

平成 30 年 6 月 22 日現在

機関番号：33111

研究種目：若手研究(B)

研究期間：2014～2017

課題番号：26860423

研究課題名(和文)チリ胆嚢がん患者の予後規定因子および本症発生に關与する遺伝要因の解明

研究課題名(英文)Genetic factors to predict the survival of gallbladder cancer patients in Chile

研究代表者

浅井 孝夫 (Asai, Takao)

新潟医療福祉大学・医療技術学部・講師

研究者番号：60612736

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：胆嚢がんの予後に遺伝的要因が關与していると考え、我々は生存期間に關連するSNPの同定を試みた。チリ胆嚢がん患者96例を対象にCancer SNP Panel (Illumina Inc.)を用いて1,421 SNPを網羅的に解析し、疾患特異的生存率に關与するSNPをスクリーニングした。その結果、6個の生存期間に有意差を有するSNPが同定され、これらリスクSNPの保有数は有意に癌特異的生存期間と關連していた($P=3.05 \times 10^{-10}$)。また、年齢、ステージを共変量とした多変量Cox回帰分析においても、リスクSNPの保有数は独立した予後予測因子であることが示された($P=2.26 \times 10^{-8}$)。

研究成果の概要(英文)：Genetic factors may contribute to the survival of gallbladder cancer (GBC) patients. We aimed to identify SNPs involved in the survival times of GBC patients. A total of 96 GBC patients in Chile were analyzed using a Cancer SNP Panel (Illumina Inc.) that contained 1,421 SNPs in 408 cancer-related genes. SNPs associated with the cancer-specific survival were screened by log-rank test. Six SNPs were identified to have statistically significant association with cancer-specific survival. The cancer-specific survival times of patients grouped according to the number of risk genotypes of 6 SNPs differed significantly (0-1 vs 2-3 vs 4-6 risk genotypes; $P=3.05 \times 10^{-10}$). The number of risk genotypes was independently associated with survival of GBC patients in a multivariate analysis that included age and stage ($P=2.26 \times 10^{-8}$).

研究分野：分子遺伝疫学

キーワード：胆嚢がん 遺伝要因 遺伝子多型 SNPジェノタイピング チリ 予後予測 生存期間

1. 研究開始当初の背景

(1) 胆嚢がんは南米アンデス山脈西側やインド北部など特定の地域や民族において多発するものの世界的には比較的稀な悪性腫瘍であり、他の消化器系がん比べて症例や研究報告が少ないことから、これまで外科的切除以外の有効な治療方法は確立されていない。また本症は自覚症状・初期症状が乏しいため早期発見が困難ながんであり、進行期で診断される例が多いことから、患者の5年生存率が低く、予後も悪い。そのため、胆嚢がんの早期発見につながる有効な腫瘍マーカーの同定や予防・治療の指針となる因子の解明が望まれている。

(2) 一方、同じ病理組織学的ステージで切除術を受けても、症例によって予後の良・不良に差がみられることが報告されている。胆嚢がんの予後を規定する因子については、進行度、リンパ節転移、壁深達度、組織型などいくつか報告されているが、予測因子として活用するには施設間差や判定者間差が生じやすく、また時間の経過で変化しうるもので統一性や普遍性に欠けるという課題が残されている。もし、客観性と普遍性に優れた遺伝的要因からなる予後規定因子が見出されれば、予後良好の患者には副作用の少ない治療から開始するなど予後に応じて適切な治療の選択が可能となり、治療方針選択に大きな材料を提供することができると期待される。

2. 研究の目的

胆嚢がんの予後に遺伝的要因が関与していると考え、我々は、がん関連遺伝子 SNP アレイを用いて、生存期間に関わる遺伝子の同定を試みた。SNP (Single nucleotide polymorphism, 一塩基多型) は先天的な遺伝子の個人差で再現性が保証され、また簡便かつ非侵襲的に検査できる。試料には FFPE (Formalin-Fixed Paraffin-Embedded, ホルマリン固定パラフィン包埋) 組織を用いることで、汎用性の高い長期保存標本から、豊富な臨床情報とともに過去に遡って参照することが容易である。これらの特徴に加え、症例数が豊富かつ患者情報を照会するシステムが整備されている胆嚢がん多発国チリにおいて、本症患者のがん関連 SNP を網羅的に解析することにより、患者の予後を規定する因子を明らかにすることを目的とした。

3. 研究の方法

(1) チリ・サンティアゴの Hospital Dr. Sótero del Río において 2007 年から 2012 年までの間に胆嚢がんを診断され胆嚢切除術を受けた患者 96 例の FFPE 組織を試料とした。また、解析に必要な情報として、患者の性別、病理組織学的ステージ、死亡届日、初回手術日、誕生日などの情報を収集した。本研究において、生存期間は、初回手術日から死亡届日までの月数とした。また、初回手術

時年齢は、誕生日から初回手術日までの年数とした。観察期間終了時点で死亡届日が入力されていない患者は打ち切りとした。

(2) FFPE 組織のブロックからマイクロトームを用いて厚さ 10 μm の切片を作製した。切片 4~8 枚を脱パラフィンした後、QIAamp DNA FFPE Tissue Kit (Qiagen, Hilden, Germany) を用いて DNA を回収し、TE 緩衝液に溶解した。DNA の品質は、Illumina 推奨の DNA 濃度測定試薬 Quant-iT PicoGreenR dsDNA Assay Kit を用いて濃度 50ng/ μl 以上、DNA 総重量 250ng 以上となることを確認し、規定値を満たしていなかった 8 サンプルについては、DNA の濃縮を行った。その結果、解析に使用可能な品質条件を満たし、解析可能なサンプルと判断した。

(3) DNA サンプルは、Illumina Golden Gate assay (Illumina, Inc., San Diego, CA) によりゲノム DNA を増幅し、Illumina Golden Gate Custom 1536SNP Beadsarray にゲノム DNA をハイブリダイゼーションし、洗浄、染色後、スキャニングを行った。その後、抽出されたファイルを Genome Studio (Illumina, Inc.) で数値化、Normalization を行い、SNP 情報のプロファイリング結果を抽出した。

遺伝子型判定に用いた Illumina Cancer SNP Panel には、SNP500 データベースにあるアポトーシス、腫瘍形成、腫瘍抑制および G タンパク質共役受容体タンパク質シグナル伝達等に関連する 408 遺伝子から 1,421 SNP が搭載されている。

(4) タイピングした SNP は前処理 (コールレート 99%未満のサンプルを除外、ノーコールを含む SNP を除外、マイナーアレル頻度 (Minor allele frequency: MAF) が 5%未満の SNP を除外等) を行ったうえで統計解析に用いた。

ログランク検定による、生存期間と SNP の関連解析を行い、同一 SNP の各変異型アレルのモデルのうち、ログランク検定 P 値が最小となるモデルを選択した。多重比較の補正のため、検定 P 値については、多重検定によるエラーとエラーの両方を考慮した false discovery rate (FDR) 法を用いて P 値の補正を行い、FDR-Q 値を計算した。FDR-Q のカットオフ値は、多重検定により棄却された全ての帰無仮説のうち、エラーが含まれている確率を表す。本解析では、FDR-Q 値が 0.3 より小さくなる SNP に絞り込んだ。

選定された SNP について、ハーディー・ワインベルグ平衡 (Hardy-Weinberg equilibrium: HWE) 検定を行い、SNP タイピング時のエラーや、標本集団の遺伝的偏りの可能性について検討した。

FDR の基準により選択された SNP を用いて各患者のリスクを推定するため、多変量

Cox 回帰分析を行って予測スコアを推定した。各患者の予測スコアが全体の中央値より小さい場合は Low risk 群、中央値以上の場合は High risk 群に分類した。本推定の精度を確認するため、LOOCV (leave-one-out cross-validation) を用いて学習データセットについて多変量 Cox 回帰分析を行った。学習データセットの予測スコアの中央値より、評価データの予測スコアが小さい場合は Low risk 群、中央値以上となる場合は High risk 群として、全体の分類との一致率を確認した。

多変量 Cox 回帰によるリスク分類に加えて、選択されたリスク SNP の保有数による分類も行った。加えて、共変量として考えられる臨床的因子として、性別、年齢、病理組織学的ステージについてもカテゴリ化を行い、それぞれ生存期間のログランク検定を行って生存期間との関連解析を行った。ここで有意と考えらる因子について、SNP によるリスク群と合せて多変量 Cox 回帰モデルを構築した。

これらの統計解析には、統計言語 R (ver.3.5.0) を使用し、SNP の遺伝モデル設定および評価には、SNPassoc (SNPs-based whole genome association studies, v.1.9-2) ログランク検定には survival (Survival Analysis, v.2.41-3)、作図には ggplot2 (Create Elegant Data Visualisations Using the Grammar of Graphics, v.2.2.1) を使用した。Cox 回帰分析には stats (The R Stats Package v.3.5.0) を使用した。

4. 研究成果

(1) SNP タイピングのコールレートは 99.7% であった。96 検体のうちサンプルのコールレートが 99% を下回った 5 件については、解析の対象外とした。また、ノーコールを含む 77 SNPs についても解析から除外した。MAF が 5% 未満となった 97 SNPs については、検出力が著しく小さくなり、過剰適合の原因になるため、解析から除外した。

(2) SNP の Quality Control として MAF やコールレートによるスクリーニングを行った結果、91 検体の 1,255 個の SNPs に限定された。ここで、SNPs データを変異型アレルのモデルに置き換え、MGF が 0.3 以上のスクリーニングを行い、dominant モデル、recessive モデル、additive モデルからそれぞれ 766 個、8 個、1 個の合計 775 個の SNPs を選択した。これらの SNPs を対象にログランク検定を行った。多重検定を考慮し、ログランク検定の P 値を補正した FDR-Q 値を算出し、FDR-Q 値が 0.3 より小さい SNPs に絞り込んだ。これにより、表 1 に示す 6 個の SNPs が選定された。

表 1.

SNP location (Gene)	Risk genotype	MGF	Log-rank P value	FDR Q value
rs2235716 (FOXC1)	C/C (vs T/C, T/T)	0.385	0.0023	0.295
rs3748302 (RERG)	C/C (vs T/C, T/T)	0.407	0.0007	0.286
rs3741211 (IGF2AS)	T/T (vs T/C, C/C)	0.495	0.0006	0.286
rs1760897 (TEP1)	T/T (vs T/C, C/C)	0.451	0.0020	0.295
rs1800871 (IL10)	C/C (vs T/C, T/T)	0.396	0.0019	0.295
rs2909786 (APC)	G/G (vs A/G, A/A)	0.374	0.0013	0.295

これら選択された 6 個の SNPs についてハーディー・ワインベルグ平衡が仮定された。選択された SNPs はそれぞれ別の染色体上にある 6 個の遺伝子 (FOXC1, RERG, IGF2AS, TEP1, IL10, APC) 上から得られた。

(3) これら 6 SNPs の保有状況と胆嚢がんリスクとの関連を予測するため、LOOCV を採用した多変量 Cox 回帰を実施した結果、46 件が高リスク群、45 件が低リスク群となった。LOOCV による正答率は 100% であった。

また、ログランク検定を実施したところ、High risk 群の生存期間の中央値は 4.4 か月、Low risk 群の中央値は 12.1 か月で、P 値 < 0.0001 であった (図 1)。

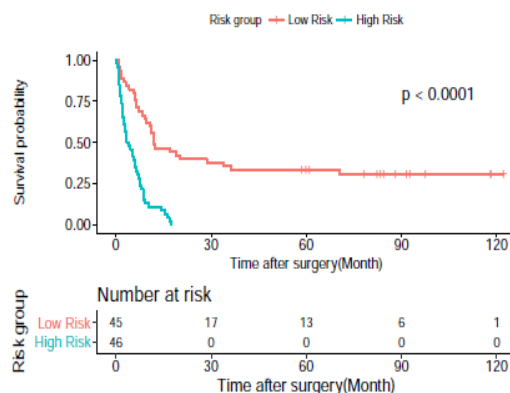


図 1.

(4) また 6 SNPs のうち、リスク遺伝子型を保有する数と生存期間の関連について解析を行った。リスク遺伝子型を 0~1 個保有する群、2~3 個保有する群、4~6 個保有する群の 3 群にカテゴリ化し、ログランク検定を実施したところ、それぞれの生存期間の中央値は 53.3 か月、6.4 か月、2.8 か月で P 値 < 0.0001 であった (図 2)。

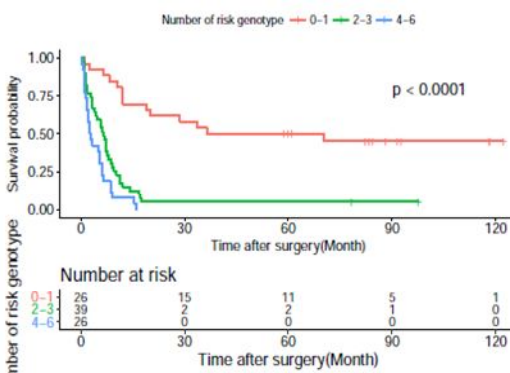


図 2.

(5)次に選定された個々のSNP、6個のリスク遺伝子型の保有数、6SNPsによるリスク分類(Low risk群 vs High risk群)、年齢(65歳未満 vs 65歳以上)、性別(男性 vs 女性)、ステージ(3以下 vs 4)と生存期間との関連について単変量Cox比例ハザード分析を行った(表2)。

表2.

Variables	Univariate analysis		
	HR (95% CI)	P	AIC
rs2235716 in FOXC1(C/C vs T/C-T/T)	2.04(1.28-3.25)	0.0028	587.9
rs3748302 in RERG(C/C vs T/C-T/T)	2.18(1.37-3.46)	0.0010	586.0
rs3741211 in IGF2AS(T/T vs T/C-C/C)	2.18(1.38-3.44)	0.0008	585.4
rs1760897 in TEP1(T/T vs T/C-C/C)	2.04(1.29-3.23)	0.0024	587.4
rs1800871 in IL10(C/C vs T/C-T/T)	2.06(1.3-3.28)	0.0023	587.5
rs2909786 in APC(G/G vs A/G-A/A)	2.1(1.32-3.32)	0.0016	587.0
sum6(0-1 vs 2-3 vs 4-6)	2.83(2.05-3.92)	3.1E-10	555.4
RiskGroup6 (Low vs High)	4.19(2.5-7)	4.8E-08	564.6
age (<65 vs >=65)	1.71(1.07-2.74)	0.0257	591.8
gender (mail vs femail)	1.54(0.91-2.62)	0.1100	594.1
stage (0-3 vs 4)	3.39(2-5.74)	5.8E-06	450.9

個々のSNPとリスク遺伝子型保有数、リスク分類を比較すると、リスク分類(RiskGroup6 (Low vs High))のハザード比は4.19(95%信頼区間=2.5-7、P値=4.8E-08)であり最も大きかった。また年齢65歳を閾値に2群に分類した変数では、高齢群のハザード比が1.71(95%信頼区間=1.07-2.74、P値=0.0257)であった。ステージについては、不明だった15件を除いたうえで、ステージ0~3とステージ4の2群に分類して解析したところ、ステージ4群のハザード比が3.39(95%信頼区間=2-5.74、P値=5.8E-06)となり、有意により短い生存期間と関連していた。性別についてはハザード比が1.54(95%信頼区間=0.91-2.62、P値=0.11)であり生存期間との関連性は示されなかった。

(6)単変量解析で有意であったリスク遺伝子型保有数、リスク分類において、年齢、ステージの因子を加えて、多変量Cox比例ハザード分析を行った。なお、ステージが不明な15件については除外し、サンプル76件を対象として解析を行った(表3)。

表3.

Variables	Multivariate (model1)			Multivariate (model2)		
	HR (95% CI)	P	AIC	HR (95% CI)	P	AIC
sum6 (0-1 vs 2-3 vs 4-6)	2.85(1.98-4.12)	2.3E-08	421.5			427.8
RiskGroup6 (Low vs High)				4.69(2.53-8.68)	9.0E-07	
ages (<65 vs >=65)	1.48(0.87-2.53)	0.1521		1.02(0.59-1.75)	0.9533	
stage (0-3 vs 4)	2.38(1.37-4.15)	0.0021		2.45(1.41-4.24)	0.0014	

6個のリスク遺伝子型の保有数を用いたモデル(model1)でのハザード比は、リスク遺伝子の保有数では2.85(95%信頼区間=1.98-4.12、P値=2.2E-08)、年齢では1.48(95%信頼区間=0.87-2.53、P値=0.1521)、およびステージでは2.38(95%信頼区間=1.37-4.15、P値=0.0021)であった。

一方、6個のSNPsを用いたリスク分類を

用いたモデル(model2)でのハザード比は、リスク遺伝子の保有数では4.69(95%信頼区間=2.53-8.68、P値=9.0E-07)、年齢では1.02(95%信頼区間=0.59-1.75、P値=0.9533)、およびステージでは2.45(95%信頼区間=1.41-4.24、P値=0.0014)であった。

以上より、年齢、ステージを因子に加えた多変量解析においても、リスク遺伝子型保有数およびリスク分類は、独立した予後予測因子であることが示された。

本研究によって、がん関連SNPアレイ解析により、胆嚢がんの生存期間に関連する6個の候補SNPsが同定された。これらのSNP解析と従来の臨床病理学的なリスク因子との併用により、より正確な予後予測や個人に合わせた治療法の選択につながると考えられる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計0件)

6. 研究組織

(1)研究代表者

浅井 孝夫 (ASAI, Takao)

新潟医療福祉大学・医療技術学部・臨床技術学科・講師

研究者番号: 60612736

(2)研究協力者

土屋 康雄 (TSUCHIYA, Yasuo)

新潟大学・医歯学総合研究科・客員研究員
研究者番号: 60334679

Sergio Báez

Hospital Dr. Sótero del Río (Santiago, Chile)

Andrea Tello Martínez

Hospital Dr. Sótero del Río (Santiago, Chile)