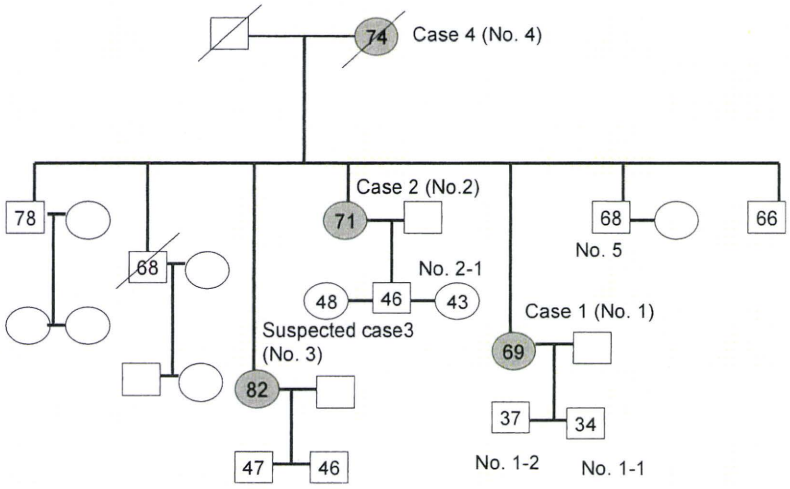


Figure 1. **Family tree. Circles indicate females and squares refer to males. Numbers inside circles and squares indicate age (years) at the time of the present study. Gray symbols indicate those affected by pulmonary MAC, while white symbols refer to subjects unaffected by pulmonary MAC. Diagonal lines drawn through the symbols indicate deceased subjects.**

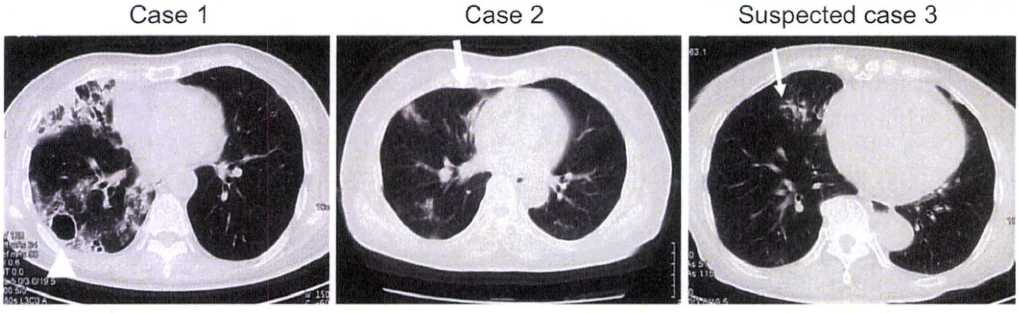


Figure 2. **Computed tomography findings of Case 1, Case 2, and Case 3. Chest computed tomography (CT) scan demonstrates multiple, small, peripheral pulmonary nodules centred on the bronchovascular tree, cylindrical bronchiectasis, a cavitary lesion, and consolidations in Case 1 at the age of 69 years. The cavitary lesion, indicated by the arrowhead, resulted from expansion of the bronchiectatic lesion in Case 1. Multiple, small, peripheral, pulmonary nodules centred on the bronchovascular tree and cylindrical bronchiectasis are seen in Case 2, indicated by the arrow (at the age of 69 years). Nodular and centrilobular granular shadows are observed in Case 3, indicated by the thin arrow (at the age of 82 years). The CT scan findings of Case 1 are more severe than those of both Case 2 and Case 3.**

pected Case 3 (No. 3), 3rd older sister of Case 1; younger brother of Case 1 (No. 5); two sons of Case 1 (No. 1-1 and No. 1-2); and a daughter of Case 2 (No. 2-1). Case 4 (No. 4), Case 1's mother, was deceased at the time of the study, but she had been diagnosed as having pulmonary MAC disease after menopause based on CT findings and sputum culture results.

Bacterial examination and chest CT findings confirmed nodular/bronchiectatic type pulmonary MAC disease in Cases 1 and 2 at 65 years of age. Chest CT showed nodular and centrilobular granular shadows in suspected Case 3 (Fig. 2); sputum cultures were negative, and bronchoscopic washing was not performed. It was thought that she might have pulmonary MAC disease associated with the same familial factors as her sisters and mother.

Delay in the time required for the taste of saccharin particles to be perceived was not observed in Case 1, Case 2, and suspected Case 3. Obstructive defects were not observed on pulmonary function testing, and elevated serum cold agglutinins were also not observed in any of the seven participating family members.

Standard commercial genetic mutation screening of the *CFTR* gene was performed in Case 1 (5); no known mutations and no novel variants were identified. The poly T polymorphisms in intron 8 (IVS8) of the *CFTR* gene were examined in all seven participating members (6). Three polymorphisms of the *NRAMP1* gene, the 5′(GT)n, D543N, and the 3′untranslated region TGTG insertion/deletion, were also genotyped in all seven family members (8). No significant differences between MAC cases and healthy members

Table 1. Characteristics of Individual Members of This Family Line

Family member (No.)	Age (at diagnosis ; years)	Sex	Diagnosis	Complication	HLA-A	*CFTR* polyT	*NRAMP1* D543N	*NRAMP1* TGTG
1 (case 1)	69 (65)	F (post memopausal)	Pulmonary *M.intracellulare*	Sinusitis DM endometriosis	26/26	7/7	G/G	ins/ins
1-1	34	M	-	-	26/24	7/7	G/G	ins/ins
1-2	37	M	-	-	26/24	7/7	G/G	ins/ins
2 (case 2)	71 (65)	F (post menopausal)	Pulmonary *M.avium*	Sinusitis	26/2	7/7	G/G	ins/ins
2-1	43	F (pre menopausal)	-	Sinusitis endometoriosis	26/24	7/7	G/G	ins/ins
3 (case 3)	82 (82)	F (post menopausal)	Suspected pulmonary MAC	Sinusitis	11/2	7/7	G/G	ins/ins
4 (case 4)	74 (N/A)	F (post menopausal)	Pulmonary MAC	N/A	N/A	N/A	N/A	N/A
5	68	M	-	-	26/2	7/7	G/G	ins/ins

HLA: human leukocyte antigen, *CFTR*: cystic fibrosis transmembrane conductance regulator, *NRAMP1*: natural resistance-associated macrophage protein one, MAC: *Mycobacterium avium* complex, DM: diabetes mellitus, N/A: not available, ins: insertion

were observed in the *NRAMP1* and *CFTR* gene polymorphisms. HLA typing showed that HLA-A26 antigen was predominant.

The characteristics of individual family members are shown in Table 1. All female family members had sinusitis, and all patients were postmenopausal females at the time of diagnosis.

Case 1 received combination chemotherapy (clarithromycin (CAM) 600 mg/day, ethambutol (EB) 750 mg/day, rifampin (RFP) 450 mg/day, and streptomycin (SM) 750 mg, 3 times a week for 3 months) because of productive cough. She continued to excrete MAC, and her CT findings worsened gradually despite adding levofloxacin (LVFX) 300 mg/day to her medication (Fig. 2, CT findings at 69 years of age). Case 1 had a history of diabetes mellitus (DM). Combination chemotherapy, including CAM (600 mg/day), EB (750 mg/day), and RFP (300 mg/day), was prescribed in Case 2. Although the patient continued to have positive sputum cultures, her symptoms and chest X-ray and CT findings showed little change with combination therapy. Since suspected case 3 was asymptomatic, she remained untreated. Drug susceptibility testing for *M. intracellulare* showed an MIC of 0.5 μg/mL for CAM in Case 1; in Case 2, the MIC of *M. avium* for CAM was 2.0 μg/mL. Case 1 had more severe symptoms and radiographic findings than Cases 2 and 3 (Fig. 2).

Abbreviation: MIC: minimal inhibitory concentration

Discussion

A familial study of the factors implicated in the development and exacerbation of pulmonary MAC disease, including testing for previously reported disease-susceptibility genes, was conducted in two pulmonary MAC cases, one suspected case, and four healthy members of this family.

Recurrent urinary tract infections occur in many postmenopausal women. In addition to epithelial atrophy, estrogen deficiency can increase vaginal pH and alter the vaginal flora, changes which may predispose to urinary tract infection (12). On the other hand, nodular-bronchiectatic type pulmonary MAC disease is well-known to occur after menopause. Estrogen may play also some role in resistance to MAC infection (13) due to its effects on various immune functions, such as macrophage activity, natural killer cell activity, and T cell-mediated immunity (13). Since all of the patients in this family were postmenopausal females, these factors might have been associated with pulmonary MAC disease onset.

In this family, all females had maxillary sinusitis, suggesting that they had sinobronchial syndrome. Hence, impaired mucociliary clearance may have been associated with pulmonary MAC disease in the females of this family. Major

sinobronchial syndromes are classified into diffuse panbronchiolitis, primary ciliary dyskinesia, and cystic fibrosis. *CFTR* mutations appear to contribute to the susceptibility and pathogenesis of pulmonary nontuberculous mycobacteria (NTM) infection (5-7). CFTR protein is a cyclic adenosine monophosphatase-regulated chloride channel located in the apical membrane of epithelial cells. CFTR deletion impairs mucociliary clearance together with changes in the airway microenvironment, and it leads to a progressive cycle of infection and inflammation (6). Although CF incidence is very low in the Japanese population because of the rare possession of a typical delta F508 mutation of the *CFTR* gene (14), Mai et al reported associations between the IVS8 5T allele in the *CFTR* gene and pulmonary MAC infection in Japanese (6). However, the IVS8 5T alleles were not identified in this family. Also, the full *CFTR* gene analysis on the most severe patient (Case 1) revealed that neither known mutations nor novel variants were present. On the other hand, Ziedalski and Kim et al. reported that the prevalence of *CFTR* mutations in their pulmonary NTM cohort ranged from 36% to 50% in the United States (5, 7). Thus, the genetic background of pulmonary MAC disease may differ between Japanese and Caucasian populations.

In this family line, all seven participating members did not meet the diagnostic criteria for diffuse panbronchiolitis, because of the results of pulmonary function testing and the serum cold agglutinin results. As for primary ciliary dyskinesia, examinations of biopsies with electron microscopy were not performed, because the results of saccharin testing did not fulfill the diagnostic criteria for primary ciliary dyskinesia in Case 1, Case 2, and suspected Case 3.

The murine *Nramp-1* determines susceptibility to intracellular pathogens. The human homologue, *NRAMP1*, might be involved in susceptibility to pulmonary MAC infection (8). All members of this family were homozygotes for major alleles of the D543N and TGTG insertion/deletion polymorphism of the *NRAMP1* gene, which occurs in 70% to 80% of the Japanese population (8). As a result, the effect of these alleles in this family was unclear.

DM might be associated with deterioration of infection because of impaired cell-mediated immunity (15). Case 1, with DM, had more severe symptoms and radiographic findings than Cases 2 and 3, who did not have DM. Therefore, DM might contribute to the severity of pulmonary MAC disease.

HLA plays an important role in immune response; previous studies have shown that NTM lung disease is associated with HLA alleles (9-11). Kubo et al suggested that HLA A-26 antigen might be associated with deterioration of pulmonary MAC infection (9). Case 1, with homozygously expressed HLA-A26, had more severe symptoms and radiographic findings than Case 2, with heterozygously expressed HLA-A26. Case 2 had more severe symptoms and radiographic findings than Case 3 without HLA-A26. Therefore, HLA-A26 might play a role in the deterioration of pulmonary MAC infection.

HLA class I antigens present peptides from inside the cell, including viral peptides if present. These peptides are produced from digested proteins that are broken down in the proteasomes. The peptides are also generally small polymers, about 9 amino acids in length. If foreign peptides are present on cells with HLA class I antigens, cells containing foreign proteins will be attacked by cytotoxic T lymphocytes. In general, HLA class I antigens are expressed on all nucleated cells, while HLA class II antigens are expressed on most immune system cells. HLA class I antigens are classified into A, B, and C. Therefore, it is speculated that certain peptides within the respiratory epithelium presented by HLA-A26 induce HLA-A26-restricted cytotoxic T lymphocytes, and this immune reaction is associated with deterioration of pulmonary MAC disease. Further studies detailing the role of HLA-A26 in the pathogenesis of pulmonary MAC disease are needed.

In conclusion, female sex and menopause might be associated with the onset of nodular/bronchiectatic type pulmonary MAC in this family. HLA-A26 antigen and DM might be involved in disease exacerbation. Thus, it is important to monitor the currently premenopausal daughters of Case 2 because of their similar backgrounds.

References

1. Griffith DE, Aksamit T, Brown-Elliott BA, et al. An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. Am J Respir Crit Care Med **175**: 367-416, 2007.
2. Tanaka E, Kimoto T, Matsumoto H, et al. Familial pulmonary *Mycobacterium avium* complex disease. Am J Respir Crit Care Med **161**: 1643-1647, 2000.
3. Inoue T, Tanaka E, Sakuramoto M, et al. Three sisters of pulmonary *Mycobacterium avium* complex disease. Kansenshogaku Zasshi **79**: 341-347, 2005 (in Japanese).
4. Sexton P, Harrison AC. Susceptibility to nontuberculous mycobacterial lung disease. Eur Respir J **31**: 1322-1333, 2008.
5. Kim RD, Greenberg DE, Ehrmantraut ME, et al. Pulmonary nontuberculous mycobacterial disease: prospective study of a distinct preexisting syndrome. Am J Respir Crit Care Med **178**: 1066-1074, 2008.
6. Mai HN, Hijikata M, Inoue Y, et al. Pulmonary *Mycobacterium avium* complex infection associated with the IVS8-T5 allele of the *CFTR* gene. Int J Tuberc Lung Dis **11**: 808-813, 2007.
7. Ziedalski TM, Kao PN, Henig NR, et al. Prospective analysis of cystic fibrosis transmembrane regulator mutations in adults with bronchiectasis or pulmonary nontuberculous mycobacterial infection. Chest **130**: 995-1002, 2006.
8. Tanaka G, Shojima J, Matsushita I, et al. Pulmonary *Mycobacterium avium* complex infection: association with *NRAMP1* polymorphisms. Eur Respir J **30**: 90-96, 2007.
9. Kubo K, Yamazaki Y, Hanaoka M, et al. Analysis of HLA antigens in *Mycobacterium avium-intracellulare* pulmonary infection. Am J Respir Crit Care Med **161**: 1368-1371, 2000.
10. Takahashi M, Ishizaka A, Nakamura H, et al. Specific HLA in

pulmonary MAC infection in a Japanese population. Am J Respir Crit Care Med **162**: 316-318, 2000.

11. Um SW, Ki CS, Kwon OJ, et al. HLA antigens and nontuberculous mycobacterial lung disease in Korean patients. Lung **187**: 136-140, 2009.

12. Caillouette JC, Sharp CF Jr, Zimmerman GJ, et al. Vaginal pH as a marker for bacterial pathogens and menopausal status. Am J Obstet Gynecol **176**: 1270-1275, 1997.

13. Tsuyuguchi K, Suzuki K, Matsumoto H, et al. Effect of estrogen on *Mycobacterium avium* complex pulmonary infection in mice. Clin Exp Immunol **123**: 428-434, 2001.

14. Yamashiro Y, Shimizu T, Oguchi S, et al. The estimated incidence of cystic fibrosis in Japan. J Pediatr Gastroenterol Nutr **24**: 544-547, 1997.

15. Plouffe JF, Silva J Jr, Fekety R, et al. Cell-mediated immunity in diabetes mellitus. Infect Immun **21**: 425-429, 1978.

Genes and Immunity (2010), 1–7

www.nature.com/gene

ORIGINAL ARTICLE

Association analysis of susceptibility candidate region on chromosome 5q31 for tuberculosis

C Ridruechai[1], S Mahasirimongkol[2,3], J Phromjai[2], H Yanai[2,4], N Nishida[5], I Matsushita[6], J Ohashi[7], N Yamada[2], S Moolphate[2], S Summanapan[8], C Chuchottaworn[9], W Manosuthi[10], P Kantipong[11], S Kanitvittaya[12], P Sawanpanyalert[3], N Keicho[6], S Khusmith[1] and K Tokunaga[5]

[1]*Department of Microbiology and Immunology, Mahidol University, Bangkok, Thailand;* [2]*TB/HIV Research Project, Research Institute of Tuberculosis, Japan Anti-Tuberculosis Association, Chiang Rai, Thailand;* [3]*National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Nonthaburi, Thailand;* [4]*Department of Clinical Laboratory, Fukujuji Hospital, Japan Anti-Tuberculosis Association, Tokyo, Japan;* [5]*Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan;* [6]*Department of Respiratory Diseases, Research Institute, International Medical Center of Japan, Tokyo, Japan;* [7]*Doctoral Programme in Life System Medical Sciences, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Ibaraki, Japan;* [8]*Chiang Rai Provincial Health Office, Office of the Permanent Secretary, Ministry of Public Health, Chiang Rai, Thailand;* [9]*Chest Disease Institute, Department of Medical Services, Ministry of Public Health, Nonthaburi, Thailand;* [10]*Bamrasnaradura Infectious Diseases Institute, Department of Disease Control, Ministry of Public Health, Nonthaburi, Thailand;* [11]*Department of Medicine, Chiang Rai Regional Hospital, Ministry of Public Health, Chiang Rai, Thailand and* [12]*Chiang Rai Regional Medical Sciences Center, Department of Medical Sciences, Ministry of Public Health, Chiang Rai, Thailand*

Chromosome 5q31 spans the T helper (Th) 2-related cytokine gene cluster, which is potentially important in Th1/Th2 immune responses. The chromosome 5q23.2–31.3 has been recently identified as a region with suggestive evidence of linkage to tuberculosis in the Asian population. With the aim of fine-mapping a putative tuberculosis susceptibility locus, we investigated a family-based association test between the dense single nucleotide polymorphism (SNP) markers within chromosome 5q31 and tuberculosis in 205 Thai trio families. Of these, 75 SNPs located within candidate genes covering SLC22A4, SLC22A5, IRF1, IL5, RAD50, IL13, IL4, KIF3A and SEPT8 were genotyped using the DigiTag2 assay. Association analysis revealed the most significant association with tuberculosis in haplotypes comprising SNPs rs274559, rs274554 and rs274553 of SLC22A5 gene ($P_{Global} = 2.02 \times 10^{-6}$), which remained significant after multiple testing correction. In addition, two haplotypes within the SLC22A4 and-KIF3A region were associated with tuberculosis. Haplotypes of SLC22A5 were significantly associated with the expression levels of RAD50 and IL13. The results show that the variants carried by the haplotypes of SLC22A4, SLC22A5 and KIF3A region potentially contribute to tuberculosis susceptibility among the Thai population.
Genes and Immunity advance online publication, 20 May 2010; doi:10.1038/gene.2010.26

Keywords: *chromosome 5q31; family-based association test; Thais; tuberculosis*

Introduction

Tuberculosis is a major burden for global public health, especially in developing countries. Infection with *Mycobacterium tuberculosis* does not always result in tuberculosis; approximately 10% of infected individuals will eventually develop clinical disease during their lifetimes.[1] The outcome of infection is determined by environmental factors, pathogen virulence and host immune response. Human genetic variations involving susceptibility to tuberculosis has initially been shown in twins in whom the concordance rate was higher in monozygotic than in dizygotic twins,[2] although reanalysis of such a survey indicated that environmental factors outweighed the importance of hereditary factors.[3] Several forward genetic approaches, including genome-wide linkage analysis and candidate gene approaches, have been applied in an attempt to map and elucidate the host susceptibility genes that contribute to controlling the immune response against *M. tuberculosis*.

Five genome-wide linkage studies have been performed to date. The first study identified the suggestive evidence of linkage to tuberculosis on chromosomes 15q11–13 and Xq26 in the Gambians and South Africans.[4] Subsequent fine-mapping of 15q11–13 locus in the same ethnicity revealed a 7-bp deletion in the *UBE3A* gene associated with tuberculosis ($P = 0.002$).[5] The positional candidate gene encoding CD40 ligand, *TNFSF5*, located on Xq26.3 was subsequently investigated, but no association with tuberculosis was found.[6] The second two-stage genome scan in Brazilians suggested linkage evidence in three regions 10q26.13, 11q12.3 and 20p12.1.[7] The focused linkage analysis in the candidate region of

Correspondence: Dr K Tokunaga, Department of Human Genetics, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-Ku, Tokyo 113-0033, Japan.
E-mail: tokunaga@m.u-tokyo.ac.jp
Received 30 September 2009; revised 6 January 2010; accepted 14 January 2010

the human chromosome 17q provided weak evidence of linkage in 17q11–21.[8] The strongest evidence of linkage was shown in the genome-wide scan conducted in Moroccan multiplex families,[9] in which a region of chromosome 8q12-13 was mapped as a major susceptibility locus for tuberculosis with non-parametric multipoint likelihood binomial logarithm (base 10) of odds score of 3.49 and a parametric logarithm (base 10) of odds score of 3.38 under a dominant model. The fourth genome-wide linkage study in South Africa and Malawi has implicated regions of 6p21–23 and 20q13.31–33 for tuberculosis susceptibility.[10] Fine-mapping of 20q13.31–33 provided association evidence of melanocortin 3 receptor (*MC3R*) and cathepsin Z (*CTSZ*) genes. Finally, the genome-wide single nucleotide polymorphism (SNP)-based linkage analysis in 93 affected sib-pairs localized a region with suggestive evidence of linkage to chromosome 5q23.2–31.3 among the Thai population (Z_{lr} score 3.01, logarithm (base 10) of odds score 2.29) and two regions on chromosome 17p13.3-13.1 and 20p13-12.3 showed linkage evidence in families with a younger age at the onset of tuberculosis (logarithm (base 10) of odds score 2.57 and 3.33, respectively).[11] Collectively, those findings imply that the allelic architecture of genetic susceptibility to tuberculosis is likely heterogeneous, partly explained by the genetic background of the studied population.

The suggestive evidence of linkage to the 5q region in Asians is of particular interest because the region spans the major T helper (Th) 2 cytokine gene cluster, which is potentially important for Th1/Th2 immune responses. This region has been identified as a susceptibility locus for various infectious and autoimmune diseases. The linkage and allelic association between *Plasmodium falciparum* blood infection levels and markers on chromosome 5q31–33 have also been reported.[12,13] The region has been found to harbor a susceptibility locus (SM1) controlling the infection intensity of *Schistosoma mansoni*.[14] In autoimmune diseases, the gene conferring susceptibility to Crohn's disease has been linked to chromosome 5q31[15] and an attempt to localize this region through linkage disequilibrium (LD) mapping identified a long-range haplotype (~250 kb) that was strongly associated with the disease.[16] The cytokine gene-rich 5q31–33 region in association with susceptibility to other inflammatory conditions had been reported in psoriasis, asthma and/or atopic phenotypes. The polymorphic *IL4* gene on chromosome 5q31 has a linkage and association with atopic dermatitis in Japanese nuclear families,[17] while the 5q31-32 region was confirmed as a susceptible locus for psoriasis.[18] Furthermore, a number of linkage analyses in ethnically diverse populations have identified the linkage signals on chromosome 5q31–33 with total serum IgE,[19] bronchial hyper-responsiveness[20] and asthma-related phenotypes (5q23-31).[21]

The prime candidate genes within this linkage region are a cluster of Th2-related cytokine genes important in immune regulation including *IL4, IL5, IL13, RAD50, KIF3A, SLC22A4* and *SLC22A5*. Interleukin-4 and interleukin-13 coordinated in various immune functions including induction of immunoglobulin class switching, synthesis of IgE and promoting the expression of chemokines, whereas interleukin-5 enhanced the production and survival of eosinophils.[22] Recently, a critical *cis*-regulatory element controlling Th2 cytokine expression has been identified in the 3′end of *Rad50*,[23] which acts as a locus control region for the expression of *IL4* and *IL13* genes. Association analysis on the SNP rs17166050, located in intron 4 of *RAD50*, showed a significant association with Crohn's disease.[24] Thus, the unidentified SNPs in the 5q31 region may influence variation of gene function among populations and contribute to the risk of Crohn's disease and other infections and inflammatory diseases. The activation of Th2 cytokine genes has been suggested to correlate with extensive tuberculosis disease.[25] Th2 cytokines were shown to inhibit autophagy, a dynamic process of subcellular degradation, thus impairing the ability of macrophages to kill mycobacteria.[26] In addition, the hyper-virulent Beijing family of *M. tuberculosis* strains that are predominant in east Asian, preferentially induced human monocytes to express Th2 cytokines.[27,28] Therefore, variations within this Th2 cytokine locus are reasonable candidates that might involve in the immunopathology of tuberculosis.

In this study, we attempted to fine-map a putative tuberculosis susceptibility locus on chromosome 5q31. A panel of SNP markers within chromosome 5q31 was developed and the family-based association between these SNP markers and tuberculosis was analyzed. The results provide evidence on the association of variants in *SLC22A4, SLC22A5* and *KIF3A* with tuberculosis.

Results

Of the 83 SNPs that were analyzed by the PBAT software and Haploview, 6 SNPs with evidence of deviation from the Hardy–Weinberg equilibrium (*P*-value <0.05) and 2 SNPs being monomorphic were excluded. Mendelian inconsistencies within families detected by PedCheck and Haploview were manually deleted. Pedigree errors among family members inspected by the Graphical Representation of Relationships (GRR) program were not observed partly because of the strong linkage disequilibria among the markers that were tested, resulting in none of the families being excluded. After the quality control process, 75 SNPs, which passed the specified criteria, were further analyzed. The genotyped SNPs within the candidate genes located on chromosome 5q31 covering *SLC22A4, SLC22A5, IRF1, IL5, RAD50, IL13, IL4, KIF3A* and *SEPT8* and their physical locations, are described in Supplementary Figure 1 and Supplementary Table 1. Linkage disequilibrium patterns across the 5q31 region were determined and 60 out of 75 SNPs were assigned into the six haplotype blocks (Supplementary Figure 1).

To examine the usefulness of the HapMap data for the association study on the 5q31 region, the patterns of LD in four populations, African (YRI), European (CEU), Chinese (CHB) and Japanese (JPT), were compared using genotype data from the HapMap and Thai trio families. In agreement with the earlier report,[29] we confirmed the similarity of LD patterns among the Chinese, Japanese and Thais in this 5q31 region.

Single marker and haplotype analyses

The associations of SNPs and haplotypes with tuberculosis under four different genetic models, additive,

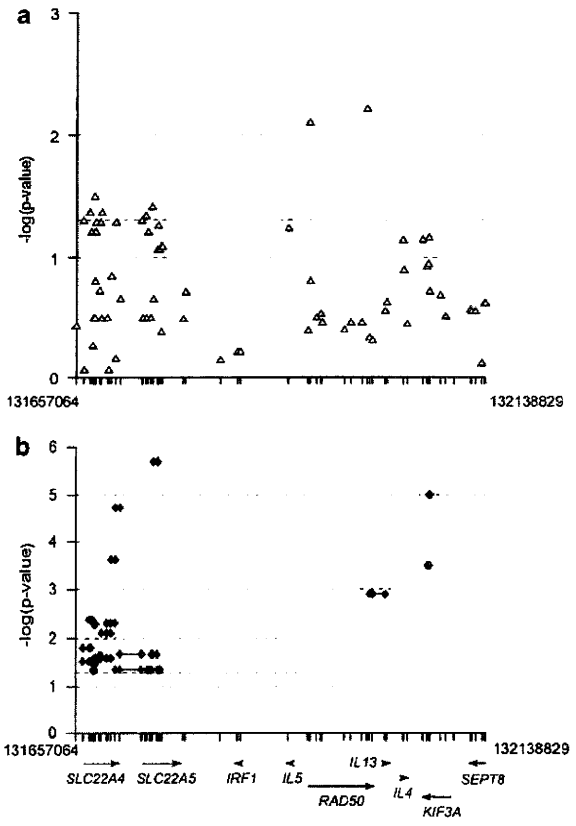


Figure 1 Association plots between single nucleotide polymorphisms (SNPs) located on chromosome 5q31 (NCBI Build 36) and tuberculosis. (**a**) The −log (*P*-values) plots for single-SNP association analysis were obtained from the best genetic model. (**b**) The global *P*-value plots for overall haplotype association analysis presented when $P < 0.05$. The filled symbols connected by solid lines indicated the haplotype.

dominant, recessive and heterozygous advantage, were tested using the PBAT software. The SNPs with a *P*-value of <0.05 are detailed in Supplementary Table 2. Under the additive genetic model, the strongest association with tuberculosis for an individual marker, rs2237060, located on the *RAD50*, was shown ($P = 0.0061$) (Figure 1a). Allele A of this A/C SNP was transferred from parents to offspring with tuberculosis. On the basis of the significant threshold at 1.74×10^{-4} estimated by taking into account the effective number of tests for single SNP analysis, the number of sliding windows for haplotype analysis and the number of genetic models included, this single-marker association evidence was not sufficient to reject the null hypothesis of no association with tuberculosis.

The most significant association with tuberculosis was observed in haplotype block 1 ($P_{Global} = 2.02 \times 10^{-6}$) under the dominant model of inheritance as shown by association analysis of a three-marker haplotype present in each haplotype block (Figure 1b and Table 1). This haplotype is located within the *SLC22A5* gene comprising rs274559, rs274554 and rs274553 (Table 1). Significant global associations with tuberculosis were also detected in haplotypes of *SLC22A4* and *KIF3A* genes (rs272873-rs2306772-rs272867 and rs2299007-rs2237057-rs2299006, respectively) under the heterozygous advantage model of inheritance (Table 1). They remained significant after multiple testing correction (P-value $< 1.74 \times 10^{-4}$).

Association analysis of the imputed untyped SNPs

A total of 895 imputed SNPs, including SNPs having association with Crohn's disease and psoriasis, were tested for association with tuberculosis. As expected, the imputation-based SNP analysis provided lower power to detect the disease association than those of typed SNPs. None of the imputed SNPs provided association evidence with tuberculosis after multiple testing corrections (Supplementary Table 3). However, the results of the imputed genotype and their frequencies can provide draft pictures of the association patterns between untyped SNPs and tuberculosis in the Thai population.

Effect of the associated haplotype on gene expression levels

To evaluate the possible effect of the significantly associated haplotype (rs274559, rs274554 and rs274553) on the expression of the neighboring genes, the UNPHASED software was used to test the association between this haplotype and the expression levels of *SLC22A4, SLC22A5, IRF1, IL5, RAD50, IL13, IL4, KIF3A, SEPT8 and PCDH1* derived from genomic expression microarray analysis of four populations, CHB, JPT, CEU and YRI, in the HapMap database. The haplotype was significantly associated with the *RAD50*, *IL13* and *PCDH1* expression (overall P-value $= 6.14 \times 10^{-4}$, 6.70×10^{-4} and 1.90×10^{-4}, respectively) and remained significant after multiple testing correction (P-value <0.005 for 10 tests performed). Different expression levels of *RAD50*, *IL13* and *PCDH1* in various combinations of three SNP genotypes were shown (Supplementary Figure 2, 3 and 4). However, when this quantitative-trait locus analysis was restricted to only CHB + JPT, it became non-significant because of the limited number of haplotypes in each population. In addition, the frequencies of this significantly associated haplotype are quite similar among the CHB, JPT and Thais, but are totally different from those in the CEU/YRI (data not shown).

Discussion

Chromosome 5q31 has been suggested as the susceptibility loci of related inflammatory diseases in ethnically diverse populations, including tuberculosis among the Thais.[11] The region has been of particular interest because it carries a cluster of Th2 cytokines and related genes including *IL4, IL5, IL13, RAD50, KIF3A, SLC22A4* and *SLC22A5*. Such results have allowed us to conduct fine-mapping of the tuberculosis susceptibility locus in this region in Thai trio families. Here we report a nominal association with tuberculosis for individual SNPs located in *SLC22A4, SLC22A5* and *RAD50* genes, although none of them remained statistically significant after multiple testing correction. The failure to provide a significant association could be explained by small sample size resulting in insufficient power to detect genetic risks with small effect size. Moreover, the presenting linkage may be the results of the multiple causative alleles within the same linkage region as shown in Crohn's disease.[30] Therefore, replication in an

Table 1 Association between three-marker haplotypes and tuberculosis

SNP ID	*Gene symbol*	*Haplotype*	*Haplotype frequency*	*Fam#*	*I_P-value*[a] *(FBAT)*	*Power (FBAT)*	*G_P-value*[b] *(WaldI)*
rs272879-rs272873-rs2306772[c]	*SLC22A4*	GCG	0.3584	120	+0.8111	0.0779	2.41×10^{-4}
		CCG	0.1202	70	−0.0435	0.1854	
		GTG	0.0108	16	−0.4926	0.0500	
		CTG	0.1729	85	−0.7681	0.8788	
		GCA	0.3136	106	−0.0054	0.6161	
		CCA	0.0111	16	−0.3867	0.0500	
rs272873-rs2306772-rs272867[c]	*SLC22A4*	CGC	0.3644	123	+0.6783	0.0601	1.86×10^{-5}
		TGC	0.0102	13	+0.6941	0.0500	
		CAC	0.3155	105	−0.0070	0.5698	
		CGT	0.1141	65	−0.0631	0.1245	
		TGT	0.1729	87	−0.8448	0.8581	
		TAT	0.0100	14	−0.0758	0.8105	
rs274559-rs274554-rs274553[d]	***SLC22A5***	CGG	0.6738	50	−0.7942	0.2869	$\mathbf{2.02 \times 10^{-6}}$
		TGG	0.1987	91	−0.9387	0.0686	
		CAC	0.0177	15	+0.7380	0.1578	
		TAC	0.0993	64	−0.0751	0.2350	
rs3798132-rs2299007-rs2237057[c]	*KIF3A*	ATC	0.3551	112	−0.6137	0.0500	3.22×10^{-4}
		GTC	0.0454	31	−0.6748	0.0977	
		ACC	0.3932	101	+0.1689	0.4564	
		GTT	0.1828	78	−0.1768	0.1451	
rs2299007-rs2237057-rs2299006[c]	*KIF3A*	TCC	0.3482	114	−0.4487	0.0500	1.00×10^{-5}
		CCC	0.3870	103	+0.0767	0.3888	
		TCG	0.0542	37	−0.7473	0.0883	
		CCG	0.0135	12	+0.9724	0.1874	
		TTG	0.1802	77	−0.1799	0.1102	

Abbreviations: Fam#, number of informative families; FBAT, family-based association test; G_*P*-value, global *P*-value; I_*P*-value, individual *P*-value; SNP, single-nucleotide polymorphism. The most significant association with tuberculosis is shown in bold.

[a]Negative/positive signs represent the direction of association between individual haplotype and tuberculosis.

[b]Uncorrected *P*-value.

[c]Haplotype carried under a heterozygous advantage model of inheritance.

[d]Haplotype carried under a dominant model of inheritance.

independent population and study in a larger number of samples are needed to prove the association.

The observed associations of three-marker haplotype comprising SNPs of *SLC22A4*, *SLC22A5* and *KIF3A* with tuberculosis, which remained significant after multiple testing correction, suggested these are the reasonable candidate genes contributing to tuberculosis susceptibility. We avoid the false positives from population stratification by the family-based association study in which the majority of samples are complete trios. Despite the significant associations observed in this study, without replication study in other populations, false-positive findings cannot be refused at the moment.

Polymorphisms in *SLC22A4* and *SLC22A5* have been suggested to be associated with Crohn's disease in the Caucasians but the variants are extremely rare in Asians.[24,31] It is possible that the same locus can be involved in susceptibility to several related phenotypes as reported in this study on different variants in *SLC22A* and *SLC22A5* contributing to tuberculosis susceptibility. The genomic neighborhood surrounding the Th2 cytokine cluster is particularly important for their gene expression or activities.[22,23] Three SNPs comprising the most significant haplotype of *SLC22A5* are located in the intron and the bioinformatics prediction of these SNPs suggested their roles in gene expression. In light of this prediction, the quantitative association of this haplotype and the expression level of genes within the Th2 cytokine cluster region were determined and the results showed that this haplotype is associated with the expression levels of *IL13* and *RAD50*. As *RAD50* contains the locus control region regulating the Th2 cytokine expression, including *IL4* and *IL13*, it is possible that variation in *RAD50* could indirectly influence the fate or magnitude of T-cell responses to infectious agents such as *M. tuberculosis*. In addition, substantial evidence indicated that increased interleukin-4 production in tuberculosis is likely associated with immunopathology.[25,32]

The associated haplotype encompassing the *SLC22A4* ergothioneine transporter, *SLC22A5* L-carnitine transporter, and the *KIF3A* gene encoding for a subunit of kinesin II, a microtubule-based motility protein,[33] suggested their role in disease susceptibility. However, in the absence of a validated biological function in tuberculosis and with regard to the heterozygous advantage model for the associations of *SLC22A4* and *KIF3A* haplotypes, it could be speculated that these haplotypes might not have true biological effects, but they are proxies to primary, but untyped, markers nearby. The heterozygous advantage model of inheritance has been described in a number of genes including the *SLC6A4* serotonin transporter gene.[34] However, as heterosis at a molecular level seems contrary to the expectation for the degree of phenotypic effect, the presence of heterosis in *SLC22A4* and *KIF3A* genes remains to be clarified.

In summary, the observations from fine-mapping and gene expression analysis enable us to confine a region of potential risk markers and haplotypes for tuberculosis within chromosome 5q31. The variants in *SLC22A4*, *SLC22A5*, *KIF3A* or flanking genes, especially genes

containing master regulatory elements for Th2 cytokine expression, may have an important role in susceptibility to tuberculosis. However, to define the causative variants, replication studies in other Asian populations and/or functional analysis of associated variants require further investigations.

Materials and methods

Sample description

Of the 203 complete trios (parents and their affected offspring) and two incomplete trio families (father or mother and affected offspring) were recruited through the tuberculosis surveillance system in Chiang Rai, the northernmost province in Thailand. Of the 203 complete trio families, 15 were part of multi-case families in a previous linkage study.[11] Tuberculosis was diagnosed on the basis of clinical presentation and bacteriological confirmation by sputum culture for *M. tuberculosis* or at least two out of three positive sputum smears for acid-fast bacilli. Diagnosis of tuberculosis in a few cases was obtained from records in the tuberculosis surveillance system accompanied by a current abnormal chest x-ray. This study was reviewed and approved by the Ethics Review Committee for Research in Human Subjects, Ministry of Public Health, Thailand and the Institutional Review Board of the International Medical Center of Japan. All patients were tested for HIV using standard serological tests. HIV-infected patients were excluded from this study. Their blood samples were collected after obtaining individual informed consent.

Genotyping and quality control

Genomic DNA was extracted from the whole blood or peripheral blood mononuclear cells using QIAamp blood midi kit (Qiagen, Hilden, Germany). The tagging SNPs were randomly selected from a 610-kb region of 5q31–33. These markers were originally chosen for a fine-mapping study for various immune disorders associated with the 5q31 region, before the release of the HapMap data. Genotyping of 96 SNPs within chromosome 5q31 was performed by Digitag2 assay described earlier.[35] The genotype calls were determined by the visual inspection of cluster plot from the SNPStar software (version 0.0.0.8, Olympus, Tokyo, Japan). The successfully typed SNPs were analyzed by Haploview version 4.1.[36] SNP genotypes of the parent data set showing significant deviation from the Hardy–Weinberg equilibrium (P-value <0.05) and being monomorphic were excluded from further association test. Mendelian inconsistency within families was detected by PedCheck[37] and Haploview version 4.1. Genotyping errors within the family were deleted manually.

Relationship determination

Pedigree error should be suspected when a number of Mendelian inconsistencies are observed more often than expected by genotyping errors. Relatedness among family members was evaluated by Graphical Representation of Relationships (GRR) program.[38] As the studied markers are highly correlated, genotype information alone does not have the power to discriminate between pedigree error and genotyping error.

Family-based association study

Family-based association tests were analyzed by the PBAT software version 3.6.[39] Single marker and haplotype analyses were performed based on the conditional power calculations. Statistical parameters for testing the null hypothesis of no linkage and no association were family-based association test-GEE and four genetic models of inheritance, namely additive, dominant, recessive and heterozygous advantage models. The plots of uncorrected P-values from the association tests were created using snp.plotter.[40]

Bonferroni correction is overly conservative in this fine-mapping analysis because of high linkage disequilibrium patterns between the markers that were tested. Thus, SNP spectral decomposition[41] was used to calculate the effective number of tests that should be corrected for single SNP analyses and the P-value threshold for rejecting null hypothesis accounting for their highly correlated allelic structure. The experiment-wide significance threshold required to keep type I error rate at 5% from SNP spectral decomposition was 0.0021. However, this threshold did not consider the number of sliding windows for haplotype tests and did not allow correction for the various genetic models tested. Taking into account the 24 effective numbers of tests corrected for single SNP analysis, 48 three-marker sliding windows calculated from 60 SNPs in the defined haplotype blocks and four genetic models tested, 288 tests were performed. This resulted in threshold to maintain the type I error rate at 5% of $P<1.74\times10^{-4}$ for the multiple testing correction in this study.

The structures of LD determined by Haploview and the pairwise LD among genotyped SNPs were measured by r^2 values. Haplotype blocks were defined based on the method described earlier[42] and the individual and overall family-based haplotype association analyses were then performed according to the defined haplotype blocks. The LD structure in the YRI, CEU, CHB and JPT was determined using genotype data from the HapMap web site (http://www.hapmap.org).

Imputation of untyped SNPs in trio samples and the HapMap reference samples

Imputation was carried out by the PLINK version 1.04 based on the HapMap genotype data of 90 JPT + CHB individuals available on the PLINK website (http://pngu.mgh.harvard.edu/purcell/plink/).[43] Individual genotype data were in forward orientation with the NCBI Build 36 reference. To prevent the problem of strand flipping in C/G and A/T SNPs, the genotyped markers with these allele configurations were removed from the input data. Removing these SNPs resulted in 54 genotyped SNPs in the trio families used in the imputation. Imputation of untyped SNPs in the trio families was based on 895 SNPs extending from 131.5 to 132.3 Mb within the HapMap individual genotype data. The imputed data set was used for empirical assessment of allele frequency of untyped SNPs and the family-based association analysis by PLINK. The association analysis of imputed untyped markers should be considered as preliminary and results having lower power and accuracy than those of typed markers were expected. We aimed to use this imputed information to provide the empirical information for the untyped SNPs within the

5q31 region that is associated with other related diseases including Crohn's disease and psoriasis.

Bioinformatics prediction of SNPs' functions
The functional SNP (F-SNP) database[44] (http://compbio.cs.queensu.ca/F-SNP/) was accessed to retrieve the functional information of SNPs that showed the lowest *P*-value in the association analysis.

Analysis of SNP effects on gene expression
The expression data of the genes on chromosome 5q31 covering *SLC22A4, SLC22A5, IRF1, IL5, RAD50, IL13, IL4, KIF3A, SEPT8 and PCDH1* were retrieved from the GENEVAR web site (http://www.sanger.ac.uk/humgen/genevar/). These normalized data sets of gene expression were produced from genome-wide expression arrays (Sentrix Human-6 Expression BeadChip, Illumina Inc., San Diego, CA, USA) in the Epstein–Barr virus-transformed lymphoblastoid cell lines of 150 offspring in phase I and phase II of the HapMap Consortium (populations: YRI, CEU, CHB and JPT). The haplotype frequencies among these populations were also determined with Haploview version 4.1. The haplotypes showing a significant association with tuberculosis were assessed for the effects on gene expression using UNPHASED version 3.0.13.[45] The box plots representing normalized expression levels of genes in various combinations of three-SNP genotypes were created with R software version 2.7.2 (http://www.r-project.org/).

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

This study was partly supported by International Cooperation Research grant, the Ministry of Health, Labor and Welfare from 2002 to 2004 and by a grant-in-aid for scientific research on priority areas 'Comprehensive Genomics' from the Ministry of Education, Culture, Sports, Science and Technology of Japan and intramural grant from Department of Medical Sciences, Ministry of Public Health, Thailand. We thank all staffs and collaborators of the TB/HIV research project, Thailand, a collaborative research project between the Research Institute of Tuberculosis (RIT) and the Japan Anti-tuberculosis Association, and the Department of Medical Sciences, Ministry of Public Health for collecting the clinical information and samples.

References

1 Bloom BR, Small PM. The evolving relation between humans and *Mycobacterium tuberculosis*. *N Engl J Med* 1998; **338**: 677–678.
2 Comstock GW. Tuberculosis in twins: a re-analysis of the Prophit survey. *Am Rev Respir Dis* 1978; **117**: 621–624.
3 van der Eijk EA, van de Vosse E, Vandenbroucke JP, van Dissel JT. Heredity versus environment in tuberculosis in twins: the 1950s United Kingdom Prophit Survey Simonds and Comstock revisited. *Am J Respir Crit Care Med* 2007; **176**: 1281–1288.
4 Bellamy R, Beyers N, McAdam KP, Ruwende C, Gie R, Samaai P *et al*. Genetic susceptibility to tuberculosis in Africans: a genome-wide scan. *Proc Natl Acad Sci USA* 2000; **97**: 8005–8009.
5 Cervino AC, Lakiss S, Sow O, Bellamy R, Beyers N, Hoal-van Helden E *et al*. Fine mapping of a putative tuberculosis-susceptibility locus on chromosome 15q11-13 in African families. *Hum Mol Genet* 2002; **11**: 1599–1603.
6 Campbell SJ, Sabeti P, Fielding K, Sillah J, Bah B, Gustafson P *et al*. Variants of the CD40 ligand gene are not associated with increased susceptibility to tuberculosis in West Africa. *Immunogenetics* 2003; **55**: 502–507.
7 Miller EN, Jamieson SE, Joberty C, Fakiola M, Hudson D, Peacock CS *et al*. Genome-wide scans for leprosy and tuberculosis susceptibility genes in Brazilians. *Genes Immun* 2004; **5**: 63–67.
8 Jamieson SE, Miller EN, Black GF, Peacock CS, Cordell HJ, Howson JM *et al*. Evidence for a cluster of genes on chromosome 17q11-q21 controlling susceptibility to tuberculosis and leprosy in Brazilians. *Genes Immun* 2004; **5**: 46–57.
9 Baghdadi JE, Orlova M, Alter A, Ranque B, Chentoufi M, Lazrak F *et al*. An autosomal dominant major gene confers predisposition to pulmonary tuberculosis in adults. *J Exp Med* 2006; **203**: 1679–1684.
10 Cooke GS, Campbell SJ, Bennett S, Lienhardt C, McAdam KP, Sirugo G *et al*. Mapping of a novel susceptibility locus suggests a role for MC3R and CTSZ in human tuberculosis. *Am J Respir Crit Care Med* 2008; **178**: 203–207.
11 Mahasirimongkol S, Yanai H, Nishida N, Ridruechai C, Matsushita I, Ohashi J *et al*. Genome-wide SNP-based linkage analysis of tuberculosis in Thais. *Genes Immun* 2009; **10**: 77–83.
12 Rihet P, Traore Y, Abel L, Aucan C, Traore-Leroux T, Fumoux F. Malaria in humans: *Plasmodium falciparum* blood infection levels are linked to chromosome 5q31-q33. *Am J Hum Genet* 1998; **63**: 498–505.
13 Flori L, Kumulungui B, Aucan C, Esnault C, Traore AS, Fumoux F *et al*. Linkage and association between *Plasmodium falciparum* blood infection levels and chromosome 5q31-q33. *Genes Immun* 2003; **4**: 265–268.
14 Marquet S, Abel L, Hillaire D, Dessein A. Full results of the genome-wide scan which localises a locus controlling the intensity of infection by *Schistosoma mansoni* on chromosome 5q31-q33. *Eur J Hum Genet* 1999; **7**: 88–97.
15 Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, McLeod RS, Griffiths AM *et al*. Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci. *Am J Hum Genet* 2000; **66**: 1863–1870.
16 Rioux JD, Daly MJ, Silverberg MS, Lindblad K, Steinhart H, Cohen Z *et al*. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat Genet* 2001; **29**: 223–228.
17 Kawashima T, Noguchi E, Arinami T, Yamakawa-Kobayashi K, Nakagawa H, Otsuka F *et al*. Linkage and association of an interleukin 4 gene polymorphism with atopic dermatitis in Japanese families. *J Med Genet* 1998; **35**: 502–504.
18 Friberg C, Bjorck K, Nilsson S, Inerot A, Wahlstrom J, Samuelsson L. Analysis of chromosome 5q31-32 and psoriasis: confirmation of a susceptibility locus but no association with SNPs within SLC22A4 and SLC22A5. *J Invest Dermatol* 2006; **126**: 998–1002.
19 Marsh DG, Neely JD, Breazeale DR, Ghosh B, Freidhoff LR, Ehrlich-Kautzky E *et al*. Linkage analysis of IL4 and other chromosome 5q31.1 markers and total serum immunoglobulin E concentrations. *Science* 1994; **264**: 1152–1156.
20 Postma DS, Bleecker ER, Amelung PJ, Holroyd KJ, Xu J, Panhuysen CI *et al*. Genetic susceptibility to asthma–bronchial hyperresponsiveness coinherited with a major gene for atopy. *N Engl J Med* 1995; **333**: 894–900.

21 A genome-wide search for asthma susceptibility loci in ethnically diverse populations. The Collaborative Study on the Genetics of Asthma (CSGA). *Nat Genet* 1997; **15**: 389–392.
22 Li-Weber M, Krammer PH. Regulation of IL4 gene expression by T cells and therapeutic perspectives. *Nat Rev Immunol* 2003; **3**: 534–543.
23 Lee GR, Fields PE, Griffin TJ, Flavell RA. Regulation of the Th2 cytokine locus by a locus control region. *Immunity* 2003; **19**: 145–153.
24 Onnie C, Fisher SA, King K, Mirza M, Roberts R, Forbes A *et al.* Sequence variation, linkage disequilibrium and association with Crohn's disease on chromosome 5q31. *Genes Immun* 2006; **7**: 359–365.
25 Seah GT, Scott M, Rock GAW. Type 2 cytokine gene activation and its relationship to extend of disease in patients with tuberculosis. *J Infect Dis* 2000; **181**: 385–389.
26 Harris J, De Haro SA, Master SS, Keane J, Roberts EA, Delgado M *et al.* T helper 2 cytokines inhibit autophagic control of intracellular *Mycobacterium tuberculosis*. *Immunity* 2007; **27**: 505–517.
27 van Soolingen D, Qian L, de Haas PE, Douglas JT, Traore H, Portaels F *et al.* Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of East Asia. *J Clin Microbiol* 1995; **33**: 3234–3238.
28 Manca C, Reed MB, Freeman S, Mathema B, Kreiswirth B, Barry III CE *et al.* Differential monocyte activation underlies strain-specific *Mycobacterium tuberculosis* pathogenesis. *Infect Immun* 2004; **72**: 5511–5514.
29 Nuchnoi P, Ohashi J, Naka I, Nacapunchai D, Tokunaga K, Nishida N *et al.* Linkage disequilibrium structure of the 5q31-33 region in a Thai population. *J Hum Genet* 2008; **53**: 850–856.
30 Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD *et al.* Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008; **40**: 955–962.
31 Peltekova VD, Wintle RF, Rubin LA, Amos CI, Huang Q, Gu X *et al.* Functional variants of OCTN cation transporter genes are associated with Crohn disease. *Nat Genet* 2004; **36**: 471–475.
32 van Crevel R, Karyadi E, Preyers F, Leenders M, Kullberg BJ, Nelwan RH *et al.* Increased production of interleukin 4 by CD4+ and CD8+ T cells from patients with tuberculosis is related to the presence of pulmonary cavities. *J Infect Dis* 2000; **181**: 1194–1197.
33 Vale RD. The molecular motor toolbox for intracellular transport. *Cell* 2003; **112**: 467–480.
34 Little KY, McLaughlin DP, Zhang L, Livermore CS, Dalack GW, McFinton PR *et al.* Cocaine, ethanol, and genotype effects on human midbrain serotonin transporter binding sites and mRNA levels. *Am J Psychiatr* 1998; **155**: 207–213.
35 Nishida N, Tanabe T, Takasu M, Suyama A, Tokunaga K. Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. *Anal Biochem* 2007; **364**: 78–85.
36 Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.
37 O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998; **63**: 259–266.
38 Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. *Bioinformatics* 2001; **17**: 742–743.
39 Lange C, DeMeo D, Silverman EK, Weiss ST, Laird NM. PBAT: tools for family-based association studies. *Am J Hum Genet* 2004; **74**: 367–369.
40 Luna A, Nicodemus KK. snp.plotter: an R-based SNP/haplotype association and linkage disequilibrium plotting package. *Bioinformatics* 2007; **23**: 774–776.
41 Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 2004; **74**: 765–769.
42 Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B *et al.* The structure of haplotype blocks in the human genome. *Science* 2002; **296**: 2225–2229.
43 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
44 Lee PH, Shatkay H. F-SNP: computationally predicted functional SNPs for disease association studies. *Nucleic Acids Res* 2008; **36**: D820–D824.
45 Dudbridge F. Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum Hered* 2008; **66**: 87–98.

Supplementary Information accompanies the paper on Genes and Immunity website (http://www.nature.com/gene)

J Gastroenterol (2011) 46:339–349
DOI 10.1007/s00535-010-0355-9

ORIGINAL ARTICLE—ALIMENTARY TRACT

CITED2 is activated in ulcerative colitis and induces p53-dependent apoptosis in response to butyric acid

Tsutomu Yoshida · Tsukasa Sekine · Ken-ichi Aisaki · Tetuo Mikami · Jun Kanno · Isao Okayasu

Received: 13 June 2010 / Accepted: 21 November 2010 / Published online: 17 December 2010

Abstract
Background In ulcerative colitis (UC), *Fusobacterium varium* is significantly detected in patients' mucosa, and butyric acid (BA), abundantly produced by the bacterium, activates the p53 system and induces epithelial apoptosis, as we previously reported. However, factors active in the link between BA and p53 have yet to be clarified. Here, we identified a gene activated by BA specifically in UC-associated cancer cell lines and ascertained the mechanism of its activation of p53.
Methods cDNA microarray analysis based on the Percellome (per cell normalization) method was performed on BA-stimulated UC-associated cancers and sporadic colorectal cancer cell lines under conditions mimicking colonic epithelium UC. For validation of microarray results, molecular, biochemical, and histopathological analyses were performed.
Results We found the CBP/p300-interacting transactivator with glutamic acid/asparagine-rich carboxy-terminal domain 2 (CITED2) to be specifically upregulated in UC-associated cancer cell lines by BA treatment, at both mRNA and protein expression levels. CITED2 could be shown to induce p53 acetylation and p53-dependent apoptosis, accompanied by binding of CBP/p300. BA-dependent apoptosis was suppressed by an inhibitor of monocarboxylate transporter-1 and an siRNA for p53. In inflammatory foci of UC, histologically evident inflammatory activity and CITED2 expression were significantly correlated.
Conclusions CITED2 was identified as UC-associated protein by cDNA microarray based on the Percellome method under UC-mimicking conditions in vitro. CITED2 activation may induce mucosal apoptosis and erosion by activating p53 and thus play a critical role in linking enteric bacteria with mucosal inflammation in UC.

Keywords Ulcerative colitis · Inflammation · Apoptosis · CITED2 · Butyric acid · p53 · CBP/p300

The GeneChip data have been deposited in the Percellome project site (http://www.nihs.go.jp/tox/TtgPublication.htm) and are accessible with no restriction.

Electronic supplementary material The online version of this article (doi:10.1007/s00535-010-0355-9) contains supplementary material, which is available to authorized users.

T. Yoshida (✉) · T. Sekine · T. Mikami · I. Okayasu
Department of Pathology, Kitasato University School of Medicine, 1-15-1 Kitasato, Minami-ku, Sagamihara, Kanagawa 252-0374, Japan
e-mail: tyoshida@med.kitasato-u.ac.jp

K. Aisaki · J. Kanno
Division of Cellular and Molecular Toxicology, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-5801, Japan

Abbreviations

BA	Butyric acid
CITED2	CBP/p300-interacting transactivator with glutamic acid/asparagine-rich carboxy-terminal domain 2
4CHC	4-Hydroxycinnamate
IL-10	Interleukin-10
MCT1	Monocarboxylate transporter-1
q-RT-PCR	Quantitative reverse transcription PCR
sCRC	Sporadic colorectal cancer
UC	Ulcerative colitis
UCCA	Ulcerative colitis-associated cancer

Introduction

Ulcerative colitis (UC) is an intractable inflammatory bowel disease whose etiology is not known in detail. With long-standing UC, adenocarcinomas tend to occur in the colon, which might be due to repeated episodes of active chronic inflammation and remission [1, 2]. In a murine model system, dextran sulfate sodium (DSS) without mutagenicity induces UC-like colitis [3], and long-term repeated treatment with DSS can induce colitic cancers [4, 5]. Therefore, UC can be considered as a model disease for carcinogenesis on a background of chronic inflammation, and elucidation of the etiology of UC is important for treatment and prevention purposes.

The bacterial flora seem to constitute an important factor for triggering inflammation in UC, in accordance with previous reports concerning spontaneous development of UC-like colitis in IL-10 knockout mice under conventional breeding but not germ-free conditions [6–8]. Previously, we detected *Fusobacterium varium* (*F. varium*) by direct culture of mucosal biopsy specimens of UC patients and found intraepithelial invasion of *F. varium* to result in production of inflammatory cytokines in colonic epithelial cell lines in vitro [9, 10]. *F. varium* produces a short-chain fatty acid, butyric acid (BA), and direct luminal injection of BA into the rectum was found to induce apoptosis and UC-like lesions in mice [11]. In addition, BA activates p53-dependent apoptosis and DNA repair in vitro, and activation of p53 by its phosphorylation has been observed in situ in inflammatory foci of UC patients [12]. Clinically, we have shown significant improvement of UC by antibiotic combination therapy targeting *F. varium* [13, 14]. Recently, this was confirmed by a double-blind, placebo-controlled, multicenter trial [15]. However, the molecular mechanisms of BA-mediated p53 activation are unknown and largely remain to be elucidated.

In order to clarify this issue, we compared the genetic profiles of UC-associated carcinoma (UCCA) and sporadic colorectal adenocarcinoma (sCRC) cell lines undergoing BA treatment. Both series of cell lines were established in our laboratory [16]. The former include both p53 wild and mutant cells, whereas the latter are p53 wild. BA upregulates early response genes especially in UCCA cell lines, acting upstream of p53. Using whole genome cDNA microarray chips and the "per cell" normalization method (Percellome method) [17], we here identified genes specifically upregulated in UCCA lines with BA treatment. By detailed validation and the further examination, we focused on one upregulated gene, encoding the CBP/p300-interacting transactivator with glutamic acid/asparagine-rich carboxy-terminal domain 2 (CITED2) [18]. The protein itself activates p53 and its downstream molecules and accelerates apoptosis, and histological analysis revealed unique expression in the colonic mucosa of UC patients. Thus, we consider that CITED2 is an important and specific mediator of inflammation in UC, and might thus serve as a target of therapy.

Materials and methods

Cell lines and clinical cases

The human UCCA cell lines UCCA21, UCCA22, UCCA23, UCCA24, UCCA25 (p53 wild), and UCCA3 (p53 mutant), and the sCRC cell lines KE24, KE43w, and KE43p (p53 wild) were established in our laboratory [16], and cultured in Dulbecco's Modified Eagle Medium (GIBCO, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum containing penicillin and streptomycin (GIBCO) at 37°C in 5% CO_2.

Microarray and data analysis

GeneChip analysis was conducted according to the Percellome method [17, 19] described in the Electronic Supplementary Material in detail. The GeneChip data have been deposited in the Percellome project site (http://www.nihs.go.jp/tox/TtgPublication.htm) and are accessible with no restriction. Differentially expressed genes were extracted by Fx algorithm [20], described in the Supplementary Methods online.

q-RT-PCR and western blotting

Total RNA was isolated from cells with an RNeasy kit (QIAGEN, Hilden, Germany), and 2-μg aliquots were used to synthesize cDNAs with a High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA). Pre-amplified DNAs generated with a TaqMan PreAmp Master Mix Kit (Applied Biosystems) were used as templates for quantitative PCR with TaqMan Gene Expression Master Mix (Applied Biosystems) and FAM-dye-labelled TaqMan probe and primer sets for β-actin, CITED2, p53, CDKN1A (p21), Bax, MCT1, p53R2 (Applied Biosystems) using the 7500 real-time PCR system (Applied Biosystems). Quantitative PCR was performed as multiplex PCR with a VIC-dye-labelled β-actin probe and the fold ratio against VIC-labeled β-actin was calculated. All the quantitative RT-PCR (q-RT-PCR) studies were performed in triplex, and representative results for three independent experiments are shown.

Aliquots of 2×10^6 UCCA24 or KE43p cells after various treatments were lysed with lysis buffer (20 mM Tris–HCl, pH 7.6, 170 mM NaCl, 1 mM EDTA, 0.5% NP-40, 1 mM dithiothreitol). After separation by 12% SDS-

polyacrylamide gel electrophoresis and protein transfer to polyvinyl membranes, blocking with 5% bovine serum albumin was performed. Antibodies for CITED2 (Novous Biologicals, Inc Littleton, CO, USA), p53 (Novocastra Laboratories, Newcastle upon Tyne, UK), acetylated-p53 (Bio Academia, Osaka, Japan), $p21^{Cip1}$ (Santa Cruz Biotechnology, Santa Cruz, CA, USA), Bax (Transduction Laboratories, Lexington, KY, USA), and CBP/p300 (Santa Cruz) were used as first antibodies for western blotting. HRP-conjugated anti-mouse or HRP-conjugated anti-rabbit antibodies were then employed as secondary antibodies with ECL Western Blotting Detection Reagents (GE Healthcare, Buckinghamshire, UK) for visualization. Densitometry was performed by using ImageJ software (National Institutes of Health, USA)

Transfection and RNAi (siRNA) assays

Human CITED2 cDNA was obtained from Open Biosystems (Thermo Fisher Scientific, Huntsville, AL, USA, clone: LIFESEQ8724402) and subcloned into pcDNA-DEST40 (Invitrogen). Human CBP/p300 cDNA was also subcloned into the pcDNA-DEST40 vector. Transfection into UCCA24 and KE43p was achieved with Lipofectamin2000 (Invitrogen). siRNAs for CDKN1A, CITED2, HOXA1, SAT, TUBA3, and p53 (Ambion, Austin, TX, USA), and for non-sense control (Ambion) were transfected with the siPORT NeoFX Transfection Agent (Ambion).

Butyric acid treatment, detection of apoptosis, and MCT1 inhibition assays

Cultured cells were exposed to 2.5 mM BA as previously described [12]. Apoptosis was detected by the TUNEL method visualized by fluorescein isothiocyanate (FITC)-labeled digoxigenin according to the manufacturer's protocol with slight modifications using an ApopTag Plus Peroxidase In Situ Apoptosis Detection Kit (Intergen, Purchase, NY, USA). Apoptotic cells were counted and the percentages versus more than 100 DAPI (4′,6-diamino-2-phenylindole) positive cells were calculated. 0.5 mM 4-hydroxycinnamate (4CHC, Sigma, Saint Louis, MO, USA) was added to the culture medium for MCT1 inhibition assays.

Immunoprecipitation

Aliquots of 2×10^7 UCCA24 and KE43p cells were lysed in 1 ml of TNE buffer (10 mM Tris–HCl, pH 7.8, 1% nonidet-P 40, 0.15 M NaCl, 1 mM EDTA). After preprecipitation with protein G-Sepharose for 1 h at 4°C, lysates were incubated with 1 μg of anti-CITED2 antibody or anti-CBP/p300 for 1 h at 4°C followed by incubation with protein G-Sepharose for 1 h at 4°C. After centrifugation at 15,000 rpm for 5 min, the precipitants were washed with TNE buffer 5 times and denatured at 95°C. After SDS-polyacrylamide gel electrophoresis, they were subjected to western blotting with anti-CBP/p300 or anti-CITED2 antibodies.

Colorectal specimens and immunohistochemistry

A total of 102 surgically resected colorectal specimens from active and inactive UC patients and 20 from diverticulitis patients were obtained from the files of Kitasato University East Hospital and Kitasato University Hospital, and the histological activity of UC was determined according to the Matts' classification [21]: score 1, normal appearance; score 2, some infiltration of neutrophils in the mucosa; score 3, much cellular infiltration or presence of cryptitis; score 4, presence of crypt abscess and cellular infiltration; and score 5, presence of ulceration, erosion or necrosis, and cellular infiltration. Diverticulitis activity in diverticulosis was evaluated as none to mild and severe. Specimens were fixed with 10% formalin and embedded in paraffin. Semi-serial 4-μm-thick sections were cut for hematoxylin–eosin (H&E) staining and CITED2 immunohistochemistry with monoclonal anti-CITED2 antibodies (JA22, 1/1000, Novus Biologicals, Inc., Littleton, CO, USA) and an EnVision + kit (DakoCytomation, Glostrup, Denmark). After blocking endogenous peroxidase with 0.3% H_2O_2 in methanol, a 15-min microwave pretreatment in citrate buffer (pH 6.0, 0.01 mol/l) was performed for antigen retrieval, and after incubation with anti-CITED2 antibodies for 1 h at room temperature. 3,3′-Diaminobenzidine was applied as the final chromogen and faint nuclear counterstaining was finally achieved with methyl green. Scoring of immunoreactivity was performed as reported by Sinicrope et al. [22], and statistically analyzed by the Spearman's rank correlation test and Mann–Whitney's *U* test.

Results

Butyric acid upregulates the transcriptional expression of CITED2 specifically in UCCA lines

To identify genes upregulated by BA treatment specifically in UCCA lines, three UCCA and three sCRC lines were treated with 2.5 mM of BA (pH 7.4 in the culture medium). We earlier reported that the concentration of BA in the wells of Vero cell culture with *F. varium* was 6.42 mM (pH 6.90) [11], and 2.5 mM induced p53-mediated DNA repair [12]. The Percellome method enabled us to compare the mRNA expression profiles among the different cell lines per cell based on DNA content of each biosample. We applied the original algorithm (Fx) [20] and upregulated and

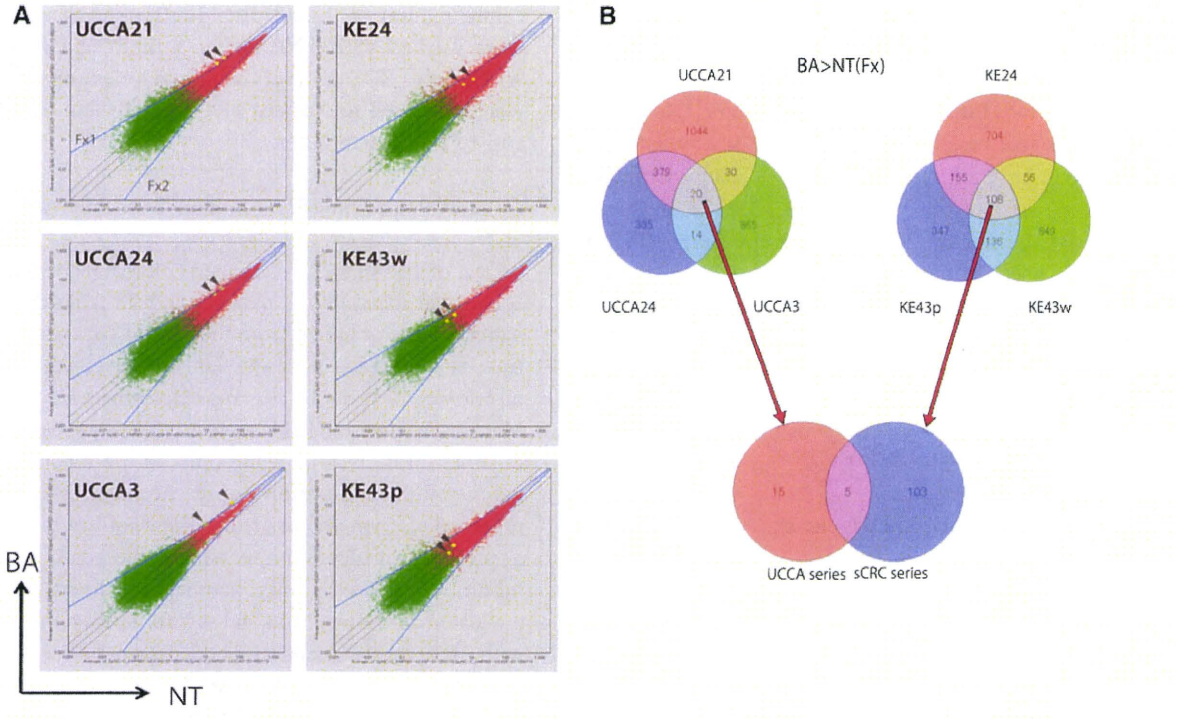


Fig. 1 **a** Genome-wide expression analysis after BA treatment. The analysis was performed by using the Affymetrix GeneChip Human Genome U133 plus 2.0 Array. Scatter plots are shown for BA treatment (BA) for 24 h and non-treatment (NT). The *open circles* show expression levels for each probe set. The *colors* indicate probabilities provided by the Affymetrix GeneChip Operation System: *red* means good and *green* means poor. The *yellow circles* show CITED2 probe sets for areas highlighted by *arrowheads*. The *black lines* show twofold, onefold, and 0.5-fold, respectively, and the *blue lines* indicate the empirical threshold levels (Fx1 and Fx2). The *spots above* the Fx1 and *below* the Fx2 line were evaluated as upregulated and downregulated, respectively ($p < 0.02$). **b** Venn diagrams of the upregulated probe sets in butyric acid (BA) treatment compared with non-treated (NT) controls in UCCA lines (UCCA21, UCCA24, and UCCA3) and sCRC lines (KE24, KE43p, and KE43w). *Numbers* of upregulated probe sets are indicated. Among the 20 and 108 probe sets upregulated in the UCCA and sCRC series, respectively, 5 were commonly upregulated in both. *Fx* Fx algorithm for the empirical threshold calculation

downregulated genes were listed at $p < 0.02$ (Fig. 1a). As compared with non-treated controls, 20 probe sets' genes were commonly upregulated in the UCCA lines, and 108 in the sCRC lines. Five were upregulated in both UCCA and sCRC lines, and 15 and 103 only in UCCA and sCRC lines, respectively (Fig. 1b). Of the 13 known genes contained in the 15 probe sets specifically upregulated only in UCCA lines, CITED2 was upregulated with 2 independent probe sets in the same GeneChip (see Table 1 for the gene list). As we previously found that the *F. varium*-BA system induces p53-dependent cellular responses [11, 12], we focused on whether p53-signaling might be influenced. As CITED2 was identified as a CBP/p300 binding protein which regulates p53 by acetylation [18, 23, 24], we considered that CITED2 might be a candidate and subjected it to further analyses.

CITED2 acts upstream of p53

q-RT-PCR showed increased upregulation of CITED2 in UCCA cell lines as compared with sCRC cell lines on BA

Table 1 Genes upregulated by BA treatment only in the UCCA series

Probe set ID	Gene symbol	GenBank ID
237156_at	–	–
204567_s_at	ABCG1	NM_004915
210431_at	ALPPL2	J04948
200982_s_at	ANXA6	NM_001155
207980_s_at	CITED2	NM_006079
209357_at	CITED2	NM_006079
202806_at	DBN1	NM_004395
215506_s_at	DIRAS3	AK021882
211538_s_at	HSPA2	U56725
203752_s_at	JUND	NM_005354
208581_x_at	MT1X	NM_005952
201599_at	OAT	NM_000274
201364_s_at	OAZ2	AF242521
35666_at	SEMA3F	U38276
204141_at	TUBB2	NM_001069

treatment, validating the microarray results (Fig. 2a). p53 downstream molecules, $p21^{Cip1}$ and Bax, were upregulated at the mRNA level, although p53 mRNA expression was not changed (Fig. 2b). The expression of monocarboxylate transporter-1 (MCT1), a transporter of BA, was also upregulated. When an siRNA for CITED2 was introduced into UCCA24 with or without BA treatment, it significantly suppressed CITED2 mRNA expression. Although p53 mRNA expression was upregulated by knockdown of CITED2 without BA treatment, this was abrogated by BA treatment with CITED2 knockdown, indicating an upstream location of CITED2 in the p53 signaling pathway (Fig. 2c). Suppression of CITED2 expression did not significantly affect the mRNA expression of p21, Bax, and p53R2, downstream molecules of p53, with or without BA treatment. CITED2 expression was also upregulated by BA treatment at the protein level in UCCA24 after 1–12 h, with a peak at 6 h, accompanied by p53 protein accumulation and followed by p21 and Bax expression (Fig. 3). Although Bax expression was high in the non-treatment state, Bcl-2 levels were also high and decreased with BA treatment. Upregulation of CITED2 was noted within 6 h of CITED2 transfection and p53 protein accumulation was observed in parallel. Thus, on treatment with BA, the p53 protein level was upregulated without transcriptional upregulation. As p53 downstream molecules were upregulated at the mRNA and protein levels, stabilization of p53 protein without transcriptional expression could account for this discrepancy. Therefore, CITED2 could be a candidate transducer of BA responsible for post-transcriptional upregulation of p53.

BA induces apoptosis through MCT1, CITED2, and p53

As we previously reported [12], BA induced apoptosis detected by the TUNEL method in the UCCA24 line (Fig. 4a, b). Apoptosis was abundantly induced by doxorubicin with wild-type p53 as previously reported [16]. It is well known that the doxorubicin-induced apoptosis is p53-dependent [25], and apoptosis was inhibited by an siRNA for p53, but not by non-sense siRNA. BA-induced apoptosis was inhibited by 4-hydroxycinnamate (4CHC), a specific inhibitor for MCT1 in both types of cell lines, and was also significantly inhibited by siRNA for p53. CITED2 overexpression also induced apoptosis in UCCA24, and this was partially and significantly reduced by p53 siRNA, but not by non-sense siRNA. As BA activates p53 by stabilization and does not induce transcriptional expression, the partial inhibition suggests reduction of the protein pool before BA treatment or CITED2 induction. The results indicate that BA induces apoptosis through MCT1, CITED2, and p53 activation. Although the BA-induced

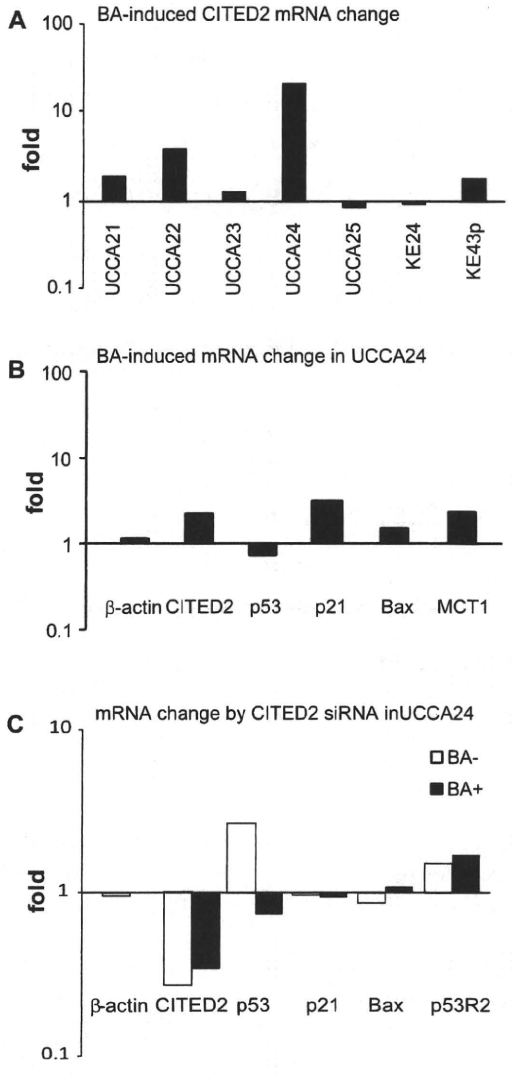


Fig. 2 q-RT-PCR for validation of CITED2 and p53-related genes, and knockdown effects of CITED2. Representative results for three independent experiments are shown. **a** CITED2 mRNA expression is induced by BA in UCCA cell lines. q-RT-PCR was performed after 2.5 mM BA was added for 24 h. UCCA cell lines, including UCCA 21 and 24 used for microarray analysis, tended to show upregulation of CITED2 mRNA expression against no BA treatment, whereas sCRC lines (KE24, KE43p) did not. **b** mRNA expression of p53-related molecules except p53 itself was increased in UCCA24 by BA treatment for 24 h, as analyzed by q-RT-PCR. Fold ratios of mRNA expression after 24 h BA treatment against no BA treatment are shown in the histogram. **c** siRNA for CITED2 was transfected 24 h before BA treatment in UCCA24. After BA treatment for 24 h, expression of *β*-actin, CITED2, p53, p21, Bax, and p53R2 was analyzed by q-RT-PCR. Fold expression of CITED2 siRNA transfected cells was compared with non-specific siRNA transfection. *Open* and *filled bars* show fold ratios for each gene against VIC-labeled *β*-actin expression with and without BA treatment

CITED2 level was almost equivalent to or greater than that in CITED2 transfected cells (Fig. 3a, b), the apoptosis level was much greater in CITED2 transfected cells (Fig. 4b).

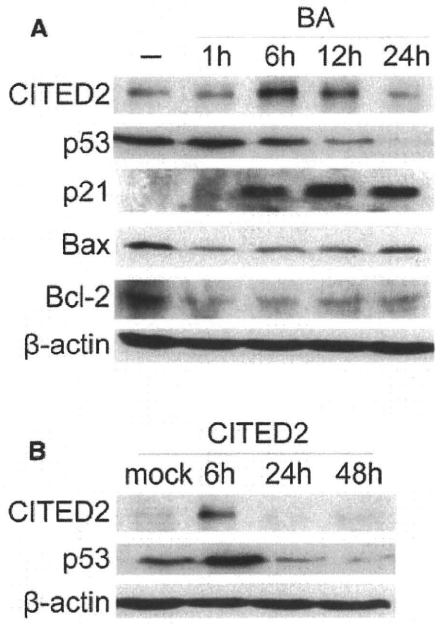


Fig. 3 CITED2, p53, and related protein expression with BA treatment and forced CITED2 expression in UCCA24. The indicated protein expression was analyzed by western blotting with BA treatment for 1, 6, 12, and 24 h (**a**) and CITED 2 transfection for 6, 24, and 48 h (**b**). CITED2 was transfected 24 h prior to BA treatment

This might be due to the direct effect of CITED2 by transfection whereas the apoptosis induced by BA could be regulated by other factors.

BA and CITED2 induce acetylation of p53 and increase protein expression

CITED2 has been identified as a co-factor for CBP/p300 [18], which induces apoptosis through acetylation of p53 [24, 26]. The amount of p53 protein was significantly increased along with the acetylated form of p53 with BA treatment (Fig. 5a) and CITED2 transfection (Fig. 5b). It is likely that CITED2 stabilizes p53 by its acetylation.

CITED2 binds to CBP/p300 in vitro

During BA treatment, we confirmed binding of CITED2 to CBP/p300 in vitro by immunoprecipitation (Fig. 5c). Binding was induced by CITED2 and CBP/p300 transfection as well as by BA treatment.

CITED2 is expressed in the surface epithelium of UC patients' colonic mucosa

Expression of CITED2 protein was evaluated in the colonic mucosa of UC patients in situ. The area (0–3) and the intensity (0–3) of CITED2 positivity were multiplied together and the correlation of the resulting CITED2 expression score with histological evidence of inflammation was assessed according to Matts (see Fig. 6a) [21]. CITED2 expression was localized in crypt bottoms of normal colonic mucosa. In contrast, it was found to be significantly increased in the surface lining epithelium and both upper and lower halves of the crypts in UC cases, whereas the increase of expression was limited to the lower halves in diverticulitis, with non-specific inflammation. In addition to the significant differences in expression score between histological grades by the Mann–Whitney's U test shown in Fig. 6b, a significant correlation of the increase of CITED2 expression with histological grade by Spearman's rank correlation test was also demonstrated, especially in lining epithelium and the upper halves of crypts ($\rho = 0.641$, $p < 0.0001$, lining epithelium, $\rho = 0.691$, $p < 0.0001$, upper crypts, $\rho = 0.460$, $p = 0.0002$, lower crypts).

Discussion

There have been many studies of the mechanisms underlying UC. Concerning roles of enterobacteria in UC development, bacterial invasion into the mucosal epithelia is considered to be associated with the production of cytokines, as a first step in inflammatory processes [10]. Thus, inflammatory cytokines produced by macrophages and lymphocytes in the stroma of the colonic mucosa are considered important factors. Interleukin-10 (IL-10)-deficient mice spontaneously develop UC-like colitis under conventional conditions, but not under specific pathogen- or germ-free conditions [6, 7]. To induce colitis in IL-10-deficient mice, commensal enteric bacteria are necessary [7]. Shkoda et al. [8] clearly revealed that IL-10 inhibits inflammation-induced endoplasmic reticulum (ER)-mediated stress responses by modulating activating transcription factor (ATF)-6 nuclear recruitment to the glucose-regulated ER stress protein-78 gene promoter in colonic epithelial cells. In inflammatory foci of UC, p53 is activated in the crypt epithelium [12], which suggests protective activity against inflammatory stress. Therefore, mutations of p53, frequently found in the early stages of UC-associated tumorigenesis, including non-tumorous regenerative mucosa, in contrast to the adenoma-carcinoma sequence [27], may greatly contribute to carcinogenesis. Short-chain fatty acids including BA are important to maintain colonic functions, and BA constitutes an energy supply for colonocytes and plays a central role in homeostasis of the colonic mucosa [28]. Previously, we detected *F. varium* by direct culture of UC patients' mucosa [14] and induced experimental UC in mice by intra-rectal administration of BA secreted from *F. varium* [9, 11]. In addition, BA induces DNA repair and apoptosis with the activation of

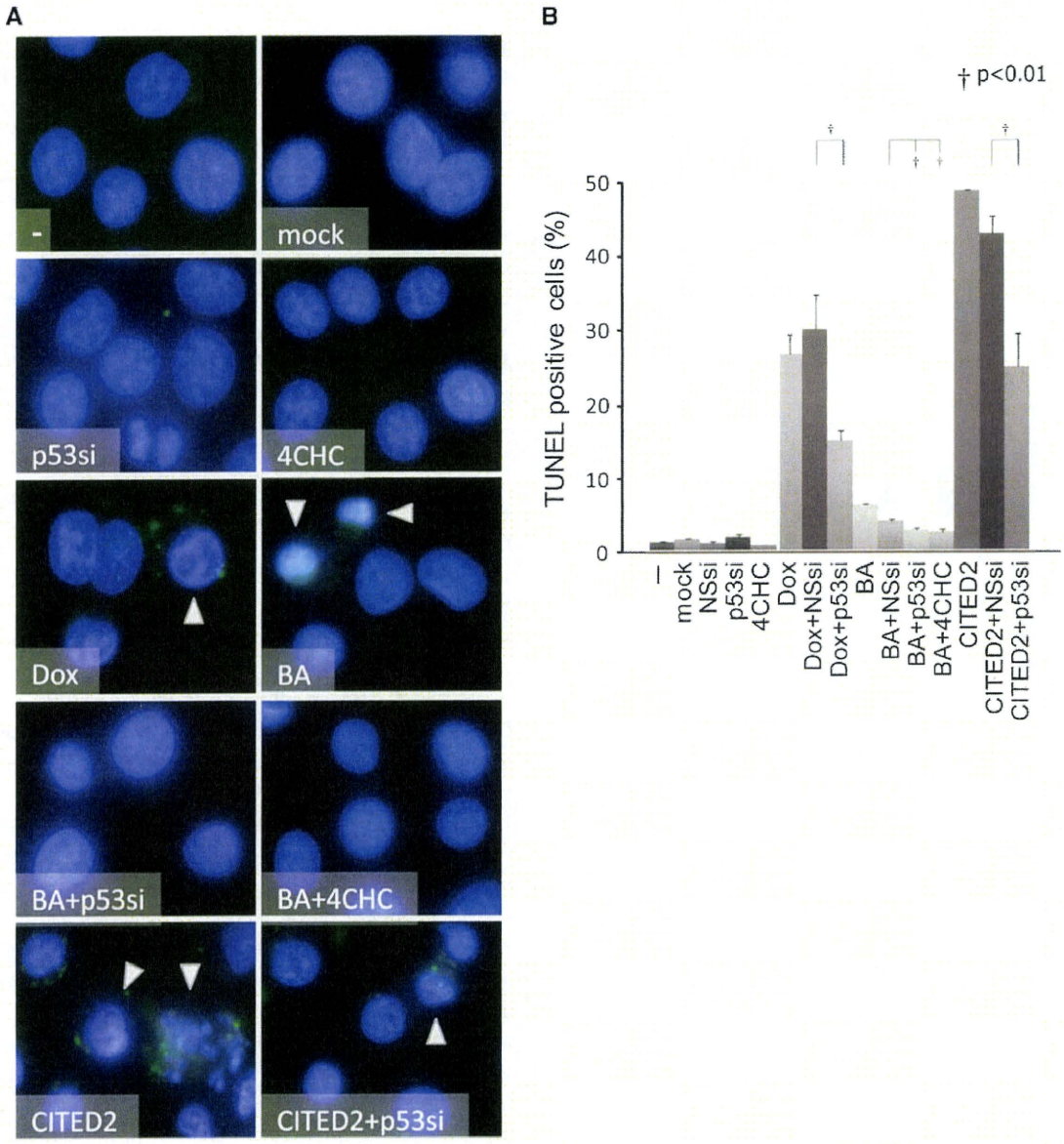


Fig. 4 Induction of p53-dependent apoptosis by BA treatment and CITED2 expression. **a** Representative images of apoptosis detected by the TUNEL method under various conditions in UCCA24 (*NT* no treatment, *Dox* treated with doxorubicin, *4CHC* treated with 4CHC, *BA* stimulated with BA, *BA + 4CHC* treated with both BA and 4CHC, *mock* mock transfection, *CITED2* CITED2 transfection). **b** Average percentages of apoptotic cells in three experiments in UCCA24 in the histogram with standard deviation bars. NSsi and p53si, siRNA for non-sense and p53 transfected, respectively. †Significant differences ($p < 0.01$) by Student's *t* test for each condition are also shown

p53–p53R2 system [12]. Therefore, BA would damage colonocytes in a high concentration condition. In order to mimic the inflammatory condition of UC in vitro, we used UCCA cell lines derived from human UC-associated adenocarcinomas [16]. Although UCCA lines are tumor cell lines and may have different characteristics from non-tumorous colonic epithelium, those used here are unique in being derived from a human UC background. As BA-induced apoptosis and activated p53-dependent DNA repair have been found in both sCRC and UCCA lines [11, 12], we here focused on their roles as a key factor for UC inflammation.

First, we established that mRNA expression of CITED2 and several other genes was upregulated by BA treatment in ulcerative colitis-associated cancer cell lines, but not in sporadic counterparts, by cDNA microarray analysis with the Percellome method [17]. Whereas two independent probe sets were listed as only upregulated in UCCA lines,

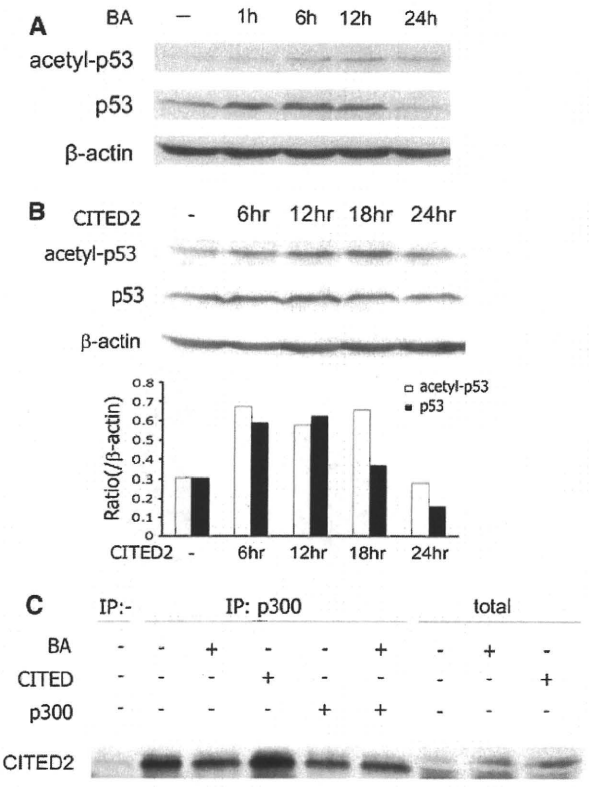


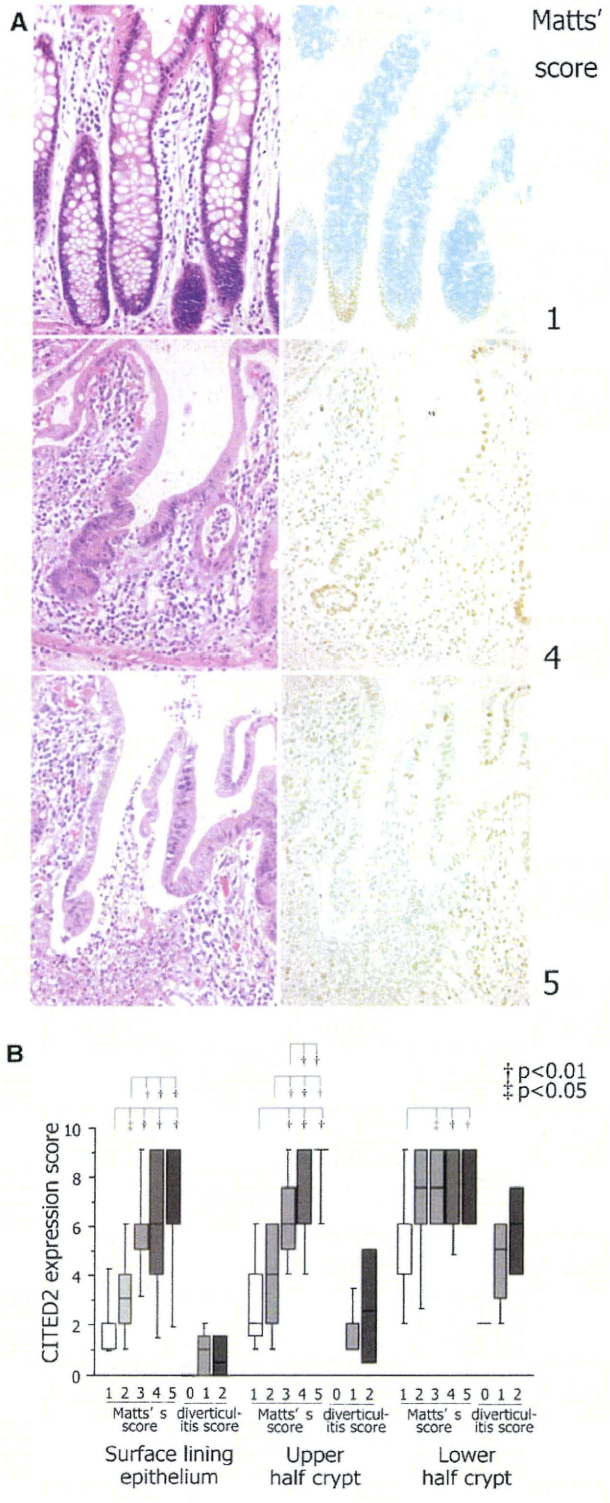
Fig. 5 Induction of p53 acetylation by BA and CITED2, accompanied by binding of CBP/p300 and CITED2 in UCCA24. **a** 2.5 mM BA treatment induced acetylation of p53 and increased the total amount of p53 from 1 to 24 h. **b** Transfection of CITED2 24 h prior to BA treatment for 24 h also induced acetylated p53 and increased the total amount of p53. A histogram of results for relative expression of acetylated p53 and p53 is shown *below*. Relative expression against *β*-actin was calculated by densitometry of western blot data. **c** Immunoprecipitation by anti-CBP/p300 antibody and immunoblotting with anti-CITED2 antibody showed binding of CITED2 and CBP/p300. *BA* indicates treatment by 2.5 mM BA for 24 h, *CITED2* and *p300* indicate the transfection of CITED2 and CBP/p300 24 h prior to the assay, respectively. A western blot from the total lysate is also illustrated

we especially focused on CITED2 among the 13 genes, because its function seemed to be closely related to p53 via CBP/p300 [18], with an upstream action. Indeed, RNAi analysis directly indicated that CITED2 is upregulated upstream of p53. Induction of CITED2 mRNA was not evident in several UCCA lines, such as UCCA23 and UCCA25 (Fig. 2). However, we confirmed the induction of CITED2 protein in those lines (data not shown) and the upregulation of CITED2 was statistically confirmed by the immunohistochemistry from the 102 inflammatory lesions of UC (Fig. 6).

Although CITED2 has been noted as an important factor for organ or tissue development [29, 30], it was originally found to inhibit the binding to hypoxia-inducible factor 1 (HIF1) [31]. Bakker et al. [32] reported CITED2 to be transcriptionally induced by FOXO3a, a member of the FOXO subfamily of Forkhead transcription factors, and to inhibit HIF1-induced apoptosis in hypoxia, which is proposed as a survival response of cancer cells. Although HIF1 induces p53-dependent apoptosis [33–35], it also causes p53-independent apoptosis via BNIP3 and NIX induction [34–36]. As Bakker et al. used p53 wild-type cell lines, it was uncertain whether CITED2 inhibited p53-dependent HIF1-induced apoptosis. In the present study, however, BA and CITED2 clearly induced apoptosis, which was inhibited by an siRNA for p53. Furthermore, the BA-induced apoptosis was inhibited by an inhibitor of MCT1, through which BA is transported into the cytoplasm [37, 38], and an siRNA for p53 partially but significantly inhibited the BA- and CITED2-induced apoptosis. Apoptosis induced by CITED2 transfection was more frequent than by BA treatment (Fig. 4b), although the BA-induced CITED2 level was almost equivalent to or greater than that in CITED2 transfected cells (Fig. 3a, b). This might be due to a direct effect of CITED2 by transfection or to apoptosis induced by BA being regulated by other factors. The dependence of BA-induced apoptosis on MCT1, demonstrated by 4CHC, an inhibitor of MCT1, should be verified by other approaches. Although protein stabilization of p53 was apparent on BA treatment with CITED2 overexpression, knockdown of p53 at the mRNA level could also inhibit the BA–CITED2 pathway. Therefore, CITED2 possibly plays different roles in BA-induced and HIF1-induced apoptosis, which might depend on modulation of the phosphorylation or acetylation sites of p53.

According to recent studies, p53 induces apoptosis via acetylation of lysine 120 on DNA damage or in response to stress signaling with CBP/p300 binding [24, 26]. The acetylated form of p53 is stable with regard to the Mdm2-ubiquitination system [39]. We here demonstrated acetylation of p53 by BA treatment and forced CITED2 expression accompanied by p53 protein accumulation (Fig. 5a, b). At the same time, binding of the CITED2 and CBP/p300 was also observed by immunoprecipitation (Fig. 5c). Therefore, we conclude that CITED2 is upregulated by BA, in turn activating p53-dependent apoptosis via acetylation of p53 protein (Fig. 7). Although transcription of p53 was not observed, the data support p53 protein stabilization by BA and CITED2. Further investigation of the half-life of p53 after BA treatment with or without CITED2 knockdown should clarify this point.

In UC, p53 gene alteration is observed even in non-tumorous regenerative epithelium [27]. In addition, p53 expression is upregulated in inflammatory foci of UC without gene mutation and the p53 and p53R2 DNA repair system is accelerated [12]. Therefore, the repair capacity may be overloaded in active UC. We previously found greater upregulation of Bax protein in UCCA as compared



◀ **Fig. 6** CITED2 expression in inflammatory foci of ulcerative colitis. **a** Immunohistochemical staining of CITED2 is shown in the *right panels* and semi-serial H&E sections are shown in the *left panels*. Examples of inactive UC (Matts' score 1, *upper panel*) and active UC (Matts' score 4 and 5, *middle* and *lower panel*) are presented. Strong CITED2 expression is remarkable in epithelial nuclei in active UC. **b** CITED2 expression correlates with activity of inflammation of UC in situ. The area (0–3) and the intensity (0–3) of CITED2 expression were evaluated and the CITED2 expression score (0–9) was calculated by multiplying the two results. The inflammatory grade of the diverticulosis cases was scored from 0 to 2 (none, mild, severe). The CITED2 expression scores are shown along with histological scores. Median values are shown by the *horizontal line*, the *box* represents values between the 25th and 75th percentiles, and the *lower* and *upper bars* indicate the 10th and 90th percentiles, respectively. $^{\dagger}p < 0.01$ and $^{\ddagger}p < 0.05$ by Mann–Whitney's *U* test

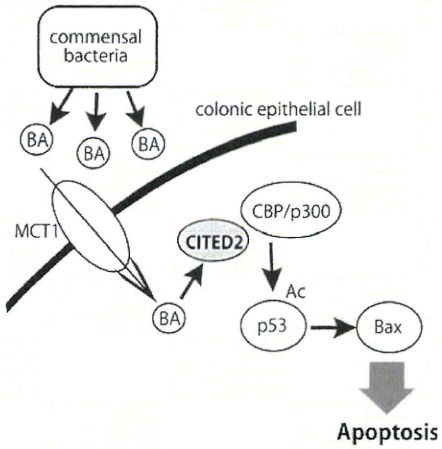


Fig. 7 Scheme for BA-induced apoptosis. BA produced by commensal bacteria, such as *F. varium*, is incorporated by MCT1 and induces CITED2 expression. The CITED2-CBP/p300 complex then acetylates p53 which stabilizes the protein, resulting in upregulated transcription of proapoptotic genes including Bax, and induction of apoptosis. Apoptosis of colonic epithelial cells leads to erosion of the mucosa, so that bacteria can more easily colonize. *Ac* acetylation; *BA* butyric acid; *MCT1* monocarboxylate transporter-1

with sCRC [40], which might have a similar effect. Matts' histological scores based on the infiltration of neutrophils and the presence of cryptitis, crypt abscess and erosion, reflect the inflammation activity of UC [21]. The fact that CITED2 expression in UC is also increased in the upper halves of the crypts and positively correlated with the histological activity of UC, evaluated by Matts' histological scores, is clear evidence of an association with UC inflammation. In contrast, although the CITED2 expression was increased in the inflammatory foci of diverticulitis, the expression score was almost at the same level as that of the inactive UC (Matts' score 1). This might indicate that the BA-CITED2 system is specifically involved in the UC inflammation, and not under conditions of non-specific colonic inflammation such as diverticulitis.

Previously, we showed that specific enterobacteria including *F. varium* are frequently detectable in UC patients' mucosa. Here we showed BA-induced apoptosis via CITED2 induction and p53 activation, and that CITED2 expression clearly correlates with histological UC