecting both sexes. Over half a century after this mutation was initially studied for mode of inheritance, Robb et al (2011) mapped the *dp-1* mutation to a 708-kb region on GGA 1 (p arm) using an SNP array (formerly 720 kb based on the galGal3 assembly).

Herein we describe our research conducted as part of a long-term strategy to identify the causative element responsible for the chicken dp-1 phenotype, which involved breeding, genetic, genomic and developmental tools of classical and modern biology. Fine-mapping and identification of recombinant individuals narrowed the region to 586 kb. To advance the aim to discover the causative element, a targeted genomic capture enrichment (a.k.a. capture array [CA]) was paired with next-generation sequencing (NGS) to sequence, in its entirety, the region linked to the UCD-Dp-1.003 mutation. Bioinformatics of the NGS results identified sequence variants (e.g., SNPs, micro-indels, gaps) specific to the *dp-1* mutation, while multiple pairwise-line comparative genomic analyses and validation tests identified those unique (UCD-Dp-1.003-specific) elements, which could be responsible for the mutant phenotype. Lastly, RNA *in situ* hybridization was utilized to assess the spatial and temporal expression patterns of known genes in the linked region.

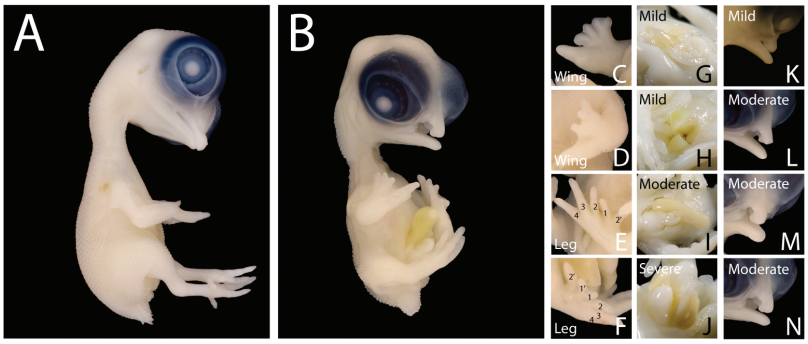

Figure 1. Developmental phenotype variation of diplopodia-1 mutant embryos. All individuals shown are at 10 days of embryogenesis. **A**. Normal (+/+) UCD-Dp-1.003 chicken embryo. **B**. UCD-Dp-1.003 mutant embryo (–/–) displaying truncation of the limbs, mild cleft-palate, mild dwarfism (shorter stature), exposed viscera (omphalocele), and pre-axial polydactyly (5 digits on both wings and legs). **C, D**. Wing digit number variation is observed in UCD-Dp-1.003 mutants (5 digits shown in C; 4 digits shown in D; normal digit number is 3). **E, F.** Hind limb (leg) digit number variation is observed in UCD-Dp-1.003 mutants. Note the inter-individual variation of digit length and individual morphology (i.e., webbing versus no webbing at same stage of development) in the wings and legs. **G–J.** Variation of visceral exposure is observed in UCD-Dp-1.003 mutant embryos. **K–N.** Variability in clefting (mild through severe) is also observed in UCD-Dp-1.003. Note that dp-1 mutant embryos displaying severe clefting, a phenotype not as common but yet observed over the course of this study, were not available on the day of photography and it was deemed inappropriate/unnecessary to set additional eggs for observations when the severe phenotype representation is available from other lines (e.g., coloboma—see Robb et al. 2013) at same age (i.e., in terms of reducing the usage of animals).

Methods

Chicken Genetic Lines

The individuals utilized for this study were from 2 genetic lines, the developmental mutant-congenic inbred line UCD-Dp-1.003 (aka Dp-1.003) and the inbred ($F > 0.99$) parent background line UCD-003 (Abplanalp 1992; Robb et al. 2011). Animals were under the care and supervision of trained staff and as per an approved protocol by the UC Davis Institutional Animal Care and Use Committee (Protocol # 15439). Associated with the studies described below, an in-depth assessment of mutant ($n = 56$) phenotypic variation, in the context of the inbred congenic line background, was conducted.

Sample Collection for Fine-Mapping

Dp-1.003 embryos were incubated to E10, a stage of development such that the phenotypes (normal: +/+ or +/–; mutant: –/–) could be easily and accurately discerned (Figure 1). Adult or embryonic blood samples were collected according to Robb et al. (2011) and pin feathers were collected from day-of-hatch chicks (Robb 2012). DNA (with an RNAse step) was isolated from blood and tissue (feather pulp, brain) sources using the DNeasy® Blood & Tissue kit (Qiagen) and examined by electrophoresis to determine integrity. See Robb and Delany (2012b) for preparation of samples used in the capture array (CA) technology.

SNP-Genotyping, Fine-Mapping Analysis, and CR Identification

Eight SNPs (rs14931758–rs13990802; Table 1), in complete linkage disequilibrium with the dp-1 trait (Robb et al. 2011), were used to fine-map the region. DNAs (10 ng/µL) isolated from Dp-1.003 homozygous normal (+/+, $n = 54$), heterozygous normal (+/–, $n = 106$), and mutant (–/–, $n = 78$) samples were used to determine and/or confirm genotypes at the 8 loci using standard polymerase chain reaction (PCR) conditions. Reactions were amplified using Phire® Hot Start II DNA Polymerase (Thermo Scientific) and purified by QIAquick® Spin Kit (Qiagen) under manufacturer conditions. Amplicons were sequenced (Davis Sequencing, Davis, CA) using ABI 3730 DNA sequencers (Applied Biosystems) and were analyzed for genotype-specific SNP differences. The SNP analysis and causative region (CR) identification (maximum and minimum CR: CR_{max} and CR_{min}, respectively) were defined as previously described (Robb et al. 2011). A subset of the mutant samples ($n = 20$) were utilized in the variant validation portion of this study described below.

Capture Array and SOLiD™ Sequencing

Genomic Enrichment Services of SeqWright, Inc., were utilized to design a custom NimbleGen capture array to enrich for the 585 451-bp chromosomal region (GGA 1: 185,695,823-186,281,274 [galGal5]; formerly 189,915,923-190,511,744 [galGal3]) linked to the UCD-Dp-1.003 developmental mutation (Robb et al. 2011; Robb and Delany 2012b). Our strategy for the capture array development was to design probes for the 586 kb CR_{max} identified post-60K SNP array and subsequent fine-mapping. Two dp-1 mutant samples (dp1-166F and dp1-168F) were utilized in the CA/NGS after genotypic confirmation at the 8 SNP loci listed in Table 1. Both samples exhibited a reduced CR_{max}, indicated by SNP analyses. Specifically, dp1-166F was utilized because the mutant reduced the region at the 5′ end, whereas dp1-168F narrowed the region at the 3′ end of the linked region (Table 2). Full details of the array setup and methods can be found in Robb and Delany (2012b). (Note that the same array was utilized to sequence 2 other UCD-congenic developmental mutations, *coloboma* [co.003] and *wingless-2* [wg-2.331]; Robb and Delany 2012b; Robb et al. 2013; Webb et al. 2018.) The other 2 congenic lines, which mapped to different chromosomes (GGA Z and 12), were utilized as controls for GGA 1 comparisons (Robb et al. 2011; Robb 2012; Robb and Delany 2012b).]

Unique Variant Identification and Causative Element Analysis

The CA/NGS data were analyzed to identify SNPs, micro-indels (1–3 nt), gaps, and chromosomal rearrangements. Reads were aligned to the *Gallus gallus* NCBI UCD-001 Red Jungle Fowl (RJF) reference genome sequence (WASHUC2, May 2006). Reference-assisted assembly and variant identification procedures are described fully in Robb and Delany (2012b); Dp-1.003-specific variant analysis is briefly reviewed below.

The discrimination of non-causative polymorphic elements (naturally occurring in the introgressed region) from the potentially causative polymorphic elements was accomplished by a comparative approach. Polymorphisms were identified by comparing the

Table 1. Diplopodia-1 fine-mapping primers used to identify carrier status, CR size, and recombination events

SNP	Position[a]	Primers (5′–3′)		
		Forward	Reverse	Product size (bp)[b]
rs13989579	185432685	GAGATGTGGAGGCCAAAAGT	CTAATGCCAACCTGGCTTCT	282
GGaluGA060936	185482293	CACAAAAATGGTGAATTCTTAACAT	GTTAGAGTTCACAAAGTAGAAGTCTGG	250
rs14931758	185587553	CAGTGACCAGCAGAAGGACA	TGGCCTGGTTAATGTGTGAA	192
rs14931949	185671027	GCCAAGTATCAACTCCCATC	CCACTCTGAAAAGGAGCATT	214
rs13989873	185695823	TTCTGCTATTCTGTGGCTCA	TGGAAATCAGAGTTGAAGCA	224
rs14932143	185805128	GTTACTTTCAGCAAGCACAG	GCACCATTTCAACACATCTA	250
rs13989974	185830327	CTGATGTGCTCAGGTAACAC	AGGAGACACTGATTTCCTTG	194
rs13990135	185909526	TCAAGAACTTGCTGGCATAA	GGTTCCTTAGTCTCCCCAGT	192
rs13990785	186281274	CCACAAATGCATTACCTGAG	TCCCCTACAGACCCTATTGA	169
rs13990802	186295291	ATCAACGCAGGGTCTCAACT	GGCTGTACAAATCTGATCATCATTAC	260

[a]Chromosomal location (bp) of SNP on GGA 1; positions are based on the December 2015 *Gallus gallus* assembly (galGal5).

[b]PCR fragment size was determined by 3 methods: 1) using the UCSC genome browser (http://genome.ucsc.edu/), 2) sizing by gel electrophoresis, and 3) DNA sequencing.

Table 2. SNP fine-mapping shows a reduced CR via assessment of linked molecular markers in the UCD-Dp-1.003 genetic line

				Original 60K SNP array mutant samples[a]			Post-60K SNP array mutants[b]			Post-60K SNP array heterozygotes[c]				Overall reduced genotype[d]	Normal (+/+)[e]
			Sample ID No:[f]	*n* = 2	*n* = 3	*n* = 5	*n* = 7[g]	*n* = 3[h]	*n* = 2	*n* = 2	*n* = 2	*n* = 3	*n* = 6		
SNP ID	SNP	Position[i]	Genotype:[j]												
rs13556505	A/G	177307260		AG	GG	AA	AG	GG	AG	GG	GG	GG	GG	AA	AA
rs13989579	A/G	185432685		GG	GG	AA	AG	GG	AG	GG	GG	GG	GG	AA	AA
GGalu-GA060936	A/G	185482293		AA	AA	GG	AG	AA	AG	AA	AA	AA	AA	GG	GG
rs14931758	A/G	185587553		GG	GG	AA	AG	GG	AG	GG	GG	GG	GG	AA	AA
rs14931949	A/C	185671027		CC	CC	CC	AC	CC	AC	CC	CC	CC	CC	AA	AA
rs13989873	A/G	185695823		GG	GG	GG	AG	GG	AG	AG	AG	GG	GG	AA	AA
rs14932143	A/C	185805128		CC	CC	CC	CC	CC	CC	AC	AC	AC	AC	CC	AA
rs13989974	A/G	185830327		AA	AA	AA	AA	AA	AA	AG	AG	AG	AG	AA	GG
rs13990135	A/G	185909526		AA	AA	AA	AA	AA	AA	AG	AG	AG	AG	AA	GG
rs13990785	A/G	186281274		GG	GG	GG	GG	AG	AG	AG	AA	AG	GG	AA	AA
rs13990802	A/G	186295291		GG	GG	GG	GG	AG	AG	AG	GG	AG	GG	GG	GG
rs13990804	A/G	186316405		AA	AA	AA	AA	AG	AG	AG	AA	AG	AA	AA	AA
60K CR_{min}:[k]	610,247		Individual CR_{min}:[l]	848,589	848,589	610,247	476,146	8,602,266	104,398	620,582	213,703	511,277	104,398	104,398	
60K CR_{max}:[k]	707,738		Individual CR_{max}:[l]	8,988,029	9,447,557[M]	707,736	599,466	9,474,014	585,449	1,145,378	610,245	1,120,582	585,451	585,449	

Samples were assessed for their 60K SNP genotyping pattern to identify recombination events, thereby reducing the size of the linked region. This CR_{max} decreased by 122-kb post-60K SNP array analysis. Reducing the size of the CR is beneficial as it: 1) eliminates candidate genes in question, 2) eliminates genomic regions (as other elements could be the cause of the mutation), and 3) allowed for the identification of samples with a narrowed CR to be used in the capture array technology.

[a]Samples originally genotyped using the Illumina 60K SNP genotyping array (Robb et al. 2011).

[b]A total of 78 mutant embryos were collected post-60K SNP analysis (yr 2009–2011). These were genotyped using the 60K SNPs (Table 1), but only 12 displayed recombination events. Two of these samples, dp1-166F and dp1-168F, were utilized in the targeted genomic capture enrichment technology.

[c]A total of 73 heterozygous chicks were hatched post-60K SNP analysis (yr 2009–2011). All were genotyped using the 60K SNPs (Table 1) and 13 displayed recombination events.

[d]The overall reduction, combining the fine-mapping results of both mutant and heterozygous samples, is displayed. Three SNPs remain linked to the Dp-1.003 mutation, thereby maintaining a CR_{min} and CR_{max} of 104,398 and 585,449 bp, respectively.

[e]SNP genotype observed in normal, control samples.

[f]Please note that all individuals used in reproduction were Dp-1.003 carriers (+/–) except for female 18539F (used in 15 matings) and male 10775M (used in 1 mating) which were UCD-003 (congenic background).

[g]Note: targeted genomic capture enrichment technology sample, dp1-166F, showed this 60K SNP genotype pattern.

[h]Note: targeted genomic capture enrichment technology sample, dp1-168F, showed this 60K SNP genotype pattern.

[i]Chromosomal location (bp) of SNP on GGA 1; positions are based on the December 2015 *Gallus gallus* assembly (galGal5). SNPs were identified using the 60K SNP array (Robb et al. 2011).

[j]Genotypes shaded in gray indicate a region which is no longer linked to the Dp-1.003 mutation.

[k]UCD-Dp-1.003 minimum and maximum CR (candidate gene region; CR_{min} and CR_{max}) identified by Robb et al. (2011) using the chicken 60K SNP genotyping array.

[l]Minimum and maximum CR (candidate gene region) identified for each sample.

[m]The recombination rate at this position of GGA 1 is 2 cM/Mb (Groenen et al. 2009). In order to best estimate the CR_{max}, 500 kb was subtracted from the 5′ sequence coordinate of the CR_{min}.

Table 3. Thirteen genes found within the Dp-1.003 260 475 nt CR identified by fine-mapping techniques

Candidate gene	Synteny chromosome[a]		Gene function[d]	Knockout mouse phenotype[e]	Affected anatomical systems[e]	Associated diseases[f]
	Mouse[b]	Human[c]				
MRE11[g]	Mmu 9	Hsa 11	Component of MRN (Mre11a, Rad50, Nbs1) complex, which is involved in DNA repair in animals. Specifically, it has a role in DNA recombination, double-strand break repair, and maintenance of telomere length and meiosis. Disruption results in defective sister chromatid cohesion.	Mortality/aging, embryogenesis, behavior, tumorigenesis; lethal: embryos developmentally arrest at e9.5 with a poorly defined body axis, no heart development	Reproductive, cardiovascular	*Human: Ataxia telangiectasia-like disorder (ATLD); breast, lymphoid, and colorectal cancers; Alzheimer's disease; Cornelia de Lange Syndrome I* *Mouse: Associated with several identified QTLs (e.g., tuberculosis severity 2, and others)*
GPR83	Mmu 9	Hsa 11	G-protein coupled receptor; endogenous GPR83 plays a role in the induction of Foxp3 expression in peripheral T cells. Gpr83 may be involved in the central control of the energy metabolism.	Immune, hematopoietic, lean, resistant to diet-induced obesity and resistant to diet-induced glucose intolerance	Normal regulatory T cell development	*Human: Associated with several identified QTLs (e.g., blood pressure, body weight, prostate tumor susceptibility, and others)* *Mouse: Associated with several identified QTLs (e.g., cholesterol 6, Salmonella enteritidis susceptibility 6, and others)*
PANX1	Mmu 9	Hsa 11	Plasma membrane protein which mediates "find-me" signal release from apoptotic cells. Associated with the initiation and propagation of calcium waves; may regulate ER Ca^{2+} homeostasis.	Nervous, immune, and respiratory systems, behavior	Impaired macrophage recruitment, ATP release by apoptotic astrocytes, hippocampal neurons, and thymocytes	*Human: Associated with several identified QTLs (e.g., insulin level, chronic obstructive pulmonary disease, and others)* *Mouse: Tumor-suppressive effects*
HEPHL1	Mmu 9	Hsa 11	Hephaestin-like protein 1. Function inferred by similarity: may be involved in copper transport and homeostasis as well as may function as ferroxidase	In progress	In progress	*Human: Associated with several identified QTLs (e.g., prostate tumor susceptibility, serum adiponectin level, and others)* *Mouse: Associated with several identified QTLs (e.g., haloperidol-induced catelepsy 5, cholesterol 6, and others)*
CP	Mmu 3	Hsa 3	Ceruloplasmin (ferroxidase enzyme) helps move iron from organs/tissues into blood stream by preparing iron for incorporation into transferrin (which transports iron to RBCs); also binds copper in plasma.	Behavior, immune system, homeostasis, hematopoietic	Vision, cardiovascular, nervous system, liver, pigmentation	*Human: Aceruloplasminemia; Wilson's disease; Obsessive compulsive disorder; Schizophrenia* *Mouse: Associated with several identified QTLs (e.g., organ weight, HIV-associated nephropathy1, and others)*
HEPH	Mmu X	Hsa X	A multi-copper oxidase critical for intestinal iron absorption as it regulates transport of dietary iron from epithelial cells of the intestinal lumen into the circulatory system.	Embryogenesis, cardiovascular, hematopoietic, digestive/alimentary, nervous system	Impaired iron transport to gut and placenta, small and pale at birth, exhibit hypochromic anemia (disappears with age)	*Human: Colorectal cancers* *Mouse: Severe microcytic, hypochromic anemia; age-related macular degeneration*

Table 3. Continued

Candidate gene	Synteny chromosome[a]		Gene function[d]	Knockout mouse phenotype[e]	Affected anatomical systems[e]	Associated diseases[f]
	Mouse[b]	Human[c]				
TM4SF1a	Mmu 3		A cell surface antigen belonging to the tetraspanin family, which is characterized by the presence of 4 hydrophobic domains. Members of this transmembrane family are cell-surface proteins, which typically mediate signal transduction events involving the regulation endothelial cell movement and proliferation and growth; commonly associated with cancer cell proliferation.	Embryonic lethal at day 9 due to vasculogenesis failure	Cardiovascular	*Human: Breast, colon, lung, and ovarian carcinomas* *Mouse: NA*
VSTM5	Mmu 9	Hsa 11	Cell adhesion membrane glycoprotein, which regulates neuronal morphology and synapse formation during cortical development in the brain	—	—	*Human: Unknown* *Mouse: Unknown*
C11orf54	Mmu 9	Hsa 11	Exhibits ester hydrolase activity on the substrate p-nitrophenyl actetate	In progress	In progress	*Human: Associated with several identified QTLs (e.g., blood pressure, body weight, insulin level, and others)* *Mouse: Associated with several identified QTLs (e.g., cholesterol 6, tuberculosis severity 2, and others)*
MED17	Mmu 9	Hsa 11	Serves as a scaffold for the assembly of a functional preinitiation complex with RNA polymerase II and the general transcription factors; without MED17 and TFIID, SP1 cannot be activated.	In progress	In progress	*Human: Microcephaly postnatal progressive with seizures and brain atrophy (MCPHSBA)* *Mouse: Associated with several identified QTLs (e.g., haloperidol-induced catelepsy 5, cholesterol 6, and others)*
SNORD6	Mmu 9	Hsa 11	Small nucleolar RNA, C/D box 6 involved in a variety of important processes such as RNA splicing, regulation of transcription factors or RNA Polymerase II, and maintaining telomeres. Exact role unknown.	—	—	*Human: Associated with several identified QTLs (e.g., body weight, insulin level, prostate tumor susceptibility, and others)* *Mouse: Unknown*
SNORA8	Mmu 9	Hsa 11	Small nucleolar RNA, H/ACA box 8 involved in a variety of important processes such as RNA splicing, regulation of transcription factors or RNA Polymerase II, and maintaining telomeres. Exact role unknown.	—	—	*Human: Associated with several identified QTLs (e.g., chronic obstructive pulmonary disease, body weight, and others)* *Mouse: Unknown*

Table 3. Continued

Candidate gene	Synteny chromosome[a]		Gene function[d]	Knockout mouse phenotype[e]	Affected anatomical systems[e]	Associated diseases[f]
	Mouse[b]	Human[c]				
SMCO4	Mmu 9	Hsa 11	Single-pass membrane and coiled-coil domain-containing protein 4; hypothetically found in membrane and interacts with carbon nanotube, tributyltin oxide, troglitazone	—	—	*Human: Unknown* *Mouse: Unknown*

[a]UCSC genome browser (http://genome.ucsc.edu/) was utilized to identify the location of each gene and the percent to the identity to human protein.

[b]Synteny alignment was made to the mouse reference genome (December 2011; GRCm38/mm10) in the UCSC genome browser.

[c]Synteny alignment was made to the human reference genome (December 2013; GRCH38/hg38) in the UCSC genome browser.

[d]References for the gene function are as follows: *MRE11*: Petrini et al. (1995), Trujillo et al. 1998, Gatei et al. 2000, Barber et al. 2008; *GPR83*: Lu et al. (2007), Müller et al. (2013); *PANX1* (aka *Pannexin 1*): Vanden Abeele et al. (2006), Chekeni et al. (2010), Kienitz et al. (2011); *HEPHL1*: Chen et al. (2010); *CP*: Holmberg and Laurell (1948), Hellman and Gitlin (2002); *HEPH*: Frazer et al. (2001), Vulpe et al. (1999), Prohaska (2011), *C11orf54* (aka mouse: *4931406C07Rik*): Manjasetty et al. (2006); *MED17*: Baek et al. (2006); *SNORD6*: Dieci et al. (2009); *SNORA8*: Dieci et al. (2009). SMCO4: https://rgd.mcw.edu/rgdweb/report/gene/main.html?id=1319742; TM4sf1a: Zukauskas et al. (2011). *VSTM5*: Lee et al. (2016).

[e]Information for the knockout mouse phenotype and affected anatomical systems were obtained from the Mouse Genome Informatics (MGI, http://www.informatics.jax.org/), the International Knockout Mouse Consortium (IKMC, http://www.knockoutmouse.org/), and the Rat Genome Database (RGD) PhysGen Knockouts (http://rgd.mcw.edu/). A dash indicates that the particular gene was not present in the database.

[f]References for the associated disease are as follows: *MRE11*: human: Stewart et al. (1999), Fukuda et al. (2001), Jacobsen et al. (2004), Sjöblom et al. (2006), mouse: http://rgd.mcw.edu/ ID=735478; GPR83: human: http://rgd.mcw.edu/ ID=1352259, mouse: http://rgd.mcw.edu/ ID=733083; PANX1: human: http://rgd.mcw.edu/ ID=1347653, mouse: Lai et al. (2007); HEPHL1: human: http://rgd.mcw.edu/ ID=1345614, mouse: http://rgd.mcw.edu/ ID=1622024; *CP*: human: Scheinberg and Gitlin (1952), Gitlin (1998), Wolf et al. (2006), Virit et al. (2008), mouse: http://rgd.mcw.edu/ ID=10384; *HEPH*: human: Brookes et al. (2006), mouse: Hahn et al. (2004), Griffiths et al. (2005); *C11orf54*: human: http://rgd.mcw.edu/ ID=1603055, mouse: http://rgd.mcw.edu/ ID=1319924; *MED17*: human: Kaufmann et al. (2010), mouse: http://rgd.mcw.edu/ ID=1322409; *SNORD6*: human: http://rgd.mcw.edu/ ID=1602744, mouse: NA; and *SNORA8*: human: http://rgd.mcw.edu/ ID=1602779, mouse: NA; *TM4SF1*: human: Marken et al. (1992).

[g]Note that the 3′ end of MRE11 (11,462 nt inclusive of 14 conserved—human, mouse, frog, opossum–exons) remains linked to the Dp-1.003 mutation.

Dp-1.003 causative (sequenced) region to variants identified in other chicken genetic lines including: the other inbred congenic lines (*wg-2*, Webb et al. 2018, and *co*, Robb et al. 2013) utilized on the capture array, the chicken reference genome (UCD-001), or any other previously reported polymorphism (e.g., within NCBI [ex. dbSNP] and the UCSC genome browser). Any polymorphism (e.g., SNP, indel, etc.) found in the Dp-1.003 GGA 1 CR, not found within any of the other sources, was deemed *dp-1*-specific (i.e., unique) and therein considered as a possible causative element for the *dp-1* mutation. Newly identified, shared variants were submitted to NCBI (accessions: ss472337944–ss472340673).

To further eliminate normal variation, not causing the dp-1 phenotype, unique sequence elements were assessed in a set of new mutant samples, the progeny from different matings. DNAs isolated from 20 mutant (–/–) and 2 control (UCD-003, +/+, $F > 0.999$) individuals were used to assess the dp-1-specific elements identified by pairwise comparisons described above within the newly reduced 260 kb region on GGA 1 identified by fine-mapping of the 2 mutant samples utilized in the CA.

Specifically, all polymorphisms located within an exon, UTR, or splice site within the original 586 kb were assessed for linkage. Primers were designed to flank the variant using sequences available through the UCSC genome browser (http://genome.ucsc.edu/) (primer sequences are available upon request). Elements were PCR amplified (standard conditions) and amplicons were sequenced. Each element was then evaluated using polymorphic marker pattern analyses described in Robb et al. (2011). Briefly, a variant can be eliminated as causative if it is present in control (+/+) samples or if the polymorphism is absent in additional mutant samples (beyond the 2 individuals sequenced as part of the CA/NGS study). Using this strategy, unique SNPs, micro-indels, and gaps found within an exon or splice site (thus potentially impacting mRNA and protein) of known and predicted genes (Table 3) were tested in an additional 20 mutant and 2 control (UCD-003) samples to verify if a particular variant remained linked to the *dp-1* mutation (and thus could be potentially causative). Table 5 outlines the number of each element found at each stage of the bioinformatics analysis.

RNA *In situ Hybridization*

Eleven candidate protein-encoding genes (*MRE11*, *GPR83*, *PANX1*, *HEPHL1*, *CP*, *HEPH*, *TM4SF1A*, *VSTM5*, *C11ORF54*, *MED17*, and *SMCO4*) within the *dp-1* 260 kb fine-mapped region (post-CA/NGS: 260,475 bp) were examined for their expression in normal (+/+) chick embryos covering developmental stages HH14 to 27 (Hamburger and Hamilton 1951). Two small nucleolar RNA genes, *SNORD6* (63 nt) and *SNORA8* (137 nt), could not be assessed using the methodology employed. A probe of at least 300 nt is needed to obtain enough signal to detect a moderate to abundantly expressed RNA. The RNA *in situ* hybridization procedures were followed as described in Darnell et al. (2007); in some cases RNA probes were created from cDNA (EST) clones (Table 4) acquired from BBSRC ChickEST Database (http://www.chick.manchester.ac.uk/) using methods adapted from Nieto et al. (1996) or by PCR amplification. The expression patterns for these genes at developmental stages other than those shown here can be found at http://geisha.arizona.edu/geisha/ and in Robb (2012). Embryos used were under the care and supervision of trained staff and as per University of Arizona regulations (Protocol # 08-133).

Results

Mutant Phenotype

Assessment of the UCD-Dp-1.003 congenic inbred line indicated that ~70% of the mutants (n = 56) display ≤11 total digits on both feet (mode: 5 digits on each foot [n = 27]), while ~30% display ≥12 digits (total for both feet). In 1959, it was reported that 17% of *dp-1* mutant embryos displayed asymmetrical numbers of toes; of these, 83% had more toes on the right leg (Abbott 1959b). We found a similar pattern, with 27% of *dp-1* mutant embryos displaying an asymmetrical number of digits between the right and left legs. However, of these asymmetrical cases only 33% had more toes on the right leg. Note that the number of digits on the foot (ranging from 4 to 6 on each foot; 9–13 total) was found to be relatively constant within a mating pair. We also observed phenotypic variability in the dp-1.003 mutants with regard to the amount of exposed viscera (Figure 1G–J), degree of dwarfism, and cleft palate (Figure 1K–N), as compared to normal embryos at the same stage of development.

Fine-Mapping Strategies Reduce Mutant-Associated Region Size and Indicate Dp-1.003 GGA 1 Recombination Rate

A 60K SNP genotyping array was initially employed to map the CR for the *dp-1* phenotype to GGA1 with a CR_{min} and CR_{max} of 610,247 and 707,738 bp (formerly 622 and 720 kb, respectively, based upon the galGal3 assembly) (Robb et al. 2011). In this study the *dp-1* mutation was further fine-mapped (region narrowed) using carrier (+/–) individuals and mutant (–/–) progeny via assessment with 8 SNPs (Table 1) to identify their genotypic pattern and detect recombination events. Based upon recombinant analysis of new progeny the CR_{min} and CR_{max} were reduced to 104,398 and 585,451 bp, respectively (Table 2).

A breeding scheme was initiated, which included the mating of those heterozygous individuals with a reduced linked region (Table 2), to decrease the CR size in subsequent generations of carrier and mutant progeny. Of the 78 mutant embryos and 106 day-of-hatch chicks (+/–) collected post-60K SNP analysis, 12 and 13, respectively, showed a decreased CR relative to that identified in the 60K SNP array (Table 2). From these results the recombination rate within the GGA 1 CR for the *dp-1.003* congenic line was calculated to be ~0.4 cM/Mb (LOD score [Z] = 21.7), thereby indicating high linkage disequilibrium in this region and contrasting with the rate of 1.0–1.5 cM/Mb calculated for this region in other chicken genetic lines (Groenen et al. 2009; Elferink et al. 2010).

Targeted Genomic Capture Enrichment Technology

A targeted-sequence genomic enrichment capture array (CA) paired with NGS was employed to sequence the entire 586 kb region (GGA 1: 185,695,823-186,281,274) linked to the *dp-1* mutation by recombinant analysis post-60K array. Bioinformatic analyses were then carried out to identify all *dp-1*-specific variants (e.g., SNPs, indels). Those elements found within an exon or splice site of known/predicted genes (Table 5) were further assessed for causation in the validation portion of this study. The CA/NGS methods (capture array format and procedures) and results (number of reads, coverage, etc.) for *dp-1* are reported in Robb and Delany (2012b). Briefly, a total of 21.0 M dp-1 reads were generated in the CA/NGS sequencing efforts. Those reads covered 96.9% of the 586 kb targeted region with an average fold coverage of 107.2*x* (Robb and Delany 2012b).

SNP Identification within the Dp-1.003 Congenic Line ~586 kb CR and ~260 kb CR

586 kb CR (GGA 1: 185,695,823-186,281,274) SNPs

A total of 2593 SNPs were found within the GGA 1 targeted region. Multiple pairwise genomic comparisons led to the elimination 1815 SNPs as these variants were shared across one or both of the other developmental congenic lines, which served as controls for GGA 1 on the array (Table 5). The transition:transversion ratio was calculated to be 1.8:1, which is 1.2x lower compared to that reported for other chicken breeds (2.2:1) (Sherry et al. 2001) and compared to the other control congenic lines used on the array (Coloboma.003: 2.2:1; Wingless-2.331: 2.1:1) for this region. Similarly, the average SNP density of the dp-1 region (any SNP identified relative to the RJF reference genome) was 4.4 SNP/kb, which is also 1.2x lower than that reported in other domestic chicken lines (5.1–5.8 SNP/kb) (ICPMC 2004).

A total of 172 SNPs were found to be heterozygous and could be eliminated as causative since mutants are homozygous recessive and require 2 of the same allele to be affected. Assessment of the heterozygous SNP pattern led to the identification of the recombination breakpoints for mutants dp1-166F and dp1-168F (Table 2), which were utilized in the CA because of their reduced linked regions. CA SNP sequence analyses further reduced the CR_{max} to 260,477 bp (herein referred to as 260 kb) (Figure 2B), specifically moving the coordinate location from chr1: 186,281,274 to chr1: 185,956,300. The new CR_{max} now spans chromosome 1 from position 185,695,823 to 185,956,300.

~260kb CR (GGA 1: 185,695,823-185,956,300) SNPs

Within the 260 kb CR, 1509 SNPs were identified with 696 of those found to be unique to *dp-1* (Table 5) by the pairwise comparison analyses. Of these SNPs, 382 are non-genic and 314 are found within genes. Specifically, 23 SNPs were found in exons or UTRs, 21 at splice sites, and 270 within introns (see Table 5 and Supplementary Table 1). The unique SNPs within an exon or at a splice site were assessed for codon and amino acid changes within 6 reading frames (+1, +2, +3, –1, –2, –3); none generated nonsense mutations (synonymous SNPs: 528; non-synonymous SNPs: 168). Those SNPs found within an exon or splice site of a known gene were further assessed in the validation portion of this study described below.

Micro-indel (1–3 nts) within the UCD-Dp-1.003-Associated CRs

~586 kb CR (GGA 1: 185,695,823-186,281,274) micro-indels

The number and location of micro-indels (insertions or deletions relative to the reference genome) of 1–3 nucleotides in length were investigated. In total, 283 indels (150 insertions, 133 deletions) were placed within the original 586 kb GGA 1 CR with an average distance between micro-indels of 2.11 kb (Table 5). On average, the distance between insertions was 3.98 kb and the average distance between deletions was 4.38 kb. This value (0.48 micro-indels/kb), is slightly higher than that reported previously (~0.4 short indels/kb) (ICPMC 2004). The deletion:insertion ratio (1:1.1) within the 586 kb region for this congenic line is lower than that observed in other breeds for the macrochromosomes (1.6:1) (Brandström and Ellegren 2007). Moreover, the average micro-indel size for Dp-1.003 was 1.3 bp. Note that the average distance between micro-indels and deletion:insertion ratio was calculated here for only the 586 kb region, while values from other studies were calculated for the entire chromosome 1 or macrochromosomes (GGA 1–5) in the case of deletion:insertion ratio.

Table 4. Probes used in whole-embryo *in situ* hybridization: Analysis of UCD-Diplopodia-1.003 candidate gene expression

Gene[ab]	Accession number	Forward primer (5′–3′)	Reverse primer (5′–3′)	Probe length (bp)
MRE11	NM_204778	TTATAAAAGCTTGCGG CCGCA GAATATACTCC AGCCCAAAAC GAACAG	GCTCTAGAAATTAACCCTCAC TAAA GGGAATGTCAT CATCTGAATCACCCG	931
GPR83	NM_001256138.1	AACAGAAGCCTTG AGGAGGTG	AATTAACCCTCACTAAAGGAA GC GAGGTTTCAGAGGGTG	467
PANX1	XM_015280513.2	GAGAACATTGCTGT GTCCGAG	AATTAACCCTCACTAAAGGAG TCTAAACTCCCAGAAGGATACG	1087
HEPHL1	XM_015280492.2	GTCAGGCTCTAC TACATCGCTG	AATTAACCCTCACTAAAGGG TCAAAT ACATCTCCTCTGTGGC	885
CP	XM_015291853.2	TTCCGAAATAAT GCCAGCCGCC	AATTAACCCTCACTAAAGGC TTC CTCTGCTCTGCTTTTTGG	1000
HEPH	XM_420165.6	GCTGCCATCT ATGAAGTCCGTC	AATTAACCCTCACTAAAGGG TCC CTTTGTTGTCAGGTTGC	1013
TM4SF1A	NM_001277830.1	TCCTGTATTTCC CCAATGGACG	AATTAACCCTCACTAAAGGCC ACACACAAAGCCAATGAGAC	469
VSTM5	XM_003640569.4	CAGCCCAACATCA ACGCAAC	AATTAACCCTCACTAAAGGCC AGT CATTTCAGACAGCATCATC	659
C11ORF54	NM_001277277.1	GTGTTCCATACCTCATACCACTTG	AATTAACCCTCACTAAAGG TCAGCCA TAAGCCCTTCAGTTAC	940
MED17	NM_001006280.1	CCTGTCTCATCAA GAGCAAGAAC	AATTAACCCTCACTAAAGGG TTATTC CATCAACCCCAAGCC	921
SMCO4	NM_001164340.3	TGCTCCCCCATTTTGGTGAAC	AATTAACCCTCACTAAAGGGG CAGA GAGTTTCTTTTGGCT	290

Gene[a]	EST ID[c]	Clone ID[c]	BLAT coordinates[d]	Probe length (bp)[e]
MRE11	603507591F1	ChEST434f24	chr1:185695679–185700608	~750
PANX1	603767553F1	ChEST695j13	chr1:185742472–185742920	~850
C11ORF54	603114284F1	ChEST65h21	chr1:185861531–185865290	~900
MED17	603799139F1	ChEST767n9	chr1:185857532–185861163	~900
SMCO4	603851519F1	ChEST848f5	chr1:185946385–185974927	~850

[a]A gene identified in any species within the Dp-1.003-linked 260,475 nt region was used in whole-embryo *in situ* hybridization*. Note: *C11orf54* encodes for protein Fn5; *Med17* is a.k.a. *Crsp6*.

[b]Probes established using PCR.

[c]ESTs were purchased from Source BioScience UK Limited geneservice (Cambridge, UK), through the BBSRC ChickEST database (http://www.lifesciences.sourcebioscience.com/). Clones were selected on carbenicillin plates (50 µg/mL) prior to growth in LB broth+ carbenicillin (50 µg/mL) and clone purification (using Qiagen's Plasmid Purification Kit). ESTs were sequenced prior to use in RNA ISH to confirm clone identity.

[d]UCSC genome browser (http://genome.ucsc.edu/) was utilized to identify the location of each gene and the EST percent identity to chicken mRNAs previously identified.

[e]Each EST clone was inserted and amplified in the pBluescript II KS+ vector, 3.0 kb (Stratagene). The estimated size of each EST was determined through standard restriction enzyme digest (NotI and EcoRI) and subsequent gel electrophoresis. NotI was used to cleave the vector for sense-strand RNA creation using T3 polymerase. Similarly, EcoRI was used, paired with T7 polymerase to generate anti-sense-strand RNA.

*Two small nucleolar RNA genes found within the region (*SNORD6* [63 nt] and *SNORA8* [137 nt]) were too small to test by the methodology employed. A probe of at least 300 nt is needed to obtain enough signal to detect a moderate to abundantly expressed RNA by these methods.

~260 kb CR (GGA 1: 185,695,823-185,956,300) micro-indels

Within the reduced 260 kb CR, 86 *dp-1* specific micro-indels (47 insertions and 39 deletions) were found (Table 5). A total of 31 and 28 insertions and deletions, respectively, were found external to a gene (non-genic), while 21 insertions and 11 deletions were found within a gene (see Table 5 and Supplementary Table 1 for details). Upon frameshift analysis of the 260 kb region, 5 of the 86 micro-indels generated non-synonymous mutations. The functional impact of micro-indels within an exon, UTR, or splice site was also examined. One insertion (+CC), located at the *MRE11* donor splice site (3′ end of exon 13), would alter protein structure by failing to remove a 856 nt intron, thereby adding 285 amino acids to the polypeptide and causing a frameshift in the downstream sequence. A second micro-indel (1 nt deletion), present in the 3′ UTR of *C11ORF54* could possibly alter the expression of this gene via miRNA-targeted translational inhibition or mRNA degradation. The micro-indels found within an exon, UTR, or splice site of a known gene were further assessed in the validation portion of this study, see below.

Identification of Gaps within the 260 kb CR

Lastly, alignment data indicated 296 gaps (i.e., putative large deletions ≥4 nts) across the 260 kb CR. The average and maximum gap sizes are 73 and 1932 nt, respectively. Those gaps localized to a known exon or splice site (Table 6 and Supplementary Table 1) were validated for legitimacy (i.e., is the gap a large deletion or the result

Table 5. Diplopodia-1.003 SNPs, micro-indels, and sequence gaps: number and genomic location

CR size[a]	Originally identified in CA[b]				Unique variants[c]			
	SNPs	Insertions[d]	Deletions[d]	Gaps[e]	SNPs	Insertions[d]	Deletions[d]	Gaps[ef]
585,451 nt	2593	150	133	581	778	82	64	581
260,477 nt	1509	73	63	296	696	47	39	296

[a]CR size = 585,451 nt (based on the galGal5 coordinate alignment) is the CR identified by the 60K SNP array, which was subsequently utilized for capture array probe creation. 260,477 nt is the CR identified by fine-mapping analysis of recombinant individuals and CA/NGS bioinformatic analyses.

[b]The original number of variants identified after bioinformatics analyses, prior to multiple pairwise-line comparisons to identify unique variants.

[c]Unique variants are those specific to Dp-1.003 only after multiple pairwise genomic comparisons. See Methods section for details as to unique variant identification.

[d]Insertions and deletions range from 1 to 3 nt in length (within the paper referred to as micro-indels).

[e]Sequence gaps (DNA that was not captured for sequencing in the CA/NGS) are those gaps greater than 4 nt in length, as identified by alignment to the 585,451 nt RJF reference genome obtained from NCBI (see Methods section).

[f]Gaps listed are those found within the region.

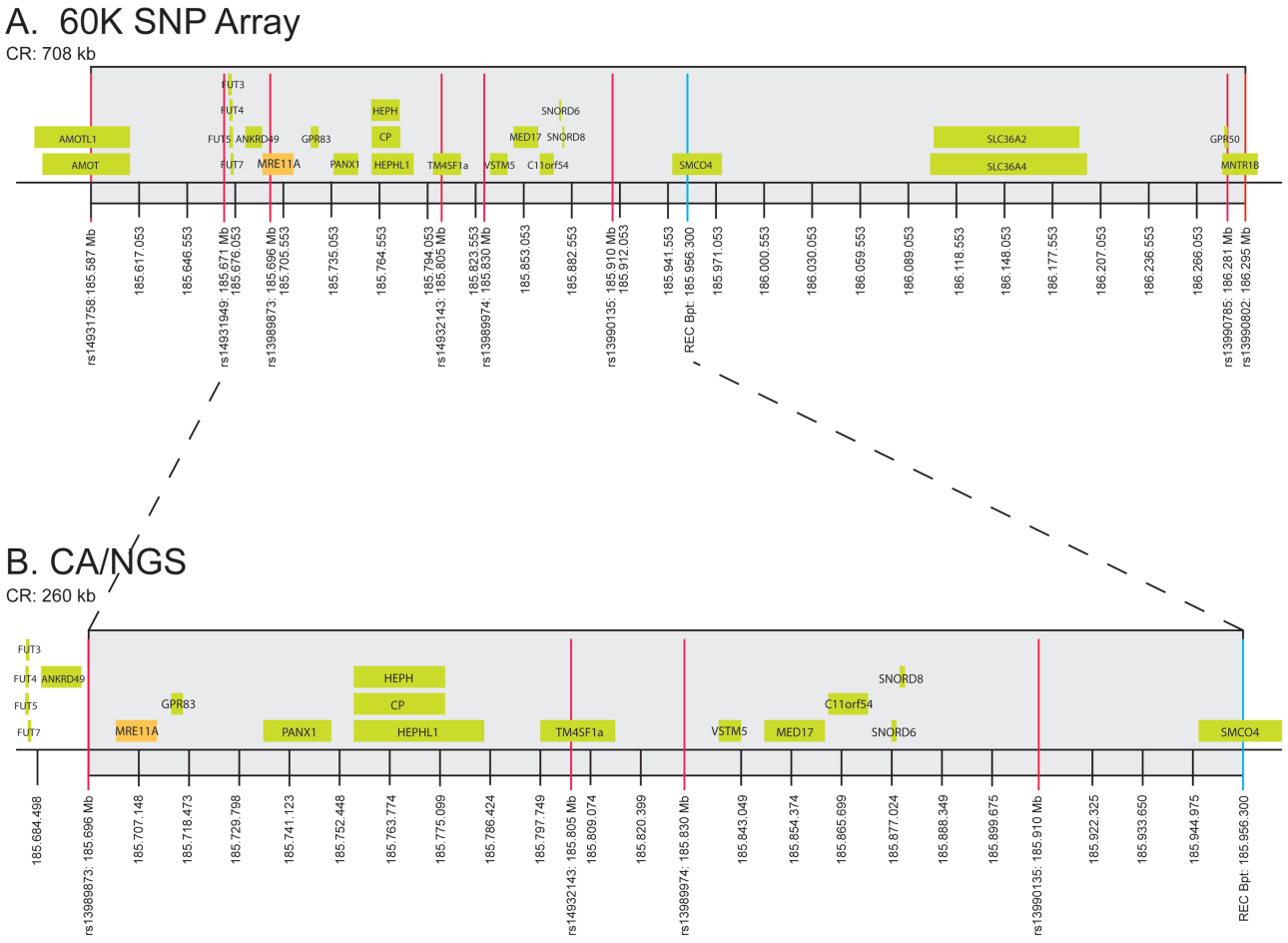

Figure 2. Chromosome 1 CR associated with the chicken diplopodia-1 mutation. **A**. *Causative/Linked Region Identified by the 60K SNP Array:* 8 polymorphic SNP markers (red lines) throughout the 707,738 bp CR identified by utilization of the Illumina 60K chicken iSelect SNP genotyping array. SNP markers rs14931758 and rs13990802 denote the boundaries of this region (GGA 1: 185,587,533-186,295,271). A total of 24 genes (green/orange boxes) are encoded within this region. The high-priority candidate gene, *MRE11* (Robb et al. 2011), is marked with an orange box. **B**. *Causative/Linked Region Identified by Analysis of the Capture Array Results (CA):* A 260,475 bp CR was established through SNP fine-mapping and capture array analyses. Assessment of the capture array data identified a unique dp-1 polymorphic SNP at GGA 1: 185,956,300 (blue line in both panels) which decreased the CR by 447 kb. This region encompasses 13 genes. *Note*: Coordinates and values shown in the figure represent those of the December 2015 *Gallus gallus* assembly (galGal5).

of a sequencing error), for linkage, and potential causation in the validation portion of this study, see below.

Validation of Elements Identified by Capture Enrichment and Assessment of Causation

It was a goal of this study to use advanced technologies to identify, and without bias, all sequence polymorphisms in the CR. Following that, an equally important aspect was to eliminate those polymorphic elements not involved in the *dp-1* mutation representing normal variation in the introgressed region. Finally, it was our goal to identify the top priority candidate elements or genes for future study. As such, the unique SNPs, micro-indels, and gaps found within an exon or splice site (thus potentially impacting mRNA and protein) of known and predicted genes (Table 3) were verified in an additional 20 mutant and 2 control ((UCD-003)) samples to confirm that a particular variant remained linked to the *dp-1* mutation. All validated elements (SNPs, micro-indels, variants found in gaps) have been submitted to NCBI (Accessions: ss475871243–ss475871302; ss475875438; ss475871304; ss475871305; ss475875441; ss475875442; ss475871308–ss475871332). Individual variant validation results are described below.

Validation of Exonic, UTR, and Splice Site SNPs, Micro-indels (1–3 nts), and Gaps in the 260 kb CR

A total of 49 SNPs were localized to an exon, UTR or splice site of a known or predicted gene (Table 6, Supplementary Table 1). The additional mutant sample analysis revealed that 41 of the 49 SNPs remain linked to the *Dp-1*.003 mutation. These variants reside within 7 (*HEPHL1*, *CP*, *HEPH*, *TM4SF1A*, *VSTM5*, *MED17*, and *C11ORF54*) of the 13 known/predicted genes in the 260 kb region.

Four micro-indels were found within an exon, UTR, or splice site (Table 6, Supplementary Table 1). Assessment of new mutant samples eliminated 2 of these variants—the insertion in *MRE11* (+CC) and the deletion in *C11ORF54*, which upon frameshift analysis would cause a disruption in the polypeptide sequences produced. Thus, 2 micro-indels remain linked to the mutation. One resides within the 3′ UTR of *C11ORF54* and the other within a splice site of *C11ORF54*.

Nine gaps (>4 nt) were identified within an exon, UTR, or splice site (Table 6, Supplementary Table 1). Of these, one gap identified no longer is localized to chromosome 1 based upon the new genome assembly update. Sequencing of the remaining 8 gaps identified the presence of sequence in 3 gaps identical to control DNA; thus, these genome gaps were likely due to a sequencing error during completion of the CA. Furthermore, validation efforts reliably identified sequence in 3 gaps; however, upon DNA alignment, 6 SNPs and one insertion reside within these regions. These newly identified variants remain linked to the *Dp-1*.003 mutation and are found within an exon and splice site of *MED17* and within an exon of *C11ORF75*. Moreover, 2 gaps remain linked due to sequencing validation complications attributed to the high G-C percentage and polymerase slippage during sequencing. These gaps (aka ≥4 nt deletions) are found within the 3′ UTR of *PANX1* and the 5′ UTR/splice site of *C11ORF54* (Supplementary Table 1).

Developmental Gene Expression of Candidates by RNA *In situ* Hybridization

To assist in our prioritization of the likelihood that a gene is involved in the *dp-1* mutation, the expression patterns of 11 protein-coding genes (*MRE11*, *GPR83*, *PANX1*, *HEPHL1*, *CP*, *HEPH*, *TM4SF1A*, *VSTM5*, *C11ORF54*, *MED17*, and *SMCO4*) mapping within the 260 kb CR were studied during early embryonic stages in normal (+/+) chicken embryos. Eight of these genes did not show expression during early embryonic developmental stages (Figure 3D–K), while one gene (*PANX1*; Figure 3C) displayed a positive staining pattern of expression but not in the regions of interest (e.g., limbs, pharyngeal arches/clefts, facial prominences). Interestingly, 2 genes, *MRE11* and *GPR83*, which are less than 5000 bp apart, show positive expression results in which appropriate, localized expression

Table 6. Genomic location of *dp-1*-specific SNPs, micro-indels, and sequence gaps within the 260 kb CR

	SNP location[a]			Insertion location[b]			Deletion location[c]			Gap location[d]		
Genes	Exon	Splice site	Intron	Exon	Splice site	Intron	Exon	Splice site	Intron	Exon	Splice site	Intron
MRE11	0	3	18	0	1	0	0	0	1	1	0	11
GPR83	1	0	1	0	0	0	0	0	0	1[e]	0	1
PANX1	0	0	7	0	0	2	0	0	1	1[f]	0	15
CP	3	0	50	0	0	2	0	0	0	0	0	14
HEPH	1	1	50	0	0	2	0	0	0	0	0	14
HEPHL1	5	1	66	0	0	2	0	0	0	0	0	20
TM4SF1	3[e]	2	40	0	0	2	0	0	3	1[e]	0	10
VSTM5	3	0	16	0	0	1	0	0	0	0	0	3
MED17	7[e]	14	54	0	0	3	0	0	1	3[ef]	0	23
C11ORF54	3	2	32	0	1	2	1[e]	1	2	1[f]	0	13
SNORD6	0	0	0	0	0	0	0	0	0	0	0	0
SNORD8	0	0	0	0	0	0	0	0	0	0	0	0
SMCO4	1[e]	0	45	0	0	7	0	0	2	1[f]	0	26

[a]All other unique SNPs (n = 454) are non-genic (not found within a gene).
[b]All other unique insertions (n = 26) are non-genic.
[c]All other unique deletions (n = 27) are non-genic.
[d]All other gaps (n = 554) are non-genic.
[e]One variant is located in the 3′ UTR of the gene.
[f]One variant is located in the 5′ UTR of the gene.

was observed during stages of embryogenesis wherein organogenesis, limb, and craniofacial development occur (Figure 3A,B).

The *dp-1* mutant phenotype displays abnormalities within several main anatomical features (i.e., somites, mandibular arch, head and limb regions). At HH24, a stage in which the limb buds and maxillary arches are very distinct, strong *MRE11* expression is found within the developing limbs, maxillary process, mandibular and hyoid arches, optic cup, somites, ventral neural tube, and brain (mesencephalon and telencephalon) (Figure 3A). Additionally, *GPR83* expression at the same stage can be clearly observed in the limbs, maxillary process, mandibular and hyoid arches, optic cup, somites, ventral neural tube, ventricle of the heart and brain (telencephalon) (Figure 3B). The expression pattern for these genes at additional developmental stages can be found at http://geisha.arizona.edu/geisha/.

Discussion

The chick embryo shares many features (anatomy, function, organization, cellular composition, and molecular pathways) in common with higher vertebrates, including mammals, during development (Le Douarin 2004; Tickle 2004; Stern 2005). This allows for valuable comparison between similar chick (e.g., dp-1) and human malformations (e.g., diplopodia). Despite the advanced state of the human genome sequence, not all heritable disorders have been linked to specific genes. Therefore, the chick embryo, specifically the developmental mutant *dp-1* described herein, is an excellent system in which to search for and identify a causative element, with the hope that this information can then be applied and developed further for screening and therapeutic applications of human malformations. Moreover, such animal model research adds knowledge to known developmental pathways and can uncover new molecular networks and gene involvement in normal vertebrate development.

Our aim was to generate new information on the genomics of the candidate region as well as the molecular, developmental, and morphological processes that contribute to the dp-1 phenotype. *Dp-1* mutants develop normal digits (digits 2 through 4 are normal on the foot; digits 2 and 3 on the wing are normal); however, the most anterior digit, digit 1, is accompanied by 2–3 digit supernumeraries (Figure 1) (MacCabe and Abbott 1974; MacCabe et al. 1975). The *dp-1* mutant also exhibits truncated limbs (micromelia), dwarfism, exposed visceral organs, and cleft palate. These conditions are seen individually or combined in reported cases of human congenital disorders (reviewed in Vlahovic et al. 2015). Interestingly, the extent of phenotypic variation observed in mutants of the dp-1.003 congenic inbred line was reduced compared to initial reports on the *dp-1* mutation. We hypothesize that through the establishment of the congenic line (breeding the *dp-1* mutation on the highly inbred UCD-003 background, $F > 0.999$), variant sequence elements (e.g., promoter, epistatic influences, etc.) contributing to the phenotypic variability observed in the 1940s through 1970s were eliminated, thereby creating the relatively uniform phenotype currently observed.

To date, *dp-1* has been well-studied as to processes involved in normal and mutant limb development (Taylor and Gunns 1947; Landauer 1956; Abbott 1959a, 1959b; Taylor et al. 1959; MacCabe and Abbott 1974; Taylor 1974; MacCabe et al. 1975; Rodriguez et al. 1996). Tissue interaction studies revealed the defect to be of mesodermal origin (MacCabe and Abbott 1974; MacCabe et al. 1975). Rodriguez et al. (1996) studied several genes in the *dp-1* mutants known to be involved in limb formation and digit outgrowth (*SHH*, *HOXD*, *BMP2*, *FGF4*). The expression patterns for *SHH* and *HOXD* were similar to normal embryos. However, *BMP2* and *FGF4* expression was expanded. Noteworthy, none of these genes reside within the *dp-1*-linked CR of GGA 1 (this study and Robb et al. 2011); thus, the causative element of dp-1 may be operating within a novel pathway or is a critical upstream element in a known pathway.

In this study, we assessed the RNA expression patterns of 11 protein-coding genes (*MRE11*, *GPR83*, *PANX1*, *HEPHL1*, *CP*, *HEPH*, *TM4SF1A*, *VSTM5*, *C11ORF54*, *MED17*, and *SMCO4*) within the 260 kb fine-mapped region. Of these, 2 proximally close genes (*MRE11* and *GPR83*) showed strong expression (Figure 3A,B) during the appropriate stages of development wherein main morphological features (limbs, facial prominences, pharyngeal arches/clefts, somites) affected in the *dp-1* mutation start to and become elaborated. To determine if the function of these genes is conserved across vertebrates, we performed *Mre11* RNA *in situ* hybridization in mouse and obtained similar results (Supplementary Figure 1). Histological analyses reported for *GPR83*, in mouse are also consistent with what was observed here (http://www.informatics.jax.org/marker/MGI:95712). In other model organisms, such as the zebrafish, gene expression was reported in the pectoral fin buds and pharyngeal arches for *mre11* (Thisse and Thisse 2004); no studies on *gpr83* in the zebrafish have been published to date. Moreover, human fetuses at 10–20 weeks of gestation shows elevated *MRE11* and *GPR83* expression in the same tissues (NCBI: https://www.ncbi.nlm.nih.gov/gene; *geneID: 4361 and 10888*). Although a gene might not be expressed at these developmental stages (as in the case of the other 9 non-expressing genes), hypothetically, a polymorphism could induce (mis)expression. It is noteworthy, however, that even at later stages of development, the non-expressing genes found in this study are also not expressed in the mutant-relevant tissue locations (NCBI: https://www.ncbi.nlm.nih.gov/gene; European Bioinformatics Institute Expression Atlas: https://www.ebi.ac.uk/gxa/home; *The Protein Atlas:* https://www.proteinatlas.org/).

The primary goal of our genomics research was to discover and establish those variants, which might be involved in dp-1 through sequencing, linkage analysis/fine-mapping and variant validation in additional mutants or invalidation in non-mutants. Interestingly, splice and/or exonic variants remain linked in 8 genes (*PANX1*, *HEPHL1*, *CP*, *HEPH*, *TM4SF1A*, *VSTM5*, *C11ORF54*, *MED17*) but not in 5 genes (*MRE11*, *GPR83*, *SNORD6*, *SNORA8*, *SMCO4*). Although exonic and splice variants were eliminated as being causal in both *MRE11* and *GPR83*, a *dp-1*-linked regulatory element could impact their gene expression (i.e., cis-regulatory element) and/or modify genes in other molecular pathways (i.e., trans-regulatory element, e.g., altering *BMP2* or *FGF4* expression). This type of gene action is possible as mutations in regulatory elements commonly affect the expression pattern of a gene and notably has already been reported in the *dp-1* mutant with the expanded expression of *BMP2* and *FGF4* (Rodriguez et al. 1996).

Based on comparative vertebrate biology (mouse, human, zebrafish), pathway role (Supplementary Figure 2), and the spatio-temporal expression profiles in this work we suggest *MRE11* as a candidate of high interest for the *dp-1* mutation. Until recently, *GPR83* had only been identified for its role in peripheral T-cell development (Lu et al. 2007). However, our RNA *in situ* results (appropriate spatio-temporal expression, see Figure 3B) and recent functional studies (Müller et al. 2013) combined suggest an additional role of *GPR83* in limb and craniofacial development, and a potential role in the *dp-1* mutant phenotype. Below we outline more on the function of *MRE11* and *GPR83*, associated human diseases,

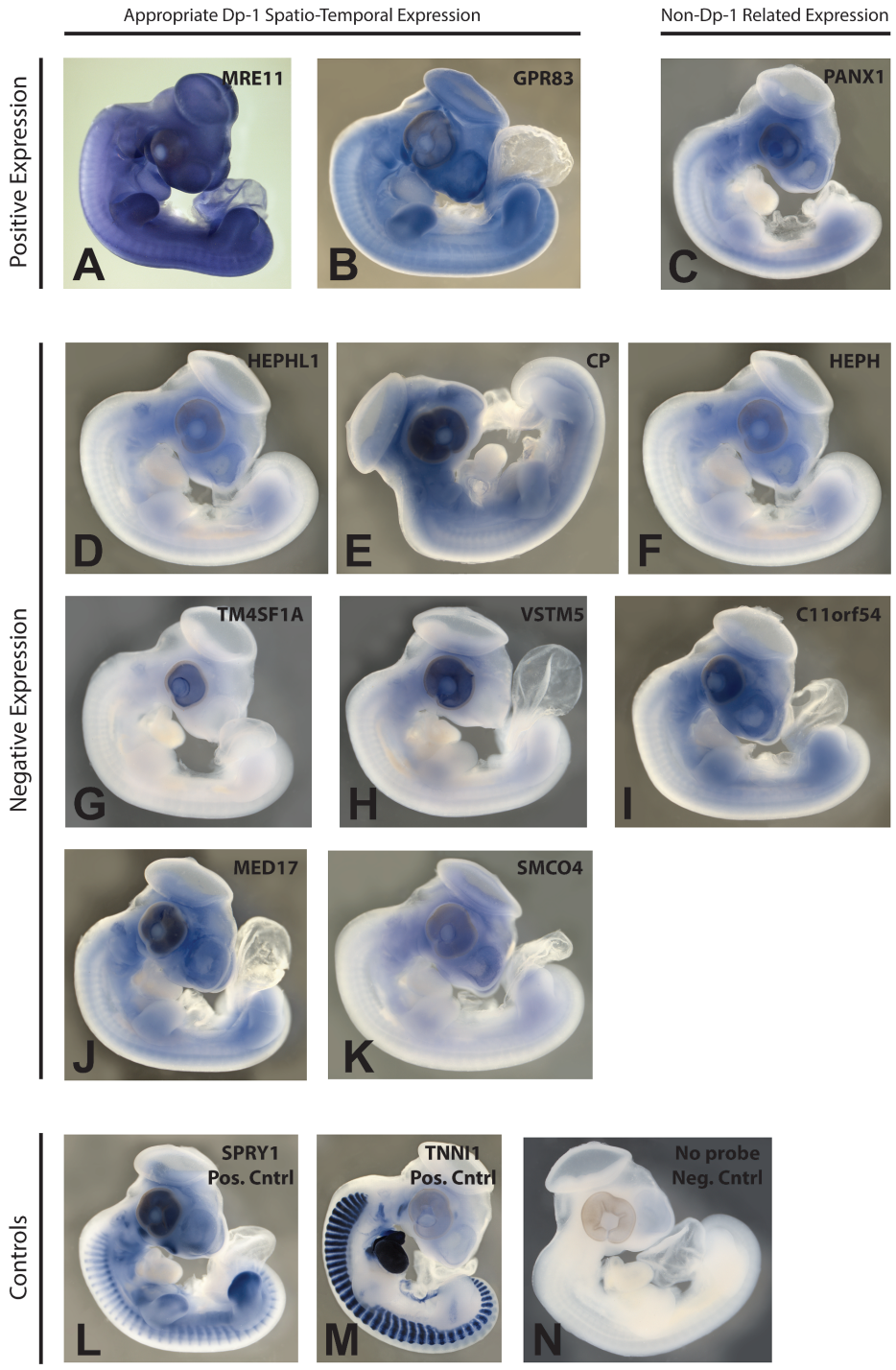

Figure 3. Priority candidate genes *MRE11* and *GPR83* show strong expression in structures leading to the formation of the face and limbs in normal chick embryos. RNA *in situ* hybridization was utilized to assess the expression of 11 candidate protein-encoding genes in the 260 kb region using standard procedures. Analysis of the results identified 3 genes with positive expression (A–C), with 2 (A,B) showing appropriate dp-1 spatio-temporal expression and the third showing expression in regions not affected in the dp-1 mutant embryo. Additionally, 8 genes showed negative expression (D–K). **A.** *MRE11* (HH24): high levels of expression in the limb buds, somites, pharyngeal arches/clefts, and brain. **B.** *GPR83* (HH24): punctate expression in the limb buds, somites, and pharyngeal arches/clefts. **C.** *PANX1* (HH22) expression is observed in the ventral neural tube (not shown: pancreatic expression identified via ventral cross-section). **D.** *HEPHL1* (HH24), only background is present. **E.** *CP* (HH25), only background staining is present. **F.** *HEPH* (HH24), only background staining is present. **G.** *TM4SF1A* (HH24), only background is present. **H.** *VSTM5* (HH24), only background staining is present. **I.** *C11ORF54* (HH23), only background staining is present. **J.** *MED17* (HH24), only background staining is present. **K.** *SMCO4* (HH24), only background staining is present. **L.** SPRY1 (HH24)—positive control: strong, highly specific expression in the limb buds, somites, pharyngeal arches/clefts, and brain. **M.** TNNI1 (HH24)—positive control: strong expression in the somites, atrium/ventricle, and pharyngeal arches. **N.** No probe (HH24)—negative control (no gene probe used): controls for background staining that may arise independent of the presence of an antisense probe. The expression patterns for these genes at developmental stages other than those shown in this figure can be found at http://geisha.arizona.edu/geisha/.

molecular pathway(s) of interest, as well as the next steps and implications of this work.

MRE11's Potential for a Contribution to the Dp-1 Phenotype

MRE11 (*m*eiotic *re*combination *11* homologue) is a protein with known involvement in DNA double-stranded break repair, homologous recombination, nonhomologous end joining, and telomere length maintenance (Lee and Paull 2005; Deng et al. 2009; Zha et al. 2009). Altered regulation and mutations in *MRE11* were shown to produce human pleiotropic disorders (e.g., Nijmegen breakage syndrome (Matsumoto et al. 2011), Ataxia Telangiectasia-like disorder (Stewart et al. 1999)) with affected developmental systems similar to that observed in *dp-1* mutants. Moreover, mutations in genes with similar functions to *MRE11* (e.g., *NIPBL*, *SMC1L1*) are associated with Cornelia de Lange syndrome (CdL) (Krantz et al. 2004; Tonkin et al. 2004; Kaur et al. 2005; Musio et al. 2006; Barber et al. 2008), which shares features in common with the *dp-1* phenotype such as limb and digit abnormalities, growth retardation, cleft palate, and organ malformations (Jackson et al. 1993; Kline et al. 2007; Liu and Krantz 2009).

Evaluation of downstream targets of MRE11 suggest 2 potential molecular mechanisms that could contribute to an aberrant phenotype such as dp-1 given potential MRE11 misexpression; these are highlighted in more detail in Supplementary Figure 2 (and corresponding Supplementary Material). Briefly, under normal circumstances MRE11 recruits ATM (a serine/threonine kinase) to a double-stranded break to initiate repair of DNA damage. In turn, ATM subsequently activates DNMT1 (Uziel et al. 2003), a DNA methyltransferase that preferentially methylates specific genomic regulatory regions, including that of the *NOGGIN* promoter (Shamma et al. 2013). Interestingly, NOGGIN is an antagonist of BMPs which bind directly to BMP receptors (Supplementary Figure 2A) and, as a result, plays an important role in many developmental processes for limb, craniofacial, and skeletal formation (Reshef et al. 1998; Wijgerde et al. 2005; Bayramov et al. 2011) by creating morphogenic gradients. Thus, misregulation of *NOGGIN* via DNMT1 methylation can lead to irregular BMP spatial and temporal expression (thereby altering craniofacial and limb development) (outlined in Supplementary Figure 2B,C), as was reported by Rodriguez et al. (1996) in their studies of *dp-1* mutants (Figure 1B–F, K–N). Additionally, ATM phosphorylates histone H2A variant, H2AX (Stiff et al. 2006), thereby inducing chromatin remodeling and subsequent apoptosis/senescence of the cells in the interdigit space of vertebrate embryos (Montero et al. 2016). Failure to trigger programed cell death results in interdigital webbing (Supplementary Figure 2A), another phenotype observed in *dp-1* mutants (Figure 1B–F). Thus, given these genetic networks, as well as the *MRE11* mutant human developmental phenotypes, we hypothesize that abnormal expression of *MRE11* (and as a result, effects on downstream pathways) is a likely contributor of the dp-1 phenotype.

GPR83's Potential for Contribution to the Dp-1 Phenotype

GPR83 (*G P*rotein-Coupled *R*eceptor *83*) is an orphan protein with an identified, but not well-understood role in regulatory T cell development via its interaction with Foxp3 (Lu et al. 2007). Additional expression analysis in the rat brain led researchers to propose a second role for *GPR83* in central processing of glucocorticoids (Sah et al. 2005). Specifically, GPR83 regulates metabolic pathways by binding to ghrelin's receptor (Ghsr1a) (Müller et al. 2013). When *Gpr83* is knocked out, mice show lower body weight, fat mass, food intake, glucose tolerance, and insulin sensitivity, as compared to control mice on the same diet (Müller et al. 2013), phenotypes which are consistent with that of a knockdown/out ghrelin receptor. These results therefore suggest that altered/mutated *Gpr83* disturbs native Ghsr1a function.

Interestingly, loss of *Ghsr1a* in mouse leads to short stature through reduced hypothalamic–pituitary–growth hormone activity (a phenotype observed in the *dp-1* mutant, Figure 1) as well as a hypoglycemic state (Pantel et al. 2006; Sun et al. 2006, 2008). Interestingly several studies outline the presence of both cleft lip/palate in children with neonatal hypoglycemia (De Leon and Stanley 2017). Moreover, neonatal hypoglycemia has been associated with Beckwith–Wiedemann syndrome, characterized by in midline abdominal wall defects (omphalocele) (Queensland Clinical Guidelines 2013) similar to that observed in the *dp-1* mutant (Figure 1G–J), as well as dysmorphic facial features (https://ghr.nlm.nih.gov/). Note that other congenital abnormalities are also found in neonatal hypoglycemic patients that are not observed in the *dp-1* mutants. Thus, based upon our spatio-temporal *in situ* expression results and comparative mouse/human studies, we suggest that *GPR83* should also be considered a priority candidate gene.

Final Thoughts and Research Implications

Although *Mre11* knockout results in early lethality (Buis et al. 2008) and *Gpr83* knockout mice survive to adulthood with no morphological abnormalities (Müller et al. 2013), our proposed hypothesis (misexpression via a polymorphism in the 260-kb region) of one, or perhaps even both genes deserve future analysis especially with regard to intronic elements within *MRE11* (18 SNPs, 1 indel) and *GPR83* (1 SNP) that remain linked to the phenotype in addition to a SNP in the intergenic region between these 2 proximally close genes.

Performing such research is of course a complicated effort. Moreover, the proximity of the 2 high priority candidate genes (5000 bp apart), makes it even more difficult (and yet interesting given the mode of inheritance) to further eliminate variants using breeding techniques and recombination/linkage analyses. And, although much has been discovered regarding gene regulation, status/knowledge of regulatory motifs and their impacts remains obscure. Functional studies will be essential to explore the remaining variants and their contributing role in *MRE11* and *GPR83* expression toward the dp-1 phenotype.

Through this work we successfully discovered unique genetic variants specific to the *dp-1* mutation, yet despite the advances of genomics and allied fields, this work affirms what other such research has shown, that finding the exact sequence element responsible for even a simply-inherited trait (single-gene autosomal recessive pattern) is not easily accomplished. In addition to more studies focusing on *MRE11* and *GPR83*, evaluation of elements found within the 260 kb-linked region, beyond those found within exons, UTRs and splice sites reported here, will be a contribution to uncovering the causative sequence. Identification of the causal mutation for dp-1 will be valuable to not only the field of developmental biology, shedding light on the mechanisms involved in normal/abnormal development, but also to the fields of genetics and genomics by elucidating the pathways involved in the regulation and expression of various genes and involved pathways. Ultimately, the application of this research in the medical field and the implementation of unique genetic lines as biomedical models to study cleft palate, dwarfism, polydactyly, and

omphalocele, could lead to screening systems for congenital malformations as well as research opportunities for therapeutic remedies.

Supplementary Material

Supplementary data are available at *Journal of Heredity* online.

Funding

This work was supported by the U.S. Department of Agriculture-National Institute of Food and Agriculture Multistate Research NC-1170 (CA-D*-ASC-6414-RR), the NRSP-8 National Animal Genome Research Support Program (CA-D*-ASC-7233-RR) and the UC Davis (UCD) John and Joan Fiddyment endowment (M.E.D.). The *in situ* hybridization studies were supported by National Institutes of Health grant HD064559 (P.B.A.). We gratefully acknowledge the UCD Department of Animal Science for graduate student fellowship support and the UCD Jastro-Shields Graduate Research Award (E.A.O.). We appreciate the infrastructure and the poultry genetic resources supported by the UC Davis Department of Animal Science, College of Agricultural and Environmental Sciences, and the California Agricultural Experiment Station.

Acknowledgments

We thank Dr Jackie Pisenti for her dedication to the preservation of the UC Davis poultry genetic resources and for her attention to the breeding and phenotypes of the UCD developmental congenic lines. We greatly appreciate the anonymous reviewers for their careful review and comments that contributed to an improved manuscript.

Data Availability

We have deposited the primary data underlying these analyses as follows:
-Validated sequence variants: NCBI accessions: ss475871243-ss475871302; ss475875438; ss475871304; ss475871305; ss475875441; ss475875442; ss475871308–ss475871332.
-RNA *in situ* hybridization gene images: GEISHA—Gallus Expression in Situ Hybridization Analysis (http://geisha.arizona.edu/geisha/).

References

Abbott UK. 1959a. Further studies on diplopodia. II. Embryological features. *J Genet*. 56:179–196.

Abbott UK. 1959b. Further studies on diplopodia. III. The relationship between expressivity and penetrance in selected lines and crosses. *J Genet*. 56:197–225.

Abplanalp H. 1992. Inbred lines as genetic resources of chickens. *Poult Sci Rev*. 4:29–39.

Antin PB, Kaur S, Stanislaw S, Davey S, Konieczka JH, Yatskievych TA, Darnell DK. 2007. Gallus expression in situ hybridization analysis: a chicken embryo gene expression database. *Poult Sci*. 86:1472–1477.

Baek HJ, Kang YK, Roeder RG. 2006. Human Mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. *J Biol Chem*. 281:15172–15181.

Barber TD, McManus K, Yuen KW, Reis M, Parmigiani G, Shen D, Barrett I, Nouhi Y, Spencer F, Markowitz S, *et al*. 2008. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proc Natl Acad Sci USA*. 105:3443–3448.

Bayramov AV, Eroshkin FM, Martynova NY, Ermakova GV, Solovieva EA, Zaraisky AG. 2011. Novel functions of Noggin proteins: inhibition of Activin/Nodal and Wnt signaling. *Development*. 138:5345–5356.

Brandström M, Ellegren H. 2007. The genomic landscape of short insertion and deletion polymorphisms in the chicken (*Gallus gallus*) Genome: a high frequency of deletions in tandem duplicates. *Genetics*. 176:1691–1701.

Brookes MJ, Hughes S, Turner FE, Reynolds G, Sharma N, Ismail T, Berx G, McKie AT, Hotchin N, Anderson GJ, *et al*. 2006. Modulation of iron transport proteins in human colorectal carcinogenesis. *Gut*. 55:1449–1460.

Brower JS, Wootton-Gorges SL, Costouros JG, Boakes J, Greenspan A. 2003. Congenital diplopodia. *Pediatr Radiol*. 33:797–799.

Buis J, Wu Y, Deng Y, Leddon J, Westfield G, Eckersdorff M, Sekiguchi JM, Chang S, Ferguson DO. 2008. Mre11 nuclease activity has essential roles in DNA repair and genomic stability distinct from ATM activation. *Cell*. 135:85–96.

Burt DW. 2007. Emergence of the chicken as a model organism: implications for agriculture and biology. *Poult Sci*. 86:1460–1471.

Chekeni FB, Elliott MR, Sandilos JK, Walk SF, Kinchen JM, Lazarowski ER, Armstrong AJ, Penuela S, Laird DW, Salvesen GS, *et al*. 2010. Pannexin 1 channels mediate 'find-me' signal release and membrane permeability during apoptosis. *Nature*. 467:863–867.

Chen H, Attieh ZK, Syed BA, Kuo YM, Stevens V, Fuqua BK, Andersen HS, Naylor CE, Evans RW, Gambling L, *et al*. 2010. Identification of zyklopen, a new member of the vertebrate multicopper ferroxidase family, and characterization in rodents and human cells. *J Nutr*. 140:1728–1735.

Darnell DK, Kaur S, Stanislaw S, Davey S, Konieczka JH, Yatskievych TA, Antin PB. 2007. GEISHA: an in situ hybridization gene expression resource for the chicken embryo. *Cytogenet Genome Res*. 117:30–35.

Davey MG, Paton IR, Yin Y, Schmidt M, Bangs FK, Morrice DR, Smith TG, Buxton P, Stamataki D, Tanaka M, *et al*. 2006. The chicken talpid3 gene encodes a novel protein essential for Hedgehog signaling. *Genes Dev*. 20:1365–1377.

Davey MG, Tickle C. 2007. The chicken as a model for embryonic development. *Cytogenet Genome Res*. 117:231–239.

Davey MG, Towers M, Vargesson N, Tickle C. 2018. The chick limb: embryology, genetics and teratology. *Int J Dev Biol*. 62:85–95.

Delany ME. 2004. Genetic variants for chick biology research: from breeds to mutants. *Mech Dev*. 121:1169–1177.

De Leon DD, Stanley CA. 2017. Congenital hypoglycemia disorders: new aspects of etiology, diagnosis, treatment and outcomes: highlights of the proceedings of the congenital hypoglycemia disorders symposium, Philadelphia April 2016. *Pediatr Diabetes*. 18:3–9.

Deng Y, Guo X, Ferguson DO, Chang S. 2009. Multiple roles for MRE11 at uncapped telomeres. *Nature*. 460:914–918.

Dieci G, Preti M, Montanini B. 2009. Eukaryotic snoRNAs: a paradigm for gene expression flexibility. *Genomics*. 94:83–88.

Duprez D, Bell EJ, Richardson MK, Archer CW, Wolpert L, Brickell PM, Francis-West PH. 1996. Overexpression of BMP-2 and BMP-4 alters the size and shape of developing skeletal elements in the chick limb. *Mech Dev*. 57:145–157.

Elferink MG, van As P, Veenendaal T, Crooijmans RP, Groenen MA. 2010. Regional differences in recombination hotspots between two chicken populations. *BMC Genet*. 11:11.

Foppiano S, Hu D, Marcucio RS. 2007. Signaling by bone morphogenetic proteins directs formation of an ectodermal signaling center that regulates craniofacial development. *Dev Biol*. 312:103–114.

Frazer DM, Vulpe CD, McKie AT, Wilkins SJ, Trinder D, Cleghorn GJ, Anderson GJ. 2001. Cloning and gastrointestinal expression of rat hephaestin: relationship to other iron transport proteins. *Am J Physiol Gastrointest Liver Physiol*. 281:G931–G939.

Fukuda T, Sumiyoshi T, Takahashi M, Kataoka T, Asahara T, Inui H, Watatani M, Yasutomi M, Kamada N, Miyagawa K. 2001. Alterations of the double-strand break repair gene MRE11 in cancer. *Cancer Res*. 61:23–26.

Gatei M, Young D, Cerosaletti KM, Desai-Mehta A, Spring K, Kozlov S, Lavin MF, Gatti RA, Concannon P, Khanna K. 2000. ATM-dependent phosphorylation of nibrin in response to radiation exposure. *Nat Genet*. 25:115–119.

Gitlin JD. 1998. Aceruloplasminemia. *Pediatr Res*. 44:271–276.

Griffiths TA, Mauk AG, MacGillivray RT. 2005. Recombinant expression and functional characterization of human hephaestin: a multicopper oxidase with ferroxidase activity. *Biochemistry*. 44:14725–14731.

Groenen MA, Wahlberg P, Foglio M, Cheng HH, Megens HJ, Crooijmans RP, Besnier F, Lathrop M, Muir WM, Wong GK, *et al*. 2009. A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Res*. 19:510–519.

Hahn P, Qian Y, Dentchev T, Chen L, Beard J, Harris ZL, Dunaief JL. 2004. Disruption of ceruloplasmin and hephaestin in mice causes retinal iron overload and retinal degeneration with features of age-related macular degeneration. *Proc Natl Acad Sci U S A*. 101:13850–13855.

Hamanishi C, Ueba Y, Iwashita Y, Yamamuro T. 1985. Diplopodia with reversed foot. Normal gait after operation at 8 years of age. *Acta Orthop Scand*. 56:439–441.

Hamburger V, Hamilton HL. 1951. A series of normal stages in the development of the chick embryo. *J Morphol*. 88:49–92.

Hellman NE, Gitlin JD. 2002. Ceruloplasmin metabolism and function. *Annu Rev Nutr*. 22:439–458.

Holmberg CG, Laurell C-B. 1948. Investigations in serum copper. II. Isolation of the copper containing protein, and a description of its properties. *Acta Chem Scand*. 2:550–556.

Hu D, Colnot C, Marcucio RS. 2008. Effect of bone morphogenetic protein signaling on development of the jaw skeleton. *Dev Dyn*. 237:3727–3737.

International Chicken Polymorphism Map Consortium (ICPMC). 2004. A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature*. 432:717–722.

Jackson L, Kline AD, Barr MA, Koch S. 1993. de Lange syndrome: a clinical review of 310 individuals. *Am J Med Genet*. 47:940–946.

Jacobsen E, Beach T, Shen Y, Li R, Chang Y. 2004. Deficiency of the Mre11 DNA repair complex in Alzheimer's disease brains. *Brain Res Mol Brain Res*. 128:1–7.

Jones D, Barnes J, Lloyd-Roberts GC. 1978. Congenital aplasia and dysplasia of the tibia with intact fibula. Classification and management. *J Bone Joint Surg Br*. 60:31–39.

Kadir KH, Abdul Rashid AH, Das S, Ibrahim S. 2011. A rare case of diplopodia and syndactyly: anatomical and surgical considerations. *J Foot Ankle Surg*. 50:252–256.

Karchinov K. 1973. Congenital diplopodia with hypoplasia or aplasia of the tibia. A report of six cases. *J Bone Joint Surg Br*. 55:604–611.

Kaufmann R, Straussberg R, Mandel H, Fattal-Valevski A, Ben-Zeev B, Naamati A, Shaag A, Zenvirt S, Konen O, Mimouni-Bloch A, *et al*. 2010. Infantile cerebral and cerebellar atrophy is associated with a mutation in the MED17 subunit of the transcription preinitiation mediator complex. *Am J Hum Genet*. 87:667–670.

Kaur M, DeScipio C, McCallum J, Yaeger D, Devoto M, Jackson LG, Spinner NB, Krantz ID. 2005. Precocious sister chromatid separation (PSCS) in Cornelia de Lange syndrome. *Am J Med Genet A*. 138:27–31.

Khan SA, Kumar A, Varhney MK. 2008. A rare association of deformities with diplopodia, aplasia of the tibia and double fibula: a case report. *J Med Case Rep*. 2:102.

Kienitz MC, Bender K, Dermietzel R, Pott L, Zoidl G. 2011. Pannexin 1 constitutes the large conductance cation channel of cardiac myocytes. *J Biol Chem*. 286:290–298.

Kline AD, Grados M, Sponseller P, Levy HP, Blagowidow N, Schoedel C, Rampolla J, Clemens DK, Krantz I, Kimball A, *et al*. 2007. Natural history of aging in Cornelia de Lange syndrome. *Am J Med Genet C Semin Med Genet*. 145C:248–260.

Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, *et al*. 2004. Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of *Drosophila melanogaster* Nipped-B. *Nat Genet*. 36:631–635.

Lai CP, Bechberger JF, Thompson RJ, MacVicar BA, Bruzzone R, Naus CC. 2007. Tumor-suppressive effects of pannexin 1 in C6 glioma cells. *Cancer Res*. 67:1545–1554.

Landauer W. 1956. A second diplopod mutation in the fowl. *J Hered*. 47:57–63.

Le Douarin NM. 2004. The avian embryo as a model to study the development of the neural crest: a long and still ongoing story. *Mech Dev*. 121:1089–1102.

Lee AR, Ko KW, Lee H, Yoon YS, Song MR, Park CS. 2016. Putative cell adhesion membrane protein Vstm5 regulates neuronal morphology and migration in the central nervous system. *J Neurosci*. 36:10181–10197.

Lee JH, Paull TT. 2005. ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. *Science*. 308:551–554.

Liu J, Krantz ID. 2009. Cornelia de Lange syndrome, cohesin, and beyond. *Clin Genet*. 76:303–314.

Liu W, Selever J, Murali D, Sun X, Brugger SM, Ma L, Schwartz RJ, Maxson R, Furuta Y, Martin JF. 2005. Threshold-specific requirements for Bmp4 in mandibular development. *Dev Biol*. 283:282–293.

Lu LF, Gavin MA, Rasmussen JP, Rudensky AY. 2007. G protein-coupled receptor 83 is dispensable for the development and function of regulatory T cells. *Mol Cell Biol*. 27:8065–8072.

MacCabe JA, Abbott UK. 1974. Polarizing and maintenance activities in two polydactylous mutants of the fowl: diplopodia and talpid. *J Embryol Exp Morphol*. 31:735–746.

MacCabe JA, MacCabe AB, Abbott UK, McCarrey JR. 1975. Limb development in diplopodia4: a polydactylous mutation in the chicken. *J Exp Zool*. 191:383–394.

Manjasetty BA, Büssow K, Fieber-Erdmann M, Roske Y, Gobom J, Scheich C, Götz F, Niesen FH, Heinemann U. 2006. Crystal structure of Homo sapiens PTD012 reveals a zinc-containing hydrolase fold. *Protein Sci*. 15:914–920.

Marken JS, Schieven GL, Hellström I, Hellström KE, Aruffo A. 1992. Cloning and expression of the tumor-associated antigen L6. *Proc Natl Acad Sci U S A*. 89:3503–3507.

Matsumoto Y, Miyamoto T, Sakamoto H, Izumi H, Nakazawa Y, Ogi T, Tahara H, Oku S, Hiramoto A, Shiiki T, *et al*. 2011. Two unrelated patients with MRE11A mutations and Nijmegen breakage syndrome-like severe microcephaly. *DNA Repair (Amst)*. 10:314–321.

Montero JA, Sanchez-Fernandez C, Lorda-Diez CI, Garcia-Porrero JA, Hurle JM. 2016. DNA damage precedes apoptosis during the regression of the interdigital tissue in vertebrate embryos. *Sci Rep*. 6:35478.

Müller TD, Müller A, Yi CX, Habegger KM, Meyer CW, Gaylinn BD, Finan B, Heppner K, Trivedi C, Bielohuby M, *et al*. 2013. The orphan receptor Gpr83 regulates systemic energy metabolism via ghrelin-dependent and ghrelin-independent mechanisms. *Nat Commun*. 4:1968.

Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S, Vezzoni P, Larizza L. 2006. X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations. *Nat Genet*. 38:528–530.

Narang IC, Mysorekar VR, Mathur BP. 1982. Diplopodia with double fibula and agenesis of tibia. A case report. *J Bone Joint Surg Br*. 64:206–209.

Nieto MA, Patel K, Wilkinson DG. 1996. In situ hybridization analysis of chick embryos in whole mount and tissue sections. *Methods Cell Biol*. 51:219–235.

Norrie JL, Lewandowski JP, Bouldin CM, Amarnath S, Li Q, Vokes MS, Ehrlich LIR, Harfe BD, Vokes SA. 2014. Dynamics of BMP signaling in limb bud mesenchyme and polydactyly. *Dev Biol*. 393:270–281.

Pantel J, Legendre M, Cabrol S, Hilal L, Hajaji Y, Morisset S, Nivot S, Vie-Luton MP, Grouselle D, de Kerdanet M, *et al*. 2006. Loss of constitutive activity of the growth hormone secretagogue receptor in familial short stature. *J Clin Invest*. 116:760–768.

Petrini JH, Walsh ME, DiMare C, Chen XN, Korenberg JR, Weaver DT. 1995. Isolation and characterization of the human MRE11 homologue. *Genomics*. 29:80–86.

Pisenti JM, Delany ME, Taylor RL, Abbott UK, Abplanalp H, Arthur JA, Bakst MR, Baxter-Jones C, Bitgood CC, Bradley FA, *et al*. 1999. *Avian genetic resources at risk. an assessment and proposal for conservation of genetic stocks in the USA and Canada*. Davis (CA): University of California Division of Agriculture and Natural Resources, Genetic Resources Conservation Program. Available from: http://grcp.ucdavis.edu/publications/index.htm. Report No. 20. (accessed January 11, 2019).

Prohaska JR. 2011. Impact of copper limitation on expression and function of multicopper oxidases (ferroxidases). *Adv Nutr*. 2:89–95.

Queensland Clinical Guidelines (QCG). 2013. Newborn hypoglycaemia. Available from: https://www.health.qld.gov.au/__data/assets/pdf_file/0019/142156/g-hypogly.pdf (accessed August 1, 2018).

Reshef R, Maroto M, Lassar AB. 1998. Regulation of dorsal somitic cell fates: BMPs and Noggin control the timing and pattern of myogenic regulator expression. *Genes Dev*. 12:290–303.

Robb EA. 2012. *Developmental syndromes in the chicken: mapping and candidate gene studies* [PhD thesis]. [Davis (CA)]: University of California, Davis. p. 463.

Robb EA, Antin PB, Delany ME. 2013. Defining the sequence elements and candidate genes for the *Coloboma* mutation. *PLoS One*. 8:e60267.

Robb EA, Delany ME. 2012a. The expression of preaxial polydactyly is influenced by modifying genetic elements and is not maintained by chromosomal inversion in an avian biomedical model. *Cytogenet Genome Res*. 136:50–68.

Robb EA, Delany ME. 2012b. Case study of sequence capture enrichment technology: identification of variation underpinning developmental syndromes in an amniote model. *Genes (Basel)*. 3:233–247.

Robb EA, Gitter CL, Cheng HH, Delany ME. 2011. Chromosomal mapping and candidate gene discovery of chicken developmental mutants and genome-wide variation analysis of MHC congenics. *J Hered*. 102:141–156.

Rodriguez C, Kos R, Macias D, Abbott UK, Izpisúa Belmonte JC. 1996. Shh, HoxD, Bmp-2, and Fgf-4 gene expression during development of the polydactylous talpid2, diplopodia1, and diplopodia4 mutant chick limb buds. *Dev Genet*. 19:26–32.

Sah R, Pritchard LM, Richtand NM, Ahlbrand R, Eaton K, Sallee FR, Herman JP. 2005. Expression of the glucocorticoid-induced receptor mRNA in rat brain. *Neuroscience*. 133:281–292.

Scheinberg IH, Gitlin D. 1952. Deficiency of ceruloplasmin in patients with hepatolenticular degeneration (Wilson's disease). *Science*. 116:484–485.

Schock EN, Chang CF, Youngworth IA, Davey MG, Delany ME, Brugmann SA. 2016. Utilizing the chicken as an animal model for human craniofacial ciliopathies. *Dev Biol*. 415:326–337.

Shamma A, Suzuki M, Hayashi N, Kobayashi M, Sasaki N, Nishiuchi T, Doki Y, Okamoto T, Kohno S, Muranaka H, *et al*. 2013. ATM mediates pRB function to control DNMT1 protein stability and DNA methylation. *Mol Cell Biol*. 33:3113–3124.

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. 2001. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 29:308–311.

Sjöblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD, Mandelker D, Leary RJ, Ptak J, Silliman N, *et al*. 2006. The consensus coding sequences of human breast and colorectal cancers. *Science*. 314:268–274.

Stern CD. 2005. The chick; a great model system becomes even greater. *Dev Cell*. 8:9–17.

Stern C. 2018. The chick model system: a distinguished past and a great future. *Int J Dev Biol*. 62:1–4.

Stewart GS, Maser RS, Stankovic T, Bressan DA, Kaplan MI, Jaspers NG, Raams A, Byrd PJ, Petrini JH, Taylor AM. 1999. The DNA double-strand break repair gene hMRE11 is mutated in individuals with an ataxia-telangiectasia-like disorder. *Cell*. 99:577–587.

Stiff T, O'Driscoll M, Rief N, Iwabuchi K, Löbrich M, Jeggo PA. 2006. ATM and DNA-PK function redundantly to phosphorylate H2AX after exposure to ionizing radiation. *Cancer Res*. 64:2390–2396.

Sun Y, Asnicar M, Saha PK, Chan L, Smith RG. 2006. Ablation of ghrelin improves the diabetic but not obese phenotype of ob/ob mice. *Cell Metab*. 3:379–386.

Sun Y, Butte NF, Garcia JM, Smith RG. 2008. Characterization of adult ghrelin and ghrelin receptor knockout mice under positive and negative energy balance. *Endocrinology*. 149:843–850.

Taylor LW. 1972. Further studies on diplopodia. V. Diplopodia-3. *Can J Genet Cytol*. 14:417–422.

Taylor LW. 1974. Further studies on diplopodia. VI. Segregation of diplopod embryos from chickens heterozygous for both dp-1 and dp-3 genes. *Can J Genet Cytol*. 16:121–135.

Taylor LW, Abbott UK, Gunns CA. 1959. Further studies on diplopodia. I. Modification of phenotypic segregation ratios by selection. *J Genet*. 56:161–178.

Taylor LW, Gunns CA. 1947. Diplopodia; a lethal form of polydactyly in chickens. *J Hered*. 38:67–76.

The Chick Embryo Model System. 2018. *Int J Dev Biol*. 62(1/2/3):1–272 (all articles).

Thisse B, Thisse C. 2004. Fast release clones: a high throughput expression analysis. ZFIN Direct Data Submission. Available from: http://zfin.org

Tickle C. 2004. The contribution of chicken embryology to the understanding of vertebrate limb development. *Mech Dev*. 121:1019–1029.

Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T. 2004. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. *Nat Genet*. 36:636–641.

Trujillo KM, Yuan SS, Lee EY, Sung P. 1998. Nuclease activities in a complex of human recombination and DNA repair factors Rad50, Mre11, and p95. *J Biol Chem*. 273:21447–21450.

Uziel T, Lerenthal Y, Moyal L, Andegeko Y, Mittelman L, Shiloh Y. 2003. Requirement of the MRN complex for ATM activation by DNA damage. *EMBO J*. 22:5612–5621.

Vanden Abeele F, Bidaux G, Gordienko D, Beck B, Panchin YV, Baranova AV, Ivanov DV, Skryma R, Prevarskaya N. 2006. Functional implications of calcium permeability of the channel formed by pannexin 1. *J Cell Biol*. 174:535–546.

Virit O, Selek S, Bulut M, Savas HA, Celik H, Erel O, Herken H. 2008. High ceruloplasmin levels are associated with obsessive compulsive disorder: a case control study. *Behav Brain Funct*. 4:52.

Vlahovic AM, Pistignjat BS, Vlahovic NS. 2015. Nine toes; Mirror Foot deformity. *Indian J Orthop*. 49:478–481.

Vulpe CD, Kuo YM, Murphy TL, Cowley L, Askwith C, Libina N, Gitschier J, Anderson GJ. 1999. Hephaestin, a ceruloplasmin homologue implicated in intestinal iron transport, is defective in the sla mouse. *Nat Genet*. 21:195–199.

Webb AE, Youngworth IA, Kaya M, Gitter CL, O'Hare EA, May B, Cheng HH, Delany ME. 2018. Narrowing the *wingless-2* mutation to a 227 kb candidate region on chicken chromosome 12. *Poult Sci*. 97:1872–1880.

Wijgerde M, Karp S, McMahon J, McMahon AP. 2005. Noggin antagonism of BMP4 signaling controls development of the axial skeleton in the mouse. *Dev Biol*. 286:149–157.

Wolf TL, Kotun J, Meador-Woodruff JH. 2006. Plasma copper, iron, ceruloplasmin and ferroxidase activity in schizophrenia. *Schizophr Res*. 86:167–171.

Zha S, Boboila C, Alt FW. 2009. Mre11: roles in DNA repair beyond homologous recombination. *Nat Struct Mol Biol*. 16:798–800.

Zukauskas A, Merley A, Li D, Ang LH, Sciuto TE, Salman S, Dvorak AM, Dvorak HF, Jaminet SC. 2011. TM4SF1: a tetraspanin-like protein necessary for nanopodia formation and endothelial cell migration. *Angiogenesis*. 14:345–354.