Submission intended as an Article, Discoveries

Elaboration of the corticosteroid synthesis pathway in primates through a multi-step enzyme

Carrie F. Olson-Manning^{1,2}

¹ Augustana University, Sioux Falls, SD 57197, ² University of Chicago, Chicago, IL 60615

Corresponding author: carrie.olsonmanning@augie.edu

Abstract

Metabolic networks are complex cellular systems dependent on the interactions among, and regulation of, the enzymes in the network. However, the mechanisms that lead to the expansion of networks are not well understood. While gene duplication is a major driver of the expansion and functional evolution of metabolic networks, the effect and fate of retained duplicates in a network is poorly understood. Here, I study the evolution of an enzyme family that performs multiple subsequent enzymatic reactions in the corticosteroid pathway in primates to illuminate the mechanisms that shape network components following duplication. The products of the pathway (aldosterone, corticosterone, and cortisol) are steroid hormones that regulate metabolism and stress in tetrapods. These steroids are synthesized by a multi-step enzyme Cytochrome P450 11B (CYP11B) that performs subsequent steps on different carbon atoms of the steroid derivatives. Through ancestral state reconstruction and *in vitro* characterization, I find the ancestor of the CYP11B1 and CYP11B2 paralogs (in primates) had moderate ability to synthesize cortisol and aldosterone. Following duplication in the primate lineage the CYP11B1 homolog specialized on the production of cortisol while its paralog, CYP11B2, maintained its ability to perform multiple subsequent steps as in the ancestral pathway. Unlike CYP11B1, CYP11B2 could not specialize on the production of aldosterone because it is constrained to perform earlier steps in the corticosteroid synthesis pathway to achieve the final product aldosterone. These results suggest that pathway context, along with tissue-specific regulation, both play a role in shaping potential outcomes of metabolic network elaboration.

Introduction

Biochemical networks consist of consecutive chemical reactions carried out by enzymes. In eukaryotes, the same pathway can be expressed across multiple tissues with each tissue possibly requiring different amounts of pathway products. Substitutions that lead to changes in enzyme function or expression that optimize pathway production in one tissue may have deleterious pleiotropic consequences in other tissues (Stern and Orgogozo 2008; Wray et al. 2003; Carroll 2005). To avoid these deleterious pleiotropic consequences, there are several mechanisms that allow for tissue specific regulation of widely expressed biochemical pathways: tissue-specific post-translational modifications (Ikegami et al. 2014), expression of different parts of the pathway in different tissues (Stearns 2010; Ramsay, Rieseberg, and Ritland 2009; Coberly and Rausher 2008), allosteric regulation where a metabolite (Snášel and Pichová 2019; Gui, Lewis, and Vander Heiden 2013) or another protein (Ikushiro, Kominami, and Takemori

1992; Nnamani et al. 2016; Perutz 1989) binds to an enzyme at a site other than its catalytic site affecting its activity.

Deleterious pleiotropic expression of a widely expressed pathway could be subverted with gene duplication followed by tissue-specific regulation and perhaps homology specialization. However, understanding how the mechanistic constraints of a pathway and its tissue-specific regulation changes following duplication requires an understanding of ancestral enzyme function, how these functions changed following duplication, and the effects of substitutions in those enzymes. Although there is much theory (Jensen 1976; Morowitz and Others 1999; Harborne 1990; Horowitz 1965) and comparative biology (Koes, Verweij, and Quattrocchio 2005; Chapman and Ragan 1980; Jensen 1985) on how biochemical pathways are built and change, to date, there is little empirical evidence on biochemical pathways change on evolutionary timescales.

The corticosteroid synthesis pathway in vertebrates provides a model to understand how a biochemical pathway resolves a duplication of one of its components that result in tissue-specific homologs with divergent functions. The corticosteroids are tetrapod hormones that regulate metabolism and stress. The corticosteroid biosynthetic pathway (Fig 1A) is a well-studied example of a pathway that uses a variety of regulatory mechanisms to produce tissue-specific expression of its products across different tissues (Vinson 2016). The enzyme in the adrenal cortex that synthesizes corticosteroids is Cytochrome P450 11B (CYP11B) (Fig 1B). CYP11B is a multi-step enzyme in that it is a single enzyme that performs distinct and consecutive reactions of a biochemical pathway. CYP11B first modifies the 11-carbon atom of the steroid precursors 11-deoxycorticosterone (DOC) and 11-deoxycortisol and subsequently oxidizes the 18-carbon of corticosterone to produce aldosterone (Kominami, Harada, and Takemori 1994).

In all tetrapods studies thus far, 18-oxidized steroids and 11-hydroxylated steroids are produced in different tissues of the adrenal cortex. Aldosterone is produced in the zona glomerulosa and the 11 β -corticosteroids (cortisol and corticosterone) are produced in the zona reticularis and zona fasciculata (Fig 1B). In addition to producing different steroid hormones, the zona have different cellular and topological morphologies (Vinson 2016). This zona specific synthesis (zonation) of corticosteroids persists in species with one CYP11B enzyme (Sanders et al. 2016; Boon et al. 1997; Schiffer, Anderko, and Hannemann 2014; Nonaka et al. 1995) or those that have experienced duplications and have two CYP11B homologs (Nishimoto et al. 2010; Domalik et al. 1991; Ogishima et al., n.d.; Malee and Mellon 1991; Miller and Auchus 2011; Payne and Hales, n.d.; Lisurek and Bernhardt 2004; Bassett, White, and Rainey 2004).

Previous work on the CYP11B homologs in primates suggests that following a duplication, the CYP11B paralogs have specialized on either 18-oxidation or 11-hydroxylation (Zöllner et al. 2008; Hobler et al. 2012; Strushkevich et al. 2013), but none have taken a historical approach (Harms and Thornton 2013) to understand how function changed in this pathway and the mechanisms that have shaped that change.

The mechanisms that allow the production of different metabolites in the different tissues of the adrenal cortex in single-CYP11B species are unclear but likely causes include tissue-specific expression of other pathway members (Nishimoto et al. 2010), allosteric regulation of CYP11B by other enzymes (Ikushiro, Kominami, and Takemori 1992), or the tissue-specific expression of CYP17A (Figure 1). In species with two CYP11Bs (some primates and, independently, some rodents (Baker, Nelson, and Studer 2015)), zonation is controlled by a combination of tissue-specific expression of the paralogs (Nishimoto et al. 2010; Domalik et al. 1991; Ogishima et al., n.d.; Malee and Mellon 1991) and enzyme substrate specialization. In humans, for example, CYP11B1 is expressed in the zona reticularis and zona fasciculata and synthesizes the 11 β -oxidized cortisol and corticosterone (Zöllner et al. 2008). CYP11B2 produces the 18-oxidized corticosteroid aldosterone (Hobler et al. 2012) and is expressed only in the zona glomerulosa.

In all accounts, it appears that the physiological function and, in particular, the zonation of the corticosteroids in the adrenal cortex has been maintained in mammals with one or two CYP11B enzymes. The misregulation of these enzymes is catastrophic. Altered cortisol levels in humans leads to metabolic syndrome, mood disorders, and bone loss (Marques, Silverman, and Sternberg 2009) and the misregulation of aldosterone leads to hyper- or hypo-tension, hypokalemia and androgen excess (Connell, Fraser, and Davies 2001; Peter, Dubuis, and Sippell 1999). Therefore, following the duplication of CYP11B, it is critical that the levels of corticosteroids produced are properly balanced as to avoid any deleterious pleiotropic effects.

In this study, I dissect the changes to the intrinsic biochemical functions of CYP11B paralogs following the duplication in the corticosteroid pathway in simian primates (infraorder Simiiformes) from their common ancestor. I focus our study on the *in vitro* functional changes in the primate lineage. To do so I characterized the specific activity prior to and after its duplication in the primate lineage, with emphasis on CYP11B2 enzyme (sometimes called aldosterone synthase), including a detailed dissection of the trajectory of evolutionary change in the CYP11B2 lineage. Following duplication of the CYP11B multifunctional enzyme, both 11- and 18-oxidation were maintained in tissues where 18-oxidation is required. These findings have

108 implications for the elaboration of biochemical pathways and on constraints on specialization
109 when the enzymes are multi-functional.

Methods

Phylogenetics and ancestral state reconstruction

We sampled 86 sequences in tetrapods, cartilaginous and bony fishes from NCBI, Uniprot, and Ensembl. I aligned the sequences first with MUSCLE (Edgar 2004), and manually checked for alignment errors and removed species-specific insertions. With PhyML (Guindon and Gascuel 2003, Guindon et al. 2010) and SPR+NNI rearrangements to infer the maximum likelihood (ML) phylogeny of our vertebrate CYP11B/C sequences under the JTT+G+I model, which was the best-fit model as evaluated by PROTTEST (Darriba et al. 2011) using AIC. Insertions and deletions were resolved with maximum parsimony (Felsenstein 2003).

There were minor incongruences between the ML phylogeny and the known mammalian species phylogeny (Sen Song et al. 2012) that led to two hypotheses about the evolution of CYP11B. A more parsimonious explanation of gene gains and losses places the duplication of CYP11B in the Old World primates (Figure S2), and the placement of the New World primate orthologs within the Old World primate CYP11B1 clade in the ML topology is incorrect. I reconstructed the ancestral states with (Yang 2007) implemented in Lazarus (Hanson-Smith, Kolaczowski, and Thornton 2010) for both topologies. The reconstructed sequences for the ancestor of all Old World primate CYP11Bs (which is also ancestral to New World primate CYP11Bs in the ML topology), the ancestor of primate CYP11Bs, and the ancestor of primate CYP11B2s are identical between topologies.

To test whether our functional inferences are robust to statistical uncertainty in the reconstruction, I made an alternative ancestor for the ancestral Old World primate CYP11B that included one site (A438T) in the reconstruction that had an alternate state with a posterior probability greater than 0.2.

Plasmid construction

Unlike (Hobler et al. 2012; Zöllner et al. 2008) I did not get high expression of the human sequence of CYP11B1 or CYP11B2 in *Escherichia coli* under published expression conditions. Instead, I ordered sequences that were codon optimized for expression in *E. coli* as IDT gblocks (Integrated DNA Technologies). The *E. coli* codon optimized human cDNA and the inferred ancestral CYP11B cDNAs were expressed in the plasmid Novagene pET17b plasmid cloned with the restriction sites NdeI/HindIII as in (Hobler et al. 2012) with slight modifications. Briefly, the mature form of the protein was expressed without the N-termini mitochondrial targeting sequence. The new N-terminus was modified from GTRAAR to MATKAAR (Hobler et al. 2012) and a hexa-histidine tag added on the C-terminus to assist with purification (Nonaka et al. 1992). Individual point mutations were introduced via site-directed mutagenesis.

The mature human ferredoxin 1 (FdX) was expressed in pET17b was amplified as in (Uhlmann et al. 1992) as a cytosolic protein. The mature peptide (sites 61-184 beginning as MSSSEDKI) was amplified with the restriction sites NheI and HindIII and cloned into the pET17b vector. The mature human FdxR (beginning as MSTQEK (Sagara et al. 1993)) was amplified with restriction sites NdeI and KpnI and cloned into the vector pET17b. The chaperones GroEL and GroES (Nishihara et al. 1998) were obtained in the pGRO12 plasmid (from Dr. Michael Waterman).

Protein expression CYP11B

pET17b-CYP11B (AmpR) plasmids were co-transformed with the pGRO12 plasmid (KanR) into competent BL21(DE3) *E. coli* cells (Zöllner et al. 2008). All P450-expression cell lines were grown in terrific broth (TB) supplemented with 50 ug/ml ampicillin and 50 ug/ml kanamycin. For expression, a scraping of frozen glycerol stock was inoculated into 5 ml TB with the same amount of antibiotics under shaking (225 rpm) overnight at 37 C. I had very poor expression when overnight cultures were started from anything other than a glycerol stock or if the antibiotic was carbenicillin instead of ampicillin. 4 ml of the overnight culture was added to 400 ml TB (ampicillin, kanamycin) in a 2.8 L baffled flask and shaken at 210 rpm at 37 C. When the optical density at 600 nm (OD₆₀₀) reached 0.7, I added 1 mM δ -aminolevulinic acid, 0.5 mM IPTG, and 4 mg/ml arabinose and changed the incubation parameters to 27.5 C and 76 rpm. Four hours later, 10 ul of 25 mg/ml chloramphenicol was added and shaking was changed to 125 rpm. Shaking continued for 16 hours before harvest.

Cells were harvested under 4 C, 4000 x g for 30 minutes. Pellets were frozen at -80 C and resuspend in CYP11B buffer A (50 mM potassium phosphate (pH 7.4), 0.1 mM EDTA, 20% glycerol, 1.5% sodium cholate, 1.5% Tween 20, 0.1 mM PMSF, 0.1 mM DTT) with sonication. The sonication conditions were: setting 8 on the Q55 Sonicator (QSonica), 15 second on, 15 seconds off four times. After resting on ice for one minute, the sonication protocol was repeated three times.

After sonication the cell debris was removed by centrifugation at 19,000 x g for 45 minutes. The supernatant was passed through a 0.45 micron low-protein binding filter, applied to a 5 ml Ni-NTA agarose column bed, and equilibrated with CYP11B Buffer A with 40 mM imidazole. Proteins were eluted with CYP11B buffer B (200 mM imidazole acetate, pH 7.4, 20% glycerol, 0.1 mM EDTA, 0.1 mM dithiothreitol, 1% sodium cholate, 1% Tween 20). Following elution, fractions containing the P450 (red fractions) were combined and concentrated in concentrating tubes with a 30,000 kDa membrane at 4 C. After one wash with CYP11B buffer A, the proteins were concentrated in a Slide-A-Lyzer dialysis cassette (Thermo Scientific) with the

concentrating solution of 40% PEG 20 (Sigma-Aldrich Product No. 81300) at 4 C for up to 2 hours. Aliquots of concentrated enzyme were flash frozen in liquid nitrogen and stored at -80 C. Frozen samples maintained activity for at least six months.

The concentration of active enzyme was calculated with the ferrous carbon monoxide (CO) versus ferrous difference spectrum (Guengerich et al. 2009). I added 10-20 μ l of concentrated P450 to CO buffer (100 mM potassium phosphate buffer (pH 7.4), 1 mM EDTA, 20% glycerol, 0.5% sodium cholate, and 0.4% Triton N-1010) to a total volume of 1 ml. After blanking with the P450 and buffer solution, CO was bubbled into the cuvette at 2 bubbles per second for 45 seconds. A few milligrams of sodium dithionite powder is added to reduce the CO-P450 complex and the cuvette was covered with parafilm and inverted ten times. In a cuvette in the Nanodrop 2000c I scanned the spectra from 400 to 500 nm every thirty seconds for up to 15 minutes, when the spectra reached equilibrium, or when the difference between absorbance at 450 nm and 490 nm was the greatest (Omura and Sato 1964). I calculated the total amount of active enzyme by taking the difference of the absorbance at 450 nm from 490 nm with an extinction coefficient of P450 $0.091 \text{ mM}^{-1}\text{cm}^{-1}$ $[(\Delta A_{450} - \Delta A_{490})/0.091 = \eta \text{mol of P450 per ml}]$.

Protein expression Reductases

The BL21(DE3) cells with FdX were grown overnight (100 μ g/ml AMP in LB) and 5 ml of overnight culture was added to 500 ml LB with no antibiotic at 37 C 210 rpm. When the OD₆₀₀ was 0.7, the cells were induced with 0.5 mM IPTG and grown at 30 C, 125 rpm for 24 hours (Modified from (Uhlmann et al. 1992)).

FdX containing cells were harvested via centrifugation 4000 x g for 25 minutes at 4 C. The pellet was resuspended in 30 ml Buffer P (10 mM KPO₄ pH 7.4, 1 mM EDTA, 1 mM PMSF, 30 μ l lysozyme and DNaseI) from (Uhlmann et al. 1992) and homogenized with sonication as outlined above for the P450 enzymes. The large cell debris was removed by centrifugation at 10,000 x g for 30 minutes at 4 C. AdX was enriched by adding ammonium sulfate to 60% at 4 C followed by centrifugation at 45,995 g for 1 hour. The supernatant was dialyzed overnight with Buffer P. The supernatant was passed through a 0.45 μ m filter, added to Source Q anion exchange column (GE Healthcare Life Sciences) and eluted with a 0-400 mM NaCl gradient in Buffer P. The brown fractions were concentrated and stored in 20% glycerol Buffer P. and quantified in a nanodrop cuvette with the extinction coefficient of $9.8 \text{ mM}^{-1}\text{cm}^{-1}$ at 415 nm.

The human FdxR was also expressed in pET17b and was co-expressed with GroEL/ES in BL21(DE3). Cells were grown overnight (37 C, 225 rpm) and inoculated into 500 ml TB with 100 μ g/ml ampicillin and 50 μ g/ml kanamycin. At OD₆₀₀ of 0.5 FdxR was induced with 0.5 mM

IPTG and 4 mg/ml arabinose. After induction, the cells were grown at 26 C, 125 rpm shaking for 72 hours (modified from (Sagara et al. 1993)).

Cells were harvested via centrifugation 4000 x g for 25 minutes at 4C and resuspend in 30 ml Buffer P (10 mM KPO₄ pH 7.4, 1 mM EDTA, 1 mM PMSF, 30 ul lysozyme and DNaseI). The cells were homogenized via sonication as described above. The cell debris was removed by centrifugation at 45,995 x g for 1 hour at 4 C. The supernatant was dialyzed overnight with Buffer P, and applied to a Source Q anion exchange column after filtration through a 0.45 micron filter. The flow through was collected and applied to a 2'5'-ADP-Sepharose column. The column was washed with Buffer P and the protein eluted with a 0-400 mM NaCl gradient in Buffer P. The flow through on the 2'5'-ADP Sepharose column was reapplied to obtain more FdxR. Yellow fractions were collected, concentrated and stored in 20% glycerol Buffer P. FdxR was quantified using the extinction coefficient 10.9 mM⁻¹cm⁻¹ at 450 nm in Buffer P.

Cytochrome P450 kinetic assays

To quantify the specific activity (μmol min⁻¹ mg⁻¹) of the CYP11B homologs, I combined 0.4 μM enzyme with 50 uM HEPES, 1 μM MgCl₂, 0.5 mM FdxR, 10 mM FdX, 5 μM glucose-6-phosphate, 0.1 units glucose-6-phosphate dehydrogenase (Sigma), and 350 μM substrate in 100 μL reaction volume. The reactions were placed at 37 C for 3 minutes and the reaction was started by adding NADPH to 2 mM. The reaction was incubated at 37 C with slight agitation for 15 minutes and stopped with 150 μL chloroform. After the reaction was stopped, I added cortisol to 100 μM as an internal standard. The steroids were extracted three times with 100 μL chloroform and dried in a hood. After the chloroform dried, I added 20% HPLC-grade acetonitrile to resuspend the steroids (modified from (Hobler et al. 2012)).

All steroids (products and substrates) were monitored with LC-MS methods. The samples were run on an Agilent 1200 Series HPLC coupled to an Agilent 6410 triple quadrupole mass spectrometer (Morton, Roux, and Bergelson 2015). 1 μl of sample was injected onto a 4.6 x 50-mm Agilent ZORBAX Eclipse Rapid Resolution HT XBD-C18 column with a 1.8 μM particle size. All samples were run with the following program: 0.3 mL/min flow rate, ramp from 60:40 0.1% formic acid in HPLC-grade water:acetonitrile to 30:70 ratio over ten minutes. The mass spectrometer was run in precursor positive-ion electrospray mode, monitoring all parent ions (Table S2). The fragmentor voltage was optimized with steroid standards for aldosterone, cortisol, corticosterone, and 11-deoxycorticosterone (DOC) obtained from Sigma-Aldrich and Steraloids. Maximum detection was achieved with collision energy (Table S2). Individual steroids were identified based on comparison to purchased standards, fragmentation pattern,

245 and retention time as compared with a cortisol internal standard. The final data were generated
246 by dividing the parent ion counts by the internal cortisol standard.

Results

CYP11B duplicated in the primates

We studied a duplication followed by functional divergence in the corticosteroid pathway in primates. As previously reported (Baker, Nelson, and Studer 2015), I found at least three independent duplications of CYP11B in mammals (Figure S1). Here I focus on the duplication and subsequent functional shift that occurred in the primate lineage (Figure 2). The phylogenetic evidence is nearly equivocal on whether the gene duplication occurred before or after the split of Old World primates (parvorder Catarrhini) and New World monkeys (parvorder Platyrrhini), with the ML tree weakly supporting the duplication occurring before this split, and a more parsimonious hypothesis of gene gain and loss favoring the duplication occurring after this split.

Of the New World monkeys for which I found sequences, the black-capped squirrel monkey (*Saimiri boliviensis boliviensis*) had two copies of CYP11B. However, this apparent second duplication is species-specific and occurred after the divergence of New and Old World primates. The closest relatives of New and Old World primates (Sen Song et al. 2012), including the Chinese tree shrew (*Tupaia chinensis*) and tarsier (*Tarsius syrichta*) have a single copy of CYP11B. I reconstructed the duplication node and the nodes that led to the two CYP11B paralogs for both phylogenies.

Human CYP11B1 and CYP11B2 functional divergence

The CYP11B paralogs in humans have functionally diverged following duplication. CYP11B increased its ability to produce either cortisol and corticosterone (11 β -hydroxylation products) and most of its ability to produce aldosterone (Figure 3) (consistent with previous studies (Zöllner et al. 2008; Hobler et al. 2012; Strushkevich et al. 2013)). These results agree with these previous studies except that I detected a small amount of 18-hydroxycorticosterone and aldosterone produced by CYP11B1 *in vitro* (Figure 3).

The human CYP11B2 paralog has increased ability to produce both 11 β -hydroxylation products and 18-oxidation products (aldosterone) (Figure 3). 11 β -hydroxylation function is always higher than 18-oxidation function and both paralogs retain the ability to complete both 11 β - and 18-oxidation.

Function of ancestral CYP11B, CYP11B1, and CYP11B2

The pre-duplication ancestor of CYP11B1 and CYP11B2, AncCYP11B, had moderate 18-oxidation activity and relatively low 11 β -hydroxylation activity *in vitro* (Fig 3). This result is robust to statistical ambiguity in the reconstruction: an alternate reconstruction incorporating all second-best states with PP > 0.2 (AncCYP11B-AltALL) has indistinguishable function from the most probable ancestor. Both of the AncCYP11B1 and AncCYP11B2 reconstruction have an

average posterior probability greater than 0.99 and have no sites with an alternate reconstruction with a posterior probability higher than 0.2.

AncCYP11B1 and AncCYP11B2 enzymes have differentiated from the AncCYP11B toward 11 β - or 11-oxidation and 18-oxidation activity, respectively (Fig 3). With respect to AncCYP11B, AncCYP11B1 had a moderate increase in 11 β -hydroxylation function and a decrease in 18-oxidation function. Between AncCYP11B1 and human CYP11B1, there was a 4-fold increase in 11 β -hydroxylation activity. AncCYP11B2 had a two-and-a-half and four-fold increase in 18-oxidation and 11 β -hydroxylation, respectively. From AncCYP11B2 to human CYP11B2, there was only a slight increase in 18-oxidation and a slight decrease in 11 β -hydroxylation. The AncCYP11B2 and the extant human CYP11B2 both maintain 11-hydroxylation and 18-oxidation activity.

However, it is important to note that CYP11B2 did not specialize in the traditional sense in that it did not lose one function in favor of another. It appears that it is necessary that in order to increase 18-oxidation activity, there must also be an increase in 11 β -hydroxylation activity. Therefore, I dissected the substitutions from the preduplicated AncCYP11B to AncCYP11B2 to understand the interplay of balancing 18-oxidation and 11 β -hydroxylation activities.

Substitutions leading to AncCYP11B2

To understand the trajectory of functional evolution in the CYP11B2 lineage, I characterized the function of all intermediates from AncCYP11B to AncCYP11B2. There are four substitutions along this lineage. I dissected the effect of those substitutions by characterizing the 11 β -hydroxylase (Fig 4A) and 18-oxidation (Fig 4B) functions of all 16 combinations. All four substitutions appear to affect either 11 β -hydroxylation or 18-oxidation activity (Fig 4). Here I depict ancestral character states as lowercase and derived as uppercase letters with the amino acid position between; the four substitutions are d83N, d302E, q472L, and s492G.

When I dissect the intrinsic function of all possible intermediates, Two of the substitutions (d83N and s492G) always increase 11 β -hydroxylation function in any genetic background tested (Table 1, Fig 5A). Consistent with this observation, on the trajectory between AncCYP11B and AncCYP11B2, only genotypes with both 82N and 492G (NEqG, NdLG, NdqG, and NELG) have the highest 11 β -hydroxylation activity. The other two substitutions (d302E and q472L) have mixed effects possibly caused by non-significant background-dependent effects (Table 2, Fig 5B). None of the substitution combinations lead to complete loss of either function, and only one genotype (NEqG) has 18-oxidation activity that is lower than the ancestral

314 AncCYP11B genotype (ddqs). Therefore, most mutational trajectories are plausible between
315 AncCYP11B and AncCYP11B2, except maybe through NEqG.

316 Structural analysis does not provide much insight but is reported nonetheless. I mapped
317 the substitutions on the human CYP11B2 crystal structure (PDB: 4dvq (Strushkevich et al.
318 2013)) and found that none of the substitutions are in the canonical substrate recognition
319 regions (Gotoh 1992) or even within 12 Angstroms of the heme or substrate. Any structural
320 modifications must have subtle effects on the substrate recognition or binding regions.

Discussion

Here I studied the evolution of the corticosteroid synthesis pathway to determine how functional evolution is resolved for a multi-step enzyme after duplication. The corticosteroids are a crucial pathway, the products of which control metabolism, salt balance, and immunity among other processes. To understand the processes that shape the augmentation of this pathway (two enzymes that do the job of one), I used ancestral state reconstruction to resurrect enzymes bracketing a gene duplication in the corticosteroid synthesis pathway in primates. Prior to its duplication, the AncCYP11B enzyme had moderate activity on both the 11 β - or 18-carbon of the precursor DOC. Following duplication, while the CYP11B1 paralog specialized on 11 β -carbon activity, CYP11B2 maintained multiple functions, similar to the that of AncCYP11B, but increased.

As biochemical pathways evolve, both mechanistic enzymatic and tissue-specific constraints shape their evolution. Previous biochemical studies have established that CYP11B 11 β -hydroxylation activity occurs first followed by 18-oxidation (Kominami, Harada, and Takemori 1994). However, if the 11 β -hydroxylated substrate (i.e. corticosterone) leaves the active site, it cannot reenter to be 18-oxidized to aldosterone. This implies that in order to have tissue-specific expression of CYP11B duplicates and still maintain zonation, CYP11B2 must maintain both functions rather than specializing on 18-oxidation activity. Our results that consider the ancestral CYP11B activity help explain why CYP11B2 is constrained to maintain 11-hydroxylation activity in addition to 18-oxidation even when the conserved function of the zG is to produce 18-hydroxylated products.

The resolution of multi-functional (more than one defined function on preferred substrates) or promiscuous (trace activity for a non-preferred substrate) enzymes have been well studied (Huang et al. 2012; Voordeckers et al. 2012; Thomson et al. 2005; Des Marais and Rausher 2008; Deng et al. 2010; Boucher et al. 2014) and the prevailing model of functional evolution following a duplication in most of these studies is the escape from adaptive conflict model (with (Boucher et al. 2014) finding a rare case of neofunctionalization). However, under the umbrella of multi-functional enzymes are multi-step enzymes. Multi-step enzymes, like CYP11B, provide an understudied avenue in which biochemical pathways change their configuration without gaining a novel synthetic step. CYP11B1 underwent classic duplication, degeneration, and complementation (DDC) model where one function was maintained (Force et al. 1999). However, none of the classic models fit for CYP11B2 as both functions were maintained, if not slightly improved following duplication. When the complexity of biochemical mechanisms of metabolism are considered with the further complication of tissue-specific

regulation, further models may be required, including one to include the maintenance of function in enzymes that are themselves short biochemical pathways like CYP11B. It is yet unknown how common or important such multi-step enzymes are in metabolism.

We tested the intrinsic activities of these enzymes *in vitro*, but predict that some of their *in vivo* activities may differ depending on the other enzymes with which they are expressed in the tissues of the adrenal cortex. For example, the activity of CYP11B in bovine (which has a single CYP11B copy - similar to state in the ancestral primate) is allosterically regulated by CYP11A *in vitro* (Ikushiro, Kominami, and Takemori 1992). This may partially explain how the adrenal cortex in bovine still expresses zonation of cortisol and aldosterone production with only a single copy of CYP11B. When co-expressed with CYP11A in the zona fasciculata and zona reticularis, bovine CYP11B preferentially makes 11 β -hydroxylated products. When bovine CYP11B is expressed alone in the zona glomerulosa of the adrenal cortex, it preferentially makes 18-hydroxylated products. However, this does not change the primary result that CYP11B2 must maintain both functions to preserve the conserved physiological configuration of the corticosteroid products.

Nonetheless, the functional results presented here should be viewed in the light that allosteric regulation by CYP11A would modify both 11- and 18- oxidation activities in the ancestral enzymes. There is some evidence that the human CYP11B paralogs are not allosterically regulated by CYP11A (Cao and Bernhardt 1999) suggesting the allosteric regulation by CYP11A may have been lost after the duplication of CYP11B in the primate lineage. It is therefore probable that their contemporaneous AncCYP11As regulated primate AncCYP11B and possibly AncCYP11B1. Specifically, I expect the AncCYP11B to have higher 11 β -hydroxylase and lower 18-oxidation activity when co-expressed with its ancestral CYP11A. It may also be that AncCYP11B1 had higher 11 β -hydroxylase activity if it had not lost allosteric regulation by CYP11A by the first speciation node after the duplication. However, the functional dissection of AncCYP11B2 is not subject to this problem because its function in the zona glomerulosa was never altered by CYP11A as they are not co-expressed. Thus, the dissection of AncCYP11B2 is not subject to this caveat.

When tracing the effects of substitutions on the CYP11B2 lineage, all four substitutions affect protein function, though there is some non-significant background-dependence in their effects. They combine to ultimately produce a CYP11B2 enzyme with higher 18-oxidation that still retains 11 β -hydroxylation activity. Both human CYP11B paralogs are capable of oxidation on both the 11 β - and 18-carbons of 11-deoxycorticosterone (DOC). The proportion of 11 β -hydroxylation to 18-oxidation activity by one enzyme may partially depend on how long the

substrate is retained in the active site to undergo subsequent oxidations (Imai, Yamazaki, and Kominami 1998), which suggests activity could be modulated by subtle shifts in the overall protein structure.

The imbalance of mineralocorticoids or glucocorticoids can lead to a variety of diseases in humans (Marques, Silverman, and Sternberg 2009; Peter, Dubuis, and Sippell 1999; Connell, Fraser, and Davies 2001) and during the evolution from AnceCYP11B to AncCYP11B1 or AncCYP11B2, it would be critical to maintain the balance of these hormones. The combinations NEqG and NdLG have higher 11-hydroxylation activity than any other variant, but it is difficult to speculate without knowing the mutational trajectory from AncCYP11B to AncCYP11B1 or without knowing its cellular expression or dependency on CYP11A. Those variants also have much lower 11-hydroxylase activity than the extant humanCYP11B1. When tracing the path of change from Anc11B to Anc11B2, nearly all mutational trajectories are open in the sense that there is not a complete loss of either function. The exception is NEqG which has much lower 18-oxidation activity than the AncCYP11B. In humans, a deficiency in aldosterone leads to hypoaldosteronism, hyponatremia, hyperkalaemia, fatal salt loss, and failure to thrive (Peter, Dubuis, and Sippell 1999). Although aldosterone likely plays a similar role in extant primates, again, it is unknowable if the NEqG variant would have been deleterious without additional pathway context. The reconstruction of the contemporaneous CYP11A with the ancestral CYP11B proteins is needed to further understand how allosteric regulation interacts with the resolution of the duplication of CYP11B in primates.

Our results suggest a complex evolutionary history of gene duplication and functional evolution with the interplay of enzyme activity within the enzymes studied. We still do not understand the relative importance and interplay of these factors during evolution of this gene family or how common these factors are generally during the evolution of biochemical pathways. However, this study suggests that biochemical networks can be shaped by mechanistic constraints of the enzyme and tissue-specific regulation with selection to avoid deleterious pleiotropy.

416 **Acknowledgement and funding sources**

417 I want to thank J. W. Thornton and all members of the Thornton lab for their comments on this
 418 manuscript and support that enhanced this work. I especially want to thank L. Picton and A.
 419 Venkat for their support. Thanks to Michael Waterman for crucial plasmids. I want to thank A.
 420 Manning, K. Ferris, and C. Smukowski-Heil for their helpful comments on this manuscript. This
 421 work was supported by the National Institutes of Health (NRSA 1 F32 GM112350-01, C.F. O-
 422 M.) and initial funding from the National Institutes of Health (NIH) grant (R01-GM104397;
 423 J.W.T.)

- Baker, Michael E., David R. Nelson, and Romain A. Studer. 2015. "Origin of the Response to Adrenal and Sex Steroids: Roles of Promiscuity and Co-Evolution of Enzymes and Steroid Receptors." *The Journal of Steroid Biochemistry and Molecular Biology* 151 (July): 12–24.
- Bassett, Mary H., Perrin C. White, and William E. Rainey. 2004. "The Regulation of Aldosterone Synthase Expression." *Molecular and Cellular Endocrinology* 217 (1-2): 67–74.
- Boon, W. C., P. J. Roche, A. Butkus, J. G. McDougall, K. Jeyaseelan, and J. P. Coghlan. 1997. "Functional and Expression Analysis of Ovine Steroid 11 β -Hydroxylase (Cytochrome P45011 β)." *Endocrine Research* 23 (4): 325–47.
- Boucher, Jeffrey I., Joseph R. Jacobowitz, Brian C. Beckett, Scott Classen, and Douglas L. Theobald. 2014. "An Atomic-Resolution View of Neofunctionalization in the Evolution of Apicomplexan Lactate Dehydrogenases." *eLife* 3 (June).
<http://elifesciences.org/lookup/doi/10.7554/eLife.02304.036>.
- Cao, Pei-Rang, and Rita Bernhardt. 1999. "Interaction of CYP11B1 (cytochrome P-45011beta) with CYP11A1 (cytochrome P-450scc) in COS-1 Cells." *European Journal of Biochemistry / FEBS* 262 (3): 720–26.
- Carroll, Sean B. 2005. "Evolution at Two Levels: On Genes and Form." *PLoS Biology* 3 (7): e245.
- Chapman, D. J., and M. A. Ragan. 1980. "Evolution of Biochemical Pathways: Evidence From Comparative Biochemistry." *Annual Review of Plant Physiology* 31 (1): 639–78.
- Coberly, L. Caitlin, and Mark D. Rausher. 2008. "Pleiotropic Effects of an Allele Producing White Flowers in *Ipomoea Purpurea*." *Evolution; International Journal of Organic Evolution* 62 (5): 1076–85.
- Connell, J. M., R. Fraser, and E. Davies. 2001. "Disorders of Mineralocorticoid Synthesis." *Best Practice & Research. Clinical Endocrinology & Metabolism* 15 (1): 43–60.
- Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2011. "ProtTest 3: Fast Selection of Best-Fit Models of Protein Evolution." *Bioinformatics* 27 (8): 1164–65.
- Deng, Cheng, C-H Christina Cheng, Hua Ye, Ximiao He, and Liangbiao Chen. 2010. "Evolution of an Antifreeze Protein by Neofunctionalization under Escape from Adaptive Conflict." *Proceedings of the National Academy of Sciences of the United States of America* 107 (50): 21593–98.
- Des Marais, David L., and Mark D. Rausher. 2008. "Escape from Adaptive Conflict after Duplication in an Anthocyanin Pathway Gene." *Nature* 454 (7205): 762–65.
- Domalik, Leslie J., David D. Chaplin, M. Sue Kirkman, Richard C. Wu, Welwin Liu, Thad A. Howard, Michael F. Seldin, and Keith L. Parker. 1991. "Different Isozymes of Mouse 11

Beta-Hydroxylase Produce Mineralocorticoids and Glucocorticoids." *Molecular Endocrinology* 5 (12): 1853–61.

Edgar, Robert C. 2004. "MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput." *Nucleic Acids Research* 32 (5): 1792–97.

Felsenstein, J. 2003. "Inferring Phylogenies," January.

Force, A., M. Lynch, F. B. Pickett, A. Amores, Y. L. Yan, and J. Postlethwait. 1999. "Preservation of Duplicate Genes by Complementary, Degenerative Mutations." *Genetics* 151 (4): 1531–45.

Gotoh, O. 1992. "Substrate Recognition Sites in Cytochrome P450 Family 2 (CYP2) Proteins Inferred from Comparative Analyses of Amino Acid and Coding Nucleotide Sequences." *The Journal of Biological Chemistry* 267 (1): 83–90.

Guengerich, F. Peter, Martha V. Martin, Christal D. Sohl, and Qian Cheng. 2009. "Measurement of Cytochrome P450 and NADPH–cytochrome P450 Reductase." *Nature Protocols* 4 (9): 1245–51.

Gui, Dan Y., Caroline A. Lewis, and Matthew G. Vander Heiden. 2013. "Allosteric Regulation of PKM2 Allows Cellular Adaptation to Different Physiological States." *Science Signaling* 6 (263): e7.

Guindon, Stéphane, Jean-François Dufayard, Vincent Lefort, Maria Anisimova, Wim Hordijk, and Olivier Gascuel. 2010. "New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0." *Systematic Biology* 59 (3): 307–21.

Guindon, Stéphane, and Olivier Gascuel. 2003. "A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood." *Systematic Biology* 52 (5): 696–704.

Hanson-Smith, V., B. Kolaczkowski, and J. W. Thornton. 2010. "Robustness of Ancestral Sequence Reconstruction to Phylogenetic Uncertainty." *Molecular Biology and Evolution* 27 (9): 1988–99.

Harborne, Jeffrey B. 1990. "Constraints on the Evolution of Biochemical Pathways." *Biological Journal of the Linnean Society. Linnean Society of London* 39 (2): 135–51.

Harms, Michael J., and Joseph W. Thornton. 2013. "Evolutionary Biochemistry: Revealing the Historical and Physical Causes of Protein Properties." *Nature Reviews. Genetics* 14 (8): 559–71.

Hobler, Anna, Norio Kagawa, Michael C. Hutter, Michaela F. Hartmann, Stefan A. Wudy, Frank Hannemann, and Rita Bernhardt. 2012. "Human Aldosterone Synthase: Recombinant Expression in E. Coli and Purification Enables a Detailed Biochemical Analysis of the

- Protein on the Molecular Level.” *The Journal of Steroid Biochemistry and Molecular Biology* 132 (1-2): 57–65.
- Horowitz, N. H. 1965. “The Evolution of Biochemical Syntheses—retrospect and Prospect.” *Evolving Genes and Proteins*, 15–23.
- Huang, Ruiqi, Frank Hippauf, Diana Rohrbeck, Maria Haustein, Katrin Wenke, Janie Feike, Noah Sorrelle, Birgit Piechulla, and Todd J. Barkman. 2012. “Enzyme Functional Evolution through Improved Catalysis of Ancestrally Nonpreferred Substrates.” *Proceedings of the National Academy of Sciences* 109 (8): 2966–71.
- Ikegami, Keisuke, Xiao-Hui Liao, Yuta Hoshino, Hiroko Ono, Wataru Ota, Yuka Ito, Taeko Nishiwaki-Ohkawa, et al. 2014. “Tissue-Specific Posttranslational Modification Allows Functional Targeting of Thyrotropin.” *Cell Reports* 9 (3): 801–10.
- Ikushiro, S., S. Kominami, and S. Takemori. 1992. “Adrenal P-450_{scc} Modulates Activity of P-450_{11β} in Liposomal and Mitochondrial Membranes. Implication of P-450_{scc} in Zone Specificity of Aldosterone Biosynthesis in Bovine Adrenal.” *The Journal of Biological Chemistry* 267 (3): 1464–69.
- Imai, Tadashi, Takeshi Yamazaki, and Shiro Kominami. 1998. “Kinetic Studies on Bovine Cytochrome P450 11β Catalyzing Successive Reactions from Deoxycorticosterone to Aldosterone †.” *Biochemistry* 37 (22): 8097–8104.
- Jensen, R. A. 1976. “Enzyme Recruitment in Evolution of New Function.” *Annual Review of Microbiology* 30 (January): 409–25.
- . 1985. “Biochemical Pathways in Prokaryotes Can Be Traced Backward through Evolutionary Time.” *Molecular Biology and Evolution* 2 (2): 92–108.
- Koes, Ronald, Walter Verweij, and Francesca Quattrocchio. 2005. “Flavonoids: A Colorful Model for the Regulation and Evolution of Biochemical Pathways.” *Trends in Plant Science* 10 (5): 236–42.
- Kominami, Shiro, Daisuke Harada, and Shigeki Takemori. 1994. “Regulation Mechanism of the Catalytic Activity of Bovine Adrenal Cytochrome P-450_{11β}.” *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1192 (2): 234–40.
- Lisurek, M., and R. Bernhardt. 2004. “Modulation of Aldosterone and Cortisol Synthesis on the Molecular Level.” *Molecular and Cellular Endocrinology*, January. <http://www.sciencedirect.com/science/article/pii/S0303720703005094>.
- Malee, M. P., and S. H. Mellon. 1991. “Zone-Specific Regulation of Two Messenger RNAs for P450_{c11} in the Adrenals of Pregnant and Nonpregnant Rats.” *Proceedings of the National Academy of Sciences of the United States of America* 88 (11): 4731–35.

- Marques, Andrea H., Marni N. Silverman, and Esther M. Sternberg. 2009. "Glucocorticoid Dysregulations and Their Clinical Correlates. From Receptors to Therapeutics." *Annals of the New York Academy of Sciences* 1179 (October): 1–18.
- Miller, Walter L., and Richard J. Auchus. 2011. "The Molecular Biology, Biochemistry, and Physiology of Human Steroidogenesis and Its Disorders." *Endocrine Reviews* 32 (1): 81–151.
- Morowitz, Harold J., and Others. 1999. "A Theory of Biochemical Organization, Metabolic Pathways, and Evolution." *Complexity* 4 (6): 39–53.
- Morton, T. C., F. Roux, and J. Bergelson. 2015. "Coselected Genes Determine Adaptive Variation in Herbivore Resistance throughout the Native Range of *Arabidopsis thaliana*." *Proceedings of the*. <http://www.pnas.org/content/112/13/4032.short>.
- Nishihara, Kazuyo, Masaaki Kanemori, Masanari Kitagawa, Hideki Yanagi, and Takashi Yura. 1998. "Chaperone Coexpression Plasmids: Differential and Synergistic Roles of DnaK-DnaJ-GrpE and GroEL-GroES in Assisting Folding of an Allergen of Japanese Cedar Pollen, Cryj2, in *Escherichia coli*." *Applied and Environmental Microbiology* 64 (5): 1694–99.
- Nishimoto, Koshiro, Ken Nakagawa, Dan Li, Takeo Kosaka, Mototsugu Oya, Shuji Mikami, Hirotaka Shibata, et al. 2010. "Adrenocortical Zonation in Humans under Normal and Pathological Conditions." *The Journal of Clinical Endocrinology and Metabolism* 95 (5): 2296–2305.
- Nnamani, Mauris C., Soumya Ganguly, Eric M. Erkenbrack, Vincent J. Lynch, Laura S. Mizoue, Yingchun Tong, Heather L. Darling, Monika Fuxreiter, Jens Meiler, and Günter P. Wagner. 2016. "A Derived Allosteric Switch Underlies the Evolution of Conditional Cooperativity between HOXA11 and FOXO1." *Cell Reports* 15 (10): 2097–2108.
- Nonaka, Y., M. Okamoto, K. Morohashi, S. Kiritani, T. Hashimoto, and T. Omura. 1992. "Functional Expression of cDNAs for Bovine 11 Beta-Hydroxylase-Aldosterone Synthases, P450(11 Beta)-2 and -3 and Their Chimeras." *The Journal of Steroid Biochemistry and Molecular Biology* 41 (3-8): 779–80.
- Nonaka, Y., H. Takemori, S. K. Halder, T. Sun, M. Ohta, O. Hatano, A. Takakusu, and M. Okamoto. 1995. "Frog Cytochrome P-450 (11 Beta,aldo), a Single Enzyme Involved in the Final Steps of Glucocorticoid and Mineralocorticoid Biosynthesis." *European Journal of Biochemistry / FEBS* 229 (1): 249–56.
- Ogishima, T., H. Shibata, H. Shimada, F. Mitani, H. Suzuki, T. Saruta, and Y. Ishimura. n.d. "Aldosterone Synthase Cytochrome P-450 Expressed in the Adrenals of Patients with Primary Aldosteronism." *Jbc.org*. <http://www.jbc.org/content/266/17/10731.short>.

- Omura, T., and R. Sato. 1964. "The Carbon Monoxide-Binding Pigment of Liver Microsomes I. Evidence for Its Hemoprotein Nature." *The Journal of Biological Chemistry*.
<http://www.jbc.org/content/239/7/2370.short>.
- Payne, Anita H., and Dale B. Hales. n.d. "Overview of Steroidogenic Enzymes in the Pathway from Cholesterol to Active Steroid Hormones." *Endocrine Reviews*.
<http://www.ncbi.nlm.nih.gov/pubmed/15583024>.
- Perutz, M. F. 1989. "Mechanisms of Cooperativity and Allosteric Regulation in Proteins." *Quarterly Reviews of Biophysics* 22 (2): 139–237.
- Peter, M., J. M. Dubuis, and W. G. Sippell. 1999. "Disorders of the Aldosterone Synthase and Steroid 11beta-Hydroxylase Deficiencies." *Hormone Research* 51 (5): 211–22.
- Ramsay, Heather, Loren H. Rieseberg, and Kermit Ritland. 2009. "The Correlation of Evolutionary Rate with Pathway Position in Plant Terpenoid Biosynthesis." *Molecular Biology and Evolution* 26 (5): 1045–53.
- Sagara, Yasuhiro, Akira Wada, Yutaka Takata, Michael R. Waterman, Kazuhisa Sekimizu, and Tadao Horiuchi. 1993. "Direct Expression of Adrenodoxin Reductase in Escherichia Coli and the Functional Characterization." *Biological & Pharmaceutical Bulletin* 16 (7): 627–30.
- Sanders, K., J. A. Mol, H. S. Kooistra, A. Slob, and S. Galac. 2016. "New Insights in the Functional Zonation of the Canine Adrenal Cortex." *Journal of Veterinary Internal Medicine / American College of Veterinary Internal Medicine* 30 (3): 741–50.
- Schiffer, L., S. Anderko, and F. Hannemann. 2014. "The CYP11B Subfamily." *The Journal of Steroid ...*, January. <http://www.sciencedirect.com/science/article/pii/S0960076014002465>.
- Sen Song, Liang Liu, Scott V. Edwards, and Shaoyuan Wu. 2012. "Resolving Conflict in Eutherian Mammal Phylogeny Using Phylogenomics and the Multispecies Coalescent Model." *Proceedings of the National Academy of Sciences* 109 (37): 14942–47.
- Snášel, Jan, and Iva Pichová. 2019. "Allosteric Regulation of Pyruvate Kinase from Mycobacterium Tuberculosis by Metabolites." *Biochimica et Biophysica Acta: Proteins and Proteomics* 1867 (2): 125–39.
- Stearns, Frank W. 2010. "One Hundred Years of Pleiotropy: A Retrospective." *Genetics* 186 (3): 767–73.
- Stern, David L., and Virginie Orgogozo. 2008. "The Loci of Evolution: How Predictable Is Genetic Evolution?" *Evolution; International Journal of Organic Evolution* 62 (9): 2155–77.
- Strushkevich, Natallia, Andrei A. Gilep, Limin Shen, Cheryl H. Arrowsmith, Aled M. Edwards, Sergey A. Usanov, and Hee-Won Park. 2013. "Structural Insights into Aldosterone Synthase Substrate Specificity and Targeted Inhibition." *Molecular Endocrinology* 27 (2):

315–24.

Thomson, J. Michael, Eric A. Gaucher, Michelle F. Burgan, Danny W. De Kee, Tang Li, John P. Aris, and Steven A. Benner. 2005. "Resurrecting Ancestral Alcohol Dehydrogenases from Yeast." *Nature Genetics* 37 (6): 630–35.

Uhlmann, H., V. Beckert, D. Schwarz, and R. Bernhardt. 1992. "Expression of Bovine Adrenodoxin in E. Coli and Site-Directed Mutagenesis of /2 Fe-2S/ Cluster Ligands." *Biochemical and Biophysical Research Communications* 188 (3): 1131–38.

Vinson, Gavin P. 2016. "Functional Zonation of the Adult Mammalian Adrenal Cortex." *Frontiers in Neuroscience* 10 (June): 238.

Voordeckers, Karin, Chris A. Brown, Kevin Vanneste, Elisa van der Zande, Arnout Voet, Steven Maere, and Kevin J. Verstrepen. 2012. "Reconstruction of Ancestral Metabolic Enzymes Reveals Molecular Mechanisms Underlying Evolutionary Innovation through Gene Duplication." *PLoS Biology* 10 (12): e1001446.

Wray, Gregory A., Matthew W. Hahn, Ehab Abouheif, James P. Balhoff, Margaret Pizer, Matthew V. Rockman, and Laura A. Romano. 2003. "The Evolution of Transcriptional Regulation in Eukaryotes." *Molecular Biology and Evolution* 20 (9): 1377–1419.

Yang, Z. 2007. "PAML 4: Phylogenetic Analysis by Maximum Likelihood." *Molecular Biology and Evolution* 24 (8): 1586–91.

Zöllner, Andy, Norio Kagawa, Michael R. Waterman, Yasuki Nonaka, Koji Takio, Yoshitsugu Shiro, Frank Hannemann, and Rita Bernhardt. 2008. "Purification and Functional Characterization of Human 11beta Hydroxylase Expressed in Escherichia Coli." *The FEBS Journal* 275 (4): 799–810.

616 Table 1: Average effect of adding derived substitution

Substitution	11 β -hydroxylation (nmol/uM450/min)	18-oxidation (nmol/uM450/min)
d82N	23.62 $\pm$ 9.13	-0.35 $\pm$ 0.29
d302E	-3.65 $\pm$ 9.97	0.38 $\pm$ 0.39
q472L	-0.47 $\pm$ 11.73	0.10 $\pm$ 0.28
s492G	18.07 $\pm$ 6.45	-0.19 $\pm$ 0.36

617

Table 2: Average effect of completing epistatic interaction

Comparison	18-oxidation (nmol/uM450/min)
xdLx to xELx	0.48 ± 0.57
xdxx (not dL) to xExx (not EL)	0.16 ± 0.56
xEqx to xELx	0.29 ± 0.66
xxqx (not Eq) to xxLx (not EL)	-0.12 ± 0.18

Figure 1 Synthesis of corticosteroids in humans (A) is accomplished by the CYP11B homologs. CYP11B1 and CYP11B2 hydroxylate the 11-carbon of 11-deoxycorticosterone (DOC) or 11-deoxycortisol (light grey arrows) to produce corticosterone and cortisol. After performing 11 β -hydroxylation on 11-deoxycorticosterone, CYP11B2 oxidizes the 18-carbon of corticosterone and 18-hydroxycorticosterone (dark grey arrows) to produce aldosterone. In humans, the two paralogs of CYP11B are expressed in different layers of the adrenal cortex (inset). Cortisol and corticosterone (produced by 11-carbon hydroxylation by CYP11B1) are preferentially produced in the *zona fasciculata* and *zona reticularis* of the adrenal cortex. Aldosterone is preferentially produced by CYP11B2 in the *zona glomerulosa* by performing both 11-carbon hydroxylation of 11-deoxycorticosterone and 18-carbon oxidation. Other Cytochrome P450s in the corticosteroid pathway are labeled and new modifications for each enzymatic step are shaded.

Figure 2 Maximum likelihood gene phylogeny of 86 vertebrate CYP11B/C sequences with evolutionary model JTT+I+G (full maximum likelihood gene phylogeny in Figure S1). The clades of bony fish CYP11C, Laurasiatheria (including bats, carnivorans, whales and even-toed ungulates), Lagomorpha (rabbit and pika) and Rodentia are collapsed for clarity. The Old World primate phylogeny is recapitulated twice on the phylogeny with the CYP11B1 clade (light grey box) and the CYP11B2 clade (dark grey small box). The approximate likelihood statistic indicates node support at each node is indicated with asterisk (*>4, **>10, ***>50). The reconstructed ancestors of the CYP11B1, CYP11B2, and the node of those clades which is the common ancestor of all primates in the infraorder Simiiformes.

Figure 3 The primate AncCYP11B (black star) and the alternate ancestral sequence (white star) had moderate 11 β -hydroxylation and 18-oxidation while the other paralogs have specialized. The inset depicts a reduced maximum likelihood phylogeny of CYP11B evolution in primates. The error bars indicate the 95% confidence intervals. The extant human paralogs CYP11B1 (light grey square) and CYP11B2 (dark grey square) have increased 11-hydroxylation and 18-oxidation activities. The intermediate ancestor, AncCYP11B1 (light grey circle), has lost most of its 18-oxidation activity and AncCYP11B2 (dark grey circle) has increased both activities.

Figure 4 Genetic dissection of 11β-hydroxylation (A) and 18-oxidation (B) activity in the CYP11B2 lineage from AncCYP11B to AncCYP11B2. The genotypes are depicted in or near the circles with lowercase indicating ancestral and uppercase as derived. Moving from left to right, single derived substitutions are added. The diameter of dark circles indicates the mean activity in nmol product/uM P450/minute for that genotype as indicated by the key in the box. The light grey circles surrounding the mean specify 95% confidence intervals.

Figure 5 The distribution of functional effect of each substitution across all tested backgrounds, calculated by subtracting derived from ancestral activity. 11 β -hydroxylation activity (A, light grey) increases in every background for the d83N and s492G substitutions. 18-oxidation activity is partially controlled by non-significant background-dependent effects d302E and q472L (B, dark grey) and its pattern is more complicated.