

• 论 著 •

Fst 统计法评估 KIF1B 基因 rs17401966 多态性与肝细胞癌风险 Meta 分析的异质性来源^{*}

苏明宽, 郭建峰, 陈宏斌, 黄建成

(福建医科大学附属闽东医院检验科, 福建宁德 355000)

摘要:目的 采用 Fst 统计法评估二分类变量 Meta 分析的异质性来源。方法 检索 Google Scholar, EMBASE, PubMed, ISI Web of Knowledge 和 CNKI 数据库, 搜集有关 KIF1B 基因 rs17401966 多态性与肝细胞癌风险关系的病例-对照研究。应用 Stata 12.0 软件进行 Meta 分析, 使用 Arlequin 3.5 软件分析研究人群间的遗传分化程度。结果 最终纳入 5 个病例-对照研究, 合计 12 个研究人群。对 12 个研究人群的 Meta 分析结果显示 KIF1B 基因 rs17401966 G 等位基因与 HCC 风险负相关($OR=0.77, 95\%CI:0.65\sim0.93; P=0.005$), 然而合并结果存在很强的异质性。通过对 12 个纳入人群进行遗传分化检验, 发现 Zhang 等的 5 个研究人群与其他研究人群存在不同程度的遗传分化, 进而根据 Fst 值进行适当的亚组分析, 把组 8 和组 9 异质性检验的 I^2 值降到了 25% 以下, 然而组 8 和组 9 的 Meta 分析结果却不一致。结论 研究显示在对 KIF1B 基因 rs17401966 多态性与肝细胞癌风险进行 Meta 分析时, 通过对纳入人群的遗传分化检验, 可以发现 Meta 分析的异质性来源。

关键词: 异质性; 多态性; Meta 分析; 遗传分化; 肝细胞癌

DOI:10.3969/j.issn.1673-4130.2016.23.005

文献标识码: A

文章编号: 1673-4130(2016)23-3252-04

Fst statistical method for evaluating KIF1B gene rs17401966 polymorphism and heterogeneity source of hepatocellular carcinoma risk meta-analysis^{*}

SU Mingkuan, GUO Jianfeng, CHEN Hongbin, HUANG Jiancheng

(Department of Clinical Laboratory, Affiliated Mindong Hospital, Fujian Medical University, Ningde Fujian 355000, China)

Abstract: **Objective** To adopt Fst statistical method to assess the heterogeneity sources of meta-analysis by dichotomous variable. **Methods** The case-control studies on the relationship between KIF1B gene rs17401966 polymorphism and hepatocellular carcinoma(HCC) risk were collected by retrieving the databases including Google Scholar, EMBASE, PubMed, ISI Web of Knowledge and CNKI. The meta analysis was performed by using the Stata12.0 software. The genetic differentiation degree among populations was analyzed and researched by using the Arlequin 3.5 software. **Results** A total of 5 case-control studies were finally included, involving 12 research populations. The meta analysis on 12 populations showed that KIF1B gene rs17401966G allele was negatively correlated with HCC risk($OR=0.77, 95\%CI:0.65-0.93; P=0.005$). However, the strong heterogeneity existed in this pooled results. The genetic differentiation test in the included 12 populations found that Zhang's five research populations had varying degrees of genetic differentiation compared to other populations. Then the proper subgroup analysis was further conducted based on Fst value, and then the I^2 value of the heterogeneity test in the group 8 and 9 was descended to less than 25%. However, the meta analysis results of the group 8 and 9 were inconsistent. **Conclusion** This study showed that conducting the meta-analysis of KIF1B gene rs17401966 polymorphism and the HCC risk can find the heterogeneity sources of meta-analysis by conducting the genetic differentiation test in the included population.

Key words: heterogeneity; polymorphism; meta-analysis; genetic differentiation; hepatocellular carcinoma

2010 年, Zhang 等^[1]开展的中国人群乙型肝炎病毒(HBV)感染相关肝细胞癌(HCC)的全基因关联分析(GWAS), 发现 1p36.22 区域的 KIF1B 基因单核苷酸多态性 rs17401966 与 HCC 风险显著相关, 其 5 个研究人群的联合 P 值达 3.4×10^{-19} 。然而, 其他学者后继重复研究并没有取得相一致结果, 可能由于遗传分化, 不同的研究设计, 较小的样本量等。Meta 分析是对具有相同研究目的多个医学研究进行综合分析的一系列过程, 通过 Meta 分析, 可以达到增加样本量进而增强统计效能。效应量的异质性检验是 Meta 分析的一个重要步骤, Meta 分析的过程中若不能对研究间存在的异质性

进行合理分析, 也没有采用适当的方法对其加以控制, Meta 分析的结果就不可靠, 其结论也不能用于指导解决相应的临床问题。目前 RevMan、Stata 等 Meta 分析软件无有效的方法对异质性的来源进行分析。固定指数(Fst)是种群分化和遗传距离的一种衡量方法^[2], 可以对不同人群之间遗传关系的远近进行量化。为此本研究就引入 Fst 统计法评估 KIF1B 基因 rs17401966 多态性与肝细胞癌风险 Meta 分析的异质性来源作一讨论。

1 资料与方法

1.1 资料来源与文献检索策略 检索 Google Scholar、EM-

^{*} 基金项目: 福建省自然科学基金项目(2016J01596); 福建省宁德市科技计划项目(20150013)。

作者简介: 苏明宽, 男, 主管技师, 主要从事分子生物学研究。

BASE、PubMed、ISI Web of Knowledge 和 CNKI 数据库,搜集有关 KIF1B 基因 rs17401966 多态性与 HCC 风险关系的病例-对照研究,检索时间为 2010 年 1 月至 2016 年 4 月。检索采取主题词与自由词相结合的方式,英文检索词包括 chronic hepatitis B、hepatocellular carcinoma、HCC、Liver cancer、KIF1B、Kinesin family member 1B、rs17401966、polymorphism 和 variant;中文检索词包括慢性乙型肝炎、肝细胞癌、肝癌、驱动蛋白家族、rs17401966、多态性、变异。纳入文献的标准如下:(1)公开发表发表的 KIF1B 基因 rs17401966 多态性与 HCC 风险相关的病例-对照研究;(2)有相应基因型的频数数据以可用来计算遗传分化(genetic differentiation)程度;(3)以慢性 HBV 感染者为对照组;(4)所研究对象均要排除合并 HCV、HIV 感染;(5)慢性 HBV 感染及 HCC 诊断标准符合中国或国际标准。排除标准如下:(1)基于家系的研究;(2)慢性 HBV 感染及 HCC 诊断标准描述不清的文献;(3)文献质量差、重复报告及数据描述不详的研究。总共检索到文献 37 篇,经过纳入和排除标准其中 32 篇文献被排除,最后纳入 Meta 分析的文献共 5 篇^[1,3-6],合计 12 个研究人群。其中中国人群为 8 个、日本人群为 2 个、韩国人群 1 个、泰国人群 1 个。12 个研究人群累计 HCC 4 886 例,对照 5 442 例。由 2 名研究者分别提取入选文献研究对象的种族来源、病例与对照组人数、基因型频数信息,纳入文献基本信息见表 1。

1.2 统计学处理 采用显性遗传模型(GG+AG *vs.* AA)来分析 rs17401966 多态性与 HCC 的风险。二分类资料采用比值比(OR)作为效应统计量,并计算其 95%可信区间(95% CI)。采用 Cochrane Q 检验对纳入研究进行异质性检验,同时结合 *I*² 定量判断异质性的^{大小}。*I*²<25%时,认为没有异质性。*I*² 值为 25%~<50%时,认为有轻度的异质性。*I*² 值为 50%~75%时,认为有中度的异质性。当 *I*²>75%时,认为有很强的异质性。当 *I*²<50%时,采用固定效应模型进行统计量的合并。否则,采用随机效应模型来合并统计量。以上统计量均使用 Stata 12.0 软件进行统计分析,双侧检验,以 *P*<0.05 为差异有统计学意义。使用 Arlequin 3.5 软件分析研究人群间的遗传分化程度,Fst>0.05 为差异有统计学意义。

2 结 果

2.1 遗传分化检验 由于不同地域的人群遗传背景各异导致

对同种疾病存在不同程度的易感性^[7],为了避免这些差异在人群合并分析后难以表现出来,需要对纳入研究的 12 个人群进行遗传分化检验。Fst 用来描述遗传分化程度,Fst<0.05 代表两个人群没有遗传分化。表 2 列出了 12 个人群两两比较的 Fst 值,发现广西与香港(Fst=0.064)、泰国(Fst=0.052)、北京 2(Fst=0.057)人群存在显著遗传分化。另外北京 1 与香港(Fst=0.056)、北京 2 (Fst=0.050)也存在显著遗传分化。还有一些人群之间存在低水平的遗传分化,虽然其 Fst<0.05。

2.2 KIF1B 基因 rs17401966 多态性与 HCC 风险的关联 对 12 个研究人群的合并结果显示,在显性遗传模型下,异质性检验 *I*² 为 77.3%,*P* 值更是达到了 1.19×10⁻⁶,提示研究存在很强的异质性,因此采用随机效应模型。Meta 分析结果显示,KIF1B 基因 rs17401966 G 等位基因与 HCC 风险负相关(OR=0.77,95%CI:0.65~0.93;*P*=0.005)。组 2、组 4、组 6、组 8、组 9 异质性检验 *I*² 值均为 0.0%,*P* 值为 0.567~0.890,因此均采用固定效应模型。组 3、组 5、组 7 异质性检验 *I*² 值为 61.4%~76.9%,因此均采用随机效应模型合并统计量。Meta 分析结果见表 3。

表 1 纳入研究基本特征(<i>n</i> , <i>n</i> / <i>n</i>)						
纳入研究	研究人群	样本量		rs17401966 基因型(病例/对照)		
		病例	对照	GG	AG	AA
Zhang 等	广西	348	359	8/26	100/141	240/192
Zhang 等	北京 1	276	266	5/24	86/109	185/133
Zhang 等	江苏	507	215	26/21	181/101	300/93
Zhang 等	广东	751	509	26/35	228/195	497/279
Zhang 等	上海	428	440	12/32	141/169	275/239
Hu 等	南京	1 300	1 344	107/118	480/500	706/716
Sawai 等	日本 1	179	769	13/45	61/261	105/463
Sawai 等	日本 2	142	251	5/14	46/91	91/146
Sawai 等	韩国	164	144	17/15	59/55	88/74
Sawai 等	香港	94	187	10/13	39/80	44/94
Sopipong 等	泰国	202	196	21/16	81/83	100/97
Chen 等	北京 2	503	772	63/65	194/309	246/398

表 2 12 个人群病例组与对照组两两比较的 Fst 值												
	广西	北京 1	江苏	广东	上海	南京	日本 1	日本 2	韩国	香港	泰国	北京 2
广西		−0.001 44	0.010 99	0.000 27	0.000 99	0.026 44 *	0.016 48	0.000 74	0.038 92 *	0.064 90 * *	0.052 35 * *	0.057 21 * *
北京 1	0.000 08		0.007 97	−0.000 72	−0.000 32	0.022 23 *	0.012 47	−0.000 85	0.032 62 *	0.056 24 * *	0.045 04 *	0.050 29 * *
江苏	0.007 92	0.001 16		0.004 95	0.003 01	0.003 20	−0.001 40	0.000 81	0.005 82	0.017 35	0.012 99	0.018 45
广东	−0.001 00	0.001 62	0.011 17		−0.000 79	0.017 80	0.008 33	−0.001 72	0.026 69 *	0.048 31 *	0.039 16 *	0.045 69 *
上海	−0.001 22	0.000 79	0.009 44	−0.001 01		0.014 36	0.005 68	−0.002 29	0.021 78 *	0.041 36 *	0.032 93 *	0.038 92 *
南京	−0.000 76	−0.000 21	0.006 53	−0.000 07	−0.000 45		0.000 06	0.011 07	−0.001 14	0.003 14	0.001 84	0.005 39
日本 1	0.003 47	0.010 74	0.026 98 *	0.002 00	0.002 75	0.005 39		0.002 90	0.001 32	0.009 87	0.006 79	0.011 48
日本 2	0.000 95	0.006 65	0.020 20 *−	0.000 06	0.000 46	0.002 63	−0.001 10		0.016 81	0.033 90 *	0.026 52 *	0.032 43 *
韩国	−0.000 70	−0.002 68	0.000 33	0.000 86	0.000 02	−0.001 00	0.010 18	0.006 05		−0.001 50	−0.001 72	0.000 77
香港	−0.001 50	−0.001 95	0.003 14	−0.000 46	−0.001 03	−0.001 38	0.006 68	0.003 29	−0.002 75		−0.003 56	−0.003 19
泰国	−0.000 47	−0.002 21	0.001 12	0.001 01	0.000 21	−0.000 70	0.009 92	0.005 91	−0.003 01	−0.002 38		−0.001 31
北京 2	−0.000 42	−0.000 98	0.004 10	0.000 63	0.000 06	−0.000 33	0.007 58	0.004 30	−0.001 78	−0.001 66	−0.001 40	

注:* 代表存在低水平的遗传分化,** 代表存在较强的遗传分化,右上角为 HCC 组 Fst 值,左下角为对照组 Fst 值。

表 3 KIF1B 基因 rs17401966 多态性与 HCC 风险 Meta 分析结果

分组	Meta 分析结果		异质性检验	
	OR(95%CI)	P	I ²	P
组 1	0.77(0.65~0.93)	0.005	77.3%	1.19×10 ⁻⁶
组 2	0.51(0.40~0.64)	6.75×10 ⁻⁹	0.0%	0.831
组 3	0.84(0.71~1.00)	0.051	70.9%	3.0×10 ⁻⁴
组 4	1.09(1.91~1.31)	0.360	0.0%	0.890
组 5	0.70(0.57~0.85)	4.0×10 ⁻⁴	76.9%	3.0×10 ⁻⁵
组 6	1.01(0.90~1.13)	0.906	0.0%	0.837
组 7	0.63(0.57~0.71)	8.97×10 ⁻¹⁵	61.4%	0.016
组 8	0.58(0.51~0.65)	2.89×10 ⁻¹⁷	0.0%	0.567
组 9	1.00(0.90~1.11)	0.955	0.0%	0.816

3 讨 论

Meta 分析的异质性分为临床异质性、方法学异质性和统计学异质性^[8-9]。临床异质性是指参与者不同、干预措施的差异及研究的终点指标不同所导致的变异。方法学异质性是指由于试验设计和质量方面的差异引起的变异。因此,降低临床异质性和方法学异质性的手段是在 Meta 分析时,首先要制订严格、统一的纳入和排除标准,只有具有相同研究目的、高质量的研究才能纳入分析^[10]。统计学异质性是研究间临床和方法学上多样性的直接结果。异质性检验的方法主要有 *I*² 统计量、Q 统计量、H 统计量、Galbraith 图^[11] 和 *L'* Abbe 图法^[12], 根据以上 5 种异质性检验结果选择固定效应模型或随机效应模型。也可将所纳入数据分成更小的单元,进而在亚组内进行比较,如按不同设计方案、研究对象地域来源等进行亚组分析。然而异质性来源于多个纳入研究对象时,仅通过人为而无一种检验方法进行分组分析存在诸多困难。为此引入 Arlequin 3.5 软件的 Fst 统计法评估二分类变量 Meta 分析的异质性来源。

从 Fst 值可以看出 Meta 分析的异质性主要源于病例组, 对照组中仅江苏人群与日本人群存在低水平的遗传分化。由于广西和北京 1 人群与香港、泰国、北京 2 人群存在较强的遗传分化,而广西与北京 1 人群无遗传分化,另外香港、泰国、北京 2 这 3 个人群之间也无遗传分化。因此,把广西和北京 1 人群分为组 2,对余下 10 个人群分为组 3,或把香港、泰国、北京 2 人群分为组 4,另外 9 个人群分为组 5,然而组 3、组 5 仍存在很强的异质性。进一步把南京、韩国、香港、泰国、北京 2 人群分为组 6,余下 7 个人群分为组 7,虽然组 7 的异质性检验 *I*² 值低于组 3 和组 5,但仍大于 50%。因为有的人群之间存在低水平的遗传分化,而 Meta 分析是通过对其纳入人群的合并分析,对低水平的遗传分化具有累加作用,而造成 *I*² 大于 50%。本文发现 Zhang 等的 5 个研究人群与其他 7 个研究人群存在不同程度的遗传分化,可能原因为这 7 个研究人群来源于不同文献,其病例的选择存在差异。另外,还发现北京 1 与北京 2 人群存在很强的遗传分化,其原因不排除存在抽样误差的可能。因此,把 Zhang 等的 5 个研究人群划分为组 8,另外 7 个人群划分为组 9,通过这种分组方式,把这 2 个亚组的异质性检验 *I*² 值降到了 25% 以下。值得注意的是,组 8 累计 HCC

2 310 例,对照 1 789 例。组 9 累计 HCC 2 576 例,对照 3 653 例。然而组 8 和组 9 的 Meta 结果却不一致,因此,认为 KIF1B rs17401966 多态性与 HCC 风险还需进一步研究。

综上所述,本研究显示,在对 KIF1B 基因 rs17401966 多态性与肝细胞癌风险进行 Meta 分析时,通过引入 Fst 统计法对纳入人群进行遗传分化检验,可以发现 Meta 分析的异质性来源,并可根据 Fst 值进行适当的亚组分析。

参考文献

[1] Zhang H, Zhai Y, Hu Z, et al. Genome-wide association study identifies 1p36.22 as a new susceptibility locus for hepatocellular carcinoma in chronic hepatitis B virus carriers[J]. Nat genet, 2010, 42(9):755-758.

[2] Ngamruengphong S, Patel T. Molecular evolution of genetic susceptibility to hepatocellular carcinoma[J]. Dig Dis Sci, 2014, 59(5):986-991.

[3] Hu L, Zhai X, Liu J, et al. Genetic variants in human leukocyte antigen/DP-DQ influence both hepatitis B virus clearance and hepatocellular carcinoma development[J]. Hepatology, 2012, 55(5):1426-1431.

[4] Sawai H, Nishida N, Mbarek H, et al. No association for Chinese HBV-related hepatocellular carcinoma susceptibility SNP in other East Asian populations[J]. BMC Med Genet, 2012, 19(13):47-52.

[5] Sopipong W, Tangkijvanich P, Payungporn S, et al. The KIF1B(rs17401966) single nucleotide polymorphism is not associated with the development of HBV-related hepatocellular carcinoma in Thai patients[J]. Asian Pac J Cancer Prev, 2013, 14(5):2865-2869.

[6] Chen K, Shi W, Xin Z, et al. Replication of genome wide association studies on hepatocellular carcinoma susceptibility loci in a Chinese population[J]. PLoS One, 2013, 8(10):e77315.

[7] Pan W, Song IS, Shin HJ, et al. Genetic polymorphisms in Na⁺-taurocholate co-transporting polypeptide (NTCP) and ileal apical sodium-dependent bile acid transporter (ASBT) and ethnic comparisons of functional variants of NTCP among Asian populations[J]. Xenobiotica, 2011, 41(6):501-510.

[8] Tao W, Li PS, Shen Z, et al. Effects of omega-3 fatty acid nutrition on mortality in septic patients; a meta-analysis of randomized controlled trials[J]. BMC Anesthesiology, 2016, 16(1):39.

[9] Tian WJ, Huang ML, Qin QF, et al. Prognostic Impact of epidermal growth factor receptor overexpression in patients with cervical cancer: a meta-analysis[J]. PLoS one, 2016, 11(7):e0158787.

[10] Brunetti ND, Santoro F, Correale M, et al. Incidence of atrial fibrillation is associated with age and gender in subjects practicing physical exercise; a meta-analysis and meta-regression analysis [J]. Intern J Cardiol, 2016, 221(10):1056-1060.

(下转第 3257 页)