



PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ
ESCOLA DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA SAÚDE

**A COMPREHENSIVE ANALYSIS OF *AHRR* GENE AS CANDIDATE FOR CLEFT
LIP/PALATE**

MESTRANDA: BIANCA DOMIT WERNER LINNENKAMP
ORIENTADOR: ROBERTO HIROCHI HERAI

CURITIBA
2019



BIANCA DOMIT WERNER LINNENKAMP

**A COMPREHENSIVE ANALYSIS OF *AHRR* GENE AS CANDIDATE FOR CLEFT
LIP/PALATE**

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Ciências da Saúde, da Pontifícia Universidade Católica do Paraná-PUCPR.
Orientador: Prof. Dr. Roberto H. Herai

CURITIBA

2019

Dados da Catalogação na Publicação
Pontifícia Universidade Católica do Paraná
Sistema Integrado de Bibliotecas – SIBI/PUCPR
Biblioteca Central
Edilene de Oliveira dos Santos CRB 9 / 1636

L758c
2019 Linnenkamp, Bianca Domit Werner
A comprehensive analysis of *AHRR* gene as candidate for cleft lip/palate /
Bianca Domit Werner Linnenkamp ; orientador, Roberto H. Herai. -- 2019
78 f. : il. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná,
Curitiba, 2019
Inclui bibliografias

1. Genética humana. 2. Fissura palatina. 3. Lábio fissurado. 4. Fenótipo.
I. Herai, Roberto Hirochi. II. Pontifícia Universidade Católica do Paraná.
Programa de Pós-Graduação em Ciências da Saúde. III. Título.

CDD 20. ed. – 575.1

DEDICATÓRIA E AGRADECIMENTOS

Gostaria de dedicar esta conquista a minha família, amigos e professores pois sem eles jamais teria me tornado a pessoa que sou hoje. Agradeço a meus pais (Rene e Tita) por sempre valorizarem a educação e sempre me proporcionarem todas as oportunidades para alcançar meus objetivos além de ouvirem minhas incansáveis apresentações sobre genética sem entenderem nada mas tentando fazer o possível para me encorajar. Agradeço também ao meu irmão Marc que me lembra constantemente que me vê como seu maior exemplo fazendo com que cada dia eu queira ser melhor e me dedicar mais. Agradeço a todas as outras pessoas da minha família que também sempre me apoiaram. Não poderia esquecer meus amigos (Gaby, Aline, Nati, Bia, Manu, Thiago, Gui, João, Henrique, Vasco e outros) que estiveram ao meu lado durante a trajetória, alguns desde a infância enquanto outros se tornaram meus amigos ao longo da faculdade mas que me ensinaram e aconselharam nos últimos anos e sempre acreditaram em mim. Agradeço ao meu orientador de PIBIC Jr Prof Fábio Faucz quem primeiro me mostrou o que era genética, o que era pesquisa e despertou em mim a vontade de ser médica geneticista e pesquisadora, que foi o que me guiou nos anos de ensino médio, nas provas de vestibular e durante a faculdade. Agradeço a todos do laboratório de genética (Giane, Fábio, Rodrigo, Elke e tantos outros) que me acolheram quando tinha apenas 15 anos e confiaram em mim permanecendo até hoje a par de minhas conquistas e me dando apoio. Agradeço ao professor Salmo por ter sido meu orientador durante tantos anos de PIBIC e ter aceitado minha proposta de iniciar um mestrado sob sua orientação enquanto ainda estava na graduação. Agradeço a professora Cristina Baena por ter pensado em mim como candidata para o PIBIC Master e sempre ter sido um ombro amigo durante a graduação e durante as dificuldades do mestrado. Agradeço a toda a equipe do PIBIC da PUC, em especial à Cleybe que acreditou em mim durante os últimos nove anos e concedeu a oportunidade de fazer PIBIC Master. Agradeço a todos os meus professores médicos que foram exemplos e também despertaram em mim a vontade de ajudar os pacientes não somente através da pesquisa mas também através da

assistência e da empatia. Por fim agradeço ao meu orientador Prof Roberto Herai que me adotou já durante a metade do mestrado e sempre me estimulou a seguir em frente, tirou minhas dúvidas, explicou quantas vezes necessário e me respondeu mesmo na noite de um sábado pós feriado. Sem seus ensinamentos e sua orientação esta conquista não aconteceria.

RESUMO

Introdução: Fissuras labiopalatais (FL/P) compreendem uma gama de deformidades congênitas com falhas na formação de lábio e/ou palato, que afetam também a fala, audição, cognição e aparência visual dos portadores. Além disso, as FL/P correspondem ao tipo mais comum de malformação congênita craniofacial e ocorrem em 1 a cada 500 a 2500 nascidos vivos. Seu desenvolvimento está associado a fatores ambientais maternos (tabagismo, álcool, medicamentos, idade materna) e genéticos. A nível genético, embora já tenham sido descritos na literatura muitos genes e regiões cromossômicas que tenham associação causal com FL/P, há ainda grande heterogeneidade. Portanto, a identificação de novos genes e variantes causais torna-se fundamental para auxiliar nas análises genéticas de indivíduos com FL/P e fazer o aconselhamento genético de suas famílias. O gene *AHRR* (*Aryl-Hydrocarbon Receptor Repressor*) foi associado a FL/P porém ainda há poucos estudos que avaliem a interação genótipo-fenótipo do *AHRR* com FL/P. Entretanto, vários estudos associam a via molecular do *AHRR* com FL/P, o que torna este gene um candidato funcional na etiologia de FL/P. **Método:** Realização de revisão sistemática de literatura a partir da análise de banco de dados públicos (PUBMED) seguindo a ordem: Questão Principal (*AHRR* e FL/P); Questões secundárias; Palavras-chave (Cleft lip ou cleft palate ou orofacial cleft e *AHRR* ou sinônimos); Bases de busca; Seleção artigos (título e abstract); Coleta de dados; Análise de resultados. **Resultados:** A revisão da literatura permitiu identificar o gene *AHRR* como candidato associado com a fisiopatologia da FL/P. Uma análise *in silico* permitiu detectar que *AHRR* interage com outros genes previamente associados com a FL/P, como *ARNT* (*Aryl Hydrocarbon Receptor Nuclear Translocator*) e *CYP1A1* (*Cytochrome P450 Family 1 Subfamily A Member 1*), o que justifica o estudo molecular deste gene em portadores de fissuras. O gene *AHRR* é codificado para uma proteína que faz parte da cascata de sinalização *AHR* (*aryl hydrocarbon receptor*), e tem como função regular a toxicidade celular causada pelas dioxinas através da interação com o citocromo P450. A exposição à dioxina TCDD em embriões animais levou ao desenvolvimento de fissura palatina causada tanto por falha na fusão do palato e como também por ruptura pós fusão. A toxicidade do TCDD é mediada via *AHR* e mesmo em baixos níveis a exposição ao TCDD leva a uma menor tolerância a outros fatores patogênicos para FL/P. Além disso, o *AHRR* participa da regulação do desenvolvimento e da diferenciação celular, vias fundamentais para o desenvolvimento embriológico do lábio/palato. O *AHRR* também possui um papel na carcinogênese e recentemente demonstrou-se que portadores de FL/P e seus familiares possuem maior risco de desenvolverem diferentes tipos de tumores. O *AHRR* também é um marcador de tabagismo devido a alterações de metilação identificadas no DNA de tabagistas. Entretanto, o uso de ácido fólico é capaz de reverter parcialmente estas alterações de metilação o que é compatível com o fato de tabagismo materno e a falta de folato complementar na gestação serem fatores de risco para FL/P. **Conclusões:** Esta pesquisa permitiu identificar que o gene *AHRR* é

extremamente importante na resposta a dioxinas e é um exemplo da interação entre genética e ambiente que é imprescindível no desenvolvimento de diferentes doenças multifatoriais.

Palavras-chave: fissura labial, fissura palatina, *AHRR*, genética, fenótipo.

SUMÁRIO

1. INTRODUÇÃO	7
2. JUSTIFICATIVA	14
3. OBJETIVO GERAL	21
3.1. OBJETIVOS ESPECÍFICOS	21
4. MÉTODOS	22
5. RESULTADOS E DISCUSSÃO	25
5.1 O GENE <i>AHRR</i>	25
5.2 FUNÇÃO DO GENE <i>AHRR</i> E VIA MOLECULAR	31
6. CONCLUSÃO	35
7. REFERÊNCIAS	37
8. ARTIGO DERIVADO DO TRABALHO	41

1. INTRODUÇÃO

O termo fissuras orofaciais inclui uma série de deformidades congênitas das quais as fissuras orofaciais típicas não sindrômicas, ou seja, fissuras que acometem apenas a região de lábio e/ou palato, são as principais representantes (SHKOUKANI, CHEN e VONG, 2013). O termo fissura compreende uma falha ou quebra entre as estruturas orofaciais (SIMIONI et al., 2014), o que leva a anormalidades no lábio superior, alvéolo e/ou palato (MOSSEY et al., 2009). As fissuras apresentam um vasto espectro de acometimento e desencadeiam não somente distúrbios na fala, audição, cognição e aparência do portador (PRABHU S et al., 2012) como também possuem repercussões psicológicas e socioeconômicas (WEHBY e CASSEL, 2010; MARCUSSON, AKERLIND e PAULIN, 2001). Segundo Christensen, o impacto sobre o indivíduo pode ser extremo a ponto de os portadores de fissuras apresentarem uma taxa de suicídio mais elevada se comparados a não afetados (CHRISTENSEN et al., 2004), possivelmente devido a uma maior dificuldade de integração social (BERK e MARAZITA, 2002).

A prevalência de fissuras em lábio/palato (FL/P) varia conforme região geográfica e etnia além de depender de elementos ambientais e socioeconômicos (DIXON et al., 2011). A nível mundial, estima-se que 4 a 20 indivíduos a cada 10 mil nascidos vivos tenha fissura (LESLIE e MARAZITA, 2013, DIXON et al., 2011) de forma que esta é a malformação congênita craniofacial mais comum (JUGESSUR et al. 2009, MOSSEY et al. 2009). Entretanto, na América Latina há grande escassez e heterogeneidade de dados sobre anomalias craniofaciais. O Estudo Colaborativo Latino-Americano de Malformações Congênitas (ECLAMC) é responsável pela vigilância epidemiológica destas patologias em maternidades voluntárias e representa a principal fonte epidemiológica sobre fissuras (MONLLEO e GIL-DA-SILVA-LOPES, 2006). Segundo o ECLAMC, a prevalência de fissuras labiopalatais (FL/P) no Nordeste e no Sul do Brasil varia entre 9,72-11,89/10 mil. Já no Sudeste do Brasil, a prevalência situa-se entre 5,39-9,71/10 mil. Já, a prevalência de fissuras palatais (FP) no Sudeste do país varia de 3,09-5,01/10 mil enquanto no Nordeste e no Sul do Brasil é de

2,41-3,08/10 mil (ECLAMC, 2016; MONLLEO e GIL-DA-SILVA-LOPES, 2006; CASTILLA, LOPES-CAMELO e PAZ, 1995).

O atendimento de uma criança com fissuras é de alta complexidade e exige uma equipe multidisciplinar com experiência em cirurgia plástica, otorrinolaringologia, ortodontia, cirurgia bucomaxilofacial, fonoaudiologia, aconselhamento genético, enfermagem, psicologia, nutrição e assistência social (MOSSEY et al., 2009). Este atendimento visa permitir uma alimentação adequada, crescimento facial e dentário, desenvolvimento de fala e linguagem e também promove um caráter estético, ao restaurar a aparência do indivíduo (SHKOUKANI, CHEN e VONG, 2013). No entanto são necessários de 5 a 10 anos com tais tratamentos para atingir resultados satisfatórios, o que torna os custos elevados (LACE et al., 2006). Em 2000, o National Institute of Dental and Craniofacial Research dos Estados Unidos da América estimou os gastos anuais em 1 bilhão de dólares para o atendimento de portadores de fissuras. Já no Reino Unido, o centro nacional de serviços de saúde avaliou que 6,4 milhões de libras anuais seriam necessárias para investir em centros multiprofissionais capazes de atender 140 novos casos de fissura por ano (MONLLEO e GIL-DA-SILVA-LOPES, 2006; WHO, 2007). Estima-se que cada portador de fissura gere um custo de duzentos mil dólares em tratamentos ao longo de sua vida (BERK e MARAZITA, 2002). Logo, devido a complexibilidade de reabilitação, custo de tratamento e impacto aos portadores, FL/P representam um problema de saúde pública (LACE et al., 2006).

As FL/P podem ser classificadas de diferente formas, avaliando-se local afetado e grau de acometimento (MOSSEY et al., 2009; DIXON et al., 2011; HARVILLE et al., 2005). Uma das principais classificações baseia-se no forame incisivo como ponto de referência, o que permite a divisão em pré-forame, transforame e pós- forame (SPINA et al., 1972). Outra forma de classificar as fissuras consiste em separá-las em unilaterais ou bilaterais e observar se há acometimento somente de lábio, lábio e palato e palato isolado, com extensão completa ou incompleta (DIXON et al., 2011). Na figura 1 é possível visualizar fissuras unilaterais que acometem apenas o palato (1a), apenas o lábio (1b), lábio e palato de forma incompleta (1c) e lábio e palato completamente

(1d). Já na linha inferior é possível ver o acometimento bilateral apenas em palato (1e), lábio (1f), lábio e palato incompleto (1g) e lábio e palato de forma completa (1h). Por outro lado, na figura 2 é possível observar pacientes com diferentes fenótipos: acometimento leve de lábio superior denominado microforma (2a), acometimento de lábio superior esquerdo (2b), lábio e palato esquerdos (2c) e acometimento bilateral envolvendo lábio e palato (2d).

Figura 1. Classificação de Fissuras

a.fissura apenas em palato, b. fissura apenas em lábio, c. fissura em lábio e palato incompleta, d. fissura em lábio e palato completa. Linha superior: acometimento unilateral. Linha inferior: acometimento bilateral.

Fonte: Dixon et al., 2009

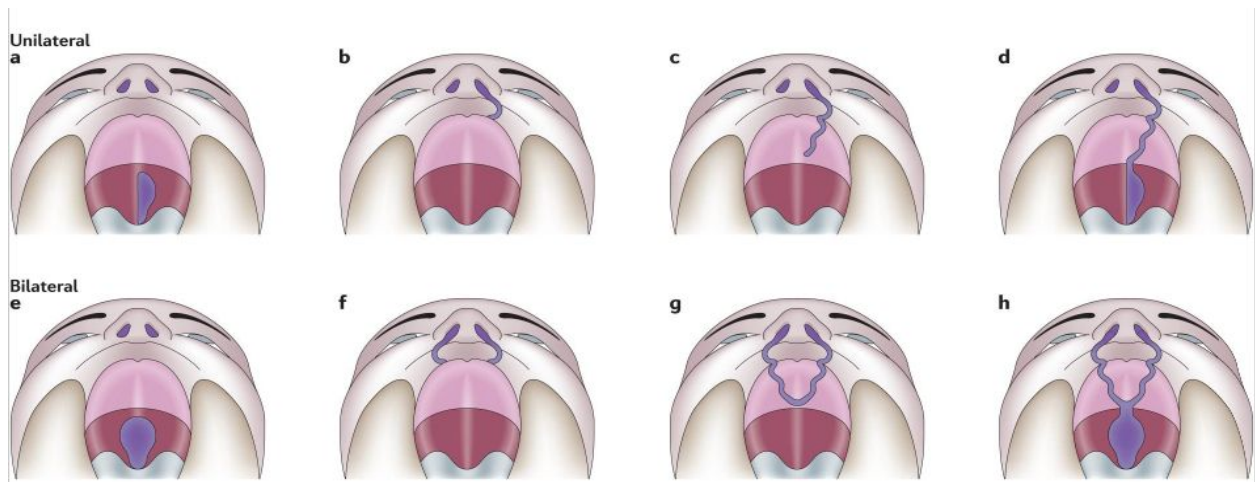
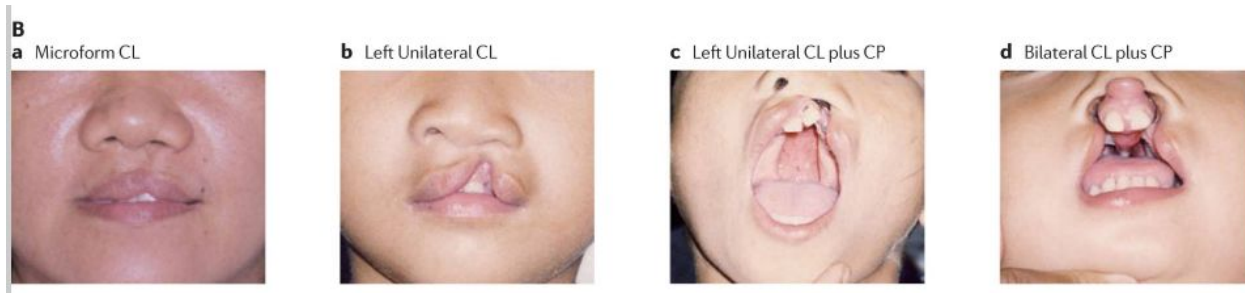


Figura 2. Fenótipos de fissuras a.microforma de fissura labial, b. fissura unilateral em lábio, c. fissura unilateral em lábio e palato, d. fissura bilateral em lábio e palato.

Fonte: Dixon et al., 2009



Enquanto fissuras labiais são duas vezes mais comuns em homens, fissuras palatais acometem mais mulheres (2:1) (JAMES, 2000; MOSSEY et al., 2009). Enquanto muitos estudos dividem as fissuras em FL/P e FP, aqueles que distinguiram FL/P de fissuras apenas em lábio (FL) demonstraram que FL/P são duas vezes mais prevalentes que FL (LESLIE e MARAZITA, 2013). Dentre as fissuras, o acometimento unilateral é mais comum de forma que em 75% dos casos de FL/P apenas um lado é acometido (LESLIE e MARAZITA, 2013). O lado predominantemente mais afetado é o esquerdo em uma proporção de 2:1 (DIXON et al., 2011; DASKALOGIANNAKIS et al., 1998).

Além desta classificação, é importante diferenciar fissuras sindrômicas de não sindrômicas. Quando a fissura ocorre isoladamente ela é classificada como não sindrômica (KOHLI SS e KOHLI VS, 2012). Por outro lado, fissuras associadas a outra malformação em padrão já reconhecido ou associadas a disfunções cognitivas são classificadas como fissuras sindrômicas (MOSSEY et al., 2009, STUPPIA et al., 2011, LESLIE e MARAZITA, 2013). Formas sindrômicas correspondem a cerca de 30% dos casos quando há fissura em lábio com ou sem acometimento de palato (FL/P). Já em fissuras palatais isoladas (FP), a presença de formas sindrômicas é mais frequente e corresponde a 50% (JONES, 1998; CALZOLARI et al., 2007; LESLIE e MARAZITA, 2013; DIXON et al., 2011; SÉTO-SALVIA e STANIER, 2014; WYSZYNSKI, SARKÖZI e CZEIZEL, 2006; (SHKOUKANI, CHEN e VONG, 2013). Entretanto, em 2014, Monlleó e colaboradores descreveram uma amostra brasileira com maiores taxas de casos sindrômicos (55% dos casos de FL/P sindrômica e 79% de FP sindrômica). Neste trabalho os autores propuseram que a presença de 4 ou mais sinais menores seria

suficiente para classificar a fissura como uma forma sindrômica, independente de ser em padrão específico reconhecido. Tal mudança pode ter contribuído para a divergência dos achados com a literatura preexistente (MONLLEO et al., 2014).

As formas sindrômicas de fissuras representam um grupo heterogêneo composto por mais de 275 condições diferentes das quais apenas 75% possuem etiologia genética conhecida (STUPPIA et al., 2011; MOSSEY et al., 2009, LESLIE e MARAZITA, 2013). Em grande parte destes casos, a etiologia pode ser atribuída a um único fator causal que pode ser monogênico, cromossômico ou, eventualmente, ambiental ou teratogênico (STUPPIA et al., 2011, (STUPPIA et al., 2011; LESLIE e MARAZITA, 2013; DIXON et al., 2011; SÉTO-SALVIA e STANIER, 2014; WYSZYNSKI, SARKÖZI e CZEIZEL, 2006; IPDTC, 2011). A síndrome de Van der Woude, por exemplo, é a forma sindrômica de fissuras mais comum (2% dos casos de FL/P) e sua ocorrência é justificada pela presença de variantes no gene *Interferon Receptor Factor 6 - IRF6* (KONDO et al., 2002; STUPPIA et al., 2011,).

Em contrapartida, formas não sindrômicas possuem etiologia multifatorial, oriundas da complexa interação entre genes e ambiente (STUPPIA et al., 2011; MURRAY, 2002; DIXON et al., 2011) em que cada fator possui um papel menor dentro de um mecanismo intrincado (STUPPIA et al., 2011). O tabagismo materno é um dos principais fatores de risco ambientais e eleva em 30 a 50% a chance do bebê desenvolver fissuras orofaciais (SHI, WEHBY e MURRAY, 2008; WYSZYNSKI e WU, 2002). Já o consumo de álcool no período gestacional pode aumentar o risco de fissuras labiopalatais em 1,5 a 4,7 vezes. (DE ROO et al., 2008; MUNGER et al. 1996). Outro fator é a idade materna avançada ao gestar (acima de 40 anos), o que eleva em até 56% o risco de FL/P se comparado a mulher na faixa etária de 20 a 29 anos (HERHKRATH et al., 2012). Além destes fatores ambientais, a administração de medicamentos durante a gestação pode desencadear um efeito teratogênico. Os principais representantes destes fármacos são os antiinflamatórios, corticosteróides, antifúngicos, antibióticos, progestágenos e anticonvulsivantes, em particular a fenitoína, (LEITE et al., 2005; HILL et al., 1988; MCBRIDE, 1961; LOFFREDO et al., 1994). Além

disso, a não administração de folato suplementar já no primeiro trimestre gestacional pode triplicar o risco de malformações faciais (SHAW et al., 2002; WILCOX et al., 2007). Outros fatores de risco adicionais incluem exposição ocupacional, hipertermia, estresse, obesidade materna, suplementação hormonal, exposição a radiação ionizante e infecções maternas (LEITE, PAUMGARTTEN e KOIFMAN, 2002; DIXON et al., 2011).

Em decorrência de sua etiologia complexa, tem-se aumentado cada vez mais o interesse por esta malformação, de forma que múltiplos estudos são realizados com o intuito de identificar seus fundamentos genéticos (RAHIMOV et al., 2012; LESLIE e MARAZITA, 2013; DIXON et al., 2011). Os estudos para a compreensão da etiologia genética das FL/P se iniciaram em 1942 por Fogh-Andersen que propôs que fatores genéticos contribuíam para seu aparecimento, a partir da observação do aumento de sua frequência em famílias que possuíam um paciente afetado. Atualmente, com o grande avanço molecular na área, são possíveis incontáveis estratégias para se verificar a associação de genes a esta malformação tais como estudos de associação de genes candidatos em larga escala; estudos em organismos modelo; sequenciamento direto de amostras de indivíduos afetados; identificação de anomalias cromossômicas ou microdeleções e estudos de associação por varredura genômica (GWAS) (SOUZA, 2015; LIU et al., 2015; RAHIMOV et al., 2012).

Novos estudos de análise genômica demonstram que FL/P não sindrômicas apresentam um caráter multigênico (MANGOLD, LUDWIG e NOETHEN, 2011; MARAZITA, 2012). Atualmente, existem genes já associados ao desenvolvimento de FL/P não sindrômicas, como *FOX1* (*forkhead box 1*), *IRF6* (*interferon regulatory factor 6*), *MAFB* (*MAF b ZIP transcription factor B*), *ABCA4* (*ATP binding cassette subfamily A member 4*), *MSX1* (*msh homeobox 1*), *VAX1* (*ventral anterior homeobox 1*) (BIRNBAUM et al., 2009). Também existem genes em potencial (genes candidatos) tais como *GLI2* (*GLI family zinc finger 2*), *MSX2* (*msh homeobox 2*), *SKI* (*SKI protooncogene*), *SATB2* (*special at- rich binding sequence protein 2*), *SPRY2* (*sprouty drosophila homolog of 2*), *TBX22* (*T- box 22*), *PVRL1* (*poliovirus receptor like 1*),

FGFR2 (fibroblast growth factor receptor 2), *BMP4* (bone morphogenic protein), *MTHFR* (methylenetetrahydrofolate reductase), *TGFA* (transforming growth factor A), *TGFB3* (transforming growth factor Beta 3), *CLPTM1* (cleft lip and palate-associated transmembrane protein 1) e a região ch8q24 (BEATY et al., 2010; OSOEGAWA et al., 2008; SUZUKI et al., 2009; KOHLI SS e KOHLI VS, 2012). Em estudo mais recente por Leslie e colaboradores, novos genes foram associados a maior risco de FL/P não sindrômicas em populações específicas, especificamente *N4BP2* (*NEDD4 binding protein 2*), *CDSN* (corneodesmosin), *PRTG* (protogenin) e *AHRR* (aryl hydrocarbon receptor repressor) (LESLIE et al., 2017).

Segundo este mesmo estudo de Leslie, algumas variantes expressas em menor frequência na população podem promover um efeito protetor contra fissuras enquanto outras variantes podem levar a uma maior predisposição a sua ocorrência. Neste mesmo estudo, o gene *AHRR* foi descrito como catalisador do risco de FL/P não sindrômicas em população europeia (LESLIE et al., 2017). Por se tratar de uma descoberta muito recente, ainda não há outros estudos que correlacionem variantes neste gene com fissuras.

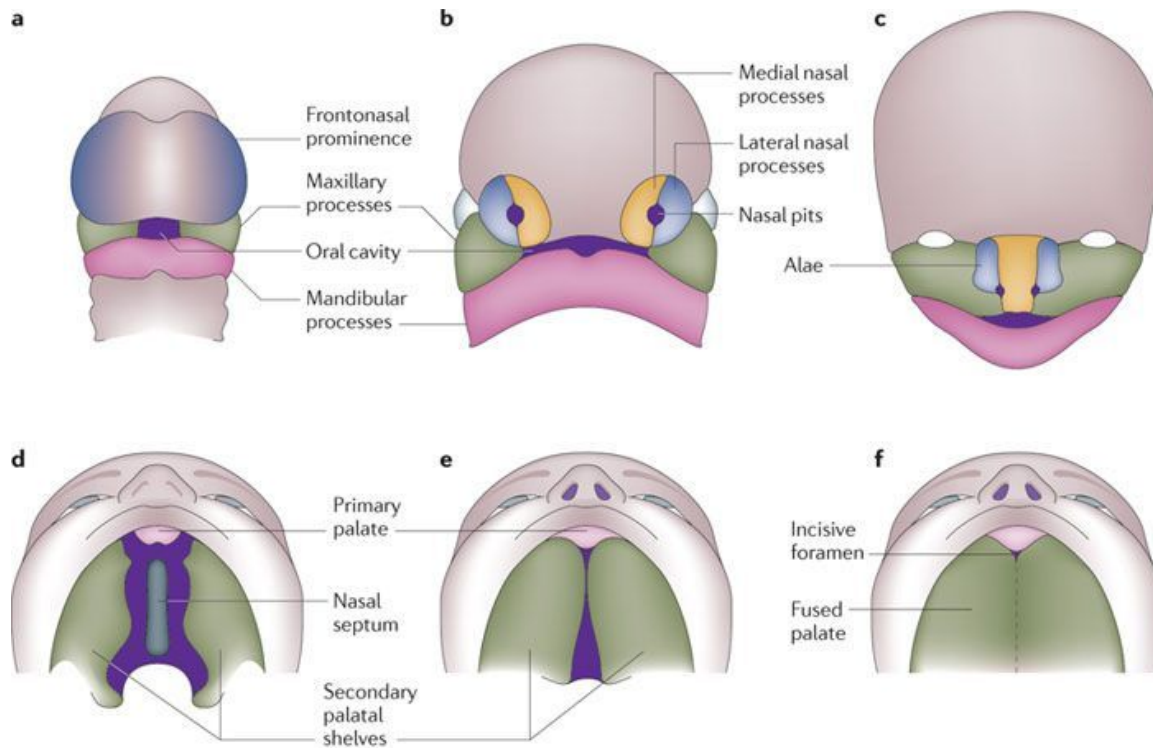
2. JUSTIFICATIVA

Fissuras orofaciais compõem um grupo heterogêneo de doenças que afetam a estrutura da face e a cavidade oral os quais foram divididos em fissuras em lábio com ou sem comprometimento de palato e apenas em palato. Durante um período, acreditou-se que FP e FL/P seriam variantes de um mesmo defeito com graus de acometimento diferentes (MARAZITA, 2012). Apesar de palato primário e secundário apresentarem origens distintas em seu desenvolvimento embrionário (figura 3), FP e FL/P compartilham o defeito do palato primário o que motivou a redivisão formando um grupo comum de FL/P (FOGH-ANDERSEN, 1942; FRASER, 1955). No entanto, estudos epidemiológicos (GROSEN et al., 2010) e biológicos posteriores (RAHIMOV et al., 2008; LUDWIG et al., 2012) sugeriram que FP e FL/P possuem etiologias genéticas distintas. Todavia, vias moleculares em comum podem atuar em ambos os grupos e levar a padrões mistos de fissuras dentro do mesmo pedigree (RAHIMOV et al., 2012).

Figura 3. Desenvolvimento do lábio e do palato. a. desenvolvimento da proeminência frontonasal, processos maxilares pareados e processos mandibulares pareados ao redor da estrutura primitiva da cavidade oral na quarta semana de desenvolvimento, b. na quinta semana, os pits nasais já estão formados, levando ao pareamento dos processos nasais laterais e mediais, c. os processos nasais mediais unem-se com os processos maxilares para formar lábio superior e palato primário ao final da sexta semana enquanto os processos nasais laterais formam as asas nasais e os processos maxilares fundem-se para formar a mandíbula, d. na sexta semana de embriogênese o palato secundário surge como crescimentos bilaterais a partir do processo maxilar e crescem verticalmente para baixo na lateral da língua, e. em sequência, as prateleiras palatinas elevam-se para uma posição horizontal acima da língua, entram em contato e iniciam o processo de fusão, f. ocorre a fusão das

prateleiras palatinas dividindo o espaço oronasal em duas cavidades separadas (oral e nasal).

Fonte: Dixon et al., 2009

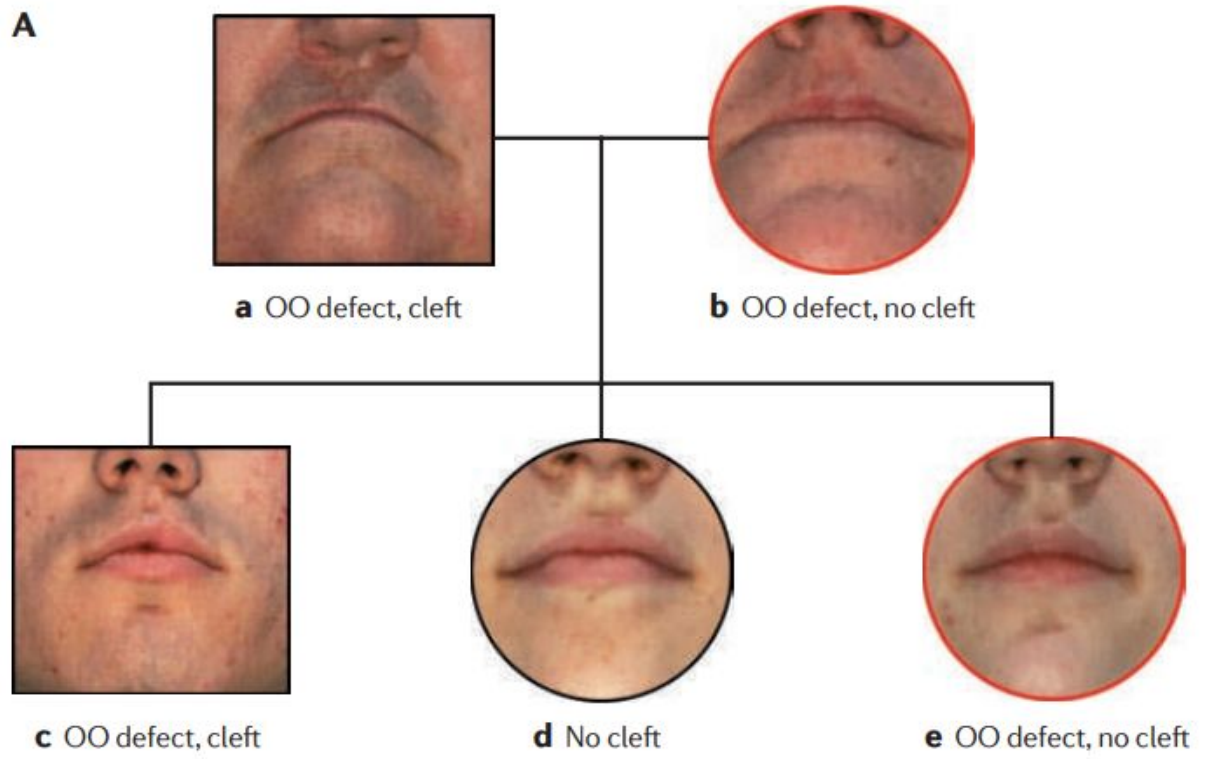


Apesar da vasta gama de apresentações fenotípicas, fissuras orofaciais são tipicamente vistas como um traço simples e qualitativo, dividido em afetados e não afetados. Porém, evidências mais recentes sugerem que indivíduos com fissuras fazem parte de um espectro de fenótipos ocultos que incluem fissuras apenas em lábio, lábio e palato e apenas em palato, além de também incluir múltiplas características subclínicas que podem estar presentes em familiares não afetados por fissuras (WEINBERG et al., 2009, LESLIE e MARAZITA, 2013). Nestes fenótipos podem ser incluídas alterações de medidas craniofaciais, anormalidades dentárias (agenesia, microdontia e dente supranumerário), diferenças cerebrais estruturais, alterações dermatoglíficas em lábios, microformas de fissuras, defeitos no músculo *orbicularis oris* (figura 4), úvula bífida, FP submucosa e insuficiência velofaríngea (LESLIE e

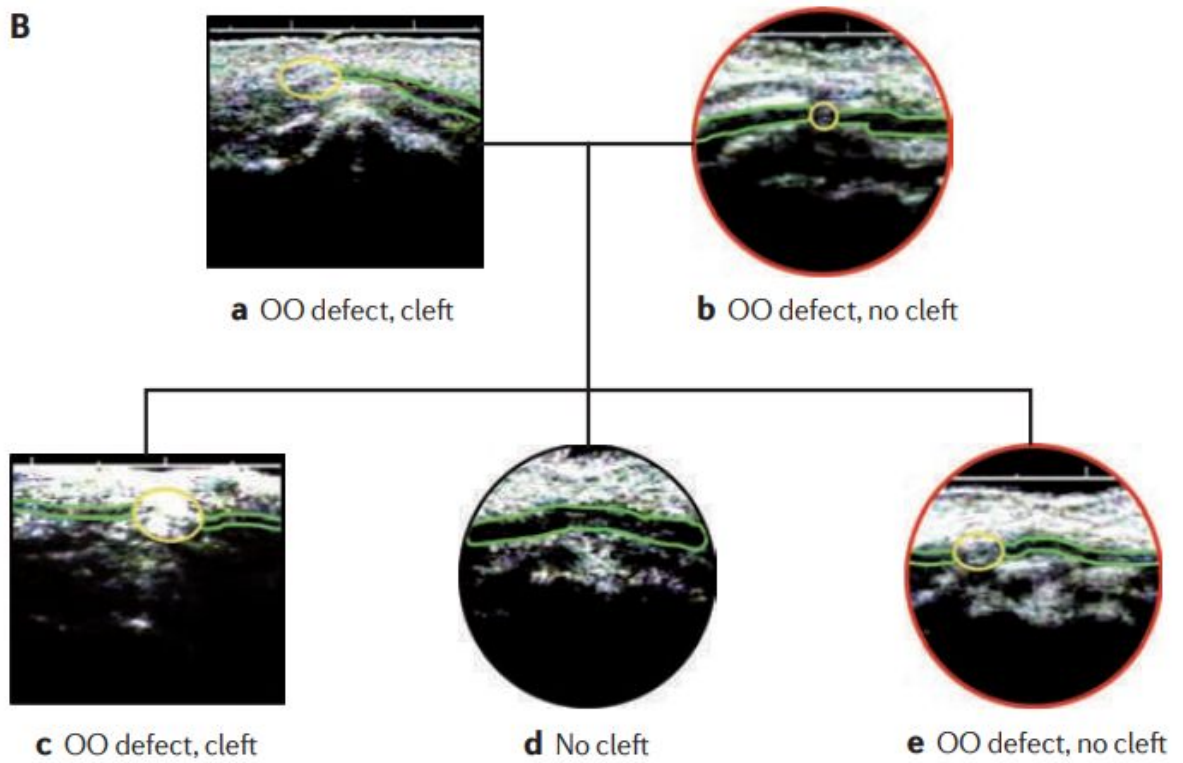
MARAZITA, 2013). Esses fenótipos subclínicos podem ajudar a explicar a penetrância incompleta ou aparente falta de herança mendeliana observada em famílias com fissuras. Tais manifestações também podem explicar a discordância clínica entre gêmeos monozigóticos apesar de o risco de seus filhos nascerem afetados ser idêntico. Logo, é importante a correta identificação destes fenótipos para dividir adequadamente estes grupos e estudar suas etiologias.

Figura 4. Fenótipos subclínicos a. Fotografias da região de lábio superior de membros de uma família com dois membros portadores de fissuras não sindrômicas corrigidas cirurgicamente (indivíduos a, c). Os três outros membros da família (b, d, e) não apresentam defeitos externos visíveis porém dois indivíduos circulados em vermelho (b, e) possuem um fenótipo subclínico caracterizado por defeito no músculo *orbicularis oris*. b. Imagens ultrassonográficas das regiões de lábio superior de todos os membros da família exibida na figura a. Nos dois membros circulados em vermelho (b, e) e nos dois membros portadores de fissura (a, c) é possível observar interrupções no músculo *orbicularis oris*. Demonstrando múltiplos fenótipos encontrados em uma única família.

A



B



Apesar de o componente familiar em fissuras já ter sido percebido há quase trezentos anos, estudos genéticos formais iniciaram-se nas últimas décadas (FOGH-ANDERSEN, 1942). Análises de segregação e estudos com gêmeos corroboram o papel da genética em fissuras e demonstraram que FL/P não sindrômicas possuem uma alta taxa de recorrência familiar. Estima-se que familiares de primeiro grau de portadores de FL/P possuem 32 vezes mais risco de fissuras se comparados a famílias sem membros afetados (SIVERTSEN et al., 2008). Outro fator demonstrado foi que a taxa de concordância entre gêmeos monozigóticos é de 40-60% enquanto em gêmeos dizigóticos esta taxa é de 3-5% o que sugere uma herança genética forte mas não completamente pura (LITTLE e BRYAN, 1986).

Enquanto múltiplos fatores ambientais já se mostraram relacionados a maiores taxas de fissuras, os dados sobre fatores genéticos predisponentes ainda são escassos (DIXON et al., 2011), principalmente em casos de fissuras não sindrômicas. A identificação de genes que contribuem para FL/P não sindrômicas tem sido objeto de estudos durante décadas. Múltiplas abordagens foram realizadas tais com estudos de ligação, rearranjos genômicos, genes candidatos, GWAS (genome wide association studies) no entanto, o sucesso atingido ainda é modesto (DIXON et al., 2011; MARAZITA, 2012) apesar de FL/P representarem um dos melhores exemplos de sucesso na aplicação de técnicas modernas de genética molecular no estudo de doenças complexas. Apesar de os estudos conseguirem identificar genes e regiões associadas, novos estudos de sequenciamento são necessários para identificar variantes funcionais e regiões regulatórias. Ainda, os estudos de ligação demonstraram diferentes regiões alteradas se comparados aos GWAS o que corrobora a necessidade de realizar sequenciamento destes genes a fim de detectar variantes adicionais (LESLIE E MARAZITA, 2013).

Embora diferentes estudos evidenciem uma ou outra causa, ainda há uma grande necessidade de atualizar o conhecimento existente e buscar novas associações que possam ter efeitos aditivos ao risco de fissura. Diferentemente do estudo de genes

como o *IRF6*, o qual já foi associado múltiplas vezes a fissuras, o estudo de um gene mais novo pode identificar variantes nunca descritas além de ampliar o conhecimento sobre o gene e seu papel na formação do lábio e do palato.

Cada vez mais torna-se necessária a inclusão da diversidade fenotípica e étnica nos estudos, assim como a investigação das vias moleculares, das interações gene-gene, das variantes funcionais e a compreensão de seu significado etiológico e fisiopatológico a fim de traduzir este resultados e aplicá-los na prática clínica mundialmente (LESLIE e MARAZITA, 2013).

Logo, a determinação acurada das variantes genéticas presentes em fissuras labiopalatais não sindrômicas poderá responder questões relacionadas a variabilidade de expressão dos fenótipos dos pacientes, assim como contribuir para uma melhor compreensão da etiologia e risco de recorrência que são importantes para a orientação e o esclarecimento dos portadores e familiares.

O gene *AHRR* tornou-se um candidato após estudos recentes demonstrarem que variantes em sua região podem implicar em maiores riscos de fissuras. Em 2017, um estudo brasileiro que avaliou CNVs (variações de números de cópias), identificou e associou múltiplos *loci* a ocorrência de fissuras não sindrômicas. Neste estudo foi descrita uma duplicação na região 5p15.33 (SILVA e OLIVEIRA et al., 2017) predispondo a maiores riscos de fissuras. Este estudo sugere que o gene *AHRR* é um gene candidato posicional uma vez que outros genes da região cromossômica 5p15.33 foram associados com fissuras não sindrômicas.

Além disso, novos estudos sugerem que pacientes portadores de fissuras e seus familiares de primeiro e segundo grau possuem maior risco de desenvolverem tumores malignos (POPOFF et al., 2013; LIMA et al., 2013). Em 2016, Dunkhase e colaboradores relataram uma sobreposição entre fatores genéticos predisponentes a cânceres e fissuras. Neste estudo, múltiplos *loci* foram identificados em ambas patologias o que sugeriu uma ação de pleiotropismo nos genes compreendidos nestas regiões. A região 5p15.33 foi uma entre várias descritas e foi associada tanto a câncer de pulmão quanto a fissuras não sindrômicas (LANDI et al., 2009; DUNKHASE et al.,

2016). Logo, o gene *AHRR* localizado neste *locus* e que também possui um papel na tumorigênese, mostra-se como um forte candidato funcional e posicional.

Ainda, a associação específica do gene *AHRR* com fissuras é extremamente recente (2017) e foi realizada apenas em um estudo de associação com 3621 participantes. Neste estudo, variantes de baixa frequência do *AHRR* foram associadas com maior risco de fissuras em população europeia (LESLIE et al., 2017).

Ademais, a via molecular do *AHRR* e outros genes na via também foram descritos em estudos de predisposição a fissuras. Entretanto ainda há uma escassez na literatura de dados capazes de permitir associação genótipo fenótipo.

3. OBJETIVO GERAL

Fazer uma revisão sistemática sobre o gene *AHRR* e sua relação com Fissura Lábio/Palatina (FL/P).

3.1 OBJETIVOS ESPECÍFICOS

1. Buscar achados sobre *AHRR* e FL/P
2. Descrever a estrutura do gene *AHRR*, isoformas e expressão tecidual
3. Descrever a estrutura proteica AHRR
4. Descrever interação molecular do gene *AHRR*
5. Buscar e descrever moléculas associadas ao *AHRR* com FL/P
6. Relacionar achados do *AHRR* com fenótipos de FL/P

4. MÉTODOS

Uma revisão sistemática da literatura (RSL) foi realizada de acordo com o protocolo de Kitchenham e Charters (KITCHENHAM e CHARTERS, 2007). O objetivo principal foi de realizar a revisão sistemática a fim de identificar e sintetizar as associações já descritas entre o gene *AHRR* e fissuras labiopalatinas a nível molecular, celular e fenotípico. A revisão constituiu-se nas seguintes etapas: análise bancos de dados públicos (PUBMED); revisão sistemática da literatura: questão principal; questões secundárias; palavras-chave; bases de busca; seleção artigos; coleta de dados e análise de resultados. A pergunta de pesquisa principal foi: qual a correlação entre o gene *AHRR* e fissura lábio/palato. A pergunta principal foi subdividida em: qual a correlação entre o gene *AHRR* com fenótipos moleculares de FL/P, quais as variantes genéticas no *AHRR* associadas a FL/P.

Critérios de inclusão e exclusão: Os seguintes critérios de inclusão foram definidos: 1) artigos publicados até maio de 2019; 2) considerados apenas artigos em inglês; 3) artigos que avaliaram FL/P e *AHRR* e genes relacionados a *AHRR*; 4) artigos que apresentaram variantes genéticas encontradas em FL/P. O critério de exclusão foi: 1) artigos que utilizaram AHRR, AHHR e AHH como iniciais com outros significados.

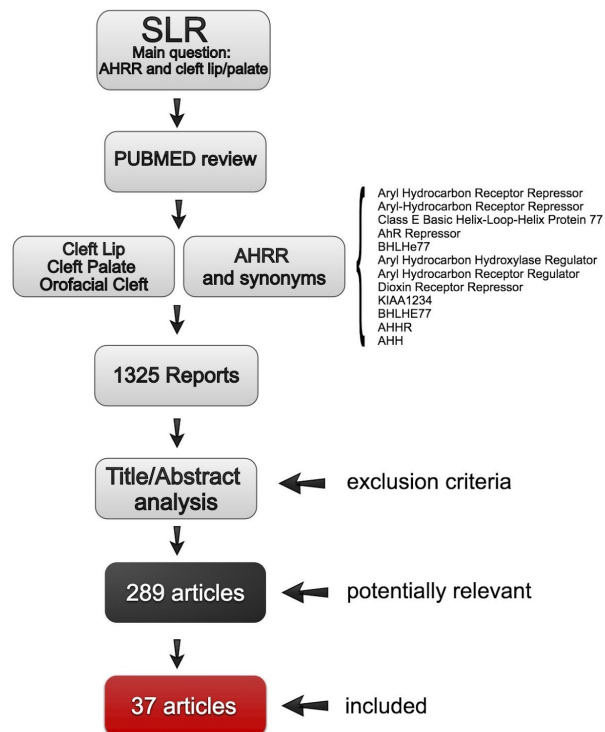
A busca de dados foi feita na base PUBMED através de uma busca eletrônica. PUBMED é uma base de busca de dados gratuita desenvolvida e mantida pelo NCBI (National Center for Biotechnology Information). Esta base possui informações sobre ciências da vida, ciências comportamentais, químicas e bioengenharia através de cerca de 25 milhões de citações e sumários nas áreas de saúde e biomedicina (<https://www.ncbi.nlm.nih.gov/pubmed/>).

A estratégia de busca utilizada a fim de se obter o maior número de artigos possível baseou-se no uso dos termos e seus sinônimos: cleft lip, cleft palate, orofacial cleft, *AHRR*, aryl-hydrocarbon receptor repressor. Em sequência, obtivemos palavras-chave que geraram a seguinte chave de busca: (((((((((((((((cleft lip[Title/Abstract]) OR cleft palate[Title/Abstract]) OR orofacial cleft[Title/Abstract]) AND ahrr[Title/Abstract]) OR Aryl Hydrocarbon Receptor Repressor[Title/Abstract]) OR

Aryl-Hydrocarbon Receptor Repressor[Title/Abstract]) OR Class E Basic Helix-Loop-Helix Protein 77[Title/Abstract]) OR AhR Repressor[Title/Abstract]) OR BHLHe77[Title/Abstract]) OR Aryl Hydrocarbon Hydroxylase Regulator[Title/Abstract]) OR Aryl Hydrocarbon Receptor Regulator[Title/Abstract]) OR Dioxin Receptor Repressor[Title/Abstract]) OR KIAA1234[Title/Abstract]) OR BHLHE77[Title/Abstract]) OR AHHR[Title/Abstract]) OR AHH[Title/Abstract]. Apesar de os termos BHLHe77 e KIAA1234 serem descritos como sinônimos para *AHRR*, tais termos não foram encontrados no PubMed. O string de busca gerado resultou em 1325 artigos. Uma triagem inicial foi conduzida primeiramente por título e em seguida por análise dos sumários. Após a triagem foram identificados 29 títulos e 1007 sumários que utilizavam as iniciais AHH com outro significado. Os seguintes significados para as iniciais AHH foram encontrados: acquired hypocalciuric hypercalcemia, "ahh" referindo a fonação de ahh durante exame do palato, acquired hypogonadotropic hypogonadism, células tipo AHH-1, atrial-his-atrial, acyl heck-heck, "ahh-choo" referindo o som de espirro, acute hypobaric hypoxia, American Heart Association, aliphatic halogenated hydrocarbons, arecanut husk hydrochar, ahh family, approved by the Danish Data Protection Agency (approval number: AHH-2017-048), axillary hyperhidrosis, apex of the hook of the hamate, acute hypervolemic hemodilution, chimeric peptide Ala-His-His, adult-onset hypogonadotropic hypogonadism, amaranth hydrolysate, acute hemophilic hemarthrosis, two-hop adaptive hop by hop vector-based forwarding (2hop-AHH-VBF), A-Helping-Hand, adult human heart, benzo(a)pyrene hydroxylase, A.hydraspica [n-hexane], acute hypoxic hepatitis, high-dose AK in alcohol/HFD group, automated hotspot-hunting procedure, acute hypobaric hypoxia, high altitude hypoxia, acquired hypocalciuric hypercalcemia. Após eliminação destes artigos, 289 estudos foram considerados potencialmente relevantes para a busca e continham informações sobre *AHRR* e sua via molecular. Os critérios de inclusão e exclusão previamente descritos foram aplicados e após leitura 37 artigos foram incluídos na revisão (Figura 5).

Figura 5. Fluxograma das etapas da revisão sistemática

Pergunta principal da revisão sistemática de literatura AHRR e fissura lábio/palato; revisão no PUBMED; termos utilizados no string de busca e sinônimos; resultado inicial de 1325 artigos; análise de títulos e sumários aplicando-se critérios de exclusão; 289 artigos restantes considerados potencialmente relevantes; 37 incluídos após leitura devido a relevância para a pergunta principal.



5. RESULTADOS E DISCUSSÃO

5.1 O GENE *AHRR*

O gene *AHRR* possui 2 isoformas descritas na base NCBI Reference sequence (tabela 1) e está localizado na banda cromossômica 5p15.3 (BABA et al., 2001).

Tabela 1. Dados sobre isoformas do gene *AHRR*

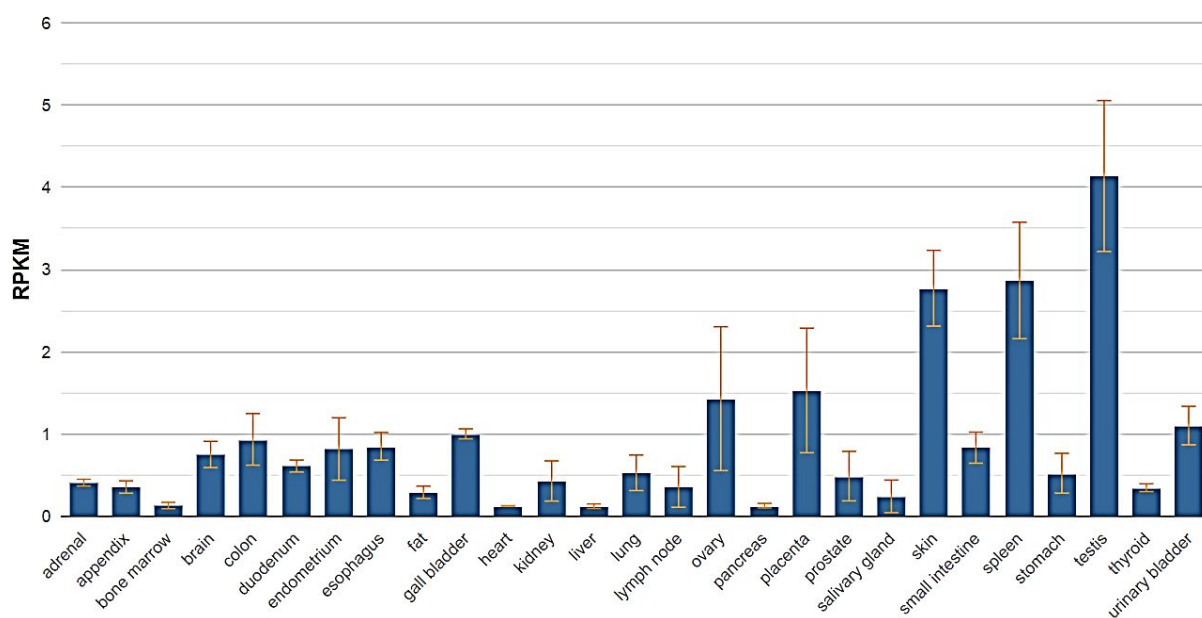
Isoforma	NCBI Reference Sequence:	Tamanho	Região CDS	Tamanho da região 5'UTR	Tamanho da região 3'UTR
Homo sapiens aryl-hydrocarbon receptor repressor (<i>AHRR</i>), transcript variant 1, mRNA	NM_020731.4	5676 bp	Posição 45..2204 (tamanho 2160)	44	3472
Homo sapiens aryl-hydrocarbon receptor repressor (<i>AHRR</i>), transcript variant 2, mRNA	NM_001242412.1	5622 bp	Posição 45..2150 (tamanho 2106)	44	3472

A isoforma NM_020731.4 possui 12 exons e 11 introns, e a segunda isoforma NM_001242412.1 11 exons e 10 introns. Embora hajam outras isoformas dependendo da base de dados considerada, estas duas correspondem às duas formas do gene anotadas como revisadas pelo banco RefSeq (NCBI).

Apesar de existirem diferentes isoformas, Fagerberg testou a expressão do gene *AHRR* em 27 tecidos humanos. Os diferentes níveis de expressão podem ser vistos na figura 6. A expressão mais elevada foi encontrada em testículo, seguido de baço, pele, placenta e ovário (FAGERBERG et al., 2014). A partir de tais dados é possível inferir que o *AHRR* possa impactar no desenvolvimento embrionário devido a sua elevada expressão em gônadas e placenta. Também é possível supor um papel no desenvolvimento da pele de forma que estas informações combinadas podem indicar a importância da expressão adequada do *AHRR* para a formação do lábio.

Figura 6. Expressão do gene *AHRR* em 27 tecidos humanos. O eixo Y está representado na unidade RPKM (reads per kilobase per million of reads); o eixo X representa o tecido analisado.

Fonte: Extraído do NCBI através de dados de Fagerberg et al., 2014.



O gene *AHRR* codifica uma proteína homônima que possui dois domínios já conhecidos e que podem ser visualizados nas figuras 7 e 8. O primeiro domínio é o bHLH - *Myc-type, basic helix-loop-helix (bHLH) domain profile* e o segundo consiste no PAS - *PAS repeat profile*. O domínio bHLH, que ocupa a posição 35 a 89, possui ação

na dimerização de proteínas (GO:0046983) o que leva a formação de uma macromolécula composta pela associação não covalente de duas subunidades idênticas ou não. Já o segundo domínio ocupa a posição 114 a 180 e interage com proteínas sinalizadoras na função de domínio sensor de sinais tais como luz, oxidação e potencial redox (TAYLOR e ZHULIN, 1999). O termo PAS origina-se de Per-period circadian protein, Arnt-Ah receptor nuclear translocator protein e Sim-single-minded protein (https://smart.embl.de/smart/show_motifs.pl?GENOMIC=1&DO_PFAM=DO_PFAM&INCLUDE_SIGNALP=INCLUDE_SIGNALP&ID=9606.ENSFP00000323816).

Figura 7. Domínios conhecidos de AHRR, AHR e ARNT

Representação dos domínios conhecidos da proteína AHRR e de outras duas proteínas relacionadas que possuem domínios similares. AHRR, AHR e ARNT possuem um domínio bHLH e um domínio PAS A que permite a ligação ao DNA e dimerização (AHR-ARNT e AHRR-ARNT). AHR tem um domínio PAS B que permite ligação a ligantes como o TCDD, uma dioxina. AHRR tem um domínio repressor que liga a correpressores e sofre SUMOilação. ARNT e AHR possuem domínios transativadores que permitem ligação a coativadores. AHR também possui um domínio C-terminal subdividido em subdomínios ricos em ácido, glutamina e serina/treonina/prolina.

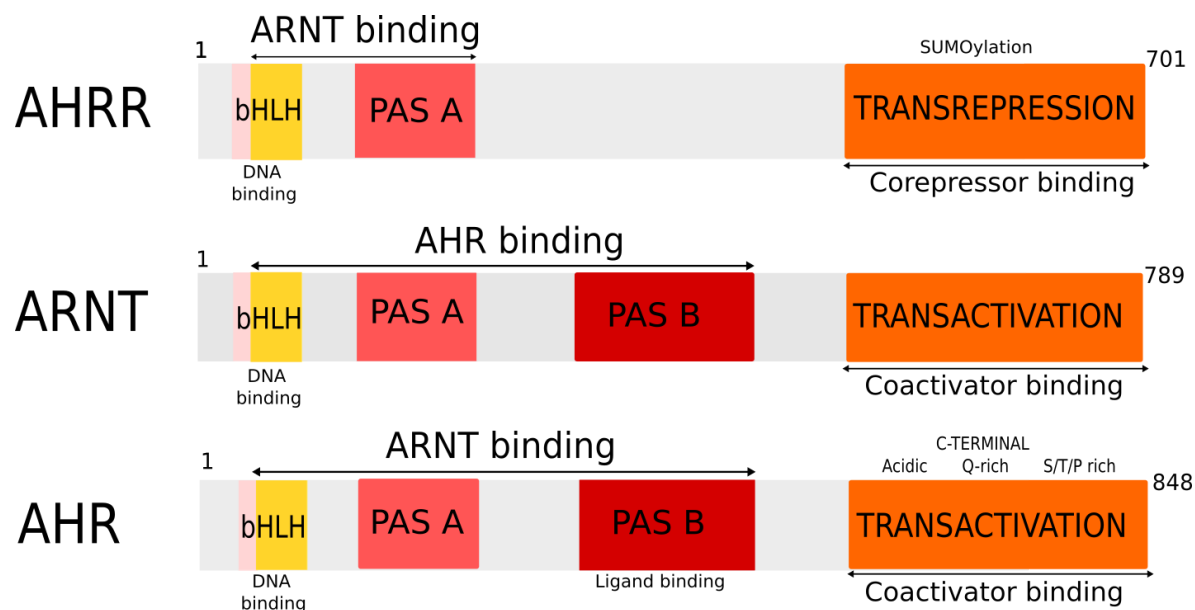
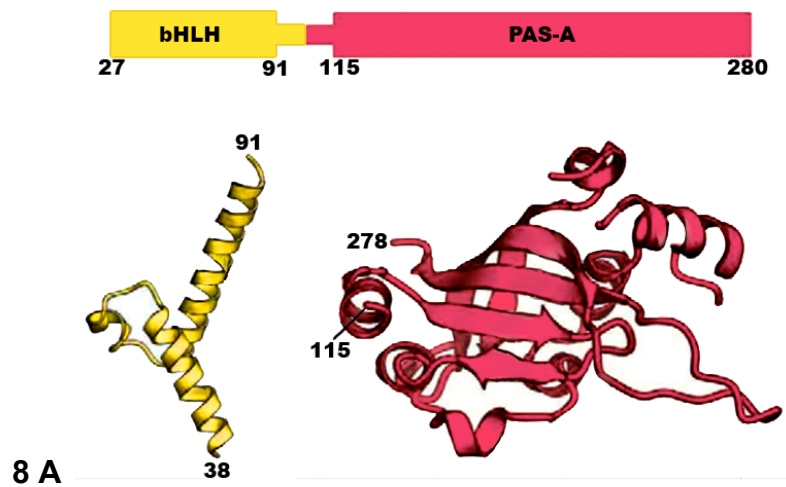
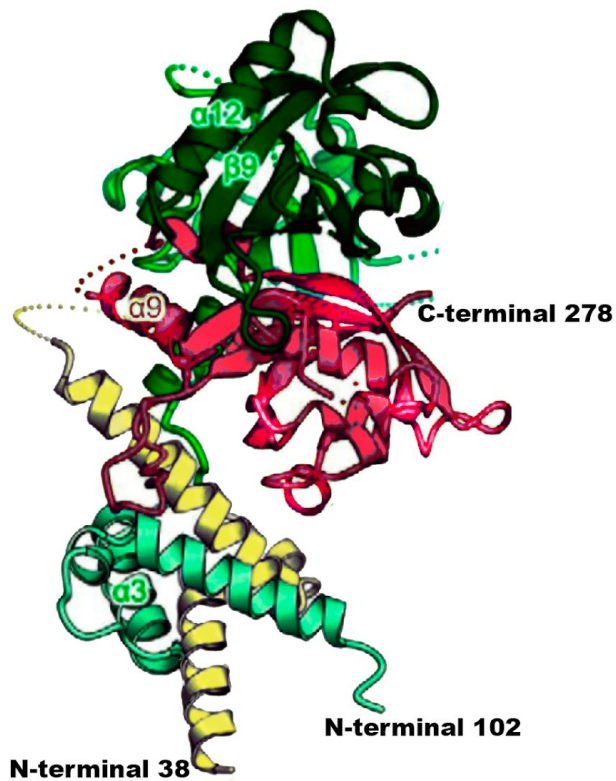
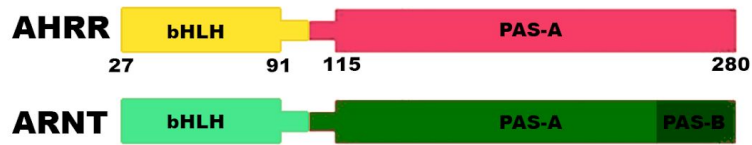


Figura 8. Estrutura 3D da proteína AHRR e respectivos domínios funcionais. a. Estrutura 3D cristalizada dos domínios ativos da proteína AHRR: bHLH e PAS A. A estrutura é composta por dois domínios, representadas pelos retângulos amarelo (bHLH) e rosa (PAS A). b. Estrutura 3D cristalizada dos domínios ativos AHRR e ARNT e sua interação. Estruturas dos genes *AHRR* (retângulos amarelo e rosa) e *ARNT* (retângulos verdes) e abaixo representação 3D da interação entre os domínios, amarelo e verde claro representam bHLH, rosa e verde representam (PAS A) e verde escuro representa PAS B.

Fonte: adaptado de Sakurai et al., 2017





8 B

A proteína AHRR também faz parte de uma rede proteica a qual pode ser visualizada através das ferramentas String (<https://string-db.org/cgi/network.pl?taskId=DzqKAEKuuCsZ> (figura 8). A rede de interação do gene *AHRR* também pode ser visualizada pelo site Pathway Commons (<https://www.pathwaycommons.org/pcviz/#neighborhood/ahrr>) (figura 10).

Figura 9. Rede de interação proteína-proteína do gene *AHRR*. As esferas coloridas indicam proteínas "query" e a primeira linha de interações. As linhas coloridas indicam interações conhecidas determinadas experimentalmente (rosa), ou de bases curadas (azul), interações previstas com genes na mesma vizinhança (verde), fusões de gene (vermelho), co-ocorrência de genes (azul escuro). A linha amarelo esverdeada indica proteínas usualmente mencionadas juntas em textos, a linha preta indica co-expressão e a linha lilás indica homologia entre proteínas.

Fonte: Plataforma String.

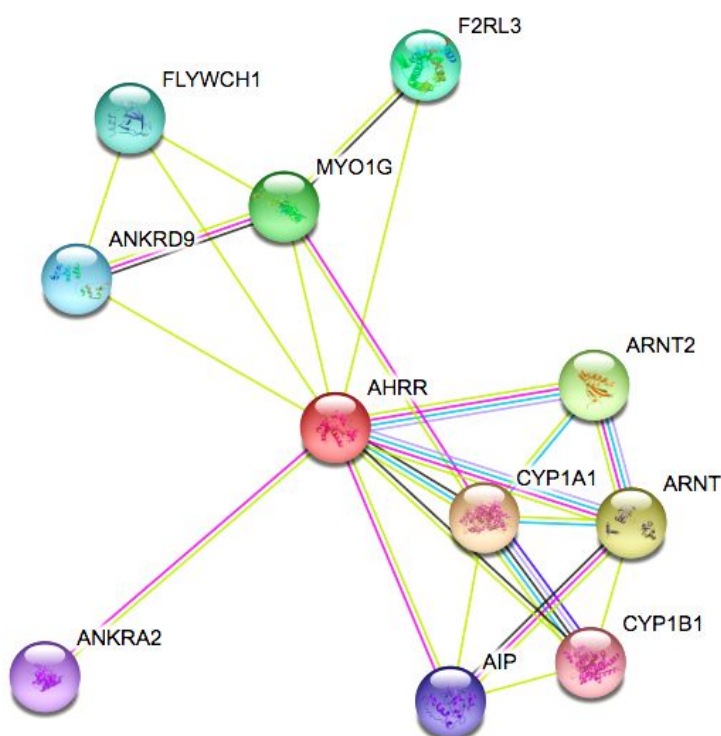
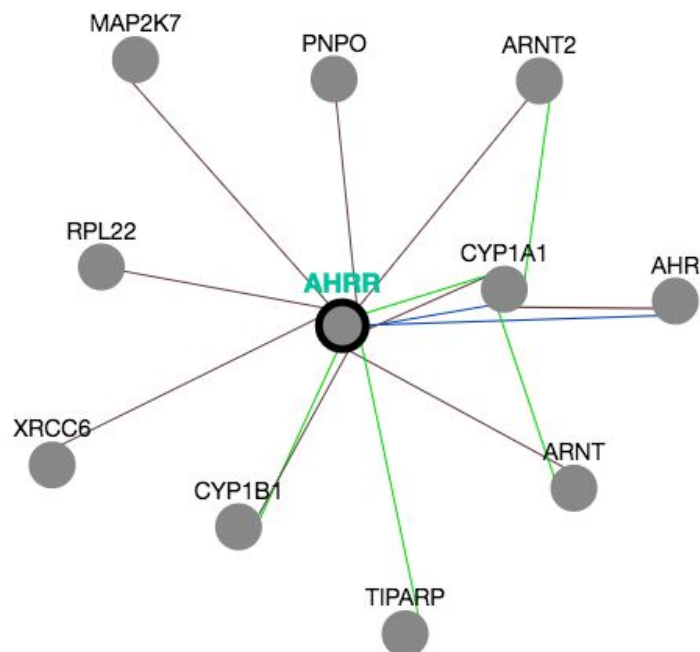


Figura 10. Rede de interação do gene *AHRR* com outros genes Rede de interação gerada pelo software PathwayCommons. A linha verde indica controle de expressão, a linha azul indica controle de mudança de estado, a linha marrom indica estruturas que fazem parte de um mesmo complexo.

Fonte: Pathway Commons, 2019.



5. 2 FUNÇÃO DO GENE *AHRR* E VIA MOLECULAR

O *AHRR* (repressor do receptor de Aril hidrocarbono) atua como mediador da toxicidade a dioxina e também está envolvido na regulação de crescimento e diferenciação celular. Sua ação no metabolismo de xenobióticos ocorre através da regulação do aril hidrocarbono (*AHR*), o qual interage com o complexo citocromo P450 (FUJII-KURIYAMA e KAWAJIRI, 2010). A proteína *AHRR* reprime a atividade de transcrição do receptor *AHR* por competir com este fator de transcrição para formar um heterodímero com o *ARNT*. Este heterodímero *ARNT-AHRR* então liga-se a sequência do elemento de resposta a xenobióticos (*XRE*) em regiões promotoras regulatórias de vários genes e também reprime a ação do *CYP1A1* por ligar-se a sequência *XRE* e recrutar as proteínas *ANKRA2* (Ankyrin Repeat Family A Member 2), *HDAC4* (Histone Deacetylase 4) e *HDAC5* (Histone Deacetylase 5) (FUJII-KURIYAMA e KAWAJIRI, 2010).

A figura 11 demonstra o mecanismo de regulação da transcrição do complexo ativador AHR-ARNT e do complexo repressor AHRR-ARNT. Neste esquema, o complexo formado por um ligante e importinas ligadas a AHR entra no núcleo celular. Dentro do núcleo ocorre a ligação de ARNT não modificado a AHR o que forma um heterodímero. O heterodímero AHR-ARNT recruta os coativadores CBP/p300 a fim de formar o complexo ativador que liga-se ao XRE e estimula a síntese de CYP1A1. Esta ligação também estimula a produção de AHRR o que promove autorregulação desta via uma vez que a proteína AHRR possui uma ação oposta à AHR dentro do núcleo. A AHRR formada também é capaz de formar um heterodímero com ARNT o que aumenta a SUMOilação destas proteínas. A SUMOilação é uma modificação proteica pós traducional que compete com ubiquitinação e consiste na conjugação de um membro da família SUMO (small ubiquitin-like modifier) com resíduos de lisina em proteínas alvo. Tal modificação é um processo reversível e dinâmico que atua na regulação da localização subcelular de proteínas, na ligação entre proteína e DNA, interações proteína-proteína, regulação transcricional, reparo de DNA e organização genômica (YANG et al., 2017). A SUMOilação do heterodímero AHRR-ARNT leva ao recrutamento de correpressores como ANKRA2, HDAC4 e HDAC5 para formar o complexo repressor que impede a síntese de CYP1A1 e consequentemente reprime a ação metabólica do citocromo P450.

Figura 11. Esquema da via do AHRR.

Quando o ligante TCDD atravessa a membrana plasmática via difusão passiva, ele se liga ao AHR via domínio PAS B. AHR geralmente é encontrada no citoplasma associada a proteínas HSP90, p23 e outras na forma de um complexo em estado dormente. O TCDD liga-se ao AHR o que permite a dissociação do complexo TCDD-AHR das demais proteínas e permite a translocação deste complexo para dentro do núcleo celular. A nível nuclear, o complexo TCDD-AHR liga-se ao ARNT e forma um heterodímero que liga-se a sequência de resposta a xenobióticos (XRE). A XRE é uma sequência específica do DNA localizada na região promotora de genes alvo e a ligação

Além deste mecanismo de ação, os principais estudos sobre o gene *AHRR* o relacionam com múltiplos cânceres. Zudaire e colaboradores descobriram que a expressão de *AHRR* foi consistentemente reduzida (*downregulation*) em tecidos malignos humanos de diferentes origens, quando comparados a tecidos normais. Os autores também identificaram hipermetilação como o mecanismo regulador do silenciamento do gene *AHRR*. Ademais, o silenciamento da expressão de *AHRR* em células de câncer pulmonar humanas levou ao aumento do crescimento celular *in vitro* e também em camundongos com imunossupressão pós transplante. A menor expressão do gene *AHRR* em tecido epitelial mamário humano normal permitiu um crescimento mais fácil destas células. Zudaire também demonstrou que *downregulation* da expressão deste gene em linhagens celulares de câncer pulmonar resultou em maior resistência aos sinais de apoptose, maior mobilidade e potencial invasivo *in vitro*, além de maior potencial angiogênico *in vivo*. Expressão ectópica do *AHRR* diminuiu o crescimento celular e a angiogênese em células tumorais. Zudaire concluiu que o gene *AHRR* tem papel essencial como supressor tumoral de múltiplos tumores em humanos (ZUDAIRE et al., 2008). Além disso, Zhang e colaboradores relataram que a expressão excessiva de uma porção do RNA produzido pelo gene *AHRR* em monócitos humanos THP-1 também atua aumentando a expressão de genes inflamatórios, o que culmina no aumento do fator de transcrição da resposta imune NF-kappaB (ZHANG et al., 2017).

Os resultados da revisão serão apresentados em maiores detalhes no artigo que se encontra após as referências da dissertação. É importante destacar que a introdução do artigo corresponde a uma versão simplificada da introdução desta dissertação. Entretanto, a apresentação e discussão dos resultados é apresentada de forma muito mais detalhada, pois corresponde aos objetivos deste projeto. O artigo está escrito no idioma Inglês e foi submetido para publicação no periódico Elsevier – Mutation Research (ISSN: 1383-5742, Fator de Impacto: 6.081).

6. CONCLUSÃO

O estudo de doenças complexas tais como fissuras orofaciais permanece um grande desafio que depende de novos trabalhos que permitam a elucidação da doença e de sua fisiopatologia. A partir da revisão realizada conclui-se que embora o mecanismo por trás da teratogenicidade das dioxinas ao palato não ser totalmente elucidado, o impacto da exposição a elas é evidente conforme os trabalhos reportados. De forma complementar, a todo momento novos genes são descritos e novas funções são descobertas e neste contexto o gene *AHRR* apresenta num contexto pleiotrópico diversos fatores genéticos e moleculares e com uma ampla gama de funções em nosso organismo que pode, conforme hipotetizado neste trabalho, ter relação causal com as fissuras orofaciais.

O estudo do *AHRR* também demonstra o importante papel da interação entre genética e ambiente que é imprescindível na fisiopatologia de doenças multicausais. Apesar de a predisposição genética a certas doenças ser um fator não modificável, outros fatores que são essenciais para a ocorrência de uma doença multifatorial podem ser modificáveis, como a exposição a dioxinas e outros teratógenos. Logo, uma alternativa na prevenção de doenças complexas como FL/P seria atuar na exposição ambiental com medidas de saúde pública. O melhor controle da liberação de dioxinas por indústrias e da exposição humana a estas toxinas, além de prevenção de tabagismo e uso de ácido fólico complementar em mulheres com desejo reprodutivo são exemplos de fatores ambientais modificáveis capazes de causar um impacto a nível genético.

Finalmente, a revisão de literatura permitiu encontrar um conjunto maior de evidências do papel do *AHRR* no desenvolvimento de FL/P de forma que um estudo de associação entre *AHRR* e pacientes com fissuras poderia auxiliar na identificação de variantes causais as quais também poderiam ser utilizadas em estudos funcionais. Como exemplo o estudo de famílias *multiplex* com múltiplos indivíduos afetados poderia permitir a identificação de uma variante nova a qual poderia ser testada com o

auxílio da tecnologia CRISPR-Cas9. Estudos funcionais como este poderiam permitir uma maior compreensão da relação genótipo fenótipo.

7. REFERÊNCIAS

1. BABA T, MIMURA J, GRADIN K, KUROIWA A, WATANABE T, MATSUDA Y, FUJII-KURIYAMA Y. Structure and expression of the Ah receptor repressor gene. **Journal of Biological Chemistry**.v.276, n.35, p.33101-33110, 2001
2. BEATY TH *et al.* A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. **Nat Genet.** v. 42, n. 6, p. 525-9, jun 2010
3. BERK NW, MARAZITA ML. Costs of cleft lip and palate: personal and societal implications. In: Wyszynski DF, editor. **Cleft lip and palate: from origin to treatment**. New York: Oxford University Press; 2002. pp. 458–467.
4. BIRNBAUM S *et al.* Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. **Nature Genetics**. 2009.
5. CALZOLARI E *et al.* Associated anomalies in multi-malformed infants with cleft lip and palate: An epidemiologic study of nearly 6 million births in 23 EUROCAT registries. **Am J Med Genet A**. v. 143A, n. 6, p. 528-37, mar. 2007.
6. CASTILLA EE, LOPEZ-CAMELO JS, PAZ JE. Atlas geográfico de las malformaciones congénitas en Sudamérica. Rio de Janeiro: **Editora Fiocruz**; 1995
7. CHRISTENSEN K, JUEL K, HERSKIND AM, MURRAY JC. Long term follow up study of survival associated with cleft lip and palate at birth. **BMJ** v. 328 n. 7453 jan. 2004. 1405.10.1136/bmj.38106.559120.7C
8. CORREA A *et al.* Diabetes mellitus and birth defects. **Am J Obstet Gynecol.**, v.199, n. 3, p. 237.e1-9. set. 2008.
9. DASKALOGIANNAKIS J *et al.* Unilateral cleft lip with or without cleft palate and handedness: is there an association? **Cleft Palate Craniofac J.**, v. 35, n. 1, p. 46-51, jan. 1998.
10. DE ROO LA *et al.* First-trimester maternal alcohol consumption and the risk of infant oral clefts in Norway: a population-based case-control study. **Am J Epidemiol.**, v. 168, n. 6, p. 638-46, set. 2008.
11. DIXON MJ, MARAZITA ML, BEATY TH, MURRAY JC. Cleft lip and palate synthesizing genetic and environmental influences. **Nat. Rev. Genet.** 2011 Mar; 12(3): 167–178;
12. DUNKHASE E, LUDWIG KU, KNAPP M, ET AL. Nonsyndromic cleft lip with or without cleft palate and cancer: Evaluation of a possible common genetic background through the analysis of GWAS data. **Genomics Data**. 2016;10:22-29.
13. FAGERBERG L, HALLSTRÖM BM, OKSVOLD P, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. **Mol Cell Proteomics**.v.13, n.2, p.397–406, 2014.
14. FUJII-KURIYAMA Y, KAWAJIRI K. Molecular mechanisms of the physiological functions of the aryl hydrocarbon (dioxin) receptor, a multifunctional regulator that senses and responds to environmental stimuli. Proceedings of the Japan Academy Series B, **Physical and Biological Sciences**. 2010;86(1):40-53.

15. HARVILLE EW *et al.* Cleft lip and palate versus cleft lip only: are they distinct defects? **Am J Epidemiol.**, v. 162, n. 5, p. 448-53, set. 2005.
16. HERKRATH AP *et al.* Parental age as a risk factor for non-syndromic oral clefts: a meta-analysis. **J Dent.**, v. 40, n. 1, p. 3-14, jan. 2012.
17. HILL L *et al.* Maternal drug histories and congenital malformations: limb reduction defects and oral clefts. **Journal Epidemiology Comm. Helth.**, p. 1-7, 1988
18. JAMES WH. Are oral clefts a consequence of maternal hormone imbalance? evidence from the sex ratios of sibs of probands. **Teratology**, v. 62, n. 5, p. 342-5, nov. 2000.
19. JONES MC. Etiology of facial clefts: prospective evaluation of 428 patients. **Cleft Palate J.**, v. 25, n. 1, p. 16-20, jan. 1988.
20. JUGESSUR A, FARLIE PG, KILPATRICK N. The genetics of isolated orofacial clefts: from genotypes to subphenotypes. **Oral Diseases**. 15 ed. 2009. 437-453;
21. KOHLI SS e KOHLI VS. A comprehensive review of the genetic basis of cleft lip and palate. **Journal of Oral and Maxillofacial Pathology : JOMFP**, v. 16, n. 1, p. 64-72, 2012. doi:10.4103/0973-029X.92976.
22. KONDO S, *et al.* Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. (Letter) **Nature Genet.** 32: 285-289, 2002;
23. LACE B *et al.* Mutation analysis of the MSX1 gene exons and intron in patients with nonsyndromic cleft lip and palate. **Stomatologija**, v. 8, p. 21-4, 2006.
24. LANDI MT, CHATTERJEE N, YU K, ET AL. A Genome-wide Association Study of Lung Cancer Identifies a Region of Chromosome 5p15 Associated with Risk for Adenocarcinoma. **American Journal of Human Genetics**. 2009;85(5):679-691.
25. LEITE IC, PAUMGARTTEN FJ E KOIFMAN S. Chemical exposure during pregnancy and oral clefts in newborns. **Cad Saude Publica.**, v. 8, n. 1, p. 17-31, jan-fev. 2002.
26. LEITE IC *et al.* Fendas orofaciais no recém nascido e o uso de medicamentos e condições de saúde materna: estudo caso-controle na cidade do Rio de Janeiro. **Rev. Bras. Saúde materna Infantil**, p. 35-43, Mar 2005.
27. LESLIE EJ, MARAZITA ML. Genetics of Cleft Lip and Cleft Palate. **Am. J. Med. Genet. Semin. Med. Genet.** v.163, n.4, p.246-258, Nov 2013.
28. LESLIE EJ *et al.* Association studies of low-frequency coding variants in nonsyndromic cleft lip with or without cleft palate. **Am J Med Genet.** v. 173 n.6 p.1531-1538, Jun 2017
29. LIMA, LS *et al.* Frequency of Cancer in First-Degree Relatives of Patients with Cleft Lip and/or Palate in the Brazilian Population. **Braz. Dent. J.**, Ribeirão Preto, v. 24, n. 3, p. 200-203, 2013.
30. LOFFREDO LC ET AL. Cleft Lip and palate: case control study. **Rev. Saúde Pública**, p. 213-7, 1994.
31. MANGOLD E, LUDWIG KU E NOETHEN M. Breakthroughs in the genetics of orofacial clefting. **Trends Mol. Med.**, v. 17, p. 1-9, 2011.
32. MARAZITA ML. The evolution of human genetic studies of cleft lip and cleft palate. **Annu. Rev. Genomics Hum. Genet.** v. 13, p. 263-283, 2012.

33. MARCUSSEON A, AKERLIND I e PAULIN G. Quality of life in adults with repaired complete cleft lip and palate. **Cleft Palate Craniofac J.**, v. 38, n. 4, p. 379, 2001.
34. MCBRIDE, WG. Thalidomide and congenital abnormalities. **Lancet**, 1961.
35. MONLLEO, Isabella Lopes; GIL-DA-SILVA-LOPES, Vera Lúcia. Anomalias craniofaciais: descrição e avaliação das características gerais da atenção no Sistema Único de Saúde. **Cad. Saúde Pública**, Rio de Janeiro , v. 22, n. 5, p. 913-922, 2006.
36. MOSSEY PA, LITTLE J, MUNGER RG, DIXON MJ, SHAW WC. Cleft lip and palate. **Lancet**. v. 374, n. 9703, p. 1773-85, 2009.
37. MUNGER RG *et al.* Maternal alcohol use and risk of orofacial cleft birth defects. **Teratology**, v. 54, n. 1, p.27-33, jul. 1996.
38. MURRAY JC. Gene/environment causes of cleft lip and/or palate. **Clin Genet.**, v. 61, n. 4, p. 248-56, abr. 2002.
39. OSOEGAWA K *et al.* Identification of novel candidate genes associated with cleft lip and palate using array comparative genomic hybridisation. **J Med Genet.**, v. 45, n. 2, p. 81-6, fev. 2008.
40. PRABHU S, JOSE M, KRISHNAPILLAI R e PRABHU V. Etiopathogenesis of orofacial clefting revisited. **J Oral Maxillofac Pathol.**, v. 16, p. 228-232, 2012
41. POPOFF, Daniela Araújo Veloso *et al.* Non-syndromic oral clefts and risk of cancer: a systematic review. **Dentistry** 3000, [S.l.], v. 1, n. 1, p. 12-18, aug. 2013. ISSN 2167-8677
42. SAKURAI S, SHIMIZUT, OHTO U. The crystal structure of the AhRR-ARNT heterodimer reveals the structural basis of the repression of AhR-mediated transcription. **J Biol Chem**, v.292, n.43, p.17609–17616, 2017.
43. SHI M, WEHBY GL, MURRAY JC. Review on genetic variants and maternal smoking in the etiology of oral clefts and other birth defects. **Birth Defects Res C Embryo Today**, v. 84, n. 1, p. 16-29, mar. 2008
44. SHKOUKANI, MA, CHEN, M e VONG, A. Cleft lip-a comprehensive review. **Frontiers in Pediatrics**. v. 1, p. 53, 2013 doi:10.3389/fped.2013.00053.
45. SIVERTSEN A, WILCOX AJ, SKJAERVEN R, VINDENES HA, ABYHOLM F, HARVILLE E, LIE RT. Familial risk of oral clefts by morphological type and severity: Population based cohort study of first degree relatives. **Br Med J** v.336, p. 432–434, 2008
46. SIMIONI M *et al.* Investigation of genetic factors underlying typical orofacial clefts: mutational screening and copy number variation. **J of Hum Genet**, v. 60, n. 1, p. 17-25, 2015 (epub 2014)
47. SPINA, V *et al.* Classificação das fissuras lábio-palatinas. Sugestão de modificação. **Rev. Hosp. Clin. Fac. Med. São Paulo**, v.27, p.5-6, 1972.
48. STUPPIA L, CAPOGRECO M, MARZO G, LA ROVERE D, ANTONUCCI I, GATTA V, *et al.* Genetics of syndromic and nonsyndromic cleft lip and palate **J Craniofac Surg.**, v. 22, n. 5, p 1722-1726, 2011.
49. SUZUKI S *et al.* Mutations in BMP4 are associated with subepithelial, microform, and overt cleft lip. **Am J Hum Genet.**, v. 84, n. 3, p. 406-11, mar. 2009.

50. TAYLOR BL, ZHULIN IB. PAS Domains: Internal Sensors of Oxygen, Redox Potential, and Light. *Microbiology and Molecular Biology Reviews*. v. 63, n. 2, p.479-506, 1999.
51. WEHBY GL e CASSELL CH. The impact of orofacial clefts on quality of life and healthcare use and costs. **Oral Dis.**, v. 16, n. 1, p.3–1010.1111, 2010. /j.1601-0825.2009.01588.x
52. WEHBY, GL. Advancing and prioritizing research on oral clefts in Brazil. **J Pediatr (Rio J)**. v.89, p. 112–5, 2013.
53. YANG Y, HE Y, WANG X, LIANG Z, HE G, ZHANG P, ZHU H, XU N, LIANG S. Protein SUMOylation modification and its associations with disease. **Open Biol.** v. 7, p. 170167, 2017.
54. ZHANG DD, WANG WT, XIONG J, XIE XM, CUI SS, ZHAO ZG, LI MJ, ZHANG ZQ, HAO DL, ZHAO X, LI YJ, WANG J, CHEN HZ, LV X, LIU DP. Long noncoding RNA LINC00305 promotes inflammation by activating the AHRR-NK-kappa-B pathway in human monocytes. **Sci. Rep.** v. 7, 46204, 2017.
55. ZUDAIRE E, CUESTA N, MURTY V, WOODSON K, ADAMS L, GONZALEZ N, Martinez A, Narayan G, Kirsch I, Franklin W, Hirsch F, Birrer M, Cuttitta F. The aryl hydrocarbon receptor repressor is a putative tumor suppressor gene in multiple human cancers. **J. Clin. Invest.** v.118, p.640-650, 2008.

Referências online disponíveis em:

56. <https://www.omim.org/entry/606517?search=ahrr&highlight=ahrr> Acesso em 2019
57. <http://www.eclamc.org/> Acesso em 2018.
58. https://genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&lastVirtModeType=default&lastVirtModeExtraState=&virtModeType=default&virtMode=0&nonVirtPosition=&position=chr5%3A304176-438291&hgid=666399513_RZ43jc6UM2NhtNztuBRly1vIWfrs Acesso em 2019
59. Pathway Commons Network Visualizer . Disponível em <https://www.pathwaycommons.org/> Acesso em 2019
60. String database. Disponível em <https://string-db.org/> Acesso em 2019
61. https://smart.embl.de/smart/show_motifs.pl?GENOMIC=1&DO_PFAM=DO_PFAM&INCLUDE_SIGNALP=INCLUDE_SIGNALP&ID=9606.ENSP00000323816 Acesso em 2019

8. ARTIGO DERIVADO DO TRABALHO

Running title: A comprehensive analysis of *AHRR* gene as candidate for cleft lip/palate

Bianca Domit Werner Linnenkamp¹, Salmo Raskin², Selene Elifio Esposito^{1,3}; Roberto H. Herai^{1,4,*}

¹ Graduate Program in Health Sciences, School of Medicine, Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba, Paraná, Brazil;

² School of Medicine, Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba, Paraná, Brazil;

³ School of Life Sciences, Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba, Paraná, Brazil;

⁴ Research Department, Lico Kaesemodel Institute (ILK), Curitiba, Paraná, Brazil;

To whom correspondence should be addressed: Bianca Linnenkamp; e-mail: biancalinnenkamp@gmail.com Phone: +55-41-32716341; Address: R. Imaculada Conceição, 1155 - Prado Velho, Graduate Program in Health Sciences, School of Medicine, Pontifícia Universidade Católica do Paraná (PUCPR), 80215-901, Curitiba, Paraná, Brazil.

1. Introduction

Cleft lip with or without cleft palate (CL/P), and cleft palate only (CPO), are among the most common craniofacial anomalies [1,2] and affect approximately 1/700 live births worldwide [3]. Caused by breaks or gaps between orofacial structures [4], clefts affect not only speaking, hearing and appearance of the bearer [5] but also cause negative psychological and socioeconomic repercussions [6,7] which can also lead to a

higher suicide rate [8]. To reduce their impact in the health of affected individuals, it is required a multidisciplinary team of surgeons, speech therapists, physiotherapists, nurses, psychologists, geneticists and other healthcare professionals [2]. It is estimated that five to ten years [9] of treatment with an average cost of two hundred thousand American dollars per patient [10] are necessary to achieve satisfactory results. These lifelong morbidities with extremely elevated costs to national healthcare centers lead to an important public healthcare issue [9,11].

Even though there is a vast variation of phenotypic presentation, orofacial clefting is usually classified as a qualitative disorder, either present or absent, and according to the degree of commitment (microform, cleft lip, cleft lip plus hard palate, cleft lip plus hard and soft palate, cleft palate only) [12]. However, recent literature suggests individuals with cleft are only part of a much larger spectrum that includes not only CL/P and CPO but also multiple subclinical characteristics that could be present in non-affected relatives [13,14]. These characteristics may include craniofacial measurement alterations, dental anomalies (tooth agenesis, microdontia and hyperdontia (extranumerary teeth)), structural brain anomalies, *orbicularis oris* muscle defect, cleft microform, bifid uvula, submucous cleft palate and velopharyngeal insufficiency [14].

Cleft lip is a result from the failure of the union between maxillary prominence on the affected with the merged medial nasal prominences. This improper fusion of the primary palate may cause the palatal shelves to be mispositioned and therefore could interfere with the fusion of the secondary palate. This explains why cleft lip can occur

with or without cleft palate (CL/P) [15]. On the other hand, several pathogenic mechanisms are proposed to be involved in the development of CPO. Palatogenesis depends on several steps which include growth, shelf elevation, contact and fusion [16], therefore the proposed mechanisms leading to CPO include failure of palatal elevation, growth perturbations and failure of fusion of the palatal shelf, all mechanisms that could prevent the mesenchymal masses of the lateral palatal process to meet and fuse with each other and/or with the nasal septum, causing a secondary palate cleft [15,17]. Another mechanism proposed by Kitamura suggests that the palate could also be ruptured after being completely fused what would result in cleft palate as well [18].

Although most clefts are unilateral and affecting the left side twice more often [12,19], a minority of patients has bilateral clefts [12]. While cleft lip is more prevalent among male individuals, cleft palate is more frequent in female individuals [2,20].

Clefts can be classified as syndromic and non-syndromic [21]. Clefts associated to other structural, developmental or cognitive anomalies are defined as syndromic [11], and represent 30% of CL/P and 50% of CPO cases [14,22,23,24,25,26]. Although syndromic forms are caused by a large group of conditions, about 75% of the genetic etiology of syndromic CL/P and CP is known [2,14,27]. On the other hand, non-syndromic clefts are caused by numerous genetic and environmental factors with smaller roles in an intricate mechanism [12,27,28]. Some predisposing environmental factors during pregnancy include maternal smoking [29,30], alcohol use [31,32], advanced maternal age (above 40 years-old) [33], lack of folate supplementation during

the first trimester [34] and use of medications such as anti-inflammatory drugs, antibiotics and anticonvulsants [35,36,37,38].

CL/P and CPO have been considered two distinct conditions with separate etiologies since the first genetic study by Fogh-Andersen due to different embryological origins of the lip and the palate [39,40]. Nonetheless, some families have both CL/P and CPO cases suggesting some etiologic overlap [41], especially in families with etiologic mutations in the genes tumor protein 63 (*TP63*), msh homeobox 1 (*MSX1*), interferon regulatory factor 6 (*IRF6*) and fibroblast growth factor receptor 1 (*FGFR1*) [12].

Many other genes and chromosomal regions have been associated to a greater risk of non-syndromic CL/P, however they represent only part of the genetics behind this condition [42,43,44]. The lack of inheritance pattern or incomplete penetrance found in some case reports could be partially explained due to the phenotypic variability and subclinical manifestations in affected families which would cause an incorrect identification of all affected family members [14]. Better knowledge of the genetic background in CL/P could result in improved prevention, treatment and prognosis for individuals with these conditions [12] but currently genetic markers are still missing, and more candidate genes are necessary to better characterize the condition. While some genes and *loci* are well known such as for the genes *IRF6*, *VAX1*, *FOX1*, *MAFB*, *ABCA4* and 8q24 chromosomal band, other genetic *loci* have insufficient supporting data correlating with non syndromic CL/P [44,45]. Most genetic variations were found by a variety of approaches including GWAS (genome wide association studies), target-based genetic screening and linkage studies.

In another recent study, Leslie and colleagues evaluated the role of low frequency coding variants in orofacial clefts predisposition and pointed *AHRR* as one of the genes related to non-syndromic CL/P [45]. Although this is the first report in literature to directly correlate *AHRR* variants to a higher non-syndromic CL/P risk, the genomic pathway of *AHRR* and its genomic locus have been described in several CL/P genetic studies. This could indicate *AHRR* has perhaps been overlooked as a candidate gene in CL/P [45,46].

In this work, we performed a systematic literature review to show the current findings associating the *AHRR* gene with cleft lip phenotype. We first describe the genetic structure of *AHRR* gene, know splicing isoforms and gene expression levels per tissue. We also show the 3D protein structure of the coded *AHRR* gene, and how it interacts at both gene and protein level with other molecules known to be associated with cleft lip. Finally, we discuss all the findings associating the gene *AHRR* with known cleft lip phenotypes.

2. Methods

The method used to execute this systematic literature review (SLR) was based on the developed protocol by Kitchenham and Charters [47]. The main objective was to identify and synthesize the state of the art vis-à-vis of the *AHRR* gene that have already been associated with cleft palate characteristics, including molecular, cellular and clinical phenotypes. In the following sections, the steps considered by this SLR will be described in detail.

Research questions

The development of this SLR was primarily focused on the research question: What is the correlation between *AHRR* gene and cleft lip/palate? This question was then subdivided into the following research questions: What is the correlation between *AHRR* gene with cleft lip/palate molecular phenotypes? What are the genetic variants in *AHRR* associated to cleft lip/palate.

Exclusion and Inclusion Criteria

The following inclusion criteria were defined: i) studies published until May 2019; ii) were considered only articles in the English language; iii) articles that consider cleft lip/palate and *AHRR* and *AHRR* related genes iv) articles that present genetic variants in found in cleft lip/palate. The exclusion criteria were: i) articles that used *AHRR*, *AHHR* and *AHH* initials for a different purpose.

Search databases

PubMed was used as database for an electronic search. PubMed is a search-free resource, developed and maintained by the National Center for Biotechnology Information (NCBI), at the U.S. National Library of Medicine (NLM), located at the National Institute of Health (NIH). PubMed covers portions of life sciences, behavioral sciences, chemical sciences and bioengineering through more

than 25 million citations and summaries from the fields of biomedicine and health (<https://www.ncbi.nlm.nih.gov/pubmed/>).

Search Strategies

To obtain as many articles as possible, we sought to use the terms and their synonyms for: cleft lip, cleft palate, orofacial cleft, *AHRR*, aryl-hydrocarbon receptor repressor. Subsequently, we obtained keywords that composed the following search string: (((((((((((((((cleft lip[Title/Abstract]) OR cleft palate[Title/Abstract]) OR orofacial cleft[Title/Abstract]) AND ahrr[Title/Abstract]) OR Aryl Hydrocarbon Receptor Repressor[Title/Abstract]) OR Aryl-Hydrocarbon Receptor Repressor[Title/Abstract]) OR Class E Basic Helix-Loop-Helix Protein 77[Title/Abstract]) OR AhR Repressor[Title/Abstract]) OR BHLHe77[Title/Abstract]) OR Aryl Hydrocarbon Hydroxylase Regulator[Title/Abstract]) OR Aryl Hydrocarbon Receptor Regulator[Title/Abstract]) OR Dioxin Receptor Repressor[Title/Abstract]) OR KIAA1234[Title/Abstract]) OR BHLHe77[Title/Abstract]) OR AHHR[Title/Abstract]) OR AHH[Title/Abstract]. Although BHLHe77 and KIAA1234 are known aliases for AHRR, these terms were not found in PubMed. The composition of the keywords resulted in 1325 combinations of terms used for research. Firstly, a triage was carried out by title, followed by analysis of abstracts, resulting in respectively 29 and 1007 articles eliminated due to initials AHH having meanings other than Aryl Hydrocarbon Hydroxylase. The following meanings were found for AHH: acquired hypocalciuric hypercalcemia, "ahh" referring to the phonation of ahh during palate examination,

acquired hypogonadotropic hypogonadism, AHH-1 cells, atrial-his-atrial, acyl heck-heck, "ahh-choo". 1 tints - ahh sound, acute hypobaric hypoxia, American Heart Association, aliphatic halogenated hydrocarbons, arecanut husk hydrochar, ahh family, approved by the Danish Data Protection Agency (approval number: AHH-2017-048), axillary hyperhidrosis, apex of the hook of the hamate, acute hypervolemic hemodilution, chimeric peptide Ala-His-His, adult-onset hypogonadotropic hypogonadism, amaranth hydrolysate, acute hemophilic hemarthrosis, two-hop adaptive hop by hop vector-based forwarding (2hop-AHH-VBF), A-Helping-Hand, adult human heart, benzo(a)pyrene hydroxylase, *A. hydaspica* [n-hexane], acute hypoxic hepatitis, high-dose AK in alcohol/HFD group, automated hotspot-hunting procedure, acute hypobaric hypoxia, high altitude hypoxia, acquired hypocalciuric hypercalcemia. This resulted in 289 articles containing information about *AHRR* gene and its pathway which were considered potentially relevant to the research. Two independent readers analyzed all potentially relevant studies which resulted in the inclusion of 37 articles (Figure 1). In all cases, the inclusion and exclusion criteria previously presented were applied.

3. Results and Discussion

To investigate whether the gene *AHRR* has correlation with the known phenotypes presented by CL/P, we performed a genetic and proteomic characterization of this gene and its coded protein. We then put together this information to create an interaction network of genes and proteins to show their interacting molecules, and whether they have previously been associated with CL/P. Finally, we performed a

literature-based investigation to describe the current findings of the *AHRR* gene and how it is associated with CL/P, at molecular and cellular levels, including involved environmental factors. These data are detailed in the next sections.

3.1 Genetic and protein characteristics of *AHRR* gene

AHRR gene is located on the short arm of chromosome 5 [5p15.3; genomic coordinates (GRCh38/hg38) chr5:304,176-438,291] [48] and is 134.116 bases long (Genecards, 2019). According to RefSeq, *AHRR* has two reviewed transcript isoforms, NM_020731.4 (5676 bp, 12 exons and 11 introns) and NM_001242412.1 (5622 bp, 11 exons and 10 introns) [49] (Figure 2).

According to Uniprot protein data bank, there are three known protein isoforms, the first has been chosen as canonical and has 701 amino acids (A9YEQ3-1). The second isoform (A9YEQ3-2) differs from the canonical at position 240 when new amino acids (LARGSQAWQLRLCCPEPLM) are added making this sequence 719 amino acids long. The third isoform (A9YEQ3-3) has the same amino acid insertion plus a deletion of the first four amino acids which changes its length to 715 [50]). Seven others isoforms exist, and have been predicted by computationally detection (D6RE68 (557 amino acids), D6RF73 (547 amino acids), E5RFG4 (157 amino acids), E5RGQ2 (81 amino acids), G3V143 (139 amino acids), H0YA43 (215 amino acids) and E5RHE0 (12 amino acids)) [50].

Furthermore, even though there are different splicing isoforms, Fagerberg tested *AHRR* gene expression in the human body and that it is expressed in most of human

tissues. In an investigation through RNA-sequencing of 27 distinct body tissue samples, it was found expression of the gene *AHRR* in all analyzed samples. Testicular tissue presented the highest expression of *AHRR* followed by spleen, skin, placenta and ovaries respectively [51] (Figure 3).

AHRR protein domains

AHRR codes for a protein also known as AHRR. There are two known AHRR domains, the bHLH and PAS (Figure 4 and 5). The domain bHLH (Myc-type, basic helix-loop-helix domain profile) starts at amino acid position 35 until 89 and acts in protein dimerization enabling the non-covalent union of two identical or non-identical subunits creating a macromolecule. PAS domain [PAS repeat profile (Per-period circadian protein, Arnt-Ah receptor nuclear translocator protein and Sim-single-minded protein)] starts at position 114 up until 180 and interacts with signaling proteins as a responder to light, oxidation and other stimuli [52]. The two domains are also present in AHRR related proteins Aryl Hydrocarbon Receptor (AHR) and AHR Nuclear Translocator also called HIF-1 β (ARNT).

Besides these two domains, AHR and ARNT proteins contain an additional PAS domain and a transactivation bHLH domain (Figure 4). The bHLH domain in AHR allows its dimerization with ARNT, binding of DNA and interaction with chaperones like HsP90 (Heat Shock Protein 90). It is important for nuclear importation and exportation as well [53]. The PAS domain with two repeats involved with ARNT dimerization (PAS-A) and allows ligand binding (PAS-B). There is an extra C-terminal domain with three

subdomains, one enriched with acidic residues, the second Q-rich (glutamine enriched) and the third serine, threonine and proline enriched (S/T/P). This domain is responsible for interactions with coactivators and corepressors [54,55]. The PAS-B domain is absent in ARNT and AHRR therefore both are unable to bind ligands [56].

3.2 Protein network, interaction and involved pathways of AHRR

AHRR protein acts as dioxin toxicity mediator and is also involved in cellular growth regulation and differentiation. Most studies correlate the *AHRR* gene to different types of cancer. Zudaire reported that *AHRR* expression is consistently downregulated in human malignant tissues mainly due to hypermethylation, when compared to healthy tissue from different sites [57]. This silencing of *AHRR* in human lung cancer cells lead to a rise in cellular growth both *in vitro* and in animal models. A similar effect was observed in healthy human mammary epithelial cells. Furthermore, the downregulation of *AHRR* in lung cancer cells has also resulted in an increase in cell mobility, apoptosis resistance, *in vivo* angiogenesis and *in vitro* invasive potential. On the other hand, ectopic expression of *AHRR* has decreased cellular growth rates and angiogenesis in tumor cells which indicates the essential part *AHRR* has as a tumor suppressor in humans [57]. Another study by Zhang has reported that overexpression of the RNA produced by AHRR in human monocytes also enhances inflammatory genes expression leading to higher levels of immune response transcription factor NF-kappaB [58]

In addition, *AHRR* has a role in xenobiotic metabolism that occurs through the regulation of *AHR* gene, which interacts with p450A1 cytochrome complex in the

genomic *AHR* pathway [59]. The *AHR* pathway begins when a complex formed by ligand [TCDD(2,3,7,8-tetrachlorodibenzo-p-dioxin)], chaperone proteins (HSP90, XAP2 and p23) and importins associated to the protein AHR enters the cellular nucleus. Inside the nuclear structure, AHR binds to ARNT and forms a heterodimer. This AHR-ARNT dimer recruits coactivators CBP/p300 in order to create an activated complex that binds to XRE (xenobiotic response element) and stimulates the synthesis of proteins CYP1A1 (Cytochrome P450 (CYP) 1A1) and AHRR [59] (Figure 6). CYP1A1 is one of the most important enzymes involved in cytochrome p450 drug metabolism and synthesis of cholesterol, steroids and other lipids. This CYP1A1 mediated metabolism pathway can generate reactive metabolites and lead to mutagenesis and tumor formation [60].

AHRR protein production autoregulates this pathway since AHRR can also form a heterodimer with ARNT and therefore competes with AHR repressing its transcription. This process leads to protein SUMOylation and recruits other co-repressors like proteins ANKRA2 (Ankyrin Repeat Family A Member 2), HDAC4 (Histone Deacetylase 4) and HDAC5 (Histone Deacetylase 5) to form a repressor complex that prevents CYP1A1 synthesis and the metabolic action of P450 cytochrome [59,61] (Figure 6).

According to String database, AHRR protein structure has recently been described and its interaction network includes ARNT (Aryl Hydrocarbon Receptor Nuclear Translocator), ARNT2 (Aryl Hydrocarbon Receptor Nuclear Translocator 2), CYP1A1 (Cytochrome P450 family 1 subfamily A member 1), MYO1G (Myosin IG), ANKRD9 (Ankyrin Repeat Domain 9) , FLYWCH1 (FLYWCH-Type Zinc Finger 1), F2RL3 (F2R Like Thrombin Or Trypsin Receptor 3), ANKRA2 , AIP (Aryl Hydrocarbon

Receptor Interacting Protein) and CYP1B1 (Cytochrome P450 Family 1 Subfamily B Member 1) (Figure 7) [62].

Part of these protein-protein interactions of AHRR are also confirmed by another independent database of genetic interactions for the gene *AHRR*, such as for the genes *CYP1A1*, *ARNT*, *ARNT2* and *CYP1B1* (Figure 8).

According to the protein interacting network, the proteins ARNT, ARNT2, RPL22 (Ribosomal Protein L22), XRCC6 (X-Ray Repair Cross Complementing 6), PNPO (Pyridoxamine 5'-Phosphate Oxidase), CYP1A1, CYP1B1 and MAP2K7 (Mitogen-Activated Protein Kinase Kinase 7) are members of the same complex as AHRR. ARNT, ARNT2 and AHRR are also known to regulate *CYP1A1* expression. and AHRR is able to change the state of AHR and CYP1A1. Another important protein that might also modulate the expression of these proteins is TIPARP [TCDD Inducible Poly(ADP-Ribose) Polymerase] through the regulation of the expression of *AHRR* [63].

3.3 AHRR and cleft lip/palate

Due to the genetic heterogeneity of non-syndromic CL/P, multiple genetic approaches have been applied to study either the whole genome or specific candidate genes. Some methods include linkage analysis with inbred or multiplex families, association studies with case-parent trios, identification of chromosomal anomalies, identification of microdeletions and CNVs (copy number variations) and direct DNA sequencing of probands [12].

In 2018, a study with probands from Brazil detected and compared different CNVs and identified multiple genomic *loci* related to non-syndromic CL/P. The study described a duplication at the chromosomal band 5p15.33 that could predispose to a higher clefting risk [46]. In 2017, a study by Leslie and colleagues tested 1995 unrelated cases with non-syndromic CL/P and 1626 unrelated controls selected from a larger set of cases, controls, and families recruited from 11 countries across North America and Latin America, Asia, and Europe. Cases included probands from unrelated families and unrelated individuals recruited as sporadic cases. Controls had no known history of orofacial clefts and other craniofacial anomalies. This study analyzed low-frequency variants in a multiethnic population and then stratified them into three ancestry groups: Europeans, Latin Americans, and Asians. In the European population group, six variants were reported in *AHRR*: rs111451538, rs200229746, rs111768223, rs35756515, rs61757546 and rs200762873 (Table 1). This was the first study to associate *AHRR* to an increased risk of non-syndromic CL/P even though *AHRR* did not seem to have supporting evidences connecting this gene with a functional role in craniofacial development [45].

Nonetheless, although the gene *AHRR* has not been previously directly connected to cleft lip/palate, the genomic pathway regulated by *AHRR* (TCDD-AHR pathway) [64]. Exogenous ligands like TCDD activating AHR pathway can result in deregulation of tissue morphogenesis and insufficient differentiation of epithelial cells causing not only cleft palate but also other malformations in TCDD-exposed mice and zebrafish [65]. Furthermore, several known targets of AHR activation are important in

the regulation of tissue and extracellular matrix (ECM) remodeling, which indicates that the AHR-pathway has an endogenous role in matrix metabolism and deposition. Besides the effects of AHR pathway on ECM remodeling during the development, it is also significant for processes that require extensive remodeling such as wound repair and regeneration. In zebrafish, it was demonstrated that expression and activity of proteases and ECM remodeling are required for fully regeneration of the tail fins after amputation [65,66].

In another work, it was also reported the TCDD administration to pregnant hamsters, mice and rats resulted in cleft palate in fetuses but did not result in cleft lip phenotypes [67]. The susceptibility peak period for the onset of cleft palate is around nine and twelve gestational days in mice. According to Couture, TCDD administration to pregnant mice resulted in 93% and 100% incidence when given on gestational day 8 and 12. However, the incidence dropped to 40% when TCDD administration was anticipated to the sixth gestational day and achieved 0% when postponed to gestational day 14. Interestingly, when TCDD exposure dose doubled in the early period, incidence reached 100% which indicated high doses may cancel the apparent window of susceptibility in the earlier period. On the other hand, the low incidence found in later periods did not depend on dosage and could be more related to palatogenesis completion timing [67,68,69].

Also, in the normal development in mice, palatal shelves elevation occurs from gestational day 14 to 14,25 and complete fusion ends by day 14,5. In TCDD exposed mice not only is this process delayed by one day but also the fusion is less efficient

inducing cleft palate on days 17-18. Even in successfully fused palate shelves, TCDD exposure in gestational day 15 lead to cleft in 17,5% of the fetuses which suggests TCDD can disturb palatogenesis and induce post-fusional rupture of the palate [67,70]. According to Yamada, the best explanation of TCDD-induced cleft palate is medial edge epithelium disruption and fusional failure. Nevertheless, an additional clefting mechanism caused by post-fusional split is suggested in some surviving palates [71].

Also, palate shelf fusion initiates through periderm cells adhesion and medial edge epithelial palate shelf makes initial contact and adheres to the contralateral shelf. Afterwards, a medial epithelium seam is formed of periderms and MEE cells. This seam degenerates by cell death and the dying cells activate basal lamina degradation leading to mesenchymal confluence of the palate and completion of palatogenesis. According to Cuervo and Covarrubias, although the exact mechanism in which TCDD leads to cleft palate is unknown, one leading hypothesis is that TCDD impairs the growth of medial edge epithelial cells in the palatal shelves [72,73,74].

In another study by Gao the effect of TCDD in cells isolated from human palate shelves was tested and revealed TCDD promoted epithelial cell growth without effect on apoptosis. His results indicated cleft palate in TCDD-exposed embryos may be due to accelerated cell proliferation [75]. Other studies have demonstrated cell-cycle-dependent effects of TCDD on gene expression and also AHR-dependent and AHR-independent effects of TCDD on cell-cycle progression [76,77]. Gao also, demonstrated that the proliferation of human palate shelves cells promoted by TCDD was mediated through an *AHR*-dependent activation of *PI3K/ Akt* pathway and its

upstream components, like *PDK1* and *PTEN*. The author suggests the changes in epithelium proliferation caused by TCDD may be responsible for the underlying mechanism by which TCDD induces palatal fusion failure since cell degradation is an important piece in palatogenesis [75]

Furthermore, when *AHR* is suppressed, TCDD exposure did not result in cleft palate, which indicates the important role of *AHR* in this pathologic process [66] and corroborates that most of the toxic effects of TCDD are mediated via *AHR* [78].

Many other efforts were made by independent institutions that made possible to reduce concentrations levels of dioxins like TCDD, since they are present at varying levels of toxicity in our environment. Still, dioxins levels should be monitored because even a low level exposure can lead to a decrease in the threshold for multifactorial pathogens of CPO [71]. In addition, in 2014 Yamada demonstrated that the mouse palatal bone is rich in *AHR* and that could enhance its sensitivity to TCDD. His study showed an increase in *AHR* expression in muscle tissues, such as the tongue and masticatory muscles after the administration of TCDD suggesting a susceptibility to TCDD in osteogenesis and myogenesis [71].

Besides studies using animal models, several other investigations were performed in humans that have also identified the *AHR* gene as a causative factor of CL/P. A whole exome sequencing (WES) study in the Honduran population has correlated *AHR* to CL/P. The study used WES to identify CNVs in non-syndromic CL/P multiplex families, which are families with two or more affected members. One of the analyzed families had two affected brothers, their unaffected mother, and unaffected

grandfather. One of the brothers had complete bilateral cleft lip and palate while the other had complete unilateral cleft lip without cleft palate. A deletion of 13.3 kb and 23.6 kb in the region of AHR was found in the affected brothers. The CNV in AHR was identified in two affected siblings as well as their unaffected mother and maternal grandfather, consistent with an autosomal dominant mode of inheritance with incomplete penetrance [79].

Moreover, recent data has suggested both CL/P patients and their first and second degree relatives had an increased risk of developing malignant tumors [80,81]. In 2016, Dunkhase reported an overlap among genetic factors predisposing to clefts and cancer. His study identified multiple loci that were related to both pathologies proposing the genes in these regions may have a pleiotropic activity. Among the described regions was 5p15.33, associated to non-syndromic CL/P and lung cancer [82,83].

Furthermore, recent epigenetic studies have shown additional data supporting maternal tobacco exposure and lack of folate supplementation as risk factors for cleft palate. Prenatal maternal smoking and passive exposure to tobacco are known risk factors for not only CL/P [12] but also preterm delivery, intrauterine growth restriction, low birth weight and potentially lifelong negative health consequences [84,85]. Cigarette smoke has numerous chemicals capable of penetrating the placenta such as polycyclic aromatic hydrocarbons which bind to AHR and activate AHR pathway [85,86]. Exposure to maternal smoking during pregnancy is associated with an increase in *AHRR* mRNA expression in cord blood mononuclear cells and is also associated with hypomethylation

in the *AHRR* gene in umbilical cord blood, peripheral blood and buccal cells. The well described association between smoking and *AHRR* hypomethylation allowed this to become a biomarker for current and former smoking status [86,87,88]. A study by Fa tested if exposure to maternal smoking in the first trimester changed DNA methylation and mRNA expression of *CYP1A1* and *AHRR*. This study detected two CpG sites for *AHRR* with an increase in methylation in placentas of female fetuses. Further, higher exposure-associated mRNA expression was observed in first trimester fetal livers. Therefore, *AHRR* expression may be increased in fetal tissues, leading to *CYP1A1* repression, as a response to sustained maternal smoking and the long-term expression change can manifest as hypomethylation of specific CpGs within *AHRR* [88].

In 2018, Zhang published a comparison between methylated areas in smoking and non-smoking african-american mothers and their newborns. The study confirmed epigenome-wide associations between maternal smoking and cord blood DNA methylation (DNAm) at multiple sites, including *AHRR* gene and suggested that these DNAm changes related to smoking are present in whites and African Americans. This study also demonstrated that the adverse effect of maternal smoking on the newborn may vary according to sex and gene suggesting the male fetus may be more vulnerable to the impact of maternal smoking on DNAm than the female. Zhang also indicates some of these effects in DNAm can be partially mitigated by sufficient maternal folate levels however he does not confirm that sufficient maternal folate levels can counteract the adverse effects of maternal smoking on child health [89].

In conclusion, *AHRR* has an important role in the response to exogenous toxins and therefore plays a significant piece in the gene-environment interaction leading to several multifactorial diseases and could also be influencing the phenotypes presented by the non-syndromic CL/P. Although additional data and investigations are required to conclude that *AHRR* has a pathogenic effect on this disease, several investigations could be proposed to support the genotype-phenotype connection. Perhaps the next step should be the identification of novel *AHRR* variants in CL/P patients and multiplex families followed by functional studies of these variants, including CRISPR/cas9 based experiments and gain-loss of function gene expression modulation which could allow a better genotype-phenotype understanding and guide other future studies.

References

- [1] Jugessur A, Farlie PG, Kilpatrick N. The genetics of isolated orofacial clefts: from genotypes to subphenotypes. *Oral Diseases*.2009;15:437-453.
- [2] Mossey PA, Little J, Munger RG, Dixon MJ, Shaw WC. Cleft lip and palate. *Lancet*.2009;374(9703):1773-85.
- [3] Mossey PA, Modell B. Epidemiology of oral clefts 2012: an international perspective. *Cleft lip and palate*.2012;16:1-18.
- [4] Simioni M et al. Investigation of genetic factors underlying typical orofacial clefts: mutational screening and copy number variation. *J of Hum Genet*.2015;60(1):17-25.
- [5] Prabhu S, Jose M, Krishnapillai R and Prabhu V. Etiopathogenesis of orofacial clefting revisited. *J Oral Maxillofac Pathol*.2012;16:228-232.

- [6] Wehby GL e Cassell CH. The impact of orofacial clefts on quality of life and healthcare use and costs. *Oral Dis.*2010;16(1):3–1010.1111.
- [7] Marcusson A, Akerlind I e Paulin G. Quality of life in adults with repaired complete cleft lip and palate. *Cleft Palate Craniofac J.*2001;38(4):379.
- [8] Christensen K, Juel K, Herskind AM, Murray JC. Long term follow up study of survival associated with cleft lip and palate at birth. *BMJ.* 2004;328(7453)1405.
- [9] Lace B et al. Mutation analysis of the *MSX1* gene exons and intron in patients with nonsyndromic cleft lip and palate. *Stomatologija.*2006;8:21–4.
- [10] Berk NW, Marazita ML. Costs of cleft lip and palate: personal and societal implications. In: Wyszynski DF, editor. Cleft lip and palate: from origin to treatment. *New York: Oxford University Press*; 2002.458–467.
- [11] Basha M, Demeer B, Revencu N, Helaers R, Theys S, Saba SB, Vikkula M. Whole exome sequencing identifies mutations in 10% of patients with familial non-syndromic cleft lip and/or palate in genes mutated in well-known syndromes. *Journal of medical genetics.* 2018;55(7):449-458.
- [12] Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate synthesizing genetic and environmental influences. *Nat. Rev. Genet.* 2011;12(3):167–178.
- [13] Weinberg SM., Naidoo SD, Bardi KM, Brandon CA, Neiswanger K, Resick JM, Marazita ML. Face shape of unaffected parents with cleft affected offspring: combining three-dimensional surface imaging and geometric morphometrics. *Orthodontics & craniofacial research.* 2009;12(4):271-281.

- [14] Leslie EJ, Marazita ML. Genetics of Cleft Lip and Cleft Palate. *Am. J. Med. Genet. Semin. Med. Genet.* 2013;163(4):246–258.
- [15] Shi M. Comprehensive gene environment analysis of the causes of orofacial clefts The University of Iowa, *ProQuest Dissertations Publishing*, 2005. 3172444.
- [16] Pratt R.M. Mechanisms of chemically-induced cleft palate. *Trends Pharmacol. Sci.* 1983;4:160–162.
- [17] Ferguson MW. Palate development. *Development.* 1988;103:41–60.
- [18] Kitamura, H. (1991). Evidence for Cleft Palate as a Postfusion Phenomenon. *The Cleft Palate-Craniofacial Journal.*28(2):195–211.
- [19] Daskalogiannakis J et al. Unilateral cleft lip with or without cleft palate and handedness: is there an association? *Cleft Palate Craniofac J.*1998;35(1)46-51.
- [20] James WH. Are oral clefts a consequence of maternal hormone imbalance? evidence from the sex ratios of sibs of probands. *Teratology.* 2000;62(5):342-5.
- [21] Kohli SS and Kohli VS. A comprehensive review of the genetic basis of cleft lip and palate. *Journal of Oral and Maxillofacial Pathology: JOMFP.*2012;16(1):64-72.
- [22] Jones MC. Etiology of facial clefts: prospective evaluation of 428 patients. *Cleft Palate J.* 1988;25(1):16-20.
- [23] Calzolari E et al. Associated anomalies in multi-malformed infants with cleft lip and palate: An epidemiologic study of nearly 6 million births in 23 EUROCAT registries. *Am J Med Genet A.*2007;143A(6):528-37.
- [24] Setó-Salvia N, Stanier P. Genetics of cleft lip and/or cleft palate: Association with other common anomalies.*European Journal of Medical Genetics.*2014;57(8):381-393.

- [25] Wyszynski DF, Sárközi A and Czeizel AE. Oral clefts with associated anomalies: methodological issues. *The Cleft palate-craniofacial journal*.2006;43(1):1-6.
- [26] Shkoukani, MA, Chen, M and Vong A. Cleft lip-a comprehensive review. *Frontiers in Pediatrics*.2013;1:53.
- [27] Stuppia L, Capogreco M, Marzo G, La Rovere D, Antonucci I, Gatta V, et al. Genetics of syndromic and nonsyndromic cleft lip and palate. *J Craniofac Surg*.2011;22(5):1722-1726.
- [28] Murray JC. Gene/environment causes of cleft lip and/or palate. *Clin Genet*.2002;61(4):248-56.
- [29] Shi M, Wehby GL, Murray JC. Review on genetic variants and maternal smoking in the etiology of oral clefts and other birth defects. *Birth Defects Res C Embryo Today*.2008;84(1):16-29.
- [30] Wyszynski DF and Wu T.Prenatal and perinatal factors associated with isolated oral clefting. *The Cleft palate-craniofacial journal*.2002;39(3):370-375.
- [31] De Roo LA et al. First-trimester maternal alcohol consumption and the risk of infant oral clefts in Norway: a population-based case-control study. *Am J Epidemiol*.2008; 168(6)638-46.
- [32] Munger RG et al. Maternal alcohol use and risk of orofacial cleft birth defects. *Teratology*.1996;54(1):27-33.
- [33] Herkrath AP et al. Parental age as a risk factor for non-syndromic oral clefts: a meta-analysis. *J Dent*.2012;40(1):3-14.

- [34] Harville EW, Wilcox AJ, Lie RT, Åbyholm F, and Vindenes H. Epidemiology of cleft palate alone and cleft palate with accompanying defects. *European journal of epidemiology*. 2007;22(6):389-395.
- [35] Leite IC et al. Fendas orofaciais no recém nascido e o uso de medicamentos e condições de saúde materna: estudo caso-controle na cidade do Rio de Janeiro. *Rev. Bras. Saúde materna Infantil*.2005:35-43.
- [36] Hill L et al. Maternal drug histories and congenital malformations: limb reduction defects and oralclefts. *Journal Epidemiology Comm. Helth*.1998;1-7.
- [37] McBride WG. Thalidomide and congenital abnormalities. *Lancet*, 1961.
- [38] Loffredo LC et al. Cleft Lip and palate: case control study. *Rev. Saúde Pública*. 1994:213-7.
- [39] Fogh-Andersen, Poul. Inheritance of Harelip and Cleft Palate *Nyt Nordisk Forlag. Copenhagen: Arnold Blstk*. 1942.
- [40] Fraser FC. Thoughts on the etiology of clefts of the palate and lip. *Acta Genet Stat Med*.1955;5(4):358–369.
- [41] Kondo S, et al. Mutations in *IRF6* cause Van der Woude and popliteal pterygium syndromes. (*Letter*) *Nature Genet*.2002;32:285-289.
- [42] Mangold E, Ludwig KU, Noethen M.Breakthroughs in the genetics of orofacial clefting. *Trends Mol. Med*.2011;17:1–9.
- [43] Marazita ML. The evolution of human genetic studies of cleft lip and cleft palate. *Annu. Rev. Genomics Hum. Genet*.2012;13:263–283.

- [44] Shaffer JR, LeClair J, Carlson JC, et al. Association of low-frequency genetic variants in regulatory regions with nonsyndromic orofacial clefts. *Am J Med Genet Part A*. 2019;179A: 467– 474.
- [45] Leslie EJ et al. Association studies of low-frequency coding variants in nonsyndromic cleft lip with or without cleft palate. *Am J Med Genet*. 2017;173(6):1531-1538.
- [46] da Silva HPV, Oliveira GHDM, Ururahy MAG, Bezerra JF, de Souza KSC, Bortolin RH, Brito MEF. Application of high-resolution array platform for genome-wide copy number variation analysis in patients with nonsyndromic cleft lip and palate. *Journal of clinical laboratory analysis*. 2018;32(6):e22428.
- [47] Kitchenham B, Charters S. Guidelines for performing Systematic Literature Reviews in Software Engineering. Version 2.3, Keele University and University of Durham, EBSE Technical Report.2007
- [48] Baba T, Mimura J, Gradin K, Kuroiwa A, Watanabe T, MatsudaY and Fujii-Kuriyama Y. Structure and expression of the Ah receptor repressor gene. *Journal of Biological Chemistry*.2001;276(35):33101-33110.
- [49]NCBI's reference sequence (RefSeq) database
(<http://www.ncbi.nlm.nih.gov/RefSeq/>)
- [50] Uniprot Online Database (<https://www.uniprot.org/uniprot/A9YTQ3#A9YTQ3-2>).
- [51] Fagerberg L, Hallström BM, Oksvold P, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics*. 2014;13(2):397–406.

- [52] Taylor BL, Zhulin IB. PAS Domains: Internal Sensors of Oxygen, Redox Potential, and Light. *Microbiology and Molecular Biology Reviews*. 1999;63(2):479-506.
- [53] Whitelaw ML, McGuire J, Picard D, Gustafsson JA, Poellinger L. Heat shock protein hsp90 regulates dioxin receptor function in vivo. *Proc Natl Acad Sci U S A*. 1995;92(10):4437–4441.
- [54] Jain S., Dolwick K.M., Schmidt J.V., Bradfield C.A. Potent transactivation domains of the Ah receptor and the Ah receptor nuclear translocator map to their carboxyl termini. *J. Biol. Chem*. 1994;269:31518–31524.
- [55] Ko HP, Okino ST, Ma Q, and Whitlock JP. Transactivation domains facilitate promoter occupancy for the dioxin-inducible CYP1A1 gene in vivo. *Molecular and cellular biology*. 1997;17(7):3497-350.
- [56] Haarmann-Stemmann T, Esser C., Krutmann J. The Janus-faced role of aryl hydrocarbon receptor signaling in the skin: consequences for prevention and treatment of skin disorders. *J. Invest. Dermatol*. 2015;135(11):2572 - 2576.
- [57] Zudaire E, Cuesta N, Murty V, Woodson K, Adams L, Gonzalez N, and Hirsch F. The aryl hydrocarbon receptor repressor is a putative tumor suppressor gene in multiple human cancers. *The Journal of clinical investigation*. 2008;118(2):640-650.
- [58] Zhang DD, Wang WT, Xiong J, Xie XM, Cui SS, Zhao ZG, Li MJ, Zhang ZQ, Hao DL, Zhao X, Li YJ, Wang J, Chen HZ, Lv X, Liu DP. Long noncoding RNA LINC00305 promotes inflammation by activating the AHRR-NK-kappa-B pathway in human monocytes. *Sci. Rep*. 2017;7:46204.

- [59] Fuji-Kuriyama Y, Kawajiri K. Molecular mechanisms of the physiological functions of the aryl hydrocarbon (dioxin) receptor, a multifunctional regulator that senses and responds to environmental stimuli. *Proceedings of the Japan Academy Series B, Physical and Biological Sciences*. 2010;86(1):40-53.
- [60] Walsh AA, Szklarz GD, Scott EE. Human cytochrome P450 1A1 structure and utility in understanding drug and xenobiotic metabolism. *J Biol Chem*. 2013;288(18):12932–12943.
- [61] Larigot L, Juricek L, Dairou J, Coumoul X. AhR signaling pathways and regulatory functions. *Biochim Open*. 2018;7:1–9. Published 2018 Jun 11.
- [62] String database (<https://string-db.org/cgi/network.pl?taskId=cNSwVwx85Il1>)
- [63] Pathway Commons: <https://www.pathwaycommons.org/pcviz/#neighborhood/ahrr>
- [64] Weber H, Harris MW, Haseman JK, Birnbaum LS. Teratogenic potency of TCDD, TCDF and TCDD-TCDF combinations in C57BL/6N mice. *Toxicology letters*. 1985;26(2-3):159-167.
- [65] Kung T, Murphy KA, White LA. The aryl hydrocarbon receptor (AhR) pathway as a regulatory pathway for cell adhesion and matrix metabolism. *Biochem Pharmacol*. 2009;77(4):536–546.
- [66] Gill SE, Parks WC. Metalloproteinases and their inhibitors: regulators of wound healing. *Int J Biochem Cell Biol*. 2008;40(6-7):1334–1347.
- [67] Yoshioka W, Tohyama C. Mechanisms of Developmental Toxicity of Dioxins and Related Compounds. *Int J Mol Sci*. 2019;20(3):617. Published 2019 Jan 31.

- [68] Yamada T., Mishima K., Fujiwara K., Imura H., Sugahara T. Cleft lip and palate in mice treated with 2,3,7,8-tetrachlorodibenzo-p-dioxin: A morphological in vivo study. *Congenit. Anom.* 2006;46:21–25.
- [69] Couture L.A., Harris M.W., Birnbaum L.S. Characterization of the peak period of sensitivity for the induction of hydronephrosis in C57BL6N mice following exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Fundam. Appl. Toxicol.* 1990;15:142–150.
- [70] Sakuma C., Imura H., Yamada T., Sugahara T., Hirata A., Ikeda Y., Natsume N. Cleft palate formation after palatal fusion occurs due to the rupture of epithelial basement membranes. *J. Craniomaxillofac. Surg.* 2018;46(12):2027-2031.
- [71] Yamada, T., Hirata, A., Sasabe, E., Yoshimura, T., Ohno, S., Kitamura, N., and Yamamoto, T. TCDD disrupts posterior palatogenesis and causes cleft palate. *Journal of Cranio-Maxillofacial Surgery.* 2014;42(1):1–6.
- [72] Cuervo, R., and Covarrubias, L. Death is the major fate of medial edge epithelial cells and the cause of basal lamina degradation during palatogenesis. *Development.* 2004;131(1):15-24.
- [73] Abbott, B.D., Buckalew, A.R., DeVito, M.J., Ross, D., Bryant, P.L., Schmid, J.E. EGF and TGF- α expression influence the developmental toxicity of TCDD: dose response and AhR phenotype in EGF, TGF- α , and EGF + TGF- α knockout mice. *Toxicol. Sci.* 2003;71:84–95.
- [74] Fujiwara, K., Yamada, T., Mishima, K., Imura, H., Sugahara, T.. Morphological and immunohistochemical studies on cleft palates induced by 2,3,7,8-tetrachlorodibenzo-p-dioxin in mice. *Congenit. Anom.* 2008;48:68–73.

- [75] Gao, Z., Bu, Y., Zhang, G., Liu, X., Wang, X., Ding, S., ... and Yu, Z. Effect of TCDD on the fate of epithelial cells isolated from human fetal palatal shelves (hFPECs). *Toxicology and applied pharmacology*. 2016;305:186-193.
- [76] Denison, M.S., Soshilov, A.A., He, G., DeGroot, D.E., Zhao, B., . Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor. *Toxicol. Sci.* 2011;124:1–22.
- [77] Faust D, Kletting S, Ueberham E, Dietrich C. Aryl hydrocarbon receptor-dependent cell cycle arrest in isolated mouse oval cells. *Toxicol. Lett.* 2013;223:73–80.
- [78] Peters JM, Narotsky MG, Elizondo G, Fernandez-Salguero PM, Gonzalez FJ and Abbott BD. Amelioration of TCDD-induced teratogenesis in aryl hydrocarbon receptor (AhR)-null mice. *Toxicological sciences: an official journal of the Society of Toxicology*. 1999;47(1):86-92.
- [79] Cai Y, Patterson KE, Reinier F, et al. Copy Number Changes Identified Using Whole Exome Sequencing in Nonsyndromic Cleft Lip and Palate in a Honduran Population. *Birth Defects Res.* 2017;109(16):1257–1267.
- [80] Popoff DAV et al. Non-syndromic oral clefts and risk of cancer: a systematic review. *Dentistry 3000*, [S.l.], 2013;1(1):12-18.
- [81] Lima, LS et al . Frequency of Cancer in First-Degree Relatives of Patients with Cleft Lip and/or Palate in the Brazilian Population. *Braz. Dent. J.* 2013;24(3):200-203.
- [82] Landi MT, Chatterjee N, Yu K, et al. A Genome-wide Association Study of Lung Cancer Identifies a Region of Chromosome 5p15 Associated with Risk for Adenocarcinoma. *American Journal of Human Genetics*. 2009;85(5):679-691.

- [83] Dunkhase E, Ludwig KU, Knapp M, et al. Nonsyndromic cleft lip with or without cleft palate and cancer: Evaluation of a possible common genetic background through the analysis of GWAS data. *Genomics Data*.2016;10:22-29.
- [84] Swanson JM, Entringer S, Buss C, Wadhwa PD. Developmental origins of health and disease: environmental exposures. *Semin Reprod Med*. 2009;27(5):391–402.
- [85] Knopik VS, Maccani MA, Francazio S, McGeary JE. The epigenetics of maternal cigarette smoking during pregnancy and effects on child development. *Dev Psychopathol*. 2012;24(4):1377–1390.
- [86] Nielsen, C.H., Larsen, A. and Nielsen, A.L. DNA methylation alterations in response to prenatal exposure of maternal cigarette smoking: A persistent epigenetic impact on health from maternal lifestyle? *Arch Toxicol*. 2016;90(2): 231-245
- [87] Gao X, Jia M, Zhang Y, Breitling LP, Brenner H. DNA methylation changes of whole blood cells in response to active smoking exposure in adults: a systematic review of DNA methylation studies. *Clin Epigenetics*. 2015;7:113. Published 2015 Oct 16.
- [88] Fa S, Larsen TV, Bilde K, Daugaard TF, Ernst EH, Lykke-Hartmann K, Olesen RH, Mamsen LS, Ernst E, Larsen A, Nielsen AL. Changes in first trimester fetal CYP1A1 and AHRR DNA methylation and mRNA expression in response to exposure to maternal cigarette smoking. *Environmental Toxicology and Pharmacology*. 2018; 57:19-27.
- [89] Zhang B, Hong X, Ji H, Tang W, Kimmel M, Ji Y, Pearson C, Zuckerman B, Surkan PJ and Wang X. Maternal smoking during pregnancy and cord blood DNA

methylation: new insight on sex differences and effect modification by maternal folate levels. *Epigenetics*. 2018;13(5):505-518.

[90] Sakurai S, Shimizu T, Ohto U. The crystal structure of the AhRR–ARNT heterodimer reveals the structural basis of the repression of AhR-mediated transcription. *Journal of Biological Chemistry*,207;292(43):17609-17616.

List of Tables and Figures and Corresponding Legends

Table and Table Legend

Table 1 - Low frequency variants (MAF<0.02) in AHRR gene found in a cohort of nonsyndromic cleft lip with or without cleft palate individuals. Mutations were reported in a GWAS investigation (Leslie et al., 2017). The columns within table are described as follows: Gene is the gene name; CHR is the chromosome number; BP is the base position; SNP is the single nucleotide polymorphism code in SNPDB; A1 is the reference nucleotide; A2 is the alternative nucleotide; MAF is the minimum allele frequency; N is the absolute amount of individuals; Cases is the probands; Controls are the neurotypical individuals; Population if the geographic location of the individuals in which the study was conducted.

GENE	CHR	BP	SNP	A1	A2	Cases		Controls		Population
						MAF	N (Chrom.)	MAF	N (Chrom.)	
AHRR	5	306807	rs111451538	A	G	0	1152	0.001796	1670	Europeans
AHRR	5	432995	rs200229746	A	G	0.000868	1152	0	1670	Europeans
AHRR	5	434146	rs111768223	T	C	0.01302	1152	0.00479	1670	Europeans
AHRR	5	434219	rs35756515	T	C	0.000868	1152	0	1670	Europeans
AHRR	5	434546	rs61757546	T	C	0.009549	1152	0.001198	1670	Europeans
AHRR	5	434581	rs200762873	G	C	0.003472	1152	0.000599	1670	Europeans

Figures and Figure Legends

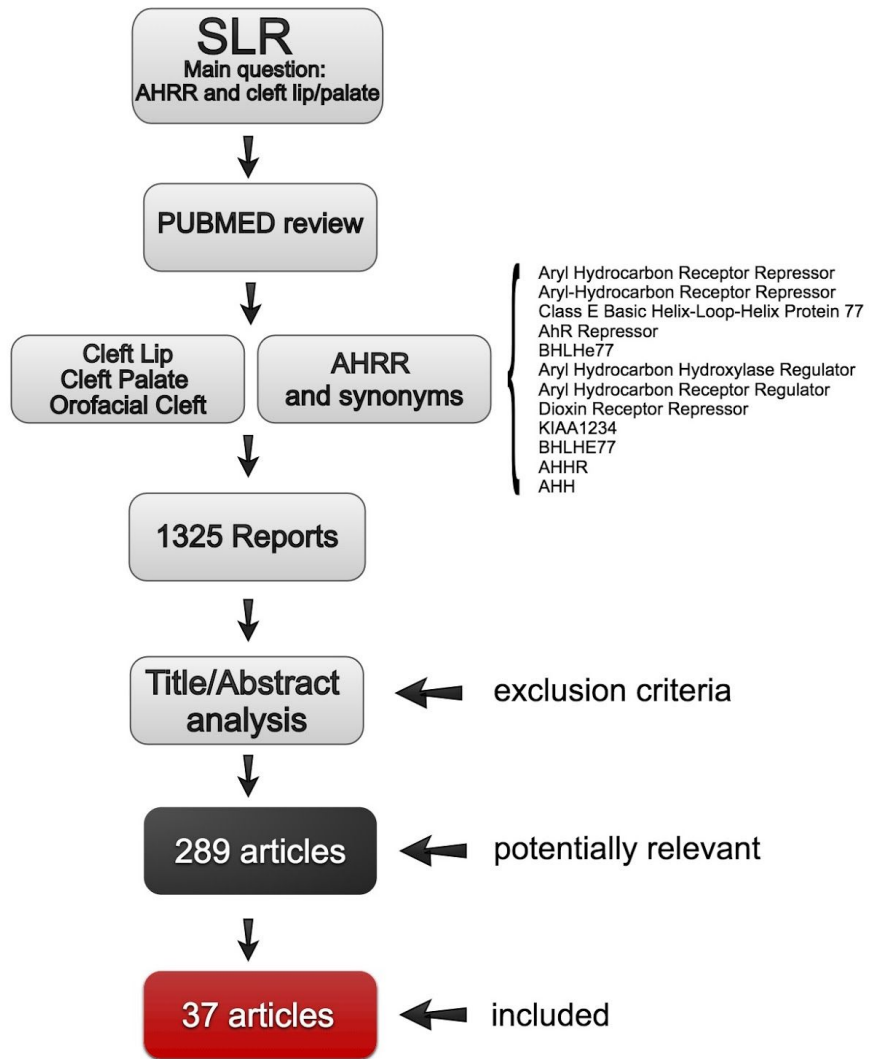


Figure 1 - Schematic representation of the systematic literature research algorithm.

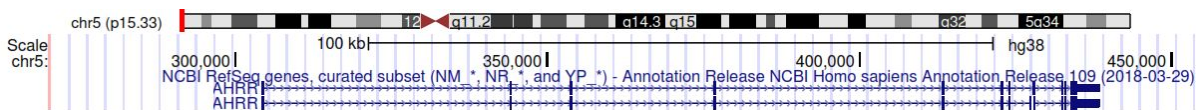


Figure 2 - Splicing isoforms of the gene *AHRR* listed in RefSeq databank. Isoforms are represented within the graph as blue lines (intronic regions) and blue boxes (exonic regions: wider are protein coding regions, and narrow are untranslated regions).

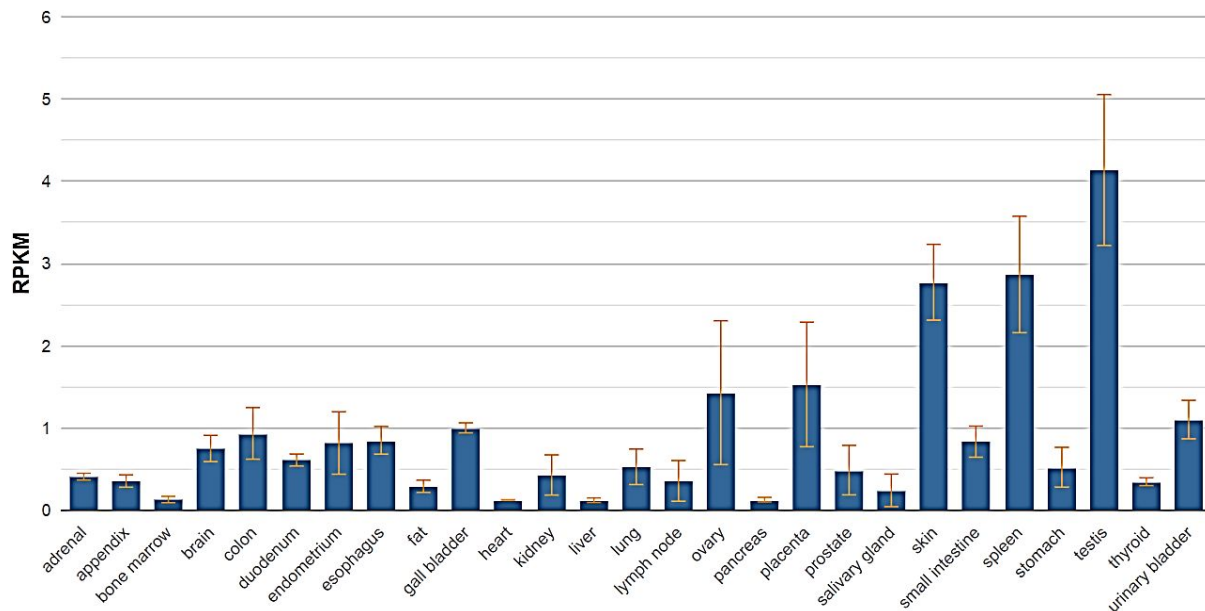


Figure 3 - AHRR gene expression through 27 human tissues. The Y axis is represented by RPKM: reads per kilobase per million of reads; the X axis if the tissue name.
Source: Extracted from NCBI with data from Fagerberg et al., 2014.

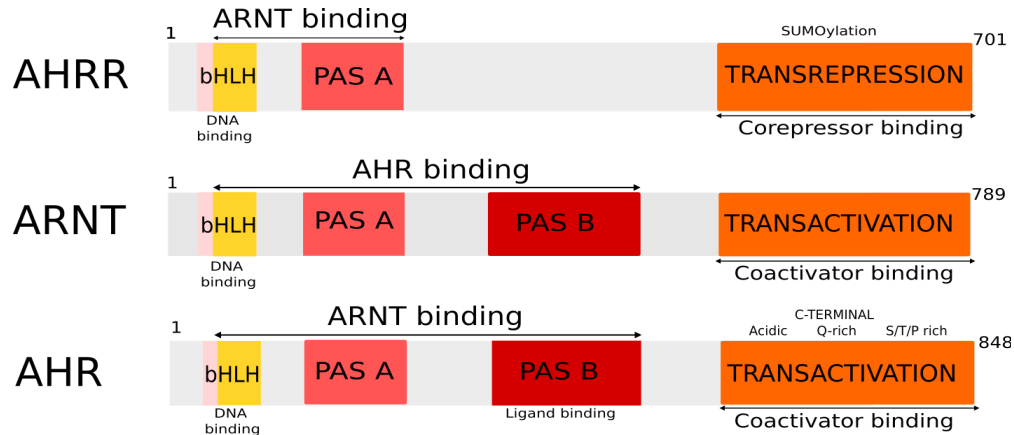
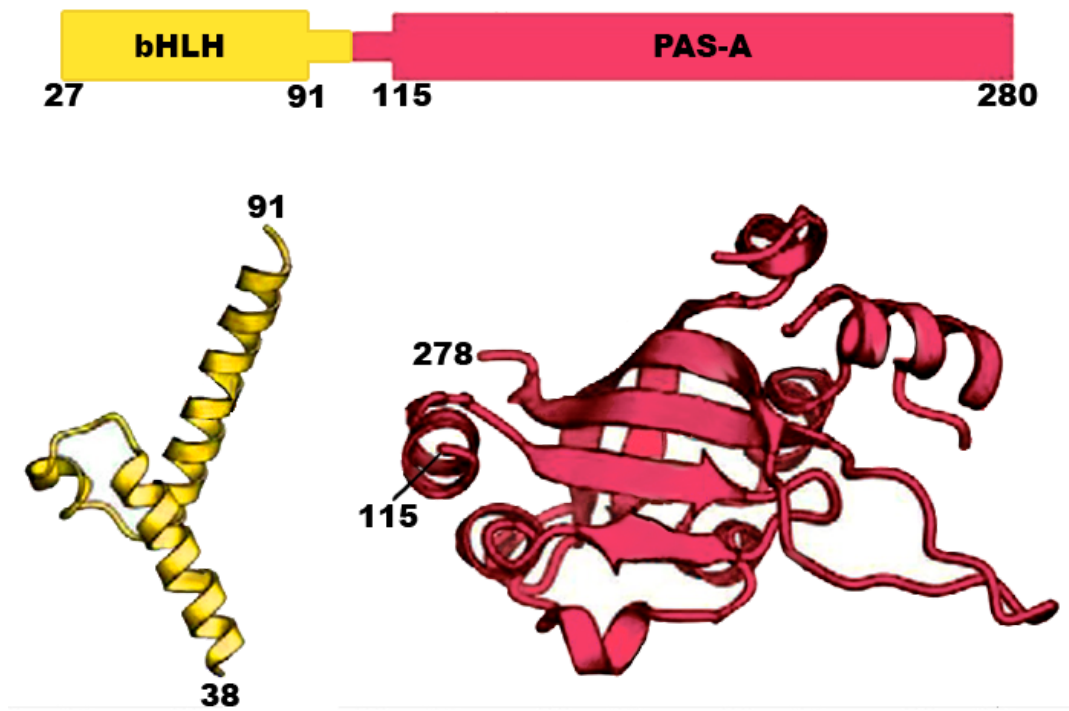
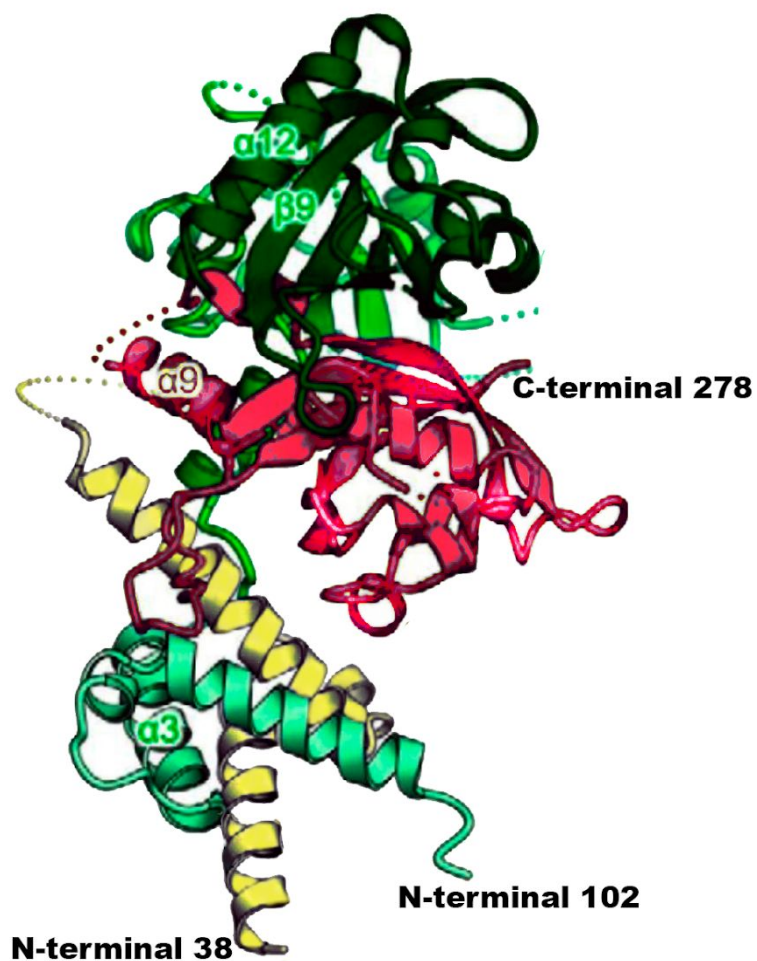
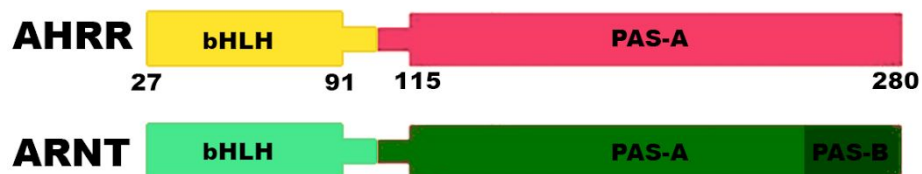


Figure 4 - AHRR, AHR and ARNT protein structure and domains. Protein domain locations through the AHRR protein and AHRR-related proteins, AHRR, ARNT and AHR have a bHLH domain and a PAS A domain which allows DNA binding and dimerization (AHR-ARNT and AHRR-ARNT). AHR has a PAS B domain which allows ligand binding such as TCDD. AHRR has a transrepressor domain which binds to corepressors and undergoes SUMOylation. ARNT and AHR have a transactivator domain which allows binding to coactivators. AHR has an additional C-terminal domain subdivided in acidic, Q-rich (glutamine enriched) and S/T/O rich (serine, threonine and proline enriched) subdomains.



A



B

Figure 5

A- Crystal structure (3D) of the active AHRR protein domains bHLH and PAS-A complexes. The gene structure (wide and narrow boxes) is composed by two domains, and below its corresponding protein 3D structure.

B- Crystal structure (3D of the active AHRR protein domains interacting with ARNT. *AHRR* (yellow and pink) and *ARNT* (green) gene structures (wide and narrow boxes) and below both 3D structures interaction. Yellow and light green represent bHLH while pink and green represent PAS A. The darkest shade of green represents PAS B.

Source: adapted from Sakurai et al., 2017

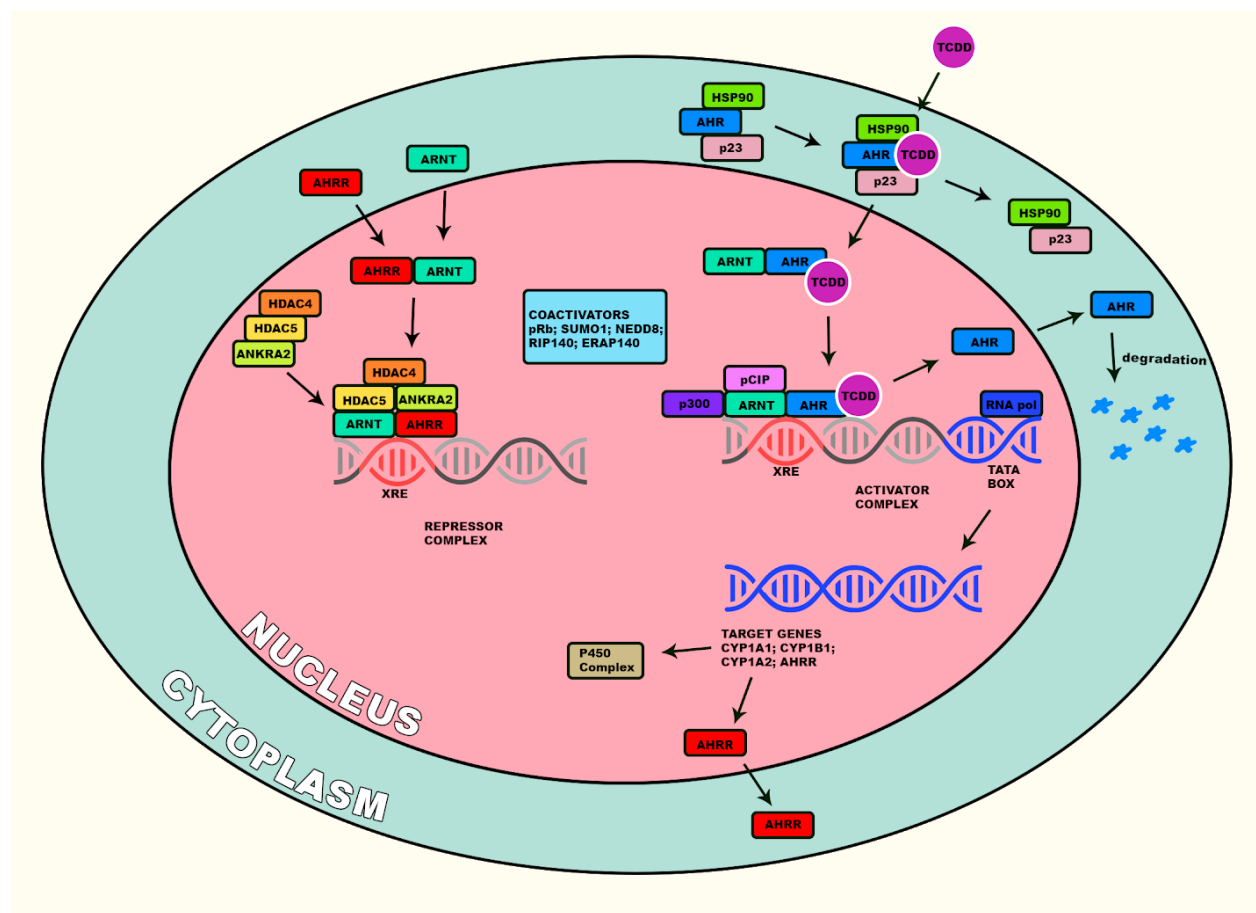


Figure 6 – A schematic model for the molecular AHR pathway and ARNT-AHRR regulation. When a ligand (TCDD) passes the plasma membrane through passive diffusion, it binds to AHR (PAS B domain)

AHR usually exists in a dormant state within the cytoplasm associated with a complex of HSP90, p23 (co-chaperone) and other proteins. This ligand binding allows the dissociation of the ligand-AHR complex from HSP90/p23 complex and translocation of this complex into the cell nucleus. Inside the nucleus, AHR complex binds to ARNT and forms a heterodimer. The AHR-ARNT dimer binds to XRE (xenobiotic response element) a specific DNA sequence located in the promoter regions of target genes recruiting components of the transcriptional machinery such as CBP/p300 (cAMP response element-binding protein binding protein) and p/CIP (p300/CBP/CoIntegrator-associated Protein) as well as other transcriptional coactivators such as RIP140 (Receptor Interacting Protein 140), pRb (retinoblastoma protein), SUMO1 (Small Ubiquitin Like Modifier 1), NEDD8 (neural precursor cell expressed, developmentally

down-regulated 8) and ERAP140 (nuclear receptor coactivator 7) . This process promotes the transcription of target genes *CYP1A1*, *CYP1B1*, *CYP1A2* and *AHRR*. ARNT forms a heterodimer with the synthesized AHRR, which enhances the SUMOylation of both proteins. SUMOylated AHRR recruits corepressors like ANKRA2, HDAC4, and HDAC5 creating a transcriptional repressor complex which also binds to XRE and represses the transcription of target genes. AHR is then exported and degraded in the cytoplasm by the proteasome.

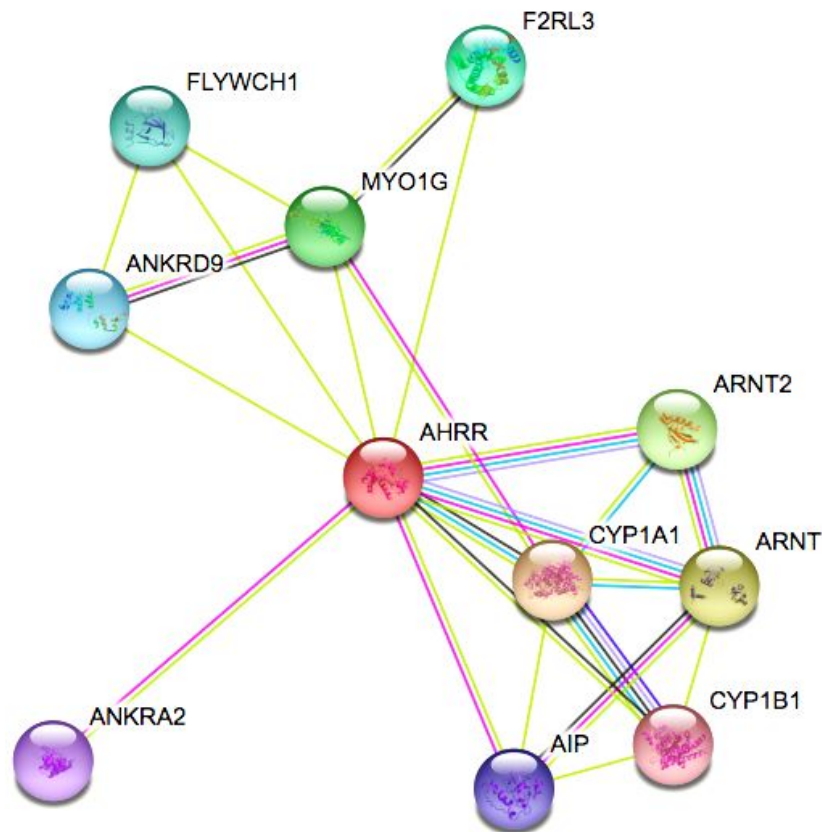


Figure 7 – Protein-protein interacting network of AHRR in humans. Interaction generated by String protein interaction database. The node colors indicate: query proteins and first shell of interactors (colored nodes). The colored lines indicate known interactions experimentally determined (pink) or from curated databases (blue); predicted interactions with genes in the same neighborhood (green), gene fusions (red) and gene co-occurrence (dark blue). The greenish yellow line indicates proteins usually mentioned together in texts (textmining), black lines stand for co-expression and the lilac line means protein homology.

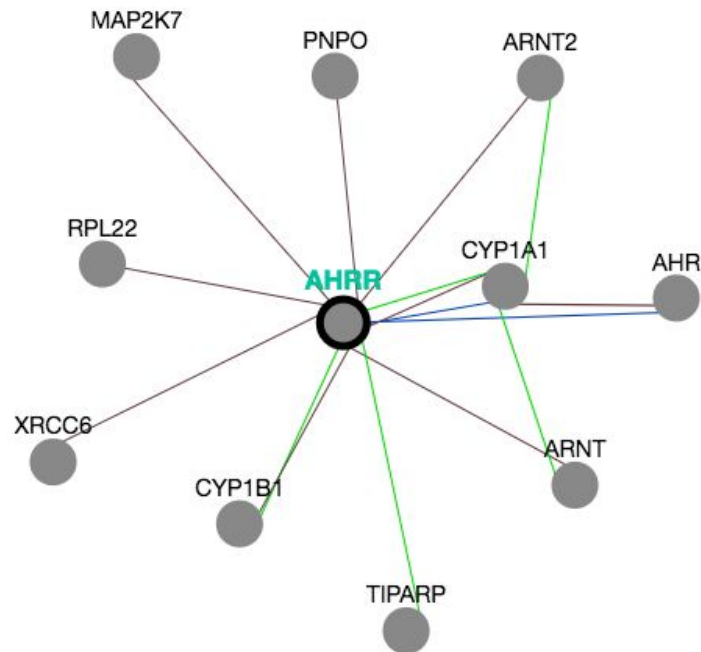


Figure 8 – Genetic interactions of AHRR gene with other known human genes. Genetic interaction was generated by PathwayCommons software. The green line indicates expression control, the blue line indicates state change control and the brown line indicates they are in the same complex.