

# Environmental determinants of *Opisthorchis viverrini* prevalence in northeast Thailand

Yi-Chen Wang<sup>1</sup>, Chen-Chieh Feng<sup>1</sup>, Paiboon Sithithaworn<sup>2</sup>

<sup>1</sup>Department of Geography, National University of Singapore, Singapore, Singapore; <sup>2</sup>Department of Parasitology, Khon Kaen University, Khon Kaen, Thailand

**Abstract.** *Opisthorchis viverrini* infection is an important, human, parasitic disease along the middle and lower Mekong River of Southeast Asia. Despite vast efforts in epidemiological research, the wide geographical variation in *O. viverrini* prevalence remains unexplained. To investigate the potential influence of village location, prevalence data recorded during the year 2009 for 90 villages in northeast Thailand were analysed by a geographical information systems approach based on a digital elevation model for altitude determination. Land use compositions of village surroundings were derived from images produced by the Thematic Mapper of the Landsat satellite and landscape metrics were applied to quantify spatial patterns. Pearson's correlation coefficients were used to evaluate the associations between the environmental factors and *O. viverrini* prevalence. In spite of no statistical difference ( $P = 0.117$ ), higher-ground villages generally had a lower level of opisthorchiasis than those situated lower down, while a significant negative correlation was detected between *O. viverrini* prevalence and the composition of farmland with low water content ( $P = 0.028$ ), indicating the potential influence of agricultural lands with drought-tolerant crops. Strengthening this hypothesis, a significant positive association with *O. viverrini* prevalence was found for the combined land use classes of water and of farmland with high water content ( $P = 0.032$ ), indicating that the dominance of a large land use patch capable of providing suitable habitats for the intermediate hosts might contribute to the disease. This study highlights the influence of landscapes on the epidemiology of *O. viverrini* infection and emphasises the need to incorporate land use planning into control strategies.

**Keywords:** spatial epidemiology, *Opisthorchis viverrini*, food-borne trematodiasis, neglected tropical disease, land use, geographical information systems, Thailand.

## Introduction

Pathogen transmission is an inherently spatial process, requiring a susceptible host to encounter an infected host or vector (Killalea et al., 2008). Factors affecting the spatial distributions of pathogens, hosts and vectors, and their likelihood of close encountering, are therefore essential to the dynamics and patterns of diseases (Ostfeld et al., 2005). Advances in geospatial techniques and availability of high-resolution remote sensing images have created new opportunities to investigate environmental factors in explaining the spatial variation in disease risk and incidence (Elliott and Wartenberg, 2004; Graham et al., 2005; Malone, 2005). Indeed, landscape epidemiology has arisen as an important discipline for understanding the causes and consequences of spatial heterogeneity in infectious diseases, particularly in recognising the environmental

effects on the transmission of many parasitic diseases (Liang et al., 2007; Meentemeyer et al., 2012).

Food-borne parasitic infections constitute a group of recognised, emerging, human diseases but are underestimated as a world health problem. Opisthorchiasis, liver fluke infection caused by *Opisthorchis viverrini* through the consumption of raw freshwater fish, is one of the most important human parasitic diseases in Southeast Asia, where raw or undercooked fish dishes are of considerable cultural and nutritional significance. Most infections cause only mild symptoms and can be treated with praziquantel (Sithithaworn et al., 2007), while heavy infections are associated with a number of hepatobiliary diseases and closely related to the incidence of cholangiocarcinoma (CCA) (Sripa et al., 2011). Opisthorchiasis is unevenly distributed at every spatial scale explored (e.g. Sripa, 2008; Sayasone et al., 2011). On continental Southeast Asia, *O. viverrini* is endemic in the Lower Mekong Basin, including Thailand, Lao People's Democratic Republic (Lao PDR), Cambodia and southern Vietnam (Sripa, 2008), with an estimated 67.3 million people at risk of infection (Keiser and Utzinger, 2005). At the smaller spatial extent within a province of a country, sometimes even between villages within the same district, wide range

Corresponding author:  
Yi-Chen Wang  
Department of Geography, National University of Singapore  
1 Arts Link, Singapore 117570, Singapore  
Tel. +65 6516-6811; Fax +65 6777-3091  
E-mail: geowyc@nus.edu.sg

of variation in *O. viverrini* prevalence has been reported (Wang et al., 2011). For example, prevalence of infection ranges from 2.1% to 70.8% in 20 districts in Khon Kaen province, northeast Thailand (Sriamporn et al., 2004). Prevalence remains high in some endemic areas, evident in a recent study in a lowland district in Champasack province, Lao PDR, for which a high prevalence of 92.0% has been reported (Sayasone et al., 2011).

Because of the high prevalence, great risk of infection and as an important risk factor for CCA, considerable efforts have devoted to *O. viverrini* medical research (e.g. Sornmani et al., 1984; Saowakontha et al., 1993; Sayasone et al., 2007; Chaiyarit et al., 2011). One remaining unexplained issue is, however, the wide geographic variation in *O. viverrini* prevalence. This underscores the need to scrutinise the environmental determinants of *O. viverrini* infection (Wang, 2012), because environmental factors affecting the ecological niches and habitats of the parasites and their hosts potentially govern the spatial distributions of parasitic diseases (Zhou et al., 2009). Regarding *O. viverrini*, its life cycle involves three hosts: snails of the genus *Bithynia* act as the first intermediate host; Cyprinidae fish serve as the second intermediate host; and fish-eating mammals, including humans, cats and dogs are the definitive hosts. At the point of transmission from human or reservoir hosts to snails, the snails ingest *O. viverrini* eggs present in the faeces of the prior host in the shallow lentic habitats of snails. The subsequent transmission between snails and fish takes place when cercariae released from *Bithynia* snails locate cyprinid fish at a freshwater habitat. The chance of transmission increases when fish has direct contact with infected snails through overlapping habitats.

As the habitats of the two *O. viverrini* intermediate hosts are freshwater environments, location and surrounding land use of human settlement can play an important role in affecting the risk of infection. Villages of which the surroundings are dominated by land use relevant to the intermediate host habitats, such as wetlands, paddies, streams, ponds and lakes, may be more susceptible to the parasite than villages dominated by forests (Wang et al., 2011). A recent study at the provincial level in southern Lao PDR showed that the risk of *O. viverrini* infection is higher for people living in close proximity to freshwater bodies, with high risk zones concentrated in the Mekong River corridor (Forrer et al., 2012). Conversely, earlier work in Loei and Nong Khai provinces in Thailand observed lower prevalence rates among people residing on the banks than those far from the rivers, despite the higher recording of raw fish

consumption in the villages on the bank (Tesana et al., 1991). The differences in these research findings suggest that a quantitative landscape epidemiological approach incorporating geographical information systems (GIS) and remote sensing to evaluate the locations of the villages, such as their altitude and the distance to water bodies, should be considered, as opposed to visually estimating the proximities to rivers as was often done previously. Additionally, efforts are needed to assess the influence of land use compositions and landscape patterns, such as the dominance and aggregation of water bodies of the surrounding landscapes, on *O. viverrini* prevalence across a broad regional scale so as to provide insights into the large human geographic variability of *O. viverrini* infection.

This study employs geospatial techniques to examine the environmental determinants of *O. viverrini* prevalence in northeast Thailand, the region with the highest incidence of the opisthorchiasis associated CCA in the world (Sriamporn et al., 2004; Shin et al., 2010). Specifically, this study addresses four research questions. First, do the locations of villages, in different catchments and altitudes, affect their *O. viverrini* rates? Second, do the proximities of villages to water bodies influence the *O. viverrini* prevalence rates of the villages? Third, do the compositions of the surrounding land use account for the differences in *O. viverrini* infections among villages? Fourth, do the spatial patterns and dominances of the landscapes contribute to the variation of *O. viverrini* infections across the villages?

## Materials and methods

### Study area

The study area of northeast Thailand (Fig. 1a) consists of 19 provinces, covering approximately 160,000 km<sup>2</sup>. It is roughly conterminous with the Khorat Plateau, which tilts from the Phetchabun mountain range in the west of the region down towards the Mekong River in the east along the border with Lao PDR. Topographically, the region is generally low and undulating; elevation ranges from 8 m to 1,800 m with an average of 227 m. It has a tropical monsoon climate with distinctive dry and rainy seasons. The rainy season occurs between May and October; the wettest period of the year is usually August to September. Average annual rainfall is 1,379 mm; about 80% of the rain takes place during the rainy season. Mean minimum and maximum temperatures are 16.6 °C and 35.9 °C, respectively (Thai Meteorological Department, 2012).

### Prevalence data

The *O. viverrini* human prevalence data used in this study were part of the 2009 national survey of parasitic infections conducted by the Thai Ministry of Public Health. The acquired data consisted of 4,632 individuals from 90 villages across the study area. Among these individuals, 734 persons were found to be *O. viverrini* positive, resulting in an overall prevalence of 15.9% for the area. At the village level, however, the prevalence varied greatly, ranging from 0% to 59.5% across the 90 villages, with a mean and median of 16.7% and 15.0%, respectively.

To visualise the spatial variation of *O. viverrini* prevalence, village coordinates were searched and then entered into a GIS (Fig. 1b). Among the 90 villages, the coordinates of 65 villages (72.2%) were verified, while the locations for 25 villages (27.8%) could only be confirmed to their sub-districts (one administrative level above village), mainly due to border change, re-allocation of villages or translation issues for village

names. For those villages, the centroids of their respective sub-districts were used as the coordinates. This positional uncertainty would have minimum impacts on most of the analyses in this study because elevation within the same sub-district did not change greatly. Nevertheless, when precise coordinates were crucial, such as investigating the proximities of villages to major water bodies, analyses were conducted using the two datasets (i.e. 90 villages with some coordinates estimated *versus* 65 villages with precise coordinates) to examine if the results varied. Additionally, because all the village locations were confirmed at the sub-district level, when quantifying the surrounding land use compositions and landscape patterns of the villages, the boundaries of sub-districts and districts (one administrative level above the sub-district) were used to define the two spatial extents of analysis, so as to evaluate their potential effects on the spatial variation of *O. viverrini* prevalence. All the GIS analyses were conducted in ArcGIS version 10 (ESRI; Redlands, USA).

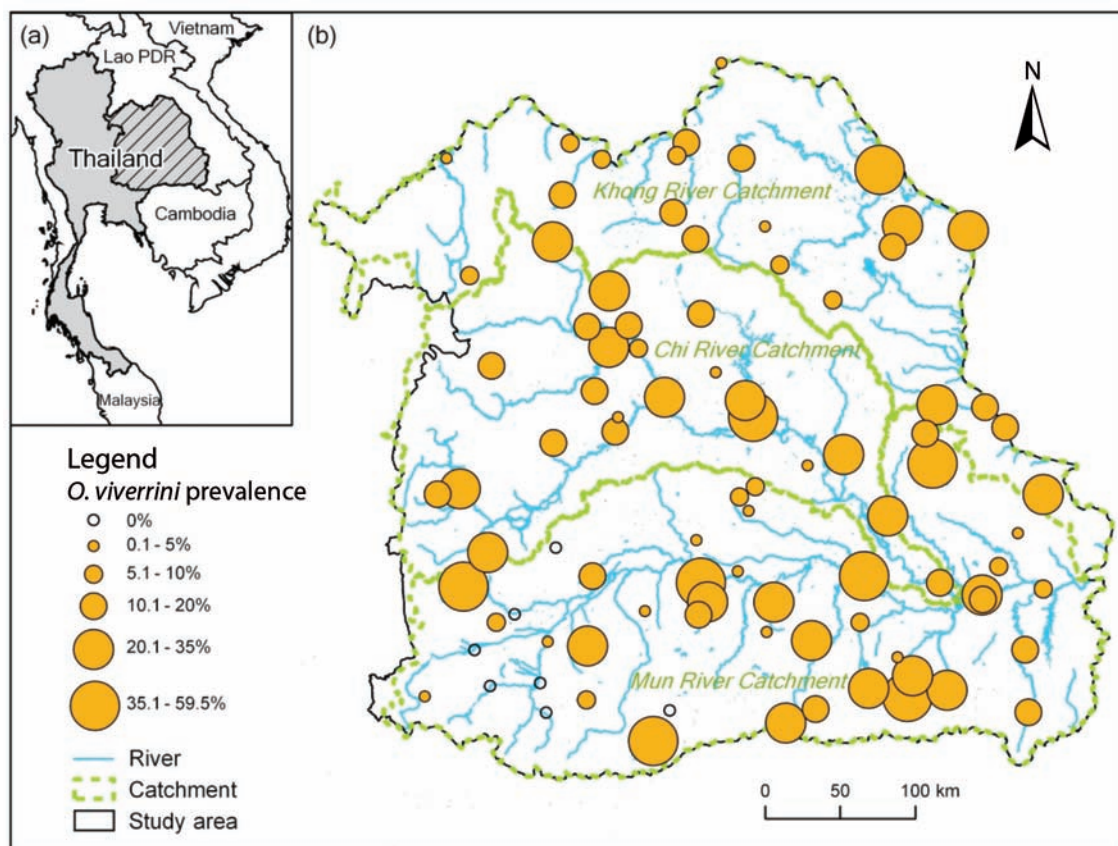


Fig. 1. Prevalence of *O. viverrini* infection for the villages analysed in northeast Thailand. (a) Location of the study area (diagonal lines) in relation to Thailand and other Southeast Asian countries. (b) The study area in northeast Thailand showing the river catchments and *O. viverrini* prevalence for the 90 villages surveyed in 2009.

### *Analysis of village locations*

To address research question one for how village locations, in different catchments and at different altitudes might influence the prevalence rates, two analyses were carried out. First, a GIS river basin data layer prepared by the Royal Irrigation Department of Thailand was used to group the villages into three main catchments: The Khong River catchment in the North, the Chi River catchment centrally and the Mun River catchment in the South (Fig. 1b). The Khong River catchment is part of the Mekong River main corridor bordering Thailand and Lao PDR. Prevalence rates for the villages by catchments were summarised using box plots. The Shapiro-Wilk test (1965) diagnosed that the prevalence data for the three catchments were not normally distributed ( $P < 0.01$ ); the Kruskal-Wallis test (1952) was hence performed to test the equality of the mean prevalence across the catchments.

The second analysis used digital elevation model (DEM) data to examine if the altitudes of the villages could affect their *O. viverrini* prevalence rates. The DEM data of the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), acquired from the Land Processes Distributed Active Archive Center at [https://lpdaac.usgs.gov/products/aster\\_products\\_table/astgtm](https://lpdaac.usgs.gov/products/aster_products_table/astgtm), were overlaid with the village locations to extract the elevation for each village. Pearson's correlation coefficient ( $r$ ) was derived to relate the *O. viverrini* prevalence rates and the village altitudes. Next, the villages were divided into two groups, based on the average elevation of the villages. An  $F$ -test was used to confirm the equal variance of the two groups of villages before a two sample  $t$ -test was employed to examine if the prevalence rates of the higher-ground and lower-ground villages were statistically significant.

To investigate research question two for whether the proximities of villages to major water bodies had any impact on the *O. viverrini* prevalence rates, the river network systems from the ASTER DEM data were derived using ArcGIS. In addition, a GIS water body data layer of major reservoir and lake distributions was obtained from the Royal Irrigation Department of Thailand. The village locations were then overlaid respectively with the derived river network data layer and the GIS water body data layer to calculate the distances between the villages and the water bodies. The distances from the villages to the nearest water bodies (i.e. rivers, reservoirs or lakes) were then obtained for correlation analysis. It was

hypothesised that the prevalence of infection would exhibit a negative relationship with the distance to the nearest water body. All the statistical tests and analyses were carried out in R version 2.15.2.

### *Classification of remote sensing images and accuracy assessment*

To scrutinise the effects of land use composition and landscape patterns on the variation of *O. viverrini* prevalence for research questions three and four, the 2009 Landsat Thematic Mapper images were used for land use classification. The images were dated between January and early February, at the beginning of the dry season.

Because rice paddies, streams, ponds and lakes are important habitats of the *O. viverrini* intermediate hosts, delineating the distributions of these freshwater environments was essential to test the hypothesis that villages surrounded by the intermediate host habitats would have high *O. viverrini* prevalence. A land use class of water, consisting of water bodies and shallow water environments (e.g. rice paddies), was therefore required. However, not all the rice paddies had standing water within to be classified into the water class based on the cloud-free images of the dry season because in northeast Thailand, rice was usually cultivated during the rainy season, and some paddies would be harvested before the dry season. To delineate these recently harvested paddy fields, which could play an important role in facilitating *O. viverrini* transmission during the rainy season, three types of farmlands, representing soils with high, medium and low water contents, were classified. It was assumed that farmlands with high water contents in soils derived from the images taken at the beginning of the dry season represented the approximate areas for paddies during the rainy season, while farmlands with low water contents in soils signified areas of other crop cultivation, such as cassava and beans.

Consequently, six land use classes were derived as follows:

- (i) water, including water bodies, such as rivers, ponds and lakes, and shallow water environments, particularly paddies with standing water;
- (ii) farmland with high water content, representing areas that were likely to be paddies or easily flooded during the rainy seasons;
- (iii) farmland with low water content, indicating areas for cultivating more drought-tolerant crops;



- (iv) farmland with medium water content, representing crop cultivations other than rice and those drought-tolerant crops;
- (v) build-up area and bare land, such as human settlements, roads and construction sites; and
- (vi) vegetation, consisting of forests, plantations and orchards.

For brevity, the six land use classes were hereafter termed water, high-water farmland, low-water farmland, medium-water farmland, build-up and bare land, and vegetation.

For classification, a total of nine input datasets were used, including six bands from the Landsat images (except for the thermal band), the normalised difference vegetation index, the modified normalised difference water index and the unsupervised classification result using the ISODATA algorithm. After layer stacking these datasets, the regions of interest (ROIs) for each land use type were selected as the training data. The ROI separability was calculated using the Jeffries-Matusita (J-M) distance (Lillesand et al., 2004); when the separability of the J-M distance for each land use type was above 1.8, the ROIs were adopted to create a decision tree with the CART algorithm. The decision tree was implemented in the RuleGen in ENVI version 4.8 (Exelis VIS, Boulder, USA) to the whole study area to create a classified image.

To assess the classification accuracy, an aligned systematic sampling approach from Wang et al. (2012) was adapted. Using a grid of 25 x 25 km, a total of 408 reference points were generated for the study area. Visual interpretation was performed using the original Landsat image facilitated by the high resolution Google Earth image for classification accuracy assessment. The overall accuracy was 93.5% with a Kappa coefficient of 0.91.

#### *Examination of land use composition and landscape patterns*

To examine the effects of the surrounding land use compositions on the variation of *O. viverrini* prevalence across villages for research question three, the outcome of the remote sensing classification was used to calculate the percent compositions of the six land use classes at two spatial extents (i.e. the sub-district and the district levels) for each village. Pearson's correlation coefficients were then used to measure the relationships between the land use compositions of the two spatial extents and the *O. viverrini* prevalence. It was hypothesised that villages surrounded by water and high-water farmland would have higher *O. viverrini*

prevalence than villages surrounded by other land use classes.

To assess research question four, four landscape metrics were computed using FRAGSTATS version 3.3 (McGarigal et al., 2002) to quantify the landscape patterns of the sub-districts and the districts where the villages were located. The largest patch index (LPI) quantified the percentage of total landscape area comprised by the largest patch, thereby signifying a measure of dominance of a specific land use class. The landscape shape index (LSI) calculated the perimeter-to-area ratio, representing the overall geometric complexity of a land use class. The clumpiness index (CLUMPY) measured the level of aggregation for a specific land use class. The contagion index (CONTAG) quantified the overall configuration of the patches of the land use classes at the landscape level; a landscape with well-interspersed land use patches would have lower contagion. These four metrics quantified different aspects of spatial patterns and dominances of the landscape, which could provide insights into the geographic variations in disease prevalence. Pearson's correlation coefficients were calculated to assess if any of the landscape patterns associated statistically significantly with the *O. viverrini* prevalence.

#### *Evaluation of the contributions of environmental factors*

The variables scrutinised in this study, including the catchment where the village in question was located, village altitude, distance to nearest water body, surrounding land use composition and landscape pattern, were then included in a stepwise regression analysis to select the variables for a multiple linear regression. The purpose of the multiple linear regression analysis was to evaluate the contributions of the environmental factors to the spatial variation of *O. viverrini* prevalence, rather than developing a model for disease prevalence prediction because such predictive models required the consideration of socioeconomic factors. Variable selected were based on the stepwise regression result with the minimum Akaike information criterion (AIC), which indicated the relative goodness of fit of the model. To minimise the problem of multi-collinearity, instead of using both the sub-district and district levels of land use composition and landscape pattern data, only the data at the sub-district level were included in the analysis. The variance inflation factor (VIF) was calculated for the variables included in the multiple regression model to indicate potential multi-collinearity issues.

## Results

### Analysis of village locations

#### By catchment

Among the 90 villages analysed in this study, 45 were located in the Mun River catchment, while 21 and 24 villages were located in the Chi River and the Khong River catchments, respectively. The highest prevalence at 59.5% was found in Non Sung village of Kalasin province, situated within the Chi River catchment. Seven villages were not infected with *O. viverrini*; they were all located in the Mun River catchment (Fig. 1b), mostly in Nakhon Ratchasima province.

The villages within the Mun River catchment generally had lower prevalence rates (Fig. 2). Indeed, the median value of the prevalence for the Mun River villages was 8.8%, much lower than the median for the Chi River (18.2%) and for the Khong River (15.2%) villages. The spread of the prevalence in the villages across the catchments was different, evident in the wide variations between the upper quartiles and the lower quartiles (Fig. 2). Alternatively, the average prevalence for the villages in the Chi, Khong and Mun catchments was 18.9%, 17.0% and 15.5%, respectively, while the Kruskal-Wallis test suggested that the average prevalence was not statistically different across the three catchments ( $P = 0.282$ ).

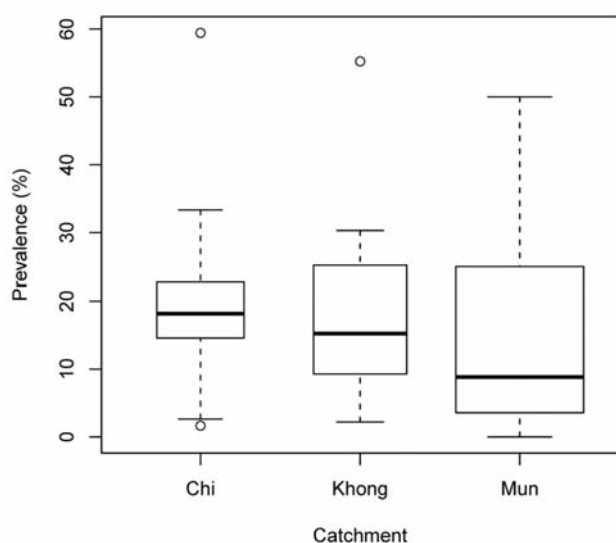


Fig. 2. Comparison of *O. viverrini* prevalence rates among villages of three major river catchment areas in northeast Thailand (boundaries depicted in Fig. 1). The width of the box plot is proportional to the number of villages within each catchment. The lines dividing the boxes horizontally represent the median of the dataset.

#### By altitude

Based on the ASTER DEM, the villages were located at altitudes ranging from 116 m to 423 m; approximately three quarters of these 90 villages were located below 200 m. The result of the correlation analysis showed that the prevalence of *O. viverrini* infection was negatively associated with elevation ( $r = -0.15$ ) but not statistically significant ( $P = 0.163$ ). Analysis of the 65 villages with confirmed coordinates did not improve the statistical significance of the result.

The average elevation of the villages was 176 m, which was used to divide the villages into two groups, higher-ground *versus* lower-ground villages, to examine if the *O. viverrini* infection rates were different. Among the 90 villages, 34 were considered as the higher-ground villages because they were located at altitudes above 176 m; the remaining 56 villages were therefore grouped as the lower-ground villages. The higher-ground villages generally had lower *O. viverrini* prevalence, marked by the following three findings. First, the average prevalence for the higher-ground villages was 13.8%, while that for the lower-ground villages was slightly higher at 18.4%. Second, a total of seven villages in the 2009 survey had no *O. viverrini* infection. Six of these villages were located at the higher-ground area (i.e. above 176 m) and only one at the lower-ground area, but the altitude at which it was situated (i.e. 174 m) was close to the average village elevation. Third, a total of three villages were found to have *O. viverrini* prevalence above 50% and they were all located at the lower-ground area. The *F*-test confirmed that the two groups of villages had equal variances of *O. viverrini* prevalence ( $F = 0.783$ ;  $P = 0.457$ ). The two sample *t*-test, however, suggested that the average prevalence rates between the higher-ground and lower-ground villages were not statistically significant ( $t = -1.581$ ;  $P = 0.117$ ), despite the differences in the prevalence rates between the higher-ground and lower-ground villages noted in the aforementioned three findings.

#### By proximity to water body

Examination of the proximities of villages to major water bodies revealed that the distances between the villages and their respective nearest major water bodies ranged greatly, from 0 m to 5,608 m. In general, villages that were nearer to major water bodies had higher *O. viverrini* prevalence rates than those further away. For example, four villages were found to be at least 1,000 m away from major rivers and lakes, and

their average *O. viverrini* prevalence was 8.9%. Conversely, eight villages located within 100 m of major water bodies had an average of *O. viverrini* prevalence at 20.5%, more than double that of those 1,000 m away. The correlation analyses exhibited negative relations between the distances to major water bodies and the *O. viverrini* prevalence. The relations derived using all the 90 villages and the 65 villages with precise coordinates were, however, both weak ( $r = -0.06$  and  $-0.09$ , respectively) and not statistically significant ( $P = 0.596$  and  $P = 0.498$ , respectively).

#### *Examination of land use composition*

Analysis of land use compositions of the sub-districts and the districts in which the 90 villages were located exhibited similar associations between land use composition and *O. viverrini* prevalence. The influences from the land use classes of water and high-water farmland, however, become more noticeable when the spatial extent of analysis was broadened to the district level. Kamalasai district in Kalasin province was found to have the highest percent value of water (32.7%) for its land use composition, compared to the average of 6.4% across the 90 districts (Table 1; Fig. 3a). Kamalasai district also had the highest percent value of high-water farmland at 11.3%, much higher than the average of 1.7% (Table 1). The high percentage of water and high-water farmland in Kamalasai district might result in high *O. viverrini* prevalence in its villages because these two land use classes provided suitable habitat conditions for the intermediate hosts. Indeed, the village reporting the highest *O. viverrini* prevalence among the 90 villages was Non Sung village, located in the Kamalasai district, with a prevalence of 59.5% (Table 1).

Conversely, vegetation, built-up and bare land and low-water farmland were least associated with the *O. viverrini* intermediate host habitats. Consequently, villages surrounded by these land use classes might experience lower *O. viverrini* infections. The district with the highest percentage of vegetation was Phu Kadung district of Loei province with 85.1% of its land use classified as vegetation (Table 1; Fig. 3e). The village surveyed within the district was Huai Pao village, for which the *O. viverrini* prevalence was 9.5% in 2009 (Table 1), slightly lower than the overall prevalence of 15.9% for the study area. For built-up and bare land, the village with the highest percent value of this land use class was Don Phlai village in Chok Chai district of Nakhon Ratchasima province; approximately 25.5%

of its surrounding land use was classified as built-up and bare land (Table 1; Fig. 3d). A closer scrutiny of its land use composition revealed that almost 60% of the district consisted of built-up and bare land, vegetation and low-water farmland, which could potentially account for the 0% *O. viverrini* infection for Don Phlai village (Table 1). Likewise, the highest percent value of low-water farmland at 24.5% was found in Kra Dian village in Trakan Phutphon district of Ubon Ratchathani province, a village that also reported a low prevalence of 1.9% (Table 1). Investigation of the land use composition of Trakan Phutphon district, in which Kra Dian village was located, showed an extensive coverage of low water and medium-water farmlands (Fig. 3c), while water and high-water farmland accounted for less than 2% of its composition.

The relationships between most of the land use classes and *O. viverrini* infections were, however, not strong, as indicated by the low  $r$  values (Table 2). At the sub-district level, only the composition of low-water farmland was significantly correlated with *O. viverrini* infections ( $P = 0.028$ ). When the land use classes of low-water farmland and medium-water farmland were further combined, the correlation became stronger ( $P = 0.025$ ) (Table 2). Alternatively, the three land use classes which exhibited positive associations with *O. viverrini* prevalence were water, high-water farmland, and vegetation, but none of them were significant at  $P = 0.05$ . Combination of the land use classes of water and high-water farmland for further analysis resulted in a stronger correlation than that with water or high-water farmland alone, although the relationship was only marginally significant ( $P = 0.066$ ) (Table 2). When broadening the area of analysis to the district level, the correlation between water and *O. viverrini* infection became slightly stronger ( $r = 0.20$ ) and marginally significant ( $P = 0.056$ ) than the sub-district level analysis. Overall, the relationships between land use composition derived at the district level and the *O. viverrini* infection in villages exhibited similar patterns to the outcomes computed using the sub-district, land use compositions (Table 2).

#### *Examination of landscape patterns*

Quantifications of the landscape patterns surrounding the villages at the spatial extents of sub-districts and districts revealed some statistically significant findings for the metrics of LPI, CLUMPY and CONTAG, but not the LSI.

Table 1. Average and the highest percentages for each of the six land use classes computed for the districts in which the 90 villages surveyed for *O. viverrini* infections were located.

Land use class	Average (%) <sup>a</sup>	Highest (%) <sup>b</sup>	Village	District	Province	<i>O. viverrini</i> (%) <sup>c</sup>
Water	6.4	32.7	Non Sung	Kamalasai	Kalasin	59.5
High-water farmland	1.7	11.3	Non Sung	Kamalasai	Kalasin	59.5
Medium-water farmland	36.8	68.3	Ban Yang	Sikhoraphum	Surin	24.5
Low-water farmland	4.0	24.5	Kra Dian	Trakan Phutphon	Ubon Ratchathani	1.9
Vegetation	38.3	85.1	Huai Pao	Phu Kadung	Loei	9.5
Build-up and bare land	12.8	25.5	Don Phlai	Chok Chai	Nakhon Ratchasima	0.0

<sup>a</sup>Mean percentage across all districts analysed; <sup>b</sup>highest percentage across all districts analysed; <sup>c</sup>village *O. viverrini* prevalence.

The LPI provided insights into whether the dominance of certain land use patch might have contributed to the level of *O. viverrini* infection of a village. The strength of the correlation for most of the land use classes increased when the spatial extent of analysis was broadened from the sub-districts to the districts (Table 3). At the district level, the LPIs for the medium-water farmland and the low-water farmland were both statistically negatively associated with *O. viverrini* prevalence, with the same correlation coefficient ( $r = -0.22$ ;  $P = 0.039$ ). Conversely, the LPIs for water and high-water farmland both exhibited positive relations with *O. viverrini* prevalence

( $r = 0.10$  and  $0.12$ , respectively), but not statistically significant ( $P > 0.05$ ). When the land use classes were combined into groups to represent the suitable habitats of the *O. viverrini* intermediate hosts (i.e. water and high-water farmland combined) as well as the agricultural land not favourable to the intermediate hosts (i.e. medium and low-water farmland combined), the correlations became stronger (Table 3). In particular, a statistically positive relation with the prevalence was detected for the LPI of the land use class, water and high-water farmland combined, at both the sub-district ( $r = 0.23$ ;  $P = 0.032$ ) and the district level ( $r = 0.29$ ;  $P = 0.006$ ).

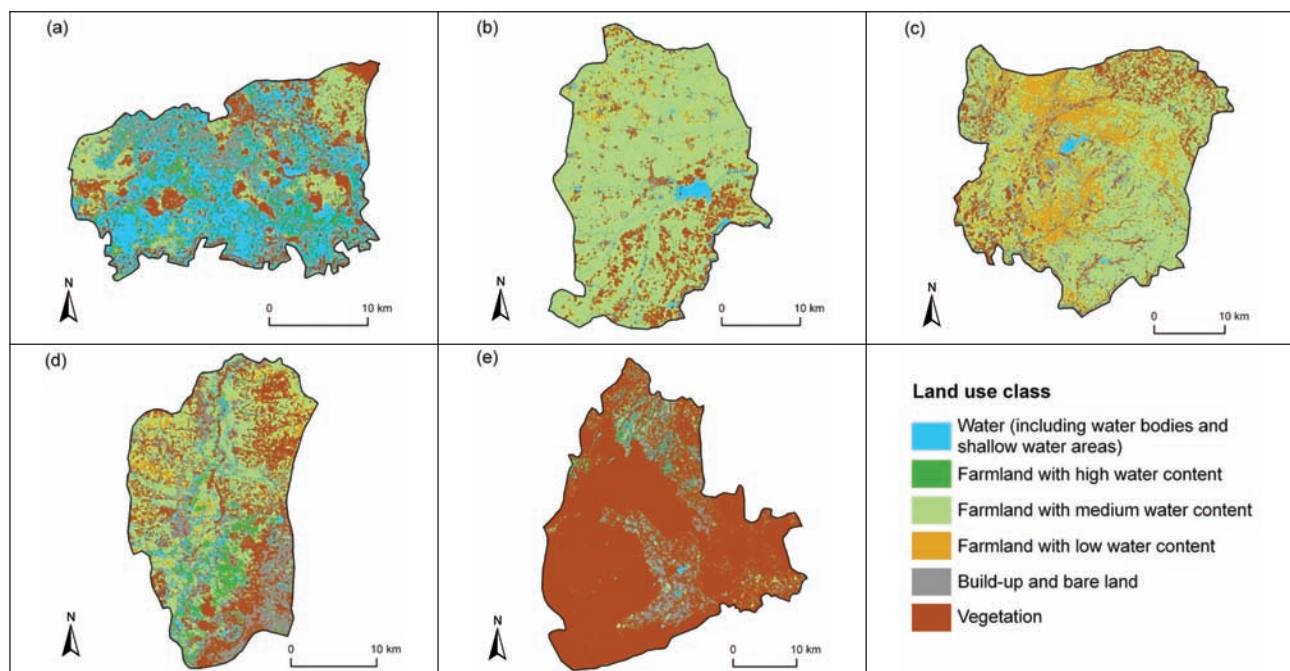


Fig. 3. Land use distributions of districts with the highest percentages of the various land use classes (water and high-water classes represent potential habitats for the *O. viverrini* intermediate hosts). (a) the Kamalasai district with the highest percentages of water and high-water farmland compared to all the districts analysed; (b) the Sikhoraphum district with medium-water farmland; (c) the Trakan Phutphon district with low-water farmland; (d) the Chok Chai district with build-up and bare land; and (e) the Phu Kadung district with vegetation.



Table 2. Relationship between percent composition of each land use class and *O. viverrini* prevalence in 90 villages in northeast Thailand<sup>a</sup>.

Land use class	Sub-district		District	
	<i>r</i>	P-value	<i>r</i>	P-value
Water	0.19	0.082	0.20	0.056
High-water farmland	0.16	0.132	0.16	0.133
Medium-water farmland	-0.20	0.062	-0.20	0.058
Low-water farmland	-0.23	0.028	-0.19	0.075
Vegetation	0.15	0.157	0.14	0.196
Build-up and bare land	-0.05	0.636	<-0.01	0.999
Water and high-water farmland combined	0.20	0.066	0.20	0.054
Medium and low-water farmlands combined	-0.24	0.025	-0.23	0.029

<sup>a</sup>Land use compositions in the surrounding areas of 90 villages were computed for the spatial extents of sub-district and districts.

The resulting values for the CLUMPY index for the 90 villages were all greater than zero at the two spatial extents, indicating that the patches of different land use classes were more aggregated than expected under a spatially random distribution. Nevertheless, the clumpiness of most land use classes did not exhibit statistically significant relationships with *O. viverrini* prevalence, except for vegetation and low-water farmland. Significant positive correlations were found between the *O. viverrini* prevalence rates of the villages and the clumpiness of vegetation land use patches at both the sub-district ( $r = 0.23$ ;  $P = 0.031$ ) and the district ( $r = 0.27$ ;  $P = 0.009$ ) levels (Table 4). In contrast, significant negative correlations were found between the *O. viverrini* prevalence rates of the villages and the clumpiness of low-water farmland at both the sub-district ( $r = -0.30$ ;  $P = 0.005$ ) and the district ( $r = -0.22$ ;  $P = 0.039$ ) levels (Table 4). Finally, quantifications of the overall landscape configuration using the CONTAG metric showed that at the district level, the CONTAG of the landscapes in which the vil-

lages were located were negatively correlated with their *O. viverrini* infections ( $r = -0.22$ ;  $P = 0.042$ ). No statistically significant association was, however, detected for the landscapes analysed at the sub-district level ( $P > 0.05$ ).

#### Contributions of the environmental factors

The stepwise regression analysis selected a total of six variables for the multiple linear regression model, including altitude of the village, the LPIs for two land use classes (i.e. vegetation and water and high-water farmland combined), and the CLUMPY index for three land use classes (i.e. vegetation, water and high-water farmland combined, and low-water farmland). The result of the multiple linear regression analysis showed that all these explanatory variables were significant ( $P < 0.05$ ), except for the CLUMPY index for the low-water farmland land use class ( $P = 0.120$ ) (Table 5). Multicollinearity was not an issue, since all of the VIFs were less than 5.

Table 3. Relationship between *O. viverrini* prevalence at the unit of the village and the landscape dominance of their surroundings<sup>a</sup>.

Land use class	Sub-district		District	
	<i>r</i>	P-value	<i>r</i>	P-value
Water	0.07	0.490	0.10	0.327
High-water farmland	0.00	0.965	0.12	0.255
Medium-water farmland	-0.11	0.303	-0.22	0.039
Low-water farmland	-0.18	0.089	-0.22	0.039
Vegetation	0.16	0.135	0.01	0.924
Build-up and bare land	-0.05	0.672	-0.10	0.336
Water and high-water farmland combined	0.23	0.032	0.29	0.006
Medium and low-water farmlands combined	-0.16	0.136	-0.24	0.025

<sup>a</sup>The landscape dominance was measured using the largest patch index (LPI) for the spatial extents of sub-district and district.

Table 4. Relationship between *O. viverrini* prevalence at the unit of the villages and the extent at which the patches of each land use class were aggregated<sup>a</sup>.

Land use class	Sub-district		District	
	<i>r</i>	P-value	<i>r</i>	P-value
Water	-0.02	0.876	-0.06	0.562
High-water farmland	0.00	0.999	-0.01	0.915
Medium-water farmland	0.09	0.393	0.17	0.103
Low-water farmland	-0.30	0.005	-0.22	0.039
Vegetation	0.23	0.031	0.27	0.009
Build-up and bare land	-0.14	0.198	-0.10	0.336
Water and high-water farmland combined	-0.06	0.591	-0.04	0.731
Medium and low-water farmlands combined	0.06	0.561	0.15	0.169

<sup>a</sup>The degree of aggregation of the land use patches was measured using the clumpiness index (CLUMPY) for the spatial extents of sub-district and district.

## Discussion

### *Locations of the villages*

Spatial analysis of the village locations provided insights into the geographical variation of human *O. viverrini* prevalence in northeast Thailand. Mapping of *O. viverrini* infections in 2009 across 90 villages illustrated that low infections were mostly found in the Mun River catchment, as opposed to the Chi River and the Khong River catchments (Fig. 1b). This observation is in agreement with earlier research indicating that areas of high prevalence across the whole northeast Thailand were concentrated along the valley of the Mekong River (Sadun, 1955). A recent study in southern Lao PDR also suggested that high risk zones of *O. viverrini* infections were concentrated in the Mekong River corridor (Forrer et al., 2012). The differences in *O. viverrini* prevalence across three catchments in northeast Thailand might be explained by different eating culture and tradition. The Mun

River catchment is adjacent to Cambodia (Fig. 1a), and its ethnic groups generally have a lower preference for raw fish consumption, compared to the ethnic groups in the Chi River and the Khong River catchments that are of Laotian origin.

Among the variables relevant to the village locations examined in this study, altitude contributed the most in explaining the spatial variation of *O. viverrini* prevalence, evident in the inclusion of the variable in the multiple linear regression analysis (Table 5). Indeed, villages with 0% prevalence were all reported in the upstream area of the Mun River catchment (Fig. 1b), while villages with *O. viverrini* prevalence above 50% were all located at relatively lower-ground areas. This observation echoed previous findings, which recorded low *O. viverrini* prevalence in highland areas (Sayasone et al., 2011; Forrer et al., 2012). Conversely, the infection rates between the higher-ground and lower-ground villages were not statistically significant in this study area ( $P = 0.117$ ). The reason might be that the altitudinal variation of the 90 vil-

Table 5. Contributions of environmental factors to the spatial variation of *O. viverrini* prevalence in northeast Thailand<sup>a</sup>.

Environmental factor	P-value	VIF
Altitude of village	0.012	1.574
LPI for water and high-water farmland combined	0.003	1.440
LPI for vegetation	0.030	1.839
CLUMPY for water and high-water farmland combined	0.017	1.357
CLUMPY for vegetation	0.023	1.129
CLUMPY for low-water farmland	0.120	1.202

<sup>a</sup>LPI and CLUMPY indicated the largest patch index and the clumpiness index, respectively, for describing the landscape patterns of the village surroundings.

lages analysed in this study was not as large as that in prior studies as low risks of infection was found at locations with an altitude of at least 500 m in Champasack province, Lao PDR (Forrer et al., 2012), while the village altitudes of the current investigation were all below 500 m.

Although the average prevalence (20.5%) for villages within 100 m of major water bodies was more than doubled than the average (8.9%) for those outside 1,000 m, no statistical significant association was detected between *O. viverrini* prevalence and proximity to water bodies. This finding contrasted to the recent work in Lao PDR in which the risk of *O. viverrini* infection was found to be higher for people living in close proximity to freshwater bodies (Forrer et al., 2012). The lack of relation in the present study suggested that additional data on the distribution of ponds and irrigation systems, at finer resolutions, should be considered. High *O. viverrini* prevalence had been reported in villages located in irrigated areas with more developed irrigation systems and water dependent land use in Khon Kaen province (Sornmani et al., 1981). Indeed, water resource development, particularly the constructions of ponds for freshwater aquaculture and irrigation ditches for rice production, might have converted the northeastern Thailand landscape vulnerable to the parasite (Wang et al., 2011).

#### *Surrounding land use compositions and landscape patterns of the villages*

Epidemiological research has shown that the surrounding land use and agricultural practices of the villages can affect the transmission and infection of parasitic infections, particularly for schistosomiasis (Spear et al., 2004; Liang et al., 2007), but the role of land use composition in human infection of *O. viverrini* has not been well characterised. The findings of this study thus shed light on the potential effects of land use composition on the variation of *O. viverrini* prevalence in northeast Thailand. Analyses of land use compositions at both the sub-district and the district levels showed that villages surrounded by land use classes of water and high-water farmland were likely to report high *O. viverrini* prevalence rates (Tables 1 and 2; Fig. 3a), probably because these two land use classes provided suitable habitat conditions for the intermediate hosts. A noticeable example was Non Sung village, which recorded the highest prevalence of 59.5% in 2009 among the 90 villages surveyed. The land use composition of its surrounding environment

suggested that high percent compositions of water and high-water farmland within the district (Table 1; Fig. 3a) might have contributed to the high *O. viverrini* prevalence in Non Sung village.

Alternatively, the percent composition of the land use class, specifically low-water farmland at the sub-district level, exhibited a statistically significant correlation with *O. viverrini* infections ( $P = 0.028$ ) (Table 2). The correlation became slightly stronger when the land use classes of medium and low-water farmlands were combined ( $P = 0.025$ ), and the relationship remained robust when the spatial extent was broadened to the district level ( $P = 0.029$ ) (Table 2). These significant negative correlations ( $P < 0.05$ ) suggested that villages surrounded by other types of farmlands, such as drought tolerant crops, might be less susceptible to the infection of the parasite, compared to those surrounded by rice paddies. These findings confirmed Wang et al. (2011) to underscore that the interactions between health and environment, particularly the land use impacts of agriculture and aquaculture, should not be overlooked when examining the spatial variation of *O. viverrini* prevalence.

In addition to the effects of land use compositions, this study provided insights into the influences of landscape patterns, in terms of dominance, aggregation and contagion, on the variation of *O. viverrini* infections. Among the environmental factors examined in this study, variables relevant to landscape patterns explained the most of the spatial variation of *O. viverrini* prevalence (Table 5). The dominance of the landscape, measured using the LPI for individual land use classes, showed statistically positive associations ( $P < 0.05$  for both spatial extents) with *O. viverrini* prevalence for the land use of water and high-water farmland combined (Table 3). As one of the preferred habitats for *Bithynia* snails was rice paddies (Lohachit 2004-2005; Suwannatrai et al., 2011), this finding confirmed that the dominance of a large land use patch which provided suitable habitat conditions for the *O. viverrini* intermediate hosts in the village surroundings might contribute to high prevalence for the village. On the other hand, the results quantified using the LPI and the CLUMPY suggested that the larger the patch size and the more aggregated the patches of low-water farmland, the lower the *O. viverrini* prevalence rates of the villages (Tables 3 and 4). These results reconfirmed the finding in the aforementioned land use composition analysis (Table 2) for the potential effect of low-water farmland, which represented agricultural land not favourable to the intermediate hosts, on low *O. viverrini* prevalence. The significant corre-

lations for the findings related to the composition, dominance and pattern of aggregation of the low-water farmland land use may be considered for control strategies involving land use planning and crop selection.

The CONTAG index has been widely used in landscape ecology as an effective summary of overall clumpiness of the patches of the landscape (McGarigal et al., 2002). At the spatial extent of the districts, landscape contagion quantified using the CONTAG indicated a negative association with the *O. viverrini* prevalence of the villages ( $r = -0.22$ ;  $P = 0.042$ ). Low CONTAG values generally characterised the landscapes with many small and dispersed patches, signifying that the land use types were more evenly distributed across the landscapes (McGarigal et al., 2002). The finding of the negative association for the present study area suggested that a well interspersed of the patches of different land use in the village surroundings could potentially contribute to high *O. viverrini* infections.

Compared to the results from the sub-district level analyses, the strengths of the correlations generally improved when the landscape patterns were quantified at the district levels. This highlights that research on factors affecting human infections of *O. viverrini* should not be tackled locally, i.e., within individual villages. Instead, the influences of the broad landscapes surrounding the villages need to be considered because many ecological processes, including the processes related to parasitic disease ecology, operate at the broader patch mosaic. On the other hand, the lack of significant relations between *O. viverrini* prevalence and some landscape patterns quantified at the sub-district level (e.g. the overall clumpiness of the patches of the landscape) should not prevent further analyses at a smaller extent. A vital topic for further investigation is the degree to which the land use patches are spatially connected by landscape features favourable to disease transmission. It will be useful to examine the distribution and connection of irrigation systems because they provide both structural and functional connections between the *O. viverrini* intermediate host habitats (Wang, 2012). High resolution images and intensive ground truthing are required to accurately delineate the distributions of irrigation systems; however, this also presents a challenge for large spatial extent analysis. As a beginning step towards better understanding of the influence of landscape connectivity, future research on this topic can be conducted at the nearby areas (e.g. sub-districts) of selected villages with different *O. viverrini* prevalence rates

to map out detailed distributions of irrigation systems and land use patches, as well as to investigate their connections. Such information can also provide empirical insights for subsequent broad scale analysis using landscape connectivity metrics, which often require the specification of a threshold to define the distance between patches below which they are deemed connected.

## Acknowledgements

The authors wish to thank the National University of Singapore for funding support (Grant #: R109-000-151-112); Nonglak LaOprom for Thai translation; and Yikang Feng, Liang Lu and Dongdong Li for data processing.

## References

- Chaiyarit P, Sithithaworn P, Thuwajit C, Yongvanit P, 2011. Detection of salivary antibodies to crude antigens of *Opisthorchis viverrini* in opisthorchiasis and cholangiocarcinoma patients. *Clin Oral Invest* 15, 477-483.
- Elliott P, Wartenberg D, 2004. Spatial epidemiology: current approaches and future challenges. *Environ Health Persp* 112, 998-1006.
- Forrer A, Sayasone S, Vounatsou P, Vonghachack Y, Bouakhasith D, Vogt S, Glaser R, Utzinger J, Akkhavong K, Odermatt P, 2012. Spatial distribution of, and risk factors for, *Opisthorchis viverrini* infection in southern Lao PDR. *PLoS Negl Trop Dis* 6, e1481.
- Graham AJ, Danson FM, Craig PS, 2005. Ecological epidemiology: the role of landscape structure in the transmission risk of the fox tapeworm *Echinococcus multilocularis* (Leuckart 1863) (Cestoda: Cyclophyllidae: Taeniidae). *Prog Phys Geog* 29, 77-91.
- Keiser J, Utzinger J, 2005. Emerging foodborne trematodiasis. *Emerg Infect Dis* 11, 1507-1514.
- Killilea ME, Swee A, Lane RS, Briggs CJ, Ostfeld RS, 2008. Spatial dynamics of lyme disease: a review. *EcoHealth* 5, 167-195.
- Kruskal WH, Wallis WA, 1952. Use of ranks in one-criterion variance analysis. *J Am Stat Assoc* 47, 583-621.
- Liang S, Seto EYW, Remais JV, Zhong B, Yang C, Hubbard A, Davis GM, Gu X, Qiu D, Spear RC, 2007. Environmental effects on parasitic disease transmission exemplified by schistosomiasis in western China. *Proc Natl Acad Sci USA* 104, 7110-7115.
- Lillesand TM, Kiefer RW, Chipman J, 2004. Remote sensing and image interpretation. 5<sup>th</sup> edition, New York: Wiley, 763 pp.
- Lohachit C, 2004-2005. Ecological studies of *Bithynia siamensis goniomphalos*, a snail intermediate host of *Opisthorchis*



- viverrini*, in Khon Kaen province, northeast Thailand. *Malacol Rev* 37/38, 1-26.
- Malone JB, 2005. Biology-based mapping of vector-borne parasites by geographic information systems and remote sensing. *Parassitologia* 47, 27-50.
- McGarigal K, Cushman SA, Neel MC, Ene E, 2002. FRAGSTATS v3: Spatial Pattern Analysis Program for Categorical Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst, USA. Available at: <http://www.umass.edu/landeco/research/fragstats/fragstats.html> (accessed on October 2012).
- Meentemeyer RK, Haas SE, Václavík T, 2012. Landscape epidemiology of emerging infectious diseases in natural and human-altered ecosystems. *Annu Rev Phytopathol* 50, 379-402.
- Ostfeld RS, Glass GE, Keesing F, 2005. Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends Ecol Evol* 20, 328-336.
- Sadun EH, 1955. Studies of *Opisthorchis viverrini* in Thailand. *Am J Hyg* 62, 81-115.
- Saowakontha S, Pipitgool V, Pariyanonda S, Tesana S, Rojsathaporn K, Intarakhao C, 1993. Field trials in the control of *Opisthorchis viverrini* with an integrated programme in endemic areas of northeast Thailand. *Parasitology* 106, 283-288.
- Sayasone S, Mak TK, Vanmany M, Rasphone O, Vounatsou P, Utzinger J, Akkhavong K, Odermatt P, 2011. Helminth and intestinal protozoa infections, multiparasitism and risk factors in Champasack province, Lao People's Democratic Republic. *PLoS Negl Trop Dis* 5, e1037.
- Sayasone S, Odermatt P, Phoumindr N, Vongsaravane X, Sensombath V, Phetsouvanh R, Choulamany X, Strobel M, 2007. Epidemiology of *Opisthorchis viverrini* in a rural district of southern Lao PDR. *Trans R Soc Trop Med Hyg* 101, 40-47.
- Shapiro SS, Wilk MB, 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52, 591-611.
- Shin HR, Oh JK, Masuyer E, Curado MP, Bouvard V, Fang YY, Wiangnon S, Sripa B, Hong ST, 2010. Epidemiology of cholangiocarcinoma: an update focusing on risk factors. *Cancer Sci* 101, 579-85.
- Sithithaworn P, Yongvanit P, Tesana S, Pairojkul C, 2007. Liver flukes. In: Food-borne parasitic zoonoses. Darwin MK, Bernard F (eds). Berlin, Springer, 3-52 pp.
- Sornmani S, Schelp FP, Vivatanaseth P, Pongpaew P, Sritabutra P, Supawan V, Vudhivai N, Egormaiaphol S, Harinasuta C, 1981. An investigation of the health and nutritional status of the population in the Nam Pong Water Resource Development Project, northeast Thailand. *Ann Trop Med Parasitol* 75, 335-346.
- Sornmani S, Vivatanaseth P, Impand P, Phatihatakorn W, Schelp FP, 1984. Infection and re-infection rates of opisthorchiasis in the water resource development area of Nam Pong project, Khon Kaen province, northeast Thailand. *Ann Trop Med Parasitol* 78, 649-656.
- Spear RC, Seto E, Liang S, Birkner M, Hubbard A, Qui D, Yang C, Zhong B, Xu F, Gu X, Davis GM, 2004. Factors influencing the transmission of *Schistosoma japonicum* in the mountains of Sichuan province of China. *Am J Trop Med Hyg* 70, 48-56.
- Sriamporn S, Pisani P, Pipitgool V, Suwanrungruang K, Kamsard S, Parkin DM, 2004. Prevalence of *Opisthorchis viverrini* infection and incidence of cholangiocarcinoma in Khon Kaen, northeast Thailand. *Trop Med Int Health* 9, 588-594.
- Sripa B, 2008. Concerted action is needed to tackle liver fluke infections in Asia. *PLoS Negl Trop Dis* 2, e232.
- Sripa B, Bethony JM, Sithithaworn P, Kaewkes S, Mairiang E, Loukas A, Mulvenna J, Laha T, Hotez PJ, Brindley PJ, 2011. Opisthorchiasis and *Opisthorchis*-associated cholangiocarcinoma in Thailand and Laos. *Acta Trop* 120, S158-S168.
- Suwannatrai A, Suwannatrai K, Haruay S, Piratae S, Thammasiri C, Khampoosa P, Kulsantiwong J, Prasopdee S, Tarbsripair P, Suwanwerakamtorn R, Sukchan S, Boonmars T, Malone JB, Kearney MT, Tesana S, 2011. Effect of soil surface salt on the density and distribution of the snail *Bithynia siamensis goniomphalos* in northeast Thailand. *Geospat Health* 5, 183-190.
- Tesana S, Sithithaworn P, Prasongwatana J, Kaewkes S, Pipitgool V, Pientong C, 1991. Influence of water current on the distribution of *Opisthorchis viverrini* infection in northeastern villages of Thailand. *Southeast Asian J Trop Med Public Health* 22, 93-98.
- Thai Meteorological Department, 2012. Climate of Thailand. Available at: <http://www.tmd.go.th/en/archive/surfacetemperature.php> (accessed on December 2012).
- Wang YC, 2012. Examining landscape determinants of *Opisthorchis viverrini* transmission. *EcoHealth* 9, 328-341.
- Wang YC, Feng CC, Sithithaworn P, Feng Y, Petney TN, 2011. How do snails meet fish? Landscape perspective needed to study parasite prevalence. *EcoHealth* 8, 258-260.
- Wang YC, Feng CC, Vu Duc H, 2012. Integrating multi-sensor remote sensing data for land use/cover mapping in a tropical mountainous area in northern Thailand. *Geogr Res* 50, 320-331.
- Zhou XN, Lv S, Yang GJ, Kristensen TK, Bergquist NR, Utzinger J, Malone JB, 2009. Spatial epidemiology in zoonotic parasitic diseases: insights gained at the 1st International Symposium on Geospatial Health in Lijiang, China, 2007. *Parasit Vectors* 2, 10.